Charge-storage related parameter calculation for Si and SiGe bipolar transistors from device simulation

M. Schroter[1),2)] and H. Tran[2)]
[1)]ECE Dept., University of California San Diego, La Jolla, CA, USA
[2)]CEDIC, Dresden University of Technology, Mommsenstr. 13, 01062 Dresden, Germany

Abstract: Various methods for calculating regional charge storage components in bipolar transistors from device simulation results are compared with respect to their suitability for compact modeling. The methods are evaluated for Si and SiGe transistors with very different doping profiles representing existing process technologies. Causes for the failure of a certain method under certain conditions are discussed. Such regional partitioning methods are also very useful for process development.

1 Introduction

Charge-storage effects determine the speed of semiconductor devices such as heterojunction bipolar transistors (HBTs). For circuit design, numerically efficient but accurate analytical equations are needed for describing charge storage in compact models. Since such analytical formulations usually can only be obtained through simplifying assumptions, they need to be compared to reference results from either measurements or device simulation. For verifying single components of a compact model or formulations for the intrinsic transistor, device simulation is the reference of choice.

Already in one of the first papers [1] describing the electrical behavior of semiconductor devices, including charge-storage elements, the analysis was simplified by subdividing the device in space-charge regions (SCRs) and neutral regions (NRs). This partitioning method became well-known under the name *Regional Approach* (RA). Practical applications started with the availability of first computer-based solutions of the semiconductor equations (e.g. [2][3]). While the RA is simple for analytical derivations due to the *a priori* assumption of abrupt boundaries between SCR and NR, the application to device simulation results is more complicated since abrupt boundaries do not exist in reality. Therefore, various methods have been suggested over time for defining the boundaries and for calculating charge-storage related parameters from device simulation.

In this paper, the most important approaches for determining important parameters related to charge storage effects in Si and SiGe HBTs are discussed and evaluated with respect to their suitability for compact modeling. The focus will be on the intrinsic, i.e. one-dimensional (1D), transistor behavior but the conclusions can be extended to the 2D and 3D case.

2 Investigated Device Structures

Fig. 1 shows the doping profiles used for 1D device simulation. Different transistor types were investigated: one with a conventional emitter doping (CED) profile and one with a low emitter concentration (LEC). The CED transistor A has a *base* profile optimized as BJT, while the base profile of CED transistor B was optimized as SiGe HBT. All profiles are representative of presently existing production processes with transit frequencies in the range of 55...115GHz (Fig. 2). The simulated results correspond to npn transistors with a unit emitter area $A_{E0} = 1\mu m^2$.

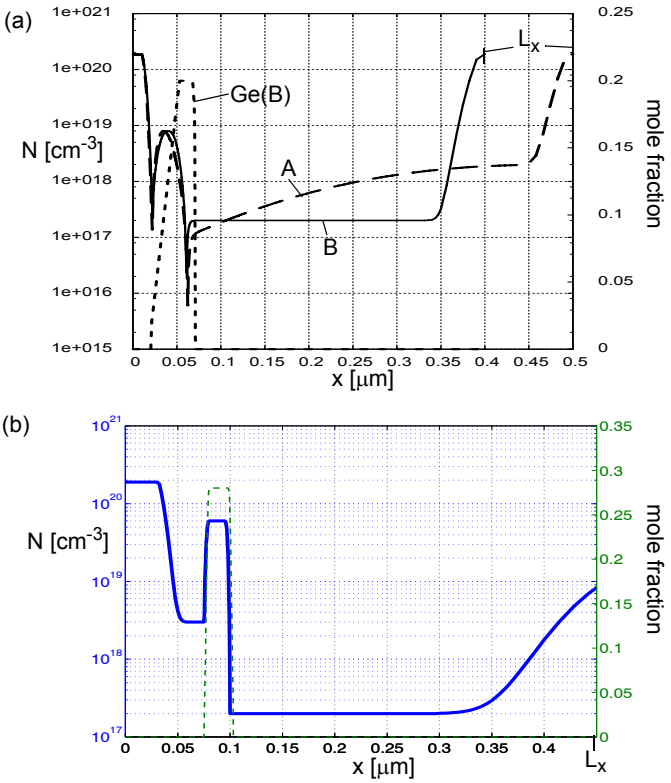

Fig. 1: 1D doping profiles under investigation: (a) CED profile A (dashed) without Ge and profile B (solid) with Ge (dotted); (b) LEC SiGe HBT profile (solid) and Ge profile (dashed).

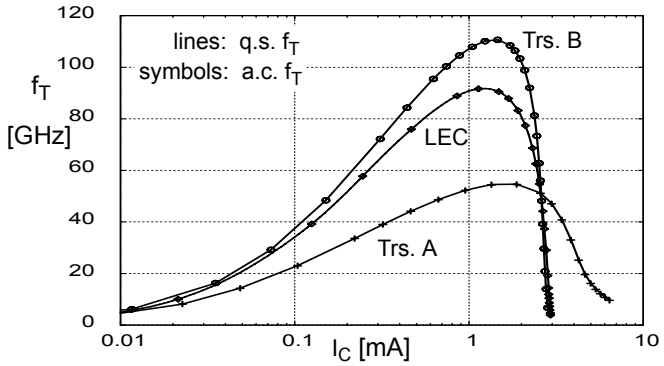

Fig. 2: Transit frequency f_T vs I_C for the investigated transistors obtained from q.s. method (solid lines) and measurement based a.c. method (symbols). $V_{C'E'} = 0.8V$.

3 Overview on regional approaches

Early approaches (e.g. [3]) started from the d.c. carrier densities and defined the SCR boundary x_r as the point where the ratio r of majority carrier density c (=p or n) to net doping density N drops below a certain value (<< 1). This definition follows the classical device analysis [1] and only works for sufficiently low injection. A more suitable criterion is based on the space(-charge) density normalized to carrier densities

$$x_r = x\left(\frac{N + p - n}{p + n} = r \ll 1\right), \tag{1}$$

where the boundary is approached *from the neutral region*. This simple method, designated as DCRA, is used in the comparisons later to calculate the desired charges within the detected region boundaries. A fundamental issue with this definition is its missing link to small-signal definitions and measurements of charge-storage elements.

A first consistent calculation for the capacitances of a SCR was given in [4], where the bias dependent change of the carriers (dp, dn) was integrated rather than the depleted doping density. Here, also the designation "neutral capacitance" was coined for the portion describing the neutralized charge *within* a SCR. However, in [4] the SCR boundaries were assumed a priori and no adjacent NRs were considered. In [5], the depletion and neutral capacitance were calculated from device simulation results. There, for the first time, *electrical* junctions were defined,

$$x_m = x(d\rho = 0) = x(dp = dn). \tag{2}$$

It was also mentioned that the "*positions of the peaks of the small signal space charge distribution at low forward bias voltages are closely associated with the depletion layer edges of the classical models*" [5]. This analysis was performed though for a pn diode only.

For compact transistor modeling a relation of the stored charges to measurable quantities is required. Charges can be measured from small-signal S-parameters via capacitances and transit times. Using charge-control theory one can show that the transit frequency f_T is given by

$$\frac{1}{2\pi f_T} = \left.\frac{dQ_p}{dI_C}\right|_{V_{C'E'}} \tag{3}$$

where

$$dQ_p = qA_E\int_0^{L_x} \partial p \, dx \tag{4}$$

is the hole charge *change* in the 1D transistor, including depletion and minority charge contributions.

Following [5], a regional analysis was proposed in [6] that for a bipolar transistor consistently defines the time constants which determine f_T *at all current density levels*:

$$\left.\frac{dQ_p}{dI_C}\right|_{V_{C'E'}} \equiv \tau_{ec} = \tau_e^* + \tau_{eb} + \tau_b^* + \tau_{bc} + \tau_c^* . \tag{5}$$

The various time constants are calculated based on the electrical junctions x_{me} and x_{mc} defined by (2). The contributions of mobile charge from emitter, base and collector then read

$$\tau_e^* = qA_E\int_0^{x_{me}} \left.\frac{\partial p}{\partial I_C}\right|_{V_{C'E'}} dx , \tag{6}$$

$$\tau_b^* = qA_E\int_{x_{me}}^{x_{mc}} \left.\frac{\partial n}{\partial I_C}\right|_{V_{C'E'}} dx , \tag{7}$$

$$\tau_c^* = qA_E\int_{x_{mc}}^{L_x} \left.\frac{\partial p}{\partial I_C}\right|_{V_{C'E'}} dx , \tag{8}$$

while the "depletion region charging times" are given by

$$\tau_{eb} = qA_E\int_0^{x_{me}} \left.\frac{\partial(n-p)}{\partial I_C}\right|_{V_{C'E'}} dx , \tag{9}$$

$$\tau_{bc} = qA_E\int_{x_{mc}}^{L_x} \left.\frac{\partial(n-p)}{\partial I_C}\right|_{V_{C'E'}} dx , \tag{10}$$

which include the depletion capacitances. It is obvious, that this method only allows the mobile charge related time constant to be divided into components associated with the *total* emitter, base and collector region, but *not* with SCRs and NRs. For compact modeling, however, it is desirable to distinguish between depletion capacitances and delay or minority storage times. Also, compared to the profiles investigated in [6], the neutral BE storage time at low current densities is not negligible anymore in modern transistors; this was actually already the case for the profiles investigated in [7][8]. As a consequence, this method is useful only for a rough component analysis but it is of *limited suitability* for developing physics-based compact models.

In order to overcome this problem and to allow a more detailed analysis of charge-storage effects, in [8] the peaks of the small-signal space-charge density $d\rho = q(dp-dn)$ were defined as boundaries between SCRs and NRs. The method starts with the "dynamic" definition of minority (mobile) carriers

$$\partial m = \begin{cases} \partial p & , |\partial p| < |\partial n| \\ \partial n & , |\partial n| \le |\partial p| \end{cases} . \tag{11}$$

Furthermore, for compact modeling the most suitable partitioning of dQ_p in (4) is

$$dQ_p = dQ_j + dQ_m \tag{12}$$

with dQ_j as depletion and dQ_m as minority charge variation in the (1D) transistor. From (4), (11) and (12) follows for the depletion charge [8]

$$dQ_j = dQ_p - dQ_m = qA_E\int_{w(\partial p < \partial n)} \partial(n-p)dx \tag{13}$$

which equals $\tau_{eb}+\tau_{bc}$ from (9) and (10) for dynamic CE short. In the latter case, however, dQ_{jC} contains the carrier modulation in the BC SCR due to a BE voltage change. In order to separate BE and BC related charge variations, the hole charge variation is written as

$$dQ_p = \left.\frac{\partial Q_p}{\partial V_{B'E'}}\right|_{V_{B'C'}} dV_{B'E'} + \left.\frac{\partial Q_p}{\partial V_{B'C'}}\right|_{V_{B'E'}} dV_{B'C'} \tag{14}$$

which holds for quasi-static (q.s.) operation and is consistent with the definition of f_T. The charge derivatives define

the capacitance (and transit times) associated with the respective junction voltage. This separation is also used for determining the SCR boundaries, which are defined by the peaks of $d\rho$ with respect to the BE or BC voltage change:

$$w_E = x\left(\left.\frac{\partial\rho}{\partial V_{B'E'}}\right|_{V_{B'C'}} = \min\right), \quad x_e = x\left(\left.\frac{\partial\rho}{\partial V_{B'E'}}\right|_{V_{B'C'}} = \max\right), \quad (15)$$

$$x_{C,c} = x\left(\left.\frac{\partial\rho}{\partial V_{B'C'}}\right|_{V_{B'E'}} = \min\right), \quad x_c = x\left(\left.\frac{\partial\rho}{\partial V_{B'C'}}\right|_{V_{B'E'}} = \max\right). \quad (16)$$

The width of the BE SCR is then given by

$$w_{BE} = x_e - w_E \ , \quad (17)$$

and the neutral base is

$$w_B = \min\{x_c, x_{jc}\} - x_e \ . \quad (18)$$

The limitation to the BC junction width is necessary at high current densities, when a possible injection zone forms, that starts at x_{jc} and ends at x_{Cc}; i.e. its width is given by

$$w_i = \max\{x_c - x_{jc}, 0\} \ ; \quad (19)$$

The width of the BC SCR under all bias conditions is then

$$w_{BC} = x_{Cc} - x_c \ . \quad (20)$$

Fig. 3a,b visualizes the above definitions by means of the dynamic space-charge density distribution at two different collector current densities - one far below peak f_T and one at peak f_T. In (a) the peaks left and right from the BE junction are clearly visible, while at the BC junction they are uniquely defined only in the low current curve. However, as shown in (b), for a BE short the peaks at the BC junction are clearly defined also at high current densities. Note the very different magnitude that a V_{BE} and V_{BC} change have on the carrier dynamics. At high current densities (dashed lines), the V_{BE} controlled modulation of the (mobile) electron charge in the BC SCR (cf. (a)) is much larger than the actual space-charge change at the BC SCR boundaries (cf. (b)). Therefore, the boundaries of the BC SCR cannot be detected from a calculation with CE short. The additional zero-crossing of $d\rho(dV_{CE} = \text{const})$ in the collector would need to be ignored to obtain meaningful results from the scheme in [6].

Once the locations of the SCR boundaries and the resulting region widths are known, the transit time components of both neutral and space-charge regions can be easily calculated from the accumulated minority transit time

$$\tau_m(x) = qA_E \int_0^x \left.\frac{\partial m}{\partial I_C}\right|_{V_{C'E'}} d\xi \ , \quad (21)$$

which now uses the minority carrier change at *CE short* in order to be consistent with (3). The final value

$$\tau_{m\Sigma} = \tau_m(L_x) \quad (22)$$

corresponds to the total minority storage time for a given bias point. Fig. 3c shows $\tau_m(x)$ for the same current densities as before and, in addition, for a very high current density, where an injection region w_i exists.

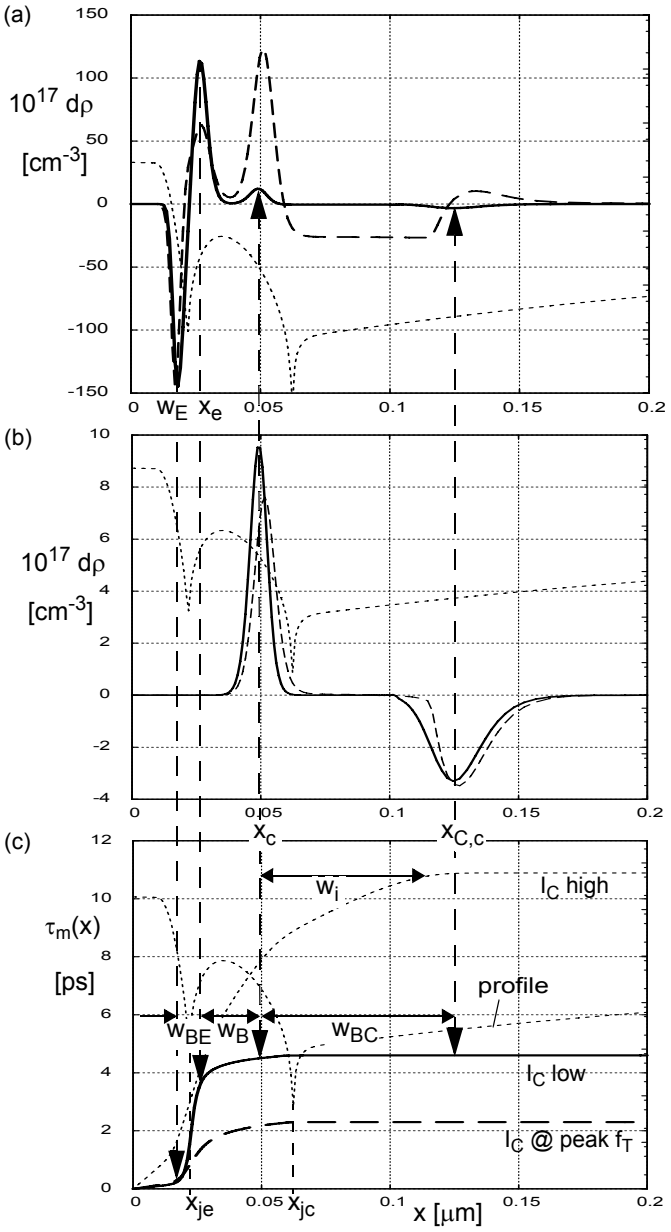

Fig. 3: Dynamic space-charge density distribution for transistor A at different collector current densities: (a) BE related change (CE short!), (b) BC related change (BE short).
(c) Accumulated transit time. Bias currents: 0.023mA (solid), 1.35mA (dashed), 5.33mA (dotted); $V_{C'E'}$ = 0.8V. The doping profile (dotted) is not drawn to scale.

From (13) the depletion capacitances are defined as

$$C_{jEi} = qA_E \int_0^{x_{me}} \left.\frac{\partial(-\rho)}{\partial V_{B'E'}}\right|_{V_{B'C'}} dx \ , \ C_{cE} = qA_E \int_{x_{mc}}^{L_x} \left.\frac{\partial(-\rho)}{\partial V_{B'E'}}\right|_{V_{B'C'}} dx \quad (23)$$

$$C_{jCi} = qA_E \int_{x_{mc}}^{L_x} \left.\frac{\partial\rho}{\partial V_{B'C'}}\right|_{V_{B'E'}} dx \ ,$$

$$C_{eC} = qA_E \int_0^{x_{me}} \left.\frac{\partial\rho}{\partial V_{B'C'}}\right|_{V_{B'E'}} dx \ . \quad (24)$$

Here, C_{jEi} and C_{jCi} correspond to the usual depletion capacitances while C_{eC} ($\approx$0) and C_{cE} are cross-controlled

capacitances. C_{cE} contains the delay through the BC SCR plus the time constant $r_{Ci}C_{jCi}$. Notice that neither (14) nor the above capacitances define the *location* of these elements in an equivalent circuit. Only C_{jEi} and C_{jCi} are physically clearly associated with the corresponding junctions.

This full RA (FRA) has enabled the development and verification not only of analytical transit time equations (e.g. [7][8][9]) but also of a physics-based transfer current relation [10] employed in compact models (such as HICUM). However, the method described above does not always work for SiGe HBTs due to the bandgap induced potential barriers at the junctions and the associated dipole layers. In most cases, this issue can be overcome by properly choosing the start and direction for the search of the relevant $d\rho$ peaks. However, in cases of a too abrupt composition profile change (which often occurs in exploratory *simulation* profiles) close to a SCR boundary the true depletion charge related peak is very difficult to distinguish from the dipole layer charge related peaks. Our implemented code tries to detect dipole layers and to correct the overall $d\rho(x)$ by the dipole layer related charge variation.

An alternative method for partitioning the device into its SCRs and NRs was devised in [11]. There, the SCR boundary is defined by the location at which the accumulated *electron* and *hole* storage time

$$\tau_n(x) = qA_E\int_0^x \frac{\partial n}{\partial I_C}\bigg|_{V_{C'E'}} d\xi \quad , \quad \tau_p(x) = qA_E\int_0^x \frac{\partial p}{\partial I_C}\bigg|_{V_{C'E'}} d\xi \qquad (25)$$

differ by 5%. Fig. 4 shows an example for the accumulated carrier storage times at different collector current densities.

At the SCR edge, the majority storage time is expected to change rapidly to support the charging of the depletion capacitance, while the minority component remains unchanged. This is clearly visible in Fig. 4 for low current densities at the BE and BC junction. However, the relative deviation tends to disappear at higher current densities. In fact, for the curve labeled "I_C high" there is almost no difference anymore between the storage times at the BE junction, and also the difference at the BC junction becomes very small, making the detection of SCR boundaries difficult or impossible.

The storage times (25) include the depletion capacitance contributions C_{jEi}/g_m and C_{jCi}/g_m. This causes the final storage time at L_x for low current densities to be significantly larger than in Fig. 3c. Furthermore, in [11] the end of the BC SCR is always set to the location L_x of the collector contact since the space-charge modulation within the BC SCR makes it difficult to apply the boundary detection criterion mentioned earlier for determining the BC SCR width. The detection criterion also usually does not work to detect an injection zone at high current densities. Although this approach provides more insight into the behavior of the different device regions than that in [6], it is not (immediately) suitable for compact modeling due to the missing partitioning into depletion and minority charge related effects and the missing detection of the physically different collector layers.

In [11], τ_E was reduced to $2/3\tau_E$ "because not all minority carriers contribute to the AC capacitance". This applies only for non-quasi-static (n.q.s) operation while storage times in compact models are defined for quasi-static operation, and n.q.s. effects are taken into account separately. Therefore, the reduction factor has been omitted here in the comparisons.

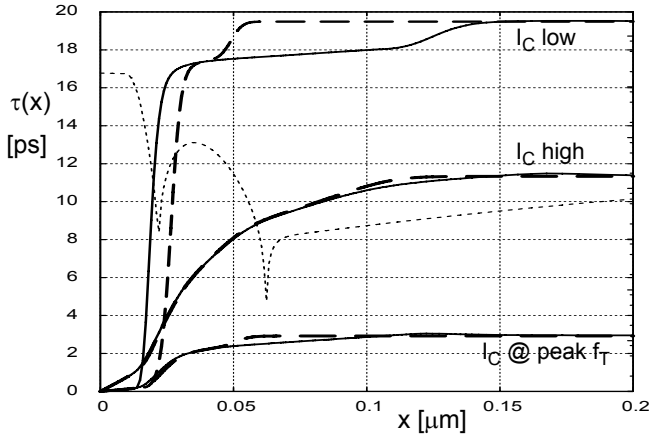

Fig. 4: Accumulated electron (solid) and hole (dashed) storage times for transistor A at different collector currents. τ (L_x) is the same for both storage times. $V_{C'E'}$ = 0.8V.

4 Results and discussion

Starting point of the dynamic RA variants was (3). The calculation of f_T from q.s. carrier distributions (via dQ_p) is verified by a comparison to the result obtained from frequency dependent small-signal (a.c.) operation, which is quite insensitive to the number of discretization points. As Fig. 2 shows, the respective curves are on top of each other. Differences can occur though from too coarse discretization and, hence, integration.

Fig. 5 contains results that are plotted in the form that is typically used for determining the transit time and high forward-bias BE depletion capacitance.

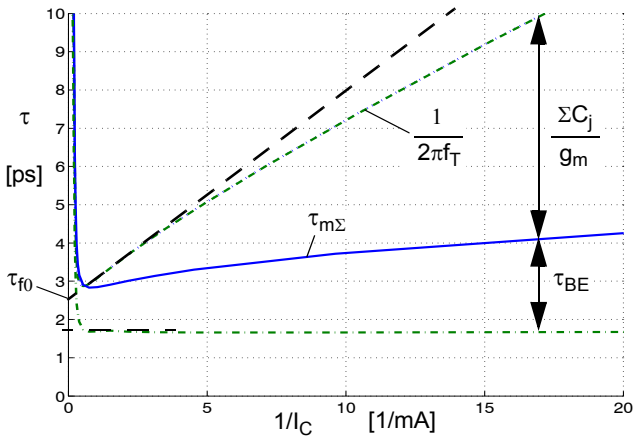

Fig. 5: Reciprocal f_T and related storage times vs $1/I_C$ for transistor A showing the partitioning into the components $\tau_{m\Sigma}$ and $\tau_{m\Sigma}$-τ_{BE}. The dashed line corresponds to the standard transit time determination method. $V_{C'E'}$ = 0.8V.

As can be seen, τ_{BE} causes an additional slight increase towards low current densities (in addition to the increase

from C_{jEi}) [12]. Thus, using the standard method (dashed line) to also determine C_{jEi} yields a larger value (τ_{f0}) which contains part of the neutral charge. If τ_{BE} is subtracted from $\tau_{m\Sigma}$, a flat line is obtained, which is consistent with the determination method, but gives a lower value than the standard transit time determination method. This comparison to the actual values provides a feeling for the validity and accuracy of the determination method.

Fig. 6a shows a comparison of the different regional approaches for transistor A. The neutral base storage time τ_B differs significantly for the different methods, even at low currents. The "bumpy" curve of DCRA reflects the discontinuities in the corresponding width detection. The discontinuous curve of DTRA at about 4mA is caused by the Kirk effect; the associated neutral BC junction region does not allow the detection of a difference between $\tau_n(x)$ and $\tau_p(x)$ anymore, so that the base *and* collector region are assigned to w_{BC}. The FRA gives a smooth τ_B curve that increases at high injection, as expected, at the same rate as the neutral collector storage time τ_C due to the Kirk effect. Neither DCRA nor DTRA are able to detect the collector injection zone and, hence, lead to $\tau_C = 0$.

The BC storage time is similar for FRA and DTRA, except at low currents, where DTRA also contains C_{jCi}/g_m and, at high injection, also the neutral collector charge. The abrupt increase of τ_{BC} for DTRA is due to the missing detection of the neutral base end x_c discussed above. The values of τ_{BC} from DCRA does not contain a.c. related charges and, thus, is much smaller.

The neutral emitter storage (cf. Fig. 6b) is fairly similar for all three methods up to quite high currents, where τ_E then rapidly drops to zero for DCRA and DTRA. This happens when the BE junction region becomes "too"neutral to detect significant differences in d.c. carrier densities and a.c. storage times. In contrast, the FRA still detects the $d\rho$ peaks on both sides of the BE junction. Finally, the methods also show large differences for the BE SCR related storage time τ_{BE}. This is understandable for DTRA which contains the C_{jEi} related contribution, which is significant at lower currents. The DCRA method is unable to detect the base-sided boundary x_e of the BE SCR. At high injection, the increasing quasi-neutrality makes DTRA unable to detect any SCR boundaries so that the τ_{BE} curves of FRA and DTRA diverge. Overall, FRA gives the most reliable results which - per definition - are also consistent with small-signal measurements.

Fig. 7a (upper figure) contains a comparison of the different regional approaches for the SiGe HBTs B in Fig. 1a. The neutral base storage time τ_B now differs much less for the various methods. At low injection DCRA and DTRA are very close, but up to 40% smaller than FRA. At high currents, the BC barrier due to the Ge drop, which is very close to the BC junction, and the resulting dipole layer cause a carrier jam within the base that leads to a very rapid increase of the base minority charge. This is properly detected by DCRA and FRA, but not by DTRA.

The BC storage time of DTRA and FRA is again very similar, and the difference at lower current densities is caused by C_{jCi}/g_m. Also, DCRA again yields much smaller values for the same reasons as already discussed before.

The results for τ_{BE} at low injection are similar to the Si transistor. However, now DCRA yields useful results that are fairly close to those of FRA. The causes for the differences to the DTRA are the same as discussed before.

The τ_E curves (lower Fig. 7a) for all methods are quite smooth up to high currents. The results still differ somewhat though, but since τ_E is extremely small for SiGe HBTs (note the different y-axis scales in Fig. 7a), the difference does not matter for practical purposes. There is a very small injection zone, that is limited by the barrier, but is detected by the FRA. Due to the carrier jam, the stored charge increases very rapidly with current.

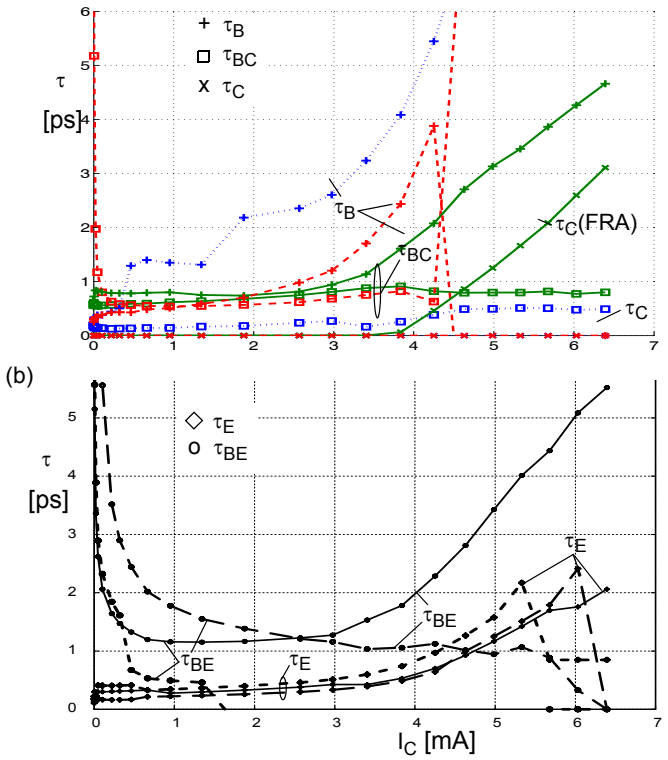

Fig. 6: Storage time components vs collector current for the Si BJT (transistor A). Comparison between DCRA (dotted), FRA (solid) and DTRA (dashed): (a) τ_B, τ_{BC}, τ_C and (b) τ_E, τ_{BE}. $V_{C'E'} = 0.8V$.

For the LEC transistor in Fig. 7b (upper figure), the base storage time dominates the transit time while the emitter storage time is negligibly small due to the box Ge profile. Both τ_B and τ_E are detected similarly by all methods. The differences in τ_{BE} between DTRA and FRA or DCRA are again caused by the C_{jEi} related storage time.

Since the Ge drop coincides with the BC junction, the associated barrier prevents the formation of an injection zone at high current densities (i.e. $\tau_C = 0$). The corresponding hole charge increase, caused by the carrier jam in front of the barrier, is accounted for by τ_{BC} in all methods. Again, DTRA and FRA are on top of each other at higher current densities, while the DCRA is unable to detect the a.c. storage time components. Knowing the BC SCR width and C_{jCi}, the FRA allows to calculate the bias dependent internal collector resistance r_{Ci} and then, from $r_{Ci}C_{jCi}$, even the actual delay time through the BC SCR.

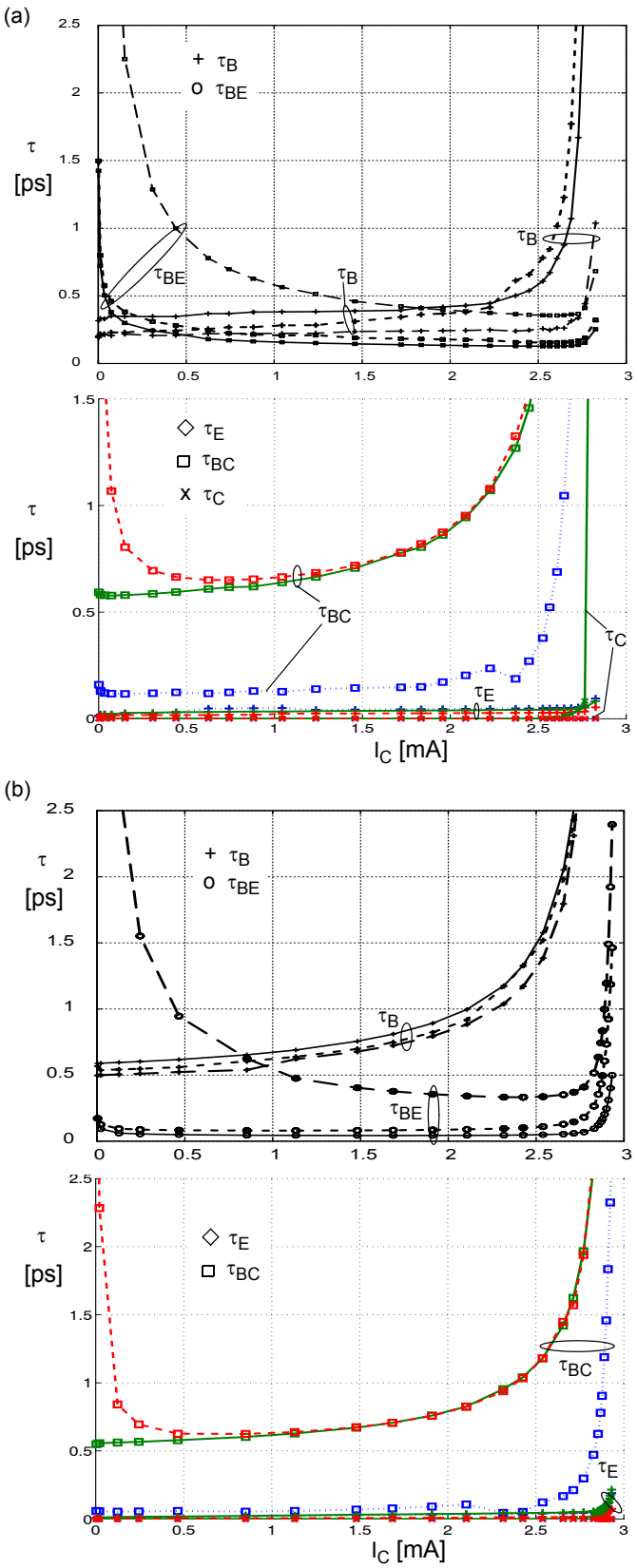

Fig. 7: Storage time components vs collector current for (a) transistor B and (b) transistor LEC. Comparison between DCRA (dotted), FRA (solid) and DTRA (dashed): τ_B, τ_{BE} (upper figure) and τ_E, τ_{BC}, τ_C (lower figure). Note that τ_C is not detected in all cases. $V_{C'E'}$ = 0.8V.

5 Conclusion

Various methods for separating neutral and space-charge regions in Si and SiGe bipolar transistors have been evaluated with respect to their usefulness for compact modeling and process development. The investigated doping profiles represent existing process technologies. 1D device simulation was used, which is sufficient for analyzing the fundamental device behavior.

The evaluation shows, that the often used method from [6] is not suitable for compact modeling since it does not distinguish between neutral and space-charge layers. This left for a comparison only a dc regional approach (DCRA) [3], a method based on the difference between accumulated electron and hole storage time (DTRA) [11], and a minority and depletion charge based full regional approach (FRA) [8]. The latter was extended for HBTs and requires a very careful implementation to enable the detection of undesired region boundaries (e.g. space-charge spikes due to bandgap variations). Overall, DCRA and DTRA turned out to be somewhat unreliable while FRA yielded correct results. Furthermore, since DTRA does not define and, hence, separate depletion capacitances it is of limited usefullness for both compact model and process development. In contrast, FRA provides clear definitions for storage times and depletion capacitances up to high current densities and seems to yield the most reliable results. For other collector voltages similar results are obtained.

The FRA can also be applied to, e.g., the evaluation of regional weighting factors of the generalized Integral Charge-Control relation.

Acknowledgments: This project has been partially supported by the German Ministry for Research and Technology within the SFB358 project. The authors are also grateful to Atmel Germany, Heilbronn, and ST Microelectronics, Grenoble (France), for financial support.

References

[1] W. Shockley, “The theory of pn junctions in semiconductors and pn junction transistors”, BSTJ, Vol. 28, pp. 435-489, 1949.

[2] R.B. Schilling, “A Regional Approach for Computer-Aided Transistor Design”, IEEE Trans. Education, Vol. E-12, 1969, pp. 152-161.

[3] D.J. Roulston, S.G. Chamberlain and J. Sehgal, “Simplified Computer-Aided Analysis of Double-Diffused Transistors Including Two-Dimensional High-Level Effects”, IEEE Trans. Electron Dev., Vol. ED-19, pp. 809-820, 1972.

[4] Morgan and Smits, “Potential distribution and capacitance of a graded pn junction“, BSTJ, Vol. 28, pp. 1573-1602, 1960.

[5] C.Y. Kwok, “Space charge and injection capacitances of pn junctions from small signal numerical solutions”, J. Appl. Phys., Vol. 16, pp263-267, 1983.

[6] J.J. van den Biesen, “A Simple Regional Analysis of Transit Times in Bipolar Transistors”, Solid-State Electronics, Vol. 29, pp. 529-534, 1986.

[7] H.-M. Rein, H. Stübing, M. Schröter, „A compact physical large-signal model for high-speed bipolar transistors at high current densities - Part I and Part II”, IEEE Trans. Electron Dev., Vol. 34, pp. 1741-1761, 1987.

[8] M. Schroter, "A compact physical large-signal model for high-speed bipolar transistors with special regard to high current densities and two-dimensional effects", Ph. D. thesis (in German), Ruhr-University Bochum, Germany, 1988.

[9] M. Schröter and T.-Y. Lee, „A physics-based minority charge and transit time model for bipolar transistors“, IEEE Trans. Electron Dev., Vol. 46, pp. 288-300, 1999.

[10] H. Tran, “Physics-based analytical modeling of SiGe heterojunction bipolar transistors for high-speed integrated circuits“, Dissertation, University of Technology Dresden, Germany, 2006.

[11] P. Pallestri, M. Mastrapasqua, A. Pacelli, C. King, “A drift-diffusion/Monte-Carlo simulation methodology for $Si_{1-x}Ge_x$ HBT design”, IEEE Trans. Electron Dev., pp. 1242-1249, 2002.

[12] M. Malorny, M. Schroter, D. Celi, D. Berger, “An improved method for determining the transit time of Si/SiGe bipolar transistors”, Proc. BCTM, pp. 229-232, 2003.