Layout to Circuit Extraction for Three-Dimensional Thermal-Electrical Circuit Simulation of Device Structures

B. H. Krabbenborg, A. Bosma, H. C. de Graaff, *Member, IEEE*, and A. J. Mouthaan, *Member, IEEE*

***Abstract*— In this paper, a method is proposed for extraction of coupled networks from layout information for simulation of electrothermal device behavior. The networks represent a three-dimensional (3-D) device structure with circuit elements. The electrical and thermal characteristics of this circuit representation are calculated with a circuit simulator. Spatial potential distributions, current flows, and temperature distributions in the device structure are calculated on the spatial coordinates. This simulation method can be placed between device simulation and (conventional) circuit simulation. It has been implemented in a circuit simulator and is demonstrated for simulation of self-heating in a bipolar low frequency power transistor. The main advantage of this simulation method is that not only the 3-D thermal behavior of the whole chip is simulated, but that this is also directly coupled to the electrical device behavior by means of the power dissipation and temperature distribution in the device. This offers the possibility for the circuit designer to simulate 3-D, coupled, thermal-electrical problems with a circuit simulator. As an example, the influence of the emitter contacting on the internal temperature and current distribution of a BJT is investigated.**

I. Introduction

NOT ONLY the increasing power densities in modern integrated circuits, but also the use of new structures (SOI) and materials (GaAs) cause self-heating to play an important role in device characteristics and reliability [1]–[3]. For analysis of self-heating effects, the heat flow equation has to be solved additionally to the conventional device equations [4]. Several two–dimensional (2-D) device simulators are available that offer this possibility [5], [6]. The heat flow equation should be solved in three dimensions (3-D) because of the 3-D nature of the temperature distributions. The simulation area in which this heat flow equation must be solved is usually much larger than the simulation area used for the other (electrical) equations (even in the 2-D case). When a device simulator is used, this results in a very high number of grid points and, therefore, long simulation times. This is especially the case when large structures, e.g., a power transistor with multiple emitter fingers, must be simulated [12]. Another well-known method for simulation of self-heating effects on a more macroscopic scale is proposed in [7]. This method uses an electrical analog for the thermal device behavior. This analog is a circuit in which currents represent heat flow and voltages represent temperatures. Power dissipation in the electrical circuit is modeled in the thermal circuit with a current source and the thermal resistance and capacitance are modeled with an electrical resistance and capacitance. This thermal circuit is coupled with the electrical circuit by means of the temperature and the power dissipation. A circuit simulator is used to solve these coupled circuits. The principle of this method is shown in Fig. 1. We used an extension of this method to make it suitable for simulation of 3-D transistor structures in which temperature distributions, current flows, and potential distributions must be solved. This is necessary if one is interested in the effects of temperature or potential gradients in the device on the overall device characteristics. Unlike the situation in Fig. 1 where only one device temperature was used (assuming one uniform temperature within the device), here the temperature distribution and the distributions of the currents in the device are calculated. Depending on the grid, the device (the transistor in this case) is split up into a number of sections. The inactive device areas are represented with resistance networks. The active transistor area is split up into a number of transistors, each with its own power dissipation, currents, temperature, etc. The thermal circuit is created the same way, but with its own grid that only coincides with the grid for the electrical circuit in the areas where heat is generated. This results in a 3-D grid representing the chip volume with capacitances from the nodes to ground, resistances from each node to its neighbors and current sources in the heat generating area. Fig. 2 shows the situation for a BJT after extension of the principle to 2-D.

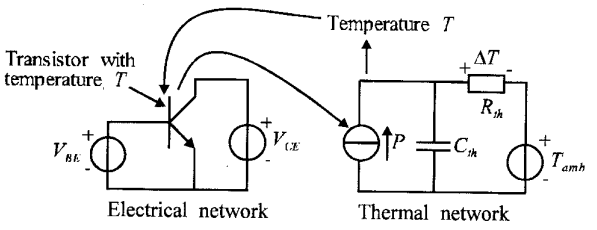


Fig. 1. Principles of the thermal-electrical circuit simulation method. In the thermal network, the symbols represent the thermal resistance $R_{\rm th}$ [K/W], the power dissipation source P [W], the thermal capacitance $C_{\rm th}$ [J/K], and the ambient temperature $T_{\rm amb}$ [K].

Manuscript received March 11, 1994; revised January 30, 1995 and September 12, 1995. This paper was recommended by Associate Editor J. White.

B. H. Krabbenborg was with the IC-Technology and Electronics Department, MESA Institute, University of Twente, Enschede, The Netherlands. He is now with Philips Semiconductors B.V., Quality and Reliability Consumer IC's, 6534 AE Nijmegen, The Netherlands.

A. Bosma was with the IC-Technology and Electronics Department, MESA Institute, University of Twente, Enschede, The Netherlands. He is now with the Dutch Institute for Animal Health Care, Deventer, The Netherlands.

H. C. de Graaff and A. J. Mouthaan are with the IC-Technology and Electronics Department, MESA Institute, University of Twente, Enschede, The Netherlands.

Publisher Item Identifier S 0278-0070(96)04966-4.

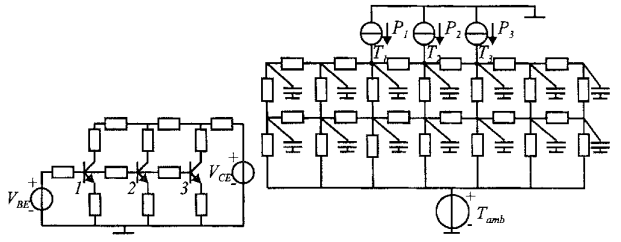


Fig. 2. Electrical and thermal network for a bipolar transistor in the 2-D case.

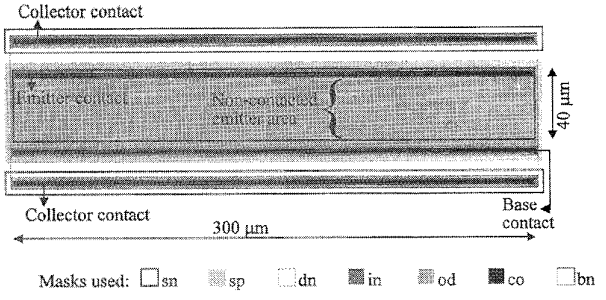


Fig. 3. Layout of the simulated bipolar NPN transistor with a partially contacted emitter area.

In the 3-D situation, the networks are usually very large. In this case, manual circuit generation is not very convenient.

This paper deals with a layout to circuit extractor developed to extract the thermal and the electrical network from the chip dimensions and the mask layout, respectively. The user determines which circuit elements are to be used for building the networks that represent the device. After simulation of these networks, a post processor couples the circuit simulation results to the original spatial coordinates so that spatial temperature and potential distributions can be examined. In Section II, the procedure for extraction of electrical and thermal circuits from the layout and the chip dimensions will be highlighted. In Section III, the use of different thermal boundary conditions is discussed, followed in Section IV by some comments on the use of the method. In Section V, the example is shown.

II. Layout to Circuit Extraction

The layout to circuit extraction process can be divided in the following parts.

- mask definition;
- grid definition;
- layer definition;
- generation of components in a layer;
- generation of components connecting layers.

All the steps will be briefly discussed in the next sections. As an example a bipolar transistor (also used for the simulation results in Section V) will be given.

A. Mask Definition

The lateral device structure is defined by the masks, and the vertical device structure in a certain point in the X-Y plane is defined by the combination of masks (and the related process steps) at that point. So the combination of masks in each point defines the device structure in that point, and what kind of

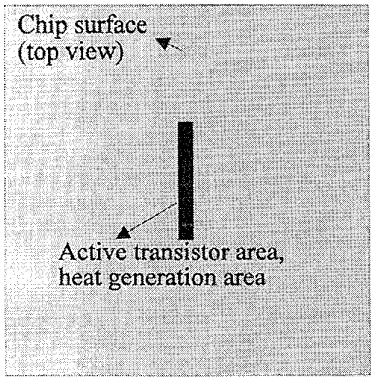


Fig. 4. "Thermal layout" of the simulated bipolar NPN transistor with a partially contacted emitter area. The "masks" in this case are the chip surface and the heat generating area (active transistor area).

circuit element must be generated for the network representing the device. In our program, each of the masks is defined with a number of rectangles. Not only the mask information used for the actual processing of the device is used, but also "masks" defining the heat generation area and the area of the chip surface have to be defined. This is necessary for the description of the thermal simulation region. Fig. 3 shows a layout of a bipolar power transistor with a partially contacted emitter. Fig. 4 shows the "masks" of the "thermal layout" of the same bipolar power transistor, the chip surface and the heat generating area.

B. Grid Definition

For extraction of electrothermal circuit models, two grids are used, one for the division of the thermal simulation region one for the division of the electrical simulation region. This shows that the regions covered by different grids may overlap each other. In the thermal-electrical case, the coordinates of the nodes of both grids in this overlap region must be equal if dissipating elements are located on the nodes (see Fig. 2). In the lateral direction, a triangular grid is used. The whole 3-D grid is built up using a number of identical stacked lateral triangular grids.

The procedure to generate the electrical grid, that determines how the layout is 'cut' into pieces, is as follows:

- First, all edges of the rectangles in the masks that have been defined must also appear as lines in the lateral electrical grid. So, when the masks have been defined, these directly give a first coarse version of the electrical grid.
- Second, the user refines the electrical grid to his demands by adding grid lines to the initial coarse electrical grid. In this way, the number of sections in which the transistor is divided, is specified and the "cut lines" are introduced.
- Third, the lateral thermal grid must be created. It must contain part of the electrical grid, and extend to the edges of the chip. The simple geometry of the chip surface, and the already available electrical grid allow an easy generation of the thermal grid (e.g., by using a quadtree method).

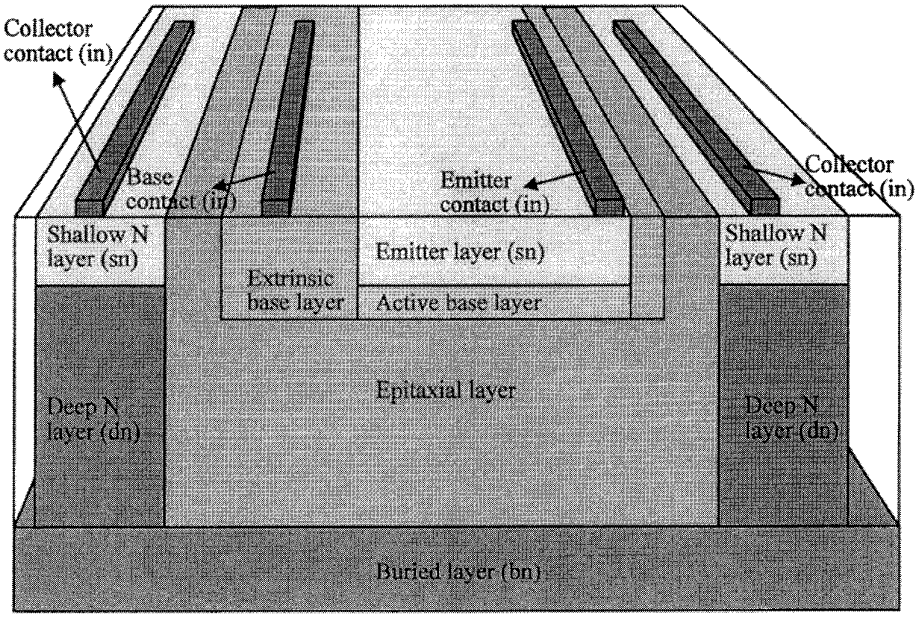


Fig. 5. The bipolar transistor built with layers. Each layer type has a different gray level.

In comparison with device simulation, the electrical grid that we use here in vertical direction (the layers), is usually very coarse. (For a bipolar transistor we usually need about four layers: one in the emitter, one in the base, one in the collector epiregion, and one in the buried layer.) This is because in device simulation, a relatively fine grid is necessary to get a good representation of the doping profile, the electric field distribution, avalanche generation, etc. In the circuit simulation method, much information concerning these quantities is already contained in the used compact device models because they have already incorporated the influence of the doping profile, avalanche generation, and electric field distribution on the current-voltage characteristics. Therefore, in device simulation, more grid points are necessary to obtain the same accuracy.

C. Layer Definition

The structure of a particular device depends on the masks used in the production process, and of course, the process itself. All specific structures in a device are represented by an abstract concept, the layer. Layers can be used to build complete devices. Fig. 5 shows how the bipolar transistor we use as an example is built up from a number of layers. A layer, or a combination of layers, represents a distinct part of the device such as the active base region of a bipolar transistor. A layer is built on a lateral grid with one or more circuit elements from each node to its neighbors. It can be connected to one or more other layers in vertical or lateral direction to create a 3-D structure. Each layer can have a number of properties such as a Z-coordinate, thickness, or sheet resistance. To start, these properties are assumed to be constant within the layer. The location and the shape of a layer in the X-Y plane is determined by the states (either on or off) of the masks used to produce a device. This process resembles the use of logical or Boolean functions. For example, the extrinsic base layer from Fig. 5 can be defined as a logical combination of the masks sp (shallow P) and sn (shallow N)

$$\text{ExtrinsicBaseLayer} = \{\text{sp } \mathbf{AND}\ (\mathbf{NOT}\ \text{sn})\}.$$

In Fig. 5, we can see that the actual transistor action is located in the active base layer. This layer is defined in the area where the base layer, emitter layer, epitaxial layer, and buried layer coincide. To determine the location of the active base layer, we can also perform logical operations on layers instead of masks

$$\text{ActiveArea} = \{\text{BaseLayer } \mathbf{AND}\ \text{EmitterLayer } \mathbf{AND}\ \text{EpitaxialLayer } \mathbf{AND}\ \text{BuriedLayer}\}.$$

This ActiveArea can be used to generate a transistor-layer that connects the base-emitter and collector layer. Layers connected in lateral direction must have the same Z-coordinate and layers connected in vertical direction must at least partly cover the same area in the X-Y plane. Layers can also be connected horizontally to model areas with different sheet resistance at the same Z coordinate. An example of this situation is the connection of the active base region and the extrinsic base region. To construct a circuit, each layer receives its own copy of the nodes defined by the global grid. All generated components in the layer will be connected to these nodes. Lateral components will be generated to connect pairs of nodes in the same layer, or horizontally connected layers. Vertical components will be generated to connect nodes from stacked layers. These components connect nodes with identical (X, Y) coordinates. Fig. 6 shows a 2-D cross section of the active part of the transistor built with layers. If a certain device area that must be represented with a layer is very thick, e.g., the thermal layer that represents the thermal resistance of the chip, it is represented with multiple layers on top of each other. This way the thickness of each of the layers defines a "grid"

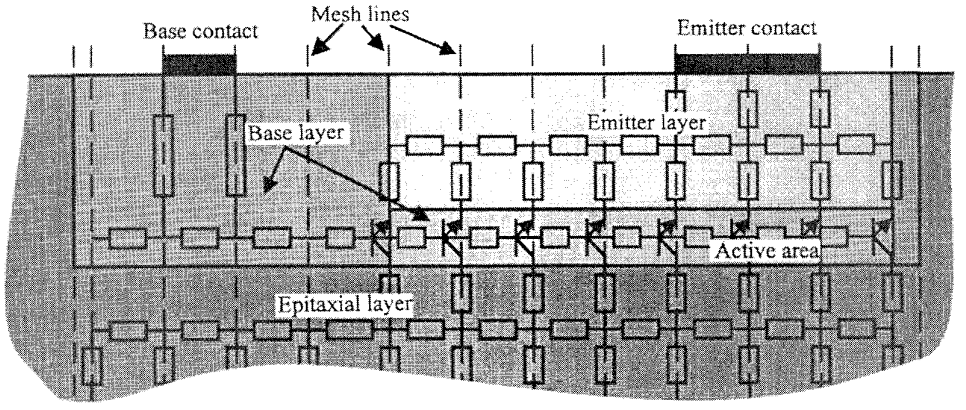


Fig. 6. Cross section of the active part of the transistor built with layers. The figure shows a part of the electrical circuit built with layers (emitter layer, active base layer, extrinsic base layer, and an epitaxial layer). In each of the layers, a number of components is drawn to indicate the layer type. The dashed lines represent the mesh lines. Note that vertical components are only located at connections between stacked layers.

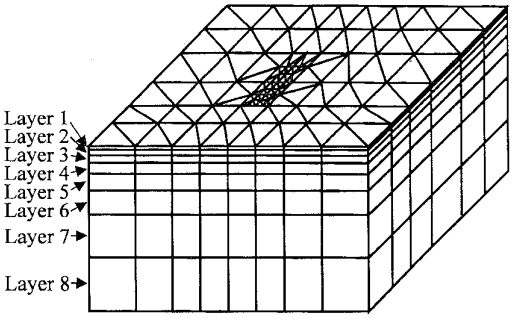


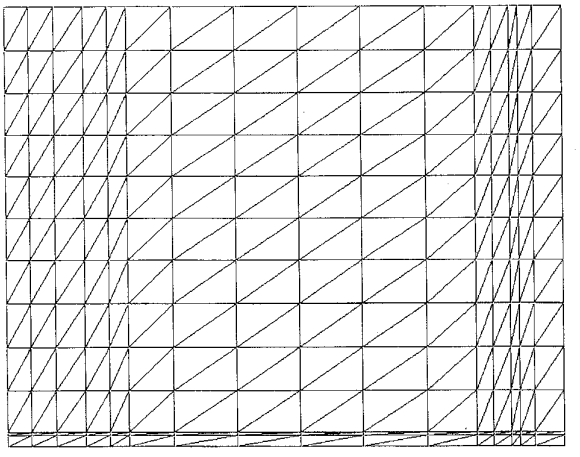
Fig. 7. Division of a layer into a number of sublayers. Because of the thickness of the chip, the thermal layer was divided into a number of sublayers all with the same material properties, but with different Z-coordinates.

in vertical direction (see Fig. 7). Fig. 8 shows the electrical and thermal grid that were used in the example in Section V.

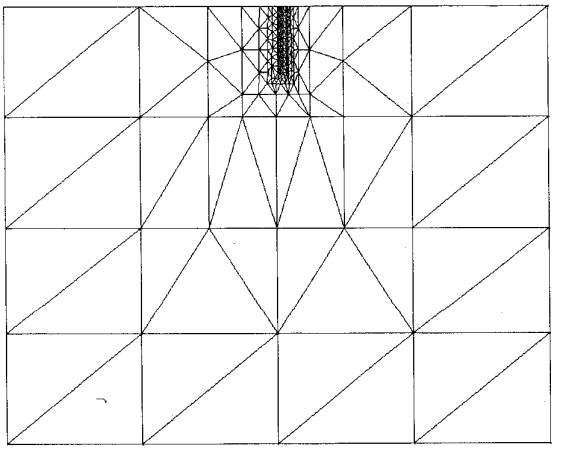
The thickness of a layer can also be variable, so that the grid dimensions vary during simulation. This is used in the thermal layer that represents the collector-base depletion region where most of the heat is generated in a bipolar transistor. The thickness of this depletion region depends on the collector-base voltage. When this voltage increases, the thickness of the thermal layer also increases. Because the volume of the whole simulation region must be constant, this increase in thickness is accompanied by an equal decrease in the thickness of the thermal layer just below. When this variable layer thickness is neglected, the calculated transistor temperatures are not correct.

D. Generation of Elements Within a Layer

The elements that are generated in a layer are a function of the grid, the thickness of the layer, and its material properties. As an example, the determination of the circuit elements in a thermal layer is given. These elements represent the heat conductance and capacitance in a silicon layer. (As in [9], but now a temperature-dependent thermal conductivity.) The heat flow equation on a triangular grid must be solved with a circuit simulator, so a representation of this equation in terms of



(a)



(b)

Fig. 8. Top views of the meshes that were used for the bipolar transistor. (a) The electrical mesh. (b) The thermal mesh.

electrical components must be found. The heat flow equation is

$$C_T \frac{\partial T}{\partial t} = H - \nabla \cdot \mathbf{S} \tag{1}$$

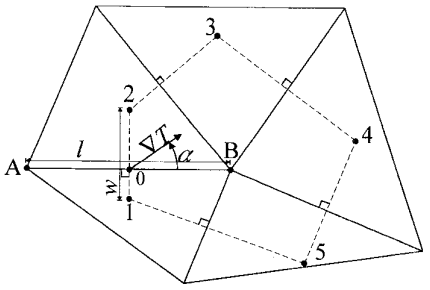


Fig. 9. Portion of a triangular network where the heat flow equation is solved.

with

$$\mathbf{S} = -\kappa \cdot \nabla T \tag{2}$$

where

C_T	thermal capacitance per unit of volume	[J/(Kcm^3)]
T	temperature	[K],
H	heat generation per unit of volume	[W/cm^3]
$\mathbf{S}$	heat flux vector	[W/cm^2]
κ	thermal conductance	[W/(Kcm)].

Now consider Fig. 9. The lines perpendicular to the edges of the triangles in Fig. 9, with an intersection in the middle of the edges, divide the region into polygons (polygon 1-2-3-4-5 in Fig. 9) surrounding each node. Now a network of thermal resistors can be constructed connecting the nodes. In silicon, κ is a function of the temperature [1]

$$\kappa = \kappa_0 T^{-4/3} \tag{3}$$

where $\kappa_0 = 3110$ [Wcm^{-1}K$^{+1/3}$]. The heat flow $\mathbf{S}$ can be rewritten to move all temperature dependent parts behind the gradient operator

$$S = \kappa_0 T^{-4/3} \nabla T = -3\kappa_0 \nabla (T^{-1/3}). \tag{4}$$

When $\nabla(T^{-1/3})$ is expanded in Taylor's series around the midpoint of the edge A-B (in point 0), and all terms are neglected except the first term, then $\nabla(T^{-1/3}) \approx \nabla(T^{-1/3})_0$. The heat flow from node A to node B is now

$$\begin{aligned} S_{AB} &= -\int_1^2 -3\kappa_0 \Delta(T^{-1/3}) \cdot \mathbf{n}\, dw \\ &= -3w\kappa_0 |\Delta(T^{-1/3})|_0 \cos(\alpha) \end{aligned} \tag{5}$$

where $\mathbf{n}$ is the unit vector normal to dw. The difference in $T^{-1/3}$ between A and B is

$$T_B^{-1/3} - T_A^{-1/3} = \int_A^B \nabla T^{-1/3} \cdot \mathrm{dl} = l\cos(\alpha)|\nabla T^{-1/3}|_0. \tag{6}$$

If the node voltages would represent $T^{-1/3}$, the value of the thermal resistance would be

$$R_{AB} = \frac{T_A^{-1/3} - T_B^{-1/3}}{S_{AB}} = \frac{-l}{3w\kappa_0} \tag{7}$$

but our node voltages represent temperatures, so we have to correct the resistance value. This results in a resistance that depends on the voltages of the nodes it is connected to (the resistance is now a voltage-controlled current source). After correction the value of the resistance is

$$R_{AB} = \frac{l(T_A - T_B)}{-3w\kappa_0 \left(T_A^{-1/3} - T_B^{-1/3}\right)} = \frac{l}{w\kappa_{\text{eff}}} \tag{8}$$

where κ_{eff} is the effective thermal conductivity on the path AB given by the integral from T_A to T_B over $\kappa(T)$ divided by the integration interval

$$\kappa_{\text{eff}} = \left(\frac{\int_{T_A}^{T_B} \kappa(T)\, dT}{T_B - T_A} \right). \tag{9}$$

A good approximation of κ_{eff} is obtained with the mean value κ_{mean}

$$\kappa_{\text{mean}} = \left(\kappa_0 \frac{T_A + T_B}{2} \right)^{-4/3}. \tag{10}$$

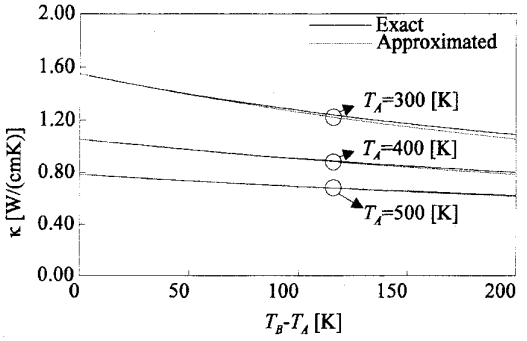


Fig. 10. Exact and approximated effective thermal conductivity as a function of the temperature difference across a thermal resistance in the case of a temperature dependent thermal conductivity.

In Fig. 10, κ_{eff} and κ_{mean} are shown as a function of the temperature difference across the thermal resistance $T_B - T_A$. It shows that the mean value can be used without introducing a significant error. In a 3-D thermal network with many thermal resistors, this approximation saves a lot of CPU time. Now the discretised version of the heat flow equation in node B is

$$C_B \frac{\partial T_B}{\partial t} = H_B + \sum_i^n S_i \tag{11}$$

with

$$S_i = \frac{T_i - T_B}{R_{i,B}} \tag{12}$$

where

C_B	heat capacitance of the area around node B	[J/K]
T_B, T_i	temperature in node B, node i	[K]
S_i	heat flow from node i to node B	[W]
$R_{i,B}$	thermal resistance from node i to node B	[W]
n	number of nodes connected with node B	[.]
H_B	total heat generation in the area around node B	[W].

The heat capacitance C_B is given by

$$C_B = CA_{\text{polygon}} \cdot d \tag{13}$$

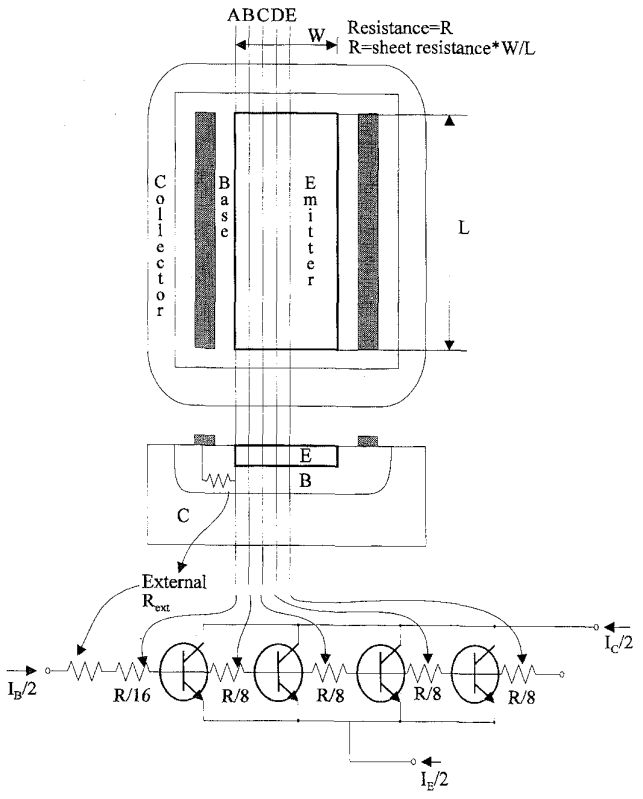


Fig. 11. Modeling the effect of base-spreading resistance [10].

where

C	is the specific heat capacity per volume	[J/(Kcm^3)]
A_{polygon}	is the volume of the area represented with node B	[cm^3]
d	is the thickness of the layer	[cm].

Obtuse interior angles triangles ($>90°$) should be avoided in triangulation. They may lead to negative values of l/w, depending on the angles of the adjacent triangle. This gives negative resistance values, and the currents in the network no longer represent physical currents in the material. For the node voltages (or temperatures in this case), this has no direct consequences.

E. Generation of Elements Connecting Layers

Vertical Connection: After the lateral elements within the layers have been generated, the vertical elements connecting the stacked layers can be generated. The following 2-D example in Fig. 11 demonstrates the partitioning of an IC-transistor into eight parts (four on each half). This example shows how the base current spreading in a bipolar transistor can be taken into account. It shows the principle of the used partitioning method. The transistors in this example give a vertical connection between three layers, the base, emitter and the collector layer. (In this example, the sheet resistance of emitter and collector layer is assumed zero.)

All necessary ingredients for a successful circuit extraction are presented in Fig. 11. A rudimentary layout is presented

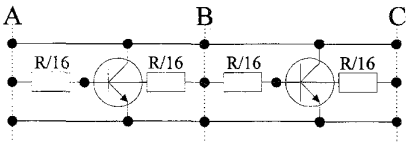


Fig. 12. A possibility to model the transistor sections from Fig. 11.

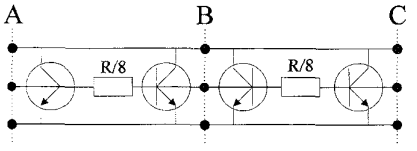


Fig. 13. A method to keep the electrical nodes on the mesh lines.

that gives information about the locations and sizes of the emitter and base. A grid is specified (the lines marked with A, B, C, D, and E), defining the sections in the IC transistor. Finally, when the value of pinched base sheet resistance and the parameters of the intrinsic transistor are given, the circuit can be generated.

The first method that could be used to construct the circuit in Fig. 11 is presented in Fig. 12. In this first method, a transistor is generated to represent the transistor action of *the area between two grid lines*. These modules (slices) can be joined to form a complete circuit. The physical properties of the transistor, like the saturation current, are proportionally scaled with the area between the grid lines. Here each submodel represents a slice of the device. A disadvantage of this method is the introduction of extra nodes in the middle of a $R/8$ resistor, necessary to introduce the transistors. No extra nodes are necessary in passive areas, however. This does not give problems when the circuits are generated manually, because this method is very intuitive. A disadvantage is that when many extra nodes have to be introduced, it is not very efficient when implemented in software.

The second method, which is actually used in our approach, is demonstrated in Fig. 13. It overcomes the disadvantages of the first method. Note that the area represented by each of the transistor symbols in this figure is only half of the area that the transistors in Fig. 12 represent. Also, note that there are two transistors connected to the same nodes, which seems very inefficient. However, multiple transistors connected to the same nodes (e.g., the two transistors connected to node B in Fig. 13) are joined to form one transistor with an area equal to the sum of the areas of the original transistors. So, in this second method, a transistor is generated to represent the transistor action in *the area around a grid line*. (In Fig. 9, this grid line is the line through node B perpendicular to the paper, and the area around this node is the polygon 1-2-3-4-5.) The resistor between nodes on grid line A and grid line B models the resistance of the area between those grid lines. With this method, electrical nodes appear only on grid lines. This is a great advantage over the previous method because a circuit can now be extracted sequentially, one type of component at a time. There is no need to divide resistances in two parts when a transistor must be connected, as in the first method.

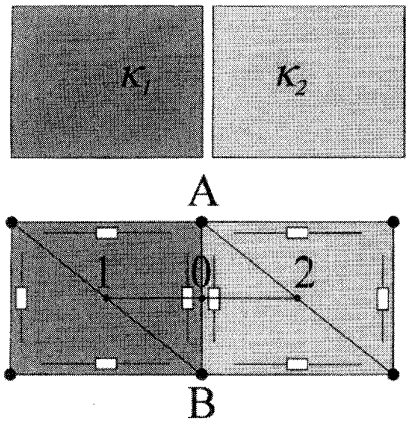


Fig. 14. Treatment of an interface between two layers with different thermal conductance.

Therefore, with this method, the order in which the structures are defined does not matter.

The method from Fig. 13 can now be used for each of the layers that has been defined. This can be done with only two commands. With one type of command line in the input file of the extractor, the generation of lateral circuit elements between nodes within a layer is done. With the other type of command line, generation of vertical circuit elements connecting a layer to one or more others is done.

When we use transistors in the circuit, a compact transistor model has to be given. In this model, the relations between the current densities, temperature, and node voltages must be defined. The dimensions of the transistor are input parameters of this model. Such a temperature dependent transistor model may contain, e.g., avalanche multiplication, high injection effects, and the Early effect. In the 2-D situation of Fig. 11, the width of each of the transistor sections is used as a parameter in the transistor model. In the extraction process, the spatial coordinates together with this model result in a transistor with parameters scaled to the dimensions of the actual active area it represents. The transistor model should not incorporate any lateral effects such as current crowding, or resistances like the base, emitter, or collector resistance, because these are modeled by in the circuit connecting all sections.

Lateral Connection: Layers with different properties can also be connected in lateral direction. Consider Fig. 14 where two materials with different thermal conductivity κ are connected. A change in κ is located on a grid line. The discontinuity of κ prevents the determination of the thermal conductivity in the midpoint between A and B, but because this discontinuity is located on a grid line, the resistor connecting A and B in Fig. 14 can be regarded as a parallel connection of two separate resistors. Therefore, the extractor also generates two parallel resistors. The the effective value R_{AB} of these two resistors between A and B is now

$$R_{AB} = \frac{l_{AB}}{\kappa_1 w_{01}} \Big\| \frac{l_{AB}}{\kappa_2 w_{02}} = \frac{l_{AB}}{\kappa_1 w_{01} + \kappa_2 w_{02}}. \qquad (14)$$

This method with separate parallel resistors is also used for the connection of layers with a different temperature dependence of the thermal conductivity. More generally speaking, it is used to handle the transitions between regions with different (temperature-dependent) thermal or electrical conductivities.

III. Thermal Boundary Conditions

An advantage of using of a circuit simulator is that various boundary conditions can be implemented relatively simple in the circuit representation of the device (e.g., heat radiation to ambient). With the layer structure, the temperature can be solved in any material (e.g., silicon, aluminum, oxide, or the passivation material). This makes it possible to incorporate the effect of contacts (Al or poly), the oxide, and the passivation layer on the thermal device behavior. Usually, the following thermal boundary conditions are used.

- A uniform temperature distribution at the bottom of the chip. Note that this does not necessarily imply a heat sink at the bottom of the chip, because the thermal resistance from chip carrier to ambient may also be significant and has to be added in that case.
- No heat flow across the sidewalls of the chip: $\mathbf{S} \cdot \mathbf{n} = 0$ where $\mathbf{S}$ is the heat flow vector and $\mathbf{n}$ is the unit vector normal to the surface of the sidewall.
- No heat flow across the top surface of the chip (unless heat radiation is taken into account as in [11]).

IV. User Interaction

The amount of user interaction necessary for using the layout extractor to create 3-D circuit models is variable. The user always has to create an input file that contains the following:

- the layout;
- the chip dimensions;
- the command lines for the circuit simulator;
- a reference to the grid-files that contain the thermal grid and the electrical grid. The grid files are created with the user directed gridding program. Because the lines in the electrical grid must cover all mask edges, an initial coarse version of the electrical grid is defined as soon as the layout has been given. By refining this initial coarse electrical grid, the user must rely on his engineering judgement to determine where the extra "cut-lines" that separate the transistor sections must be introduced. This determines the accuracy and complexity of the resulting network representation;
- a reference to the model-file. This model-file contains all definitions that are usually constant for a given process and device. Only if the user wants to simulate structures that are not present in the available model-files, he has to create his own model-file (e.g., by modifying an existing model-file). It will be clear that this enlarges the amount of necessary user interaction. The model-file contains the following:

  -The definition of the layers. The shape is defined with logical functions on the masks. The thermal and electrical components in the layer are defined in almost the same way as in any input file for the circuit simulator. The only difference is that the parameter values are

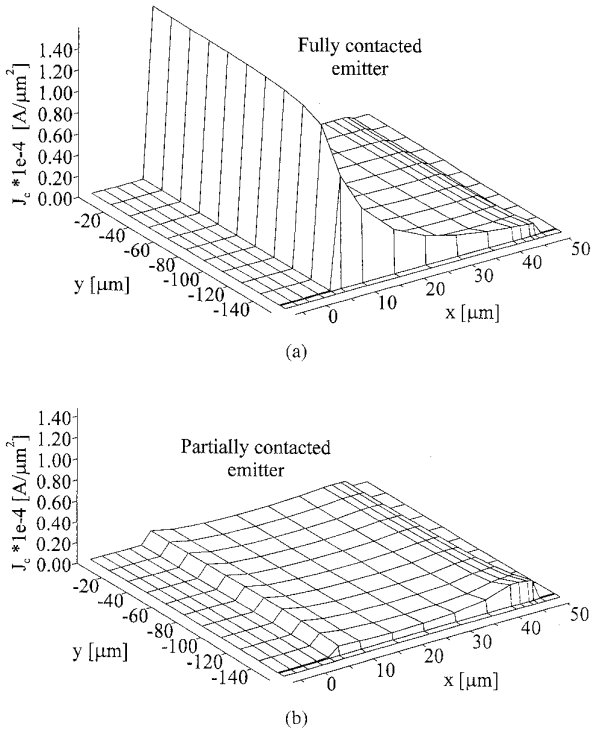


Fig. 15. Collector current distribution in transistor with fully contacted emitter area (a) and partially contacted emitter area (b) for $V_{BE} = 0.8$ [V] and $V_{CE} = 2.0$ [V]. Note that the collector current distribution is only plotted in the active emitter region.

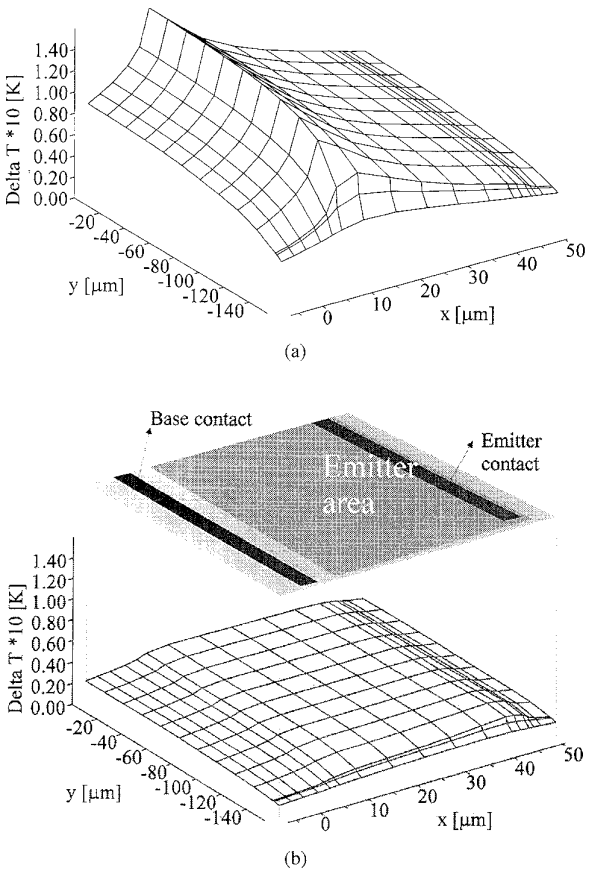


Fig. 16. Lateral temperature profile in transistor with fully contacted emitter area (a) and partially contacted emitter area (b) for $V_{BE} = 0.8$ [V] and $V_{CE} = 2.0$ [V].

still functions of spatial dimensions. These functions are evaluated during the extraction process;

-The definition of the connections between the layers;

-The models for all thermal and electrical circuit elements that are used in the layers and the connections between layers;

-The parameters for all these models (that do not depend on spatial dimensions).

The extraction program generates an output file that can be run with a circuit simulator. A post processor is used to visualise the results.

V. Application: The Effect of the Emitter Contacting on the Temperature Distribution in a Bipolar Transistor

The circuit extraction method described in the previous section has been used for a BJT to determine the influence of the base and emitter contacting on the thermal stability of the transistor. Fig. 3 shows the symmetric layout (the left half is the mirror image of the right half) of the bipolar transistor that was simulated. This figure shows that the emitter is only partly contacted. This is usually done to introduce an extra emitter series resistance for the transistor part close to the base contact. Because the ratio of emitter current and base current is approximately the same as the ratio of emitter region conductance and base region conductance, this series resistance introduces a lateral gradient in the emitter potential, approximately equal to the base potential gradient. In this way, the base-emitter junction voltage is more uniform over the whole emitter area and current crowding is reduced. It is expected that this also improves the thermal stability and the Safe Operating Area of the transistor because the temperature profile is also expected to be more uniform. Fig. 5 shows how the electrical network was built with a number of layers (each layer with a different gray scale). Both transistors (one with a fully contacted emitter and one with a partially contacted emitter) were simulated with a base emitter voltage V_{BE} of 0.8 [V] and a collector emitter voltage V_{CE} of 2 [V]. Because of symmetry only one half of the transistor was simulated ($-150\ \mu m < y < 0\ \mu m$). In the I_C-V_{CE} characteristics the influence of the emitter resistance is evident. In the transistor with the partially contacted emitter, the potential gradient in the pinched base is nearly completely compensated by the potential gradient in the emitter area, so there is less current crowding at the emitter edges than in the transistor with a fully contacted emitter. The collector current distributions for both transistors are shown in Fig. 15, and the corresponding temperature distributions are shown in Fig. 16.

In Figs. 15–18, the collector current distributions (for the collector current flowing in vertical direction) are only shown in the *active* transistor area. In the *nonactive* transistor area, there is no transistor action and, therefore, the collector current is zero there.

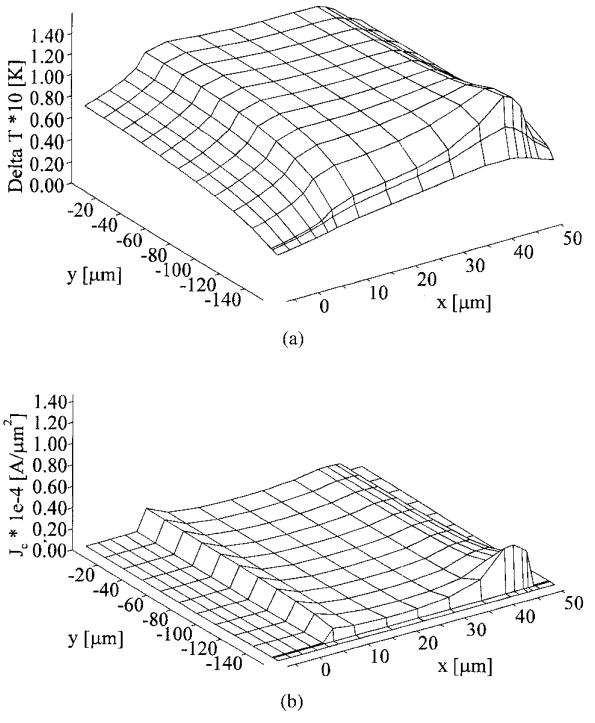


Fig. 17. Lateral temperature profile (a) and collector current distribution (b) in the transistor with the partially contacted emitter area for $V_{BE} = 0.83$ [V] and $V_{CE} = 5.0$ [V]. The power dissipation, in this case, is equal to the power dissipation (250 [mW]) in the transistor with a fully contacted emitter area in Fig. 16(a). Note that the collector current distribution is only plotted in the active emitter region.

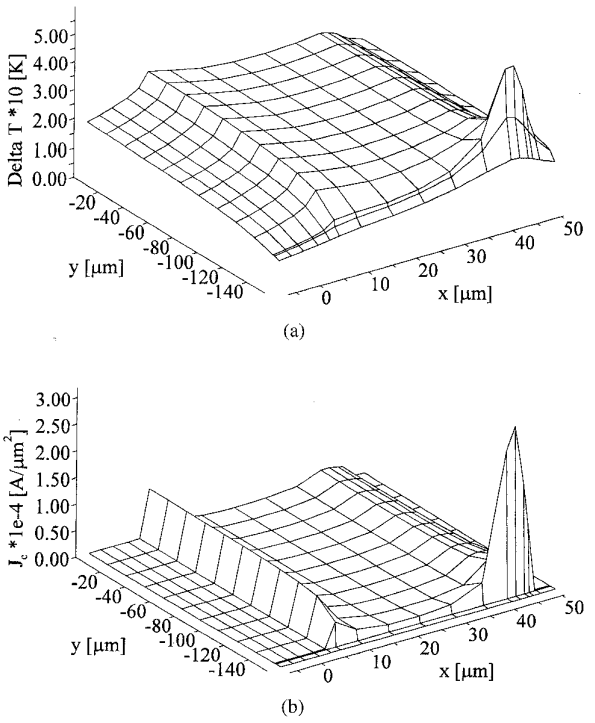


Fig. 18. Lateral temperature profile (a) and collector current density (b) in transistor with partly contacted emitter area with $V_{BE} = 1$ [V] and $V_{CE} = 5.0$ [V]. Note that the collector current distribution is only plotted in the active emitter region.

In Fig. 16, the transistor with the partially contacted emitter not only shows a lower temperature rise (because of the lower total collector current), but also a more uniform temperature profile. For a good comparison, however, the temperature profiles in the transistors should be compared under equal power dissipation. This is shown in Fig. 17(a) and 17(b) where the temperature distribution and collector current distribution of the transistor with the partially contacted emitter are shown in the case of 250 [mW] dissipation [as for the transistor with a fully contacted emitter in Fig. 16(a)]. A comparison of Fig. 17(a) with Fig. 16(a) shows that the temperature distribution in the partly contacted transistor is more uniform under equal power dissipation conditions. When it is considered that the highest peak in the temperature distribution usually determines the location at which thermal runaway is initiated, the situation with a flat temperature distribution (partly contacted emitter) is to be preferred for maximum thermal stability. When the current and voltage are further increased, the transistor with the partly contacted emitter also develops a peak in the current and temperature distribution (see Fig. 18). This peak that indicates where thermal runaway will eventually occur is located at the short edge of the emitter contact. The reason for this location is that the internal emitter potential is the lowest under the emitter contact and the internal base potential under the emitter is the highest at the short edge of the emitter because the extrinsic base region around the emitter causes a relatively low resistive path from the base contact to the base region under the short emitter edge. Therefore, the base-emitter voltage under the emitter contact has a maximum at this edge, and this is where the base and collector current concentrate. This 3-D effect proves to be quite important for the transistor behavior in the high current regime. For many practical transistor geometries, 2-D simulation methods will not be accurate here. The influence of the third dimension on the effective base resistance was also shown in [8].

VI. Conclusion

A method for circuit extraction from layout information for simulation of electrothermal device behavior has been described. This simulation method can be placed between nonisothermal device simulation and (conventional) circuit simulation. It is useful for electrothermal simulation of device structures that are too complex for 3-D nonisothermal device simulation. As an example, the simulation of self-heating effects in a bipolar transistor and the influence of base and emitter contacting has been shown.

Our simulation method has a number of advantages and compared to the alternative, 3-D thermal-electrical device simulation (using the semiconductor transport equations).

1) It allows circuit designers to simulate the electrothermal device characteristics with the circuit simulator they are used to.
2) The layout-to-circuit extraction makes it very easy to generate circuits and to study device behavior for different layouts.

3) The amount of necessary user interaction can be kept low. When a bipolar transistor is simulated, the only necessary input consists of the *layout*, the refinement of the coarse initial *electrical grid*, the *chip dimensions*, and the *thermal grid*. The whole layer structure used to build the 3-D device structure is constant for a given process, and can be read from a library file.
4) Next to the application for thermal-electrical simulation of standard devices like a bipolar transistor, the circuit simulation method with the layout extractor can be useful for many other 3-D problems (e.g., the determination of base resistances in 3-D structures or calculation of 3-D potential distributions). For these problems that do not concern standard devices, the user can define all models, layers, and connections between layers to create the desired 3-D structure.
5) For complex device structures (e.g., for power transistors with multiple emitter fingers as in [12]), the amount of CPU time for this method is much lower than for a device simulator. This makes it suitable for simulation of characteristics with many bias points.
6) Because a circuit simulator is used, the generated 3-D circuit model can be simulated together with external components.
7) The accuracy of the simulation method for a given grid is relatively good. This is because the values and parameters of the circuit elements (e.g., resistors and transistors) that are used in the circuit representation are directly obtained from measured quantities like the sheet resistance and the measured transistor parameters.

ACKNOWLEDGMENT

The authors wish to thank A. van der Scheer from the University of Twente for his support in the simulation work.

REFERENCES

[1] P. B. M. Wolbert, A. J. Mouthaan, and H. Wallinga, "Electrical and thermal analysis of ultra thin SOI MOSFET's," in *Proc. Nasecode VII*, Apr. 1991, pp. 183–184.
[2] P. W. Webb, "Thermal modeling of power Gallium arsenide microwave integrated circuits," *IEEE Trans. Electron Devices*, vol. 40, pp. 867–877, May 1993.
[3] J. A. Higgins, "Thermal properties of power HBT's," *IEEE Trans. Electron Devices*, vol. 40, pp. 1271–2177, Dec. 1993.
[4] G. K. Wachutka, "Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling," *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 1141–1149, Nov. 1990.
[5] *MEDICI User's Manual*, Technology Modeling Associates, Inc., Palo Alto, CA, Mar. 1992.
[6] P. B. M. Wolbert, "Modeling simulation of semiconductor devices in TRENDY," Ph.D. thesis, Univ. Twente, Enschede, The Netherlands, 1991.
[7] M. Latif and P. R. Bryant, "Multiple equilibrium points and their significance in the secondary breakdown of bipolar transistors," *IEEE J. Solid-State Circuits*, vol. SC-16, pp. 8–15, Feb. 1981.
[8] A. Sadovnikov and D. J. Roulston, "Quasithree-dimensional modeling of transistor characteristics," *IEEE Trans. Computer-Aided Design*, vol. 12, pp. 1742–1748, Nov. 1993.
[9] R. H. MacNeal, "An asymetrical finite difference network," *Q. Appl. Math.*, vol. 11, no. 3, pp. 295–310, 1953.
[10] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed. New York, Wiley, 1986, p. 334.
[11] N. R. Swart and A. Nathan, "Coupled electrothermal modeling of microheaters using SPICE," *IEEE Trans. Electron Devices*, vol. 41, pp. 920–925, June 1994.
[12] B. H. Krabbenborg, H. C. de Graaff, A. J. Mouthaan, H. Boezen, A. Bosma, and C. Tekin, "3-D thermal-electrical simulation of breakdown in a BJT using a circuit simulator and a layout-to-circuit extraction tool," in *Proc. SISDEP'93*, Vienna, Austria, Sept. 6–8, 1993, pp. 57–60.

B. H. Krabbenborg was born in Lichtenvoorde, The Netherlands, in 1966. He received the M.Sc. and the Ph.D. degrees in electrical engineering from the University of Twente, Enschede, The Netherlands, in 1990 and 1994, respectively. His Ph.D. dissertation dealt with multidimensional modeling and simulation of coupled thermal-electrical (breakdown) mechanisms in bipolar transistors.

He has been working on reliability aspects of processes and devices with the Quality and Reliability Department of Consumer IC's, Philips Semiconductors, Nijmegen, The Netherlands, since 1995.

A. Bosma was born in Achtkarspelen, The Netherlands, in 1967. He received the M.Sc. degree in electrical engineering from the University of Twente, Enschede, The Netherlands, in 1993. His Masters degree project dealt with the extraction of 3-D coupled electrical-thermal circuit models for simulation of self-heating effects in bipolar transistors.

He joined the Dutch Institute for Animal Health Care where he works on computer models to predict the spread of contagious diseases in 1993.

H. C. de Graaff (M'92) was born in Rotterdam, The Netherlands, in 1933. He received the M.Sc. degree in electrical engineering from the University of Delft, The Netherlands, in 1956, and the Ph.D. degree from the University of Technology, Eindhoven, The Netherlands, in 1975.

He joined Philips Research Laboratories, Eindhoven, The Netherlands, in 1964, and has been working on thin film transistor, Most, bipolar devices, and materials research on polycrystalline silicon. His present field of interest is device modeling for circuit simulation. Since his retirement from Philips Research in Nov. 1991, he has been a Consultant to the University of Twente, Enschede, The Netherlands, and the Technical University of Delft.

A. J. Mouthaan (M'89) received the B.Sc., M.Sc., and Ph.D. degrees, all in electrical engineering from the University of Twente, Enschede, The Netherlands, in 1973, 1977, and 1986, respectively. His Ph.D. research dealt with high energy ion implantation for the fabrication of vertical devices.

After a short period as a Research Assistant with the University of Twente, he joined the University of Zambia as a Lecturer to establish a Department of Electronics there. After two years with the University of Zambia as a Senior Lecturer, he returned to the University of Twente as an Associate Professor, heading research groups in modeling and simulation, and CAD for the synthesis of analog circuits. He also lectures on IC-Technology, and was a member of the Board of Deans of the Faculty of Electrical Engineering.