



# Heat transfer in a vertically integrated semiconductor device with unequal layer widths

Girish Krishnan, Ankur Jain\*

Mechanical and Aerospace Engineering Department, University of Texas at Arlington, 500 W First St, Rm 211, Arlington, TX 76019, USA

## ARTICLE INFO

### Keywords:

Vertically integrated semiconductor devices  
Thermal management  
Heterogeneous integration  
Heat transfer analysis

## ABSTRACT

Vertically integrated semiconductor devices and interposer-based heterogeneous integration technologies are being commonly investigated for advanced microelectronic chips. Such multilayer structures often integrate layers of unequal lateral sizes. While most of the past literature on thermal modeling of multilayer semiconductor chips considers layers of uniform widths, heat transfer in multilayer systems with unequal widths orthogonal to the thickness direction has not been investigated much. This remains an important challenge towards effective thermal management of vertically integrated ICs and other semiconductor chips. This work presents an analytical model to determine the temperature distribution in a multilayer device with unequal layer widths. A series solution for temperature distribution in each layer is derived, with an independent set of eigenvalues for each layer. The coefficients of the series solutions are determined through the derivation of a set of linear algebraic equations in the coefficients. Results are found to be in good agreement with numerical simulations, and also agree well with past work for special cases. An analysis of the impact of unequal widths on heat transfer and temperature distribution is presented. It is shown that the lower the ratio of the layer widths, the greater is the peak temperature. On the other hand, layer thickness is found to have an interesting, non-monotonous impact on peak temperature, which is explained on the basis of the two-dimensional nature of heat transfer in this problem. It is found that for a representative SiGe-Si device, peak temperature increases by around 9% as the SiGe layer width is reduced to one-third compared to a baseline equal width case. The novelty of the present work is that it presents the thermal modeling of an important multilayer semiconductor architecture, which has not been addressed much in the past. Results presented here may help improve the thermal design of vertically integrated semiconductor devices and other similar multilayer devices.

## 1. Introduction

Thermal management of semiconductor devices and systems continues to remain a key technological challenge that affects performance and reliability [1,2]. Traditionally, low-power semiconductor chips have been cooled simply by natural convection or air cooling along with a heat spreader and heat sink [3]. Liquid cooling [4] and two-phase cooling [5] have also been investigated for chips with higher power dissipation. In general, key challenges in semiconductor thermal management include dynamic non-uniform heat generation [6], interfacial thermal contact resistances [7] and trade-offs between thermal and electrical performance [8].

The continued increase in transistor density along with vertical heterogeneous integration of multiple chips makes the challenge of heat removal even more formidable [9]. In particular, heat transfer pathways

within vertically integrated devices and interposer based System-in-Package (SiP) are more complicated and limited compared to traditional chips [10]. Moreover, features specific to such technologies, such as through-Silicon vias (TSVs) [11] and inter-die bond pads [12] offer additional challenges such as thermal contact resistance between adjacent die [13]. Finally, the integration of materials with differing temperature requirements also necessitates careful thermal management.

There remains a continued need for accurate tools for predicting the temperature distribution in multilayer semiconductor chips. Experimental temperature measurement is time consuming and expensive, and does not offer the ability of rapid iterative improvement, due to the prohibitive cost of Silicon design and fabrication. Therefore, temperature prediction tools are valuable for pre-fabrication design. While the calculation of the full transient temperature distribution is desirable [14], thermal design is often carried out simply on the basis of

\* Corresponding author.

E-mail address: [jaina@uta.edu](mailto:jaina@uta.edu) (A. Jain).

Nomenclature			
$Bi$	Biot number, $Bi = \frac{h z_M}{k_M}$	$T$	temperature (K)
$\bar{g}$	non-dimensional thermal contact resistance, $\bar{g}_m(\eta) = \frac{R_m(y)k_M}{z_M}$	$w$	width of the body in the $y$ direction (m)
$h$	convective heat transfer coefficient ( $Wm^{-2}K^{-1}$ )	$\bar{w}$	non-dimensional width of the body in the $y$ direction, $\bar{w}_m = \frac{w_m}{z_M}$
$k$	thermal conductivity ( $Wm^{-1}K^{-1}$ )	$y, z$	spatial coordinates (m)
$\bar{k}$	non-dimensional thermal conductivity, $\bar{k}_m = \frac{k_m}{k_M}$	$\eta, \xi$	non-dimensional spatial coordinates, $\eta = \frac{y}{z_M}$ ; $\xi = \frac{z}{z_M}$
$M$	number of layers	$\gamma$	non-dimensional interface location, $\gamma_m = \frac{z_m}{z_M}$
$N$	number of eigenvalues	$\theta$	non-dimensional temperature, $\theta_m = \frac{T_m - T_{amb}}{T_{ref} - T_{amb}}$
$q^*$	heat flux ( $Wm^{-2}$ )	$\lambda$	non-dimensional eigenvalue
$q$	heat (W)	<b>Subscripts</b>	
$\bar{q}$	non-dimensional heat flux, $\bar{q}(\eta) = \frac{q^*(y)z_M}{k_M(T_{ref} - T_{amb})}$	<i>amb</i>	ambient
$R$	thermal contact resistance ( $Km^2W^{-1}$ )	<i>m</i>	layer number
		<i>M</i>	total number of layers
		<i>ref</i>	reference

steady-state concepts such as junction-to-ambient thermal resistance [15]. A compact thermal modeling approach has been shown to result in accurate temperature prediction for use in electrical design [14]. A thermal resistance network based thermal model of a 3D IC has been shown to result in around 10% increase in peak temperature rise compared to an equivalent traditional IC [15]. Computation of the temperature field in a two-layer 3D IC using Green's function method has been used for thermal-electrical co-optimization [16]. The underlying energy conservation equations have been solved in order to derive closed-form expressions for the temperature field for equally-sized die [17,18]. While an iterative model was used to demonstrate 13% reduction in peak temperature through thermally-aware floorplanning, an exact non-iterative solution was used to quantify the increase in temperature rise with increasing number of die in the stack and increasing thermal resistance between die [18]. An 8% increase in peak temperature for a stack of unequally-sized die compared to a uniform die stack was reported using an iterative method that suffered from convergence challenges [19]. Resistance-capacitance network based thermal models for a 3D IC are also available [14]. A key conclusion from such papers is that heat generated in one layer often flows through, and causes additional temperature rise in multiple other layers. The integration of temperature prediction tools with electrical design models and tools has been widely explored for co-design and co-optimization. For example, thermal-aware 3D placement has been shown to result in 34% reduction in peak temperature [20]. A similar 30–60 °C reduction in peak temperature was reported by temperature-aware routing in 3D ICs using a resistance-based temperature computation [21]. It is important for thermal models, whether analytical or numerical, to account for various complications that occur in a vertically integrated device, including dynamic and spatially-varying heat generation, inter-die thermal resistance, presence of TSVs and unequally sized die.

A key design flexibility offered by vertical integration technology in semiconductors is the ability to integrate multiple die of different sizes. The die-on-wafer integration technique makes it possible to separately fabricate and then bond individual die on to a large wafer [13]. This enables improved yield due to the use of known good die (KGD) as well as the ability to partly fabricate at cheaper technology nodes, such as memory die that can be fabricated using older, depreciated technology. Similarly, interposer-based package integration makes it possible to integrate multiple chips/chiplets of different sizes on to the same interposer [22]. In such technologies, the two or more die being integrated are not of the same size, which offers unique thermal challenges. In general, the flow of heat between two unequally-sized die results in thermal spreading, and, therefore, increased temperature rise. While such thermal spreading effects have been studied in the context of electronics packaging, where the heat spreader is usually larger in size

than the die [23], there is a lack of work in the context of a multilayer chip in which each die itself is of different size. Since such designs are now becoming common due to heterogeneous integration technologies such as vertically integrated devices and interposers, it is important to develop an understanding of thermal transport in multilayer devices with unequally sized layers.

Analytical thermal modeling of a stack of unequally-sized die is complicated because each substrate may have a different set of eigenvalues, and, therefore, satisfying the interface conditions is not straightforward. Moreover, modeling of interfacial conditions is not trivial because the die-to-die interface exists along only along the intersecting surface, while a different – usually adiabatic – boundary condition must be applied in the overhanging region. As a result of these complications, only limited literature is available on thermal modeling of a multilayer device with unequally-sized die. An iterative technique has been proposed for solving such a problem [19], which, however, is very cumbersome, and suffers from convergence challenges. This problem is similar to the thermal spreading resistance problem, in which, heat conducts from a source into a larger body, for which, a sizable literature is available [24]. However, such a model completely ignores thermal conduction in the smaller die, and treats it, instead, only as a boundary condition imposed on the larger die [23,24]. A limited work does account for thermal conduction in both mating bodies of unequal size, but the bodies are both considered to be semi-infinitely thick [25], which is clearly an unrealistic model for semiconductor chips. Finally, the literature cited above is limited only to steady state analysis and does not account for transient effects such as dynamic changes in heat dissipation, as well as thermal resistance between layers [12,26], which itself may be a function of space due to spatial distribution of bond pads over the interface between adjacent die [7,27]. There is, clearly, a need to comprehensively examine thermal transport in a multilayer body with unequally-sized layers and account for the various realistic complications that arise in a vertically integrated semiconductor device.

This work presents an analytical model for thermal conduction in a multi-layer body in which the layers may be of unequal width. Spatial variation in heat flux at one end, as well as in thermal contact resistance between adjacent die is also accounted for. Series solutions for the steady state temperature field with independent eigenvalues in each layer are written. Coefficients appearing in the series solution are determined by deriving a set of linear algebraic equations on the basis of interface conditions between the layers. It is shown that for special cases, the general model considered in the present work agrees well with results from previous work, in addition to an excellent agreement with numerical simulations. The key novelty of this work is in addressing the heat transfer problem in multilayer devices with unequal layer widths, which is representative of several state-of-the-art semiconductor devices.

The next section defines, non-dimensionalizes and presents a solution for the problem considered in this work. This is followed by a discussion of key results in Section 3, including convergence analysis, comparison with past work and numerical simulations, and the impact of various problem parameters on the temperature distribution. Finally, Section 4 presents key conclusions of this work.

## 2. Mathematical modeling

Fig. 1 presents a schematic of a general  $M$ -layer semiconductor device with layers of unequal widths, as may be the case in a vertically integrated semiconductor device. The width of each layer is given by  $2w_m$  and the thickness of the  $m$ th layer is  $z_m - z_{m-1}$ , where  $m=1,2,..M$ . Each layer has its own uniform and isotropic thermal conductivity denoted by  $k_m$  ( $m=1,2,..M$ ).  $R_m(y)$  ( $m=1,2,..M-1$ ) denote the thermal contact resistances between adjacent layers, assumed, in general to vary along the interface. The bottom face of the body is assumed to experience a spatially varying heat flux  $q''(y)$ , which may be representative of heat generation due to transistor operation. Convective cooling occurs on the top surface, which is modeled by a constant heat transfer coefficient  $h$ . This boundary condition represents, for example, cooling of the device by a heat sink. All other boundaries in the problem, including side walls and overhanging surfaces are assumed to be adiabatic in nature, as is commonly the case in semiconductor devices [17–19]. Additionally, symmetry is assumed as shown in Fig. 1, which allows consideration of only one half of the problem. Therefore, the steady state governing equation for temperature distribution in one half of the body is given by [17,28]:

$$\frac{\partial^2 T_m}{\partial z^2} + \frac{\partial^2 T_m}{\partial y^2} = 0 \quad (z_{m-1} \leq z \leq z_m; 0 \leq y \leq w_m; m = 1, 2, ..M) \quad (1)$$

Based on the problem statement discussed above, the  $z$ -direction boundary conditions are as follows [17,28]:

$$-k_1 \frac{\partial T_1}{\partial z} = q''(y) \quad (z=0) \quad (2)$$

$$k_M \frac{\partial T_M}{\partial z} + h(T_M - T_{amb}) = 0 \quad (z=z_M) \quad (3)$$

The following interface condition models thermal contact resistance between layers [29]:

$$T_m = T_{m+1} - k_m \frac{\partial T_m}{\partial z} R_m(y) \quad (0 < y < \min(w_m, w_{m+1}); z = z_m; m = 1, 2, ..M - 1) \quad (4)$$

Here,  $\min(w_m, w_{m+1})$  represents the smaller of the widths of the two

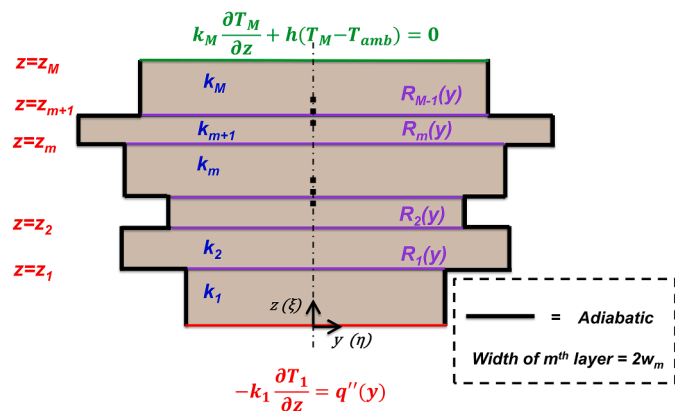


Fig. 1. Schematic of the general  $M$ -layer semiconductor device with layers of unequal widths. The device is assumed to be symmetric about the centerline, as shown.

adjacent layers, over which the layers intersect.

For each interface,  $m = 1, 2, ..M - 1$ , the following flux condition is imposed at  $z = z_m$  if  $w_m \leq w_{m+1}$  [17,28]:

$$k_{m+1} \frac{\partial T_{m+1}}{\partial z} = \begin{cases} k_m \frac{\partial T_m}{\partial z} & (0 < y \leq w_m) \\ 0 & (w_m < y \leq w_{m+1}) \end{cases} \quad (z = z_m) \quad (5a)$$

On the other hand, if  $w_m > w_{m+1}$  [17,28],

$$k_m \frac{\partial T_m}{\partial z} = \begin{cases} k_{m+1} \frac{\partial T_{m+1}}{\partial z} & (0 < y \leq w_{m+1}) \\ 0 & (w_{m+1} < y \leq w_m) \end{cases} \quad (z = z_m) \quad (5b)$$

The two-part nature of Eq. (5a) and (b) ensures that the boundary is modeled as adiabatic for  $y > \min(w_m, w_{m+1})$  at each interface, as shown by the thick black lines in Fig. 1. The adiabatic nature of the overhanging boundary is justified because in practical semiconductor devices, most of the heat removal occurs from one of the end faces that has a heat sink attached [3]. Coolant air flow is mostly directed at the heat sink rather than the overhanging surface. Moreover, it is common for multilayer semiconductor devices to be encapsulated in low thermal conductivity epoxy [13] that prevents significant heat transfer from the overhanging surface.

Note that for the special case of equal layer widths ( $w_m = w_{m+1}$ ), Eqs. (4) and (5) reduce to the usual temperature and flux continuity equations between layers of equal width.

Boundary conditions along the symmetry line and on the side walls are given by [3]:

$$\frac{\partial T_m}{\partial y} = 0 \quad (y = 0, w_m; m = 1, 2, ..M) \quad (6)$$

Eq. (6) is justified because most of the heat removal in practical semiconductor devices occurs through the top or bottom surface where a heat sink may be attached. In contrast, there is minimal air flow around the sidewalls, which also have relatively very small surface area, all of which results in heat removal from the side walls [3].

The problem defined by Eqs. (1)–(6) is first non-dimensionalized as follows:

$$\begin{aligned} \theta_m &= \frac{T_m - T_{amb}}{T_{ref} - T_{amb}}; \xi = \frac{z}{z_M}; \eta = \frac{y}{z_M}; \gamma_m = \frac{z_m}{z_M}; \bar{w}_m = \frac{w_m}{z_M}; \bar{k}_m = \frac{k_m}{k_M} \bar{g}_m(\eta) \\ &= \frac{R_m(y)k_M}{z_M}; \bar{q}(\eta) = \frac{q''(y)z_M}{k_M(T_{ref} - T_{amb})}; Bi = \frac{hz_M}{k_M} \end{aligned} \quad (7)$$

Here,  $\gamma_m$  are the interface locations,  $\bar{w}_m$  and  $\bar{k}_m$  are the layer widths and thermal conductivities, respectively. Further,  $\bar{g}_m$ ,  $\bar{q}$  and  $Bi$  represent the non-dimensional thermal contact resistances at interfaces, heat flux and convective heat transfer coefficient, respectively. Note that  $T_{ref}$  is an arbitrary temperature such that  $T_{ref} \neq T_{amb}$ .

The non-dimensional set of equations based on Eq. (7) is obtained as shown below:

$$\frac{\partial^2 \theta_m}{\partial \xi^2} + \frac{\partial^2 \theta_m}{\partial \eta^2} = 0 \quad (\gamma_{m-1} \leq \xi \leq \gamma_m; 0 \leq \eta \leq \bar{w}_m; m = 1, 2, ..M) \quad (8)$$

$$-\bar{k}_1 \frac{\partial \theta_1}{\partial \xi} = \bar{q}(\eta) \quad (\xi = 0) \quad (9)$$

$$\frac{\partial \theta_M}{\partial \xi} + Bi \cdot \theta_M = 0 \quad (\xi = 1) \quad (10)$$

$$\theta_m = \theta_{m+1} - \bar{k}_m \frac{\partial \theta_m}{\partial \xi} \bar{g}_m(\eta) \quad (0 < \eta < \min(\bar{w}_m, \bar{w}_{m+1}); \xi = \gamma_m; m = 1, 2, ..M - 1) \quad (11)$$

$$\bar{k}_{m+1} \frac{\partial \theta_{m+1}}{\partial \xi} = \begin{cases} \bar{k}_m \frac{\partial \theta_m}{\partial \xi} & (0 < \eta \leq \bar{w}_m) \quad (\bar{w}_m < \eta \leq \bar{w}_{m+1}) \quad (\xi = \gamma_m; \text{ if } \bar{w}_m \leq \bar{w}_{m+1}) \end{cases} \quad (12a)$$

$$\bar{k}_m \frac{\partial \theta_m}{\partial \xi} = \begin{cases} \bar{k}_{m+1} \frac{\partial \theta_{m+1}}{\partial \xi} & (0 < \eta \leq \bar{w}_{m+1}) \quad (\bar{w}_{m+1} < \eta \leq \bar{w}_m) \quad (\xi = \gamma_m; \text{ if } \bar{w}_m > \bar{w}_{m+1}) \end{cases} \quad (12b)$$

$$\frac{\partial \theta_m}{\partial \eta} = 0 \quad (\eta = 0, \bar{w}_m; m = 1, 2..M) \quad (13)$$

This problem in non-dimensional form is solved using the separation of variables technique by seeking a series solution for the temperature field in each layer as follows [28]:

$$\theta_m(\xi, \eta) = A_{m,0} + B_{m,0}\xi + \sum_{n=1}^{\infty} (A_{m,n} \cosh(\lambda_{m,n}\xi) + B_{m,n} \sinh(\lambda_{m,n}\xi)) \cos(\lambda_{m,n}\eta) \quad (14)$$

Note that the  $A_{m,0} + B_{m,0}\xi$  term corresponds to the zero eigenvalue, and is necessitated by the adiabatic boundary conditions at both  $\eta = 0$  and  $\eta = \bar{w}_m$ . The eigenfunctions  $\cos(\lambda_{m,n}\eta)$  and eigenvalues  $\lambda_{m,n} = \frac{n\pi}{\bar{w}_m}$  are obtained using the symmetry and side wall boundary conditions, where  $n = 1, 2..N$ . Note that, in general, each layer has a unique set of eigenvalues.

While the series solution given by Eq. (14) requires, in theory, an infinite number of terms, for practical computation, it is commonly truncated to a sufficiently large number of terms,  $N$ . Based on such

$$\bar{k}_{m+1} \left[ B_{m+1,0} + \sum_{n=1}^N \lambda_{m+1,n} (A_{m+1,n} \sinh(\lambda_{m+1,n}\gamma_m) + B_{m+1,n} \cosh(\lambda_{m+1,n}\gamma_m)) \cos(\lambda_{m+1,n}\eta) \right] = \begin{cases} \bar{k}_m \left[ B_{m,0} + \sum_{n=1}^N \lambda_{m,n} (A_{m,n} \sinh(\lambda_{m,n}\gamma_m) + B_{m,n} \cosh(\lambda_{m,n}\gamma_m)) \cos(\lambda_{m,n}\eta) \right] & (0 < \eta \leq \bar{w}_m) \quad (\bar{w}_m < \eta \leq \bar{w}_{m+1}) \end{cases} \quad (21)$$

truncation, a total of  $2(N+1)M$  coefficients  $A_{m,0}$ ,  $B_{m,0}$ ,  $A_{m,n}$  and  $B_{m,n}$  ( $m = 1, 2, ..M; n = 0, 1, 2, ..N$ ) are needed to complete the solution, given by Eq. (14). These are determined by deriving a sufficient number of linear algebraic equations using the boundary and interface conditions along the  $\xi$  direction, given by Eqs. (9)–(12).

To begin with, using Eq. (14) in Eq. (9) results in:

$$\bar{k}_1 \left( B_{1,0} + \sum_{n=1}^N \lambda_{1,n} B_{1,n} \cos(\lambda_{1,n}\eta) \right) = -\bar{q}(\eta) \quad (15)$$

Simply integrating Eq. (15) between  $\eta = 0$  and  $\eta = \bar{w}_1$  results in:

$$B_{1,0} = -\frac{1}{k_1 \bar{w}_1} \int_0^{\bar{w}_1} \bar{q}(\eta^*) d\eta^* \quad (16)$$

Further, multiplying Eq. (15) with  $\cos(\lambda_{1,n}\eta)$  for  $n' = 1, 2..N$  and integrating between  $\eta = 0$  and  $\eta = \bar{w}_1$  yields:

$$B_{1,n'} = -\frac{2}{k_1 \bar{w}_1} \lambda_{1,n'} \int_0^{\bar{w}_1} \bar{q}(\eta^*) \cos(\lambda_{1,n'} \eta^*) d\eta^* \quad n' = 1, 2..N \quad (17)$$

Eqs. (16) and (17) together constitute  $(N+1)$  equations. Note that for the special case of uniform heat flux  $\bar{q}$ ,  $B_{1,0} = -\frac{\bar{q}}{k_1}$  and  $B_{1,n'} = 0$  for  $n' = 1, 2..N$ .

The next set of  $(N+1)$  equations are obtained by using the boundary condition at  $\xi = 1$ . Using Eq. (10), one may write:

$$B_{M,0} + \sum_{n=1}^N \lambda_{M,n} (A_{M,n} \sinh(\lambda_{M,n}) + B_{M,n} \cosh(\lambda_{M,n})) \cos(\lambda_{M,n}\eta) + Bi \left[ A_{M,0} + B_{M,0} + \sum_{n=1}^N (A_{M,n} \cosh(\lambda_{M,n}) + B_{M,n} \sinh(\lambda_{M,n})) \cos(\lambda_{M,n}\eta) \right] = 0 \quad (18)$$

Term-by-term comparison in Eq. (18) yields the second set of  $(N+1)$  equations as follows:

$$(1 + Bi)B_{M,0} + Bi \cdot A_{M,0} = 0 \quad (19)$$

and

$$A_{M,n} (\lambda_{M,n} \sinh(\lambda_{M,n}) + Bi \cosh(\lambda_{M,n})) + B_{M,n} (\lambda_{M,n} \cosh(\lambda_{M,n}) + Bi \sinh(\lambda_{M,n})) = 0 \quad n = 1, 2..N \quad (20)$$

The procedure for deriving the remaining equations based on boundary conditions at  $\xi = \gamma_m$  depends on whether  $\bar{w}_m$  is less than or greater than  $\bar{w}_{m+1}$ . Depending on the relative widths of the intersecting layers, either one of two distinct Cases apply at each interface – Case A ( $\bar{w}_m \leq \bar{w}_{m+1}$ ) or Case B ( $\bar{w}_m > \bar{w}_{m+1}$ ), as illustrated in Fig. 2.

When Case A applied, firstly, the flux condition given by Eq. 12(a) results in:

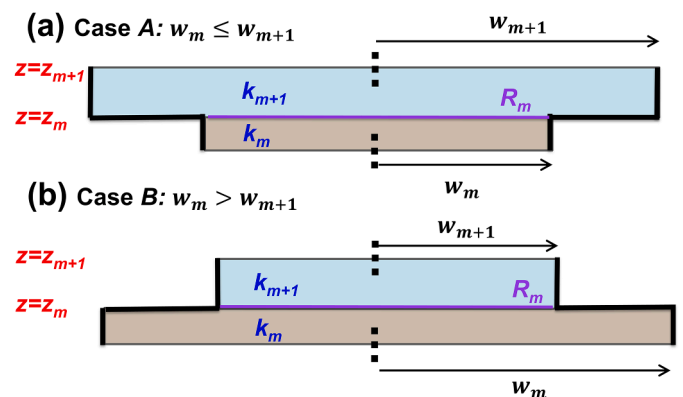


Fig. 2. Illustration of (a) Case A and (b) Case B scenarios at the interface between the  $m^{th}$  and  $(m + 1)^{th}$  layers.

Multiplying Eq. (21) with  $\cos(\lambda_{m+1,n}\eta)$  for  $n' = 0$ , and integrating between  $\eta = 0$  and  $\eta = \bar{w}_{m+1}$  yields the following equation:

$$\begin{aligned} & \bar{k}_{m+1} \left[ B_{m+1,0} \bar{w}_{m+1} + \sum_{n=1}^N \lambda_{m+1,n} (A_{m+1,n} \sinh(\lambda_{m+1,n} \gamma_m) + B_{m+1,n} \cosh(\lambda_{m+1,n} \gamma_m)) \int_0^{\bar{w}_{m+1}} \cos(\lambda_{m+1,n} \eta^*) d\eta^* \right] \\ & = \bar{k}_m \left[ B_{m,0} \bar{w}_m + \sum_{n=1}^N \lambda_{m,n} (A_{m,n} \sinh(\lambda_{m,n} \gamma_m) + B_{m,n} \cosh(\lambda_{m,n} \gamma_m)) \int_0^{\bar{w}_m} \cos(\lambda_{m,n} \eta^*) d\eta^* \right] \end{aligned} \tag{22}$$

The integrals in Eq. (22) drop out, leading to a considerably simplified equation as follows:

$$\bar{k}_{m+1} B_{m+1,0} \bar{w}_{m+1} - \bar{k}_m B_{m,0} \bar{w}_m = 0 \tag{23}$$

Similarly, multiplying Eq. (21) with  $\cos(\lambda_{m+1,n}\eta)$  for  $n' = 1, 2..N$ , followed by integration between  $\eta = 0$  and  $\eta = \bar{w}_{m+1}$  results in the following  $N$  equations at the interface:

$$\begin{aligned} & \bar{k}_{m+1} \left[ B_{m+1,0} \int_0^{\bar{w}_{m+1}} \cos(\lambda_{m+1,n} \eta^*) d\eta^* + N_{m+1,n'} (\lambda_{m+1,n'} (A_{m+1,n'} \sinh(\lambda_{m+1,n'} \gamma_m) + B_{m+1,n'} \cosh(\lambda_{m+1,n'} \gamma_m))) \right] \\ & = \bar{k}_m \left[ B_{m,0} \int_0^{\bar{w}_m} \cos(\lambda_{m+1,n} \eta^*) d\eta^* + \sum_{n=1}^N \lambda_{m,n} (A_{m,n} \sinh(\lambda_{m,n} \gamma_m) + B_{m,n} \cosh(\lambda_{m,n} \gamma_m)) \int_0^{\bar{w}_m} \cos(\lambda_{m,n} \eta^*) \cos(\lambda_{m+1,n} \eta^*) d\eta^* \right] \end{aligned} \tag{24}$$

where  $N_{m+1,n'} = \int_0^{\bar{w}_{m+1}} \cos^2(\lambda_{m+1,n'} \eta^*) d\eta^*$ . Additional equations may be derived using the interface condition associated with the thermal contact resistance. Using Eq. (11), one may write:

$$\begin{aligned} & A_{m,0} + B_{m,0} \gamma_m + \sum_{n=1}^N (A_{m,n} \cosh(\lambda_{m,n} \gamma_m) + B_{m,n} \sinh(\lambda_{m,n} \gamma_m)) \cos(\lambda_{m,n} \eta) \\ & = A_{m+1,0} + B_{m+1,0} \gamma_m \\ & + \sum_{n=1}^N (A_{m+1,n} \cosh(\lambda_{m+1,n} \gamma_m) + B_{m+1,n} \sinh(\lambda_{m+1,n} \gamma_m)) \cos(\lambda_{m+1,n} \eta) \\ & - \bar{k}_m \left[ B_{m,0} \right. \\ & \left. + \sum_{n=1}^N \lambda_{m,n} (A_{m,n} \sinh(\lambda_{m,n} \gamma_m) + B_{m,n} \cosh(\lambda_{m,n} \gamma_m)) \cos(\lambda_{m,n} \eta) \right] \end{aligned} \tag{25}$$

Eq. (25) is also treated similar to Eq. (21). Specifically, Eq. (25) is multiplied with  $\cos(\lambda_{m,n}\eta)$  for  $n' = 0, 1, 2..N$ , and integrated between  $\eta = 0$  and  $\eta = \bar{w}_m$  instead. The  $n' = 0$  case results in:

$$\begin{aligned} (A_{m,0} + B_{m,0} \gamma_m) \bar{w}_m & = (A_{m+1,0} + B_{m+1,0} \gamma_m) \bar{w}_m - B_{m,0} \bar{k}_m \int_0^{\bar{w}_m} \bar{g}_m(\eta^*) d\eta^* \\ & - \sum_{n=1}^N \bar{k}_m \lambda_{m,n} (A_{m,n} \sinh(\lambda_{m,n} \gamma_m) \\ & + B_{m,n} \cosh(\lambda_{m,n} \gamma_m)) \int_0^{\bar{w}_m} \bar{g}_m(\eta^*) \cos(\lambda_{m,n} \eta^*) d\eta^* \\ & + \sum_{n=1}^N (A_{m+1,n} \cosh(\lambda_{m+1,n} \gamma_m) \\ & + B_{m+1,n} \sinh(\lambda_{m+1,n} \gamma_m)) \int_0^{\bar{w}_m} \cos(\lambda_{m+1,n} \eta^*) d\eta^* \end{aligned} \tag{26}$$

On the other hand,  $n' = 1, 2..N$  results in:

$$\begin{aligned}
 N_{m,n'}(A_{m,n'}\cosh(\lambda_{m,n'}\gamma_m) + B_{m,n'}\sinh(\lambda_{m,n'}\gamma_m)) &= \sum_{n=1}^N (A_{m+1,n}\cosh(\lambda_{m+1,n}\gamma_m) + B_{m+1,n}\sinh(\lambda_{m+1,n}\gamma_m)) \int_0^{\bar{w}_m} \cos(\lambda_{m+1,n}\eta^*) \cos(\lambda_{m,n'}\eta^*) d\eta^* \\
 - B_{m,0}\bar{k}_m \int_0^{\bar{w}_m} \bar{g}_m(\eta^*) \cos(\lambda_{m,n'}\eta^*) d\eta^* &- \sum_{n=1}^N \bar{k}_m \lambda_{m,n} (A_{m,n}\sinh(\lambda_{m,n}\gamma_m) + B_{m,n}\cosh(\lambda_{m,n}\gamma_m)) \int_0^{\bar{w}_m} \bar{g}_m(\eta^*) \cos(\lambda_{m,n}\eta^*) \cos(\lambda_{m,n'}\eta^*) d\eta^*
 \end{aligned} \tag{27}$$

where  $N_{m,n'} = \int_0^{\bar{w}_m} \cos^2(\lambda_{m,n'}\eta^*) d\eta^*$

This completes the treatment for Case A.

When Case B applies, i.e., when  $\bar{w}_m > \bar{w}_{m+1}$ , the flux condition given by Eq. (12b) is used to obtain

$$\begin{aligned}
 &\bar{k}_m \left[ B_{m,0} + \sum_{n=1}^N \lambda_{m,n} (A_{m,n}\sinh(\lambda_{m,n}\gamma_m) + B_{m,n}\cosh(\lambda_{m,n}\gamma_m)) \cos(\lambda_{m,n}\eta) \right] \\
 &= \left\{ \bar{k}_{m+1} \left[ B_{m+1,0} + \sum_{n=1}^N \lambda_{m+1,n} (A_{m+1,n}\sinh(\lambda_{m+1,n}\gamma_m) + B_{m+1,n}\cosh(\lambda_{m+1,n}\gamma_m)) \cos(\lambda_{m+1,n}\eta) \right] \right\} (0 < \eta \leq \bar{w}_{m+1}) \quad (0 < \bar{w}_{m+1} < \eta \leq \bar{w}_m)
 \end{aligned} \tag{28}$$

Eq. (28) is multiplied with  $\cos(\lambda_{m,n}\eta)$  for  $n' = 0, 1, 2, \dots, N$ , and integrated between  $\eta = 0$  and  $\eta = \bar{w}_m$ , resulting in Eq. (23) for the  $n' = 0$  case, and the following equation for  $n' = 1, 2, \dots, N$ :

$$\begin{aligned}
 &\bar{k}_m [N_{m,n'}\lambda_{m,n'}(A_{m,n'}\sinh(\lambda_{m,n'}\gamma_m) + B_{m,n'}\cosh(\lambda_{m,n'}\gamma_m))] \\
 &= \bar{k}_{m+1} \left[ B_{m+1,0} \int_0^{\bar{w}_{m+1}} \cos(\lambda_{m,n'}\eta^*) d\eta^* + \sum_{n=1}^N \lambda_{m+1,n} (A_{m+1,n}\sinh(\lambda_{m+1,n}\gamma_m) \right. \\
 &\quad \left. + B_{m+1,n}\cosh(\lambda_{m+1,n}\gamma_m)) \int_0^{\bar{w}_{m+1}} \cos(\lambda_{m+1,n}\eta^*) \cos(\lambda_{m,n'}\eta^*) d\eta^* \right]
 \end{aligned} \tag{29}$$

Further, multiplication of Eq. (11) with  $\cos(\lambda_{m+1,n}\eta)$  and integration between  $\eta = 0$  and  $\eta = \bar{w}_{m+1}$  for  $n' = 0, 1, 2, \dots, N$  yields the following  $N + 1$  equations at the interface:

$$\begin{aligned}
 (A_{m,0} + B_{m,0}\gamma_m)\bar{w}_{m+1} + \sum_{n=1}^N (A_{m,n}\cosh(\lambda_{m,n}\gamma_m) + B_{m,n}\sinh(\lambda_{m,n}\gamma_m)) \int_0^{\bar{w}_{m+1}} \cos(\lambda_{m,n}\eta^*) d\eta^* &= (A_{m+1,0} + B_{m+1,0}\gamma_m) \\
 \bar{w}_{m+1} - B_{m,0}\bar{k}_m \int_0^{\bar{w}_{m+1}} \bar{g}_m(\eta^*) d\eta^* - \sum_{n=1}^N \bar{k}_m \lambda_{m,n} (A_{m,n}\sinh(\lambda_{m,n}\gamma_m) + B_{m,n}\cosh(\lambda_{m,n}\gamma_m)) \int_0^{\bar{w}_{m+1}} \bar{g}_m(\eta^*) \cos(\lambda_{m,n}\eta^*) d\eta^* &
 \end{aligned} \tag{30}$$

and

$$\begin{aligned}
 & \sum_{n=1}^N (A_{m,n} \cosh(\lambda_{m,n} \gamma_m) + B_{m,n} \sinh(\lambda_{m,n} \gamma_m)) \int_0^{\bar{w}_{m+1}} \cos(\lambda_{m,n} \eta^*) \cos(\lambda_{m+1,n} \eta^*) d\eta^* \\
 &= N_{m+1,n} (A_{m+1,n} \cosh(\lambda_{m+1,n} \gamma_m) + B_{m+1,n} \sinh(\lambda_{m+1,n} \gamma_m)) \\
 &- B_{m,0} \bar{k}_m \int_0^{\bar{w}_{m+1}} \bar{g}_m(\eta^*) \cos(\lambda_{m+1,n} \eta^*) d\eta^* \\
 &- \sum_{n=1}^N \bar{k}_m \lambda_{m,n} (A_{m,n} \sinh(\lambda_{m,n} \gamma_m) + B_{m,n} \cosh(\lambda_{m,n} \gamma_m)) \int_0^{\bar{w}_{m+1}} \bar{g}_m(\eta^*) \cos(\lambda_{m,n} \eta^*) \cos(\lambda_{m+1,n} \eta^*) d\eta^*
 \end{aligned} \tag{31}$$

where Eq. (31) is written for  $n = 1, 2, \dots, N$ .

At each interface, either Eqs. (23), (24), (26) and (27) (Case A) or Eqs. (23), (29)–(31) (Case B) apply. Additionally, Eqs. (16), (17), (19) and (20) also apply. Taken together, these constitute a total of  $2(N+1)M$  linear algebraic equations for the entire  $M$ -layer problem, which are sufficient for determining the  $2(N+1)M$  unknown coefficients  $A_{m,0}$ ,  $B_{m,0}$ ,  $A_{m,n}$  and  $B_{m,n}$  ( $m = 1, 2, \dots, M; n = 1, 2, \dots, N$ ). The solution for the temperature distribution is complete once the unknown coefficients are determined by solving this set of linear algebraic equations, for example, by matrix inversion. The final temperature distribution in each layer in non-dimensional form is given by Eq. (14).

### 3. Results and discussion

This section discusses key results based on the theoretical model presented above. Particular emphasis is given to a two-layer body, as it appears commonly in vertically integrated semiconductor technology and other applications.

#### 3.1. Convergence analysis

The analytical solution derived in the previous section is in the form of an eigenfunction-based infinite series. In practice, only a finite number of terms can be computed. Specifically, the technique used in the present work is based on truncating the infinite series in Eq. (14) to a

finite number of terms,  $N$ , and then determining the coefficients through a set of linear algebraic equations. Therefore, it is important to investigate how the value of  $N$  affects the accuracy of the computed temperature distribution. For a representative two-layer body with layer 2 wider than layer 1 (Case A), Fig. 3(a) plots temperature as a function of  $\xi$  across both layers at  $\eta = 1$ , and Fig. 3(b) plots temperature as a function of  $\eta$  at the heat flux face,  $\xi = 0$ . In each case, the temperature distribution is computed and plotted for different values of  $N$ . The problem parameters are  $\bar{w}_1 = 2$ ,  $\bar{w}_2 = 4$ ,  $\bar{k}_1 = 4$ ,  $\gamma_1 = 0.5$ ,  $\bar{q} = 2$ ,  $\bar{g}_1 = 0.1$ ,  $Bi = 1$ . The vertical dotted line in Fig. 3(a) represents the interface between layers 1 and 2. These plots indicate convergence of the temperature field at around  $N = 15$ , with negligible difference between  $N = 15$  and  $N = 20$  curves. Therefore, for all subsequent analysis presented in this work, a total of 15 terms are considered for computation. It is found that a similar convergence criterion also holds for Case B. Note that the discontinuity in the temperature profiles in Fig. 3(a) is due to the presence of non-zero thermal contact resistance between the layers.

#### 3.2. Numerical validation

Further investigation of the accuracy of the analytical model presented here is carried out through comparison with numerical simulations based on the finite-element method carried out in ANSYS Fluent. The geometry of a two-layer body is generated and meshed. Boundary conditions and interface thermal resistance is applied, consistent with

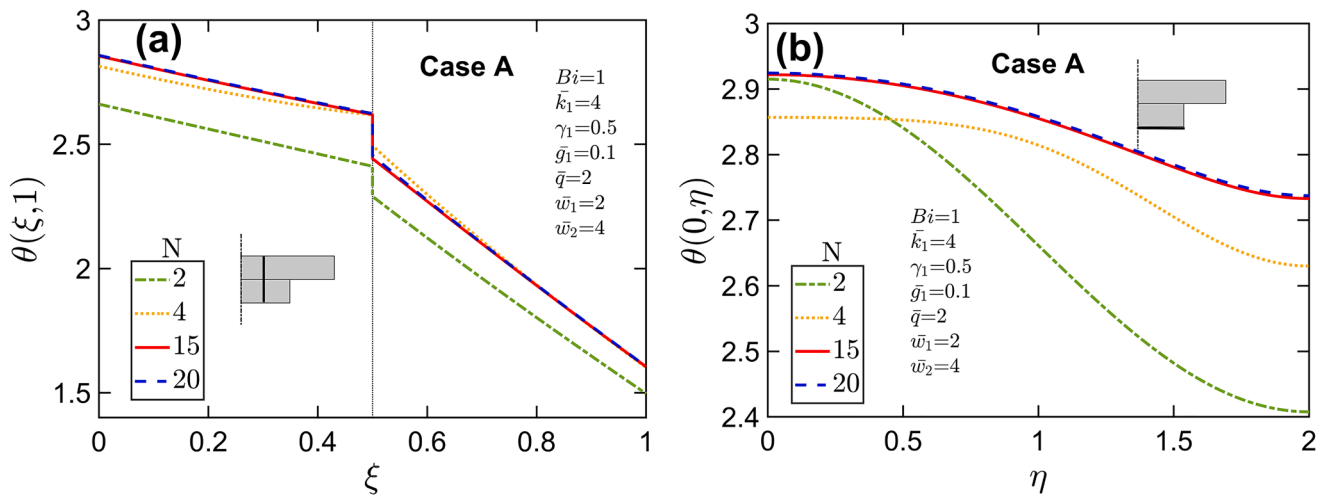
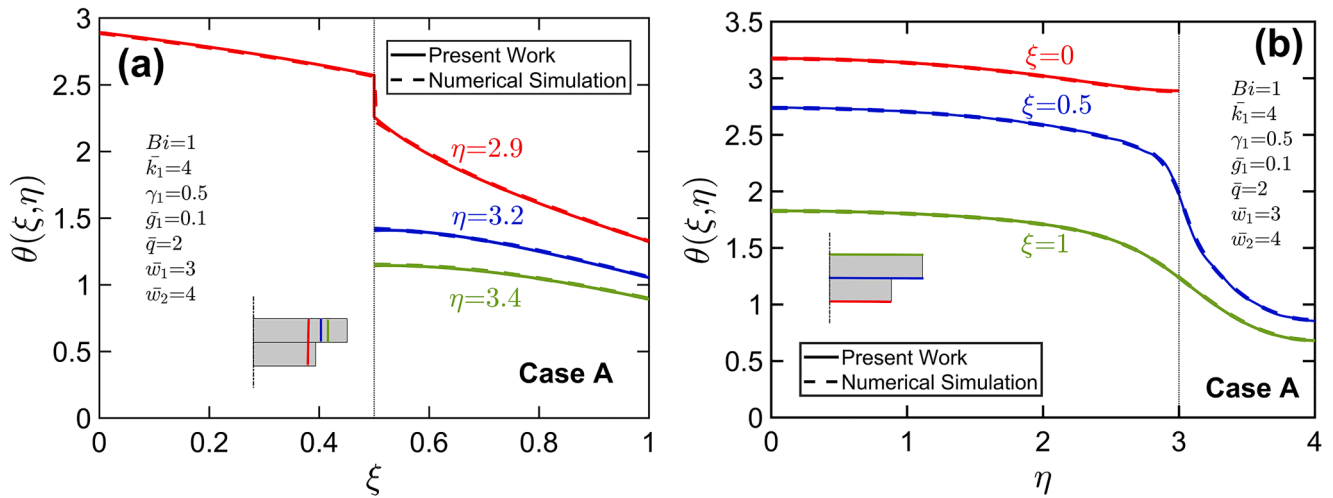


Fig. 3. Effect of number of terms on temperature distribution for a two-layer body with Case A geometry: (a)  $\theta$  as a function of  $\xi$  at  $\eta = \bar{w}_1/2$ ; (b)  $\theta$  as a function of  $\eta$  at  $\xi = 0$ . Curves are plotted for multiple number of eigenvalues. Problem parameters are  $\bar{w}_1 = 2$ ,  $\bar{w}_2 = 4$ ,  $\bar{k}_1 = 4$ ,  $\gamma_1 = 0.5$ ,  $\bar{q} = 2$ ,  $\bar{g}_1 = 0.1$ ,  $Bi = 1$ .





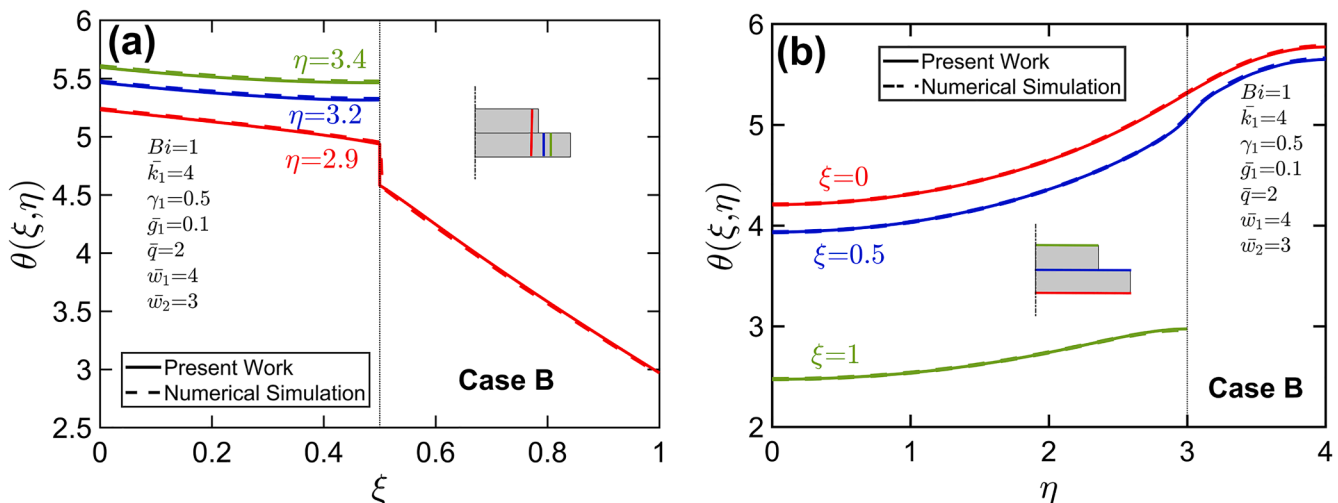
**Fig. 4.** Comparison with numerical simulations for a two-layer body with Case A geometry: (a)  $\theta$  as a function of  $\xi$  at  $\eta = 2.9, 3.2$  and  $3.4$ , (b)  $\theta$  as a function of  $\eta$  at  $\xi = 0, 0.5$  and  $1$ . Problem parameters are  $\bar{w}_1 = 3, \bar{w}_2 = 4, \bar{k}_1 = 4, \gamma_1 = 0.5, \bar{q} = 2, \bar{g}_1 = 0.1, Bi = 1$ .

the problem considered here. Grid sensitivity of the numerical simulation is established by progressively refining the grid until further refinement does not result in significant change in the computed temperature distribution. Based on mesh sensitivity results summarized in Supplementary Information, a total of around 3500 element are used for this 2D simulation.

Since the present work considers two distinct cases, Figs. 4 and 5 present results for Case A ( $\bar{w}_1 = 3, \bar{w}_2 = 4$ ) and Case B ( $\bar{w}_1 = 4, \bar{w}_2 = 3$ ), respectively. Both Figs. 4(a) and 5(a) plot temperature as a function of  $\xi$  at  $\eta = 2.9, 3.2$  and  $3.4$ , with the distinction that in Fig. 4(a), heat flows from the narrower layer into the wider layer and vice versa in Fig. 5(a). In each plot, the vertical dotted line shows the interface between the layers. Similarly, Figs. 4(b) and 5(b) plot temperature as a function of  $\eta$  at  $\xi = 0, 0.5$  and  $1$  for Cases A and B, respectively. The vertical dotted line in these plots represents the side wall of layer 1 (Fig. 4(b)) and the side wall of layer 2 (Fig. 5(b)). All problem parameters other than layer widths are identical to Fig. 3. Both Figs. 4 and 5 demonstrate excellent agreement between numerical simulations and the present work, with a worst-case deviation of less than 0.1% between the two.

In addition to consistency with numerical simulations, Figs. 4 and 5 also illustrate several interesting features of the temperature distribution for Cases A and B. Figs. 4(a) and 5(a) show that temperature profiles at  $\eta = 3.2$  and  $\eta = 3.4$  are both flat at  $\xi = 0.5$ , which is consistent with the adiabatic boundary condition imposed on the extended portions of layer 2 in Case A and layer 1 in Case B. In Fig. 5(a), the gradual reduction in temperature with increasing  $\xi$  at  $\eta = 3.2$  and  $\eta = 3.4$  is a consequence of two-dimensional heat flow from the extended portion into layer 2. On the other hand, in Fig. 4(a), the curves are initially quite flat and begin to drop steeply as one approaches the convective cooling boundary. The temperature profile at  $\eta = 2.9$  in Figs. 4(a) and 5(a) has a discontinuity at  $\xi = 0.5$ , which is a result of the non-zero interfacial thermal contact resistance,  $\bar{g}_1 = 0.1$  in this case. All curves in Figs. 4(b) and 5(b) are flat at the large  $\eta$  end due to the imposed adiabatic boundary conditions imposed at the side walls. In Fig. 4(b), along the heat flux face, the temperature drops towards the right side wall, whereas in Fig. 5(b), a significant rise in temperature is observed close to the right side wall. This rise in temperature in Fig. 5(b) is due to an increase in the inflow of heat from the extended portion of layer 1 and due to being further away from the layer 2 heat sink. On the other hand, the reduction in temperature close to  $\eta = 3$  seen in Fig. 4(b) is a result of the right side wall of layer 1 being closer to the extended portion of layer 2 that is at a lower temperature. In Fig. 4(b), the curve at  $\xi = 0.5$  is plotted initially along the interface and extends along the bottom surface of layer 2 beyond the side wall of layer 1. In contrast, the curve in Fig. 5(b) is plotted along the

side wall of layer 1. In contrast, the curve in Fig. 5(b) is plotted along the



**Fig. 5.** Comparison with numerical simulations for a two-layer body with Case B geometry: (a)  $\theta$  as a function of  $\xi$  at  $\xi = 2.9, 3.2$  and  $3.4$ , (b)  $\theta$  as a function of  $\eta$  at  $\xi = 0, 0.5$  and  $1$ . Problem parameters are  $\bar{w}_1 = 4, \bar{w}_2 = 3, \bar{k}_1 = 4, \gamma_1 = 0.5, \bar{q} = 2, \bar{g}_1 = 0.1, Bi = 1$ .



interface and then the top surface of layer 1. In Fig. 4(b), at  $\xi = 0.5$  and 1, the temperature initially reduces slowly, and then much faster as one gets closer to the right sidewall of layer 2. This is because of lower penetration of heat in the  $\eta$  direction towards the right sidewall, which is a result of higher thermal resistance along the  $\eta$ -direction in the extended portion of layer 2 compared to the thermal resistance in the  $\xi$ -direction. In Fig. 5(b), the opposite effect is observed at  $\xi = 0$  and 0.5 due to the reversed geometry of Case B compared to Case A. In the curve corresponding to  $\xi = 1$ , the temperature rise is due to penetration of additional heat from the extended portion of layer 1.

### 3.3. Comparison with past work

In addition to numerical comparison, it is also instructive to compare results from the present work with past papers for special cases. Specifically, results are compared with a past paper [27], in which, an analytical model was presented for multilayer heat transfer with layers of equal widths and with spatially varying thermal contact resistance between layers. In this past work, two-dimensionality of the temperature field arises from spatial variation in the imposed heat flux and the spatially varying thermal contact resistance, whereas the present work analyzes a more general case where the layers do not have the same width. For comparison between the two, Fig. 6(a) plots temperature as a function of  $\xi$  at  $\eta = 4.5$  for multiple values of  $\bar{w}_2$  for a two-layer body with Case A geometry, while Fig. 6(b) plots temperature as a function of  $\bar{w}_2$  at two locations  $L_1$  ( $\xi = 0, \eta = 0$ ) and  $L_2$  ( $\xi = 0.6, \eta = 0$ ) for Case B geometry.  $\bar{w}_1 = 5$  in both Figures. Heat flux  $\bar{q} = 2.0$  and thermal contact resistance  $\bar{g} = 0.1$  are both considered to be uniform. All other problem parameter values are the same as Fig. 3. Fig. 6(a) clearly shows that temperature profile from the present work aligns well with result from past work [27] as  $\bar{w}_2$  reduces and approaches the value of  $\bar{w}_1$ . For larger values of  $\bar{w}_2$ , the temperature distribution is lower than the equal width case due to extra heat dissipation through the extended portion of layer 2. Similarly, Fig. 6(b) shows that as  $\bar{w}_2$  increases and approaches the value of  $\bar{w}_1$ , temperature computed by the present model for unequal widths reduces and approaches the value from past work on equal width layer [27]. Good agreement is observed for both in the two layers.

### 3.4. Typical temperature colormaps

Fig. 7 presents typical temperature colormaps for three different problems in a two-layer body. Fig. 7(a) and 7(b) pertain to Cases A and B, respectively, with parameter values identical to Figs. 4 and 5,

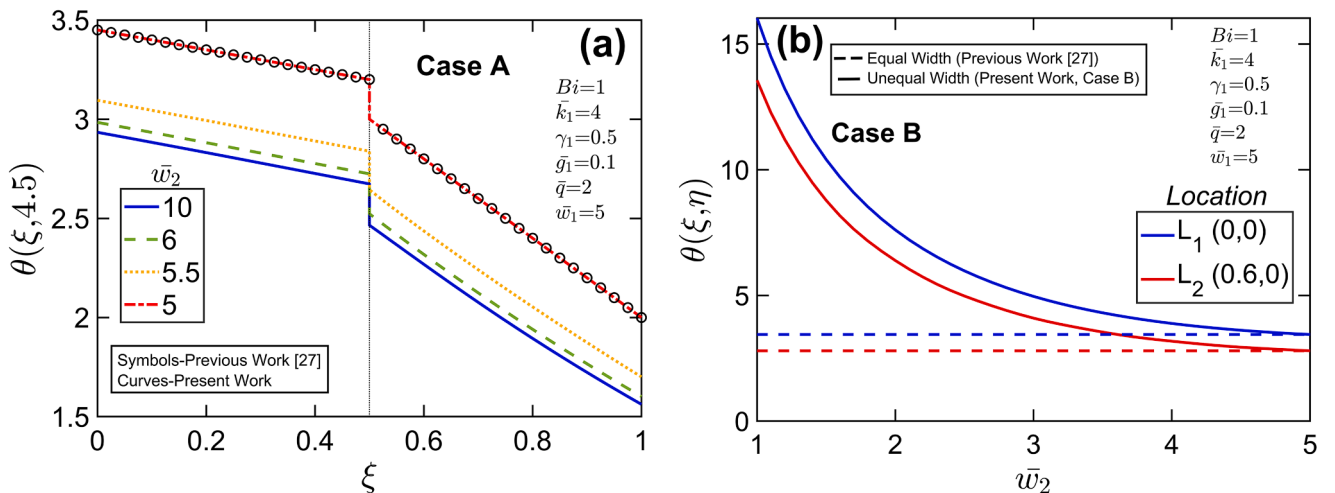


Fig. 6. Comparison with past work: (a)  $\theta$  as a function of  $\xi$  at  $\eta = \bar{w}_1/2$  for multiple values of  $\bar{w}_2$ ; (b)  $\theta$  as a function of  $\bar{w}_2$  at two locations  $L_1$  ( $\xi = 0, \eta = 0$ ) and  $L_2$  ( $\xi = 0.6, \eta = 0$ ). Problem parameters are  $\bar{w}_1 = 5$ ,  $\bar{k}_1 = 4$ ,  $\gamma_1 = 0.5$ ,  $\bar{q} = 2$ ,  $\bar{g}_1 = 0.1$ ,  $Bi = 1$ . Values from past work on equally sized layers [27] are shown for comparison.

respectively. Fig. 7(c) considers the equal width case,  $\bar{w}_1 = \bar{w}_2 = 4$ . The maximum temperature observed in Fig. 7(a) is lower than in Fig. 7(b) due to greater cooling through the extended portion of layer 2, as well as an increased convective heat transfer from the boundary along the extended portion. In contrast, there is greater heat constriction in Fig. 7(b) and an increased heat influx caused by a wider layer 1. The location of the peak temperature is also different in the two cases, as expected. The peak temperature in Fig. 7(a) occurs at the origin, whereas in Fig. 7(b), the peak temperature occurs at the right-most corner of the heat flux boundary in layer 1, which is further away from the sink layer (layer 2) compared to the origin. This may explain the shift in the maximum temperature location. Compared to the unequal width cases A and B shown in Fig. 7(a) and 7(b), respectively, heat transfer becomes completely one-dimensional when  $\bar{w}_1$  is equal to  $\bar{w}_2$ , as shown in Fig. 7(c). This is expected in such a case, because of no temperature gradient to drive heat flow in the lateral direction when both layers are of equal width and the applied heat flux and contact resistance are uniform as well. In this case, the peak temperature occurs all along the heat flux face. Comparing Fig. 7(b) and 7(c), for the same amount of heat influx, the peak temperature observed in Fig. 7(c) is lower because of additional cooling as a result of an increase in the size of the heat sink layer.

In order to illustrate the technique developed here for more complicated geometries, two representative three-layer problems are also solved, and the temperature colorplots are presented in Fig. 8(a) and 8(b). In the first problem,  $\bar{w}_1 = 2$ ,  $\bar{w}_2 = 4$ ,  $\bar{w}_3 = 2$ , so that Cases A and B apply at the interface between layers 1 and 2, and between layers 2 and 3, respectively. On the other hand, in the second problem,  $\bar{w}_1 = 2$ ,  $\bar{w}_2 = 4$ ,  $\bar{w}_3 = 6$ , so that Case A applies at both interfaces. The interface contact resistance is assumed to be spatially varying, as defined in the caption of Fig. 8. Comparing the two scenarios plotted in Fig. 8, the second one has lower peak temperature, which may be explained on the basis of greater area of the cooling boundary due to the larger size of layer 3, which prevails over the additional thermal spreading resistance that may be encountered. In both Fig. 8(a) and 8(b), heat penetration into layer 2 may be observed in the lower thermal contact region. In Fig. 8(b), due to different widths, one of the lower thermal contact resistance regions between layers 2 and 3 is directly aligned with the only lower thermal contact resistance region between layers 1 and 2. The relative size of this region is also higher because of different widths. This results in most of the heat from layer 1 flowing directly into layer 3 through this region. On the other hand, in Fig. 8(a), the higher thermal contact resistance region between layers 2 and 3 is directly aligned with the lower thermal contact resistance region between layers 1 and 2. As

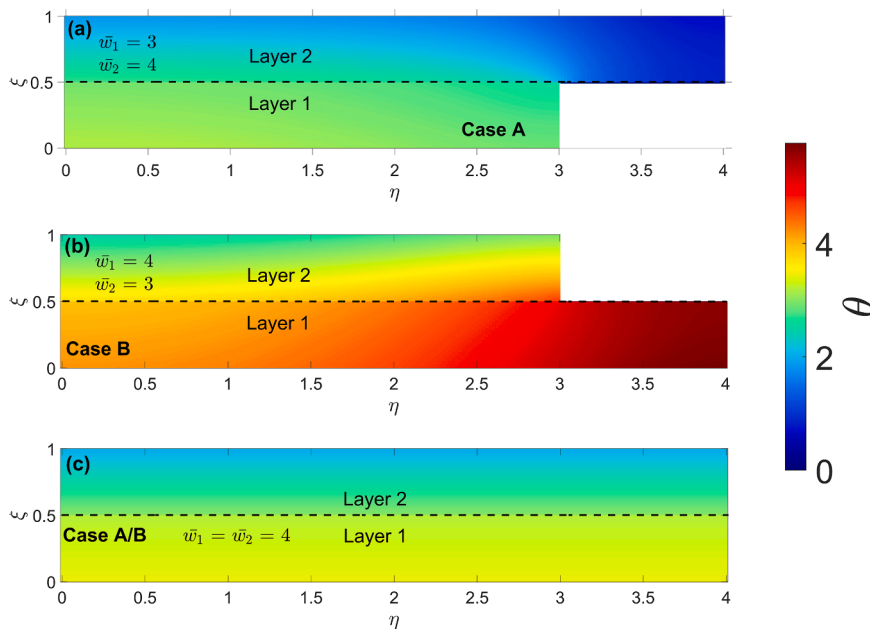


Fig. 7. A representative colorplot of the temperature field in a two-layer body: (a)  $\bar{w}_1 = 3, \bar{w}_2 = 4$  (Case A); (b)  $\bar{w}_1 = 4, \bar{w}_2 = 3$  (Case B); (c)  $\bar{w}_1 = \bar{w}_2 = 4$  (uniform width). Other problem parameters are identical to Fig. 4.

expected, greater heat penetration into layer 3 is observed in the lower thermal contact resistance region towards the left.

3.5. Effect of unequal layer widths

The effect of unequal layer widths  $\bar{w}_1$  and  $\bar{w}_2$  is studied further in Figs. 9 and 10. Fig. 9(a) and 9(b) plot temperature as a function of  $\eta$  at the heat flux face for multiple values of  $\bar{w}_2$  and  $\bar{w}_1$ , respectively, in a two-layer body. In Fig. 9(a),  $\bar{w}_1 = 1$ , so that all curves correspond to a two-layer body with Case A geometry, whereas in Fig. 9(b),  $\bar{w}_2 = 1$ , so that all curves correspond to a two-layer body with Case B geometry. All other problem parameters are identical to the ones used in Fig. 3. In Fig. 9(a), as  $\bar{w}_2$  increases, temperature across the heat flux face reduces by a moderately small amount. This effect can be attributed to greater cooling as the extended portion of layer 2 increases with increasing  $\bar{w}_2$ . Beyond a certain value of  $\bar{w}_2$ , however, this effect is expected to diminish as a result of an increase in the spreading resistance in layer 2 as well as an increase in thermal resistance along the  $\eta$  direction towards the

sidewall. This diminishing effect can be clearly observed from the curves for  $\bar{w}_2 = 4$  and 5, which are found to be very close to each other. On the other hand, in Fig. 9(b), the overall temperature across the heat flux face goes up much more strongly as  $\bar{w}_1$  is increased. This is attributable to an increase in the total amount of heat entering the system, given that the heat flux is fixed, while the size of the cooling boundary remaining unchanged. Unlike Fig. 9(a), no saturation effect is expected in this case because, as seen in Fig. 9(b), the peak temperature goes up rapidly as the spreading resistance in layer 1 increases with increasing  $\bar{w}_1$ , in addition to the factors mentioned above.

These effects are further illustrated by temperature colormaps in Fig. 10 for some of the cases considered above. Fig. 10(a) and 10(b) correspond to Case A, whereas Fig. 10(c) and 10(d) are associated with Case B. Between Fig. 10(a) and 10(b), a lower heat penetration in the extended portion of layer 2 is observed as  $\bar{w}_2$  is increased, which is due to an increase in the spreading resistance. Between Fig. 10(c) and 10(d), layer 1 gets hotter as  $\bar{w}_1$  is increased, especially in the extended portion of layer 1, which is due to an increase in the total heat entering the

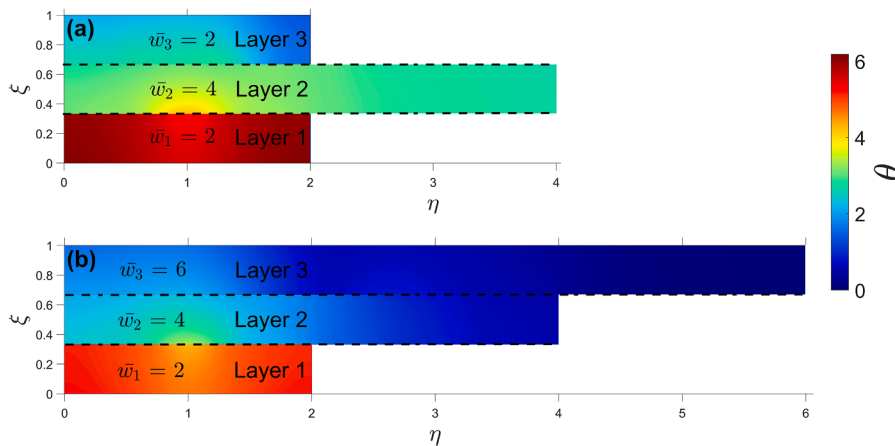


Fig. 8. Representative colorplots of the temperature field in a three-layer body: (a)  $\bar{w}_1 = 2, \bar{w}_2 = 4, \bar{w}_3 = 6$ ; (b)  $\bar{w}_1 = 2, \bar{w}_2 = 4, \bar{w}_3 = 2$ . Other problem parameters are  $\bar{k}_1 = 4, \bar{k}_2 = 2, \gamma_1 = 0.25, \gamma_2 = 0.75, \bar{q} = 2, Bi = 1$ . Further,  $\bar{g}_1(\eta) = 0.1$  between  $0.4\bar{w}_1$  and  $0.6\bar{w}_1$ , and 5 everywhere else. For (a),  $\bar{g}_2(\eta) = 10$  between  $0.4\bar{w}_3$  and  $0.6\bar{w}_3$ , and 0.1 elsewhere. For (b),  $\bar{g}_2(\eta) = 10$  between  $0.4\bar{w}_2$  and  $0.6\bar{w}_2$ , and 0.1 elsewhere.

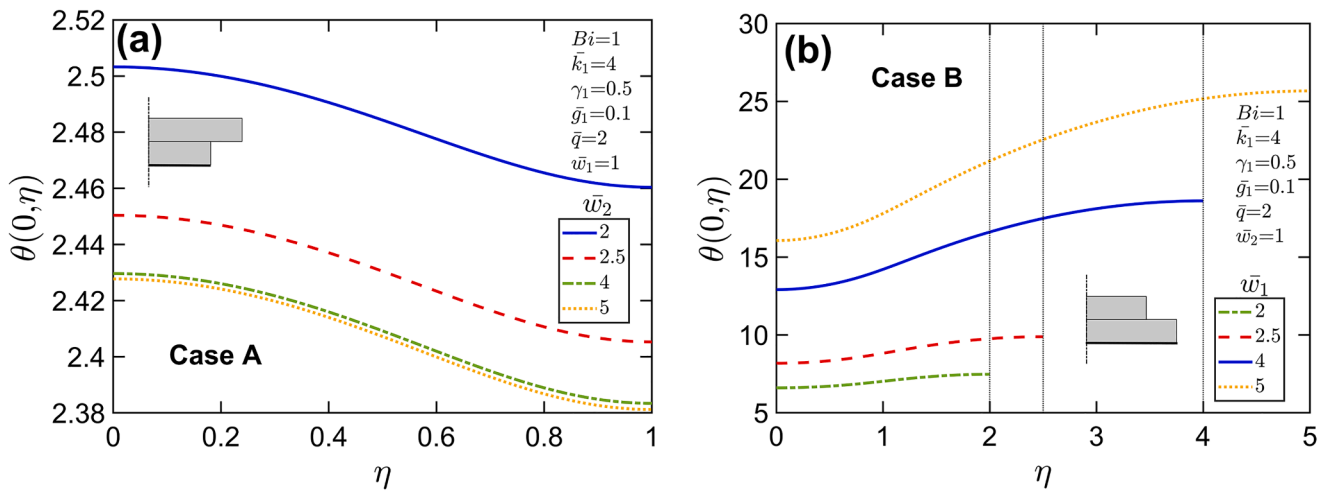


Fig. 9. Effect of  $\bar{w}_1$  and  $\bar{w}_2$  in a two-layer body: (a)  $\theta$  as a function of  $\eta$  at  $\xi=0$  for  $\bar{w}_1 = 1$  and multiple values of  $\bar{w}_2$  (Case A); (a)  $\theta$  as a function of  $\eta$  at  $\xi=0$  for  $\bar{w}_2 = 1$  and multiple values of  $\bar{w}_1$  (Case B). Other problem parameters are the same as Fig. 3.

system and spreading resistance of, while the heat sink size remains constant.

A key interpretation of results discussed in this sub-section is that thermal considerations should be taken into account in multilayer chip design in terms of the sizes of each die. While using die of unequal sizes may offer design flexibility, this must be weighed against potential deterioration in thermal performance, and hence reliability. For example, Fig. 9 quantifies the extent of temperature rise to be expected for a number of different layer widths.

3.6. Effect of thermal contact resistance

The effect of thermal contact resistance is presented next in Fig. 11 for Case A, assuming  $\bar{w}_1 = 2$  and  $\bar{w}_2 = 4$ . Fig. 11(a) plots temperature as a function of  $\xi$  across both layers at  $\eta = 1$ , while Fig. 11(b) plots temperature as function of  $\eta$  in layer 2 at  $\xi = 0.6$ , very close to the interface. In each case, multiple curves corresponding to different values of the inter-layer thermal resistance  $\bar{g}_1$  are plotted, including the  $\bar{g}_1 = 0$  perfect thermal contact case. All other problem parameters are the same as Fig. 3. As expected, Fig. 11(a) clearly shows that the temperature of layer 1 goes up as the contact resistance magnitude is increased. On the other hand, the temperature in layer 2 has a much weaker dependence on  $\bar{g}_1$ , as seen in the curves between  $\xi = \gamma_1$  and  $\xi = 1$  in Fig. 11(a). In

Fig. 11(b), similar to the observation for layer 2 in Fig. 11(a), the temperature profiles show a weak dependence on  $\bar{g}_1$ , which diminishes further in the extended portion of layer 2.

A key interpretation of results discussed in this sub-section is that thermal contact resistance, which is well known to occur in vertically integrated semiconductor devices may adversely affect temperature [13, 26]. Note that  $\bar{g}_m(\eta) = \frac{R_m(\eta)k_M}{z_M}$  for a Silicon chip of thickness of the order of 1 mm,  $\bar{g}_1$  of the order of 0.5 as considered in Fig. 11 corresponds to a thermal contact resistance of around  $3 \times 10^{-6} \text{ Km}^2/\text{W}$ , which is close to the reported value for vertically integrated semiconductor devices [26].

3.7. Effect of Biot number

Biot number is the key non-dimensional parameter that governs boundary cooling. The impact of  $Bi$  is investigated in Fig. 12. Fig. 12(a) plots temperature as a function of  $\xi$  at  $\eta = \bar{w}_1/2$  for multiple values of  $Bi$ . On the other hand, Fig. 12(b) plots the maximum temperature as a function of  $Bi$  for multiple values of  $\bar{k}_1$ . Other problem parameters in this Figure are identical to Fig. 3. As expected, Fig. 12(a) shows a reduction in temperature distribution with increasing  $Bi$ , which may be attributed to more effective boundary cooling. This effect is also observed in Fig. 12 (b), in which, the observed peak temperature drops as  $Bi$  increases. This is a particularly pronounced effect at small  $Bi$ , where even minor

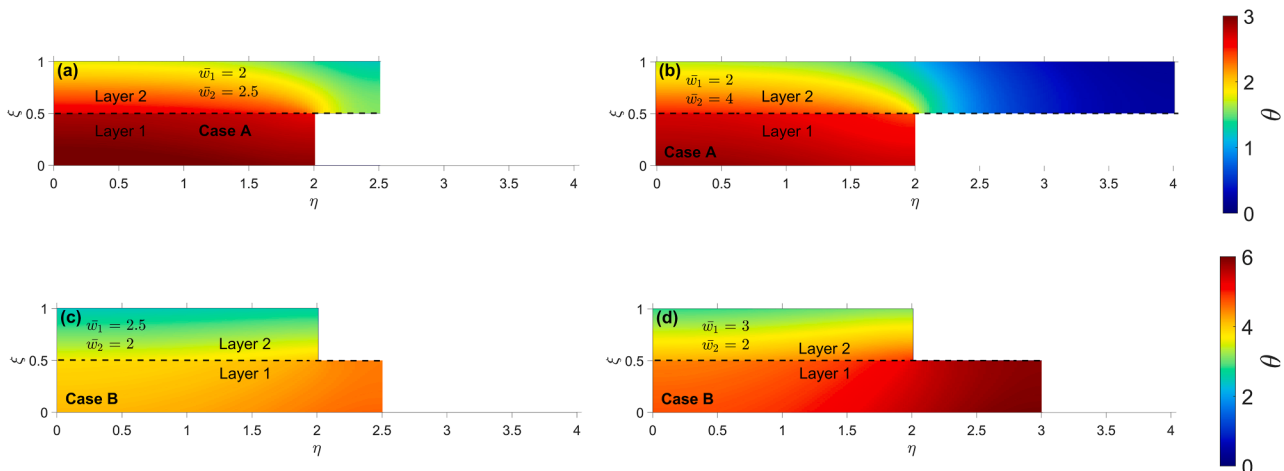


Fig. 10. Temperature colorplots in a two-layer body: (a)  $\bar{w}_1 = 2$  and  $\bar{w}_2 = 2.5$  (Case A); (b)  $\bar{w}_1 = 2$  and  $\bar{w}_2 = 4$  (Case A); (c)  $\bar{w}_1 = 2.5$  and  $\bar{w}_2 = 2$  (Case B); (c)  $\bar{w}_1 = 3$  and  $\bar{w}_2 = 2$  (Case B). Problem parameters are  $\bar{k}_1 = 4$ ,  $\gamma_1 = 0.5$ ,  $\bar{q} = 2$ ,  $\bar{g}_1 = 0.1$ ,  $Bi = 1$ .

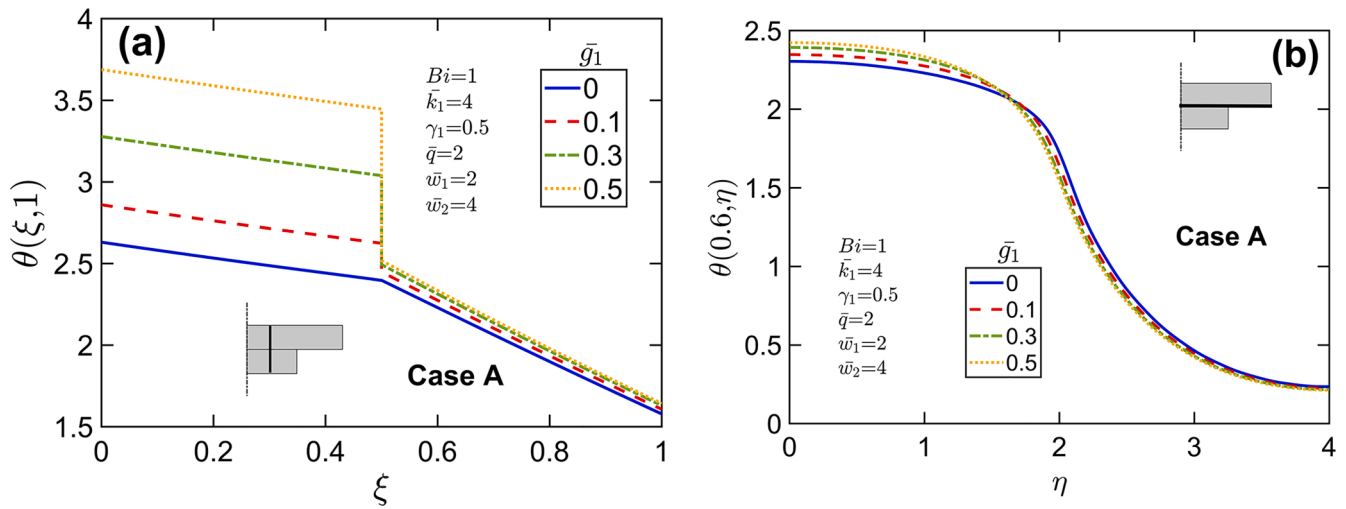


Fig. 11. Effect of  $\bar{g}_1$  in a two-layer body: (a)  $\theta$  as a function of  $\xi$  at  $\eta = \bar{w}_1/2$  for  $\bar{g}_1 = 0, 0.1, 0.3$  and  $0.5$  (b)  $\theta$  as a function of  $\eta$  at  $\xi = 0.6$  for  $\bar{g}_1 = 0, 0.1, 0.3$  and  $0.5$ . Problem parameters are  $\bar{w}_1 = 2, \bar{w}_2 = 4, \bar{k}_1 = 4, \gamma_1 = 0.5, \bar{q} = 2, Bi = 1$ , resulting in Case A geometry.

improvement in boundary cooling results in significant improvement in peak temperature. The saturation in peak temperature at large  $Bi$  is indicative of isothermal conditions, wherein further increase in  $Bi$  does not bring about significant incremental improvement. Fig. 12(b) also shows an important impact of  $\bar{k}_1$ , wherein the peak temperature is higher for lower values of  $\bar{k}_1$ . This is mainly due to increased thermal resistance within the two-layer body as  $\bar{k}_1$  decreases.

From a physical perspective, a Biot number of 1.0 corresponds to a convective heat coefficient of around  $1000 \text{ Wm}^{-2}\text{K}^{-1}$  for a glass interposer chip of 1 mm total thickness. Such a convective heat transfer coefficient is realistically obtainable from aggressive air or liquid cooling methods. Depending on the specifics of a problem of interest, the non-dimensional results in Fig. 12 may be converted to dimensional quantities.

### 3.8. Illustration of a practical problem

The capability of the present model to solve practical problems of interest is demonstrated next. This analysis also identifies an interesting and practically relevant impact of the two-dimensional nature of heat flow in this problem. A two-layer vertically integrated semiconductor device comprising Si and SiGe substrates is considered. Two scenarios

are considered – in the first scenario, layer 1 (next to the heat flux face) is assumed to be SiGe, whereas layer 2 (next to the heat dissipation face) is assumed to be Si. In the second scenario, the two materials are reversed. Thermal conductivities of Si and SiGe are assumed to be  $150 \text{ W/mK}$  and  $4.6 \text{ W/mK}$ , respectively [7]. In each scenario, the second layer is assumed to have a fixed thickness and width of  $z_2 = 1 \text{ mm}$  and  $w_2 = 15 \text{ mm}$ , respectively. Further, values of the total heat, convective heat transfer coefficient and ambient temperature are assumed to be  $7 \text{ W}, 2000 \text{ W/m}^2\text{K}$  and  $298 \text{ K}$ , respectively. The model presented in the previous section is used to investigate the impact of changing layer 1 thickness  $z_1$  on peak temperature. This is of much practical relevance, since one of the layers in a vertically integrated semiconductor device, usually the smaller-width layer, is often thinned down significantly in order to accommodate high-density TSVs of very small diameters [13]. In doing so, it is important to understand the impact of such thinning on temperature rise. Fig. 13(a) plots peak temperature as a function of layer 1 width, expressed as the ratio  $w_1/w_2$  for the first scenario. A similar plot for the second scenario is presented in Fig. 13(b). Both plots are presented for multiple values of the first layer thickness  $z_1$ , down to the likely limits of die thinning in vertically integrated semiconductor technology [13]. In each case, circles represent results from past work based on single-layer heat transfer analysis [24], in which, only heat

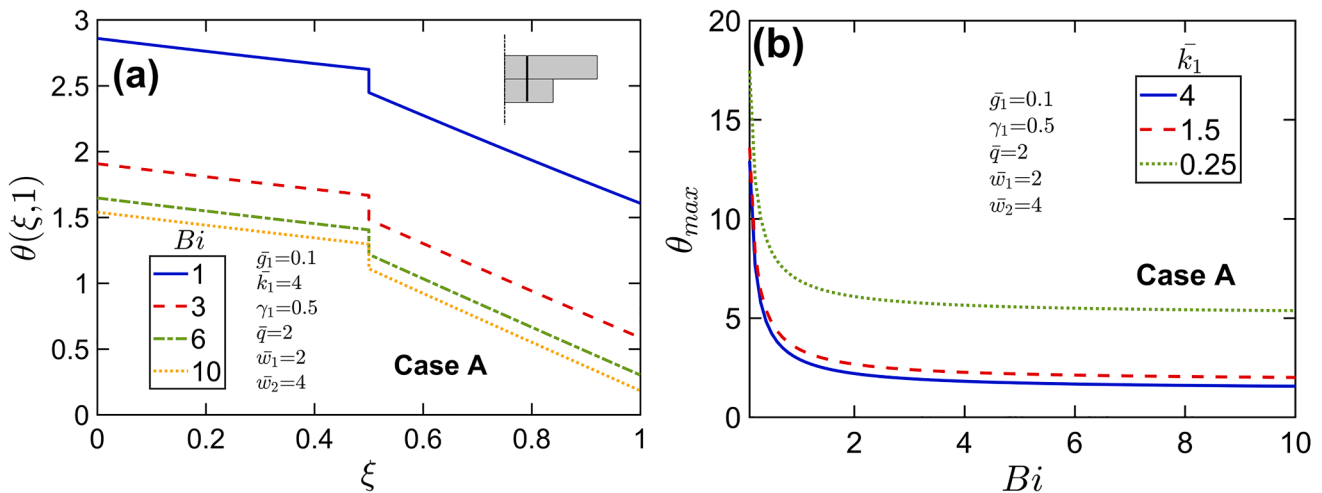
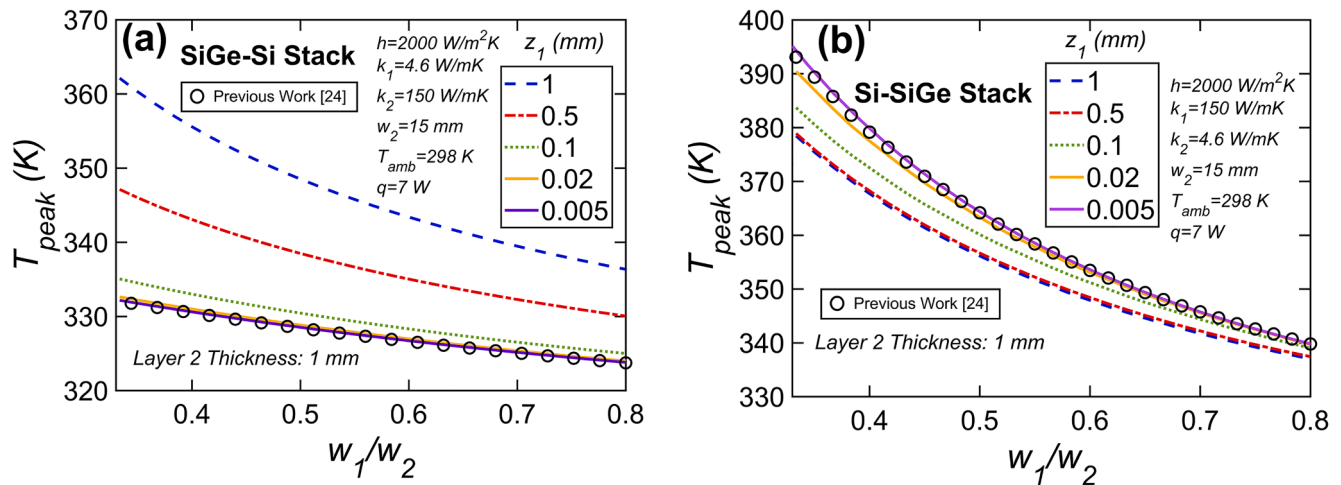


Fig. 12. Effect of  $Bi$  in a two-layer body: (a)  $\theta$  as a function of  $\xi$  at  $\eta = \bar{w}_1/2$  for  $Bi = 1, 3, 6$  and  $10$  for  $\bar{k}_1 = 4$  (b)  $\theta_{max}$  as a function of  $Bi$  for  $\bar{k}_1 = 4, 1.5$ . Problem parameters are  $\bar{w}_1 = 2, \bar{w}_2 = 4, \gamma_1 = 0.5, \bar{q} = 2, \bar{g}_1 = 0.1$ , resulting in Case A geometry.



**Fig. 13.** Peak temperature of a two-layer 3D IC stack comprising Si and SiGe substrates as a function of  $w_1/w_2$  for multiple values of  $z_1$  with constant total heat and other parameters: (a) SiGe substrate next to the heat source (layer 1), (b) Si substrate next to the heat source (layer 1). Results from single-layer analysis [24] are presented for comparison.

transfer in the second layer is accounted for, and the thickness of the first layer is ignored, i.e., heat flux is applied directly on a width  $w_1$  on the bottom face of the second layer. Curves in both Fig. 13(a) and 13(b) show, as expected, that as  $w_1$  decreases, i.e., when there is greater mismatch in layer widths, the peak temperature goes up. This is expected due to concentration of heat combined with increased two-dimensional heat flow and subsequent thermal spreading resistance as heat flows through the two layers. As the thickness  $z_1$  becomes smaller and smaller, as would occur during die thinning for ultra-fine TSV processing, it is found that the curves in Fig. 13(a) and 13(b) both converge, and approach single-layer results from past work [24]. For die thickness of 20  $\mu\text{m}$  or lower, Fig. 13(a) and 13(b) show that the simplified single-layer analysis from past work [24] is sufficient, but for larger layer 1 thicknesses, the more detailed analysis presented here is necessary for accurate temperature prediction.

An interesting aspect of the impact of layer 1 thickness  $z_1$  on peak temperature shown for the two scenarios is that when layer 1 comprises the lower thermal conductivity material SiGe, peak temperature goes up as  $z_1$  increases, as shown in Fig. 13(a). This is because an increase in  $z_1$  results in an increase in thermal resistance offered by layer 1, and, therefore, greater peak temperature. This is a particularly strong effect when layer 1 thermal resistance is relatively larger, and, therefore, the rate-limiting step. In contrast, when layer 1 comprises the higher thermal conductivity material Si, peak temperature actually decreases as  $z_1$  increases, as shown in Fig. 13(b). This somewhat counter-intuitive effect is explained on the basis of the two-dimensional nature of heat flow and the unequal widths of the two layers. Mainly, an increase in  $z_1$  improves lateral thermal resistance in layer 1 while increasing thermal resistance in the  $z$  direction. This allows greater heat flow in the  $y$  direction, which is eventually able to access the convective cooling conditions along the  $z = z_2$  boundary of the larger-width layer 2. This impact is important only in the second scenario (Fig. 13(b)) when layer has high thermal conductivity because in the opposite scenario (Fig. 13(a)), this effect is overwhelmed by the impact of increased  $z$ -direction thermal resistance due to the rate-limiting nature of layer 1 thermal resistance.

The opposite effects of changing layer thickness on temperature rise depending on the order of Si and SiGe substrates is an important insight that must be carefully considered in the practical design of such vertically integrated semiconductor devices.

#### 4. Conclusions

The key novelty of the theoretical model presented here is in

accounting for the impact of unequal widths on heat transfer in a multi-layer body with emphasis on the commonly encountered two-layer case. Such an architecture is relevant for advanced semiconductors such as vertically integrated devices, in which, unequally-sized chips are often integrated with each other to maximize design flexibility. Thermal modeling of a multilayer body with unequal widths is not straightforward. This challenge is addressed here through a series solution for the temperature field in each layer, and the unequal widths of the two layers are accounted for by deriving a sufficient set of algebraic equations to determine the coefficients of the series solutions. While this work is presented in the context of a two-dimensional body, extension to a three-dimensional body is conceptually straightforward.

Key findings of the present work include quantification of the non-monotonous impact of unequal widths on peak temperature rise in steady state, as well as the contributions of other non-dimensional parameters, such as the Biot number. For a representative two-layer semiconductor chip problem, it is found that the peak temperature may rise by 9% due to the unequal widths of the layers.

A few assumptions/limitations of the work presented here must be recognized. Firstly, the theoretical model is steady-state in nature, and, therefore, transient changes, such as a dynamic heat load are not accounted for. Derivation of the transient temperature field is identified as an important direction for future work. Further, this work assumes constant and uniform thermal properties that do not change with temperature. While temperature-dependence is usually not important for practical semiconductor chips, due to the relatively small temperature change, accounting for spatial variation in thermal properties, for example, to account for TSVs, may be of interest for future work.

Theoretical models such as the one presented here are particularly important for pre-fabrication design of semiconductor chips due to the considerable expense involved in experiments-based design and optimization. This work helps understand the impact of device architecture on temperature rise, and, therefore, may contribute towards the design and optimization of multilayer semiconductor devices with unequal widths.

#### CRedit authorship contribution statement

**Girish Krishnan:** Conceptualization, Formal analysis, Investigation, Data curation, Visualization, Methodology, Writing – original draft, Writing – review & editing. **Ankur Jain:** Conceptualization, Methodology, Supervision, Project administration, Writing – original draft, Writing – review & editing.



## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.ijheatmasstransfer.2023.124708](https://doi.org/10.1016/j.ijheatmasstransfer.2023.124708).

## References

- [1] S.S. Salvi, A. Jain, A review of recent research on heat transfer in three-dimensional integrated circuits (3-D ICs), *IEEE Trans. Compon. Packag. Manuf. Technol.* 11 (2021) 802–821, <https://doi.org/10.1109/TCPMT.2021.3064030>.
- [2] A. Bar-Cohen, Thermal management of microelectronics in the 21st century, in: *Proceeding of the IEEE EPTC*, 1997, pp. 29–33, <https://doi.org/10.1109/EPTC.1997.723880>.
- [3] R. Mahajan, C.P. Chiu, G. Chrysler, Cooling a microprocessor chip, *Proc. IEEE*. 94 (2006) 1476–1485, <https://doi.org/10.1109/JPROC.2006.879800>.
- [4] D.B. Tuckerman, R.F.W. Pease, High-performance heat sinking for VLSI, *IEEE Electron Dev. Lett.* 2 (1981) 126–129, <https://doi.org/10.1109/EDL.1981.25367>.
- [5] M.S. El-Genk, Immersion cooling nucleate boiling of high power computer chips, *Energy Convers. Manag.* 53 (2012) 205–218, <https://doi.org/10.1016/j.enconman.2011.08.008>.
- [6] W. Huang, K. Sankaranarayanan, K. Skadron, R.J. Ribando, M.R. Stan, Accurate, pre-RTL temperature-aware design using a parameterized, geometric thermal model, *IEEE Trans. Comput.* 57 (2008) 1277–1288, <https://doi.org/10.1109/TC.2008.64>.
- [7] G. Krishnan, A. Jain, Transient temperature distribution in a multilayer semiconductor device with dynamic thermal load and non-uniform thermal contact resistance between layers, *Int. J. Heat Mass Transf.* 214 (2023) 1–12, <https://doi.org/10.1016/j.ijheatmasstransfer.2023.124374>, 124374.
- [8] A. Jain, S.M. Alam, S. Pozder, R.E. Jones, Thermal–electrical co-optimisation of floorplanning of three-dimensional integrated circuits under manufacturing and physical design constraints, *IET Comput. Digit. Tech.* 5 (2011) 169–178, <https://doi.org/10.1049/iet-cdt.2009.0107>.
- [9] V. Venkatadri, B. Sammakia, K. Srihari, D. Santos, A review of recent advances in thermal management in three dimensional chip stacks in electronic systems, *J Electron. Packag.* 133 (2011) 1–17, <https://doi.org/10.1115/1.4005298>, 041011.
- [10] S.G. Kandlikar, Review and projections of integrated cooling systems for three-dimensional integrated circuits, *J Electron. Packag.* 136 (2014) 1–11, <https://doi.org/10.1115/1.4027175>, 024001.
- [11] J.H. Lau, T.G. Yue, Thermal management of 3D IC integration with TSV (Through silicon via), in: *Proceedings of the IEEE ECTC*, 2009, pp. 624–640, <https://doi.org/10.1109/ECTC.2009.5074080>.
- [12] S. Pozder, A. Jain, R. Chatterjee, Z. Huang, R.E. Jones, E. Acosta, B. Marlin, G. Hillmann, M. Sobczak, G. Kreindl, S. Kanagavel, H. Kostner, S. Pargfrieder, 3D Die-to-wafer Cu/Sn microconnects formed simultaneously with an adhesive dielectric bond using thermal compression bonding, in: *Proceedings of the IEEE ITC*, 2008, pp. 46–48, <https://doi.org/10.1109/IITC.2008.4546921>.
- [13] E.G. Colgan, P. Andry, B. Dang, J.H. Magerlein, J. Maria, R.J. Polastre, J. Wakil, Measurement of microbump thermal resistance in 3D chip stacks, in: *Proceedings of the IEEE SEMI-THERM*, 2012, pp. 1–7, <https://doi.org/10.1109/STHERM.2012.6188818>.
- [14] W. Huang, S. Ghosh, S. Velusamy, K. Sankaranarayanan, K. Skadron, M.R. Stan, HotSpot: a compact thermal modeling methodology for early-stage VLSI design, *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* 14 (2006) 501–513, <https://doi.org/10.1109/TVLSI.2006.876103>.
- [15] A. Jain, R.E. Jones, R. Chatterjee, S. Pozder, Z. Huang, Analytical and numerical modeling of the thermal performance of three-dimensional integrated circuits, *IEEE Trans. Compon. Packag. Technol.* 33 (2010) 56–63, <https://doi.org/10.1109/TCAPT.2009.2020916>.
- [16] A. Jain, S.M. Alam, S. Pozder, R.E. Jones, Thermal–electrical co-optimisation of floorplanning of three-dimensional integrated circuits under manufacturing and physical design constraints, *IET Comput. Digit. Tech.* 5 (2011) 169–178, <https://doi.org/10.1049/iet-cdt.2009.0107>.
- [17] L. Choobineh, A. Jain, Analytical solution for steady-state and transient temperature field in vertically integrated three-dimensional integrated circuits (3D ICs), *IEEE Trans. Compon. Packag. Manuf. Technol.* 2 (2012) 2031–2039, <https://doi.org/10.1109/TCPMT.2012.2213820>.
- [18] L. Choobineh, A. Jain, An explicit analytical model for rapid computation of temperature field in a three-dimensional integrated circuit (3D IC), *Int. J. Therm. Sci.* 87 (2015) 103–109, <https://doi.org/10.1016/j.ijthermalsci.2014.08.012>.
- [19] L. Choobineh, A. Jain, Determination of temperature distribution in three-dimensional integrated circuits (3D ICs) with unequally-sized die, *Appl. Therm. Eng.* 56 (2013) 176–184, <https://doi.org/10.1016/j.applthermaleng.2013.03.006>.
- [20] J. Cong, G. Luo, Y. Shi, Thermal-aware cell and through-silicon-via Co-placement for 3D ICs, in: *Proceeding of the ACM DAC*, 2011, pp. 670–675, <https://doi.org/10.1145/2024724.2024876>.
- [21] T. Zhang, Y. Zhan, S.S. Sapatnekar, Temperature-aware routing in 3D ICs, in: *Proceeding of the IEEE ASPDAC*, 2006, pp. 309–314, <https://doi.org/10.1145/1118299.1118377>.
- [22] V.S. Rao, H.S. Wee, L.W.S. Vincent, L.H. Yu, L. Ebin, R. Nagarajan, C.T. Chong, X. Zhang, P. Damaraganath, TSV interposer fabrication for 3D IC packaging, in: *Proceeding of the 11th IEEE EPTC*, 2009, pp. 431–437, <https://doi.org/10.1109/EPTC.2009.5416509>.
- [23] S. Lee, S. Song, V. Au, K.P. Moran, Constriction/spreading resistance model for electronics packaging, in: *Proceeding of the 4thASME/JSME Thermal Engineering Joint Conference*, 1995, pp. 199–206, 1995.
- [24] M.M. Yovanovich, W.M. Rohsenow, J.P. Hartnett, Y.I. Cho, *Conduction and thermal contact resistances (conductances)*, *Handbook of Heat Transfer*, 3rd Ed., McGraw-Hill, 1998.
- [25] M.M. Yovanovich, Y.S. Muzychka, J.R. Culham, Spreading resistance of isoflux rectangles and strips on compound flux channels, *J. Thermophys. Heat Transf.* 13 (1999) 495–500, <https://doi.org/10.2514/2.6467>.
- [26] L. Choobineh, J. Jones, A. Jain, Experimental and numerical investigation of interdie thermal resistance in three-dimensional integrated circuits, *J. Electronic Packag.* 139 (2017), 020908, <https://doi.org/10.1115/1.4036404>.
- [27] G. Krishnan, A. Jain, Heat transfer in a multi-layered semiconductor device with spatially-varying thermal contact resistance between layers, *Int. Communic. Heat Mass Transf.* 140 (2023) 1–11, <https://doi.org/10.1016/j.icheatmasstransfer.2022.106482>, 106482.
- [28] D.W. Hahn, M.N. Özisik, *Heat Conduction*, John Wiley & Sons, 2012.
- [29] M.G. Cooper, B.B. Mikic, M.M. Yovanovich, Thermal contact conductance, *Int. J. Heat Mass Transf.* 12 (1969) 279–300, [https://doi.org/10.1016/0017-9310\(69\)90011-8](https://doi.org/10.1016/0017-9310(69)90011-8).