Solid-State Electronics 54 (2010) 796–800

Contents lists available at ScienceDirect

Solid-State Electronics

journal homepage: www.elsevier.com/locate/sse

Statistical modeling of inter-device correlations with BPV

Ivica Stevanović [a,b,*,1], Xin Li [c,2], Colin C. McAndrew [d], Keith R. Green [e], Gennady Gildenblat [f]

[a] ABB Switzerland Ltd., Corporate Research, CH-5405 Baden-Dättwil, Switzerland
[b] Ecole Polytechnique Fédérale de Lausanne, Station 11, CH-1015 Lausanne, Switzerland
[c] GLOBALFOUNDRIES, Sunnyvale, CA 94085, USA
[d] Freescale Semiconductor, Tempe, AZ 85284, USA
[e] Texas Instruments Inc., Dallas, TX 75266, USA
[f] Arizona State University, Tempe, AZ 85287, USA

ARTICLE INFO

Article history:
Received 31 January 2010
Received in revised form 22 March 2010
Accepted 26 March 2010
Available online 18 April 2010

The review of this paper was arranged by Prof. S. Cristoloveanu

Keywords:
Backward propagation of variance
Device correlation
Semiconductor device modeling
Statistical modeling

ABSTRACT

The backward propagation of variance (BPV) technique for statistical modeling has proven to be efficient and effective in practice. In this paper we extend the BPV formalism to explicitly include modeling of the correlations between electrical performances. The new formulation is verified by applying it to two different MOSFET models and two different manufacturing technologies.

1. Introduction

Statistical modeling and simulation are important for the design of manufacturable integrated circuits [1–8]. Multivariate numerical techniques such as factor analysis or principal component analysis have been developed for statistical modeling [9–17]. However, physically based statistical modeling has proven substantially more accurate than strictly numerical modeling, and is significantly more efficient and more accurate for extrapolation for technology evolutions and for retargeting, when the mean and/or the variation in a process shifts.

MOS transistor models typically include independent physical parameters like oxide thickness t_{ox}, flatband voltage V_{FB}, and offsets Δ_L and Δ_W between design (or drawn) and effective electrical length and width, respectively, and these have been established as an appropriate basis for statistical MOSFET modeling [18]. Bipolar transistor models include many correlated parameters that depend on underlying quantities such as the base doping, base width, and collector doping, but are not explicitly tied to these physical parameters. However, using detailed knowledge of the device structure and device physics it is relatively straightforward to derive relationships between these physical parameters and bipolar model parameters [19], and to implement them in model parameter files.

Given a physically based model it is generally possible to develop test structures and measurement techniques that enable direct characterization of the statistical variations in the process parameters [19]. However, different measurement techniques do not always give the same values for the parameters, which then give different values of simulated electrical performances from one SPICE model, and the same parameter value used in different SPICE models gives different values for simulated electrical performances. The goal of statistical modeling is to accurately represent device electrical performance variation, not to model parameter variations, so “direct” measurement of process parameters does not align with that goal.

A general technique for statistical modeling, built on physically based SPICE models, is backward propagation of variance (BPV) [20,21]. In BPV, the variations in the process parameters, which are denoted $\boldsymbol{p} = (p_1, p_2, \ldots, p_{N_p})$, are calculated so that the variations in measured device electrical performances, denoted $\boldsymbol{e} = (e_1, e_2, \ldots, e_{N_e}) = \boldsymbol{e}(\boldsymbol{p})$, are modeled accurately (in fact modeled exactly if $N_e = N_p$).

Although the BPV procedure is general, and can be used to model both global and local (mismatch) statistical variations [22] and

* Corresponding author at: ABB Switzerland Ltd., Corporate Research, CH-5405 Baden-Dättwil, Switzerland. Tel.: +41 58 586 7103; fax: +41 58 586 7312.
E-mail addresses: ivica.stevanovic@ch.abb.com, ivica.stevanovic@epfl.ch (I. Stevanović), Xin1.Li@globalfoundries.com (X. Li), Colin.McAndrew@freescale.com (C.C. McAndrew), krg@ti.com (K.R. Green), gennady.gildenblat@asu.edu (G. Gildenblat).
[1] Former address: Freescale Semiconductor, 1211 Geneva, Switzerland.
[2] Former address: Arizona State University, Tempe, AZ 85287, USA.

0038-1101/$ - see front matter
doi:10.1016/j.sse.2010.03.018

can include device and circuit level electrical performances [23], it has two limitations: it assumes that the $\boldsymbol{e}(\boldsymbol{p})$ dependencies are linear; and it does not explicitly fit correlations between component of $\boldsymbol{e}$. Extension of the BPV technique to handle nonlinearities was reported in [24]. In this paper we extend BPV to incorporate explicit modeling of correlations between components of $\boldsymbol{e}$.

In our analysis we consider only global variations. Local (mismatch) variations, which are by definition uncorrelated between devices, can be separately characterized [22], and then if they are significant (i.e. for small dimension processes) can be removed from the total observed fluctuations via

$$\sigma^2_{e_m,\mathrm{global}} = \sigma^2_{e_m,\mathrm{total}} - \sigma^2_{e_m,\mathrm{local}}. \tag{1}$$

2. Analytic example

At first it seems somewhat counterintuitive that modeling of correlations can be done with independent (i.e. uncorrelated) parameters. In this section we provide a concrete, but not general, example to show how this can be done.

The effective channel lengths L_P and L_N of PMOS and NMOS devices are correlated. During fabrication, the gate poly formation steps, such as patterning and etch, are the same, so the critical dimension C_D of the poly, the difference between design and on-wafer size, is the same for NMOS and PMOS devices (here we do not consider local variation, i.e. mismatch). The effective electrical channel length differs from the poly dimension because of out-diffusion of the implants used to form the source and drain regions, which are different for PMOS and NMOS devices (although the thermal anneal cycling may be common). Denoting (twice) the out-diffusion length as O_D , and using added subscripts 'P' and 'N' to denote PMOS and NMOS devices respectively, we have [25]

$$\Delta_{L_\mathrm{P}} = C_\mathrm{D} + O_\mathrm{DP}, \tag{2}$$
$$\Delta_{L_\mathrm{N}} = C_\mathrm{D} + O_\mathrm{DN}. \tag{3}$$

For the purpose of statistical characterization, we now form the difference between these effective channel lengths

$$\Delta_{\Delta L} = \Delta_{L_\mathrm{P}} - \Delta_{L_\mathrm{N}}. \tag{4}$$

At first it would seem that this does not provide any additional information over that contained in Δ_{L_P} and Δ_{L_N} alone. Indeed, for any single measurement this is the case, we have

$$\begin{bmatrix}\Delta_{L_\mathrm{P}}\\ \Delta_{L_\mathrm{N}}\\ \Delta_{\Delta L}\end{bmatrix} = \begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & -1 & 0\end{bmatrix}\begin{bmatrix}O_\mathrm{DP}\\ O_\mathrm{DN}\\ C_\mathrm{D}\end{bmatrix}, \tag{5}$$

which is obviously singular and has an infinite number of possible solutions for C_D, O_DP, and O_DN. However, when we look at variances (5) becomes

$$\begin{bmatrix}\sigma^2_{\Delta_{L_\mathrm{P}}}\\ \sigma^2_{\Delta_{L_\mathrm{N}}}\\ \sigma^2_{\Delta_{\Delta L}}\end{bmatrix} = \begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & 1 & 0\end{bmatrix}\begin{bmatrix}\sigma^2_{O_\mathrm{DP}}\\ \sigma^2_{O_\mathrm{DN}}\\ \sigma^2_{C_\mathrm{D}}\end{bmatrix}, \tag{6}$$

which is clearly nonsingular. It may seem counterintuitive that we can transform an indeterminate specific measurement (5) into a deterministic statistical system; there is in fact an implicit third piece of information in the statistical data, the correlation

$$\rho(\Delta_{L_\mathrm{P}}, \Delta_{L_\mathrm{N}}) = \frac{\sigma^2_{C_\mathrm{D}}}{\sqrt{(\sigma^2_{O_\mathrm{DP}} + \sigma^2_{C_\mathrm{D}})(\sigma^2_{O_\mathrm{DN}} + \sigma^2_{C_\mathrm{D}})}}, \tag{7}$$

that supplements the knowledge embodied in $\sigma_{\Delta_{L_\mathrm{P}}}$ and $\sigma_{\Delta_{L_\mathrm{N}}}$ alone.

Table 1
Effective channel length analysis (units of σ are µm). MC (Monte Carlo) results are from a 10,000 sample simulation.

Parameter	Value	Performance	Fab data	MC
σ_{C_D}	0.0195	$\sigma_{\Delta_{\Delta L}}$	0.0196	0.0196
σ_{O_DP}	0.0145	$\sigma_{\Delta_{L_\mathrm{P}}}$	0.0243	0.0244
σ_{O_DN}	0.0132	$\sigma_{\Delta_{L_\mathrm{N}}}$	0.0235	0.0237
$\rho(L_\mathrm{P}, L_\mathrm{N})$	0.6654	$\rho(L_\mathrm{P}, L_\mathrm{N})$	0.6654	0.6698

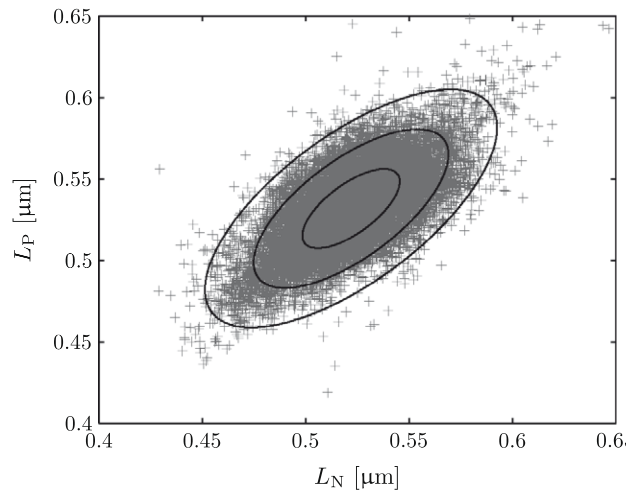

Fig. 1. Correlation between L_P and L_N from measured fab data (symbols) and model (lines, which are ellipses of 1, 2, and 3σ yield curves from 10,000 MC samples).

Applying (6) to measured data from a 0.5 µm BiCMOS process (the technique is independent of technology node, for this node the local variation is much smaller than the global variation so fab data can be directly taken to represent the latter), gives the results of Table 1. Fig. 1 shows the measured fab data compared to simulation results. The accuracy with which the correlation between the effective channel lengths L_P and L_N can be modeled is clear, which is not possible if they are modeled separately (the correlation is zero in that case).

3. General theory and formalism

Let the mean values of the process parameters be $\bar{\boldsymbol{p}} = (\bar{p}_1, \bar{p}_2, \ldots, \bar{p}_{N_p})$. Expanding one component e_m of $\boldsymbol{e}$ around this point, assuming that the elements of $\boldsymbol{p}$ are independent and normally distributed, gives

$$e_m(\boldsymbol{p}) \approx e_m(\bar{\boldsymbol{p}}) + \sum_{i=1}^{N_p} s_{m,i}(p_i - \bar{p}_i),\ \ s_{m,i} = \left.\frac{\partial e_m}{\partial p_i}\right|_{\boldsymbol{p}=\bar{\boldsymbol{p}}}, \tag{8}$$

and the mean and variance of e_m follow as

$$\mu_{e_m} = e_m(\bar{\boldsymbol{p}}), \tag{9}$$
$$\sigma^2_{e_m} = \sum_{i=1}^{N_p} s^2_{m,i}\sigma^2_i. \tag{10}$$

In the standard BPV technique, (9) for all components of $\boldsymbol{e}$ are solved for $\bar{\boldsymbol{p}}$ using nonlinear least squares optimization, where we must have $N_e \geqslant N_p$ and the parameters need to be observable in the electrical performances. The sensitivities $s_{m,i}$ are then computed from the SPICE models by differencing, and (10) are then solved for the variances of the $\boldsymbol{p}$.

To explicitly model the correlations between electrical performances, we now include the covariances