On the feasibility of 500 GHz Silicon-Germanium HBTs

A. Pawlak, M. Schröter, J. Krause,
G. Wedel
CEDIC
Dresden University of Technology
Dresden, Germany
pawlak@iee.et.tu-dresden.de

M. Schröter
ECE Dept.
University of California
San Diego, USA

C. Jungemann
University of Armed Force Munich
Neubiberg, Germany

Abstract--**A procedure for rapid TCAD based evaluation of device design alternatives is presented. It employs 1D device simulation in combination with a physics-based compact model and a corresponding model generator. This enables to provide libraries with geometry scalable models for mm-wave circuit optimization. Based on an experimentally calibrated baseline the procedure is applied to demonstrate the feasibility of SiGe HBTs with $(f_T, f_{max}) = (420, 520)$ GHz at realistic vertical and lateral dimensions. The proposed method is suitable for improving the existing ITRS 2007 SiGe HBT roadmap towards emerging mm-wave applications.**

Keywords-Device simulation, SiGe heterojunction bipolar transistors, HICUM, Device scaling

I. INTRODUCTION

SiGe BiCMOS technology is becoming increasingly attractive for the emerging mm-wave market and applications at the lower limit of the so-called THz gap [1, 2]. For the latter, corresponding *circuit* operation speed targets are presently in the range of 120 to 160 GHz. This requires maximum transistor operation speed in the order of 500 GHz. The development of an associated process technology is being addressed by the European joint research project DOTFIVE [3]. Within this project also demonstrator circuits for the above mentioned speed range will be designed. To enable an early exploration of the design space and to provide feedback for steering process development compact models need to be available that represent the final (yet unknown) process technology as accurately as possible. Such models not only have to include the physical effects and parasitics of the entire three-dimensional (3D) device structure (cf. Fig. 1) but also need to be geometry scalable for allowing circuit optimization. Scalability includes variations in both dimensions (i.e. layout) and contact configuration (i.e. location and number of base, emitter and collector fingers).

Meeting these design kit requirements demands a significant effort on both compact modeling *and* device simulation. The latter is being employed for evaluating possible, mostly one-dimensional (1D), device design options and for predicting the corresponding electrical characteristics while the former allows to include 3D parasitics and geometry scalability in an accurate but computationally efficient way. The goal of the device design here is to create a doping profile that on one hand meets the electrical specifications for circuit design and on the other hand is realistic in terms of fabrication. The circuit design specifications include certain figures of merit (FoMs) that will be discussed below and that have a significant impact

on the device design choices. Therefore, this paper describes a methodology for predictive physics-based geometry scalable model libraries enabling early circuit design (a.k.a. concurrent engineering) and target doping profiles for process development.
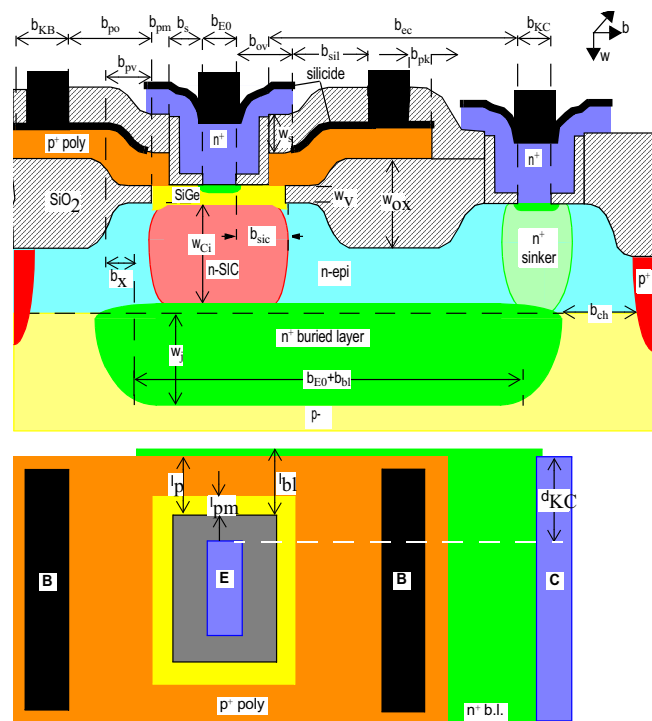

Figure 1. Schematic cross-section and layout of an advanced SiGe HBT with relevant dimensions for generating geometry scalable compact models.

II. METHODOLOGY

The device optimization is based on the evaluation of FoMs that not only properly represent the target circuit applications but also can be calculated with minimum computational effort. Very often just the transit frequency f_T is chosen as FoM (e.g. [4]) and attempted to be maximized since it can be quickly obtained from device simulation with reasonable accuracy[1]. This is more difficult for, e.g., the maximum (unilateral) oscil-

---

[1] Note that device simulation often still misses parasitic capacitances and 3D effects.

lation frequency f_{max} and nearly impossible for *circuit* FoMs such as the CML ring-oscillator delay time τ_{CML}. However, both f_{max} and τ_{CML} are generally more suitable than f_T for estimating the performance of actual high-speed circuits. For these reasons, f_{max} and τ_{CML} have also been chosen as main FoMs for DOTFIVE and this paper.

The predictive modeling procedure pursued here is shown in Fig. 2. It starts with a measured doping profile and design rules from an existing process that are as advanced as possible. In this work, the process described in [6] served as baseline. Its measured electrical characteristics are compared to device simulation results. In case of significant deviations the parameters of the physical models (e.g. mobility, bandgap) are properly adjusted to suitable measured characteristics and relevant FoMs.

The device design optimization is subdivided into two major loops, recognizing that the intrinsic operation of bipolar transistors is mainly determined by 1D profile and current flow, while the overall impact of 2D and 3D effects can be quite well estimated by e.g. using lumped elements. Therefore, in the first loop 1D doping profile options are evaluated w.r.t. given specifications. These are derived from those of 3D structures and from existing information on the impact of 2D and 3D effects on electrical device behavior. Analytical doping profiles are used due to, among others, their flexibility for making adjustments and the high uncertainty of process simulation for new advanced processes. Once the specified FoM target is reached a 2D profile is derived from the 1D profile. The electrical characteristics of both 1D and 2D simulation are then used for compact model parameter extraction. However, in contrast to usual approaches, so-called process-specific model parameters are extracted, such as sheet resistances, capacitances per unit area and length etc. Combined with design rules these allow the construction of process-based scalable compact models.
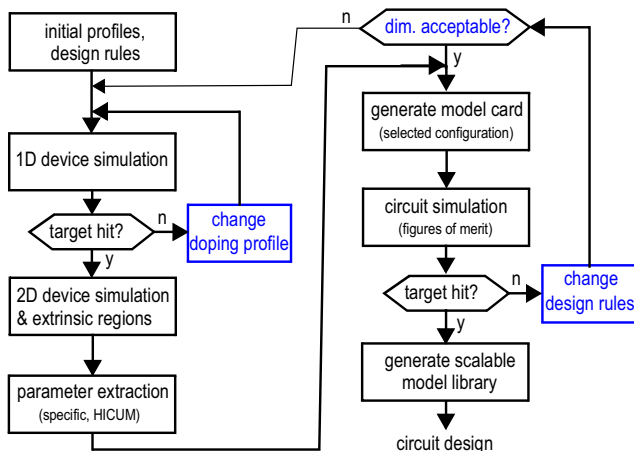

Figure 2. Flow-chart of the device design and modeling methodology employed.

The second loop starts with generating compact model parameters for selected complete 3D transistor structures that are required for the subsequent evaluation of device and circuit characteristics. Initially, the design rules of the baseline process are employed. Based on the obtained FoMs the lateral dimensions are gradually reduced by a certain factor. For a

given set of FoMs and target values there is generally a large number of possible solutions. This is especially true for bipolar transistors, in which (i) the vertical dimensions are only loosely related to the lateral dimensions, and (ii) the variety of contact configuration is much larger than for MOS transistors. A too aggressive scaling of the 1D profile usually results in high f_T and too low f_{max}, which then can only be increased by reducing the lateral dimensions beyond realistic values for the targeted process generation (i.e. given fabrication equipment capability). In this case, a different 1D profile needs to be created or selected that yields a more balanced set of FoMs. If the reduced dimensions are acceptable, a viable solution has been found and the compact model library can be generated for subsequent circuit design.

The procedure described above partitions the often used single loop, in which 2D or even 3D device simulation is used, into two loops that are computationally by orders of magnitude more efficient. However, this can only be exploited with a well-defined (preferably automated) model parameter extraction infrastructure, a suitable compact model, and a geometry scaling tool. In this work, the standard compact model HICUM/L2 [7] and the model generator TRADICA [8] were used, which both have been employed in industry for many years. TRADICA also contains a certain set of parameter extraction building blocks, eventually allowing automation[1]. Most importantly, though, the program has built-in compact models and calculates a variety of FoMs such as f_T, f_{max} and τ_{CML} directly from transistor parameters based on accurate *analytical* equations [9]. These expressions, which have been repeatedly verified by circuit simulations, allow to evaluate with TRADICA FoMs as function of bias, frequency, dimensions, specific electrical data or, within reasonable ranges, even doping concentrations [10]. Thus, circuit simulations[2] can be replaced by a much faster and more versatile evaluation.

III. DEVICE SIMULATION

The most accurate transport model for advanced device structures, such as SiGe HBTs with f_T beyond 300GHz, is the Boltzmann transport equation (BTE). However, existing solutions using the Monte-Carlo method are computationally far too expensive, even for 1D HBT structures. Nevertheless, the BTE solution can be used as reference for selected 1D doping profiles and also for calibrating physical models of the drift-diffusion (DD) and hydrodynamic (HD) simulation. Thus, DD and HD simulations were used as workhorse for profile optimization. Their calibration was performed for special bulk structures and for a given doping profile, but may lose its validity once the scaled profile becomes too different from the initial one. All DD and HD simulations were performed with in-house programs [11, 12] the results of which were also compared to other simulators such as MINIMOS [13], ATLAS [14], SDEVICE [15], and Galene [16]. The BTE was solved with MONJU [17] using the Monte-Carlo (MC) method.

It has been well-known that DD simulation does not accurately capture (1D) transport effects in advanced SiGe HBTs anymore. The resulting f_T values as well as the open-base

---

[1] For HICUM parameter extraction in industry either proprietary or commercial tool kits are available that are based on interactive user interfaces.
[2] Note that circuit simulators do not contain accurate geometry scalable BJT and HBT models.

breakdown voltage BV_{CEO} tend to be too low compared to the BTE. Adding a carrier energy balance and transport equation is believed to significantly improve the description of transport effects. However, the resulting HD simulation tends to yield too large f_T values compared to the BTE. An example is shown in Fig. 3 for a profile in [4] that is supposed to yield a SiGe HBT with f_T beyond 1 THz. Note that the HD results of [4] were obtained with default values while our HD and DD results are based on calibrations with the BTE solution. Since the base current is not well represented by device simulation due to process dependent effects at the emitter contact, the breakdown voltage was calculated from a 1% increase of I_C due to avalanche. Comparison to measured data of BV_{CEO} (= 1.53 V) gave an energy relaxation length of 40 nm. As a pragmatic approach to obtain as realistic as possible f_T values the average $(f_{T,DD} + f_{T,HD})/2$ was used in this work since it was always closer to the BTE results than either DD or HD simulation.
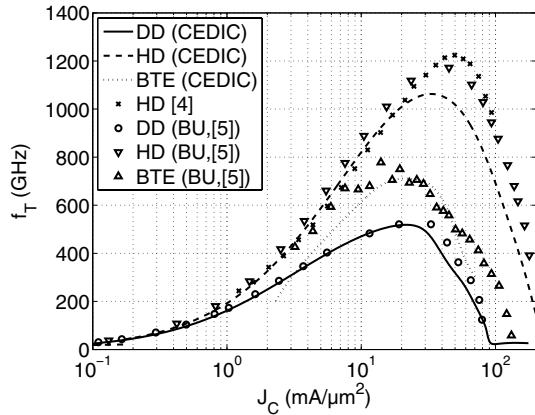

Figure 3. Comparison of DD, HD, and BTE simulation for the "THz" SiGe HBT proposed in [4].

Fig. 4 shows a comparison between measured and simulated results for f_T and $I_C(V_{BE})$ of the baseline process. The simulation corresponds to a complete HICUM with parameters obtained from combining 1D and 2D device simulation results as well as the design rules of the existing process. The overall agreement is fairly good and establishes the baseline simulation models.
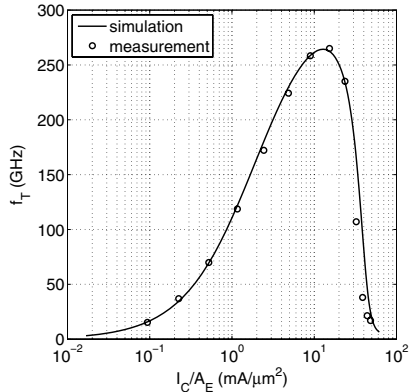

Figure 4. Transit frequency vs. collector current density: comparison between measurement and complete model (HICUM).

IV. DEVICE DESIGN CONSIDERATIONS

Starting from the baseline, the doping profile was scaled successively within the 1D simulation loop in Fig. 2 until for f_T the target value of around 500GHz was reached. This target was determined from rough calculations taking into account the expected influence of parasitics and 2D/3D effects, and considering a balanced design with not too low f_T compared to f_{max}. As a result, the profile alternatives shown in Fig. 5 were obtained. Their fundamental differences are the lightly doped emitter region (for preventing tunneling) and a spatially variable collector profile (for reducing the BC depletion capacitance). A non-local calculation yielded for all three cases BV_{CEO} = 1.37 V, since the collector is shorter than the energy relaxation length.
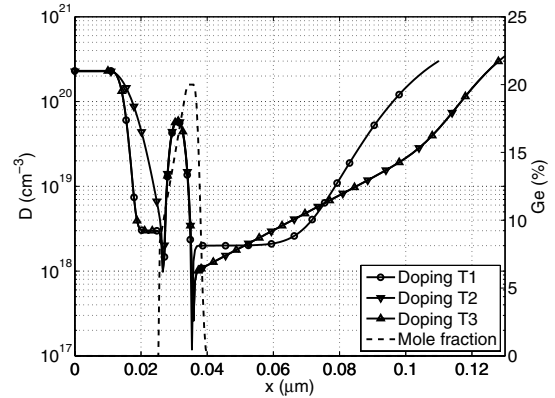

Figure 5. 1D device design alternatives for a 500GHz SiGe HBT.

Moreover, a sensitivity analysis was performed in order to obtain a feel for the variation of f_T and f_{max} as a function of process variables. Fig. 6 shows one of the results vs. internal collector width w_{Ci} and doping N_{Ci}. As expected, the opposite direction is observed for the dependence on collector doping. In this 1D case, f_{max} was calculated analytically but with the internal base resistance and capacitances calculated for typical emitter dimensions, and the external base resistance calculated for two different BE spacer widths. As can be seen, the spacer related base resistance significantly influences f_{max} in advanced HBTs, and so do the emitter dimensions and contact configuration. From these evaluations it was determined that with the profiles shown in Fig. 5 the target of f_{max} = 500GHz should be achievable with a properly chosen layout. Note that f_{max} in Fig. 6 is larger than the target value due to the yet missing parasitic elements.
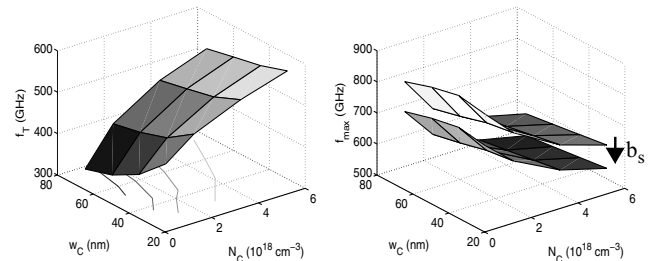

Figure 6. Sensitivity of (a) f_T and (b) f_{max} as function of internal collector width and doping. In (b) the BE spacer width was reduced from 50nm to 35nm.

The profile T3 was selected for further consideration since it has the lowest internal base sheet resistance and base-collector capacitance, and thus will allow for the most relaxed lateral dimensions at a desired f_{max} of 500GHz. The result of the HICUM parameter extraction is shown in Fig. 7, demonstrating the excellent accuracy of the model.
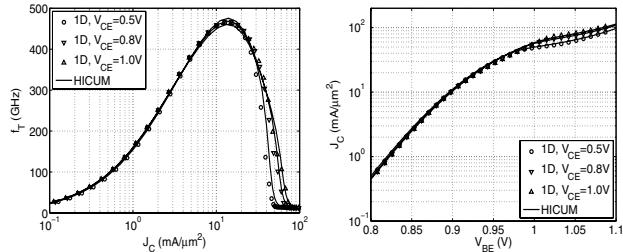

Figure 7. Comparison between HICUM and 1D HD device simulation for V_{CE}/V = 0.5, 0.8, 1: (a) transit frequency (simulation results show average between DD and HD), (b) collector current density.

Starting the second loop in Fig. 2, the lateral dimensions were reduced by the square root of the 1D f_T ratio (= 310/470). As shown in Fig. 8a, the corresponding result yielded for different transistor configurations f_{max} values around 500GHz. Fig. 8b exhibits the corresponding CML gate delay τ_{CML} along with the best presently existing experimental data [18]. The latter were obtained for an $(f_T, f_{max}) = (300, 350)$ GHz. The obtained 1.85 ps correlates very well with the experimental data for the slower process. Note that τ_{CML} can be further improved by reducing the voltage swing ΔV.
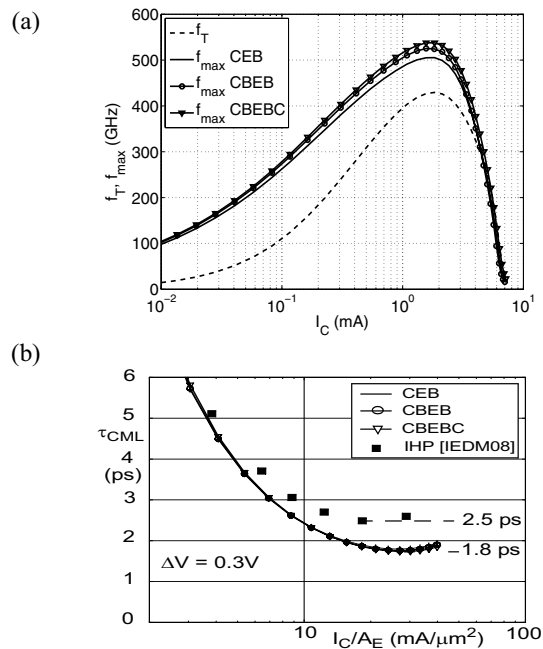
(a)

(b)

Figure 8. Predicted results for (a) f_T, f_{max} and (b) CML gate delay for different transistor configurations (0.1*0.7mm^2). The filled circles in (b) indicate measured data from [18].

V. CONCLUSIONS

A two-step procedure for a rapid TCAD based evaluation of device design alternatives has been presented. The use of 1D

DD and HD device simulations in combination with a physics-based compact model and a corresponding model generator meets the requirements for providing early libraries with geometry scalable models for circuit optimization. The procedure has been applied to study the feasibility of SiGe HBTs for mm-wave applications beyond 120 GHz by realistic scaling of the vertical and lateral dimensions. Instead of f_T the more circuit design and application related parameters f_{max} and τ_{CML} are used as target figures of merit. Based on an experimentally calibrated baseline, the results show that SiGe HBTs with $f_{max} \approx 520$ GHz and $\tau_{CML} = 1.85$ ps are within reach of presently existing process capabilities. The proposed method can also be used for producing a SiGe HBT roadmap and explore the limitations of this technology. Considering the difference between HD, DD and BTE results it is felt though that there is still a need for better calibration of physical device simulation models.

REFERENCES

[1] U. Pfeiffer, E. Ojefors, A. Lisauskas, H. Roskos, "Opportunities for silicon at mmWave and Terahertz frequencies", Proc. IEEE BCTM, pp. 149-156, 2008.

[2] M. Schroter, "Future Opportunities for bipolar transistor technology", inv. pres. at IEEE SCD, Germany, April 29, 2009.

[3] http://www.dotfive.eu/

[4] Y.Shi, G.Niu, "Vertical Profile Design and Transit Time Analysis of Nano-Scale SiGe HBTs for Terahertz f_T", Proc. IEEE BCTM, pp. 213-216, 2004.

[5] C. Jungemann, "Physics-based device simulation", DOTFIVE Report 2008.

[6] G. Avenier et al. „0.13μm SiGe BiCMOS Technology for mm-Wave Applications", Proc. IEEE BCTM, pp. 89-92, 2008.

[7] M. Schroter, "High-frequency circuit design oriented compact bipolar transistor modeling with HICUM", (inv. paper), IEICE Trans. Electronics, Special Issue on Analog Circuit and Device Technologies, Vol. E88-C, No. 6, pp. 1098-1113, 2005.

[8] M. Schröter, H.-M. Rein, W. Rabe, R. Reimann, H.-J. Wassener and A. Koldehoff, „Physics- and process-based bipolar transistor modeling for integrated circuit design", IEEE Journal of Solid-State Circuits, Vol. 34 , pp. 1136-1149, 1999.

[9] A. Pawlak, M. Schroter, "Accurate analytical representation of device figures of merit for process optimization", internal report.

[10] M. Schroter, H. Wittkopf, W. Kraus, "Statistical modeling of bipolar transistors", Proc. IEEE BCTM, pp. 54-61, 2005.

[11] M. Schröter, "Transient and small-signal high-frequency simulation of numerical device models embedded in an external circuit", COMPEL, Vol. 10, No. 4, pp. 377-378, 1991.

[12] G. Wedel, "Influence of semiconductor transport modeling on the simulation of advanced SiGe HBTs", Dipl-.Ing. thesis (in German), CEDIC, TU Dresden, 2008.

[13] Minimos-NT - Device & Circuit simulator. 2.1, TU Vienna, IME, 2004.

[14] ATLAS User's Manual, Silvaco Int., 2000.

[15] Sentaurus Device User Guide V. Z-2007.03, Synopsys, 2007.

[16] GALENE III User's Guide, ITEM, Univ. Bremen, 3.2 ed., 2002.

[17] MONJU User's Guide V. 1.6.5, ITEM, TU Braunschweig, 2006.

[18] A. Fox et al., "SiGe HBT module with 2.5ps gate delay", IEEE IEDM Techn. Digest, pp. 731-734, 2008.