Efficient Statistical BJT Modeling, Why β is More Than I_c/I_b

Colin C. McAndrew, John Bates, Richard T. Ida and Patrick Drennan

Motorola Inc., 2100 East Elliot Road, Tempe AZ 85284

ABSTRACT: This paper presents a new, efficient, accurate method for statistical BJT modeling, using backward propagation of variance (BPV). The method uses process control (PC) data to infer variations in process parameters, including geometry variations, and takes only minutes to run on an engineering workstation.

1. Introduction

Statistical SPICE modeling and simulation are necessary for the design of ICs in modern bipolar and BiCMOS technologies. Unfortunately, existing approaches to statistical BJT modeling have significant drawbacks. Simple perturbation of SPICE model parameters gives no guarantee of accurate modeling of the distributions of important electrical performances, and does not account for the interdependence of the SPICE model parameters. Extracting models from corner lots is expensive and time consuming, does not provide distributional models (only case or skew files), and is inherently approximate as corner lots are only estimates of true process limits and are themselves subject to statistical manufacturing variations. Approaches that use numerical data fitting or Principal Components are inefficient as they are based on extractions of SPICE model parameters from a large number of devices, and as they are not physically based they cannot be reliably used for extrapolations to different device layouts, or modified for new technologies.

In addition, as far as we are aware, previously published BJT statistical models do not explicitly account for the fact that emitter size variation is a major contributor to the variability in BJT behavior. We will show here that this is necessary to understand BJT statistical variations properly.

In this paper we present a new method for statistical BJT modeling based on process control (PC) data, process and geometry level modeling, sensitivity analysis, and propagation of variance. The approach has a solid theoretical foundation, and allows accurate modeling of BJT electrical behavior with a minimum of characterization effort.

The analyses we present are for a doubly diffused vertical NPN transistor. Different formulations are necessary for polysilicon emitter transistors, lateral PNPs, etc., however the principles are the same. We assume statistical variations are characterization by normal distributions.

2. Statistical Relationship Between β, I_c and I_b

By definition

$$\beta = I_c/I_b \qquad (1)$$

therefore it appears that there is no additional information in β than in I_c and I_b. Sensitivity analysis of Eq. (1) gives

$$\delta\beta/\beta = \delta I_c/I_c - \delta I_b/I_b \qquad (2)$$

and applying propagation of variance yields

$$\sigma^2_{\delta\beta/\beta} = \sigma^2_{\delta I_c/I_c} + \sigma^2_{\delta I_b/I_b}. \qquad (3)$$

Process control (PC) data for one manufacturing process gives $\sigma_{\delta I_c/I_c} = 0.142$ and $\sigma_{\delta I_b/I_b} = 0.194$, which would seem to imply $\sigma_{\delta\beta/\beta} = 0.240$. The measured value was 0.098. Why is there such a large discrepancy? We will show that this is because of underlying statistical correlations not included in this simple analysis, and present a method that allows this apparently incongruous statistical data to be modeled accurately.

3. Types of Statistical Models

The term "statistical models" is generally used without being explicit about what type of statistical modeling is being addressed, for case simulation or for a Monte Carlo type simulation. In addition, a particular case file is generally only applicable to a specific measure of circuit performance for a specific circuit. Therefore we differentiate between 3 types of statistical models:

Distributional models characterized by means and variances, suitable for Monte Carlo type simulation;

Specific case models that give the most extreme values of specific measure of circuit performance for a specific circuit for a specific probability of being manufactured; and

Generic case models that give specified variations in key device electrical performances. These are the case files commonly used for IC design, although they cannot guarantee accurate modeling of the variation of all measures of circuit performance for all circuits.

We present techniques to determine each of these types of statistical models.

4. Basis for Statistical Modeling

IC manufacturing variations can be viewed as being manifest at 4 levels, as Table 1 shows. Clearly, statistical modeling is best done at the level of process parameters. This approach was pioneered in [1].

Level	Examples	No.	Comment
process inputs	doses, energies times, temperatures	100's	uncorrelated
process parameters	ρ_{sbe} Δ J_{bei} N_{epi} T_{ox} N_{sub} V_{fb} Δ_L	10's	nearly uncorrelated
SPICE parameters	I_S B_F C_{JE} V_{AF} T_{ox} N_{sub} V_{to} Δ_L	100's to 1000's	highly correlated
device/circuit performances	I_c I_b β g_o V_{th0} τ_p V_{OL} V_{OS} ϕ_m		

Table 1. 4 Level Model of Statistical Variations

0-7803-3916-9/97/$5.00 ©1997 IEEE

For vertical NPNs the key process parameters are the pinched base sheet resistance ρ_{sbe}, the variation in emitter size Δ, and the relative ideal base current density J_{bei}, which depends on lifetimes. Other process parameters such as the collector epi doping and the substrate doping also affect BJT behavior, and are detailed in [1]. For reasons of simplicity and space we only consider the above 3 process parameters, but that does not affect the generality of our theoretical formulation and our techniques.

Because Δ is a key statistical variable that controls BJT behavior ***statistical BJT modeling must be based on both process and geometry level models.*** Unlike SPICE MOSFET models, BJT models and model parameters are commonly formulated for a single geometry device, therefore for statistical modeling the SPICE BJT model parameters must be formulated as functions of process parameters and geometry. Example mappings that we use are

$$A_e = (L_e + \Delta)(W_e + \Delta) \tag{4}$$

$$P_e = 2(L_e + W_e + 2\Delta) \tag{5}$$

$$I_S = \rho_{sbe}(I_{SA}A_e + I_{SP}P_e + I_{SC}) \tag{6}$$

$$\text{(SGP) } B_F = \frac{\rho_{sbe}(I_{SA}A_e + I_{SP}P_e + I_{SC})}{J_{bei}(I_{BEIA}A_e + I_{BEIP}P_e + I_{BEIC})} \tag{7}$$

$$\text{(VBIC) } I_{BEI} = J_{bei}(I_{BEIA}A_e + I_{BEIP}P_e + I_{BEIC}) \tag{8}$$

where L_e and W_e are the designed emitter length and width, respectively. These mappings include area A_e, perimeter P_e, and fixed (corner) components. As noted above, these mappings depend on device structure and must be formulated for each type of device to be modeled.

In [1] the mappings from the process parameters **p** to the SPICE parameters **s** were determined from theoretical analyses and TCAD simulations. The **p** were then directly extracted from wafer probing, and the statistical distributions of **p** were used to simulate distributions in electrical performances **e** using process level models to map **p** to **s** and SPICE simulations to map **s** to **e**. This procedure can be considered to be a forward propagation of variance.

However, as Figure 1 shows, because the mappings from **p** to **s** and from **s** to **e** are approximate this approach does not guarantee that the distributions of **e** are modeled accurately, which is the real goal of statistical modeling. In addition, the mappings can be different for different models, e.g. the SPICE Gummel-Poon (SGP) model [2,3] and VBIC [4], so feeding the same **p** distributions into different models gives different predicted distributions for **e**.

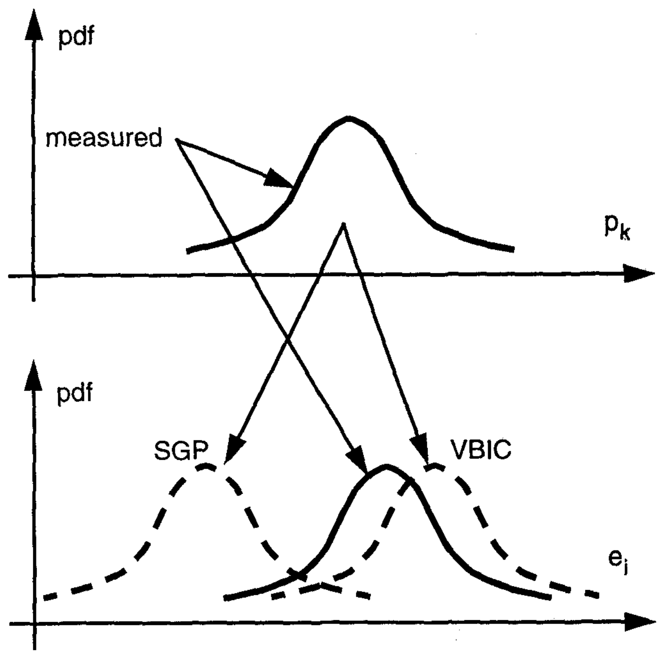

Figure 1. Forward Propagation of Variance

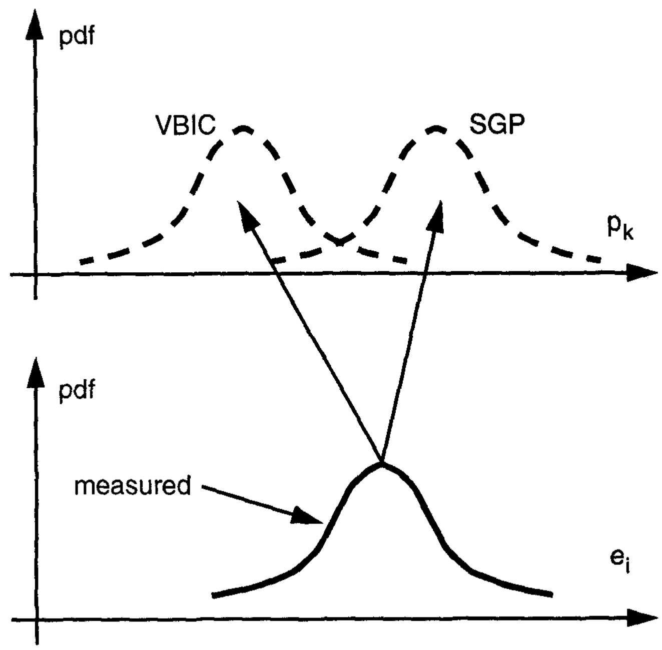

Figure 2. Backward Propagation of Variance

5. Distributional Statistical Modeling

As Figure 2 shows, what is desired for accurate statistical modeling is to be able to calculate a distribution for **p** so that **e** are modeled accurately. Our technique for this is to apply backward propagation of variance (BPV). Sensitivity analysis gives

$$\delta e_i = \sum_k \frac{\partial e_i}{\partial p_k} \delta p_k \tag{9}$$

as the variation in one element e_i of the vector **e** caused by perturbations δp_k in the vector **p**. Using propagation of variance gives

$$\sigma^2_{\delta e_i} = \sum_k \left(\frac{\partial e_i}{\partial p_k}\right)^2 \sigma^2_{\delta p_k}. \tag{10}$$

Eq. (10) defines a set of linear equations that allow the variances of **p** to be directly calculated from the variances of **e**. The sensitivities in the matrix equations (10) are calculated by differencing.

For our procedure to work, **e** must be mathematically observable in **p**, but not necessarily with a one-to-one correspondence. This is equivalent to the matrix of squared sensitivities in Eq. (10) being mathematically well conditioned, which is simple to check. Here, I_c, I_b, and β, measured at some specific bias, make ρ_{sbe}, Δ, and J_{bei} observable. Additional or different PC data may be required for distributional characterization of other devices, e.g. output conductance g_o for lateral PNPs.

Our algorithm for distributional statistical modeling is:

1. derive process and geometry level mappings for the device being statistically modeled
2. define PC measurements **e** that make the process parameters observable, verify this using Eq. (10)
3. extract SPICE model parameters and parameters of the **p** to **s** mappings from "reasonable" data (this does not have to be exactly "nominal")
4. obtain distributional information for **e**
5. calculate the "typical" case **p** by optimizing to fit the mean of **e**, see Section 7 below
6. determine the sensitivities in Eq. (10) by perturbing **p** and simulating the changes in **e**
7. solve Eq. (10) for the variances of **p**

The variations and sensitivities are normalized where appropriate, generally for all quantities except lateral geometry variations and MOSFET V_{th}.

The exact equations solved in the 3 parameter case here are,

$$\begin{bmatrix} \sigma^2_{\frac{\delta I_c}{I_c}} \\ \sigma^2_{\frac{\delta I_b}{I_b}} \\ \sigma^2_{\frac{\delta\beta}{\beta}} \end{bmatrix} = \begin{bmatrix} \left(\frac{\rho_{sbe}}{I_c}\frac{\partial I_c}{\partial \rho_{sbe}}\right)^2 & \left(\frac{J_{bei}}{I_c}\frac{\partial I_c}{\partial J_{bei}}\right)^2 & \left(\frac{1}{I_c}\frac{\partial I_c}{\partial \Delta}\right)^2 \\ \left(\frac{\rho_{sbe}}{I_b}\frac{\partial I_b}{\partial \rho_{sbe}}\right)^2 & \left(\frac{J_{bei}}{I_b}\frac{\partial I_b}{\partial J_{bei}}\right)^2 & \left(\frac{1}{I_b}\frac{\partial I_b}{\partial \Delta}\right)^2 \\ \left(\frac{\rho_{sbe}}{\beta}\frac{\partial \beta}{\partial \rho_{sbe}}\right)^2 & \left(\frac{J_{bei}}{\beta}\frac{\partial \beta}{\partial J_{bei}}\right)^2 & \left(\frac{1}{\beta}\frac{\partial \beta}{\partial \Delta}\right)^2 \end{bmatrix} \begin{bmatrix} \sigma^2_{\frac{\delta\rho_{sbe}}{\rho_{sbe}}} \\ \sigma^2_{\frac{\delta J_{bei}}{J_{bei}}} \\ \sigma^2_{\delta\Delta} \end{bmatrix} \quad (11)$$

where variations in all but Δ have been normalized. If I_c and I_b are measured in the ideal region, and we only consider the area component of Eqs. (4) through (8), then

$$\begin{bmatrix} \sigma^2_{\delta I_c/I_c} \\ \sigma^2_{\delta I_b/I_b} \\ \sigma^2_{\delta\beta/\beta} \end{bmatrix} \approx \begin{bmatrix} 1 & 0 & \left(\frac{L_e + W_e + 2\Delta}{(L_e+\Delta)(W_e+\Delta)}\right)^2 \\ 0 & 1 & \left(\frac{L_e + W_e + 2\Delta}{(L_e+\Delta)(W_e+\Delta)}\right)^2 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \sigma^2_{\delta\rho_{sbe}/\rho_{sbe}} \\ \sigma^2_{\delta J_{bei}/J_{bei}} \\ \sigma^2_{\delta\Delta} \end{bmatrix} \quad (12)$$

which is clearly non-singular and well conditioned.

The reason for the apparent discrepancy between the PC data in Section 2 is now clear. Variations in I_c are caused by variations in ρ_{sbe} and Δ, variations in I_b are caused by variations in J_{bei} and Δ, yet variations in β are caused by variations in ρ_{sbe} and J_{bei}, there is a cancellation of effective emitter size variations because an increase in the emitter size increases both I_c and I_b, but does not greatly affect β. From the data of Section 2 the standard deviations of ρ_{sbe}, J_{bei}, and Δ are 5.38%, 6.38%, and 0.17μm. The variation in β is less than expected from the

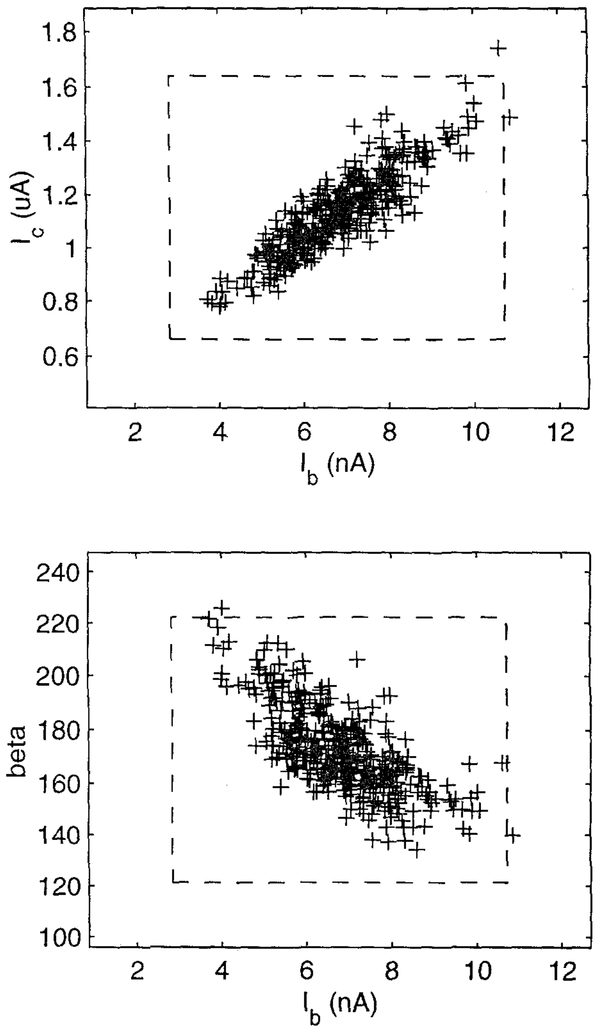

Figure 3. Modeled (+ Monte Carlo) and Fab Data (– – box)

simple analysis of Section 2 because of the correlation between I_c and I_b through Δ,

$$\sigma^2_{\delta\beta/\beta} < \sigma^2_{\delta I_c/I_c} + \sigma^2_{\delta I_b/I_b}. \quad (13)$$

The closer $\sigma^2_{\delta\beta/\beta}$ is to $\sigma^2_{\delta I_c/I_c} + \sigma^2_{\delta I_b/I_b}$ the less device behavior variation is determined by emitter size variation.

Statistically speaking therefore, there is additional information in the variance of β beyond that contained in the variances of I_c and I_b. And the statistics of the variations in effective emitter size can be obtained from statistical information for a single geometry device. We are not aware that this has been pointed out previously.

Figure 3 shows Monte Carlo simulations from our statistical models (which included a full geometry dependence, not just an area component), along with $\pm 3\sigma$ limits from manufacturing PC data. The accuracy of the statistical modeling is clear. Note that both the mean and the variance of the PC data are modeled well.

6. Specific Case Statistical Modeling

The goal of specific case statistical modeling is as follows (see Figure 4). For a specified probability of being manufactured (i.e. n·σ in **p**-space), what are the extreme values possible for each component of **e**? This is the

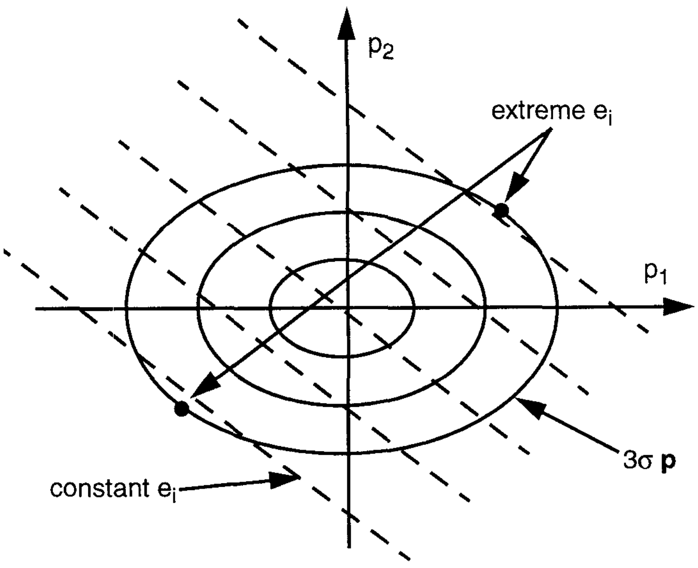

Figure 4. Constant e_i in **p**-space

inverse on the statistical design problem formulated and solved in [5]. Mathematically this is equivalent to

$$\max_{\mathbf{p}} \ \delta e_i = \sum_k \frac{\partial e_i}{\partial p_k} \delta p_k \ \text{subject to} \ \sum_j \left(\frac{\delta p_j}{\sigma_{\delta p_j}} \right)^2 = n^2 \qquad (14)$$

and using the method of Lagrange multipliers gives

$$\delta p_j = \pm \frac{n \sigma^2_{\delta p_j} (\partial e_i / \partial p_j)}{\sqrt{\sum_k \sigma^2_{\delta p_k} (\partial e_i / \partial p_k)^2}}. \qquad (15)$$

The procedure is based on linearization, which is an approximation. Table 2 shows differences between linearly predicted and simulated extreme case values from our procedure. The close agreement shows that the approximation we use is reasonable.

case →	$-3\sigma\ I_c$	$+3\sigma\ I_c$	$-3\sigma\ I_b$	$+3\sigma\ I_b$	$-3\sigma\ \beta$	$+3\sigma\ \beta$
$\%I_c$ error	-2.19	-1.07	0.42	0.05	2.73	3.05
$\%I_b$ error	0.68	0.06	-5.16	-1.86	-0.92	-3.58
$\%\beta$ error	1.31	0.41	0.09	0.15	-0.69	0.04

Table 2. Linearization Errors

Note that Eq. (15) can be extended to account for correlations between process parameters, and that the variances used are calculated as in Section 5.

7. Generic Case Statistical Models

Generic case models are intended to model specified variation levels, typically $\pm 3\sigma$, in key device electrical parameters. They do ***not*** guarantee $\pm 3\sigma$ modeling of every measure of circuit performance e_i for every circuit, however they can provide a useful means of estimating circuit sensitivities to process variations, help determine if a circuit is near a "cliff," and for some e_i can provide reasonable estimates of $\pm 3\sigma$ manufacturing variations.

Generic case files generation is commonly done by introducing $\pm 3\sigma$ variations in process parameters. However, as Figure 1 shows this does not guarantee accurate $\pm 3\sigma$ modeling of key device electrical parameters.

We determine generic case models by using nonlinear least squares optimization to compute the values of process parameters that give $\pm 3\sigma$ variations in key device electrical parameters. This ***guarantees*** accurate modeling of what is important for design. In Figure 4 the $\pm 3\sigma$ PC limits for I_c and β exactly match those simulated by our generic case files. Note that this optimization can also be done to "tune" a model to the mean of the PC distribution. This efficiently allows generation of a "typical" case file from a model generated from "reasonable" silicon, which is very useful as there is no such thing a exactly "typical" silicon.

Note that for case files, unlike for variances, there is no additional information in β than is found in I_c and I_b. Generic case files cannot force $\pm 3\sigma$ values in all of these quantities. This is intuitively obvious, and follows mathematically from the perturbational form of Eq. (12),

$$\begin{bmatrix} \delta I_c / I_c \\ \delta I_b / I_b \\ \delta \beta / \beta \end{bmatrix} \approx \begin{bmatrix} 1 & 0 & \dfrac{L_e + W_e + 2\Delta}{(L_e + \Delta)(W_e + \Delta)} \\ 0 & 1 & \dfrac{L_e + W_e + 2\Delta}{(L_e + \Delta)(W_e + \Delta)} \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \delta \rho_{sbe} / \rho_{sbe} \\ \delta J_{bei} / J_{bei} \\ \delta \Delta \end{bmatrix} \qquad (16)$$

which is clearly singular. Squaring the sensitivities for propagation of variance analysis removes the singularity!

8. Conclusions

In this paper we have presented new techniques to generate distributional statistical models, specific case statistical models, and generic case statistical models. The techniques are based on process and geometry level modeling, which can be added to the SGP and VBIC models. Application of our technique has resolved an apparent discrepancy between variances observed in PC data, and underscored the need to include emitter size variations in statistical BJT modeling. This cannot be done without having geometry dependent models, and shows that it is not possible to do statistical BJT modeling at the level of SPICE parameters for a single geometry device.

Our techniques are in use at Motorola for both BJT and MOSFET statistical modeling.

References

[1] W. F. Davis and R. T. Ida, pp. 262-265, Proc. 1989 BCTM.

[2] H. K. Gummel and H. C. Poon, Bell System Tech. J., vol. 49, pp. 827-852, May 1970.

[3] L. W. Nagel, Memorandum ERL-M520, University of California, Berkeley, May 1975.

[4] C. C. McAndrew, J. Seitchik, D. Bowers, M. Dunn, M. Foisy, I. Getreu, S. Moinian, J. Parker, P. van Wijnen, and L. Wagner, Proc. 1995 BCTM, pp. 170-177.

[5] S. R. Nassif, in *Statistical Approach to VLSI*, S. W. Director, W. Maly and A. Strowjas Eds., 1994.