The Effects of Material Properties on Heat Dissipation in High Power Electronics

T. J. Lu
Department of Engineering,
University of Cambridge,
Cambridge CB2 1PZ
England
TJL21@eng.cam.ac.uk

A. G. Evans

J. W. Hutchinson

Division of Engineering and
Applied Sciences,
Harvard University,
Cambridge, MA 02138

The role of the substrate in determining heat dissipation in high power electronics is calculated, subject to convective cooling in the small Biot number regime. Analytical models that exploit the large aspect ratio of the substrate to justify approximations are shown to predict the behavior with good accuracy over a wide range of configurations. The solutions distinguish heat spreading effects that enable high chip-level power densities from insulation effects that arise at large chip densities. In the former, the attributes of high thermal conductivity are apparent, especially when the substrate dimensions are optimized. Additional benefits that derive from a thin layer of a high thermal conductivity material (such as diamond) are demonstrated. In the insulating region, which arises at high overall power densities, the substrate thermal conductivity has essentially no effect on the heat dissipation. Similarly, for compact multichip module designs, with chips placed on both sides of the substrate, heat dissipation is insensitive to the choice of the substrate material, unless advanced cooling mechanisms are used to remove heat around the module perimeter.

1 Introduction

Heat dissipation in high power electronics provides major challenges for the integration of materials selection with thermal design, circuit design, and manufacturing technology (Bar-Cohen and Kraus, 1988; Blodgett and Barbor, 1982; Mahalingam, 1985; Nakayama, 1986). The relevant properties include the thermal conductivity, k, the thermal expansion coefficient, α, and the dielectric constant. The objective is to select materials that enable the Si chip to operate with high power density (up to 100 W per chip), while maintaining its temperature below that needed to ensure acceptable reliability (usually 90°C). Moreover, the thermally induced stresses in the Si must be small enough to avert fracture. This design challenge is exacerbated by constraints on cooling, dictated by the application; typically, forced air convection. The situation is addressed by applying materials selector principles.

A prototypical configuration of a doubly periodic multichip module, cooled from below by forced air convection (Fig. 1), motivates the present analysis. The top of the module is thermally insulated by a protective cover (not shown in Fig. 1). By symmetry and/or by design constraint, there is no heat flux across the dashed planes in the figure. For analytical purposes, a cylindrical unit cell is used to mimic the representative square cell of the module. The chip area fraction relative to the substrate (heat spreader) is chosen to equal that encompassed within the square (Hingorani et al., 1994). An allowable temperature difference, $\Delta T = T_c - T_o$, is imposed between the Si chip and the cooling medium (typically 60°C). A maximum stress is permitted in the Si (say, 100 MPa). It is required that the dielectric constant be less than 10, such that the capacitive loss of high-frequency signals be small. It is also required that the device be compact (small w and b, Fig. 1) both for reduced manufacturing costs and for shorter signal-transmission lengths among circuits.

Heat transfer into the cooling medium occurs through the lower surface and is controlled by the magnitude of the Biot number (Holman, 1976)

$$B_i = h_o a/k, \tag{1}$$

where $2a$ is the chip width, usually 1–10 cm. For forced air cooling, the heat transfer coefficient h_o is in the range $10 \rightarrow 50$ W/m^{2}K (Incropera, 1988; Holman, 1976). Thermal conductivities, k, range between 1 and 1000 W/mK (Ashby, 1992). For this combination of properties, B_i is always small relative to unity, such that heat spreading governs dissipation into the cooling medium. That is, heat must be conducted efficiently to the lower substrate surface such that the temperature difference between this surface and the coolant is maximized over the largest possible surface area. The objective is to choose the material that maximizes the power density per chip. With Q as the power generated by the chip, the scaling relationship for the power density at large chip spacing is shown below to have the form,

$$Q/\pi a^2 = \Delta T[bh_o k/a^2]^{1/2} F, \tag{2}$$

where F is a dimensionless function of B_i, b/a, and w/a; detailed calculations will be presented later in this article (see (31)). The stresses, σ, in the Si chip are as follows (Hutchinson and Suo, 1992):

$$\sigma = E_s \Delta\alpha \Delta T_s/(1 - \nu_s), \tag{3}$$

where $\Delta\alpha$ is the difference in the thermal expansion between Si and the substrate, E_s is the Young's modulus for Si, ν_s its Poisson ratio, and ΔT_s is the cooling from the die attachment temperature. Maximum power density at the lowest stress suggests a merit index, $Q/(|\sigma|\pi a^2)$. This index is obtained from (2) and (3) as

$$Q/(|\sigma|\pi a^2) = (1 - \nu_s)[bh_o/a^2]^{1/2}[k^{1/2}/|\Delta\alpha|E_s](\Delta T/\Delta T_s)F. \tag{4}$$

When $B_i \ll 1$, F is only weakly dependent on B_i (or k). Thus, for best performance satisfying thermomechanical integrity, regardless of size and cost, $k^{1/2}/|\Delta\alpha|$ dictates substrate material selection. For material properties evaluated at room temperature, this parameter is plotted on Fig. 2, as a function of the dielectric constant. AlN, SiC, diamond, and BeO exhibit the highest rankings. The choice between these relies on detailed

Contributed by the Electrical and Electronic Packaging Division for publication in the JOURNAL OF ELECTRONIC PACKAGING. Manuscript received by the EEPD January 1, 1997; revision received January 27, 1998. Associate Technical Editor: L. Goldmann.

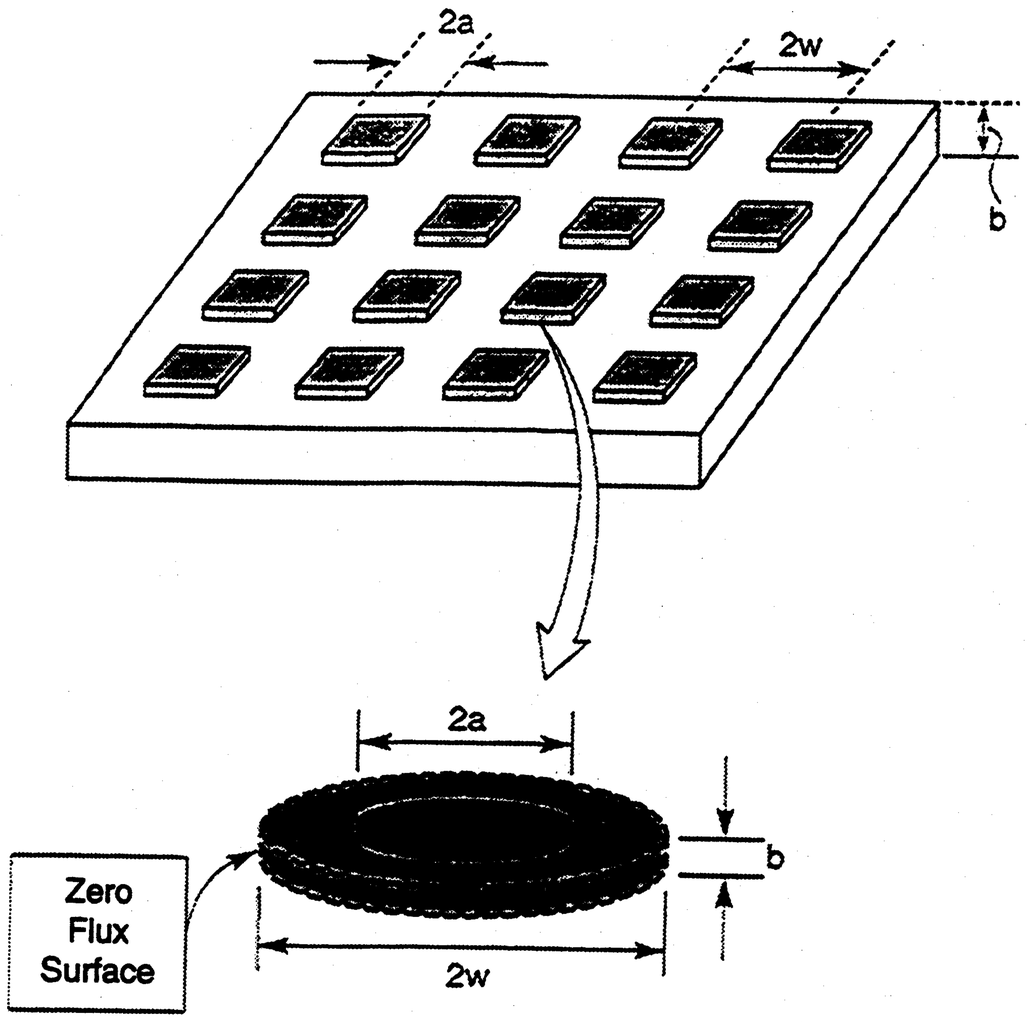

Fig. 1 Prototypical multichip module design cooled from below by forced convection. The representative cell is shown.

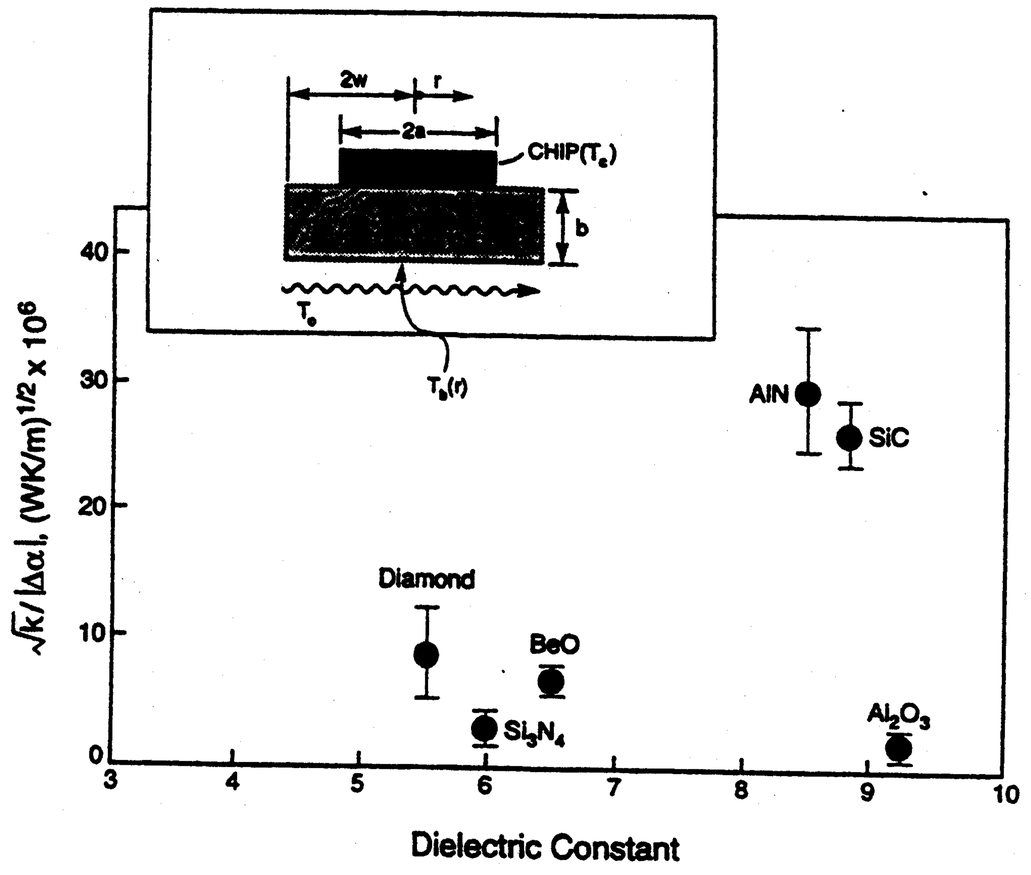

Fig. 2 A material selection diagram for high heat dissipation substrates. Large values of the ordinate and small value of the dielectric constant constitute the preferred materials. The insert shows the configuration that determines the material selection index.

thermal and electrical performance issues to be examined below, as well as manufacturing cost and size.

As the chip spacing decreases, the expression governing the power density becomes appreciably more complex than (2). Then, the thermal conductivity has a diminished role in the dissipation. One purpose of this article is to elaborate this role through detailed thermal analysis, and to identify changes in ranking from that given by Fig. 2. Monolithic substrates comprising one dielectric material, as well as Bimaterial substrates, are considered.

Manufacturing costs are assessed through their dependencies on the part volume, $V \approx 4bw^2$ (Evans et al., 1998). These dependencies are governed by both the amount of material needed in the device and the throughput for the most costly manufacturing step. Another objective of this study is to derive relationships between the performance and the part volume that can be used for cost modeling (Evans et al., 1998).

There have been many analytical and numerical calculations of related phenomena (Bar-Cohen and Kraus, 1988; Incropera, 1988; Nakayama, 1986; Peterson and Ortega, 1990). The previous study having the closest connection with the present objective is that conducted by Hingorani et al. (1994). A thorough review of the field has been made by these authors, and, hence, it will not be repeated here. In general, the problem of a heat source on a single-layer substrate with varying heat transfer boundary conditions has been solved by the finite element method, finite difference method, infinite series method, and the integral method. The classical infinite series method, coupled with separation of variable technique, was used by Hingorani et al. (1994). They demonstrated the applicability of an axisymmetric cell (Fig. 1), and established the essence of substrate heat spreading effects. That is, heat dissipation is facilitated by increasing the effective area over which heat is removed from the substrate. They also identified a critical substrate thickness that maximizes the heat dissipation. (This critical thickness is of order the chip size.) A much broader range of results is provided in this article by developing an analytical approach analogous to the shear lag concept used in stress analysis (Laws and Dvorak, 1988; Lu and Hutchinson, 1995a, b), with finite element validation. These results enable optima to be found in the overall substrate dimensions for a wide range of chip level power densities. They also enable comparisons between substrate materials that direct thermal dissipation approaches and materials concepts for high power electronics. Compared to the

Nomenclature

$a, \bar{a}$ = chip size, "super chip" size
A = temperature parameter (28)
a_i, A_i = temperature parameters (11a, 25)
b = substrate thickness
b_l = layer thickness
B_i = Biot number
$\bar{B}_i = B_i b/a$
c_1, c_2 = temperature parameters (17, 24)
C_l = temperature parameter (47b)
F = dimensionless function
h_i = chip/substrate interfacial heat transfer coefficient
h_o = heat transfer coefficient
I_0, I_1 = modified Bessel functions of first kind
k = substrate thermal conductivity
K_0, K_1 = modified Bessel functions of second kind
n = number of chips
M, N = temperature parameters (28)
N_l = temperature parameter (48)
q = heat flux
Q = heat dissipation per chip
$\bar{Q}$ = total amount of heat dissipation
$T(r, z)$ = substrate temperature
$\hat{T}(r)$ = average substrate temperature
$T_b(r)$ = temperature of substrate bottom surface
T_c = chip temperature
T_i = chip/substrate interface temperature
T_o = convective medium temperature
$w, \bar{w}$ = substrate width
α = thermal expansion coefficient
δ = die attach thickness
γ = material anisotropy factor
λ = scaling index
$\Re$ = diminution in heat dissipation due to die attach
Π_1 = power density index based on chip
Π_2 = power density index based on substrate
r, z = cylindrical coordinates

analytical and numerical methods used by other researchers, the shear-lag type method developed in this paper leads to much simpler solutions of the temperature field and power densities, with the evaluation of the Bessel function being the only numerical task. Conversely, the solutions obtained by Hingorani et al. and others involve the evaluation of infinite series whose eigenvalues must be determined numerically for each Biot number, which is not trivial. It is also noticed that the heat transfer boundary conditions considered in this study are different from those adopted by Hingorani et al. (1994). With reference to Fig. 1, the temperature of the chip is prescribed throughout the present study whereas the flux of heat to the chip is specified in Hingorani et al. (1994).

2 The Thermal Model

2.1 Concepts. Two basic designs are analyzed (Fig. 3). In the first, designated A, cooling is achieved by forced air convection passing over the lower surface, with fin designs that elevate heat transfer coefficients into the range 10–50 W/m^{2}K (Incropera, 1988; Nakayama, 1986). Type A designs based on both uni and bi-material substrates are considered (Figs. 3(a)–(b)). The chip surface and the upper surface of the package are regarded as thermally insulating because of design constraints.[1] In the second design, designated B, chips are placed on both surfaces and cooling is achieved around the perimeter (Fig. 3(c)). For this design, advanced cooling mechanisms with relatively high heat transfer coefficients at the edge are addressed in order to achieve reasonable power densities.

A steady-state analysis of the cylindrical cell model is performed wherein the chip is prescribed to be at essentially uniform temperature, T_c, generating power Q that dissipates into the substrate. In design A, the fluid medium used for cooling the lower substrate surface is maintained at temperature T_o by convective flow. Heat transfer into this medium occurs subject to a coefficient, h_o, with zero heat flux at the outer cylindrical surface. Cylindrical polar coordinates (r, θ, z) are chosen such that the z-axis is normal to the chip surface, with $z = 0$ at the center of the chip/substrate interface. By symmetry, the temperature and gradient fields in the cell depend only on r and z. For substrates with transversely isotropic properties, Fourier's law dictates that (Holman, 1976)

$$\begin{Bmatrix} q_r \\ q_z \end{Bmatrix} = -\begin{bmatrix} \gamma k & 0 \\ 0 & k \end{bmatrix} \begin{Bmatrix} T_{,r} \\ T_{,z} \end{Bmatrix}, \tag{5}$$

where $q_r(r, z)$, $q_z(r, z)$ are flux components in the radial and axial directions; γk, k are transverse and through-thickness thermal conductivities of the substrate, respectively, and $T(r, z)$ is the temperature in the substrate, with $T_{,r} \equiv \partial T/\partial r$, etc. In the analysis, γk and k are taken to be temperature independent. In the absence of internal heat sources, conservation of heat in association with Fourier's law (5) leads to the following steady-state heat conduction in the substrate

$$\frac{1}{r}\frac{\partial}{\partial r}\left(\gamma r \frac{\partial T}{\partial r}\right) + \frac{\partial^2 T}{\partial z^2} = 0. \tag{6}$$

For isotropic substrates, (6) reduces to the more familiar Laplace equation.

The interface between the chip and the substrate is assumed to be perfectly bonded and exerts no thermal resistance, requiring[2]

$$T_i = T_c, \quad (\text{on } z = 0, r \le a). \tag{7}$$

[1] Improved thermal performance could be achieved were it possible to remove heat from these surfaces.

[2] The case $T_i < T_c$ attributed to the thermal resistance of the die attach will be discussed later.

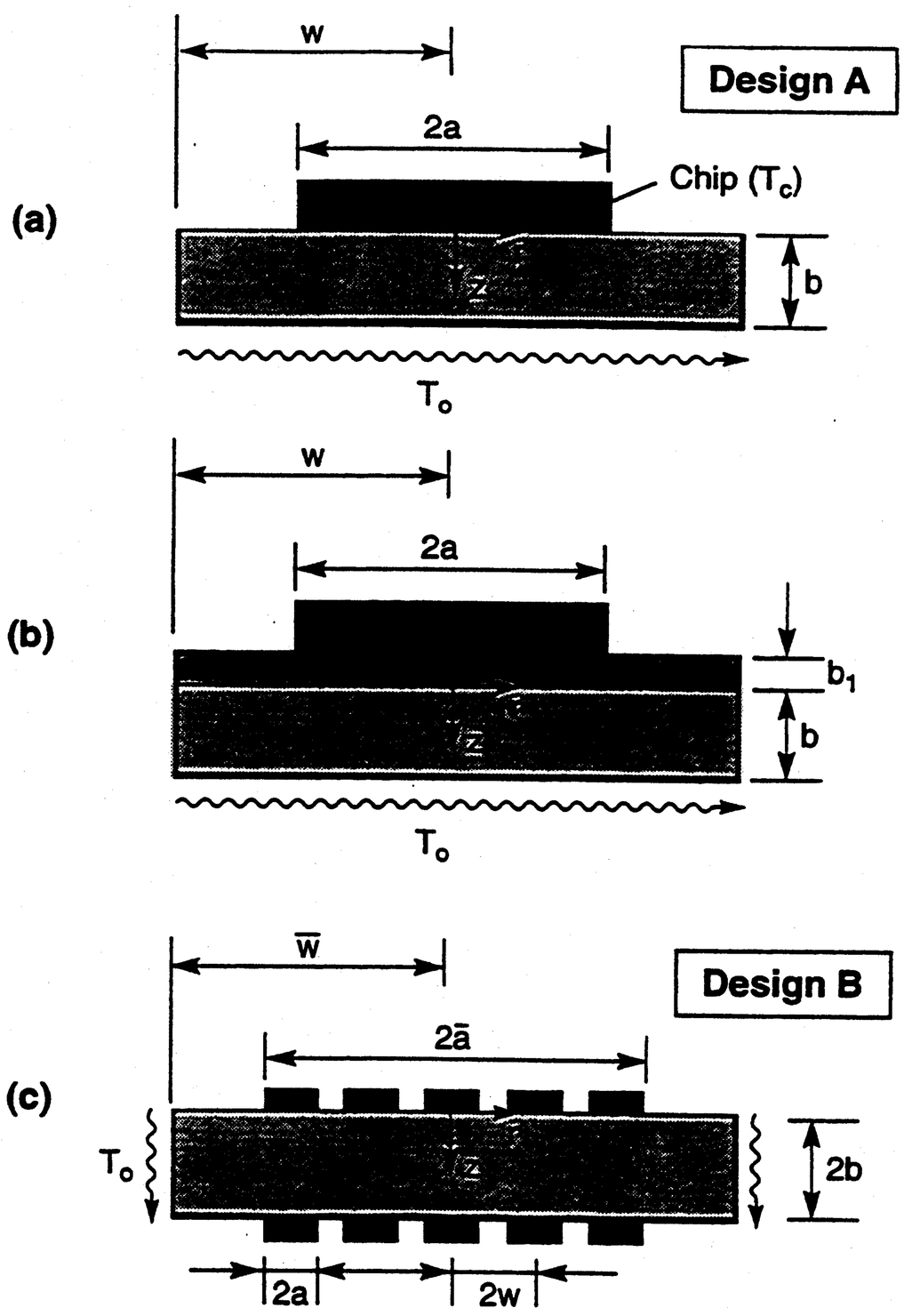

Fig. 3 Multichip module designs considered in the present analysis: (*a*) representative unit cell for a Monomaterial substrate with convective cooling from the lower surface, (*b*) bimaterial substrate with convection cooling, and (*c*) compact multichip module with perimeter cooling.

Symmetry with respect to the z-axis implies

$$T_{,r} = 0, \quad (\text{on } r = 0), \tag{8a}$$

while the periodic structure dictates that there be no heat conduction across the outer surface of the cell

$$T_{,r} = 0, \quad (\text{on } r = w). \tag{8b}$$

The upper surface of the substrate is assumed to be thermally insulated (due to protective cover), with

$$T_{,z} = 0, \quad (\text{on } z = 0, a \le r \le w). \tag{9}$$

At the lower substrate surface, heat is transferred to the environment by convection

$$kT_{,z}(r, b) = -h_o(T(r, b) - T_o), \tag{10}$$

where the heat transfer coefficient h_o is taken to be temperature independent.

2.2 Temperature Distributions. The set of governing equations of steady-state heat conduction in the cylindrical unit cell (5, 6) is formally analogous to that of the elasticity problem for the same cylindrical cell, if q_i, $T_{,i}$, k_{ij} are, replaced by ϵ_{ij}, σ_{ij}, S_{ijkl}, respectively. The so-called shear lag approximation (Laws and Dvorak, 1988; Lu and Hutchinson, 1995a, b), widely used to calculate shear stress distributions under analogous small thickness to width conditions, can be adapted to reduce the governing heat conduction equation to a one-dimensional set of ordinary differential equations. This approach is facilitated

further by distinguishing two zones in the substrate, separated by $r = a$ (Fig. 3(a)), following a related concept used by Hingorani et al. (1994). Zone I occurs directly beneath the chip, $r \leq a$. In this zone, the interface with the chip is at constant temperature, T_o. It is only required to determine the temperature of the lower surface, $T_b(r)$. Zone II occurs outside the chip ($r > a$), with the temperature now varying with position on both surfaces; $T_s(r)$ on the upper and $T_b(r)$ on the lower.

The shear lag method produces approximations for the variation of the temperature field in both zones, which are later validated through selected finite element results. In zone I, a quadratic approximation of the temperature distribution across the substrate is invoked such that

$$T(r, z) = a_1 + a_2 z + a_3 z^2. \tag{11a}$$

The coefficients a_i ($i = 1, 2, 3$) are determined by the conditions (7), (10) and the requirement that $T(r, b) = T_b(r)$, yielding

$$T(r, z) = T_c + (z/b)[(2 + \bar{B}_i)T_b(r) - 2T_c - \bar{B}_i T_o] + (z/b)^2[T_c + \bar{B}_i T_o - (1 + \bar{B}_i)T_b(r)], \tag{11b}$$

where $\bar{B}_i = h_o b/k = B_i b/a$. The temperature averaged over the substrate thickness is

$$\hat{T}(r) = \frac{1}{b}\int_0^b T(r, z)dz \tag{12}$$

such that

$$\hat{T}_{,r}(r) = \frac{4 + \bar{B}_i}{6} T_{b,r}(r). \tag{13}$$

Heat balance in the substrate element bounded by r and $r + \Delta r$ (Fig. 3(a)) requires that

$$-(2\pi br)\gamma k\hat{T}_{,r}(r) - (2\pi r\Delta r)kT_{,z}(r, 0) = -(2\pi b(r + \Delta r))\gamma k\hat{T}_{,r}(r + \Delta r) - (2\pi r\Delta r)kT_{,z}(r, b). \tag{14}$$

In the limit $\Delta r \to 0$, (14) reduces to

$$\frac{d}{dr}(\gamma r\hat{T}_{,r}) = -\frac{2r}{b^2}[T_c + \bar{B}_i T_o - (1 + \bar{B}_i)T_b], \tag{15}$$

where (11b) has been used. (This result can also be obtained by directly substituting (11b) and (12) into (6)). With the aid of (13), Eq. (15) simplifies to

$$r^2 T_{b,rr} + rT_{b,r} - (c_1 r/b)^2 T_b = -(c_1 r/b)^2 \frac{T_c + \bar{B}_i T_o}{1 + \bar{B}_i}, \tag{16}$$

where

$$c_1 = \left[\frac{12(1 + \bar{B}_i)}{\gamma(4 + \bar{B}_i)}\right]^{1/2}. \tag{17}$$

Subject to the constraints (7) and that the temperature be bounded at $r = 0$, (16) has a solution of the form

$$T_b(r) = A_1 I_0(c_1 r/b) + \frac{T_c + \bar{B}_i T_o}{1 + \bar{B}_i}, \quad (r \leq a) \tag{18}$$

where I_0 is the modified Bessel function of the first kind, and A_1 is a constant to be determined. Note that the second term on the right-hand side of (18) represents the solution in the limit $w = a$ where zone II vanishes.

In zone II, the proposed temperature distribution across the substrate is

$$T(r, z) = T_s(r) + (z/b)^2(T_b(r) - T_s(r)), \tag{19}$$

which satisfies the insulation condition (9) as well as the following surface conditions: $T(r, 0) = T_s$ and $T(r, b) = T_b$. The heat transfer boundary condition (10) dictates that the two substrate surface temperatures, T_s and T_b, be related by

$$T_s(r) = (\bar{B}_i/2 + 1)T_b(r) - \bar{B}_i T_o/2. \tag{20}$$

Since the heat flowing into the substrate element between r and $r + \Delta r$ must equal to that flowing out

$$-(2\pi br)\gamma k\hat{T}_{,r}(r) = -(2\pi b(r + \Delta r))\gamma k\hat{T}_{,r}(r + \Delta r) - (2\pi r\Delta r)kT_{,z}(r, b), \tag{21}$$

which, in the limit $\Delta r \to 0$, becomes

$$\frac{d}{dr}(\gamma r\hat{T}_{,r}) = -rT_{,z}(r, b). \tag{22}$$

With the assumed substrate temperature distribution (19), Eq. (22) may be re-expressed in the form

$$r^2 T_{b,rr} + rT_{b,r} - (c_2 r/b)^2 T_b = -(c_2 r/b)^2 T_o, \tag{23}$$

where

$$c_2 = \left[\frac{3\bar{B}_i}{\gamma(3 + \bar{B}_i)}\right]^{1/2}. \tag{24}$$

The solution to Eq. (23) is

$$T_b(r) = A_2 I_0(c_2 r/b) + A_3 K_0(c_2 r/b) + T_o, \quad (a \leq r \leq w), \tag{25}$$

where K_0 is the modified Bessel function of the second kind, and A_2 and A_3 are constants to be determined below.

The three unknown constants, A_i ($i = 1, 2, 3$) can be determined by the following three conditions:

(i) the lower surface temperature at $r = a$ is continuous:

$$T_b(a^-) = T_b(a^+); \tag{26a}$$

(ii) the total heat flowing from zone I to II must be conserved:

$$-(2\pi ab)\gamma k\hat{T}_{,r}(a^-) = -(2\pi ab)\gamma k\hat{T}_{,r}(a^+); \tag{26b}$$

(iii) there is no average heat flux across the substrate perimeter:

$$-(2\pi wb)\gamma k\hat{T}_{,r}(w) = 0. \tag{26c}$$

The resulting constants are as follows:

$$A_1 = -\frac{A\Delta T}{1 + \bar{B}_i}, \quad A_2 = \frac{\Delta T}{(1 + \bar{B}_i)N},$$

$$A_3 = \frac{I_1(c_2 w/b)}{K_1(c_2 w/b)} A_2, \tag{27}$$

where I_1, K_1 are the modified Bessel functions of the first and second kind, respectively, and

$$A = \frac{M}{N}\left[-I_1(c_2 a/b) + \frac{I_1(c_2 w/b)}{K_1(c_2 w/b)} K_1(c_2 a/b)\right]$$

$$M = \frac{2c_2}{c_1 I_1(c_1 a/b)} \frac{3 + \bar{B}_i}{4 + \bar{B}_i}$$

$$N = I_0(c_2 a/b) - MI_0(c_1 a/b)I_1(c_2 a/b) + \frac{I_1(c_2 w/b)}{K_1(c_2 w/b)} \times [K_0(c_2 a/b) + MI_0(c_1 a/b)K_1(c_2 a/b)]. \tag{28}$$

Equations (28), together with equations (11a), (18–20), and

(25), completely specify the temperature and temperature gradient fields everywhere in the substrate. It is emphasized that the model, although approximate, exactly conserves energy; that is, the heat entering the substrate is the same as that leaving its lower surface. However, there is a discontinuity of temperature gradient at the corner ($r = a$) on the upper surface. This discontinuity is also revealed by full numerical analysis using the finite element method. Such behavior has negligible effect on temperature distribution and power dissipation, even when $a/b \approx 1$.

2.3 The Power Density. The power per chip that dissipates into the convection medium is given by

$$Q = -2\pi k \int_0^a r T_z(r, 0) dr$$

$$= \left(\frac{\pi a^2 h_o \Delta T}{1 + \bar{B}_i}\right)\left[1 + \left(\frac{2 + \bar{B}_i}{\bar{B}_i}\right)\frac{2AI_1(c_1 a/b)}{c_1 a/b}\right]. \quad (29a)$$

In the limit $\bar{B}_i \to \infty$, the lower surface temperature becomes T_o in both zones, with Q reaching its maximum, Q_{max} as follows:

$$Q_{max} \approx \pi a^2 k \Delta T/b. \quad (29b)$$

In general, the following power density index may be defined[3]:

$$\Pi_1 \equiv \frac{Q}{\pi a^2 h_o \Delta T}, \quad (30a)$$

which, from (29a), is

$$\Pi_1 = \frac{1}{1 + \bar{B}_i}\left[1 + \left(\frac{2 + \bar{B}_i}{\bar{B}_i}\right)\frac{2AI_1(c_1 a/b)}{c_1 a/b}\right]. \quad (30b)$$

In the limit, when zone II disappears ($w = a$), Eq. (30b) reduces to

$$\Pi_1 = 1/(1 + \bar{B}_i), \quad (31a)$$

which is the exact limit corresponding to a constant gradient of temperature through the substrate. In the other limit, when zone II dominates ($w \gg a$),

$$\Pi_1 = \frac{1}{1 + \bar{B}_i}\left[1 + \left(\frac{b}{a}\right)\left(\frac{2 + \bar{B}_i}{1 + \bar{B}_i}\right)\sqrt{\frac{3 + \bar{B}_i}{3\bar{B}_i}}\frac{K_1(c_2 a/b)}{K_0(c_2 a/b)}\right]. \quad (31b)$$

If $B_i \ll 1$, (31b) can be further simplified to

$$\Pi_1 = 2\left(\frac{kb}{h_o a^2}\right)^{1/2}\frac{K_1(c_2 a/b)}{K_0(c_2 a/b)}, \quad (31c)$$

whereupon the function F in (2) becomes

$$F = 2K_1(c_2 a/b)/K_0(c_2 a/b),$$

which depends weakly on B_i.

2.4 Bimaterial Substrate. One concept for facilitating heat dissipation is to use a coating layer as heat spreader (Hussein et al., 1990; Beck et al., 1993; Hingorani et al., 1994), thickness b_1, having high thermal conductivity k_c (diamond, for example). If the layer is thin, its temperature is essentially uniform beneath the chip and equal to the chip temperature T_c. The temperature T_s in the layer beyond the chip is a function of r only. With $z = 0$ now chosen to coincide with the center of the interface between the two substrates (Fig. 3(b)), the solution procedures are nearly identical to those described for the monomaterial substrate, and, hence, will not be repeated. Here, the only difference is that, in zone II, heat transfer from the layer across the upper substrate surface must be included in energy balance considerations. In zone I, the lower surface temperature $T_b(r)$ is still found to have the form specified in (18), with c_1 defined in (17). In zone II, (25) still holds for $T_b(r)$ except that the coefficient c_2 is replaced by

$$c_2^* = \left[\frac{3}{\gamma}\frac{\bar{B}_i}{(3 + \bar{B}_i) + 3b_1 k_c (2 + \bar{B}_i)/(2bk)}\right]^{1/2}. \quad (32)$$

It follows that the coefficients A_i ($i = 1, 2, 3$) are the same as those listed in (28), subject to the new coefficient c_2^*. The power density index is thus,

$$\Pi_1 = \frac{1}{1 + \bar{B}_i}\left\{1 + \frac{2(2 + \bar{B}_i)}{(a/b)\bar{B}_i}\left[\frac{AI_1(c_1 a/b)}{c_1} + \frac{b_1 k_c c_2^*}{bkN}\right.\right.$$

$$\left.\left. \times \left(-I_1(c_2^* w/b) + I_1(c_2^* a/b)\frac{K_1(c_2^* a/b)}{K_1(c_2^* w/b)}\right)\right]\right\}, \quad (33)$$

which reduces to (30b) when $b_1 = 0$.

2.5 The Effect of Die Attachment. In the presence of die attachment (adhesive), heat transfer into the substrate occurs in accordance with a coefficient, h_i, that causes the substrate beneath the chip to attain a temperature, T_i, which is less than the chip temperature, T_c. When the thickness of the die attach system, δ, is small relative to the chip size, this heat transfer coefficient is

$$h_i = k_i/\delta, \quad (34a)$$

with k_i being the effective thermal conductivity of the die attach. In the analysis heretofore, perfect thermal contact ($h_i \to \infty$) has been assumed such that $T_i = T_c$. Under the assumption that the interfacial resistance, $1/h_i$ is not too large, T_i is essentially uniform with

$$T_i = T_c - Q/(\pi a^2 h_i). \quad (34b)$$

Let Q and Q_o denote separately the power dissipation per chip, with and without interfacial thermal resistance. The ratio $\Re = Q/Q_o$ is the diminution in heat dissipation caused by the die attach. It is straightforward to show that

$$\Re = 1/(1 + \Pi_1 h_o/h_i), \quad (35)$$

where Π_1 is either (30b) or (33). The ratio $\Re$ can be used as a "knock down" factor on all subsequent calculations that relate Q to the heat dissipation enabled by the substrate.

2.6 Perimeter Cooling. .For package design B (Fig. 3(c)), symmetry requires that only the upper half of the package needs be analyzed; therefore, the same cylindrical cell shown in Fig. 3(a) may be used with the understanding that the cell contains a single "super chip" which encompasses all the chips attached on one side of the package. The effective width of the "super chip" is denoted by $2\bar{a}$, while its temperature is taken to be uniform and equal to T_c. The substrate has width $2\bar{w}$ and thickness $2b$. The dimensions of the axisymmetrical super cell model are chosen such that $\pi\bar{a}^2$ equals the area of the square chip array and $\pi\bar{w}^2$ the area of the substrate. For simplicity, the substrate material is taken to be thermally isotropic ($\gamma = 1$). In accordance with the symmetry, the lower surface of the cell is thermally insulated,

$$T_z = 0, \quad (\text{on } z = b), \quad (36)$$

[3] The inverse, $1/\Pi_1$; is commonly known as the thermal resistance of the chip/substrate system (Bar-Cohen and Kraus, 1988).

while heat is transferred to the coolant by convection through the perimeter,

$$kT_{,r}(\bar{w}, z) = -h_o[T(\bar{w}, z) - T_o]. \quad (37)$$

The governing heat conduction equation and the other boundary conditions are the same as those used above, subject to the replacement of a by $\bar{a}$ and w by $\bar{w}$, respectively.

In zone I, the temperature distribution is taken as

$$T(r, z) = T_c + [T_b(r) - T_c][2z/b - (z/b)^2], \quad (38)$$

which automatically satisfies conditions (7), (36), and $T(r, b) = T_b(r)$ on the symmetry plane $z = b$. It follows from (38) that

$$\hat{T}_{,r}(r) = (2/3)T_{b,r}. \quad (39)$$

Substitution of (38–39) into Eq. (6) and rearranging, yields

$$r^2T_{b,rr} + rT_{b,r} - (\sqrt{3}r/b)^2T_b = -(\sqrt{3}r/b)^2T_c. \quad (40)$$

The solution for T_b is

$$T_b(r) = C_1I_0(\sqrt{3}r/b) + T_c, \quad (41)$$

where T_b is bounded at $r = 0$ and C_1 is an unknown constant.

In zone II, in order to satisfy the conditions that $T(r, 0) = T_s(r)$, $T(r, b) = T_b(r)$, $T_{,z}(r, 0) = 0$, and $T_{,z}(r, b) = 0$, the temperature distribution is taken as

$$T(r, z) = T_s + [T_b(r) - T_s][3(z/b)^2 - 2(z/b)^3]. \quad (42)$$

The resulting temperature averaged over the substrate thickness is

$$\hat{T}(r) = [T_b(r) + T_s(r)]/2. \quad (43)$$

Energy conservation dictates that, on every cylindrical surface, the following heat balance must hold:

$$-(2\pi rb)k\hat{T}_{,r} = (2\pi\bar{w}b)h_o[\hat{T}(\bar{w}) - T_o], \quad (44)$$

which has a solution of the form

$$\hat{T}(r) = \hat{T}(\bar{w}) + (h_o\bar{w}/k)[\hat{T}(\bar{w}) - T_o] \ln (\bar{w}/r). \quad (45)$$

The remaining two unknowns, $\hat{T}(\bar{w})$ and C_1, may be determined by satisfying the conditions that the temperature be continuous at $r = \bar{a}$ and that the heat leaving zone I is the same as that entering zone II:

$$\hat{T}(\bar{a}^-) = \hat{T}(\bar{a}^+) \quad (46a)$$

$$-(2\pi\bar{a}b)k\hat{T}_{,r}(\bar{a}^-) = -(2\pi\bar{a}b)k\hat{T}_{,r}(\bar{a}^+). \quad (46b)$$

The results are

$$\hat{T}(\bar{w}) = T_o + \Delta T/[1 + (\bar{w}/\bar{a})\sqrt{3}\bar{B}_iN_1/2I_1(\sqrt{3}\bar{a}/b)] \quad (47a)$$

$$C_1 = [\hat{T}(\bar{w}) - T_c]N_1^{-1}, \quad (47b)$$

where

$$N_1 = (2/3)I_0(\sqrt{3}\bar{a}/b) + (2/\sqrt{3})(\bar{a}/b)I_1(\sqrt{3}\bar{a}/b) \ln (\bar{w}/\bar{a}). \quad (48)$$

Note that, $\hat{T}(\bar{w}) \approx T_o$ if $B_i \gg 1$ while $\hat{T}(r) \approx T_c$ and $T_b(r) \approx T_c$ if $B_i \ll 1$. Note that $C_1 \leq 0$ since $\hat{T}(\bar{w})$ is always less or equal to T_c according to (47). It follows from (38) and (41) that the temperature everywhere in the substrate beneath the chip satisfies $T(r, z) \leq Tc$ (ensuring that heat never flows back to the chip from the substrate).

The power dissipated into the coolant is obtained with (47a) as

$$\bar{Q} = (2\pi\bar{w}b)h_o(\hat{T}(\bar{w}) - T_o) = (2\pi\bar{w}b)h_o\Delta T/[1 + (\bar{w}/\bar{a})(\sqrt{3}\bar{B}_iN_1)/2I_1(\sqrt{3}\bar{a}/b)]. \quad (49)$$

For cooling with $\bar{B}_i \ll 1$, this simplifies to

$$\bar{Q} \approx (2\pi\bar{w}b)h_o\Delta T. \quad (50)$$

If cooling were efficient ($\bar{B}_i \gg 1$), the maximum power the system could dissipate into the cooling environment becomes

$$\bar{Q} \equiv \bar{Q}_{max} \approx (4\pi k\bar{a}\Delta T/\sqrt{3})I_1(\sqrt{3}\bar{a}/b)N_1^{-1}. \quad (51)$$

For the practically relevant case, $\bar{w} = \bar{a} \gg b$, (49) reduces to

$$\bar{Q} \approx (2\pi\bar{a}b)h_o\Delta T/(1 + \bar{B}_i/\sqrt{3}) \quad (52)$$

such that the maximum possible heat dissipation (51) becomes

$$\bar{Q}_{max} = (2\sqrt{3})\pi k\bar{a}\Delta T. \quad (53)$$

Since the number of chips placed on one side of the substrate is approximately $n = (\bar{a}/w)^2$, the power generated per chip is $Q = \bar{Q}/n$. The power density index is therefore

$$\Pi_1 = 2(b/\bar{a})(w/a)^2/(1 + \bar{B}_i/\sqrt{3}). \quad (54)$$

For $B_i \ll 1$ and for practical ranges of w/a and $b/\bar{a}$, the power density, Π_1 is less than unity and well below that enabled by design A. However, if $B_i \gg 1$, the power density attains its maximum:

$$\Pi_{1,max} = (2/\sqrt{3})(kw^2/a^2h_o\bar{w}), \quad (55)$$

which is comparable to that for design A with typical ranges of properties.

2.7 Enhanced Perimeter Cooling. Cooling may be improved by allowing the upper substrate surface outside zone I ($r > \bar{a}$) to participate in convective heat transfer, in addition to perimeter cooling at $r = \bar{w}$. If cooling is adequate and $\bar{w}/\bar{a}$ not too large, it may be assumed that the temperature of the substrate beyond the "super chip" is uniform and equal to $\hat{T}(\bar{a})$, since the substrate conducts heat efficiently. The analysis of heat transfer then exactly follows that of the perimeter cooling case, subject to a change on the effective convective condition at $r = \bar{a}$ that accounts for the additional area available for heat transfer:

$$-(2\pi\bar{a}b)k\hat{T}_{,r}(\bar{a}) = [2\pi\bar{w}b + \pi(\bar{w}^2 - a^2)]h_o[\hat{T}(\bar{a}) - T_o]. \quad (56)$$

The average temperature beyond the chip, $\hat{T}(\bar{a})$, is then found to be

$$\hat{T}(\bar{a}) = T_c - \Delta T[1 + (\sqrt{3}/\bar{B}_i)(2\bar{a}b)/(2\bar{w}b + \bar{w}^2 - \bar{a}^2)]^{-1}, \quad (57)$$

which equals T_c if $B_i \ll 1$ and T_o if $B_i \gg 1$. Again, it can be checked that the substrate temperature $T(r, z)$ never exceeds T_c. The resulting total heat dissipation is

$$\bar{Q} = \pi h_o(2\bar{w}b + \bar{w}^2 - \bar{a}^2)/[1 + (\bar{B}_i/\sqrt{3})(2\bar{w}b + \bar{w}^2 - \bar{a}^2)/(2\bar{a}b)], \quad (58)$$

and the power density index follows as

$$\Pi_1 = (w/a)^2[2\bar{w}b/\bar{a}^2 + (\bar{w}/\bar{a})^2 - 1]/[1 + (\bar{B}_i/\sqrt{3})(2\bar{w}b + \bar{w}^2 - \bar{a}^2)/(2\bar{a}b)]. \quad (59)$$

It can be readily verified that the maximum heat dissipation $\bar{Q}_{max}$ and the maximum power density Q_{max} are the same as

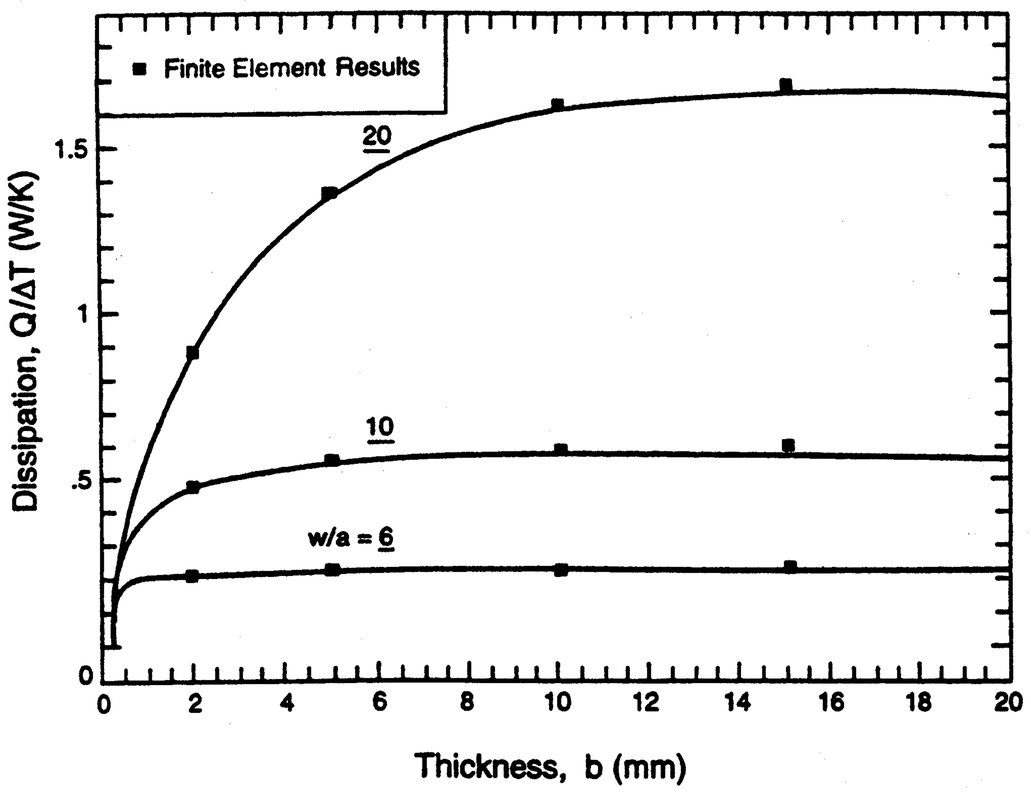

Fig. 4 Variations in power with substrate dimensions comparing magnitudes calculated using the shear lag model with those determined using finite elements. Similar levels of agreement are found over the full range of parameter space explored in this study.

those given by (53) and (55), respectively. Furthermore, when $\bar{w} = \bar{a}$, (58) reduces to (52) whereas (59) becomes (54).

3 The Results

Extensive finite element calculations using ABAQUS have been undertaken to check the accuracy of the analytical solutions. For the sake of brevity, details of these numerical calculations are not reported here. In general, the results establish that both the heat dissipation and the temperature distribution are accurately predicted by the analytical model. Comparison of selected finite element results with the analytical predictions, for a range of substrate thicknesses and chip spacings (Fig. 4), illustrates the level of accuracy.

3.1 Monomaterial Substrates. The temperatures T_b that develop along the lower surface of the substrate provide insight about the dominant heat conduction and heat transfer effects. The results are presented as temperature profiles ($0 \le r \le w$) for a range of relative thicknesses and chip spacings (Figs. 5 and 6). Beneath the chip in zone I ($r \le a$), the lower surface becomes cooler as the substrate thickness increases because the substrate behaves as an insulator. But, the effect is small and the thermal conductivity has a minimal effect on the temperatures. Beyond the chip ($r \ge a$), the opposite occurs. That is, the substrate surface becomes hotter as b increases. This trend reflects the heat spreading effect, which enables more heat to be conducted along the substrate as the path becomes wider. Also, the greater the chip spacing, the lower the temperature in the intervening substrate. The temperatures reached in this intervening region are strongly influenced by the substrate thermal conductivity. As k increases, B_i decreases causing T_b to increase. In the limit, T_b approaches the temperature T_i beneath the chip. The specific comparison of Figs. 5(a) and 6(a) with 5(b) and

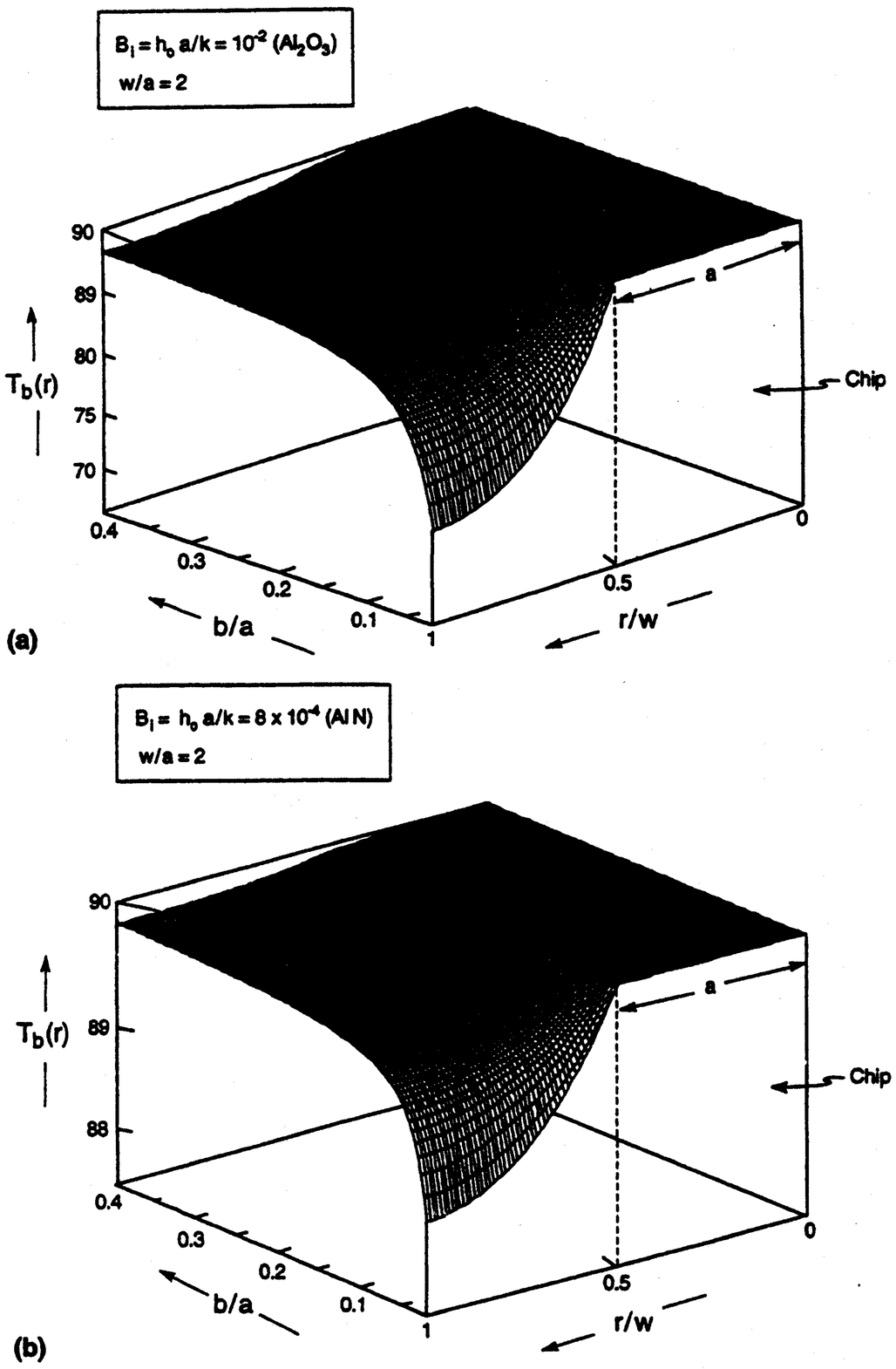

Fig. 5 Temperature distributions on the lower substrate surface indicating the role of substrate thickness, b/a for a chip density $a/w = 0.5$. Results are presented for two Biot numbers: (a) $B_i = 10^{-2}$ representative of Al_2O_3 and (b) $B_i = 8 \times 10^{-4}$ representing AlN.

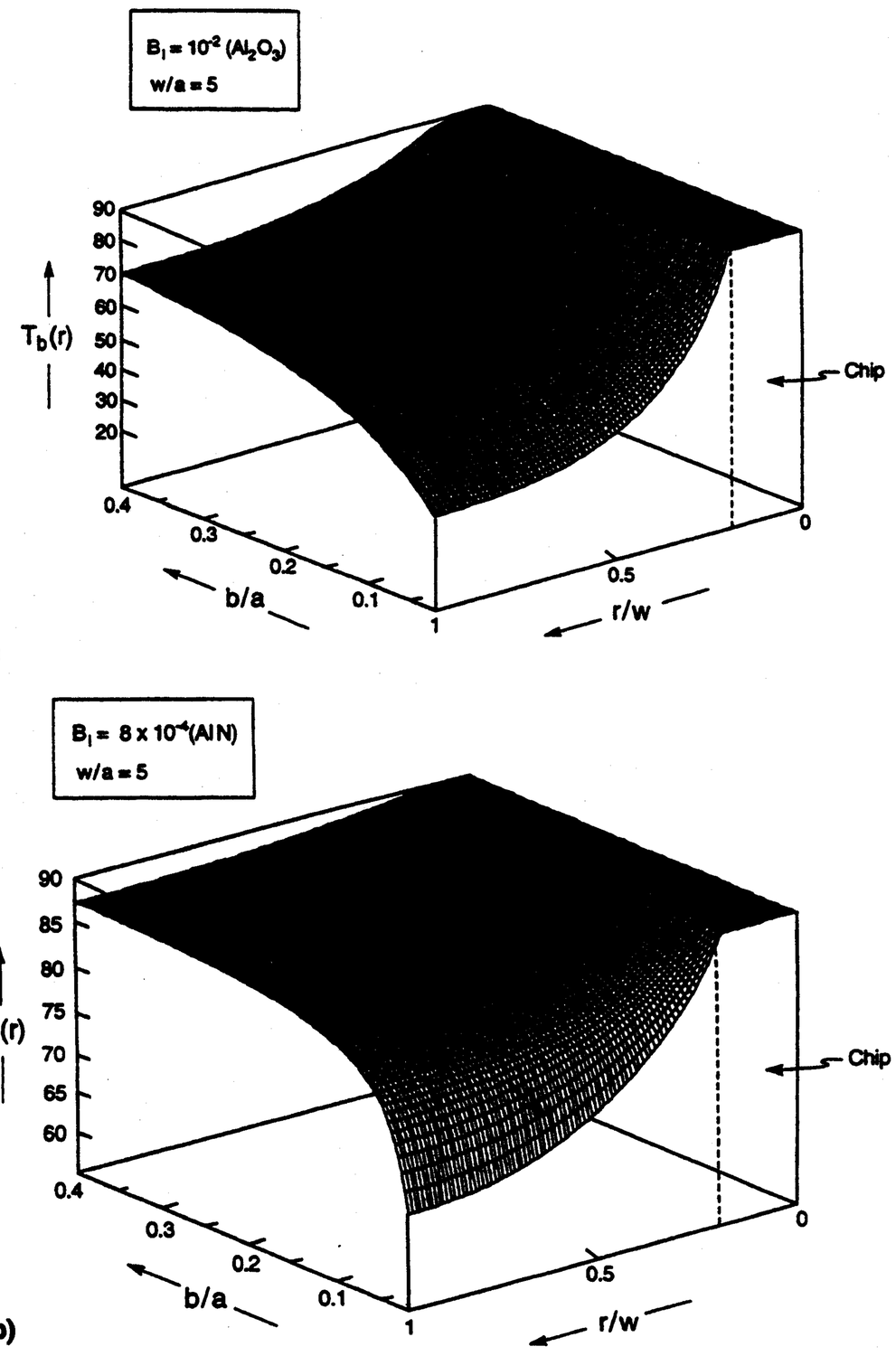

Fig. 6 Temperature distributions for a chip density, $a/w = 0.2$: (a) $B_i = 10^{-2}$ (Al_2O_3) and (b) $B_i = 8 \times 10^{-4}$ (AlN).

6(b) illustrates the advantages of AlN over Al_2O_3 for heat transfer conditions wherein heat spreading is critical to the dissipation.

Relationships between the power density and the substrate dimensions are derived using choices of the Biot number (1) that encompass the materials and heat transfer conditions of interest. The coordinates are the chip density, a/w, the relative substrate thickness, b/a, and the power density indices based on either the chip (29),

$$\Pi_1 = (Q/\pi a^2)(h_o \Delta T)^{-1}, \qquad (60a)$$

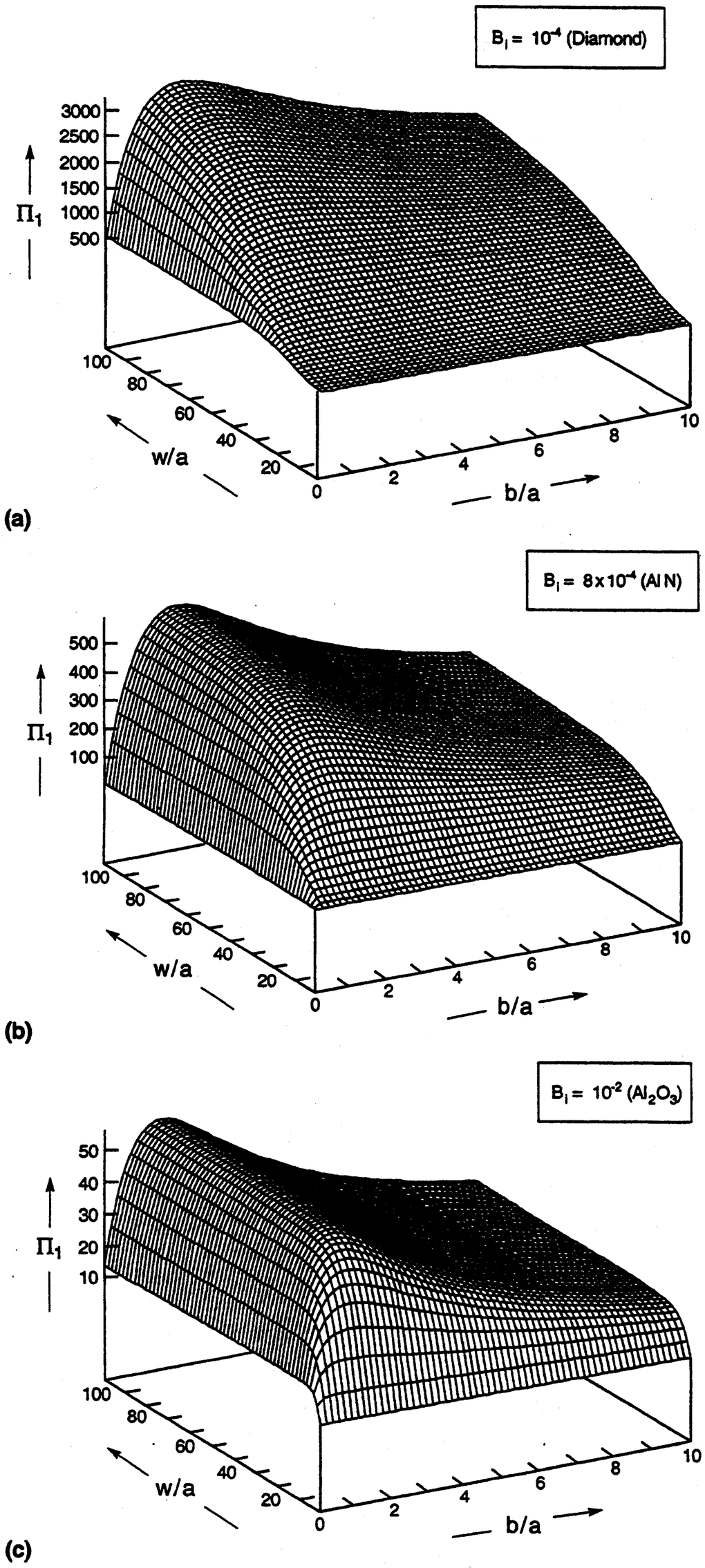

Fig. 7 Power density maps illustrating effects of chip density and substrate thickness at three levels of Biot number: (*a*) $B_i = 10^{-4}$ (diamond), (*b*) $B_i = 10^{-2}$ (Al_2O_3), and (*c*) $B_i = 8 \times 10^{-4}$ (AlN).

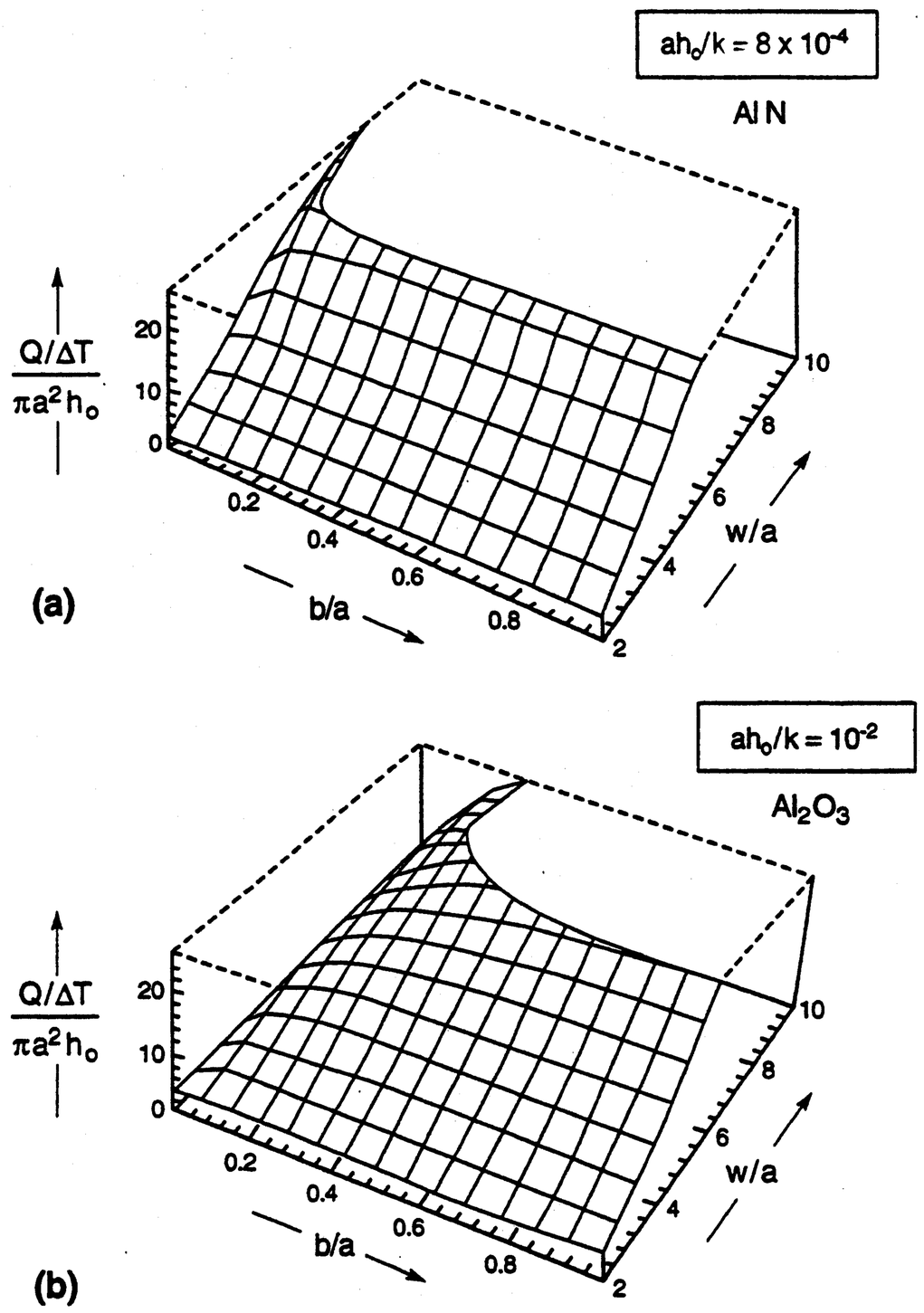

Fig. 8 Power density maps at relatively high chip densities indicating the range of substrate dimensions needed to achieve a power level, $\Pi_1 = 25$: (*a*) $B_i = 8 \times 10^{-4}$ (AlN) and (*b*) $B_i = 10^{-2}$ (Al_2O_3).

or on the substrate,

$$\Pi_2 = (Q/\pi w^2)(h_o \Delta T)^{-1}. \qquad (60b)$$

A few results are given over the full range of possible substrate dimensions in order to elaborate the overall trends and the underlying principles (Fig. 7). A ridge of maximum heat dissipation, Π_1, is evident within a relatively narrow range of substrate thickness, around $b/a \approx 1$. This effect has been identified previously (Hingorani et al., 1994) and has been designated the critical substrate thickness. But, relatively large chip spacings ($w/a \rightarrow 40$) are needed to approach the peak capacity. This capacity is strongly affected by the thermal conductivity of the substrate, increasing dramatically as k increases. The results compare Al_2O_3 with AlN and diamond. Note that the peak is almost directly proportional to k. These trends arise because heat transfer into a fluid medium at small Biot number occurs in accordance with the scaling index

$$\lambda = w^2 \int_o^1 [T_b(\rho) - T_o] \rho d\rho, \qquad (61)$$

where $\rho = r/w$. That is, as λ increases, the effective heat flux into the cooling medium increases. This accounts for the initial rise in Π_1 as the substrate thickness increases (Fig. 7). As b/a becomes larger, heat spreading is more efficient allowing λ to increase. This is offset at larger b/a by the insulating effect of the substrate beneath the chip (at $r/a < 1$). Conversely, Π_1 increases monotonically with chip spacing because λ scales with w^2, up to an asymptote at large w/a.

Further analysis emphasizes smaller chip spacings. Power density contours (Fig. 8) demonstrate the increases in power

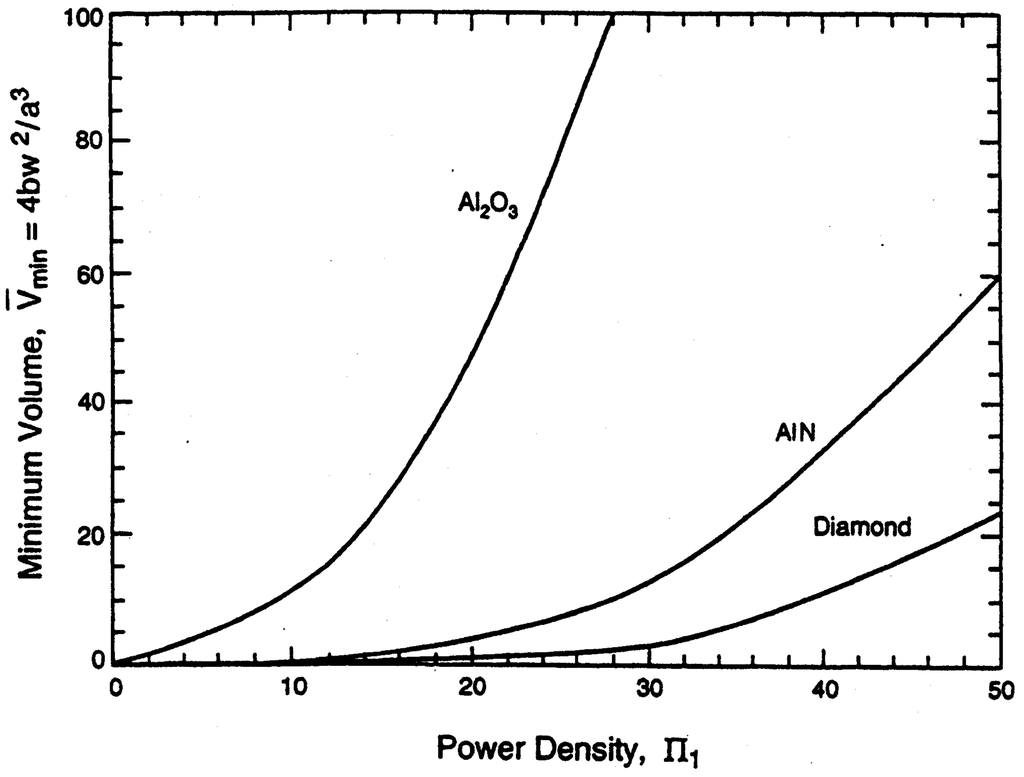

Fig. 9 Nondimensional minimum substrate volume V_{min}/a^3 as a function of power density, Π_1, for Al_2O_3 ($B_i = 10^{-2}$), AlN ($B_i = 8 \times 10^{-4}$), and diamond ($B_i = 10^{-4}$).

density, Π_1, that can be achieved by increasing the thermal conductivity of the substrate in certain ranges of substrate thickness and chip spacing. Truncating these plots at specific Π_1

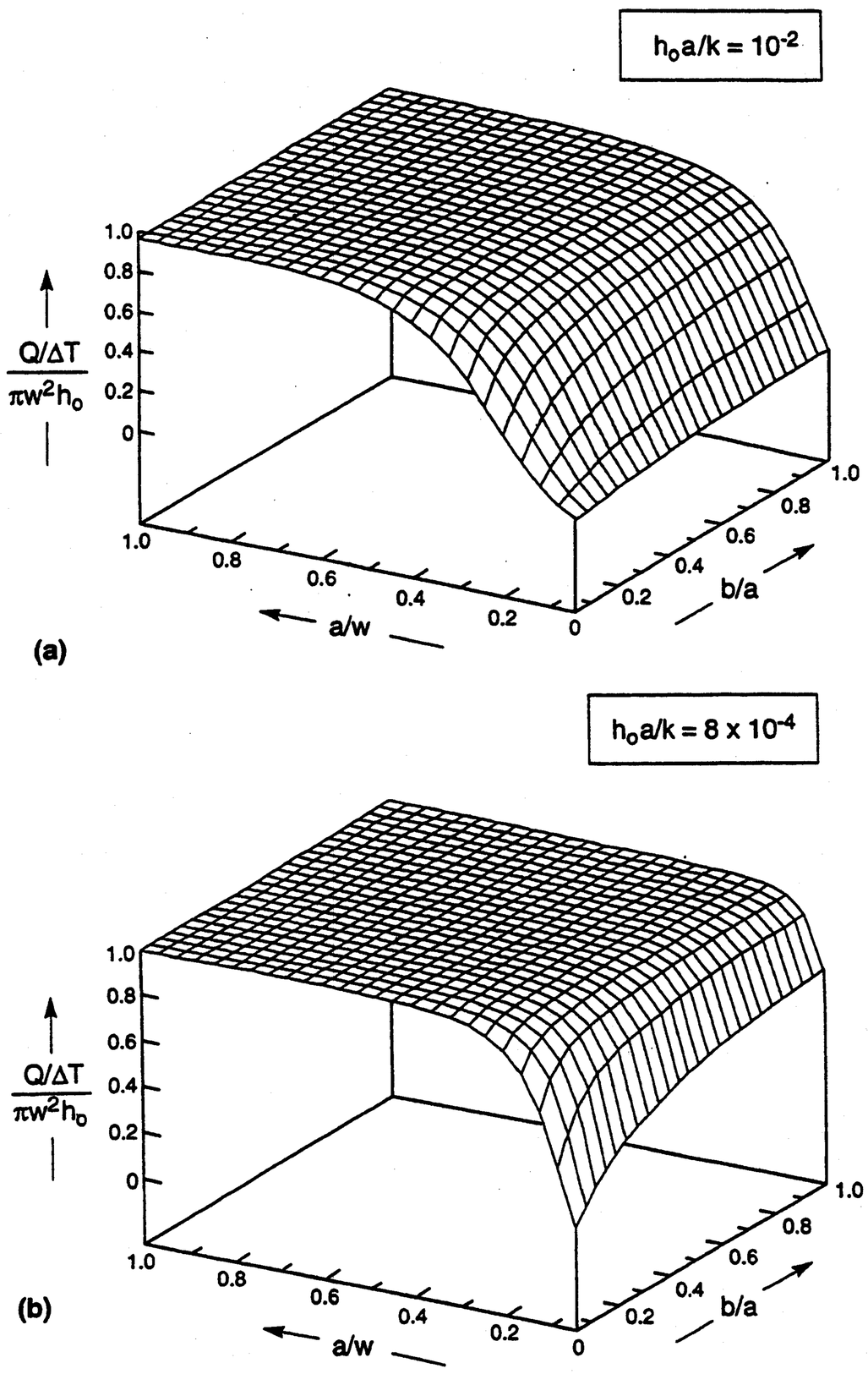

Fig. 10 The effect of substrate dimensions on the overall power density, Π_2: (*a*) $B_i = 10^{-2}$ and (*b*) $B_i = 8 \times 10^{-4}$.

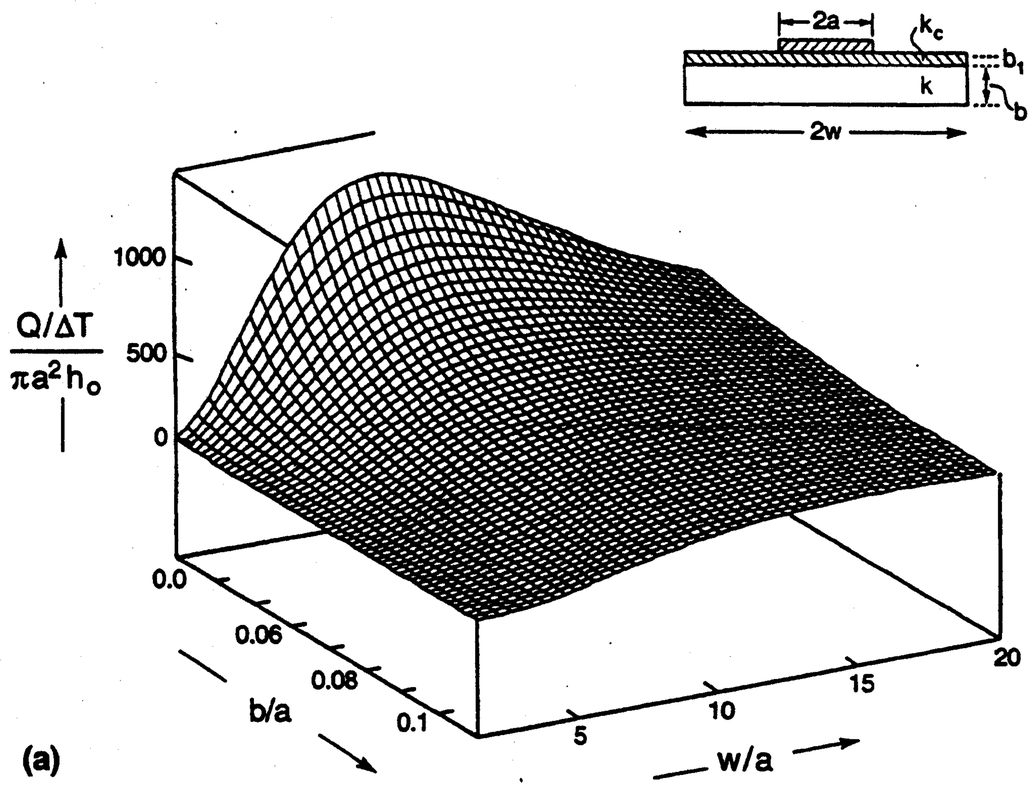

Fig. 11 The effect of a highly conducting layer on the power density, Π_1: (*a*) trend with substrate dimensions and (*b*) the influence of layer thickness, b_1. These results have been obtained by using the properties of AlN for the substrate and diamond for the layer.

(Fig. 8) vividly illustrates the effects of thermal conductivity on the substrate volume needed to achieve designated chip performance. Note that for AlN, the power requirements can be achieved with relatively narrow substrates, $w/a \approx 5$, for a wide range of thicknesses, b/a down to 0.2. However, for Al_2O_3, a large substrate is needed, $w/a \approx 7$, and, moreover, it needs to be relatively thick, $b/a \approx 0.7$. The minimum part volume V_{min} needed to achieve specified performance can be computed from these power density contours. This is plotted on Fig. 9 for several different materials in the normalized form, V_{min}/a^3. This volume is an input function to cost models (Evans et al., 1998), because of its effect on the manufacturing throughput. Note that relatively smaller substrate volumes are needed as the thermal conductivity increases (or, equivalently, as the Biot number B_i decreases), especially at high power density.

The overall power density, Π_2 (Fig. 10), attains its maximum value of unity at small w/a. This behavior is unaffected by the thermal conductivity and is insensitive to the thickness. In this regime, the power density can only be increased by enhancing the heat transfer coefficient, h_o, upon using more elaborate cooling schemes.

3.2 Bimaterial Substrates. The influences of thin layers of a high thermal conductivity material between the chip and the substrate are illustrated using material properties pertinent to AlN with a diamond layer (Figs. 11(*a*) and (*b*). The major findings are as follows. The power density index, Π_1, can be dramatically increased, but only in dimensional ranges where heat spreading dominates the dissipation. That is, when the chip density, a/w, is quite small. Moderately thick ($\sim$100 μm)

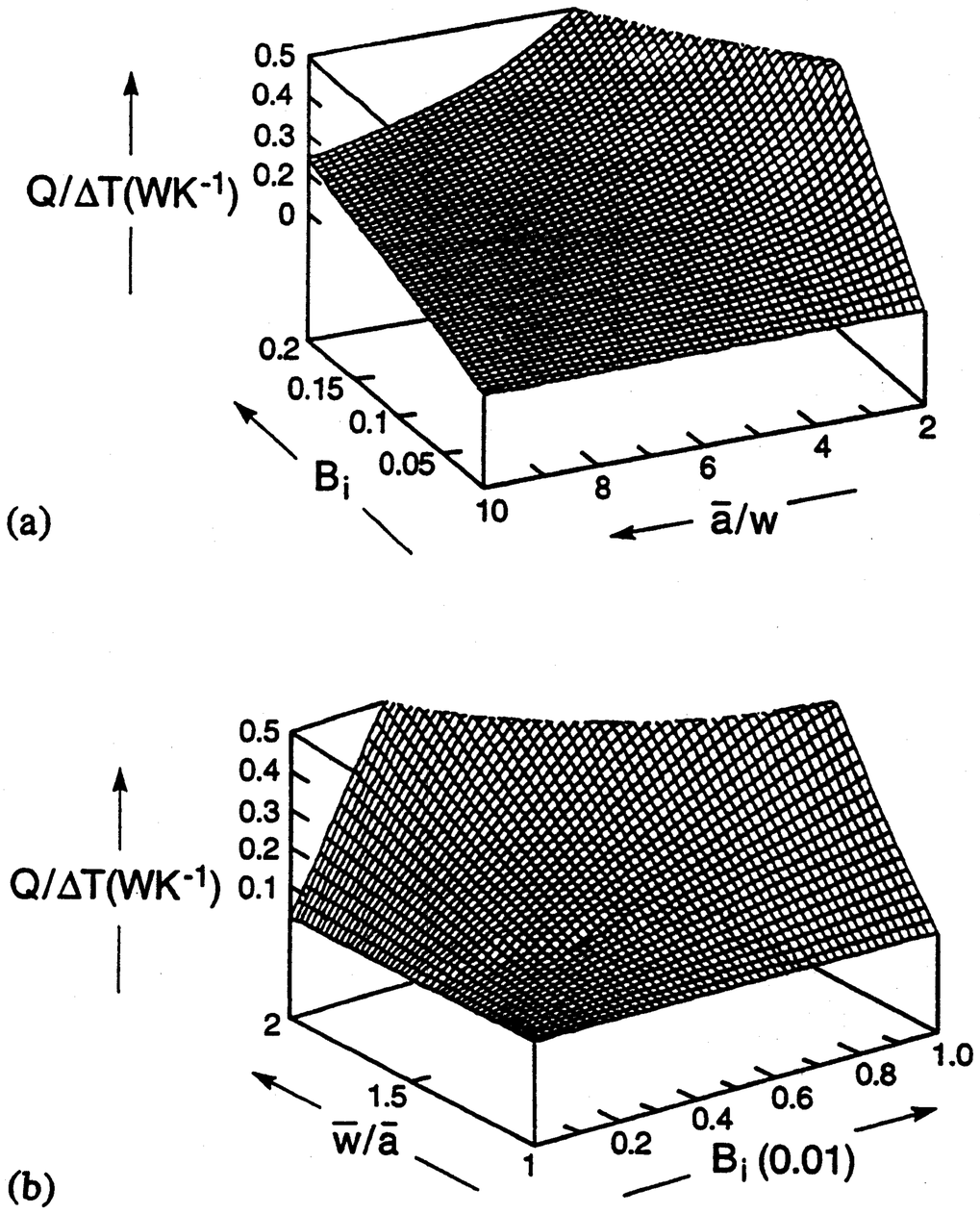

Fig. 12 (*a*) Power levels achieved with perimeter cooling, as functions of the Biot number and chip density ($a/b = 5$, $w/a = 2$, and $k = 500$ W/mK); (*b*) the effect of enhanced cooling on the power levels for perimeter cooling ($a/b = 5$, $w/a = 2$, $\bar{a}/w = 10$, and $k = 500$ W/mK).

diamond layers increase the power density index, Π_1. Moreover, the highest Π_1 arises when b/a is small (of order 0.3). This has a positive influence on the manufacturing cost via the substrate volume V_{min}. At either high or low chip densities, a/w, the diamond layer has a much diminished effect on the power density index. Then, the use of a diamond layer has little merit and, again, the cooling strategies that increase the heat transfer coefficient have the principal influence on heat dissipation.

3.3 Peripheral Cooling Design. In the peripherally cooled design, the chip level power densities are much smaller than in lower face cooling, even when high thermal conductivity substrates (AlN or diamond) are used (Fig. 12(*a*)). These results demonstrate that a successful symmetrical design requires cooling strategies. The role of improved cooling is demonstrated by Fig. 12(*b*). Here, heat could be dissipated from a 30W chip by using a heat transfer coefficient $h_o \approx 400$ W/m^{2}K.

4 Concluding Remarks

Calculations of heat dissipation in multichip modules have been used to highlight two basic regimes. One arises at low chip densities and demonstrates heat spreading effects. The other occurs at high chip densities and involves insulation. In the former, high thermal conductivity enhances the heat dissipation, and there are optima in both the thickness and the width of the substrate which maximize the chip level power density, Π_1. At the optima, Π_1 increases in proportion to k. There are also advantages to be gained by using high thermal conductivity layers, such as diamond, within specific dimensional ranges. That is, higher power density levels can be reached and thinner substrates may be used.

In the insulation range at high chip densities, the effects of the substrate are minimal. Improved heat dissipation is dictated entirely by the heat transfer coefficient through the cooling strategy. An important conclusion is that enhanced thermal conductivity of the substrate can only be used by designing to take advantage of heat spreading effects, manifest in Π_1 when w/a is large.

An assessment of compact multichip designs with peripheral cooling has indicated inferior heat dissipation, even when high thermal conductivity substrates with high heat transfer strategies (such as water cooling) are used.

The present work may be extended to cover the practically important case where the MCM has significant differences in power dissipation among the chips, such as a high powered chip surrounded by lower powered chips. In addition, the use of novel substrate materials such as open-celled metal foams with air or water cooling could be explored. A preliminary study (Lu et al., 1998) has shown that the compact multichip design of Fig. 3(*c*) may be rejuvenated if open-celled metal foams are attached to the substrate.

Acknowledgments

TJL would like to thank the Division of Engineering and Applied Sciences, Harvard University, IBM Corporation, and Carborundum Company for partially supporting this work. Support was also provided by NSF through the MRSEC at Harvard University.

References

Ashby, M. F., 1992, *Materials Selection in Mechanical Design,* Pergamon Press, Oxford.

Bar-Cohen, A., and Kraus, A. D., eds., 1988, *Advances in Thermal Modeling of Electronic Components and Systems,* Vol. 1, Hemisphere, NY.

Beck, J. V., Osman, A. M., and Lu, G., 1993, "Maximum Temperatures in Diamond Heat Spreaders Using the Surface Element Method," *ASME Journal of Heat Transfer,* Vol. 115, pp. 51–57.

Blodgett, A. J., and Barbour, D. R., 1982, "Thermal Conduction Module: A High-Performance Multilayer Ceramic Package," *IBM J. Res. Develop.,* Vol. 26, pp. 30–36.

Evans, A. G., Hutchinson, J. W., Hutchinson, R. G., Sugimura, Y., and Lu, T. J., 1998, "A Technical Cost Framework for High Temperature Manufacturing of Small Components and Devices," *J. Am. Ceram. Soc.,* in press.

Hingorani, S., Fahrner, C. J., Mackowski, D. W., Gooding, J. S., and Jaeger, R. C., 1994, "Optimal Sizing of Planar Thermal Spreaders," *ASME Journal of Heat Transfer,* Vol. 116, pp. 296–301.

Holman, J. P., 1976, *Heat Transfer,* McGraw-Hill Book Company, NY.

Hussein, M. M., Nelson, D. J., and Elshabini-Raid, A., 1990, "Thermal Management of Hybrids Circuits: Effect of Metallization Layer, Substrate Material and Thermal Environment," *Proc. International Society of Hybrid Microelectronics,* ISHM, Chicago, IL, pp. 389–394.

Hutchinson, J. W., and Suo, Z., 1992, "Mixed Mode Cracking in Layered Materials," *Advances in Applied Mechanics,* Vol. 29, pp. 63–191.

Incropera, F. P., 1988, "Convection Heat Transfer in Electronic Equipment Cooling," *ASME Journal of Heat Transfer,* Vol. 110, pp. 1097–1111.

Laws, N., and Dvorak, G. J., 1988, "Progressive Transverse Matrix Cracking in Composite Laminates," *J. Comp. Mater.,* Vol. 22, pp. 900–916.

Lu, T. J., Stone, H. A., and Ashby, M. F., 1998, "Heat Transfer in Metal Foams With Open Cells," *Acta Met. Mater.,* in press.

Lu, T. J., and Hutchinson, J. W., 1995a, "Thermal Conductivity and Expansion of Cross-Ply Composites with Matrix Cracks," *J. Mech. Phys. Solids,* Vol. 43, pp. 1175–1198.

Lu, T. J., and Hutchinson, J. W., 1995b, "Effect of Matrix Cracking on the Overall Thermal Conductivity of Fiber-Reinforced Composites," *Phil. Trans. R. Soc. Lond.,* Vol. A351, pp. 595–610.

Mahalingam, M., 1985, "Thermal Management in Semiconductor Device Packaging," *Proc. IEEE,* Vol. 73, pp. 1380–1387.

Nakayama, W., 1986, "Thermal Management of Electronic Equipment: A Review of Technology and Research Topics," *Appl. Mech. Review,* Vol. 39, pp. 1847–1868.

Peterson, G. P., and Ortega, A., 1990, "Thermal Control of Electronic Equipment and Devices," *Adv. Heat Transfer,* Vol. 20, pp. 181–314.