Statistical Circuit Modeling

Colin C. McAndrew

Motorola, Inc.,
2100 East Elliot Road, Tempe AZ, 85284 U.S.A.

Abstract

This paper reviews common approaches to statistical circuit modeling, and details their limitations. A simple, efficient, and generic approach to statistical circuit modeling is presented. Backward propagation of variance (BPV) is used to *guarantee* that the statistical circuit models match variations in key device performances. Examples are provided for MOSFETs and BJTs.

1. Introduction

This paper reviews some common approaches to statistical modeling, and details their limitations. In particular, problems with using process and device simulation (TCAD) for statistical circuit modeling will be highlighted. An efficient and accurate method based on backward propagation of variance (BPV) is presented. This provides a simple yet mathematically rigorous foundation for statistical modeling.

The scope is statistical models for circuit simulation, in SPICE type simulators [1]. Two types of statistical models will be covered: **distributional statistical models**, characterized by means and variances and used for Monte Carlo type simulations; and **generic case statistical models** that give specified variations in key device electrical performances. The latter are the normal "case" files used for circuit simulation. There exist different "worst" and "best" case models for every distinct measure of circuit performance, circuit topology, and selection of device geometries. It is possible to determine "case" statistical models for each specific performance/topology/geometry combination [2], but such specific case models are at present not in common use and will not be treated here.

MOSFET and BJT circuit simulation (compact) models have many parameters, but a small number of physical "process" parameters $\boldsymbol{p}$ control device electrical behaviors $\boldsymbol{e}$. $\boldsymbol{p}$ can include oxide thickness, lateral geometry variation, junction depth, doping density (or sheet resistance), etc. The process parameter level is the best basis for statistical circuit modeling. MOSFET models are based on process parameters, however this is not true for BJT models, and mappings from process (and geometry) parameters to SPICE model parameters must be developed as a basis for BJT statistical modeling, see [3].

The $\boldsymbol{p}$ are separated into two categories, those for which an absolute variation is more natural, e.g. flatband voltage δV_{fb} and lateral geometry variation $\delta\Delta_L$, and those for which a relative variation is more natural, e.g. oxide thickness $\delta T_{ox}/T_{ox}$ and most other parameters, including vertical geometry variations.

The statistical modeling techniques presented here are based on physical process and geometry parameter level models, and use PC (in-line process control) data statistics to characterize the process parameter level statistical models. Note that the statistics of $\boldsymbol{p}$ are *not* assumed to be directly measured as part of the PC data, but are computed base on BPV. This is a key difference between the techniques detailed here and those reported previously.

2. Existing statistical modeling approaches

2.1. Numerical data fitting

Numerical modeling techniques have been applied to statistical modeling, in the form of principle components analysis [4] and response surface modeling [5]. These approaches provide distributional statistical models and could provide case statistical models, although no definite method is presented to generate case files.

The main drawback of numerical approaches is the significant amount of effort required to construct the numerical models. E.g. [5] required SPICE model extractions from 100 die. In addition, process changes are made during the life of a technology, and the confidence that the effects of these changes can be quickly and accurately reflected in models is significantly greater for physically based models than for purely numerical models.

2.2. Forward propagation of variance

Perhaps the most common method to generate statistical circuit models is by direct measurement of p as part of PC data. This provides distributional statistical models. Generic case statistical models are then defined by $\pm 3\sigma$ variations in p. Because the variations in p propagate through the SPICE models to simulated e this approach may be termed forward propagation of variance (FPV).

There are several significant problems with this approach:

1. the $e(p)$ mappings are only approximate;
2. the $e(p)$ mappings are different for different SPICE models; and
3. different methods for measuring p directly give different values.

Therefore the FPV approach provides no guarantee about the accuracy of the modeling of e, see Fig. 1, and misses the real goal of statistical modeling. Further, introducing $\pm 3\sigma$ variations in p does not generate a known level of perturbation in any e,

$$\delta e_i = \sum_k \frac{\partial e_i}{p_k} \delta p_k \qquad (1)$$

so δe depends both on the sensitivities and the number of elements in p. Adding a variation in some p_k is useless if it does not affect e_i, e.g. Δ_L for a long channel MOSFET, and introducing $\pm 3\sigma$ variations in a large number of p leads to a pessimistic value for e_i (if the sensitivities were 1, $\pm 3\sigma$ variations in n process parameters gives a $\pm 3\sqrt{n}\sigma$ change in e_i).

2.3. Extreme case data

Data from either from split manufacturing lots or from TCAD simulations of manufacturing extremes can be used to extract extreme case SPICE model parameters.

Limitations of this approach are: it provides only generic case files, and not distributional statistical models; the effort to generate the models is directly proportional to number of case files required for circuit simulation; and such models cannot be easily re-targeted after process changes are made.

The most significant drawback is that variations, in either split lots or TCAD simulations, are done in only a small number of the p that control device electrical behavior, and the amount of each variation is only an estimate, because exact statistical data on the process parameters is not known. It is misdirected to spend significant effort to generate models that only roughly approximate the observed variations in device electrical performance.

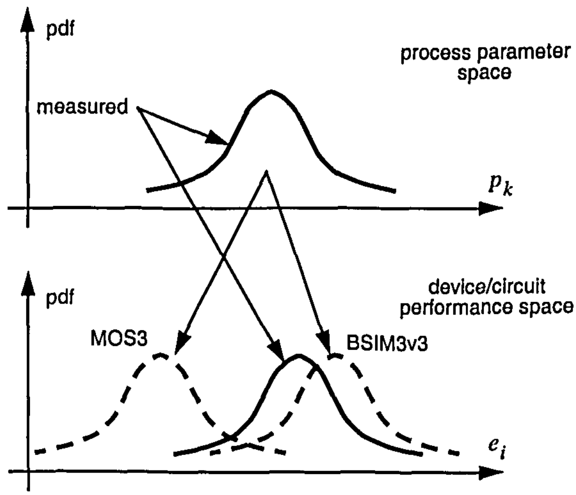

Figure 1: Errors involved in forward propagation of variance.

3. BPV for generic case statistical modeling

IC design requires generic case files that embody $\pm 3\sigma$ variations in measures of circuit performance e_c (speed, power, phase margin, etc.). e_c are determined by the device performances e, threshold voltage, drive current, transconductance g_m, output resistance r_o, etc. The BPV approach recognizes that accurate modeling of the distribution of e is paramount for circuit modeling, and the p should provide an indirect, not direct, method for modeling e properly, see Fig. 2.

The BPV procedure for generating generic case statistical models is:

1. develop process and geometry mappings for the device being modeled
2. extract model parameters from a "representative" wafer
3. define e that make p observable (see the next section for how to do this)
4. obtain distributional information on e
5. define target e for each simulation case based on how they affect circuit performance
6. optimize p to minimize $\|e - e_{target}\|$ where $\|\ldots\|$ is the Euclidean norm

4. BPV for distributional statistical modeling

Applying propagation of variance to the sensitivity relation in eq. (1) gives

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

which defines a set of linear equations that relate the variances of e to the variances of p. The sensitivities in eq. (2) are computed by differencing, and the variances in e are specified, as the distributions of the PC data. Eq. (2) can therefore be directly solved for the variances of the process parameters, hence the name backward propagation of variance, from e to p rather than from p to e, see Fig. 2

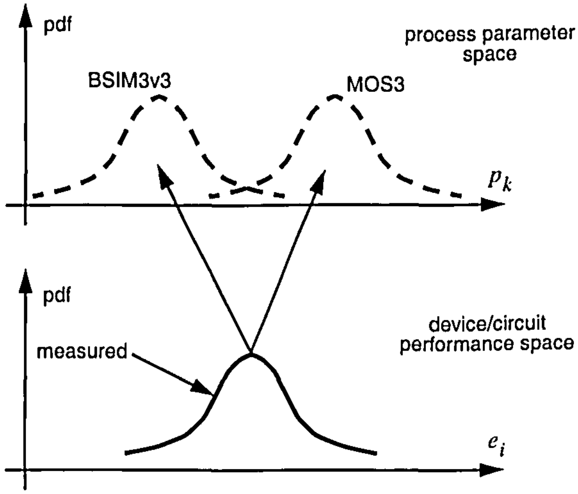

Figure 2: Statistical characterization using backward propagation of variance.

The p must be mathematically observable in e. This means that the matrix of squared sensitivities in eq. (2) must be well conditioned, at the very least nonsingular. Consequently, the e used to statistically characterize the p for each type of device being modeled must be selected to make the p observable. With a physical understanding of the way the p affect device behavior this is generally easy to do, and e suitable for MOSFET and BJT modeling are detailed below. A major feature of the BPV approach is being able to determine whether the selected e are sufficient for statistical modeling.

Some p can be characterized directly, and are thus FPV parameters. Entries in eq. (2) that are FPV are subtracted from the left hand side variances. If the specified variances of the FPV variables are inconsistent with the variances of e (as can happen for T_{ox} for MOSFETs), the left hand side of eq. (2) becomes negative, which is clearly absurd. It is a significant feature of the BPV approach that it can detect such inconsistencies.

Eq. (2) is based on linearization and the assumption that the p and e are normally distributed. These are reasonable approximations. However, slight nonlinearities imply that the differencing used to compute the sensitivities depends on the perturbation range for δp. $\pm 3\sigma$ perturbations are used. Given that the variances of p are not known *a priori* eq. (2) is solved iteratively for a self-consistent set of variances and sensitivities.

Quantities (both p and e) that vary by some ratio rather than by some additive amount, e.g. BJT $\beta = I_c/I_b$ which can vary from say 0.5 to 2.0 times its nominal value, should be transformed logarithmically prior to application of the BPV procedure.

5. MOSFET example

The key p that control the variation of MOSFET electrical behavior are T_{ox}, V_{fb}, channel length reduction Δ_L, channel width variation Δ_W, effective substrate doping N_b, low field mobility μ_0, and a parameter V_{tl} that models the change in threshold voltage with decreasing channel length, and depends on junction depth, profile shape, etc. For most MOSFET models there is a direct correspondence between these p and the SPICE model parameters s. For some s mappings may need to be defined from p, e.g.

$$V_{th0} = V_{fb} + \phi_B + \frac{T_{ox}\sqrt{2q\varepsilon_{Si}N_b}}{\varepsilon_{ox}}\sqrt{\phi_B}, \quad \phi_B = 2\frac{kT}{q}\ln\left(\frac{N_b}{n_i}\right). \tag{3}$$

Here T_{ox} is used directly, by FPV. For brevity, variations in Δ_W and N_b will be ignored. Statistical modeling of V_{fb}, Δ_L, μ_0, and V_{tl} are based on BPV. These p are observable in wide/long and wide/short threshold voltage (V_{tr} and V_{ts}) and saturated drain current (I_{dr} and I_{ds}). If the body effect is included, the statistics of N_b can be characterized.

Fig. 3 shows characteristics of generic case statistical models generated by optimizing p to model $\pm 3\sigma$ limits in the above e, for a PMOS device in a 0.5μm BiCMOS process. Clearly both g_m and r_o track with the cases, although they are not explicitly included as part of e. Note that the model limits for e *exactly* match the $\pm 3\sigma$ fab PC limits.

The BPV equations (with appropriate normalization) are

$$\begin{bmatrix} \sigma^2_{\delta V_{tr}} - \left(T_{ox}\dfrac{\partial V_{tr}}{\partial T_{ox}}\right)^2 \dfrac{\sigma^2_{\delta T_{ox}}}{T_{ox}} \\ \dfrac{\sigma^2_{\delta I_{dr}}}{I_{dr}} - \left(\dfrac{T_{ox}}{I_{dr}}\dfrac{\partial I_{dr}}{\partial T_{ox}}\right)^2 \dfrac{\sigma^2_{\delta T_{ox}}}{T_{ox}} \\ \sigma^2_{\delta V_{ts}} - \left(T_{ox}\dfrac{\partial V_{ts}}{\partial T_{ox}}\right)^2 \dfrac{\sigma^2_{\delta T_{ox}}}{T_{ox}} \\ \dfrac{\sigma^2_{\delta I_{ds}}}{I_{ds}} - \left(\dfrac{T_{ox}}{I_{ds}}\dfrac{\partial I_{ds}}{\partial T_{ox}}\right)^2 \dfrac{\sigma^2_{\delta T_{ox}}}{T_{ox}} \end{bmatrix} = \begin{bmatrix} \left(\dfrac{\partial V_{tr}}{\partial V_{fb}}\right)^2 & \left(\mu_0\dfrac{\partial V_{tr}}{\partial \mu_0}\right)^2 & \left(\dfrac{\partial V_{tr}}{\partial \Delta_L}\right)^2 & \left(\dfrac{\partial V_{tr}}{\partial V_{tl}}\right)^2 \\ \left(\dfrac{1}{I_{dr}}\dfrac{\partial I_{dr}}{\partial V_{fb}}\right)^2 & \left(\dfrac{\mu_0}{I_{dr}}\dfrac{\partial I_{dr}}{\partial \mu_0}\right)^2 & \left(\dfrac{1}{I_{dr}}\dfrac{\partial I_{dr}}{\partial \Delta_L}\right)^2 & \left(\dfrac{1}{I_{dr}}\dfrac{\partial I_{dr}}{\partial V_{tl}}\right)^2 \\ \left(\dfrac{\partial V_{ts}}{\partial V_{fb}}\right)^2 & \left(\mu_0\dfrac{\partial V_{ts}}{\partial \mu_0}\right)^2 & \left(\dfrac{\partial V_{ts}}{\partial \Delta_L}\right)^2 & \left(\dfrac{\partial V_{ts}}{\partial V_{tl}}\right)^2 \\ \left(\dfrac{1}{I_{ds}}\dfrac{\partial I_{ds}}{\partial V_{fb}}\right)^2 & \left(\dfrac{\mu_0}{I_{ds}}\dfrac{\partial I_{ds}}{\partial \mu_0}\right)^2 & \left(\dfrac{1}{I_{ds}}\dfrac{\partial I_{ds}}{\partial \Delta_L}\right)^2 & \left(\dfrac{1}{I_{ds}}\dfrac{\partial I_{ds}}{\partial V_{tl}}\right)^2 \end{bmatrix} \begin{bmatrix} \sigma^2_{\delta V_{fb}} \\ \dfrac{\sigma^2_{\delta \mu_0}}{\mu_0} \\ \sigma^2_{\delta \Delta_L} \\ \sigma^2_{\delta V_{tl}} \end{bmatrix} \tag{4}$$

where the variation caused by T_{ox} is subtracted from the left hand side, as it is an FPV parameter. Because of normalization and the weak dependence of some e_i on p_k many entries of the matrix in eq. (4) are close to either 1 or 0. E.g. in the first row of the coefficient matrix $\partial V_{tr}/\partial V_{fb} \approx 1$ and all the other sensitivities are close to 0, and in the second row the first entry is small, $(\mu_0/I_{dr})\partial I_{dr}/\partial\mu_0 \approx 1$, and the remaining entries are close to 0. (One of the reasons for normalizing quantities for which a relative variation is natural is that it leads to good conditioning, and stable numerical solution, of the BPV equations).

Fig. 4 shows Monte Carlo simulations from distributional statistical models characterized by the BPV approach, for a 0.5μm BiCMOS process. Also shown are generic case files (at the corners of the bounding hexagons, additional cases to the best and worst case files of Fig. 3 are included, to better model the observed correlations in, and distribution of, PC data). Again, note the model limits for e *exactly* match the $\pm 3\sigma$ fab PC limits.

6. BJT example

The key p that control the behavior of vertical NPNs are the effective base doping N_{beff}, effective base width w_b, variation in emitter size Δ, and recombination/generation lifetimes, here incorporated into effective non-ideal and ideal base-emitter current density parameters J_{ben} and J_{bei}. BJT behavior over geometry can be modeled well by including area, perimeter, and corner components for most SPICE model parameters, although

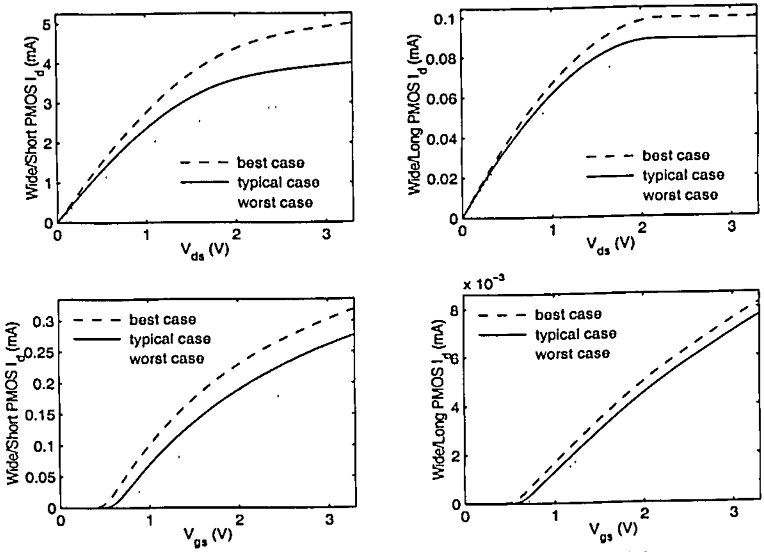

Figure 3: $I_d(V_d)$ and $I_d(V_g)$ variation in generic case statistical models.

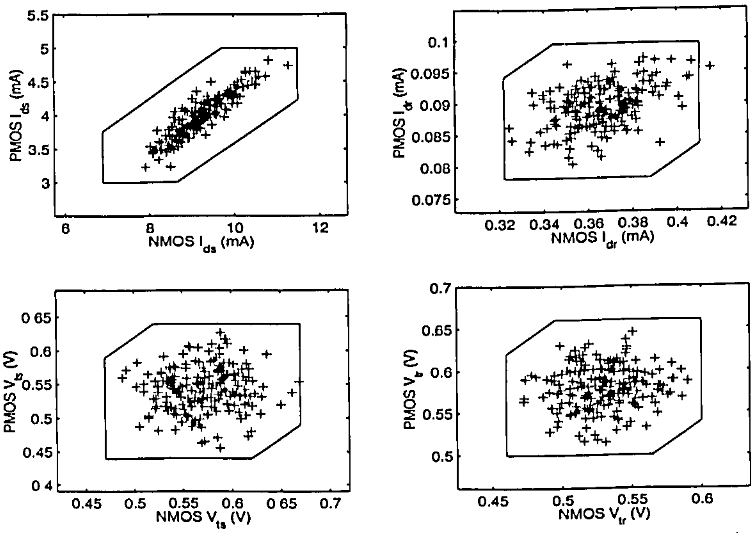

Figure 4: Monte Carlo (+) and generic case (hexagon corners) MOSFET models.

resistance modeling requires more complicated analyses. For a single emitter BJT with emitter width and length of w_e and l_e, the emitter area, perimeter, and the relative pinched base sheet resistance are

$$A_e = (l_e + \Delta)(w_e + \Delta), \quad P_e = 2(l_e + w_e + 2\Delta), \quad \frac{\delta\rho_{sbe}}{\rho_{sbe}} = \frac{1}{(\delta N_{beff}/N_{beff})(\delta w_b/w_b)}. \quad (5)$$

Representative mappings for VBIC parameters are then, following the approach of [3],

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

$$I_{BEN} = (I_{BENA}A_e + I_{BENP}P_e + I_{BENC})\delta J_{ben}/J_{ben} \tag{7}$$

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

$$\frac{1}{V_{EF}} = \frac{(I_{SA}A_e/V_{EFA} + I_{SP}P_e/V_{EFP} + I_{SC}/V_{EFC})}{(I_{SA}A_e + I_{SP}P_e + I_{SC})} \frac{1}{(\delta\rho_{sbe}/\rho_{sbe})\sqrt{\delta N_c/N_c}} \tag{9}$$

$$C_{JE} = (C_{JEA}A_e + C_{JEP}P_e + C_{JEC})/\sqrt{\delta\rho_{sbe}/\rho_{sbe}} \tag{10}$$

$$T_F = \frac{(T_{FA}I_{SA}A_e + T_{FP}I_{SP}P_e + T_{FC}I_{SC})}{(I_{SA}A_e + I_{SP}P_e + I_{SC})}\left(\frac{\delta w_b}{w_b}\right)^2 . \tag{11}$$

Note that not all area, perimeter, and corner components are needed for every parameter for every technology, and the above mappings are asymptotically correct as w_e and l_e increase [6]. In eq. (9) an extra process parameter, the collector doping N_c, is included, but for simplicity is considered fix for the example presented here.

The 5 BJT process parameters can be observed in the following 5 measures of BJT electrical performance e: the collector current I_c measured in the ideal region; current gains β_n and β_i measured in the non-ideal and ideal regions, respectively; extrapolated Early voltage V_A (note that this is significantly different from the Early voltage SPICE model parameter [6]); and the unity-gain cutoff frequency f_T, measured at relatively high current density.

Fig. 5 shows how the p affect the e for BJTs. Note that there is not a one-to-one correspondence between the p and the e. However, because the p are observable in the e, the BPV procedure allows numerically robust characterization of the p. For BJT analyses, in contrast to MOSFETs, there are no FPV parameters.

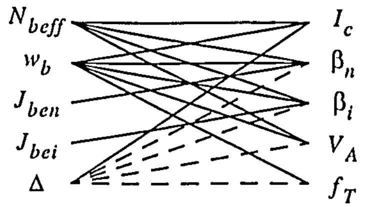

Figure 5: Dependencies between p and e for BJTs (dashed lines are weak dependence).

Fig. 6 shows results from BPV statistical characterization, for a double poly emitter NPN in a 0.5μm BiCMOS technology. Here only best and worst case models are defined, in terms of $\pm 3\sigma$ for each of the measures of device electrical performance e. As with the MOSFET example, the model limits for e ***exactly*** match the $\pm 3\sigma$ fab PC limits.

The expected underlying physical correlations between the e are apparent in Fig. 6. Early voltage is roughly inversely proportional to the collector current, whereas current gain and unity-gain cutoff frequency are roughly proportional to the collector current. These correlations are explicit in the physically based process parameter and geometry mappings given above.

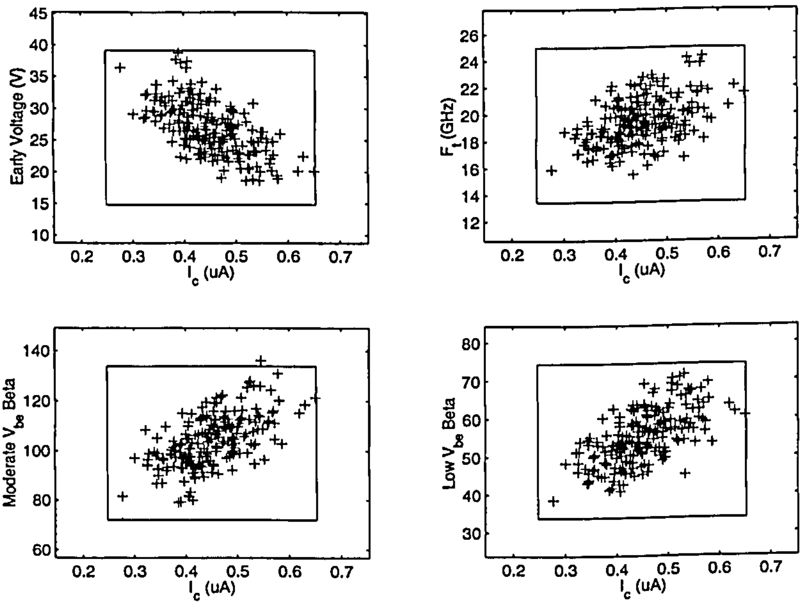

Figure 6: Monte Carlo (+) and generic case (solid lines) BJT models.

7. Conclusions

The BPV approach to statistical circuit modeling is completely generic, and has a rigorous mathematical foundation. At Motorola a single program is used to generate statistical circuit models for all devices, MOSFETs, vertical NPNs, lateral PNPs, etc., even though the models and associated process and geometry mappings are very different. The program runs in several minutes on an engineering workstation, generates both distributional and generic case statistical models, and guarantees that the resulting models accurately embody the measured distributions of device electrical performances.

References

[1] L. W. Nagel, *SPICE2: A computer program to simulate semiconductor circuits*, Memo. no. ERL-520, University of California, Berkeley, 1975.

[2] S. R. Nassif, "Statistical worst-case analysis for integrated circuits" in *Statistical Approach to VLSI*, North-Holland, pp. 233-53, 1994.

[3] W. F. Davis and R. T. Ida, "Statistical IC simulation based on independent wafer extracted process parameters and experimental designs," *Proc. IEEE BCTM*, pp. 262-265, 1989.

[4] J. A. Power *et al*, "Relating statistical MOSFET model parameter variabilities to IC manufacturing process fluctuations enabling realistic worst-case design," *IEEE Trans. Semicond. Manufact.*, vol. 7, pp. 306-318, 1994.

[5] J. A. Power, S. Kelly, E. Griffith, D. Doyle, and M. O'Neill, "Statistical modeling for 0.6μm BiCMOS technology," *Proc. IEEE BCTM*, 24-27, 1997.

[6] C. C. McAndrew, "Practical modeling for circuit simulation," *IEEE J. Solid-State Circuits*, vol. 33, no. 3, pp. 439-448, Mar. 1998.