Uniqueness Problems in Compact HBT Models Caused by Thermal Effects

Matthias Rudolph, *Member, IEEE*

***Abstract*—This paper identifies possible numerical instabilities in compact HBT models, which are introduced by a physical meaningful self-heating description. Removing these possible sources of nonconvergence would deteriorate model accuracy since they originate from the device physics. It is, therefore, necessary to be aware of them during parameter extraction and circuit simulation.**

***Index Terms*—Equivalent circuit, HBT, semiconductor device modeling.**

I. INTRODUCTION

IF A COMPACT transistor model should be useful for circuit design, it has to fulfill several requirements, among which are the overall accuracy, the possibility to determine all (physical significant) parameters unambiguously, and finally, it must be numerically robust so that the simulation converges fast in all cases the circuit designer might think of. While these three requirements are equally important, the majority of publications is dedicated to accuracy issues, while only a small minority addresses questions of numerical stability.

In order to achieve numerical robustness, the model is to be formulated to allow iterative calculation. In [1] and [2], some "common sense" rules to define stable models are listed and explained, which are as follows.

- The functions employed must be defined for all possible input values.
- They must be continuous and have continuous derivatives.
- No numerical overflow or underflow should appear.
- The model must work with every parameter set, not only for a specific parameter range.

While the first three points can be evaluated easily for each electrical branch individually, this no longer is the case for advanced models that account for self-heating. The mutual interaction between operating temperature and branch currents forms a feedback structure that has to be investigated together. This issue is addressed in [4] and [5] for field-effect transistor (FET) models. In the case of HBT models, the focus thus far has been on the thermal instability in multifinger devices, which leads to formation of hot spots [3]. The approach in this case is to form a thermally distributed model by combining lumped models for each emitter finger. However, to the author's knowledge, no investigation of the impact of self-heating on the lumped model's numerical stability exists thus far.

Manuscript received August 26, 2003; revised February 4, 2004.

The author is with the Ferdinand-Braun-Institut für Höchstfrequenztechnik, D-12489 Berlin, Germany (e-mail: rudolph@fbh-berlin.de).

Digital Object Identifier 10.1109/TMTT.2004.827007

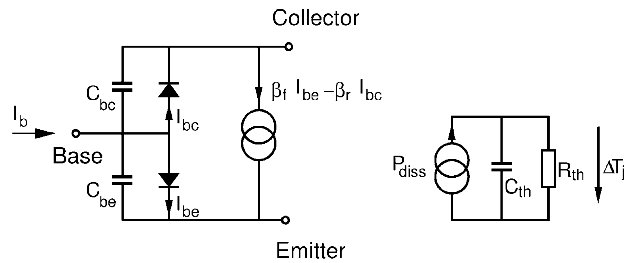

Fig. 1. Simplified intrinsic electro-thermal HBT model.

This paper attempts to identify possible multiple solutions in electro-thermal HBT models that arise from physical meaningful formulas and arbitrarily chosen model parameters. It is demonstrated that besides the well-known thermal runaway, the temperature dependence of the reversely biased base–collector diode can also cause convergence problems.

II. ELECTRO-THERMAL MODEL

To get a clear picture of the basic electro-thermal interactions, this investigation focuses on the core of every bipolar model, as shown in Fig. 1. It only contains the base–collector and base–emitter diodes and the current source that represents the forward and reverse current gain. The dissipated power is calculated and fed to the thermal equivalent circuit on the right to determine the rise in temperature due to self-heating. Of course, full-featured HBT models (e.g., [6]–[10]) have a more complex equivalent circuit. They at least account for extrinsic resistances, an extrinsic base–collector diode, and one or more additional base–emitter diodes to model nonideal base currents. However, since the instabilities already arise from the core, those elements are neglected for the moment to clarify our point.

The electro-thermal model will now be explained regarding the example of a $3 \times 30\ \mu\text{m}^2$ HBT from the Ferdinand-Braun-Institut für Höchstfrequenztechnik (FBH), Berlin, Germany, foundry [10]. Measured and simulated output *IV* curves are shown in Fig. 2. The negative slope of I_c with V_{CE} results from the temperature dependence of the forward current gain β_f, which can be modeled in a broad range of operating temperatures to decrease linearly with junction temperature

$$\beta_f(\Delta T_j) = \beta_0 - \kappa_\beta \Delta T_j \tag{1}$$

with the parameter κ_β and the temperature rise due to self-heating ΔT_j. The dashed lines in this figure show a model with

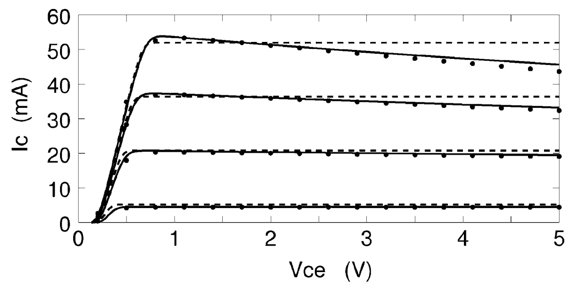

Fig. 2. Output *IV* curves measured (symbols), simulated with full model (solid lines), and simulated neglecting temperature dependence of β_f (dashed lines).

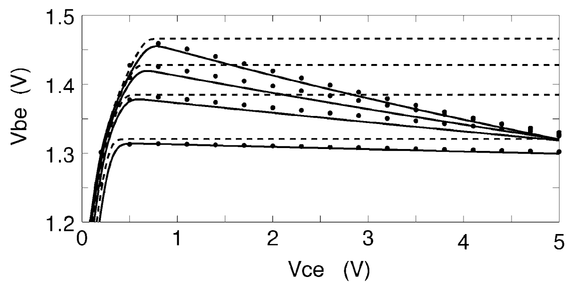

Fig. 3. $V_{\rm BE}$ obtained from output *IV* curves (parameter: I_b) measured (symbols), simulated with full model (solid lines), and simulated neglecting temperature dependence of $I_{\rm cf}$ (dashed lines).

constant β_f for comparison. Please note that the slope of the saturation region (low collector–emitter voltages) is given by the collector resistance, which is omitted in the simplified model for the reasons given above.

The diodes also depend on temperature, as the diode currents at a fixed voltage increase with temperature. Vice versa, the voltages decrease with temperature if the diode current is fixed. The latter is shown in Fig. 3. Here, the base–emitter voltage $V_{\rm BE}$ is determined while measuring the *IV* curves shown in Fig. 2. The temperature dependence of the pn-junction currents is approximately given by

$$I(V,T) = I_s e^{(V_g/V_{\rm th0} - V_g/V_{\rm th})}\left(e^{V/(nV_{\rm th})} - 1\right) \qquad (2)$$

with $V_{\rm th} = (kT)/q$ and $V_{\rm th0} = 25$ mV. V_g denotes an activation energy, n denotes the emission factor, and I_s denotes the saturation current. The dashed lines in Fig. 3 are determined by a model that does not account for this behavior. The formula is used to calculate $I_{\rm be}$ and $I_{\rm bc}$. The dominant parameter that determines the thermal behavior of the diode current is V_g. Since it is mainly given by the bandgap, the parameter for the base–emitter heterojunction will be different from that describing the base–collector junction.

Since the base–collector junction is reversely biased in normal operation, the temperature dependence of $I_{\rm bc}$ is usually of minor influence in the forward active region. However, this parameter is important to model the saturation region, as shown in Fig. 4. Setting $I_{\rm bc}$ as temperature independent results in incorrect simulated values for the turn-on voltage. While Fig. 4 only shows a change in temperature of 40 K, self-heating or environment conditions can result in junction temperature differences, which are more than three times this value. Although the investigation of numerical instabilities will focus on the forward active region, the value V_{gr} will play an important role in model stability, which is especially important since it seems to not be of any influence on device performance there.

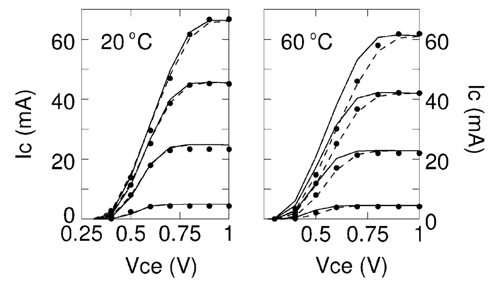

Fig. 4. Saturation region of output *IV* curves at an ambient temperature of 20 °C and 60 °C, measured (symbols), simulated with full model (dashed lines), and simulated neglecting temperature dependence of $I_{\rm cr}$ (solid lines).

The dissipated power is given by the current multiplied by the voltage across each branch. The thermal equivalent circuit generally has a low-pass behavior showing multiple time constants. However, it is common in compact models, only to account for a single time constant $C_{\rm th}R_{\rm th}$, as shown in the equivalent circuit of Fig. 1. The thermal resistance $R_{\rm th}$ is generally nonlinear, and it increases with temperature. While the reduction of the thermal conductivity of GaAs with temperature only leads to slight changes, inhomogeneous temperature distribution and formation of hot spots in the device show up as exponential increasing values of $R_{\rm th}$. It has been shown that it is possible to use a constant value of $R_{\rm th}$ in absence of hot-spot formation [6]–[8]. Since the relative error at low dissipated powers will be small, it is then useful to choose a mean $R_{\rm th}$ value at high dissipated powers.

For the stability investigation, the dissipated power in the active forward region can be approximated by

$$P_{\rm diss} \approx \beta_f I_{\rm be} V_{\rm CE} + I_{\rm bc} V_{\rm BC}. \qquad (3)$$

The first term represents the main contribution, which is the collector–emitter voltage multiplied by the collector current. All the other currents, i.e., $I_{\rm be}, I_{\rm bc},$ and $\beta_r I_{\rm bc}$, are much lower than I_c, provided that the value of β_f is high enough and the base–collector junction is reversely biased. Therefore, the second term is usually negligible. It is only accounted for since it will turn out that the base–collector diode can show thermal runaway independently of the base–emitter diode, which results in excessively high currents. Self-heating is finally defined by

$$\Delta T_j = P_{\rm diss} Z_{\rm th} \qquad (4)$$

which simplifies to $\Delta T_j = P_{\rm diss} R_{\rm th}$ in the dc case.

III. Possible Sources of Numerical Instabilities

A. *Method*

Possible numerical instability can be investigated in the following way [3]–[5]. The currents are given as functions of voltages and temperature $I = f(V,T)$, e.g., (2), while the temperature is determined by ambient temperature and dissipated power $T = T_a + R_{\rm th}P_{\rm diss} = g(V,I)$. This equation can be solved for the current I. At given voltages, the temperature dependence

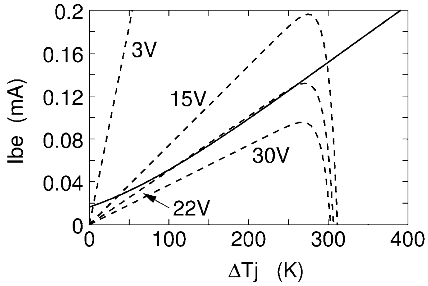

Fig. 5. $I_{\rm BE}$ calculated by (5) (solid line) and (6) (dashed lines). Parameter is $V_{\rm CE}$. The resulting value of $I_{\rm BE}$ is given by the intersection of the lines. The parameters are $V_{\rm BE} = 1.4$ V, $I_{s,\rm be} = 1.5 \cdot 10^{-27}$ A, $V_g = 1.5$ V, $I_{s,\rm bc} = 4 \cdot 10^{-22}$ A, $V_{gr} = 2.6$ V, $\beta_f = 150, \kappa_\beta = 0,$ and $R_{\rm th} = 600$ K/W.

of a current can be determined from both equations. Since both have to be fulfilled, the solution is given by the intersection of the curves. More than one (or no solution) can deter convergence not only in dc, but also in large-signal RF simulations.

The thermal runaway observed in bipolar transistors is well known. It stems from a positive feedback caused by the fact that $I_{\rm be}$ (and thereby $P_{\rm diss}$) increases with temperature, when $V_{\rm BE}$ is fixed. However, the reverse-biased base–collector junction can also show thermal runaway, as shown in the following.

B. Forward Thermal Runaway

The thermal runaway effect will first be discussed, which results from positive feedback between collector current and temperature when the base–emitter voltage is fixed.

Calculating $I_{\rm BE}$ at fixed $V_{\rm BE}$ directly from the diode characteristics, and indirectly from self-heating (4), yields the following set of equations:

$$I_{\rm BE} \approx I_{s,\rm be} e^{(V_g/V_{\rm th0} - V_g/V_{\rm th})} \left(e^{V_{\rm BE}/(nV_{\rm th})} - 1\right) \tag{5}$$

$$I_{\rm BE} \approx \frac{\Delta T_j - I_{\rm bc} V_{\rm BC} R_{\rm th}}{\beta_f V_{\rm CE} R_{\rm th}} \approx \frac{\Delta T_j - I_{s,\rm bc} e^{(V_{gr}/V_{\rm th0} - V_{gr}/V_{\rm th})} V_{\rm BC} R_{\rm th}}{\beta_f V_{\rm CE} R_{\rm th}}. \tag{6}$$

Solutions for different values of $V_{\rm CE}$ are shown in Fig. 5. Equation (5) yields a curve that increases exponentially, while (6) results in a straight line at lower temperatures if β_f is held constant, as in the example.

A single intersection corresponding to a unique solution is only observed at $V_{\rm CE} = 3$ V. Even in this case, a second solution will be found at much higher temperatures, but it is likely that the simulator finds the physical meaningful. Increasing $P_{\rm diss}$ by means of $V_{\rm CE}$ pushes the two solutions closer together ($V_{\rm CE} =$ 15 V) until the two curves follow each other smoothly over a wide range of temperatures ($V_{\rm CE} = 22$ V). Further increasing $V_{\rm CE}$ leads to no solution at all. This corresponds to the case of thermal runaway, where no stable operation is found.

The thermal runaway occurs whether β_f is modeled temperature-dependent or not. If it decreases with rising temperature, the slopes of the dashed curves increase with temperature, as shown in Fig. 6. This corresponds to a negative feedback between temperature and I_c introduced by β_f, thus enhancing the space between the two solutions. Thermal runaway is also moved to higher dissipated powers, but it will occur as long as $I_{\rm BE}$ has an exponential thermal dependence, while the slope of $1/\beta_f$ is lower. Taking into account that $R_{\rm th}$ increases with temperature, however, would lead to the opposite result. It enhances the positive feedback, and thereby reduces the stability of the model.

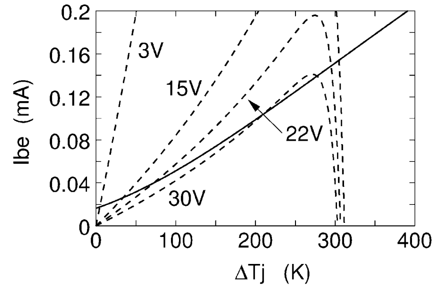

Fig. 6. $I_{\rm BE}$ calculated by (5) (solid line) and (6) (dashed lines), as in Fig. 5, except for $\kappa_\beta = 0.18$.

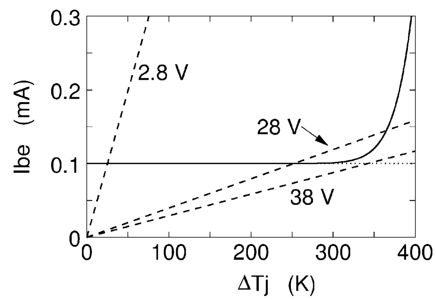

Fig. 7. $I_{\rm be}$ calculated by (8) (solid line), (6) (dashed lines), and constant $I_{\rm be}$ (dotted line). Parameter is $V_{\rm CE}$. The resulting value of $I_{\rm be}$ is given by the intersection of the lines. The parameters are $I_b = 0.1$ mA, $I_{s,\rm bc} = 4 \cdot 10^{-22}$ A, $V_{gr} = 1.8$ V, $\beta_f = 150, \kappa_\beta = 0,$ and $R_{\rm th} = 600$ K/W.

The shape of the dashed curves in Figs. 5 and 6 beyond $\Delta T_j = 250$ K comes from thermal runaway of $I_{\rm bc}$, which is discussed in Section III-C.

C. Reverse Thermal Runaway

The sharp decrease seen in the dashed curves of Figs. 5 and 6 results from the negative contribution of $I_{\rm bc}$ in (6). Ill-chosen parameters for the normally reverse-biased base–collector diode can lead to a second solution much closer to the one at low temperatures than expected. In extreme cases, the base–collector diode can show a thermal runaway effect even though it is reverse biased.

The influence of $I_{\rm bc}$ runaway is also visible in the case of constant base current I_b. I_b is given by the sum of the base–emitter current $I_{\rm be}$ and the reverse base–collector current $I_{\rm bc}$. The currents at the base node are given by

$$I_b = I_{\rm be} + I_{\rm bc} \tag{7}$$

$$\rightsquigarrow I_{\rm be} \approx I_b + I_{s,\rm bc} e^{(V_{gr}/V_{\rm th0} - V_{gr}/V_{\rm th})}. \tag{8}$$

The second equation to be fulfilled is again (6), as in the forward case. Solutions are shown in Fig. 7 as functions of ΔT_j. The solid line shows the solution of (8), the dashed lines show solutions of (6) for different values of $V_{\rm CE}$. While all parameters are reasonably chosen, it can be observed that, due to $I_{\rm bc}$ runaway, multiple solutions are possible around $V_{\rm CE} = 28$ V, and the curves have no intersection at all at higher voltages, as shown for $V_{\rm CE} = 38$ V.

D. Discussion

The previous sections showed evidence that in addition to the well-known thermal runaway effect, the reverse-biased base–collector diode can also show a thermal runaway that can possibly lead to convergence problems.

The first question that arises is: How relevant is the effect? It obviously does not play a role in real transistors. Also in Fig. 7, only very high voltages (beyond 28 V) seem to be dangerous. However, this is only one order of magnitude away when simulating at 3 V. GaAs HBTs can have collector–emitter breakdown voltages well above this value. Common 3-V HBTs can also often be operated up to 12 V, and then the problem becomes quite close. On the other hand, the parameter V_{gr} used in the examples was chosen much too high to make the effect easy visible. However, since the influence of V_{gr} is disguised in the normal active region, one usually is not aware of the danger that is posed by the fact that V_{gr} is found in the argument of an exponential function. One could argue (see the discussion in [5]) that with a nonphysical parameter set, the model cannot be expected to converge. While this is generally true, robustness of the model against parameter variations is crucial for practical use. Since the simulation is based on an iterative search for the solution, the model needs to be stable and defined for all values of voltages. It is for the same reason as it is necessary to limit the exponential function in the diode definition to prevent numerical overflow even though only the range of a few milliamperes is of interest.

What can be done to remove the possible sources of numerical instabilities? The only way to get rid of the multiple solutions is to formulate the whole model with linear equations only in order to get straight lines in Figs. 5 and 6, which intersect only at one point. This is, of course, not desirable since it does not reflect the device physics. This example also shows a general problem with model modifications that are meant to enhance numerical stability: it is, in general, not possible to tell *a priori* that it will not disturb model accuracy. A means that is sometimes applied is the definition of maximum currents $I_{\max}$ or temperatures $T_{\max}$ that can not be exceeded. Both would lead to $I_{\rm be}$ reaching a constant in Figs. 5 and 6 instead of increasing steadily. Under certain circumstances, the thermal runaway can be prevented by introducing those bounds. However, defining, for example, a maximal $I_{\rm be}$ of 0.12 mA in Fig. 5 will not yield a solution for $V_{\rm CE} = 30$ V. The second problem, that of multiple solutions, will also not be improved, as pointed out already by Maas [4]. Defining maximum values can introduce a third solution or reduce the spacing between the desired solution and the numerical unwanted one. An enhancement in numerical stability, therefore, cannot be expected from these measures.

The interaction between base–collector and base–emitter diode runaway is special to HBTs. Since it brings the normally negligible reverse base–collector current into sight, it can be advantageous in this case in the author's opinion, to limit the maximum reverse current. This can be achieved by defining maximum temperatures so that every reverse current $I_s e^{(V_g/V_{\rm th0} - V_g/V_{\rm th})}$ is limited individually to a value several orders of magnitude above its value at room temperature. While this will prevent the reverse thermal runaway by keeping the reverse current very small, it will not affect the operation if the diode is biased in a forward direction. However, it is necessary for the simulation software to inform the user if it converges to the nonphysical solution.

What remains is to be aware of the problem when extracting parameters and simulating. A physical meaningful set of parameters will yield good convergence under realistic operation conditions. The real-world measures of emitter and base ballasting—feedback by emitter or base resistance R_e or R_b—also enhances convergence.

IV. Conclusion

The fact that the diode saturation currents increase with temperature poses the threat of multiple solutions and positive feedback effects. These can lead to numerical instability, which, in the latter case, correspond to thermal runaway in real devices. Since the physical nature of the devices is reflected in these numerical problems, no simple refinement of model functions can solve the issue completely. However, realistic circuit setup and appropriate model parameters should keep the simulation safe in almost all cases.

Acknowledgment

The author would like to thank the Material and Process Technology Departments, Ferdinand-Braun-Institut für Höchstfrequenztechnik (FBH), Berlin, Germany, for providing the HBTs, S. Schulz, FBH, for performing measurements, and Dr. W. Heinrich, FBH, for helpful discussions and continuous encouragement. The reviewer's recommendations were most helpful to improve the quality of this paper's manuscript, which is also gratefully acknowledged by the author.

References

[1] C. C. McAndrew, "Practical modeling for circuit simulation," *IEEE J. Solid-State Circuits*, vol. 33, pp. 439–448, Mar. 1998.

[2] S. Maas, "Fixing the Curtice FET model," *Microwave J.*, vol. 45, no. 3, pp. 68–80, Mar. 2002.

[3] K. Lu and C. M. Snowden, "Analysis of thermal instability in multi-finger power AlGaAs/GaAs HBT's," *IEEE Trans. Electron Devices*, vol. 43, pp. 1799–1805, Nov. 1996.

[4] S. Maas, "Ill conditioning in self-heating FET models," *IEEE Microwave Wireless Comp. Lett.*, vol. 12, pp. 88–89, Mar. 2002.

[5] A. E. Parker and S. Maas, "Comments on 'Ill conditioning in self-heating FET models'," *IEEE Microwave Wireless Comp. Lett.*, vol. 12, pp. 351–352, Sept. 2002.

[6] Q. M. Zhang, H. Hu, J. Sitch, R. K. Surridge, and J. M. Xu, "A new large signal HBT model," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 2001–2009, Nov. 1996.

[7] C.-J. Wei, J. C. M. Hwang, W.-J. Ho, and J. A. Higgins, "Large-signal modeling of self-heating, collector transit-time, and RF-breakdown effects in power HBT's," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 2641–2647, Dec. 1996.

[8] M. Rudolph, R. Doerner, K. Beilenhoff, and P. Heymann, "Scalable GaInP/GaAs HBT large-signal model," *IEEE Trans. Microwave Theory Tech.*, vol. 48, pp. 2370–2376, Dec. 2000.

[9] M. Iwamoto, P. M. Asbeck, T. S. Low, C. P. Hutchinson, J. B. Scott, A. Cognata, X. Qin, L. H. Camnitz, and D. C. D'Avanzo, "Linearity characteristics of GaAs HBT's and the influence of collector design," *IEEE Trans. Microwave Theory Tech.*, vol. 48, pp. 2377–2388, Dec. 2000.

[10] M. Rudolph, R. Doerner, K. Beilenhoff, and P. Heymann, "Unified model for collector charge in heterojunction bipolar transistors," *IEEE Trans. Microwave Theory Tech.*, vol. 50, pp. 1747–1751, July 2002.

Matthias Rudolph (M'99) was born in Stuttgart, Germany, in 1969. He received the Dipl.-Ing. degree in electrical engineering from the Technische Universität Berlin, Berlin, Germany, in 1996, and the Dr.-Ing. degree from the Technische Universität Darmstadt, Darmstadt, Germany, in 2001.

He currently is a Research Scientist with the Ferdinand-Braun-Institut für Höchstfrequenztechnik (FBH), Berlin, Germany. His research focuses on characterization and modeling of FETs and HBTs and on the design of broad-band monolithic microwave integrated circuits (MMICs) for optical communication applications.