A Thermal-Fully Hydrodynamic Model for Semiconductor Devices and Applications to III-V HBT Simulation *

Augusto Benvenuti W. M. Coughran, Jr. Mark R. Pinto

Abstract

Because of the interaction between self-heating and hot carriers effects, neither isothermal nor conventional macrothermal models are adequate for the simulation of state-of-the-art power devices; instead, a detailed electro-thermal model accounting for non-stationary transport, such as the *Thermal-Fully Hydrodynamic* model, is required. We apply a one-dimensional (1D) implementation of such model to the simulation of AlGaAs/GaAs and InP/InGaAs Heterojunction Bipolar Transistors (HBTs), comparing the results with those provided by simplified models, and highlighting how deeply both non-local transport and self-heating affect the predicted device performance. The importance of the convective terms is assessed, and a new non-thermal mechanism for the output Negative Differential Resistance (NDR) is proposed.

I Introduction

To reduce the impact of self-heating on microwave device performance and circuit reliability, several technological and design precautions, such as the adoption of interdigitated structures, possibly with variable ballasting resistances on each finger [1], selective substrate thinning [2], use of via holes and *bathtub* heat sinks [3, 4], or even growth on Si substrates [5], can be undertaken.

A better understanding of the physical mechanisms involved in thermal-related phenomena can both suggest the use of new device concepts or different materials and can drive the designer towards more effective layer structures or layout geometries: for these purposes, a reliable and efficient device simulation program, taking into account both self-heating and the effect of the external ambient temperature, can be a very valuable tool.

Different approaches aiming at this goal have been proposed in the past, depending on the temporal and spatial domain scale and on the level of complexity/accuracy of the corresponding model (see [6] for a review).

*©Numerical Analysis Manuscript, No. 96/0-1, Bell Labs, Lucent Technologies, Murray Hill, NJ 07974. August 1996. Available via WWW at http://cm.bell-labs.com/cs/doc/96.

In particular, due to the intrinsic imbalance between the characteristic dimensions of the electrical and thermal problems, decoupled approaches [7, 8, 9, 10, 11] are well-suited for the thermal analysis of large, three-dimensional regions, while coupled models provide a deeper insight in the physical processes taking place on a device-scale.

In this work, we will deal with coupled electro-thermal models only: they can be further classified, according to the capability to separately describe the contributions of the charge-carrier gases (electrons and holes) to the total thermal energy storage and heat conduction, as *macroscopic* or *detailed* electro-thermal models.

Most thermal simulations are instead based on a macroscopic approach, whereby the thermal contributions arising from electrons, holes and phonons are considered as a whole; these models [12, 13] can be considered as extensions of the drift-diffusion formalism to non-isothermal conditions.

The detailed thermal approach has been proposed in 1985 by Wang [14], and further developed by Roberts and Chamberlain [15]; it was first applied in its simpler form (no convective terms) by Szeto and Reif [16], and more recently by Pierantoni and coworkers [17] and by Chen and coworkers [18], and in its full form by Benvenuti *et al.* [19].

In the present work we will highlight the hierarchical relationship existing between the advocated Thermal-Fully Hydrodynamic model [19] and simpler models, and investigate the main limitations of the macro-thermal approach: the neglect of non-local transport and of the role of the carrier energy fluxes in the field–lattice interaction.

The paper is structured as follows. First, the six transport models considered are introduced, and the adopted nomenclature is explained (Sect. II); in Sect. III some relevant numerical issues are discussed; finally, the results from the different models are critically compared (Sect. IV) and some conclusions are drawn.

II A consistent reformulation of coupled electro-thermal models

It has long been recognized that the drift-diffusion model is unable to reproduce non-stationary transport effects [20]; several improved approaches have been devised in order to circumvent this limitation, including — besides statistical methods [21] — enhanced drift-diffusion [22], energy balance [23], and fully hydrodynamic [24] models. In much the same way, conventional macro-thermal models [13] cannot correctly describe the carrier contribution to heat transport nor the spatial distribution of the lattice heat sources.

As we will show, both drawbacks are due to the same approximation, i.e., the neglect of the carrier energy fluxes, which implies local equilibrium between the field and the carrier temperature.

A more accurate description of such phenomena can be obtained with a detailed coupled electro-thermal model, such as the *Thermal-Fully Hydrodynamic* one. In order to draw a meaningful comparison between this reference model and some simplified approaches, we re-examine, in the paragraphs to follow, the simplifying assumptions leading from the former to the latter.

A The Thermal-Fully Hydrodynamic model

Since our goal is to treat in a unified manner electrical transport and thermal effects, we have to describe the charge carriers and the phonons populations by means of equations of the same character: one suitable choice could be in principle the use of a set of coupled Boltzmann equations

$$\frac{\partial f_i}{\partial t} + \nabla_{\underline{p}} f_i \cdot \frac{\mathrm{d}\underline{p}}{\mathrm{d}t} + \nabla_{\underline{r}} f_i \cdot \frac{\mathrm{d}\underline{r}}{\mathrm{d}t} = \left.\frac{\mathrm{d}f_i}{\mathrm{d}t}\right|_{\mathrm{coll}}, \tag{1}$$

each describing the time evolution in phase space $\underline{r}$–$\underline{k}$ of the distribution function f_i pertaining to a different population, labeled by index i.

However, the numerical solution of such a system of partial differential equation in a six-dimensional space is a prohibitively computation-intensive task; thus suitable approximations are sought for. As proposed by Roberts and Chamberlain [15], a viable approach is to apply the method of moments in $\underline{k}$ space [25].

If we assume the phonons to possess a vanishing momentum, and if we are not interested in their number, but just in the associated energy density, which we can relate to the lattice temperature T_l through the per unit volume lattice specific heat ρc_l, the dynamical behavior of the lattice is described by the lattice heat equation alone:

$$\frac{\partial(\rho c_l T_l)}{\partial t} + \nabla\cdot\underline{Q}_l = \left.\frac{\mathrm{d}(\rho c_l T_l)}{\mathrm{d}t}\right|_{\mathrm{coll}}, \tag{2}$$

where $\underline{Q}_l = -\kappa_l \nabla T_l$ is the lattice heat-flux.

On the other hand, in the single electron gas approximation [26], we get for the moments of order 0–2 of the electron distribution function the balance equations

$$\frac{\partial n}{\partial t} + \nabla\cdot(n\underline{v}_n) = \left.\frac{\mathrm{d}n}{\mathrm{d}t}\right|_{\mathrm{coll}} \tag{3}$$

$$\begin{aligned} \frac{\partial(n m_n^* \underline{v}_n)}{\partial t} &+ \nabla[n(k_B T_n + m_n^* v_n^2)] = -qn\underline{\mathcal{E}}_n \\ &+ n w_n \nabla \log m_n^* + \left.\frac{\mathrm{d}(n m_n^* \underline{v}_n)}{\mathrm{d}t}\right|_{\mathrm{coll}} \end{aligned} \tag{4}$$

$$\begin{aligned} \frac{\partial(n w_n)}{\partial t} &+ \nabla\cdot[n\underline{v}_n(w_n + k_B T_n) + \underline{Q}_n] \\ &= -qn\underline{v}_n\cdot\underline{\mathcal{E}}_n + \left.\frac{\mathrm{d}(n w_n)}{\mathrm{d}t}\right|_{\mathrm{coll}}, \end{aligned} \tag{5}$$

expressing conservation of electron charge, momentum and energy densities respectively [19]. The quasi-field $\underline{\mathcal{E}}_n = \nabla E_c/q$ depends on the gradient of the conduction band edge to account for the additional driving force due to spatial variations of the electron affinity (which may give an important contribution to transport in compositionally-graded regions), and reduces to the electric field $\underline{\mathcal{E}} = -\nabla\psi$ in homogeneous materials.

A similar set of conservation equations is obtained for the holes, with $\underline{\mathcal{E}}_p = \nabla E_v/q$.

The coupling between the three sub-systems (lattice, electrons and holes) is expressed through the collision terms, which have been indicated on the right hand sides of equations (3)–(5) in symbolic form, and describe particle, momentum and energy exchange due to scattering events. As such, they can be expressed by the superposition of a term accounting for intra-band scattering events, described by the Relaxation Time Approximation (RTA),

$$
\begin{aligned}
\left.\frac{\mathrm{d}n}{\mathrm{d}t}\right|_{\mathrm{RTA}} &= 0 && (6)\\
\left.\frac{\mathrm{d}(nm_n^*\underline{v}_n)}{\mathrm{d}t}\right|_{\mathrm{RTA}} &= -\frac{nm_n^*\underline{v}_n}{\tau_{mn}} && (7)\\
\left.\frac{\mathrm{d}(nw_n)}{\mathrm{d}t}\right|_{\mathrm{RTA}} &= -n\frac{w_n - w_0}{\tau_{en}} && (8)\\
\left.\frac{\mathrm{d}(\rho c_l T_l)}{\mathrm{d}t}\right|_{\mathrm{RTA}} &= n\frac{w_n - w_0}{\tau_{en}} + p\frac{w_p - w_0}{\tau_{ep}}, && (9)
\end{aligned}
$$

and a term related to inter-band scattering, which can be approximated taking into account evaluating the charge, momentum and energy exchanges related to each generation-recombination (GR) rate [15]. We took into account surface and bulk trap-related recombination (described with the conventional SRH model), direct band-to-band and Auger recombination; as to carrier generation, we tested [28] average velocity [14] and average energy [15] dependent impact ionization coefficients, as well as the conventional field dependence. Then for the lattice the inter-band contribution reads:

$$
\begin{aligned}
\left.\frac{\mathrm{d}(\rho c_l T_l)}{\mathrm{d}t}\right|_{\mathrm{GR}} = \frac{3k_B}{2}\Big[&(2T_n + T_p - T_l + T_g)\, R^n_{\mathrm{Aug}}\\
&+ (T_n + 2T_p - T_l + T_g)\, R^p_{\mathrm{Aug}}\\
&+ (\hat{T}_n - 3T_l + T_g)\, G^n_{\mathrm{ii}} + (\hat{T}_p - 3T_l + T_g)\, G^p_{\mathrm{ii}}\\
&+ \left(T_n + \frac{T_g}{2}\right) R^n_{\mathrm{SRH,dir}} + \left(T_p + \frac{T_g}{2}\right) R^p_{\mathrm{SRH,dir}}\\
&+ \left(T_l + \frac{T_g}{2}\right) (G^n_{\mathrm{SRH,dir}} + G^p_{\mathrm{SRH,dir}})\Big], \qquad (10)
\end{aligned}
$$

where T_g is a temperature corresponding to the energy band-gap, $T_g = 2E_g/(3k_B)$, and $\hat{T}_n$, $\hat{T}_p$ have been heuristically related to the impact ionization threshold energies, $\hat{T}_{n/p} = \max(T_{n/p}, 2w_{iiT}^{n/p}/(3k_B))$.

For the relaxation times, we assumed the simplest analytical model reproducing the experimental saturated velocity v_{sn} at high fields under homogeneous conditions [24], slightly modified to avoid an increase in mobility in the presence of a retarding field:

$$\tau_{mn} = \frac{m_n^*}{q}\mu_n = \frac{m_n^*}{q}\mu_{n0}\left[(1-f) + f\frac{T_l}{T_n}\right], \qquad (11)$$

$$\tau_{en} = \mu_{n0}T_l\left[\frac{3}{2}\frac{T_n}{T_n+T_l}\frac{1}{v_{\mathrm{sn}}^2} + \frac{1}{2}\frac{m_n^*}{T_n}\right], \qquad (12)$$

where $f = 1/2[1+\tanh((T_n - T_l)/T_s)]$ is a smoothing function, the low-field mobility μ_{n0} depends in the usual way from the lattice temperature and doping concentration N

$$\mu_{n0} = \left(\frac{T_l}{T_0}\right)^{-2}\left[\mu_{n1} + \frac{\mu_{n2}-\mu_{n1}}{1+(N/N_{rn})^{\theta_n}}\right], \qquad (13)$$

and the parameters μ_{n0}, v_{sn}, μ_{n1}, μ_{n2}, N_{rn} and θ_n are assumed to depend linearly on the material composition.

We restricted ourselves to the steady-state, 1D case, and closed the system by expressing the electron gas heat flux with the Fourier law:

$$\underline{Q}_n = -\kappa_n \nabla T_n, \qquad (14)$$

where the electron thermal conductivity κ_n was assumed to obey the Wiedemann-Franz relationship:

$$\kappa_n = \zeta\left(\frac{k_B}{q}\right)^2 T_n\,\sigma_n = \zeta\,\frac{k_B^2}{q}\,T_n\,\mu_n\,n. \qquad (15)$$

According to Stettler and Lundstrom [27], ζ, which controls the diffusive part of the total electron energy flux $\underline{S}_n = n\underline{v}_n(w_n + k_BT_n) + \underline{Q}_n$, is related to the coefficient of the drift part of $\underline{S}_n$.

Putting together the charge, momentum and energy conservation equations for the electrons, a similar set for the holes, the Poisson equation and the lattice heat equation, the so-called *Thermal-Fully Hydrodynamic* (T-FH) model is obtained [19].

B The Thermal-Energy Balance model

Under steady-state conditions, neglecting the so-called *convective contributions* ($\propto v_n^2, v_p^2$) to the diffusivity/mobility ratios

$$\beta_n = \frac{qD_n}{\mu_n} = k_BT_n + m_n^*v_n^2 \approx k_BT_n, \qquad (16)$$

$$\beta_p = \frac{qD_p}{\mu_p} = k_BT_p + m_p^*v_p^2 \approx k_BT_p, \qquad (17)$$

and to the average kinetic energies

$$w_n = \frac{3k_BT_n}{2} + \frac{1}{2}m_n^* v_n^2 \approx \frac{3k_BT_n}{2}, \tag{18}$$

$$w_p = \frac{3k_BT_p}{2} + \frac{1}{2}m_p^* v_p^2 \approx \frac{3k_BT_p}{2}, \tag{19}$$

the momentum conservation equations can be explicitly solved for v_n, v_p, and the moments of order 0–1 can be joined into second-order current continuity equations.

We will refer to transport models in which two partial differential equations (PDEs) are sufficient to describe each carrier population as *Energy Balance* formulations. Since convective terms [29] appear in the drift part of the electron energy flux $n\underline{v}_n(w_n + k_BT_n)$ and in the electron energy relaxation term $-n(w_n - w_0)/\tau_{en}$ as well, several slightly different versions of energy balance are conceivable: in what follows we will label *Thermal-Energy Balance* (T-EB) model [16] the version taking into account the convective terms in the relaxation term only.

C The Thermal-Drift Diffusion model

To further simplify the transport model, the differential equation expressing conservation of the electron kinetic energy must be eliminated as well. Usually, this is implicitly accomplished resorting to the drift-diffusion formalism of carrier transport, i.e. assuming local field-dependent mobilities (T-DD model). In fact, if the carrier energy fluxes are neglected, the energy conservation equations for electrons and holes degenerate to algebraic relationships between the corresponding quasi-fields and temperatures:

$$\frac{3}{2}\frac{k_B}{q}\frac{T_{\mathrm{eff,n}} - T_l}{\tau_{en}} = -\underline{v}_n \cdot \underline{\mathcal{E}}_n = -\mu_n(\mathcal{E}_n)\,\mathcal{E}_n^2, \tag{20}$$

$$\frac{3}{2}\frac{k_B}{q}\frac{T_{\mathrm{eff,p}} - T_l}{\tau_{ep}} = -\underline{v}_p \cdot \underline{\mathcal{E}}_p = -\mu_p(\mathcal{E}_p)\,\mathcal{E}_p^2. \tag{21}$$

In general, the energy relaxation times for electrons and holes, τ_{en} and τ_{ep}, are non-linear functions of the corresponding effective temperatures, $T_{\mathrm{eff,n}}$, $T_{\mathrm{eff,p}}$. The non-linear equations (20)–(21) can then be solved to yield $T_{\mathrm{eff,n}}$ and $T_{\mathrm{eff,p}}$ as a function of the quasi-fields.

Neglecting the carrier energy fluxes is equivalent to neglecting the non-locality of the energy–field relationship: this is a crude approximation under inhomogenous conditions, especially in regions where the quasi-fields increase abruptly to a high value, such as near the base-collector junction of HBTs. Since under these circumstances the carrier temperature lag the quasi-field, the effective temperatures given by Eqn. (20)–(21) strongly over-estimate the "real" T_n, T_p: as a consequence, the T-DD model yield saturated velocities, while the T-EB and T-FH models can result in a velocity overshoot region. For these reasons we will use the terminology "non-stationary transport" (or "non-local transport") referring to both the T-EB and the T-FH models; in this context such expression does not imply a time-dependent analysis, but rather refers to (the implications of) the *non-local* nature of the energy–field relationship.

Another approximation is necessary to eliminate the dependencies on the carrier temperatures in the T-DD model: the built-in voltages entering in the boundary conditions and the diffusivity/mobility ratios β_n, β_p have to be assumed as $T_{\rm eff,n}$, $T_{\rm eff,p}$ or as T_l dependent. In HBTs, the former choice would result in a huge thermal current of electrons diffusing from the collector back into the base, leading to an unphysically high concentration of minority carriers in the base, that is to a far too pessimistic device performance estimate; on the other hand, assuming "conventional" T_l-dependent Einstein relationships implies an underestimate of the thermal diffusion currents, but still gives a qualitatively correct simulated device behavior.

In conclusion, we resorted to the effective temperatures, in the electrical part of the T-DD model, only to evaluate the mobilities: everywhere else T_l was used instead. Note that this is consistent with the commonly accepted definition of the DD transport model, where only the carrier mobilities are assumed as field-dependent, while the diffusivity/mobility ratios are taken to be constant.

As to the thermal part of the T-DD model, if optical generation-recombination mechanisms are neglected, by applying a global energy balance principle for the collision terms [17, 30]

$$\left.\frac{\mathrm{d}(\rho c_l T_l)}{\mathrm{d}t}\right|_{\mathrm{GR}} = E_g R - \left[\left.\frac{\mathrm{d}(n w_n)}{\mathrm{d}t}\right|_{\mathrm{GR}} + \left.\frac{\mathrm{d}(p w_p)}{\mathrm{d}t}\right|_{\mathrm{GR}}\right], \tag{22}$$

a lumped heat equation is obtained, which describes the time evolution of the total energy density stored in the semiconductor:

$$\begin{aligned} \frac{\partial(\rho c_l T_l + n w_n + p w_p)}{\partial t} &- \nabla\cdot(\kappa_l \nabla T_l + \kappa_n \nabla T_n + \kappa_p \nabla T_p) \\ &= \frac{5 k_B}{2q} \nabla\cdot\left[\underline{J}_n T_n + \underline{J}_p T_p\right] \\ &+ \underline{J}_n \cdot \underline{\mathcal{E}}_n + \underline{J}_p \cdot \underline{\mathcal{E}}_p + E_g R. \end{aligned} \tag{23}$$

The widely-used *macro-thermal* model [12, 13] is recovered assuming $T_n \simeq T_p \simeq T_l$, and neglecting the term $5k_B/(2q)\nabla\cdot(T_l \underline{J})$:

$$\frac{\partial(\rho c T_l)}{\partial t} - \nabla\cdot(\kappa \nabla T_l) = \underline{J}_n \cdot \underline{\mathcal{E}}_n + \underline{J}_p \cdot \underline{\mathcal{E}}_p + E_g R, \tag{24}$$

where $c = c_l + 3k_B(n+p)/(2\rho)$ and $\kappa = \kappa_l + \kappa_n + \kappa_p$ are the lumped specific heat and thermal conductivity, respectively.

D Isothermal models

By allowing $\kappa_l \to \infty$ (i.e., modeling the lattice as a distributed, perfect heat sink at the ambient temperature T_a), the isothermal versions of the Fully Hydrodynamic (FH), Energy Balance (EB) and Drift-Diffusion (DD) model are recovered from the corresponding thermal-coupled counterparts (T-FH, T-EB, and T-DD, respectively).

III Implementation

All of the mentioned thermal models (T-FH, T-EB, T-DD) and their isothermal counterparts (FH, EB, DD) have been implemented in the framework of a single, highly flexible device simulation code [19], exploiting generalized continuation, grid adaption and automatic jacobian calculation, and allowing for an arbitrary number of differential/algebraic equations in one spatial dimension and time. Several discretization schemes (including Scharfetter-Gummel [31], pure 1-sided and optimal upwinding [29]) may be applied, while all equations are solved self-consistently applying the full-Newton method.

A Boundary conditions

The T-FH formulation is completed by imposing a set of coupled, mixed boundary conditions for the electron, hole and lattice temperatures T_n, T_p and T_l, accounting for the three-dimensional heat spreading through the metallization and the substrate by means of a geometrical transformation.

As it is not clear what are the correct interface conditions for *e.g.* the carrier temperatures and heat fluxes at an abrupt heterojunction, this case has been simply treated as a limiting case of a heterojunction graded over a very short distance; no attempt has been made to account for thermoionic emission and tunneling. Altough this very simplistic assumption may lead to an overestimate of the emitter injection efficiency (and thus of the DC current gain) for abrupt HBTs, we are confident that even for those devices the conclusions we draw are qualitatively correct, as they are confirmed by the results obtained for graded HBTs where these limitations do not apply.

B Discretization

The spatial discretization of advection-diffusion partial differential equations like Eqn. (4) is difficult, and canonical finite elements methods results in spurious oscillations in regions of high fields unless very fine meshes are used [31]. For the drift-diffusion model, a stable yet accurate discretization technique known as Scharfetter-Gummel scheme after their proponents has been almost universally adopted ever since its formulation in 1969, with extensions to deal with energy balance model.

However the Scharfetter-Gummel approach cannot be applied to the hydrodynamic model unless the convective term are treated as a perturbation, i.e. their dependence on the state variables is frozen at each step of the Newton method, which will degrade convergence in cases where their variation in space or time is important.

To avoid these difficulties, we used an alternative approach known in computational fluid dynamics as *upwinding* [31], consisting of a Petrov-Galerkin method with asymmetric weight functions: this is equivalent to an explicit addition of numerical diffusivity to the original equations [32], followed by the application of a standard Galerkin scheme.

In fact, recalling that we confine ourselves to steady-state, and denoting with β_n the generalized diffusivity/mobility ratio (see par. C), Eqn. (4) can be written in scaled form as:

$$\nabla(\beta_n\, n) \;= n\,[-\underline{\mathcal{E}}_n \,+\, w_n \nabla \log m_n^*] \,+\, \frac{\underline{J}_n}{\mu_n}. \tag{25}$$

We artificially enhanced the diffusivity discretizing with the Galerkin scheme the modified PDE

$$\nabla \left[\beta_n\, n \left(1 + u\frac{\mathcal{R}}{2}\right)\right] = -n\,\underline{F}_n \,+\, \frac{\underline{J}_n}{\mu_n}, \tag{26}$$

where $\underline{F}_n = \underline{\mathcal{E}}_n - w_n \nabla \log m_n^*$ is a generalized driving force, and u is the *local upwinding coefficient*

$$u = \mathrm{sign}(\mathcal{R})\,(1-\alpha) + \alpha\,[\coth(\mathcal{R}/2) - 2/\mathcal{R}]\,. \tag{27}$$

It is useful to introduce the generalized potential $\tilde{\psi} = \psi + \chi + w_n \log m_n^*$, where χ is the electron affinity, such that, approximating the contribution due to the gradient of the effective mass, $\underline{F}_n \approx -\nabla\tilde{\psi}$. This approximation is not used to replace $\underline{F}_n$ in the r.h.s. of Eqn. (26), but only to evaluate the cell Reynolds number $\mathcal{R}$; for a generic mesh element between nodes 1 and 2,

$$\mathcal{R} = \frac{\Delta\tilde{\psi} - \Delta\beta_n}{\Delta\beta_n} \ln\left(\frac{\beta_{n2}}{\beta_{n1}}\right), \tag{28}$$

where Δ denotes a difference between node 2 and node 1, and the constant $0 \le \alpha \le 1$ can be tuned to control the amount of artificial diffusivity in low field regions.

For $\alpha = 0$, a pure 1-sided scheme is obtained; $\alpha = 1$ (*optimal* scheme) corresponds to the analogous for first order differential equations of the Scharfetter-Gummel discretization, providing zero nodal errors [31] for a uniform grid and constant coefficients.

As a test we simulated a Si *nin* diode at 77 K: the optimal scheme was found [19] to effectively suppresses the wiggles with a better accuracy than the pure 1-sided scheme, since it locally minimizes the amount of numerical diffusivity required to reach stability. Our experience suggests easier Newton convergence when a nearly-optimal scheme with $\alpha = 0.7$–0.9 is used: with such a choice, for both the canonical *nin* simulations and the much more numerically challenging HBT simulations, the reduction in the current $\underline{J}_n$ induced by the discretization error is much lower than with the pure 1-sided scheme used in [33].

C Software approach

Traditionally, transistor designers have relied upon specialized device simulation programs. Such programs are well suited to production runs but it is often difficult to add or modify new model equations and terms. This is especially true for equations of a fundamentally different character like those in the FH model. As an experiment, we applied a "user friendly" interface to a very general one-dimensional PDE solver, which supports mixed systems of PDEs and ordinary differential equations (ODEs) as well as more general operators [33]. Here we outline the approach and discuss its advantages and disadvantages.

We use a modification of Schryer's POST program that includes finite-difference Jacobians, simple parameter continuation, appropriate upwinding, and adaptivity [34]. The formulation [35] is based upon a mixed set of PDEs/ODEs of the form:

$$\begin{aligned} &\partial_x a(t, x, u, \partial_x u, \partial_t u, \partial_{xt} u, v, \partial_t v) = \\ &\qquad\qquad f(t, x, u, \partial_x u, \partial_t u, \partial_{xt} u, v, \partial_t v) && (29) \\ &d(t, \partial_t u, v, \partial_t v) = 0, && (30) \end{aligned}$$

subject to boundary conditions

$$b_L(t, u(t, L), \partial_x u\ (t, L), \partial_{xt} u\ (t, L), v, \partial_t v) = 0, \tag{31}$$

$$b_R(t, u(t, R), \partial_x u\ (t, R), \partial_{xt} u\ (t, R), v, \partial_t v) = 0. \tag{32}$$

Here $u(t, x)$ is the vector-valued time and spatially dependent PDE solution and $v(t)$ is the corresponding ODE solution, while ∂_x and ∂_t denote differentiation with respect to space and time, respectively. The a and f functions define the coupled system of PDEs while the d equation represents ODE conditions. The nonlinear boundary conditions at the left and right ends are represented by b_L and b_R; for "hyperbolic" transport equations only involving first-order derivatives, the code associates a single boundary condition with one end of the interval without a corresponding boundary condation at the other end.

The i-th component of the PDE solution is represented in terms of B-splines of order k as

$$u_i(t, x) = \sum_{j=1}^{N} U_{ij}(t) B_j^{(k)}(x); \tag{33}$$

then the Galerkin method, possibly with some form of upwinding, is applied to discretize the system (29)–(30), and the resulting non-linear algebraic system is solved with the Newton method. For $k = 2$ this approach is essentially equivalent to second-order finite differences.

This general PDE/ODE solver is powerful but can be difficult to use in its "native" mode (*i.e.* calling it via a FORTRAN or C driving program). A "one liner" interface based on the macro capabilities of the Bourne shell, AWK, and the MAPLE symbolic algebra package was developed [34, 36]. By using a symbolic algebra system, it is possible to derive explicit FORTRAN or C code for complex algebraic expressions; this obviates the need for the user to derive analytic derivative expressions and avoids the numerical accuracy problems associated with finite-difference derivative approximations. Here is a simple example; to solve a simple heat equation, $T_t = T_{xx}$, with T specified at the left and right, on the interval $[0, 1]$ for time from 0 to 1, and with an initial state of 0, the user has to supply to the interface the following commands:

```
vars "T"
af "d Tx = Tt"
bc "T=1 | T=0"
x on 0 1
t on 0 1
u "0"
ndx 25 dt 1e-10.
```

This formulation and the underlying code has been used successfully in several applications, so we have experimented its use in the field of device simulation implementing the transport models described in Sect. II.

For our applications, we wanted to simulate bipolar transistors in 1D and needed to treat the base contact. The base current density was specified with an ODE variable, v in the notation above; an algebraic condition

$$\log(|p(x_b)|) = \nu(x_b) - \psi(x_b) \tag{34}$$

was applied via the ODE mechanism, where x_b is the base contact location, p is the hole density, ν accounts for the heterojunction, and ψ is the potential. The hole continuity PDE was modified to smoothly apply the associated base current across the width of the base region.

IV Results and discussion

We applied our simulator to a variety of graded and abrupt HBTs based on different material systems, including Si/SiGe, AlGaAs/GaAs, and InP/InGaAs.

All the results presented here have been obtained with steady-state simulations, taking into account surface and bulk SRH, direct and Auger recombination. Impact ionization was found to have a minor effect for all devices under the investigated bias conditions ($V_{ce} < 2.5V$), and has therefore been neglected.

We have previously shown [37] that, even for a power interdigitated HBT where multi-dimensional geometrical effects might play an important role, the T-FH model reproduces fairly well the measured device behaviour provided that parasitic effects (extrinsic base resistance and base-collector capacitance, emitter and collector series resistances) are properly taken into account.

However, the simulator described in the previous section is not intended as a predictive tool for routinely device CAD, but rather as a flexible environment to easily perform comparisons between several PDE-based device models and to explore the implications of their approximations on the interpretation of some specific features.

	type	doping cm^{-3}	width Å	x
emitter contact	n^+	5×10^{18}	2000	0
graded emitter contact	n^+	4×10^{17}	200	$0 \rightarrow 0.3$
emitter	n	4×10^{17}	800	0.3
graded emitter	n	4×10^{17}	200	$0.3 \rightarrow 0$
base	p^+	4×10^{19}	1000	0
collector	n^-	3×10^{16}	5000	0
collector contact	n^+	5×10^{18}	550	0

Table 1: Structure of the AlGaAs/GaAs HBT. The composition is given in terms of the Aluminum fraction x, where $Al_xGa_{(1-x)}As$. The grading is linear in x.

Here we present three examples of the simulator application in that direction:

A. analysis of the impact of non-local transport and of a detailed thermal description;

B. assessment of the importance of the convective terms;

C. investigation of output Negative Differential Resistance (NDR).

The first and the third topics are more closely related to thermal phenomena: thus they have been investigated on a power AlGaAs/GaAs HBT with a base-emitter junction linearly graded over 20 nm (see Table 1 for the layer sequence description).

As substantial self-heating tends to partially mask the effects of the convective terms, we also include results obtained with isothermal models on two different high-speed InP/InGaAs SHBTs. Both devices have a built-in field of about 27.5 kV/cm in the base; HBT a has an abrupt base-emitter heterojunction (see Table 2), while HBT b has a base-emitter junction linearly graded over 25 nm.

Since hot holes effects are expected to play a minor role in the electro-thermal behavior of such npn devices, we adopted the drift-diffusion model to describe hole transport.

In all the examples, the model formulations have substantial effects on potential, charge and/or thermal distributions within the device. These internal differences can manifest themselves to varying degrees on externally observable measures of device performance, and we concentrate primarily on comparisons on terminal characteristics, as these are ultimately the most relevant for device design.

A Deficiencies of simpler models

As far as the purely electrical (*i.e.* isothermal) behaviour is concerned, the main limitation of the Drift-Diffusion model is, as well known, its inability

	type	doping cm^{-3}	HBT a width Å	HBT a β	HBT b width Å	HBT b β
em. contact	n^+	5×10^{19}	500	1	500	1
emitter	n	1×10^{18}	2000	1	1750	1
graded em.	n	1×10^{18}	$\cdots$	$\cdots$	250	$1 \to 0.25$
graded base	p^+	7×10^{19}	500	0	500	$0.25 \to 0$
collector	n^-	2×10^{16}	5500	0	5500	0
coll. buffer	n	2×10^{18}	1000	0	1000	0
coll. contact	n^+	5×10^{19}	500	0	500	0

Table 2: Structure of the InP/InGaAs HBTs. The composition is given in terms of the Phosphorous fraction β, where $Ga_{0.47(1-\beta)}In_{0.53+0.47\beta}As_{1-\beta}P_{\beta}$. The grading is linear in β.

to account for non-stationary transport, due to the implicit assumption of local and instantaneous equilibrium between the carrier temperatures and the electric field.

In particular for III-V HBTs velocity overshoot in the base and in the collector space charge regions considerably reduce the corresponding transit times, causing a higher cutoff frequency f_T in the Fully Hydrodynamic model than in the Drift-Diffusion model (compare the curves labelled FH and DD respectively in Fig. 1). The total transit time has been calculated as

$$\tau = q \, \frac{\delta \displaystyle\int n(x)\, dx}{\delta J_c}; \tag{35}$$

the cutoff frequency is then given by $f_T = \dfrac{1}{2\pi\tau}$.

For this AlGaAs/GaAs HBT, the impact of self-heating on the peak cutoff-frequency is even more dramatic (see the curves labelled T-FH and T-DD in Fig. 1); similar results have been found for the InP/InGaAs HBTs, except the role of non-stationary transport is enhanced, while that of self-heating was lower.

Conventional macro-thermal models do not correctly locate the lattice heat sources, for the same reason that they cannot describe non-stationary transport, *i.e.*, because they neglect the energy flux related to the electron and hole populations (Fig. 2). Due to such fluxes, the lattice heat source profile is much smoother than suggested by the Joule term $\underline{J}_n \cdot \underline{\mathcal{E}}_n + \underline{J}_p \cdot \underline{\mathcal{E}}_p$, with a maximum substantially displaced from the base-collector (B-C) to the collector-subcollector (C-SC) junction.

However it must be remarked that, even if the heat sources distributions are so remarkably different, under equal total power dissipation conditions the two model provide quite similar lattice temperature profiles; this is due to the spatial filtering action of the lattice heat equation, that is to the long lattice energy relaxation length. Because of this consideration, the main limitation of

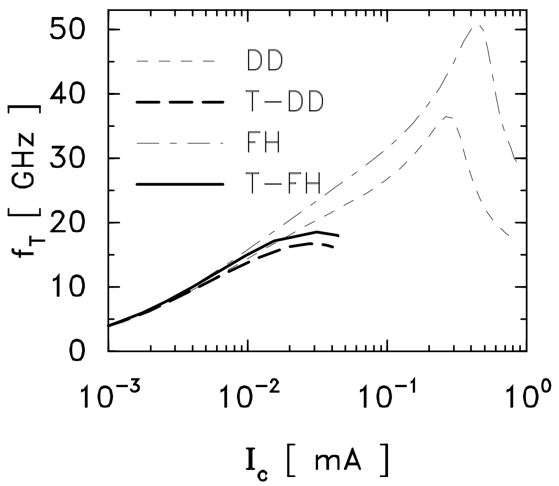

Figure 1: Cutoff frequency for the different models in the AlGaAs/GaAs HBT.

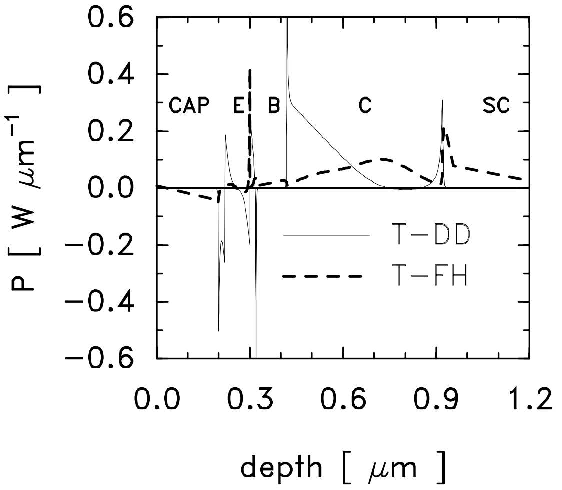

Figure 2: Lattice thermal sources profiles for the T-FH and T-DD models in the InP/InGaAs HBT.

the macro-thermal models has to be considered their inappropriate description of carrier transport, rather than of self-heating.

In contrast with the results of Liou and co-workers [13], no $T_l < T_a$, where T_a is the ambient temperature, is observed in the emitter, even though a substantial lattice cooling takes place.

To summarize, neither macro-thermal models (which neglect velocity overshoot) nor isothermal energy-balance or hydrodynamic models (which neglect self-heating) are suitable for an accurate simulation of state-of-the-art power HBTs.

B Effect of convective terms

To simplify the comparison between the different models, we have not shown in the last paragraph the results for the Energy Balance model, which is similar to the Fully Hydrodynamic model, except for the fact that it neglects the convective terms.

It is important to quantitatively assess the difference in the estimated device performance between the FH model and the EB model, since the latter is considerably easier to implement and faster to use within a general-purpose device simulation code.

Introducing the convective parameter γ, such that

$$\beta_n = k_B T_n + \gamma\, m_n^* v_n^2, \tag{36}$$

$$w_n = \frac{3 k_B T_n}{2} + \gamma\, \frac{1}{2} m_n^* v_n^2, \tag{37}$$

it is possible to recover the EB model from the FH model simply reducing γ from 1 to 0. Due to the flexibility of the underlying software, the whole procedure is handled automatically, allowing a parametrization of the mixed PDE/ODE system where all the terms are treated consistently, and exactly the same physical models and discretization schemes are used for any value of γ. In this way a meaningful comparison of the model results can be performed, and it is also possible to analyze "intermediate" situations ($0 < \gamma < 1$).

It has to be remarked that for $\gamma > 0$ the numerical solution of the mixed PDE/ODE system is a very challenging task, on which ordinary finite element discretization methods will fail or produce spurious oscillations. Only the application of the nearly-optimal upwinding discretization scheme with a careful choice of α allowed us to reach convergence to a physical solution under such conditions.

Changing γ at a constant base-emitter voltage, we found a very strong difference between the internal electron concentration and average velocity profiles for the two models, particularly in the base and in the collector space charge regions [29].

However, this effect was found to be mainly due to a change of the effective injection conditions induced by γ.

This phenomenon can be observed in Fig. 3a, which shows the Gummel plot for the graded device (HBT b): it is clear that, for a given base emitter voltage V_{be}, when the convective terms are considered (filled symbols), both the base and the collector current are reduced with respect to the EB model results (hollow symbols). Since HBT devices are basically current-driven devices, such a reduction in the current for a given V_{be} can be more properly interpreted as an increase of V_{be} for a given current density, which in turn can be quantitatively

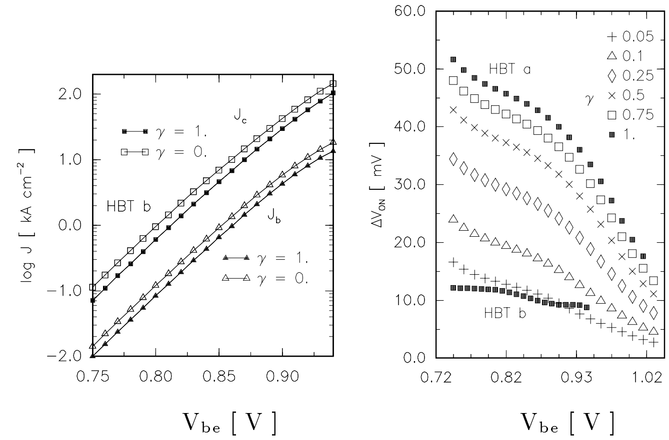

Figure 3: Effect of the convective parameter γ on the Gummel plot for the graded InP/InGaAs HBT (a), and corresponding shift of the turn-on voltage (b).

expressed as an effective shift of the turn-on voltage $\Delta V_{\rm ON}$. In Fig. 3b, we show the amount of $\Delta V_{\rm ON}$ as a function of the injection level, with γ as a parameter. As expected from visual inspection of Fig. 3a, the shift is nearly constant (about 10 mV) for the graded device, while decreses at higher injection currents for the abrupt device. Also, the dependence on the convective parameter γ is strongly non-linear.

We found that if γ is changed at a fixed collector current density, rather than at a fixed base-emitter voltage, the differences in the internal density and velocity profiles are much lower [29].

Since for a given V_{be} the reduction in J_b with γ is lower than for J_c, the predicted DC current gain $\beta_{\rm DC}$ is smaller for the FH model than for the EB model. In particular, for the abrupt InP/InGaAs HBT (Fig. 4a), the gain is reduced by nearly 25 % at low-medium injection, but no influence is found in high injection conditions, where the Kirk effect dominates. It has to be reminded that our model does not take into account the thermionic and tunneling contributions to the current at the abrupt base-emitter heterojunction, which would presumably have a strong impact on the absolute value of the current gain for the two models for this particular device; however, we expect that this approximation would not alter the qualitative trend shown.

A similar behaviour was found also for the graded HBT (Fig. 4b): the EB model predicts a DC current gain $\beta_{\rm DC}$ only 10 % too high at low injection, but a sensible reduction persists also in high injection conditions (nearly 4 % at $J_{\rm c}$ = 52 kA cm^{-2}, which corresponds to the peak cutoff frequency).

On the other hand, the convective parameter γ has only a modest influence on the cutoff frequency for the graded device (Fig. 5b), whose peak $f_{\rm T}$ is overestimated by nearly 7 % from the EB model, and almost no influence at all on

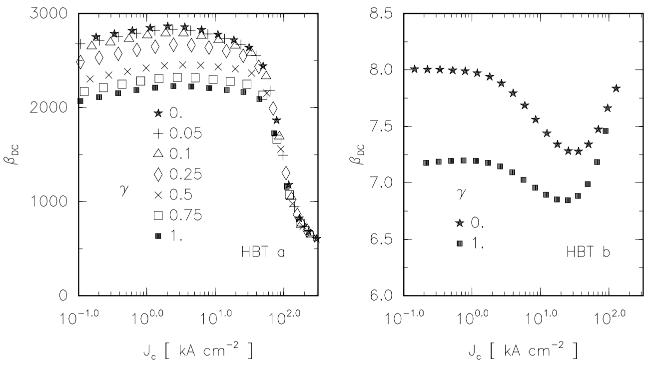

Figure 4: Comparison between the predicted DC current gain β_{DC} of the abrupt (a) and graded (b) InP/InGaAs HBT with the EB model ($\gamma = 0$) and with the FH model ($\gamma = 1$).

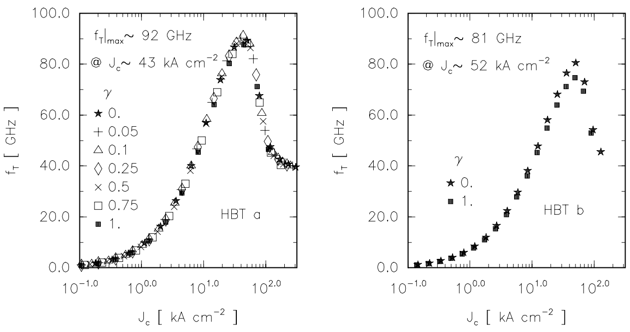

Figure 5: Comparison between the predicted cutoff frequency f_T of the abrupt (a) and graded (b) InP/InGaAs HBT with the EB model ($\gamma = 0$) and with the FH model ($\gamma = 1$).

the abrupt device (Fig. 5a).

In conclusion, the EB and FH model give similar internal velocity profiles when compared under the same current injection conditions; the simpler EB model underestimate the turn-on voltage, and provides too optimistic predictions for the DC current gain and slightly too high cutoff frequencies.

C Output NDR analysis

The performance of high-power HBTs can be severely limited by the onset of a DC Negative Differential Resistance (NDR) in the output characteristic.

Although this effect is commonly ascribed to device heating, our simulations (see Figures 6–9) point out that NDR might not be entirely due to thermal effects.

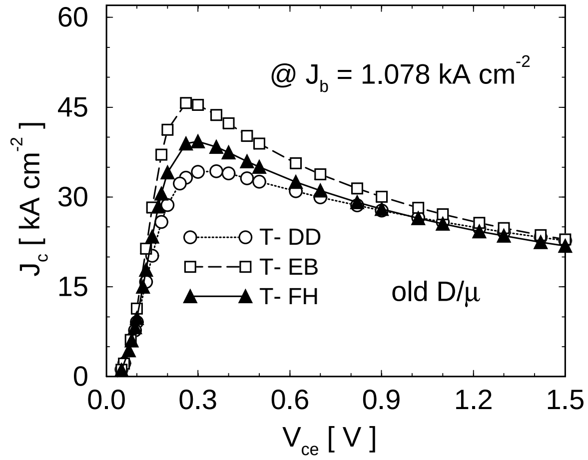

Figure 6: Output I-V curves predicted by the thermal models for the AlGaAs/GaAs HBT. DD=Drift- Diffusion, EB=Energy Balance, FH=Fully Hydrodynamic.

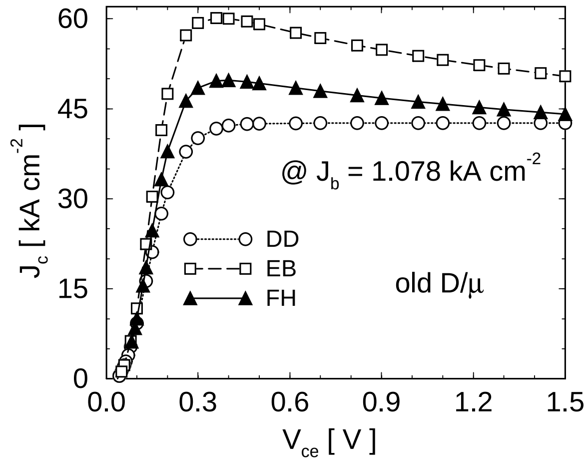

Figure 7: Output I-V curves predicted by the isothermal models for the AlGaAs/GaAs HBT.

By comparing the common-emitter I-V curves for different non-isothermal transport models (Fig. 6), we noticed that the non-stationary formulations appear to enhance the output NDR. Furthermore, the corresponding isothermal models (Fig. 7) result in a persistent NDR, thus suggesting that self-heating alone cannot explain such an effect. A detailed analysis of the different contributions to the current inside the collector layer revealed a significant component due to diffusion towards the base (driven by the strong electron temperature gradient), whose magnitude increases with the base-collector bias, while the drift current directed toward the collector contact is nearly constant. The collector current simulated by the isothermal DD model is flat due the assumption $T_n \equiv T_l \equiv 300$ K.

Recently [38], an electrical mechanism for output NDR based on the V_{ce} dependence of the conduction band barrier height at the base-emitter junction has been proposed. However we found that, for a constant base drive current, the change in $V_{\rm ce}$ mainly appears at the base-collector junction, thus leaving the base-emitter barrier height nearly unchanged.

We investigated the origin of the electrical NDR by performing some numerical experiments under isothermal conditions. We noticed that, if in the EB model the carrier temperature T_n is replaced by the lattice temperature T_l in the high-field diffusivity/mobility ratio, Eqn. (16), NDR completely disappears; conversely, if the effective temperature $T_{\rm eff,n}$ is used instead than T_l in the DD model, a very strong NDR is observed.

In fact, in the latter case, the amount of NDR was even larger than for the EB model, since $T_{\rm eff,n}$, as provided by the static field-temperature relationship (20), strongly overestimates T_n near the base-collector junction, and thus the amount of the thermal back-diffusion from the collector into the base.

Therefore, NDR appears to be closely related to the high-field diffusivity/mobility ratio.

In order to better describe the high-field diffusivity D_n, the diffusivity/mobility ratio has been modified according to the empirical relationship

$$\beta_n = \frac{qD_n}{\mu_n} = k_B \left[T_n(1-\phi) + T_l\phi \right] + \gamma\, m_n^* v_n^2, \tag{38}$$

where $\phi = \exp(-\Delta_{\Gamma L}/(k_B T_n))$ is the fraction of electrons in the upper (L) valley. As a result, the previously mentioned back-diffusion current (and thus the electrical NDR) is significantly reduced (see Fig. 8). Finally we evaluated the impact of the modified Einstein relationship, Eqn. (38), in the non-isothermal case; we found that for this device the reduction of the electrical NDR is largely masked by the strong thermal NDR (Fig. 9), so that the peak total (electrical + thermal) negative conductivity is reduced only from about -24.6 to about -22.6 mS (for a 100 μm^2 emitter area device) according to the T-FH model (Table 3).

To summarize, for the investigated AlGaAs/GaAs HBT the output NDR is partly due to non-thermal causes, but the mechanism proposed in [38] plays a minor role; instead, the D_n/μ_n ratio in the collector layer has been found to significantly affect the amount of electrical NDR. According to this new interpretation, non-stationary electron dynamics is ultimately responsible for the I_c decrease at high $V_{\rm ce}$; we expect that, for devices with substantial velocity oveshoot in the low-doped collector region, NDR cannot be entirely suppressed by a careful thermal design.

	old D/μ ratio			new D/μ ratio		
	DD	EB	FH	DD	EB	FH
electrical	—	-11.8	-6.6	—	-7.4	-3.4
thermal	-14.1	-23.1	-18.0	-14.1	-25.3	-19.2
total	-14.1	-34.9	-24.6	-14.1	-32.7	-22.6

Table 3: Peak output negative conductance, in mS, for the AlGaAs HBT, according to the various transport models.

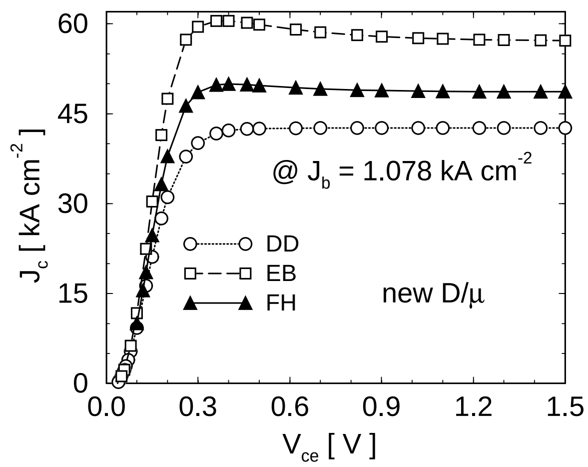

Figure 8: Output I-V curves predicted by the isothermal models with the modified Einstein relationship (38) for the AlGaAs/GaAs HBT.

V Conclusions

A 1D, non-isothermal hydrodynamic transport model has been applied to the simulation of a AlGaAs/GaAs and InP/InGaAs HBTs.

Comparisons with simpler models confirm that accounting for the interplay between non-local and thermal effects is essential in understanding the device behavior and predicting its performance; however, even for high-speed devices, the inclusion of the convective terms results in a relatively minor change of the predicted current gain and cutoff frequency.

We have shown that NDR might be partially due to non-thermal mechanisms, and critically depends on the diffusivity/mobility ratio in the collector under non-local transport conditions.

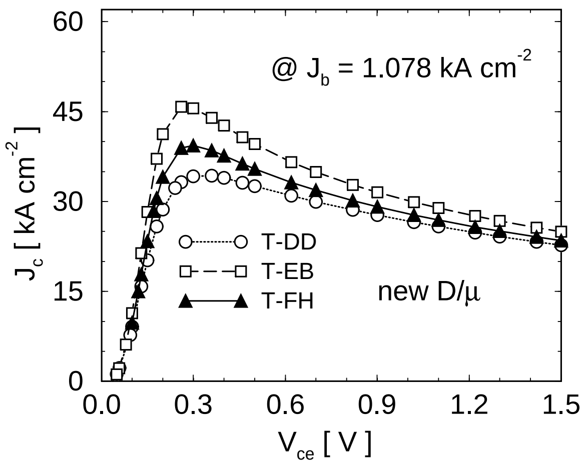

Figure 9: Output I-V curves predicted by the thermal models with the modified Einstein relationship (38) for the AlGaAs/GaAs HBT.

ACKNOWLEDGMENT

We gratefully acknowledge the fundamental contribution of N. L. Schryer of AT&T in providing the basic PDE solver code and developing the appropriate discretization schemes for device simulation. Thanks are also due to G. Ghione and C. U. Naldi of Politecnico di Torino, for their comments and suggestions on the thermal analysis, and to A. Grinberg of Lucent Technologies and S. Luryi of the State University of New York for stimulating discussions.

References

[1] W.Liu, B.Bayraktaroglu, "Theoretical Calculations of Temperature and Current Profiles in Multi-Finger Heterojunction Bipolar Transistors," *Solid State Electronics*, Vol.36, No.2, pp.125-132, 1993.

[2] A.Marty, T.Camps, J.Tasselli, D.L.Pulfrey, J.P. Bailbe, "A Self-Consistent DC-AC Two-Dimensional Electrothermal Model for AlGaAs/GaAs Microwave Power HBT's," *IEEE Trans. on Electron Devices*, Vol.ED-40, No.7, pp.1202-1210, 1993.

[3] L.W.Yang, J.J.Komiak, M.Y.Kao, D.E.Houston, D.P. Smith, K.J.Norheden, "E-Beam Re-Aligned HBTs and A New Broadband MMIC Power Amplifier Using Bathtub as Heat Sink," *IEDM Tech. Dig.-94*, pp.203-206, 1994.

[4] T.Ishikawa, K.Okaniwa, M.Komaru, K. Kosaki, Y. Mitsui, "A High-Power GaAs FET Having Buried Plated Heat Sink for High-Performance MMIC's," *IEEE Trans. on Electron Devices*, Vol.ED-41, No.1, pp.3-8, 1994.

[5] G.B.Gao, M.Z.Wang, X.Gui, H.Morkoç, "Thermal design studies of high power heterojunction bipolar transistors", *IEEE Trans. on Electron Devices*, Vol.ED-36, No.5, pp.854-862, 1989.

[6] A.Benvenuti, F.Bonani, G.Ghione, C.Naldi, "Thermal modelling of III-V devices," *Proc. 4th Int. Work. GaAs Tel.*, pp. 139-156. 1993.

[7] H.Fukui, "Thermal resistance of GaAs field-effect transistors," *IEDM Tech. Dig.-80*, pp.118-121, 1980.

[8] H.F.Cooke, "Precise technique finds FET thermal resistance," *Microwaves & RF*, pp.85-87, 1986.

[9] G.Ghione, C.Naldi, "High-resolution self-consistent thermal modelling of multi-gate power GaAs MESFET's," *IEDM Tech. Dig.-89*, pp.147-150, 1989.

[10] G.N.Ellison, "Methodologies for thermal analysis of electronic components and systems", in *Advances in thermal modelling of electronic components and systems*, A.Bar-Cohen and A.D.Kraus ed., IEEE Press, New York, 1993.

[11] F.Bonani, G.Ghione, M.Pirola, C.U.Naldi, "A new, efficient approach to the large-scale thermal modeling of III-V devices and integrated circuits," *IEDM Tech. Dig.-93*, pp.101-104, 1993.

[12] A.Chryssafis, W.Love, "A computer-aided analysis of one-dimensional thermal transients in n-p-n power transistors," *Solid-State Electronics*, Vol.22, No.3, pp.249-256, 1979.

[13] L.L.Liou, J.L.Ebel, C.I.Huang, "Thermal Effects on the Characteristics of AlGaAs/GaAs Heterojunction Bipolar Transistors Using Two-Dimensional Numerical Simulation," *IEEE Trans. Electron Devices*, Vol.ED-40, No.1, pp.35-43, 1993.

[14] C.T.Wang, "A new set of semiconductor equations for computer simulation of submicron devices," *Solid-State Electronics*, Vol.28, pp.783-788, 1985.

[15] J.W.Roberts, S.G.Chamberlain, "Energy-momentum transport model suitable for small geometry silicon device simulation," *COMPEL*, Vol.9, No.1, pp.1-22, 1990.

[16] S.Szeto, R.Reif, "A unified electrothermal hot-carrier transport model for silicon bipolar transistor simulation," *Solid-State Electronics*, Vol.32, No.4, pp.307-315, 1989.

[17] A.Pierantoni, A.Liuzzo, P.Ciampolini, G.Baccarani, "Three-Dimensional Implementation of a Unified Transport Model," *Proc. SISDEP-V*, pp.125-128, 1993.

[18] D.Chen, Z.Yu, K.-C.Wu, R.Goossens, R.W.Dutton, "Dual Energy Transport Model with Coupled Lattice and Carrier Temperatures," *Proc. SISDEP-V*, pp.157-160, 1993.

[19] A.Benvenuti, G.Ghione, M.R.Pinto, W.M.Coughran, Jr., N.L.Schryer, "Coupled thermal-fully hydrodynamic simulation of InP-based HBTs," *IEDM Tech. Dig.-92*, 1992.

[20] R.Stratton, "Diffusion of Hot and Cold Electrons in Semiconductor Barriers," *Phys. Rev.*, Vol.126, No.6, pp.2002-2014, 1962.

[21] A.Abramo *et al.*, "A Comparison of Numerical Solutions of the Boltzmann Transport Equation for High-Energy Electron Transport in Silicon," *IEEE Trans. Electron Dev.*, Vol.ED-41, No.9, pp.1646-1654, 1994.

[22] D.Chen, E.C.Kan, U.Ravaioli, "An Analytical Formulation of the Length Coefficient for the Augmented Drift-Diffusion Model Including Velocity Overshoot," *IEEE Trans. on Electron Devices*, Vol.ED-38, No.6, pp.1484-1490, 1991.

[23] D.Chen, E.C.Kan, U.Ravaioli, C.-W.Shu, R.W. Dutton, "An Improved Energy Transport Model Including Nonparabolicity and Non-Maxwellian Distribution Effects," *IEEE Electron Dev. Lett.*, Vol.13, pp.26-28, 1992.

[24] B. Baccarani and M. R. Wordeman, "An Investigation of Steady-State Velocity Overshoot in Silicon," *Solid-State Electronics*, Vol.28, No.4, pp.407-416, 1985.

[25] A.Bringer, G.Schön, "Extended moment equations for electron transport in semiconducting submicron structures," *J. Appl. Phys.*, Vol.64, No.5, pp.2447-2455, 1988.

[26] K.Bløtekjær, "Transport Equations for Electrons in Two-Valley Semiconductors," *IEEE Trans. on Electron Devices*, Vol.ED-17, No.1, pp.38-47, 1970.

[27] M.A.Stettler, M.A.Alam, M.S.Lundstrom, "A Critical Examination of the Assumptions Underlying Macroscopic Transport Equations for Silicon Devices," *IEEE Trans. Electron Devices*, Vol.ED-40, No.4, pp.733-740, 1993.

[28] A.Benvenuti, G.Ghione, C.U.Naldi, "Interaction between hot carrier transport and thermal effects in power HBTs," *Proc. VII GaAs Simulation Group Meeting*, 1993.

[29] A.Benvenuti, W.M.Coughran, Jr., M.R.Pinto, N.L. Schryer, "Hierarchical PDE simulation of nonequilibrium transport effects in semiconductor devices," *Proc. NUPAD-IV*, pp.155-160, 1992.

[30] C.C.Mc Andrew, E.L.Heasell, K.Singhal, "Modeling thermal effects in energy models of current transport in semiconductors," *Sem. Sci. Technol.*, Vol.3, pp.758-765, 1988.

[31] J.P.Kreskovsky, "A hybrid central difference scheme for solid-state simulation," *IEEE Trans. Electron Dev.*, Vol.ED-34, No.5, pp.1128-1133, 1987.

[32] T.Tang, M.Ieong, "Discretization of flux densities in device simulations using optimum artificial diffusivity," *IEEE Trans. CAD*, Vol.14, No.11, pp.1309-1315, 1995.

[33] A.Benvenuti, M.R.Pinto, W.M.Coughran, Jr., N.L. Schryer, C.U.Naldi, G.Ghione, "Evaluation of the influence of convective energy in HBTs using a fully hydrodynamic model," *IEDM Tech. Dig.-91*, pp.499-502, 1991.

[34] W.M.Coughran, Jr., N.L.Schryer, "Faster device modeling using adaptive spatial meshes and continuation," *Proc. NUPAD-III*, pp.85-86, 1989.

[35] N.L.Schryer, "Designing software for one dimensional partial differential equations," *ACM Trans. Math. Software*, Vol.16, pp.72-85, 1990.

[36] J.L.Bentley, Maria F.Fernandez, B.W.Kernighan, N.L.Schryer, "Template-driven Interfaces for numerical subroutines," *ACM TOMS*,Vol.19, No.3, pp.265–287, 1993.

[37] A.Benvenuti, F.Bonani, G.Ghione, C.U.Naldi, M. Kärner, K.Schaper, "Analysis of output NDR in power AlGaAs/GaAs HBTs by means of a Thermal-Fully Hydrodynamic model," *Proc. ISDRS-93*, pp.499-502, 1993.

[38] T.W.Lee, P.A.Houston, "New nonthermal mechanism for negative differential resistance in heterojunction bipolar transistors", *Appl. Phys. Lett.*, Vol.62, pp.1777-1779, 1993.