A Novel Methodology for Statistical Parameter Extraction*

Kannan Krishna and S. W. Director

Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA -15213

Abstract

IC manufacturing process variations are typically expressed in terms of joint probability density functions (jpdf's) or as worst case combinations/corners of the device model parameters. However, since device models can only provide approximations to actual device behavior, the difference between the two being the modelling error, only a part of the measured variation in device behavior can be modelled using device model parameter variations and the remaining appears as modelling error variation. In this paper, we present a novel statistical parameter extraction methodology that accounts for the effect of modelling error on device model parameter statistics and can be used to quantify the statistical suitability of conventional MOS device models.

1. Motivation

Inevitable fluctuations in the IC manufacturing process result in manufacturing yields of less than 100% [1]. However, statistical design techniques can be used to improve the yield, at a minimal cost, by modifying the circuit design. These techniques require a characterization of manufacturing process fluctuations, which can be expressed as a joint probability density function (jpdf) or as worst-case combinations of the device model parameters.

Unfortunately, the greatest detriment to the practical use of statistical design techniques in industry has been the fact that no general methodology for determining the jpdf of device model parameter exists. This paper addresses this issue by developing a general methodology for the characterization of manufacturing fluctuations.

2. Introduction

Recall that a *device model* is a mathematical representation of device behavior, expressed by a set of *device model equations*, written in terms of a set of circuit variables (e.g. device terminal voltages) and a set of device model parameters. Device models represent a trade-off between compu-

*. This work has been supported in part by the Semiconductor Research Corporation

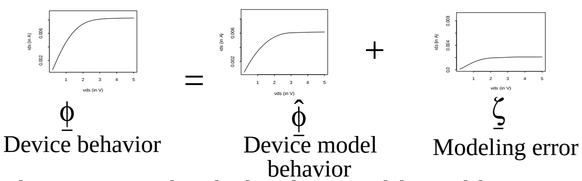

Figure 1 Device behavior partitioned into two components

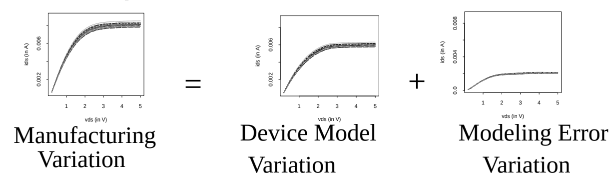

Figure 2 Two components of manufacturing variation

tational complexity and accuracy, which determines the form of the device model equations. Device model parameters are typically determined via *deterministic parameter extraction*, which involves the measurement of device behavior and the use of an optimization-based curve-fitting procedure that minimizes the *modelling error* (Fig. 1).

Since device model parameters are used to characterize the nominal manufacturing process, device model parameter variations are typically used to characterize manufacturing process fluctuations. We may, therefore, define a *statistical device model* in terms of a set of device model equations and device model parameter variations, expressed in terms of a jpdf or as worst-case combinations of the device model parameters.

Statistical parameter extraction is the process of obtaining the device model parameter jpdf or worst-case corners of a statistical device model. It consists of two steps. The first step, called *statistical parameter sample extraction*, involves obtaining a sample of device model parameter sets that correspond to a sample of manufactured devices. A straight-forward approach to statistical parameter sample extraction is to perform a deterministic extraction for each device to obtain a sample of device model parameter sets. However, since each device model parameter set has a different modelling error, the real device behavior variation will be only partially represented by the variation in the device model parameters; the remainder is in the form of modelling error variation (Fig. 2). As a result, this direct approach will provide an erroneous characterization of manufacturing variations if the modelling error component of the real device behavior variation is significant and is ignored. In previous work

[2][3][4], the effect of the modelling error has been largely ignored with the implicit assumption that the modelling error variation is insignificant. We will see that such an assumption is not always valid.The second step, termed *parameter statistic estimation*, is used to condense the information contained in this sample of device model parameter sets into a set of worst case combinations of the device model parameters or an estimate of the jpdf.

In this paper, we shall present a novel methodology for statistical parameter sample extraction that takes into account the effect of modelling error on parameter statistics. Our treatment of the parameter statistic estimation step will be brief and be restricted to the examples since it is accomplished using well-known techniques.

3. Notation & problem definition

Let $\underline{\phi}$ be a *k*-vector of measured device behavior, $\hat{\underline{\phi}}$ be the *k*-vector of the corresponding behavior predicted by the device model, and $\underline{\zeta}$ be the *k*-vector of modelling error. Then,

$$\underline{\phi} = \hat{\underline{\phi}} + \underline{\zeta} \qquad (1)$$

Let $\underline{p}$ be the m-vector of parameters of the device model. Both $\underline{\phi}$ and $\hat{\underline{\phi}}$ are functions of $\underline{p}$, and while the dependence of $\underline{\phi}$ on $\underline{p}$ is not known explicitly, the functional dependence of $\hat{\underline{\phi}}$ on $\underline{p}$ is explicitly known. Let J, $\hat{J}$, and J_{ζ} be the *k*x*m* dimensional Jacobians, respectively, of $\underline{\phi}$, $\hat{\underline{\phi}}$, and $\underline{\zeta}$ with respect to $\underline{p}$.

We can state the statistical parameter sample extraction problem as follows: Given a set of measurements from identically designed devices in a manufactured sample and a device model, determine a sample of device model parameter sets that can be used to mirror the measured variations in device behavior. If such a sample cannot be determined, ascertain the same.

Fig. 3 is a flowchart of the proposed statistical parameter sample extraction methodology; each step is described more fully below.

4. Typical device

Given a sample of devices, the *typical device* is defined as a device from the sample that is representative of the sample. It may be chosen to be the device whose measured characteristics are closest to the median of the characteristics of the input devices. Its parameter vector, $\underline{p}_0$, can be extracted using a deterministic parameter extractor [9].The

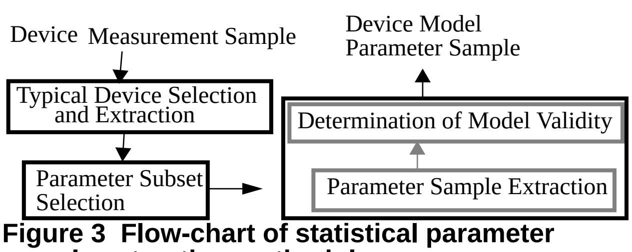

Figure 3 Flow-chart of statistical parameter sample extraction methodology

parameter vector deviation, $\Delta\underline{p}$, for each device in the sample is defined as the deviation of the sample device parameter vector from the typical parameter vector.

5. Parameter subset selection

While a large number of parameters may be required to describe the nominal behavior of a device, manufacturing variations can usually be adequately characterized in terms of only a subset of these, called the *significant statistical parameters*. There are two reasons for this:

- Device model parameters are functions of basic physical characteristics of the device. Typically, these basic physical characteristics are outnumbered by the device model parameters. Consequently, only a subset of device model parameters are truly independent.
- Since the sensitivities of the behavior of the device to different parameters vary widely, the most sensitive parameters may be used to account for most of the manufacturing-induced variations in behavior.

The *parameter subset selection* procedure identifies the set of significant statistical parameters as the parameters corresponding to a *sufficiently* linearly independent subset of the columns of $\hat{J}$ using an algorithm based on the singular value decomposition (SVD) [5] of $\hat{J}$, which is given as:

$$U^T \hat{J} V = \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \qquad (2)$$

Since the diagonal elements of the m-dimensional diagonal matrix, Σ, are the singular values of $\hat{J}$, we can use them to determine the rank of $\hat{J}$ as n (<= m) if $\sigma_n \gg \sigma_{n+1}$, where σ_n is the n-th largest singular value of $\hat{J}$. Then we can partition V as: $V^T = \begin{bmatrix} V_1^T & V_2^T \end{bmatrix}$, where V_1 is nxm. We wish to partition $\hat{J}$ as, $\hat{J} = \begin{bmatrix} \hat{J}_1 & \hat{J}_2 \end{bmatrix}$, such that $\hat{J}_1$ contains a set of sufficiently linearly independent columns of $\hat{J}$. In the SVD, an interchange of columns of $\hat{J}$ corresponds to an interchange of the corresponding rows of *V*. Therefore, to determine a set of sufficiently independent columns of $\hat{J}$, we can perform a QR decomposition, using column pivoting, of V_1 and force a set of independent columns into its first n positions. This ordering of the rows can then be used to order the columns of $\hat{J}$ to get $\hat{J}_1$. Thus, the dimensionalities of $\underline{p}$, $\underline{p}_0$, and $\Delta\underline{p}$ are reduced to n.

6. Parameter sample extraction

To obtain values for the device model parameters for each device in the input sample, henceforth referred to as the *manufactured device*, it is convenient to introduce a few variables.

- $\Delta\underline{\phi} = \underline{\phi}(\underline{p}_0 + \Delta\underline{p}) - \underline{\phi}(\underline{p}_0)$:manufacturing-induced variation in behavior of manufactured device.

- $\Delta\hat{\underline{\phi}} = \hat{\underline{\phi}}(\underline{p}_0 + \Delta\underline{p}) - \hat{\underline{\phi}}(\underline{p}_0)$: *device model component* of the manufacturing variation - component that is accounted for by device model.
- $\Delta\underline{\zeta} = \underline{\zeta}(\underline{p}_0 + \Delta\underline{p}) - \underline{\zeta}(\underline{p}_0)$: *modelling error component* of the manufacturing variation - component that is not accounted for by the device model.

We intend to determine $\Delta\underline{p}$ from $\Delta\hat{\underline{\phi}}$. However, since $\Delta\hat{\underline{\phi}}$ is unknown, a first step involves determining $\Delta\hat{\underline{\phi}}$ from $\Delta\underline{\phi}$, by solving the following non-linear program:

$$\min_{\Delta\underline{p}} \langle \Delta\hat{\underline{\phi}} - \Delta\underline{\phi} \rangle^T \langle \Delta\hat{\underline{\phi}} - \Delta\underline{\phi} \rangle \qquad (3)$$

Since the above non-linear program involves parametric fluctuations [6], it is only mildly non-linear. Therefore, the Gauss-Newton method [7] can be used to solve it.

In order to illustrate some significant points, we shall ignore the value of $\Delta\underline{p}$ obtained from (3) and determine $\Delta\underline{p}$ from $\Delta\hat{\underline{\phi}}$, as follows: Expanding $\hat{\underline{\phi}}(\underline{p}_0 + \Delta\underline{p})$ in a Taylor series about $\underline{p}_0$ and neglecting terms higher than first order yields $\hat{\underline{\phi}}(\underline{p}_0 + \Delta\underline{p}) = \hat{\underline{\phi}}(\underline{p}_0) + \hat{J}\Delta\underline{p}$. So that, $\Delta\hat{\underline{\phi}} = \hat{J}\Delta\underline{p}$ This over-determined system of linear equations can be solved in a least squares sense to obtain $\Delta\underline{p}$.

7. Determining model validity

A strong condition for a device model to be a good statistical model is that the modelling error component of the manufacturing variation is small compared to the device model component i.e., $\|\Delta\hat{\underline{\phi}}\| \gg \|\Delta\underline{\zeta}\|$. This means that most of the manufacturing variation is accounted for by the device model component. If the l_2 norm is used, this is equivalent to, $\Delta\hat{\underline{\phi}}^T\Delta\hat{\underline{\phi}} > \mathrm{N}\,\Delta\underline{\zeta}^T\Delta\underline{\zeta}$, where N is a large positive number. Since the focus of this work is on studying parametric variations, as per the assumptions stated earlier, $\Delta\hat{\underline{\phi}} = \hat{J}\Delta\underline{p}$, $\Delta\underline{\phi} = J\Delta\underline{p}$, and $\Delta\underline{\zeta} = J_\zeta\Delta\underline{p}$. Therefore, the above condition is equivalent to $\Delta\underline{p}^T B \Delta\underline{p} > 0$ where, $B = \hat{J}^T\hat{J} - \mathrm{N}\,J_\zeta^T J_\zeta$, is the *validity matrix*. A necessary and sufficient condition for this to hold is that the eigenvalues of B be positive.Then the device model is suitable for use as a statistical model for all values of $\Delta\underline{p}$. Otherwise, its validity is restricted to the *region of validity,* a region in the $\Delta\underline{p}$ space bounded by $\Delta\underline{p}^T B \Delta\underline{p} = 0$, and illustrated in Fig. 4 for a hypothetical 2-dimensional case.

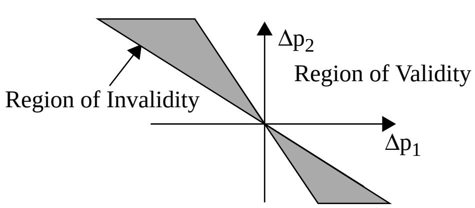

Figure 4 Regions of validity and invalidity for a hypothetical device model.

In order to use the above condition to determine the statistical validity of a device model, the validity matrix needs to be estimated from an estimate of J_ζ obtained using the $\Delta\underline{\phi}$ and the $\Delta\underline{p}$ values of the devices in the input sample.We define $\Delta P = [\Delta\underline{p}_1 \ldots \Delta\underline{p}_r]$ and $\Delta\Phi = [\Delta\underline{\phi}_1 \ldots \Delta\underline{\phi}_r]$, where r is the number of devices in the input sample. Since $\Delta\underline{\phi} = J\Delta\underline{p}$, it can be seen that $\Delta\Phi = J\Delta P$, which consists of k over-determined systems of r linear equations in n unknowns. These can be solved in a least squares sense to obtain an estimate of the elements of J. J_ζ can then be estimated using the relation, $J_\zeta = J - \hat{J}$.

For any application, as N gets larger, the region of validity shrinks. N is typically is chosen to insure that the manufacturing-induced variation can be accurately represented by the device model parameter variations. If the region of invalidity encloses the range of values of $\Delta\underline{p}$ that characterize all or most of the manufacturing variations of interest then the model is not of much use.

8. Examples

The following examples used an input device sample consisting of 61 dies, each containing ten NMOS devices of varying combinations of widths and lengths. In each of the examples, one of these dies was chosen as the typical die and the typical parameter vector was extracted from the measurements made on all of the devices within the typical die using a commercial parameter extractor [9].

8.1 MOS linear region model

In this example, devices with W/L ratios of 20μm/20μm, a width-length combination at which short channel effects are negligible, and the SPICE MOS level 3 model [8]were used. Drain current measurements were made on the devices with zero Vbs and a low Vds of 0.1V, a range in which the devices operate in their linear regions.

The mobility parameter, μ, the threshold voltage, *Vto*, and the mobility degradation parameter, θ, were identified as significant statistical parameters. This agrees very well with what would be determined using knowledge of device physics.

Two-way scatterplots of the extracted values of the three significant statistical parameters are shown in Fig. 5.

A test of the goodness of the linear approximation used in determining these parameters was performed by computing the norm of the second order term in the Taylor series expansion of $\hat{\underline{\phi}}$ that was ignored. It was found that,

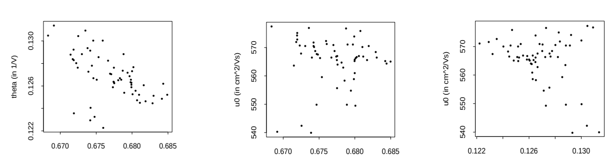

Figure 5 Two-way scatterplots of the extracted parameters

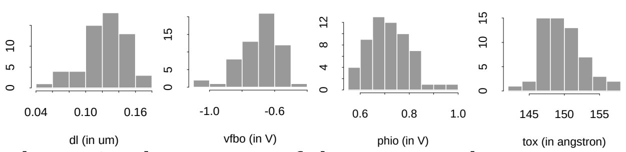

Figure 6 Histograms of the extracted parameters

the second order term was much smaller than the first order term for all the devices and, therefore, the linear approximation is justified.

The eigenvalues of the validity matrix, using N=10, were positive and, therefore, this model was determined to be a good statistical model. For the parameter statistic estimation step, we can obtain a set of worst case combinations of the device model parameters using Fig. 5. Identifying extreme combinations of the parameters gives us the following set of worst-case corners: (Vto,θ,μ)= (0.685,0.125,565.061); (0.669,0.131,540.187); (0.668,0.130,577.422); (0.676, 0.122,571.210). In practice, the specific worst-case combinations chosen will be application-dependent.

From the outset, we expected this model to prove to be a good statistical model because, for a wide-long device with negligible short and narrow channel effects, it mainly models a single physical phenomenon - the behavior of the MOS device as a resistor - and does so very accurately.

8.2 The MOS BSIM model

In this example, devices with widths of 20μm and minimum lengths and the BSIM model [8], were used. The extracted model had 61 parameters, of which 38 were length and width dependence parameters. A total of 1632 measurements, covering the Vgs-Vds-Vbs space, were made for each device. The parameter subset selection algorithm identified the following 4 parameters: ΔL, the length reduction, tox, the oxide thickness, Vfbo, the flat-band voltage, and phio, the surface potential at strong inversion. These results agree very well with the observations made in [10]. In [10], the authors have identified Vfbo, tox, ΔL, and ΔW as being the significant statistical parameters for use in MOS digital design. Since we are dealing with only one width-length combination, ΔL and ΔW are not independent parameters. Also, since our measurements span a large range of back bias voltages, unlike in digital circuits, our method identified phio as a significant statistical parameter.

Fig. 6 shows histograms of the extracted values of the 4 significant statistical parameters. Verification of the adequacy of the linear approximation, produced results that were similar to those obtained in the previous example.

The eigenvalues of the validity matrix, using N=10, indicated that there was a non-trivial region of invalidity for this device model. A physical interpretation of a value of 10 for N is that, we are allowing up to 25% of the manufacturing variation to be not modelled by the device model. Two tests were performed to identify the devices in the input sample for which the extracted parameters were in the region of invalidity. In the first test, the condition, $\Delta \underline{p}^T B \Delta \underline{p} > 0$, was used. Ten devices were identified as having values of $\Delta \underline{p}$ which were in the region of invalidity. The second test used the condition, $\Delta\hat{\underline{\phi}}^T \Delta\hat{\underline{\phi}} > N\,\Delta\underline{\zeta}^T \Delta\underline{\zeta}$, and identified 4 additional devices as being in the region of invalidity. Clearly, the second test is more accurate, since it deals with the actual values of the device model and the modelling error components of the manufacturing variation. Also, the value of the validity matrix used for the test was only an estimate, since its true value is unknown.

Since about 25% of the extracted parameter sets were found to lie in the region of invalidity, meaningful results cannot be obtained from the parameter statistic estimation step. Thus we have an example of the inability of some device models to model the measured variations in device behavior. Our conclusion is that the BSIM model is inappropriate for use as a statistical model here.

9. Conclusions

In this paper, a statistical parameter extraction method has been presented. Its use of only the significant statistical parameters, obtained via the SVD method in this paper, to characterize manufacturing fluctuations, can help avoid the problems with spurious correlations that often occur with conventional methods.

The concept of the typical device used in this method, provides an easy mechanism to separate the components of manufacturing variation that are and are not modelled by device model parameter variations and, thus, can be used to determine the statistical goodness of a device model. To the best of our knowledge, this is the first attempt to quantify the statistical suitability of device models.

10. Bibliography

[1]S.W. Director et. al.; Optimization of Parametric Yield: A Tutorial; Intl. Journal of High Speed Electronics, Vol. 3, No. 1 (1992) 95-136.

[2]C. Michael et.al.; Statistical Modeling of Device Mismatch for Analog MOS Integrated Circuits; IEEE Journal of Solid-State Circuits, Vol. 27, No. 2, (1992) 154-166.

[3]E. Felt et. al.; Measurement and Modeling of MOS Transistor Current Mismatch in Analog IC's; Proceedings of ICCAD-94, 272-277.

[4]J.A. Power et. al.; MOSFET Statistical Parameter Extraction Using Multivariate Statistics; Proc. of the IEEE Intl. Conf. on Microelectronic Test Structures, Vol. 4, No. 1, March 1991, 209-214.

[5]G.H. Golub and C.F. Van Loan; Matrix Computations, 2nd Edition; John Hopkins University Press, 1989.

[6]W. Maly et. al.; VLSI Yield Prediction and Estimation: A Unified Framework; IEEE TCAD, Vol. CAD-5, No. 1, (1986) 114-130.

[7]R. Fletcher; Practical Optimization, 1987.

[8]HSPICE User's Manual; Meta-Software Inc., 1992.

[9]AURORA User's Manual; Technology Modeling Associates, Inc., 1994.

[10]P. Cox, et. al.; Statistical Modeling for Efficient Parametric Yield Estimation of MOS VLSI Circuits; IEEE Transactions on Electron Devices, Vol. ED-32, No. 2 (1985) 471-478.