Unified Statistical Modeling for Circuit Simulation

Colin C. McAndrew and Patrick G. Drennan

Motorola, Tempe, AZ, Colin.McAndrew@motorola.com

ABSTRACT

Accurate statistical simulation and modeling are important for IC design. Different types of statistical simulation require different types of statistical models. In this paper a unified approach to statistical modeling and characterization is presented. Based on physical process parameters and propagation of variance, it allows modeling of process extremes, distributional modeling for Monte Carlo type simulation, and modeling of mismatch.

Keywords: Statistical modeling, SPICE modeling, circuit simulation.

1 PARAMETERS AND MODELING

Parameters are the means whereby a model can be adjusted to represent a specific component type in a particular technology, without having to alter the underlying model equations embedded in a simulator. A parameter is specific to a particular model, and is not an independent, sacrosanct, "physical" entity. Techniques to determine parameters commonly start with a model formulation, generally a very simple and reduced form of a more accurate model, make approximations reasonable for some particular biasing scheme, and define some measurement specification and data analysis algorithm, often quite clever, that leads to a simple and direct calculation of the parameter value. What is overlooked is that the parameter value is then only consistent with the simplified model, and is not representative of a complex model or a "physical" value.

The goal of modeling for circuit simulation is to accurately represent the electrical behaviors of individual components, so that when these components are combined to form a circuit the important electrical behaviors of the circuit can be accurately simulated. The goal of modeling is not to accurately represent "parameter" values and variations.

For example, many methods to determine MOSFET effective electrical channel length $L_{eff} = L - \Delta_L$ have been published. Different methods give different values. If the values of Δ_L from one method are used in the level 3 MOSFET model and the BSIM3v3 [1] model, they will give different results because the models are different.

This means that values for parameters, including their statistical variations, must be determined specifically for a given model, and cannot be calculated generically, independent of a model. This gives rise to a quandary: the primary source of statistical data for semiconductor manufacturing processes is electrical test (ET) data measured continually in fabs on production wafers. Yet it is not feasible, or desirable, to include in ET extraction of all key parameters for all models used for circuit simulation. So how can you do accurate yet generic statistical modeling? An efficient technique for this is presented here.

2 TYPES OF STATISTICAL MODELS

There is no single "statistical model" for circuit simulation. Different statistical simulation techniques are used to analyze circuits, and they require different types of statistical model. There are 3 major types of statistical analyses done at present: process extreme simulations; distributional simulations; and mismatch simulations (which may be combined with a distributional simulation, or may be done separately). These different types of statistical simulations require different sorts of models. The first requires corner model files. The second and third require models characterized by means and variances, of global and local variations, respectively. We present a model suitable for all of these simulation needs, and techniques to characterize the models.

3 FOUNDATION

For a given manufacturing process, a complete set of circuit simulation models, often called SPICE models, includes thousands of parameters S. Many of these, especially for BJTs, are correlated. So it is not desirable to do statistical modeling at the level of model parameters. It either introduces physical and prediction inaccuracies if the correlations are not accounted for, or requires an extremely large effort in both modeling and simulation to account for the correlations.

The electrical behavior of components is controlled by physical characteristics of silicon (or other materials), commonly called process parameters P. These include layer sheet resistances, layer thicknesses, junction depths, and critical dimension variations. A small number of essentially independent process parameters affect a single component type, and so provide the most efficient basis for statistical modeling.

Some types of models, especially of MOSFETs and resistors, are formulated in terms of process parameters. Others, like BJT models as noted above, are not. In this case the model parameters S must be written in the model

files as appropriate functions of P, and perhaps also of geometric size and other layout attributes. Even for MOSFETs, if the threshold voltage V_{TH} rather than flatband voltage V_{FB} is used as a model parameter, then V_{TH} needs to be formulated as a function of V_{FB}, oxide thickness t_{OX}, and effective substrate doping N [2]. The inclusion of V_{FB} as a model parameter for BSIM3v3.2 [1] was of significant benefit for statistical modeling. For BJTs the mappings are formulated as in [3].

Numerical techniques, like principal components, to reduce the size of the parameter space from S have also been reported [4]. These have the advantage that they are generic, and do not require physical understanding to formulate. However they require a large effort to generate as S need to be extracted from a statistically significant sample, give no insight into the physical cause of statistical variation and therefore what to attack in manufacturing to reduce variation, and are neither predictive nor portable. In addition, if a process changes you have to start the statistical characterization from the beginning. The procedure described below allows retuning of statistical parameters within minutes when new information of process variations is available, primarily because it has a basis in physical models.

For a particular parameter P its variation is modeled as

$$P = P_0 + \delta P_G + \delta P_L(g) \tag{1}$$

where P_0 is the mean value of the parameter, δP_G its global variation, and $\delta P_L(g)$ its local variation [5]. g are device attributes, including geometry, orientation, etc. The global variation is sometimes termed the interdie component, is correlated across a die, and is generally independent of geometry. The local variation is often termed the intradie or mismatch variation, is uncorrelated across a die, and strongly depends on geometry, increasing as geometry decreases.

4 COMMENTS ON MODELING

One very important observation needs to be made from (1). Analysis of variance gives

$$\sigma_P^2 = \sigma_{PG}^2 + \sigma_{PL}^2(g) \tag{2}$$

as the total variance of P. Conventionally ET data, which embody σ_P, are assumed to characterize σ_{PG}, and σ_{PL} are characterized separately from mismatch measurements. This is reasonable if $\sigma_{PL} << \sigma_{PG}$. However, σ_{PL} increases as geometry decreases, and for modern technologies can no longer be considered small compared to σ_{PG} for some parameters. For example, across chip line width variation (ACLV) is becoming a dominant cause of total variation. This has several ramifications.

First, for ET data this phenomenon is rarely, if ever, taken into account. The ET data depend on the sizes of the test structures that are measured, but are assumed to be independent of geometry. Because σ_{PL} varies reciprocally with geometry, σ_P increases as geometry decreases.

Second, simulations based only on global process parameter variations, i.e. non-mismatch analyses, do not account for the fact that σ_P varies with geometry, and can therefore be inaccurate.

Third, when mismatch analyses are included with global variation analysis, the effect of mismatch can be double counted because it is explicit in σ_{PL} and implicit in σ_{PG}. This gives inaccurate (too large) estimates of the standard deviation σ_E of circuit electrical performances E.

Fourth, if σ_{PL} is significant compared to σ_{PG}, then any global statistical simulation (based on σ_{PG} assumed to be equal to σ_P) on a circuit with multiple components can predict too large a σ_E because much of the variation in P is uncorrelated between components, whereas in σ_{PG} it is explicitly assumed to be correlated.

Fifth, and last, if statistical simulation is done via σ_{PG} and correlations between devices, rather than directly through independent $\delta P_L(g)$, then again the geometry dependence of σ_P is not accounted for, because σ_P is then implicitly assumed to be constant.

There are other disadvantages to using correlations between P, rather than independent values for $\delta P_L(g)$, to model mismatch. To model local variations between n different components using correlations requires $n(n+1)/2$ variances and covariances, compared to n variances if they are treated independently. More important, the correlation coefficient between two devices with attributes g_1 and g_2 is

$$\rho = \frac{\sigma_{PG}^2}{\sqrt{(\sigma_{PG}^2 + \sigma_{PL}^2(g_1))(\sigma_{PG}^2 + \sigma_{PL}^2(g_2))}}. \tag{3}$$

If the local component dominates so that in modeling $\sigma_{PG} = 0$, then the correlation is zero, and specifying local variations via σ_{PG} and ρ does not work.

5 CHARACTERIZATION

As noted above, the goal of modeling is to accurately represent component electrical behavior E, and because of model differences values of P best for one model need not be the best values for another model. So instead of relying on direct measurements of P, consider that we choose some E that are most important for circuits (and can be easily and directly measured). Examples include g_m, current at a typical bias, and output resistance. Then for each performance E_i

$$\delta E_i = \sum_j \frac{\partial E_i}{\partial P_j} \delta P_j \tag{4}$$

and the variance follows as

$$\sigma_{Ei}^2 = \sum_j \left(\frac{\partial E_i}{\partial P_j} \right)^2 \sigma_{Pj}^2 \ . \tag{5}$$

So by measuring σ_E and calculating the sensitivities $\partial E_i / \partial P_j$ from the underlying circuit simulation models, the σ_P can be calculated. The process is termed backward propagation of variance (BPV) because it takes measurements in variances of important electrical quantities and then calculates the variances in the process parameters necessary to fit the measured data, rather than directly using measured variances in P and then predicting, in a forward fashion, σ_E.

Note that the calculation in (5) requires the matrix of squared sensitivities to be well conditioned. Testing the condition number of this matrix allows verification of the suitability of the tests E to make the P mathematically observable.

Note that it has not been specified whether the variances in (5) are from mismatch measurements or from ET data. The same formalism is applicable to both mismatch modeling and characterization, via $\sigma_{PL}^2(g)$, and global statistical modeling and characterization, via σ_{PG}^2. If σ_{PL} is significant with respect to σ_{PG}, then calculation of σ_{PG} is done based on $\sigma_{PG}^2 = \sigma_P^2 - \sigma_{PL}^2(g)$, with the mismatch variance removed from the ET data variance. This is the reason for the term "unified" for this modeling approach, not only is it simple and physically correct, but it is valid for both global and local variation modeling.

The final type of statistical model required is corner files. These are generated as follows. Targets for variations in E are specified, generally at the $\pm 3\sigma$ level. Nonlinear least squares optimization is then done to adjust the P so that the targets are modeled. Again, provided the E are selected to be key quantities that effect circuit performance, generating case files this way gives models that are guaranteed to bracket electrical performance. This is not true if the variations are specified in terms of P directly. This gives different variations in E for different models, different variations in E if different techniques are used to determine P, and does not account for the fact that models are, by definition, inaccurate. Further, the common practice of introducing $\pm 3\sigma$ variations in each of n process parameters P gives a probability that this combination of parameters will be seen in practice as $\pm 3\sqrt{n}\sigma$, which becomes progressively unlikely as the number of parameters used in a model increases. (This does not account for the sensitivity of E with respect to each P_j, which makes the variation in E completely unpredictable; generally it is too large).

This same procedure also obviates the need for a "golden" wafer for characterization of typical case models (which in practice is impossible to come across, a wafer with every parameter at its typical value). All that is needed is a reasonably representative wafer. The models extracted from this are tuned to the typical specifications for the process using nonlinear least squares optimization (by adjusting the P, not model parameters S).

6 GLOBAL VARIATION

Fig. 1 shows measured and simulated data for the saturated drain current of wide/long PMOS and NMOS devices. The primary causes of variations in these performances are variation in T_{OX}, which is correlated between PMOS and NMOS, and variation in mobility, which is uncorrelated. (There is also some influence of V_{TH}, which is partially correlated via T_{OX} and partially uncorrelated via V_{FB} and N). The accuracy of modeling both process extremes and distributions is clear.

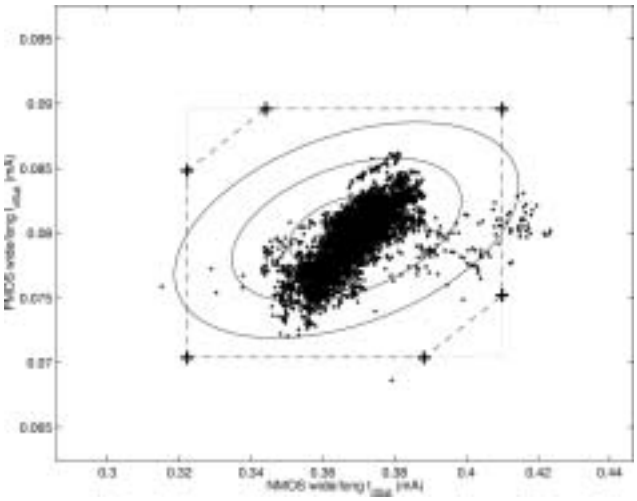

Fig. 1 Measured Data and Simulated Results from Wide/Long MOSFETs

The dotted line represents the fab spec limits. The vertices of the dashed hexagon are the case corner files. The symbols are the measured data. The solid line ellipses are 1, 2, and 3σ contours, derived from a 500 sample Monte Carlo simulation.

Similarly, Fig. 2 shows measured and simulated results for wide/short MOSFETs. Part of the Δ_L, associated with the poly critical dimension (CD) is correlated between NMOS and PMOS, and part, associated with junction lateral diffusion, is mostly uncorrelated. Again, the absolute value of the process upper and lower specs and the correlation structure of the data are accurately captured in the model. Note that for Figs. 1 and 2 the models were generated before the manufacturing data was collected.

Note that the difference in correlation structure between the behavior of long and short devices is captured properly. This is a consequence both of having the correct physical basis P for the underlying statistical model (including mappings from P to S, here only necessary for V_{TH}), and using the BPV technique to characterize σ_{PG}.

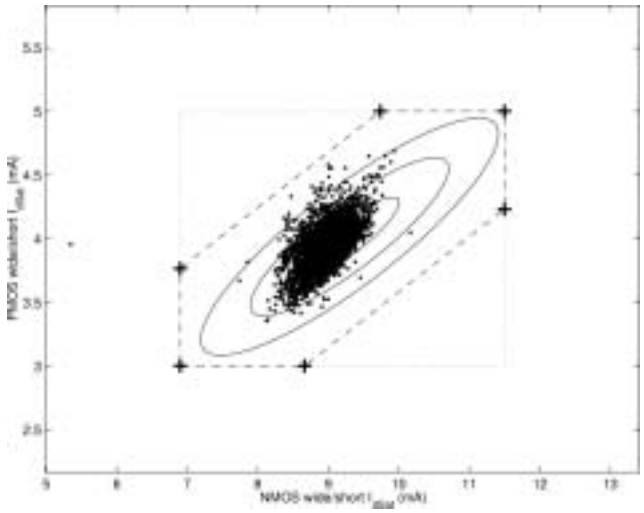

Fig. 2 1 Measured Data and Simulated Results from Wide/Short MOSFETs

7 LOCAL VARIATION

As noted previously, an identical procedure can be used to characterize σ_{PL} for mismatch modeling. The underlying P and model are identical with that used for global variation modeling. The electrical performances E are in this case mismatch in current, conductance, capacitance etc. (but not in parameters like V_{TH}) over bias and geometry. Eqs. (5) are solved, using BPV, for parameters of the $\sigma_{PL}^{\sim}(g)$ model, which includes appropriate physical $1/WL$, $1/L$, $1/W$ etc. components.

Fig. 3 shows I_d mismatch from a 0.28µm CMOS process. Note the decrease in σ_E as L decreases. A conventional MOS mismatch model, based on simple $1/WL$ variations of V_{TH} and gain factor (dotted curves in Fig. 3), is not able to even qualitatively capture some of the behavior seen in the data. Physically, as L reduces so does the body effect, and the mismatch becomes less sensitive to variations in N, even though variations in N increase. So, especially near threshold (top curves and data in Fig. 3), the I_d mismatch decreases for smaller lengths. The reduced sensitivity comes from the circuit simulation model on which the statistical model is built.

8 CONCLUSIONS

We have given details of a unified approach to statistical modeling and characterization that naturally encompasses both global and local variation. A key to accurate modeling is having a sound physical basis for a model. The model does not have to be perfect. The BPV process essentially "takes up the slack" in slight model inaccuracies when fitting data, as it does not depend on directly measured parameter values. Different values of both global and local variations ensue for similar parameters of different models, e.g. Δ_L of different MOSFET models. However, the resulting modeling of variations in electrical performances E is substantially more accurate, and more consistent between different underlying circuit simulation models, than if the parameters are characterized directly.

We have observed many different types of "anomalous" behavior, such as Fig. 3 shows, for different device structures over many generations and types of technologies (CMOS, BiCMOS, SmartPower). In all cases we have been able to find a physical reason for the behavior, and to be able to model it qualitatively and quantitatively, because of the physical basis of the models.

REFERENCES

[1] http://www-device.EECS.Berkeley.EDU/~bsim3/

[2] P. G. Drennan and C. C. McAndrew, *IEEE IEDM Tech. Digest*, pp. 167-170, 1999.

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

[4] J. A. Power *et al*., *IEEE Trans. Semicond. Manufact*., vol. 7, pp. 306-318, 1994.

[5] J. B. Shyu, G. C. Temes, and K. Yao, *IEEE J. Solid-State Circuits*, vol. SC-17, no. 6, pp. 1070-1076, Dec. 1982.

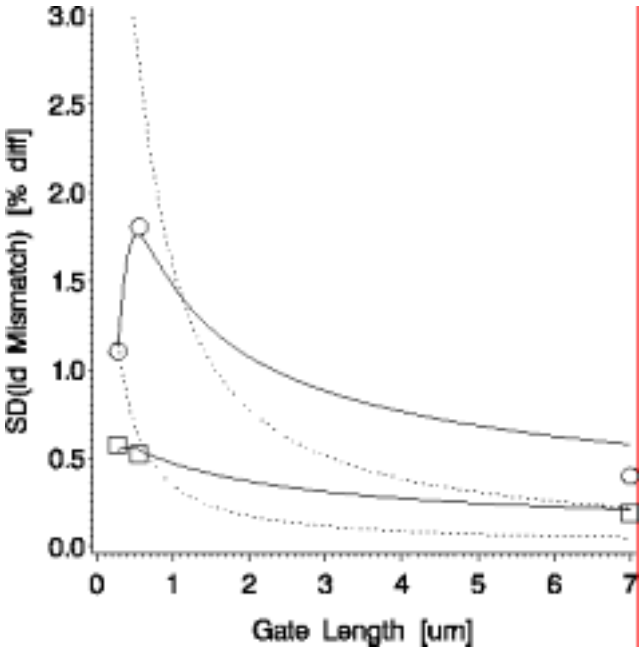

Fig. 3 I_d mismatch over length, symbols are data, dashed lines are conventional model, solid lines are unified model. Top curves are V_{gs}=1.8V, bottom curves are V_{gs}=2.5V