Advanced thermal simulation of SiGe:C HBTs including back-end-of-line

Vincenzo d'Alessandro [a,*], Alessandro Magnani [a], Lorenzo Codecasa [b], Niccolò Rinaldi [a], Klaus Aufinger [c]

[a] Department of Electrical Engineering and Information Technology, University Federico II, 80125 Naples, Italy
[b] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, I-20133 Milan, Italy
[c] Infineon Technologies AG, Neubiberg 85579, Germany

ABSTRACT

Advanced 3-D thermal simulations of state-of-the-art SiGe:C HBTs are performed, which ensure improved accu-racy with respect to conventional approaches. The whole back-end-of-line architecture is modeled so as to account for the cooling effect due to the upward heat flow. Moreover, a nonuniform power density is considered to describe the heat source, and thermal conductivity degradation effects due to germanium, doping profile, and phonon scattering in narrow layers are implemented. The numerical thermal resistances are compared with those experimentally evaluated by means of a robust technique relying on the temperature dependence of the base-emitter voltage.

1. Introduction

The performance of bipolar transistors for high-frequency applications is steadily growing thanks to aggressive lateral scaling and device architecture improvement. This is especially true for the silicon germanium (SiGe) bipolar technology, which represents a good alternative to the gallium arsenide counterpart while being based on Si-like fabrication processes. Nowadays SiGe heterojunction bipolar transistors (HBTs) enjoy beneficial properties [1] that favor their adoption in a large variety of mm-wave and near-THz applications, like high-speed communications, optical transmission, automotive radar modules, industrial sensors, and medical equipment [2,3]. Important contributions to this trend have been offered by the European DOTFIVE [4,5] and DOTSEVEN projects, the latter being expected to push the maximum oscillation frequency f_{max} to 0.7 THz.

Unfortunately, thermal issues have become a serious concern in SiGe HBTs due to the concurrent impact of the following factors: (i) the intrinsic device shrinking has induced a growth in power density within the base-collector depletion region; (ii) the trench isolation – exploited to reduce parasitics and increase f_{max} – gives rise to a limited heat spreading since trenches are filled with low conductivity materials [6–9]. This mechanism is even exacerbated by the lateral scaling, which results in a horizontal reduction of the Si volume embraced by trenches; (iii) HBTs are operated at high current densities to boost the frequency performance, which entails a further increase in dissipated power density [2]. Owing to these considerations, thermal effects can be reviewed

* Corresponding author.
  E-mail address: vindales@unina.it (V. d'Alessandro).

as an undesired, yet unavoidable, by-product of the technology evolution. The enhanced heat generation for a given dissipated power and the reduction in heat removal have pushed the thermal resistances R_{TH} of SiGe HBTs into the thousands of K/W [10–12] and even beyond 10^4 K/W for small emitter windows, as evidenced by recent experimental campaigns conducted on transistors fabricated by STMicroelectronics [7] and Infineon Technologies [13] (hereinafter referred to as STM and IFX, respectively) within the framework of the DOTFIVE project.

Thermal effects can severely degrade both the dc bias (e.g., [14]) and the low-frequency (up to 1 MHz) behavior, affect the long-term reliability, and even trigger destructive instability phenomena. Consequently, care must be taken in assessing the impact of technology on the thermal behavior. A viable strategy involves the adoption of 3-D thermal simulations based on the finite element method (FEM). In [15], nonlinear steady-state, large signal, and sinusoidal thermal analyses of an STM SiGe:C HBT (with drawn emitter equal to $0.27 \times 10\ \mu m^2$) were carried out with Sentaurus TCAD from Synopsys [16] by accounting for the back-end-of-line (BEOL) architecture; the R_{TH} was found to be in a fairly good agreement with the one measured following the method presented in [17], although no thermal conductivity degradation mechanisms (e.g., due to high doping) were accounted for. The BEOL was found to play a marginal role due to the absence of the metal-via stack above the emitter. This analysis has been recently extended to examine the influence of BEOL on the thermal behavior of multi-finger devices, with emphasis on the coupling among fingers [18]. In [7], we adopted the software package COMSOL [19] to analyze self-heating (SH) in several STM SiGe:C HBTs. In spite of their geometrical complexity, the devices were reproduced with a very high accuracy;

however, the BEOL was not included, the structures being superiorly limited to the emitter, base, and collector contacts with adiabatic top. Widely accepted literature values were considered for the material parameters. Unfortunately, although the upward heat flow was improperly suppressed, the numerical R_{TH}s were found to underestimate by about 20–25% the experimental values determined through the technique proposed in [13], independently of technological stage and emitter window.

In this paper, we improve the approach exploited in [7] through a multi-fold action: (i) the whole BEOL architecture is taken into account; (ii) a more realistic nonuniform power density distribution computed with numerical device simulations is considered to model the power dissipation region; (iii) thermal conductivity reduction mechanisms due to high doping and phonon scattering along the edges of narrow paths for the heat flow are implemented. The advanced simulation approach is applied to various state-of-the-art IFX SiGe:C HBTs with different emitter windows, yet corresponding to the same technological stage. The resulting FEM R_{TH}s are compared to those measured with a more accurate variant of the extraction technique presented in [13].

The remainder of this work is articulated as follows. In Section 2, a short description of the devices under test (DUTs) and experimental setup is provided. The R_{TH} extraction technique is explained in Section 3. The simulation approach is described in Section 4. Section 5 addresses and discusses the results; in particular, the individual impact of all the modeled effects is examined. Conclusions are then drawn in Section 6.

2. Devices under test and experimental setup

Measurements were performed on single-emitter SiGe:C NPN HBTs with one base contact and one collector contact (BEC configuration) manufactured by IFX within the DOTFIVE project; more specifically, the DUTs belong to the latest technology stage of the project development, also denoted as set #3 in [13,20]. Further details – including the key figures of the sets – can be found in [13] and are omitted here for the sake of brevity.

The on-wafer experimental campaign was carried out by means of a PM5 Karl Suss probe station. The electrical signals were handled by an HP4142B parameter analyzer. The dc common-emitter (CE) and common-base (CB) measurements were performed by using PH100 probeheads equipped with tungsten needles. Countermeasures were taken to prevent the possibly destructive oscillations that may arise under CB conditions; in particular, ferrite beads were hooked up to the cables in the close proximity of the DUTs. The backside temperature T_B was set to assigned values thanks to a thermochuck controlled by an ATT heating/cooling system.

3. Thermal resistance extraction approach

The most widespread dc extraction approach for the thermal resistance R_{TH} of bipolar transistors involves the measurement of the base-emitter voltage V_{BE} as a function of T_B to employ the resulting dependence as a *thermometer*. However, applying this standard method (presented in slightly different variants in [6,17,21]) to HBTs suffering from significant thermal effects would lead to a thermometer calibration jeopardized by SH, in turn inducing an R_{TH} overestimation. Techniques to prevent this issue were proposed in [13,22] (the latter used in e.g., [23]). Here we adopt an improved version of the approach in [13] that allows purifying the extraction from the influence of the Early effect (which, in spite of its *electrical* nature, can be misinterpreted as additional SH [24]), and can be described as follows.

In the absence of high-injection (HI) and impact-ionization (II) effects, the collector current I_C of a bipolar transistor can be described by

the simple model

$$I_C = \left(1 + \frac{V_{CB}}{V_{AF}}\right) A_E J_{S0} \exp\left[\frac{V_{BEj} + \phi(I_E) \cdot \Delta T_j}{\eta V_{T0}}\right] \tag{1}$$

where

- V_{BEj} is the internal (junction) base-emitter voltage, given by

$$V_{BEj} = V_{BE} - R_E I_E - R_B I_B \tag{2}$$

I_B, I_E, R_B, R_E being the base and emitter currents [A] and the parasitic base and emitter resistances [Ω];
- $A_E = W_E \times L_E$ [μm^2] is the effective emitter area, W_E and L_E being the emitter width and length, respectively[1];
- J_{S0} [$A/\mu m^2$] is the reverse saturation current density, η (≥ 1) is the ideality factor, and V_{T0} ($= 25.86$ mV) is the thermal voltage, all at temperature $T_0 = 300$ K;
- V_{AF} [V] is the forward Early voltage, which can be easily extracted from the output characteristics in a preliminary stage;
- ϕ [V/K] (>0) is the temperature coefficient of V_{BEj} (in absolute value), and ΔT_j is defined as $T_j - T_0$, T_j [K] being the (average) temperature over the base-emitter junction. This implies that Eq. (1) accounts for the temperature dependence of I_C ($\approx I_E$) by means of a V_{BEj} shift, by keeping J_{S0}, η, and V_{T0} at their T_0 values (e.g., [25]). The ϕ dependence on I_E is described with [13,26]

$$\phi(I_E) = \phi_0 - \eta\frac{k}{q} \ln\left(\frac{I_E}{A_E J_{S0}}\right) \tag{3}$$

Parameters J_{S0}, η in Eq. (1) and ϕ_0 in Eq. (3) can be optimized by resorting to the following procedure. First, the I_C–V_{BE} characteristic of the DUT is measured at various T_B under CE conditions by keeping V_{CE} small and sweeping V_{BE} up to values sufficiently low to safely neglect SH, HI, II, Early, and resistive effects; as a result, the I_C–V_{BE} curves can be modeled by

$$I_C = A_E J_{S0} \exp\left[\frac{V_{BE} + \phi(I_E) \cdot (T_B - T_0)}{\eta V_{T0}}\right] \tag{4}$$

which can be derived from Eq. (1) by considering $T_j = T_B$, $V_{CB}/V_{AF} = 0$, and $V_{BEj} = V_{BE}$. Parameters J_{S0} and η are then tailored to match the experimental curve at $T_B = T_0$ with

$$I_C = A_E J_{S0} \exp\left(\frac{V_{BE}}{\eta V_{T0}}\right) \tag{5}$$

and ϕ_0 is calibrated so as to ensure a good agreement between all the I_C–V_{BE} characteristics (at different T_B) and the model given by Eq. (4) with Eq. (3). Fig. 1 shows this comparison for the DUT with $A_E = 0.13 \times 2.73$ μm^2, chosen as a reference throughout the whole analysis.[2]

Once ϕ_0 is known, Eq. (3) is assumed to hold also at medium current levels [13].

A CB V_{BE}–V_{CB} characteristic is then measured at $T_B = T_0$ by sweeping V_{CB} up to values low enough to neglect II (0.8 V in our case), and keeping I_E sufficiently high to give rise to perceptible SH, yet small enough to

---

[1] Details on the definition of "effective" emitter width W_E, length L_E, and area A_E for the DUTs are given in [13].
[2] The experimental R_{TH}s of the DUTs are relatively lower than those reported in [13] due to the following concurrent reasons: here (i) the Early effect is accounted for; (ii) the applied I_E is lower, thus fully suppressing nonlinear thermal effects. In addition, the experimental campaign was conducted on another die.
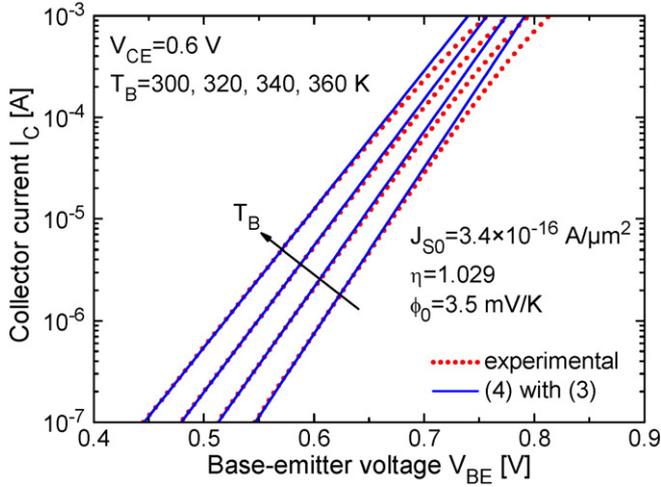
Fig. 1. Collector current I_C versus base-emitter voltage V_{BE} at $V_{CE} = 0.6$ V and $T_B = 300$, 320, 340, 360 K for the reference transistor; experimental results (dotted lines) are compared with the calibrated model (4) using Eq. (3) (solid).

prevent HI and nonlinear thermal effects. From Eq. (1) with $I_C \approx I_E$, the following expression is derived for V_{BE}:

$$V_{BE} \approx R_E I_E - \phi(I_E) \cdot \Delta T_j + \eta V_{T0} \ln\left(\frac{I_E}{A_E J_{S0}} \cdot \frac{1}{1 + \frac{V_{CB}}{V_{AF}}}\right) \quad (6)$$

where ΔT_j is given by the thermal equivalent of Ohm's law

$$\Delta T_j = R_{TH} \cdot (V_{BE} I_E + V_{CB} I_C) \approx R_{TH} \cdot (V_{BE} + V_{CB}) \cdot I_E \quad (7)$$

By substituting Eq. (7) into Eq. (6), it is obtained that

$$V_{BE} \approx \frac{1}{1 + \phi(I_E) R_{TH} I_E} \cdot \begin{bmatrix} R_E I_E - \phi(I_E) R_{TH} V_{CB} I_E + \\ + \eta V_{T0} \ln\left(\frac{I_E}{A_E J_{S0}} \cdot \frac{1}{1 + \frac{V_{CB}}{V_{AF}}}\right) \end{bmatrix} \quad (8)$$

whence the (negative) slope γ of the V_{BE}–V_{CB} curve is (if $V_{AF} \gg V_{CB}$, which holds for the DUTs)

$$\gamma = \frac{dV_{BE}}{dV_{CB}} \approx \frac{1}{1 + \phi(I_E) R_{TH} I_E} \cdot \left[-\phi(I_E) R_{TH} I_E - \frac{\eta V_{T0}}{V_{AF}} \right] \quad (9)$$

From Eq. (9) it can be straightforwardly found that

$$R_{TH} = \frac{|\gamma| - \frac{\eta V_{T0}}{V_{AF}}}{(1 - |\gamma|)\phi(I_E) I_E} \quad (10)$$

which reduces to [13]

$$R_{TH} = \frac{|\gamma|}{(1 - |\gamma|)\phi(I_E) I_E} \approx \frac{|\gamma|}{\phi(I_E) I_E} \quad (11)$$

for bipolar transistors subject to negligible Early effect. The accuracy of the approach was verified by means of the following tests:

- in [27], Eq. (11) was applied to an IFX SiGe:C HBT simulated with an advanced 2-D tool accounting for SH in the Boltzmann transport equations of electrons and holes, the solver being based on Spherical Harmonic Expansion [28]. The extracted R_{TH} was found to be consistent with the average ΔT_j over the base-emitter junction;
- Eqs. (10) and (11) were applied in ADS [29] to HICUM [30] models provided with an assigned (known) R_{TH} and calibrated on dc I—V

curves measured on IFX SiGe:C HBTs. It was found that Eq. (10) ensures an excellent accuracy (the errors amounted to about -0.1%), yet also Eq. (11) is fairly reliable ($+4\%$), since these devices show $V_{AF} > 100$ V. The adoption of Eq. (11) for bipolar transistors subject to significant Early effect would instead lead to intolerable R_{TH} overestimations.

The thermal resistance can thus be determined from Eq. (10) by measuring γ and making use of coefficient $\phi(I_E)$ given by Eq. (3). The extraction of γ for the reference transistor is depicted in Fig. 2.

4. Simulation approach

Thermal simulations were performed with COMSOL. The 3-D geometry of the DUTs was created from GDS masks layout by exploiting Griesmann's GDS II libraries [31] and information on technology (e.g., the thickness of the layers).

The whole BEOL structure, comprising five levels of metal (copper) and related interconnections (tungsten), was also built, as well as the external pads. This allows quantifying the cooling influence due the upward heat flow (often disregarded in literature), which is expected to be relevant since – differently from [15] – a metal-via stack is located over the emitter. In [23], it is indeed suggested to conceive and design improved wiring techniques to significantly reduce R_{TH}.

Owing to the presence of regions with horizontal dimensions much larger than the thickness, critical meshing issues were encountered, which were tackled by domain decomposition and calibration of scaling parameters. The distribution of grid points was optimized by resorting to selective refinement features.

The schematic cross-section of the DUTs – assumed to lie in the (x, z) plane – is represented in Fig. 3, which evidences the shallow and deep trenches surrounding the heat generation region, hereinafter also referred to as heat source (HS). Fig. 4 shows the 3-D mesh of the reference transistor, which is composed by 1.35 million tetrahedra of grossly different dimensions, corresponding to about 1.8 million degrees of freedom.

A. Heat source

In bipolar transistors the power dissipation occurs at the base-collector depletion region (also denoted as space-charge region or SCR). In conventional approaches for thermal simulations, if the emitter window is rectangular, this region is modeled as either a rectangular or a parallelepiped HS (e.g., [7,15]), both with uniform power density. However,
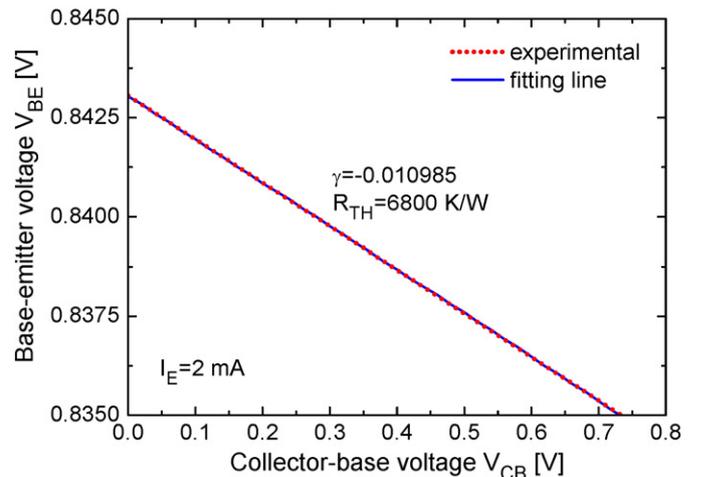

Fig. 2. Extraction of the slope γ of the measured I_E-constant V_{BE}–V_{CB} curve for the reference HBT.

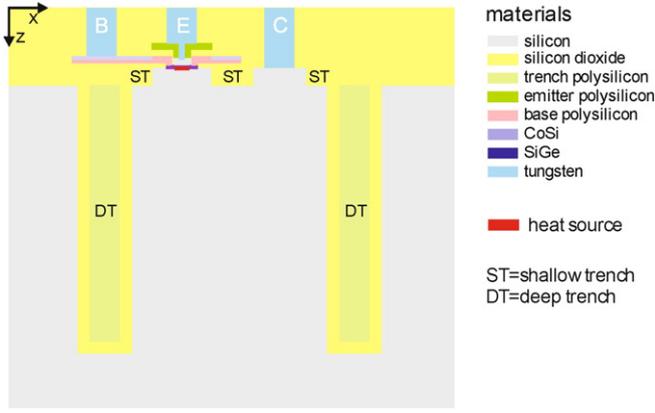**Fig. 3.** Schematic representation (limited to the innermost tungsten contacts) of the typical cross-section of the BEC DUTs.

Fig. 4. Detail of the COMSOL mesh of the reference HBT.

Fig. 5. Nonuniform power density resulting from a 2-D HD device simulation for the reference transistor; evidenced are the positions of the base-emitter (where the Peltier effect occurs) and base-collector metallurgical junctions.

these simplified choices may lead to inaccurate results: in the first case, a significant temperature overestimation is obtained at the depth where the HS is located, while in the second an underestimation is achieved if the HS geometry is assumed to coincide with the whole lightly-doped collector [32].

Here we accurately model the heat dissipation region by resorting to simulations of the IFX DUTs performed with Sentaurus according to the following procedure. The 2-D domains corresponding to the DUTs (with $W_E = 0.13, 0.23, 0.27, 0.55\ \mu m$) were realized with the Structure Editor tool by tuning the doping and Ge profiles so as to match those measured by secondary ion mass spectrometry. Simulations were carried out by using a calibrated hydrodynamic (HD) model with transport parameters optimized for SiGe:C HBTs [33,34] to accurately capture non-local and non-quasi-static effects due to vertical scaling; all mechanisms playing a role were accounted for, including SRH recombination with high field enhancement and doping dependence [35]. A good agreement was obtained between the experimental and simulated Gummel plots (the latter under isothermal conditions at T_0) of the DUTs in the V_{BE} ranges where SH is negligible. The nonuniform power density distribution in the (x, z) plane was determined as the scalar product between electric field and current density[3] under $V_{BE} = 0.8$ V and $V_{CB} = 0.4$ V to ensure an acceptable description of the real power density profile arising along the dc V_{BE}–V_{CB} curve used for the experimental R_{TH} extraction. The power density for the reference transistor is shown in Fig. 5.

The HSs in the COMSOL structures were built with the power density pattern obtained by reproducing the distribution computed by Sentaurus in the (x, z) plane, and assuming a uniform density along the device length (i.e., along the y axis orthogonal to the cross-section depicted in Fig. 3).

B. Thermal conductivity reduction due to germanium

Thermal simulations are usually performed by setting the thermal conductivities k [W/m K] of the materials to values measured from bulk samples (listed in Table 1). However, in practical cases, many effects concur to reduce k, which can be even position-dependent within the same material.

In the SiGe alloy, k can be determined by the following analytical law that accounts for the z-dependent Ge mole fraction x_{Ge} [37]:

$$k_{SiGe} = \left[\frac{1-x_{Ge}}{k_{Si}} + \frac{x_{Ge}}{k_{Ge}} + \frac{(1-x_{Ge})x_{Ge}}{c_k}\right]^{-1} \quad (12)$$

where k_{Si} and k_{Ge} are the thermal conductivities of pure Si and Ge,

respectively, and c_k is a bowing factor equal to 2.8 W/m K. Due to the k lowering imposed by Eq. (12), the SiGe layer behaves as a barrier for the heat flow from the HS to the emitter [9]. Eq. (12) was implemented in COMSOL, as it was in [7].

C. Thermal conductivity reduction due to doping

The thermal conductivity is dependent on doping due to the enhanced phonon-impurity scattering, as experimentally observed in [38–40]. Recent molecular dynamics results provide a straightforward way to account for this effect with [40]

$$k_{Si,doped}(k_{SiGe,doped}) = \frac{k_{Si,bulk}(k_{SiGe})}{1 + A \cdot \left(\frac{N}{N_{norm}}\right)^{\alpha}} \quad (13)$$

where N [cm^{-3}] is the position-dependent total doping concentration (acceptors and donors), $N_{norm} = 10^{20}$ cm^{-3}, while the fitting parameters are $A = 0.74186$, $\alpha = 0.7411$ for boron [40], and $A = 1.698$, $\alpha =$

Table 1
Bulk thermal conductivity values adopted in the simulations.

Material	Thermal conductivity k_{bulk} [W/m K]
Silicon	148
Germanium	60
Silicon dioxide	1.4
Tungsten	177
Copper	390
Emitter polysilicon	40
Base polysilicon	30
Trench polysilicon	20
Cobalt silicide	9.6

[3] It must be remarked that more advanced approaches to model the power density (e.g., [36]) can in principle be used instead of the one proposed in this work.
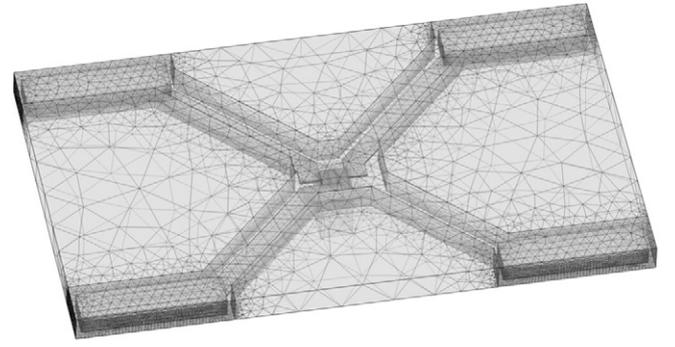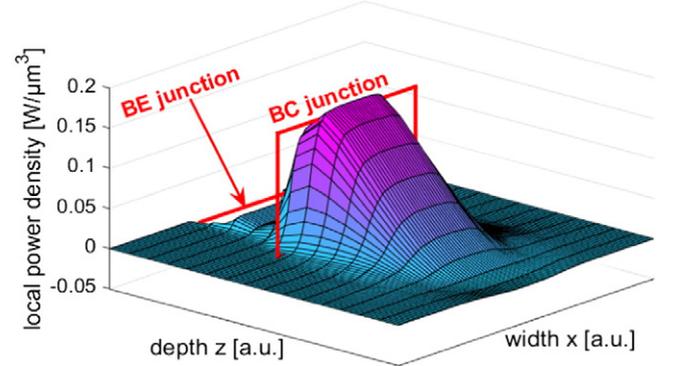
0.8251 for arsenic, as obtained with a calibration procedure relying on experimental results [39]. The thermal conductivities of boron- and arsenic-doped Si are plotted in Fig. 6 as a function of N. It can be observed that a significant k reduction with respect to the low-N $k_{Si,bulk}$ value (148 W/m K) takes place above 10^{19} cm^{-3}, which reveals that the assumption of $k = k_{Si,bulk}$ within the Si regions can provide unreliable results. Eq. (13) was enabled in the overall structure of the DUTs.

D. Thermal conductivity reduction in narrow layers

The heat propagation through laterally thin layers can be significantly jeopardized by the phonon scattering with the layer boundaries [41, 42]. In SiGe:C HBTs, where the heat flow is mostly vertical, scattering mechanisms – expected to be exacerbated in narrow (low-W_E) transistors – can take place along device portions like (from the top) emitter tungsten contact, Si emitter, SiGe base, and Si volume surrounded by shallow trench. We accounted for this deleterious effect by making use of the simple analytical method proposed in [43], which leads to a reduced anisotropic thermal conductivity, the components of which are x-dependent and given by

$$\frac{k_{y,z}(x')}{k_{Si,doped}(k_{SiGe,doped}, k_{contact})} = 1 - \frac{1}{2} \exp\left[-\left(\frac{x'}{x_{charyz}}\right)^{0.75}\right] - \frac{1}{2} \exp\left[-\left(\frac{1-x'}{x_{charyz}}\right)^{0.75}\right]$$
$$x_{charyz} = 0.32 \cdot \frac{\Lambda}{W} \tag{14}$$

$$\frac{k_x(x')}{k_{Si,doped}(k_{SiGe,doped}, k_{contact})} = 1 - \frac{1}{2} \exp\left[-\left(\frac{x'}{x_{charx}}\right)^{0.95}\right] - \frac{1}{2} \exp\left[-\left(\frac{1-x'}{x_{charx}}\right)^{0.95}\right]$$
$$x_{charx} = 0.72 \cdot \frac{\Lambda}{W} \tag{15}$$

where x' is defined as x/W, W being the layer width (along x), Λ is the mean free path for phonons, assumed to be equal to 300 nm in Si, SiGe layers, and 40 nm in the tungsten emitter contact. It must be emphasized that for layers narrower than 1 μm along x, applying Eqs. (14) and (15) implies a noticeable conductivity reduction compared to the bulk value; as an example, for W = 0.2 μm, it is obtained that $k_y/k_{bulk} = k_z/k_{bulk} = 0.64$, $k_x/k_{bulk} = 0.38$ at the center (along x) of the layer, and $k_y/k_{bulk} = k_z/k_{bulk} = 0.41$, $k_x/k_{bulk} = 0.3$ at the lateral edges, where the phonon scattering occurs.

Fig. 7 shows k_z versus z along a line crossing the center of the reference HBT, as obtained by concurrently including all the aforementioned
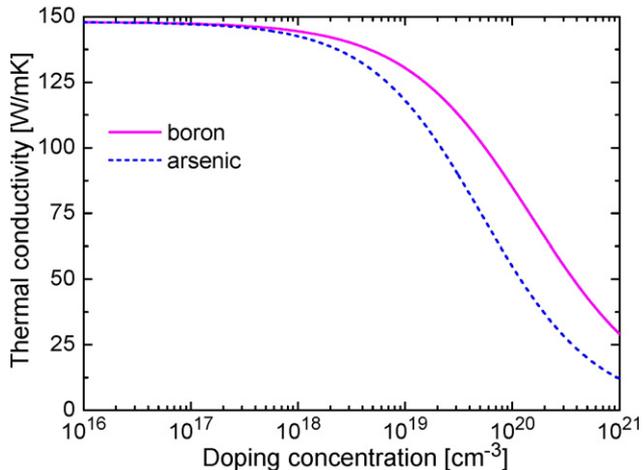

Fig. 6. Thermal conductivity as a function of doping concentration for boron- (solid line) and arsenic-doped (dashed) Si.
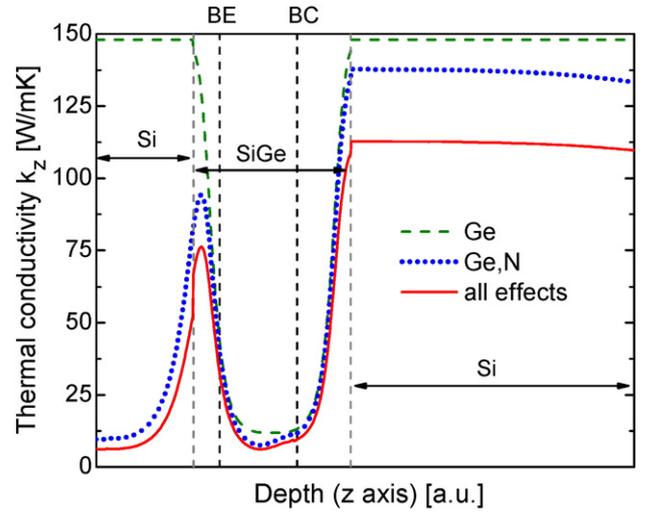

Fig. 7. Thermal conductivity component k_z along a vertical cut crossing the center of the reference HBT by enabling Eqs. (12) (dashed line), (12) and (13) (dotted), (12)–(15) (solid). Evidenced are the positions of the base-emitter and base-collector metallurgical junctions, as well as the thickness of the SiGe layer.

effects, by solely considering the impact of Ge, and by accounting for both Ge and doping (in the latter cases, $k_z = k$). It can be inferred that

- the SiGe region is impacted by all mechanisms; as a consequence, it exhibits extremely low thermal conductivity. If only the Ge influence is taken into account, its minimum amounts to 11.8 W/m K, which reduces to 7.6 W/m K by considering also the doping; as the model (14), (15) is enabled, the minimum of component k_z becomes equal to 6.1 W/m K (along the device center) and 3.6 W/m K (on the lateral edges of the SiGe layer);
- a conductivity peak is detected at the border between SiGe and Si due to the low doping in this region;
- owing to the aforementioned reasons, the narrow highly-doped Si emitter region suffers from a significant conductivity reduction;
- the thermal conductivity of the highly-doped sub-collector is halved with respect to the bulk value (not shown in Fig. 7 since the corresponding z range is beyond the represented one).

5. Results and discussion

COMSOL 3-D steady-state simulations were performed by applying an adiabatic boundary condition at the top and lateral faces of the structure,[4] and an isothermal condition on the backside ($T_B = T_0$). The thermal resistance R_{TH} was determined by evaluating the average of the temperature rise ΔT_j over the base-emitter junction, which has the most relevant impact on the behavior and performance of the device [25], and normalizing to the dissipated power. Nonlinear thermal effects were not accounted for since the experimental R_{TH}s were extracted by carefully keeping ΔT_j low. Each simulation required <10 min with 16 GB RAM occupation on a PC equipped with a quad-core i7-3820QM and 32 GB RAM.

Results are reported in Fig. 8, which shows

- the R_{TH}s determined through the improved experimental technique outlined in Section 3;
- the R_{TH}s simulated with COMSOL by considering the full advanced approach described in Section 4 (denoted as approach A), i.e., by including the BEOL and accounting for the nonuniform power density pattern and the conductivity degradation mechanisms;

---

[4] The lateral size of the simulated domain was chosen to be much larger in comparison with the transistor (located at the domain center) to safely neglect an unrealistic heating effect due to the adiabatic lateral faces.
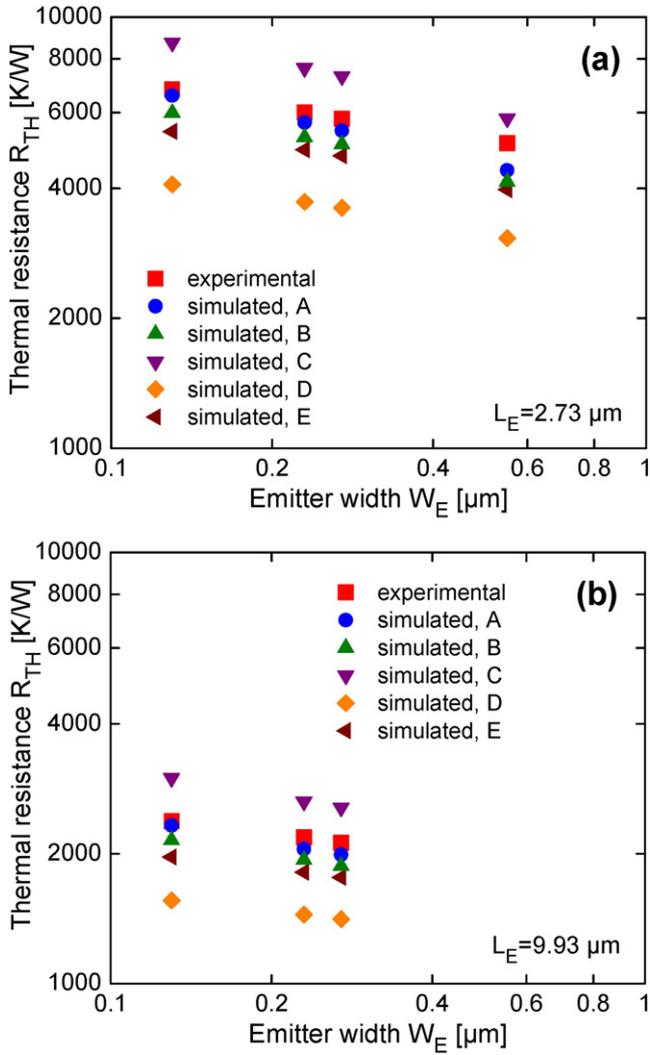
Fig. 8. Thermal resistances as a function of emitter width for (a) $L_E = 2.73\,\mu m$ and (b) $L_E = 9.93\,\mu m$: experimental (squares) values are compared with those calculated through the simulation approaches A (circles), B (triangles), C (flipped triangles), D (rhombi), and E (left-oriented triangles).

- the R_{TH}s calculated with COMSOL by modeling the HS through a standard parallelepiped-shaped HS with uniform power density geometrically coinciding with the whole power-dissipating region,[5] while considering all other effects and the BEOL structure (approach B);
- the R_{TH}s computed with COMSOL by accounting for an HS with nonuniform power density and replacing the metal with silicon dioxide (SiO_2) in the BEOL architecture so as to virtually exclude it, the tungsten contacts being instead included (approach C);
- the R_{TH}s evaluated with COMSOL by restoring the BEOL, and considering "uncorrected" bulk values for the thermal conductivities and a standard parallelepiped-shaped HS (approach D);
- the R_{TH}s computed with COMSOL by disregarding the above effects and excluding the BEOL so as to emulate a traditional approach (approach E).

Similar results were obtained for the short ($L_E = 2.73\,\mu m$) and long ($L_E = 9.93\,\mu m$) DUTs; thus, the comments are mainly focused on the short ones for the sake of brevity.

By using approach A, the R_{TH} of the reference transistor was calculated to be 6571 K/W, which is in excellent agreement (-3.4%) with the experimental value (6800 K/W); conversely, a relatively high underestimation

---

[5] More specifically, the HS was chosen to vertically cover 75% of the self-aligned ion-implanted collector.

(-13.6%) was obtained for the widest ($W_E = 0.55\,\mu m$) device, the numerical and measured R_{TH}s being 4408 and 5100 K/W, respectively. A post-processing analysis revealed a markedly nonuniform temperature distribution along x over the base-emitter junction compared to low-W_E transistors, which can be ascribed to the concurrent action of the low k_{SiGe} and the narrow tungsten emitter contact (the width of which does not scale with W_E). As a consequence, the evaluation of R_{TH} with a standard geometrical average of ΔT_j over the whole junction is likely to be incorrect. The accuracy for wide devices can be thus improved by exploiting more sophisticated approaches for the ΔT_j average (e.g., [44]) that will increase the FEM R_{TH}.

If the approach B (with the traditional HS representation) is adopted, the FEM R_{TH} lowers (compared to A) from -8.9% for the reference device to -5.9% for the HBT with $W_E = 0.55\,\mu m$, where the base-emitter temperature is uneven. Hence, it can be stated that the HS representation plays a non negligible role.

By making use of the BEOL-free approach C (upward heat flow almost annihilated), the numerical R_{TH} of the reference DUT grows to 8712 K/W, which corresponds to $+28\%$ with respect to the experimental value; this means that, although the low-conductivity SiGe base and Si emitter concur to limit the upward heat flow, the BEOL effectively extracts heat from the emitter. This effect is also enforced by the doping-affected conductivity of sub-collector, which counteracts the downward heat propagation. Similar considerations hold for the other narrow HBTs, whereas for the device with $W_E = 0.55\,\mu m$ the lower overestimation ($+14\%$) can be again attributed to the geometrical averaging procedure for the evaluation of the numerical R_{TH}.

By exploiting approach D, the DUTs enjoy an exacerbated cooling effect dictated by the BEOL architecture and the adoption of $k_{Si,bulk}$, which favor both the downward and upward heat flow. As a consequence, the FEM R_{TH}s are far lower (about -40% and -34% for $L_E = 2.73$ and $9.93\,\mu m$, independently of W_E) than the experimental counterparts.

As expected, employing the traditional approach E leads to an underestimation of about -20% regardless of W_E, since the deactivation of the k reduction mechanisms (which would imply a heating effect) prevails over the BEOL absence (which would instead cool down the device).

Another analysis was performed to quantify the individual cooling influence of the metal layers embedded in the BEOL architecture; this was done by selectively replacing the layers with SiO_2 in the reference DUT. It was found that excluding the layers from 3rd to 5th leads to $R_{TH} = 6803$ K/W, corresponding to $+3.5\%$ with respect to the value obtained with all layers (6571 K/W). This means that the cooling action of
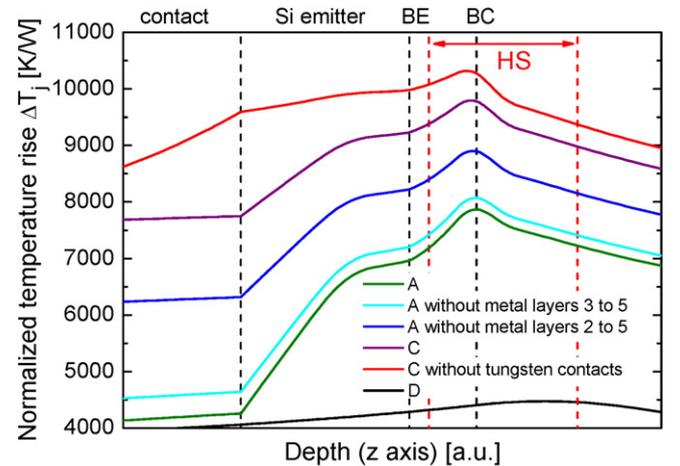

Fig. 9. Temperature rise above ambient normalized to the dissipated power versus device depth along a cut crossing the center of the reference transistor for various cases. Identified are the HS described according to the approach in Section 4.A (and corresponding to the region with power density >0 W/μm^3), as well as the base-emitter and base-collector metallurgical junctions.

the BEOL is mostly determined by the heatspreading favored by the 1st and the 2nd metal layers (by replacing the latter with SiO_2, $R_{TH} =$ 7761 K/W with $+18\%$ is evaluated). Fig. 9 illustrates the ΔT_j field normalized to the dissipated power (which corresponds to a local, i.e., position-dependent, thermal resistance) along the depth of the reference device for various cases. It can be observed that the peak of the normalized ΔT_j occurs in the close proximity of the base-collector metallurgical junction, which is an expected result. The base-emitter junction is instead at a perceptibly reduced temperature due to the interspersing low-conductivity SiGe layer. The higher temperature field for case C is due to the absence of the upward heat flow attracted by the BEOL, which is instead witnessed for case A by the sharper temperature gradient in base and emitter. Using approach D, the massive propagation of the heat emerging from the base-collector SCR to the backside and to the BEOL (both eased by the high $k_{Si,bulk}$) significantly, yet unrealistically, cools down the whole device.

6. Conclusions

In this paper, an advanced 3-D thermal simulation approach has been applied to state-of-the-art SiGe:C HBTs. In the endeavor of improving the accuracy with respect to traditional methods, (i) the whole BEOL architecture featuring five levels of metal and related interconnections has been included; (ii) a nonuniform power density pattern, preliminarily determined by electrical device simulations, has been considered to describe the heat generation region; (iii) Si thermal conductivity degradation mechanisms due to Ge mole fraction, high doping, and heat flow slowing in narrow layers have been taken into account. The resulting thermal resistances have been compared with experimental data extracted for state-of-the-art SiGe:C HBTs manufactured within the DOTFIVE project. The analysis has allowed shedding light on the thermal behavior of these geometrically-complex devices; it has been found that, although most of the heat flows toward the backside, disregarding the cooling influence of the BEOL structure leads to a large thermal resistance overestimation (in the span $+25$ to $+30\%$ for narrow devices), whereas an intolerable underestimation (-40%) is obtained by neglecting the mechanisms impacting the thermal conductivity of Si. Confirmation of the relevance of a more accurate heat source modeling has been also provided: using a power density pattern taken from device simulations in lieu of a conventional parallelepiped with uniform power density can increase the thermal resistance up to 9%.

References

[1] J.D. Cressler, G. Niu, Silicon-Germanium Heterojunction Bipolar Transistors, Artech House, Inc., 2003

[2] J.D. Cressler, A retrospective on the SiGe HBT: what we do know, what we don't know, and what we would like to know better, Proc. IEEE SiRF 2013, pp. 81–83, http://dx.doi.org/10.1109/SiRF.2013.6489439.

[3] M. Schröter, T. Rosenbaum, P. Chevalier, B. Heinemann, S.P. Voinigescu, E. Preisler, J. Böck, A. Mukherjee, SiGe HBT technology: future trends and TCAD-based roadmap, Proc. IEEE (2016)http://dx.doi.org/10.1109/JPROC.2015.2500024 (in press).

[4] P. Chevalier, F. Pourchon, T. Lacave, G. Avenier, Y. Campidelli, L. Depoyan, G. Troillard, M. Buczko, D. Gloria, D. Céli, C. Gaquière, A. Chantre, A conventional dou-ble-polysilicon FSA-SEG Si/SiGe:C HBT reaching 400 GHz f_{MAX}, Proc. IEEE BCTM 2009, pp. 1–4, http://dx.doi.org/10.1109/BIPOL.2009.5314250.

[5] P. Chevalier, T.F. Meister, B. Heinemann, S. Van Huylenbroeck, W. Liebl, A. Fox, A. Sibaja-Hernandez, A. Chantre, Towards THz SiGe HBTs, Proc. IEEE BCTM 2011, pp. 57–65, http://dx.doi.org/10.1109/BCTM.2011.6082749.

[6] J.-S. Rieh, D. Greenberg, Q. Liu, A.J. Joseph, G. Freeman, D.C. Ahlgren, Structure opti-mization of trench-isolated SiGe HBTs for simultaneous improvements in thermal and electrical performances, IEEE Trans. Electron Devices 52 (12) (Dec. 2005) 2744–2752, http://dx.doi.org/10.1109/TED.2005.859652.

[7] V. d'Alessandro, I. Marano, S. Russo, D. Céli, A. Chantre, P. Chevalier, F. Pourchon, N. Rinaldi, Impact of layout and technology parameters on the thermal resistance of SiGe:C HBTs, Proc. IEEE BCTM 2010, pp. 137–140, http://dx.doi.org/10.1109/BIPOL.2010.5667912.

[8] S. You, S. Decoutere, S. Van Huylenbroeck, A. Sibaja-Hernandez, R. Venegas, K. De Meyer, Impact of isolation scheme on thermal resistance and collector-substrate ca-pacitance of SiGe HBTs, Proc. IEEE ESSDERC 2011, pp. 243–246, http://dx.doi.org/10.1109/ESSDERC.2011.6044189.

[9] K.O. Petrosyants, R.A. Torgovnikov, Electro-thermal modeling of trench-isolated SiGe HBTs using TCAD, Proc. IEEE SEMI-THERM 2015, pp. 172–175, http://dx.doi.org/10.1109/SEMI-THERM.2015.7100156.

[10] A. El Rafei, A. Saleh, R. Sommet, J.M. Nébus, R. Quéré, Experimental characteri-zation and modeling of the thermal behavior of SiGe HBTs, IEEE Trans. Electron Devices 59 (7) (Jul. 2012) 1921–1927, http://dx.doi.org/10.1109/TED.2012.2196765.

[11] A.K. Sahoo, S. Frégonèse, M. Weiß, N. Malbert, T. Zimmer, A scalable electrothermal model for transient self-heating effects in trench-isolated SiGe HBTs, IEEE Trans. Electron Devices 59 (10) (Oct. 2012) 2619–2625, http://dx.doi.org/10.1109/TED.2012.2209651.

[12] I. Hasnaoui, A. Pottrain, D. Gloria, P. Chevalier, V. Avramovic, C. Gaquière, Self-heating characterization of SiGe:C HBTs by extracting thermal impedances, IEEE Electron Device Lett. 33 (12) (Dec. 2012) 1762–1764, http://dx.doi.org/10.1109/LED.2012.2220752.

[13] V. d'Alessandro, G. Sasso, N. Rinaldi, K. Aufinger, Influence of scaling and emitter layout on the thermal behavior of toward-THz SiGe:C HBTs, IEEE Trans. Electron Devices 61 (10) (Oct. 2014) 3386–3394, http://dx.doi.org/10.1109/TED.2014.2349792.

[14] L. La Spina, V. d'Alessandro, S. Russo, N. Rinaldi, L.K. Nanver, Influence of concurrent electrothermal and avalanche effects on the safe operating area of multifinger bipo-lar transistors, IEEE Trans. Electron Devices 56 (3) (Mar. 2009) 483–491, http://dx.doi.org/10.1109/TED.2008.2011574.

[15] A.K. Sahoo, S. Frégonèse, M. Weiß, C. Maneux, N. Malbert, T. Zimmer, Impact of back-end-of-line on thermal impedance in SiGe HBTs, Proc. IEEE SISPAD 2013, pp. 188–191, http://dx.doi.org/10.1109/SISPAD.2013.6650606.

[16] Synopsys TCAD Software User's Guide, Release 2010.12.

[17] M. Pfost, V. Kubrak, P. Brenner, A practical method to extract the thermal resistance for heterojunction bipolar transistors, Proc. IEEE ESSDERC 2003, pp. 335–338, http://doi.org/10.1109/ESSDERC.2003.1256882.

[18] A.D.D. Dwivedi, A. Chakravorty, R. D'Esposito, A.K. Sahoo, S. Frégonèse, T. Zimmer, Effects of BEOL and thermal coupling in SiGe multi-finger transistors under real operating conditions, Solid State Electron. 115 (2016) 1–6, http://dx.doi.org/10.1016/j.sse.2015.09.016.

[19] Comsol Multiphysics 3.5a User's Manual, COMSOL AB, Stockholm, Sweden, 2008.

[20] V. d'Alessandro, G. Sasso, N. Rinaldi, K. Aufinger, Experimental dc extraction of the base resistance of bipolar transistors: application to SiGe:C HBTs, IEEE Trans. Electron Devices (2016)http://dx.doi.org/10.1109/TED.2016.2565203 (in press).

[21] D.E. Dawson, A.K. Gupta, M.L. Salib, CW measurement of HBT thermal resistance, IEEE Trans. Electron Devices 39 (10) (Oct. 1992) 2235–2239, http://dx.doi.org/10.1109/16.158793.

[22] T. Vanhoucke, H.M.J. Boots, W.D. van Noort, Revised method for extraction of the thermal resistance applied to bulk and SOI SiGe HBTs, IEEE Electron Device Lett. 25 (3) (Mar. 2004) 150–152, http://dx.doi.org/10.1109/LED.2004.824242.

[23] V. Jain, B. Zetterlund, P. Cheng, R.A. Camillo-Castillo, J.J. Pekarik, J.W. Adkisson, Q. Liu, P.B. Gray, V. Kaushal, T. Kessler, D. Harame, Study of mutual and self-thermal resistance in 90 nm SiGe HBTs, Proc. IEEE BCTM 2013, pp. 17–20, http://dx.doi.org/10.1109/BCTM.2013.6798134.

[24] J.J. Sparkes, Voltage feedback and thermal resistance in junction transistors, Proc. IRE 46 (Jun. 1958) 1305–1306.

[25] Q.M. Zhang, H. Hu, J. Sitch, R.K. Surridge, J.M. Xu, A new large signal HBT model, IEEE Trans. Microwave Theory Tech. 44 (11) (Nov. 1996) 2001–2009, http://dx.doi.org/10.1109/22.543955.

[26] N. Nenadović, V. d'Alessandro, L.K. Nanver, F. Tamigi, N. Rinaldi, J.W. Slotboom, A back-wafer contacted silicon-on-glass integrated bipolar process–part II: a novel analysis of thermal breakdown, IEEE Trans. Electron Devices 51 (1) (Jan. 2004) 51–62, http://dx.doi.org/10.1109/TED.2003.820654.

[27] H. Kamrani, T. Kochubey, D. Jabs, C. Jungemann, Electrothermal simulation of SiGe HBTs and investigation of experimental extraction methods for junction tempera-ture, Proc. IEEE SISPAD 2015, pp. 108–111, http://dx.doi.org/10.1109/SISPAD.2015.7292270.

[28] S.-M. Hong, A.-T. Pham, C. Jungemann, Deterministic Solvers for the Boltzmann Transport Equation, Springer-Verlag Wien GmbH, 2011.

[29] Advanced Design System User's Guide, Keysight Technologies, 2011.

[30] M. Schröter, Advanced compact bipolar transistor models - HICUM(Chapter 8.4) in: D. Cressler (Ed.), Silicon Heterostructure Handbook, CRC Press, NY 2005, pp. 807–823.

[31] https://sites.google.com/site/ulfgri/numerical/gdsii-toolbox.

[32] V. d'Alessandro, N. Rinaldi, A critical review of thermal models for electro-thermal simulation, Solid State Electron. 46 (4) (2002) 487–496, http://dx.doi.org/10.1016/S0038-1101(01)00323-9.

[33] G. Sasso, N. Rinaldi, G. Matz, C. Jungemann, Accurate mobility and energy relaxation time models for SiGe HBTs numerical simulation, Proc. IEEE SISPAD 2009, pp. 241–244, http://dx.doi.org/10.1109/SISPAD.2009.5290206.

[34] G. Sasso, N. Rinaldi, G. Matz, C. Jungemann, Analytical models of effective DOS, saturation velocity and high-field mobility for SiGe HBTs numerical simulation, *Proc. IEEE SISPAD 2010*, pp. 279–282, http://dx.doi.org/10.1109/SISPAD.2010.5604505.

[35] G. Niu, J.D. Cressler, A.J. Joseph, Quantifying neutral base recombination and the ef-fects of collector-base junction traps in UHV/CVD SiGe HBTs, IEEE Trans. Electron Devices 45 (12) (Dec. 1998) 2499–2504, http://dx.doi.org/10.1109/16.735727.

[36] E. Pop, Energy dissipation and transport in nanoscale devices, Nano Res. 3 (3) (2010) 147–169, http://dx.doi.org/10.1007/s12274-010-1019-z.

[37] V. Palankovski, R. Quay, Analysis and Simulation of Heterostructure Devices, Spring-er-Verlag Wien GmbH, 2004.

[38] G.A. Slack, Thermal conductivity of pure and impure silicon, silicon carbide, and di-amond, J. Appl. Phys. 35 (12) (1964) 3460–3466, http://dx.doi.org/10.1063/1.1713251.

[39] A.D. McConnell, K.E. Goodson, Thermal conduction in silicon micro- and nanostruc-tures, Annual Review of Heat Transfer, 14 2005, pp. 129–168, http://dx.doi.org/10.1615/AnnualRevHeatTransfer.v14.120.

[40] Y. Lee, G.S. Hwang, Mechanism of thermal conductivity suppression in doped silicon studied with nonequilibrium molecular dynamics, Phys. Rev. B 86 (7) (2012) 075202-1–075202-6, http://dx.doi.org/10.1103/PhysRevB.86.075202.

[41] G. Chen, Nanoscale Energy Transport and Conversion: A Parallel Treatment of Elec-trons, Molecules, Phonons, and Photons, Oxford University Press, 2005.

[42] W. Liu, M. Asheghi, Thermal conduction in ultrathin pure and doped single-crystal silicon layers at high temperatures, J. Appl. Phys. 98 (12) (2005) 123523, http://doi.org/10.1063/1.2149497.

[43] O. Tornblad, P. Sverdrup, D. Yergeau, Z. Yu, K.E. Goodson, R.W. Dutton, Modeling and simulation of phonon boundary scattering in PDE-based device simulators, Proc. IEEE SISPAD 2000, pp. 58–61, http://dx.doi.org/10.1109/SISPAD.2000.871206.

[44] L. Codecasa, D. D'Amore, P. Maffezzoni, Compact modeling of electrical devices for electrothermal analysis, IEEE Trans. Circuits Syst. I 50 (4) (Apr. 2003) 465–476, http://dx.doi.org/10.1109/TCSI.2003.811422.