Realistic Worst-Case Modeling by Performance Level Principal Component Analysis

Alessandra Nardi[1]
DEI, Universita' di Padova
Via Gradenigo 6A,
I-35131 Padova, Italy
nardi@dei.unipd.it

Andrea Neviani
DEI, Universita' di Padova
Via Gradenigo 6A,
I-35131 Padova, Italy
neviani@dei.unipd.it

Carlo Guardiani[2]
PDF Solutions, Inc.
333, West San Carlos St.
San Jose, CA 95110, USA
carlo@pdf.com

Abstract

A new algorithm to determine the number and value of realistic worst-case models for the performance of module library components is presented in this paper. The proposed algorithm employs Principal Components Analysis (PCA) at the performance level to identify the main independent sources of variance for the performance of a set of library modules. Response Surfaces Methodology (RSM) and Propagation Of Variance (POV) based algorithms are used to efficiently compute the performance level covariance matrix and non-linear maximum likelihood optimization to trace back worst case models at the SPICE level. The effectiveness of the proposed methodology has been demonstrated by determining a realistic set of worst case models for a 0.25μm CMOS standard cell library.

1.Introduction

The determination of realistic worst case models for synthesis and simulation of VLSI circuits and systems in deep submicron technologies is a critical task because it directly affects the performance of the fabricated circuits. It has been observed [1] that a relatively small increase of circuit performance can be obtained by further scaling the minimum feature size because of the saturation of current, power supply scaling and deep sub-micron effects such as increased interconnect resistance and crosstalk.

Overly pessimistic worst case models, by further tearing down a consistent part of the performance margin, may cause wrong design decisions or dictate the need for newer, expensive technologies also in cases when a more conservative and cheaper fabrication process could have been used instead [2].

These or similar observations have motivated a number of research activities on this subject, which yielded a set of tools and methodologies for the realistic worst case modeling of integrated electronic devices [5]-[8]. In particular, authors in [7] have addressed the issue of using cell level performance correlation in order to reduce the pessimism of worst-case simulation. They analyzed the empirical performance distribution resulting from SPICE model extraction and simulation of basic building blocks. The extraction was performed on the routinely collected process fab electrical test data. The resulting performance correlation was empirically evaluated and circuit primitives grouped accordingly, thus identifying one worst case model for each group of primitives. Similarly, authors in [8] proposed to use performance correlation measured from extensive Monte-Carlo (MC) simulations of cell library timing performance in order to identify clusters of percentile points corresponding to a pre-defined probability value. Response Surface Methodology (RSM) [3] was used in order to speed-up the Monte-Carlo runs and maximum likelihood in the process variable space was applied in order to identify a unique SPICE model for every cluster. Both these techniques are based upon the empirically observed value of sample correlation in order to identify the number and position of performance clusters relative to basic circuit primitives.

Our methodology is based on the fact that the number and value of worst case models may be uniquely determined by the cardinality of the space associated with a set of cell level performances, defined as the number of truly independent performance factors. We propose to use Principal Component Analysis (PCA) in order to determine this cardinality and the corresponding worst case models. The structure of the paper is as follows. In Section 2 we describe the basic statistical tools used in our work. The description of the proposed methodology is given in Section 3, and the experimental results obtained on a 0.25μm CMOS standard cell library are discussed in Section 4. Finally in Section 5 we draw some conclusions and present future work.

2.Background

In this section we will synthetically describe some useful statistical tools that have been used in the derivation of the

1. Employed at STMicroelectronics as an intern until 9/99
2. Formerly with STMicroelectronics

This work has been partially funded by MEDEA-401 cooperative project, and by CNR - MADESS II project, code 2.1.3

proposed worst-case modeling methodology. The three main statistical tools that we have used are: Response Surface Methodology (RSM) in order to compute simple macromodels of the circuit performance as a function of process disturbances, Quadratic Propagation Of Variance (Q-POV) that has been used to compute analytically the entries of the performance-level covariance matrix and Principal Components Analysis (PCA) that has been used to find the main, linearly independent factors explaining most of the modules performance variance.

2.1. Response Surface Methodology

RSM is a well known statistical tool that has been applied in several fields of experimental science to analyze the unknown cause-effect relation existing between a set of *input factors* and a set of *responses*. A good tutorial overview of RSM can be found in [3]. Several authors have proposed different flavors of RSM applied to the design for manufacturability and worst case modeling of electronic circuits and systems, and a good tutorial survey can be found in [11]. The basic common steps of all these different approaches are the following: i) definition of an experiment plan, or Design Of Experiments (DOE) ii) evaluation of the circuit response iii) polynomial fitting of the responses based on linear or non-linear regression analysis.

The output of the RSM flow is usually a set of low-order polynomial equations of the type:

$$\zeta_i = c_{\zeta_i} + \underline{b}_{\zeta_i}^T \underline{x} + \underline{x}^T \mathbf{A}_{\zeta_i} \underline{x} + \dots \tag{1}$$

where $\zeta_i, i = 1, \dots, n$ represents the *i-th* response, $\underline{x}$ the m-dimensional vector of input factors and $c_{\zeta_i}, \underline{b}_{\zeta_i}^T$ and $\mathbf{A}_{\zeta_i}$ represent respectively a constant term, a vector of linear coefficients and a matrix of quadratic coefficients.

2.2. Quadratic Propagation Of Variance (Q-POV)

The statistical properties of the set of random variables $\underline{\xi}$ associated with the system responses is best characterized in terms of its vector mean $\underline{\mu}_\zeta$ and of its covariance matrix Ξ_ζ. The most simple way of computing $\underline{\mu}_\zeta$ and Ξ_ζ is by evaluating them from their definition:

$$\begin{aligned} \mu_{\zeta_i} &= \int \zeta_i f_{\zeta_i}(\zeta_1, \dots, \zeta_n) d\underline{\zeta} \\ \{\Xi_\zeta\}_{i,j} &= \int (\zeta_i - \mu_{\zeta_i})(\zeta_j - \mu_{\zeta_j}) f_{\underline{\zeta}}(\zeta_1, \dots, \zeta_n) d\underline{\zeta} \end{aligned} \tag{2}$$

where $f_{\underline{\zeta}}(\zeta_1, \dots, \zeta_n)$ is the joint probability density function of the set of random variables $\underline{\zeta}$. The integrals in (2) can be estimated by using Monte-Carlo analysis. This however has the drawback of being either computationally expensive or quite inaccurate, because of large number of samples required to stabilize MC estimations. As an alternative it is possible to solve (2) in terms of the known input factors probability moments. The problem is simplified with the aid of a couple of mild assumptions on the statistical properties of the random vector $\underline{x}$, i.e. by assuming that: (i) the input factors are linearly independent; (ii) the set of random variables $\underline{\zeta}$ is a second (or lower) order polynomial function of the random vector $\underline{x}$.

In order to simplify the notation, and without any loss of generality,it is possible to assume that A is symmetric, as it is always possible to reduce A to symmetric form otherwise. In this case, applying the transformation $\underline{\omega} = \underline{x} - \underline{\mu}_x$, it is possible to show that:

$$\mu_{\zeta_i} = c_{\zeta_i} + \underline{b}_{\zeta_i}^T \underline{\mu}_x + \underline{\mu}_x^T A_{\zeta_i} \underline{\mu}_x + tr(A_{\zeta_i} \Xi_x) \tag{3}$$

$$\begin{aligned} \{\Xi_\zeta\}_{i,j} &= \underline{b}_{\zeta_i}^T \Xi_x \underline{b}_{\zeta_j} + E[\underline{\omega}^T A_{\zeta_i} \underline{\omega}\underline{\omega}^T \underline{b}_{\zeta_j}] \\ &+ E[\underline{\omega}^T A_{\zeta_j} \underline{\omega}\underline{\omega}^T \underline{b}_{\zeta_i}] + 2\underline{\mu}_x^T A_{\zeta_i} \Xi_x \underline{b}_{\zeta_j} + 2\underline{\mu}_x^T A_{\zeta_j} \Xi_x \underline{b}_{\zeta_i} \\ &+ E[\underline{\omega}^T A_{\zeta_i} \underline{\omega}\underline{\omega}^T A_{\zeta_j} \underline{\omega}] + 4\underline{\mu}_x^T A_{\zeta_i} \Xi_x A_{\zeta_j} \underline{\mu} \\ &+ 2E[\underline{\mu}^T A_{\zeta_i} \underline{\omega}\underline{\omega}^T A_{\zeta_j} \underline{\omega}] + 2E[\underline{\mu}^T A_{\zeta_j} \underline{\omega}\underline{\omega}^T A_{\zeta_i} \underline{\omega}] \\ &- tr(A_{\zeta_j} \Xi_x) \cdot tr(A_{\zeta_i} \Xi_x) \end{aligned} \tag{4}$$

Furthermore, if the random vector $\underline{x}$ has a gaussian distribution then equation (4) can be further simplified as follows:

$$\begin{aligned} \{\Xi_\zeta\}_{i,j} &= \underline{b}_{\zeta_i}^T \Xi_x \underline{b}_{\zeta_j} + 2\underline{\mu}_x^T A_{\zeta_i} \Xi_x \underline{b}_{\zeta_j} + 2\underline{\mu}_x^T A_{\zeta_j} \Xi_x \underline{b}_{\zeta_i} \\ &+ 4\underline{\mu}_x^T A_{\zeta_i} \Xi_x A_{\zeta_j} \underline{\mu} + 2tr(A_{\zeta_i} \Xi_x A_{\zeta_j} \Xi_x) \end{aligned} \tag{5}$$

2.3. Principal Component Analysis

The purpose of PCA is to find m standard linear combinations of an n-dimensional vector of correlated random variables (with $m \le n$), that have the following properties:

i) among all possible standard linear combinations, the vector of principal components V_{PC} is the one that has maximal variance ii) each component of V_{PC} is linearly independent from all the others.

The vector of principal components can be obtained by applying an orthogonal transformation to the original correlated variables vector, i.e. [4]:

$$y = \Gamma^T(\underline{x} - \underline{\mu}) = \Gamma^T \underline{\omega} \tag{6}$$

where Γ is an orthogonal $n \times n$ matrix such that: $\Gamma^{\mathrm{T}} \Xi \Gamma = \Lambda$ is diagonal and all the eigenvalues $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_n$ are greater or equal than zero. It is possible to show that no other standard linear combination of ω has a variance greater than y_1, the first principal component, and that the ratio:

$$\frac{(\lambda_1 + \lambda_2 + \ldots + \lambda_p)}{(\lambda_1 + \lambda_2 + \ldots + \lambda_n)}$$

represents the fraction of the total variance of x that is explained by the first p components of the vector of principal components y [4]. Note that, as defined in (6) the PCA transformation is $\Re^n \times \Re^n$ as Γ is a square matrix, and the reduction of dimensionality occurs only by selecting the first m components that explain a predefined fraction of the total variance.

3.PCA BASED WORST CASE MODELING

3.1. Problem Definition

The scope of our work is limited to the derivation of realistic worst case models for the performances of a library of components (e.g. a standard cell library) in a given technology.

The objectives of the proposed methodology are:

- To determine the minimum number of combinations of electrical device model parameters sufficient to characterize the worst case performance of a set of library cells.
- To find the value of the worst case model parameters, given a pre-defined assessment of the acceptable level of risk for a worst performance to occur.

3.2. Methodology flow

The flow of the proposed worst case modeling methodology is shown in Fig. 1.

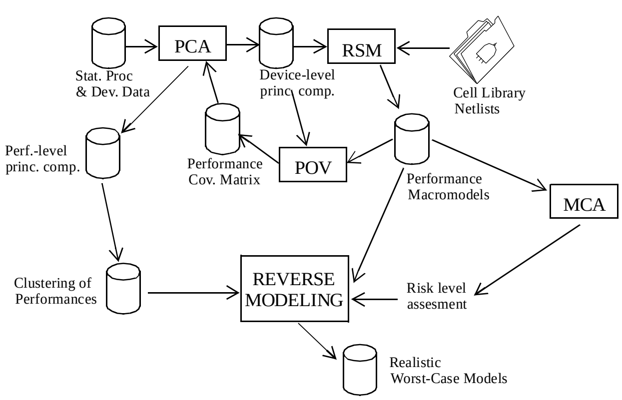


Figure 1. Flow of the PCA based realistic worst case modeling strategy

The input data consist of a set of measurement results including electrical tests (I-V curves, V_T, etc.) and in-line tests (e.g. ρ_s, T_{ox}, etc.) data used to extract a large statistical set of SPICE models.

Following this procedure, a possibly very large $\{n_1 + n_2 + \ldots + n_m\} \times \{n_1 + n_2 + \ldots + n_m\}$ covariance matrix is obtained, where n_i represents the number of first-order parameters for the i-th device, and m is the number of different devices (NMOS, PMOS, etc.) instantiated in the library cells. By using PCA a small number of independent, device-level principal components is derived [9]. Then the RSM macromodeling of the cell library performance parameters (timings, power, etc.) is performed. As different performance parameters are likely to show different process sensitivities, it is reasonable to expect multiple different worst-case corners. Once that the RSM macromodels have been obtained, POV can be used to generate a performance level covariance matrix, as explained in Section 2.2. This is the input for a second step of PCA which will thus generate a vector of *Performance-Level Principal Components*. At the same time Monte-Carlo analysis based on the performance macromodels is used to determine a vector of marginal percentile points for the cell performance $\underline{\zeta}_u$, i.e. a vector whose i-th component ζ_{ui} represents the *u*-percentile of the *i-th* performance. Note that, due to performance correlation, the joint probability of $\underline{\zeta}$ may be zero. Next we identify the number of principal components that are needed to explain a predefined fraction (e.g. 95%) of the total performance variance. Assuming that m PC are sufficient to explain the desired fraction of the total performance variance, then it possible to express each ζ_i as:

$$\zeta_i = \mu_{\zeta_i} + \sigma_{\zeta_i} \cdot \sum_{k=1}^{m} a_{\zeta_i k} \cdot y_k \tag{7}$$

where μ_{ζ_i} and σ_{ζ_i} are respectively the mean and the standard deviation of the performance ζ_i and $a_{\zeta_i k}$ represents the projection of ζ_i along the *k-th* principal component y_k. As performance measures that have maximal projections in the same PCA sub-space are highly correlated, by clustering them according to $a_{\zeta_i k}$ values it is possible to identify the cardinality of the performance space, and thus the number of required worst case corners. The actual value of worst-case model parameters can now be effectively computed by performing separate reverse modeling of the ζ_{ui} in each cluster. Reverse modeling is performed as described in [8].

4.Experimental results

We applied our methodology to a 0.25μm CMOS standard library. Statistical MM9 SPICE models [10] with empirically determined principal components (T_{ox}, ΔL, ΔW, V_{TN}, V_{TP}) have been extracted. We considered 54 performances of a subset of cells representative of the library (AND, NAND, NOR, OR, IV, EXOR), including propagation delay, internal energy per switching, transition time measurements. An RSM macromodel for each performance measure has been created at $V_{DD} = 2.5V$, $C_L = 63fF$, and input slew time $TT = 0.05ns$ by using a CCD [3] design of experiment, requiring 27 SPICE simulation for each performance, and achieving an R^2_{Adj} value (accuracy) always greater than 0.95.

The performance covariance matrix was calculated by using POV on the RSM macromodels. In Table 1 the minimum and maximum correlation coefficient values for all the possible combinations of different types of performance measure are shown. Not surprisingly, similar performance measures show very high correlation coefficients. This is due to the fact that digital library cells are characterized by similar architecture and sizing properties and, therefore, show similar sensitivities to process disturbances, as heuristically observed in [2], [7], [8]

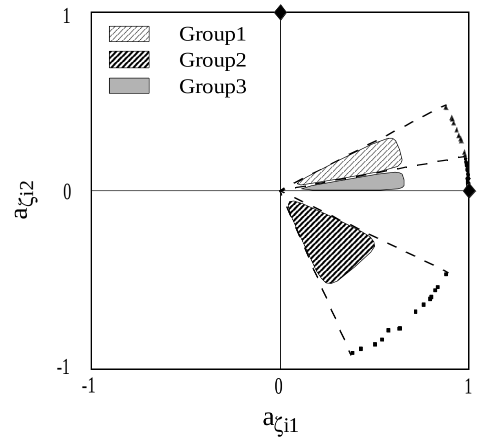
.

Table 1. Minimum and maximum values of correlation coefficients

	Propagation Delay	Power	Output Transition Time
Propagation Delay	0.8977 1	-0.0392 0.8658	0.8676 0.9996
Power		0.7528 0.9993	-0.1072 0.8405
Output Transition Time			0.8812 0.9998

Note that non homogeneous performance measures (out-diagonal terms) may still show quite high correlation coefficients, but also very low values. This information is still not sufficient to determine the actual cardinality of the performance space, which can be determined by performance level PCA as explained in Section 3. In our case the first two performance-level PC were sufficient to explain 99.14% of the total variance: 82.25% is explained by y_1 and 16.89% by y_2. As in (7) each performance measure was expressed as $\zeta_i = \mu_{\zeta_i} + \sigma_{\zeta_i} \cdot (a_{\zeta_i 1} \cdot y_1 + a_{\zeta_i 1} \cdot y_2)$. By clustering performances according to their projection coefficients the following 3 groups of performance measures are obtained. G1 with $a_{\zeta_i 1} » a_{\zeta_i 2}$, G2 with $a_{\zeta_i 1}$ comparable to $-a_{\zeta_i 2}$, G3 with $a_{\zeta_i 1}$ comparable to $a_{\zeta_i 2}$.

G1 includes all the rising transition times and all the propagation delays, except those related to falling output transitions of single stage cells (IV, NAND, NOR). G2 includes all the power measurements. G3 includes all the falling transition times and all the propagation delays not in G1. The plot in Fig. shows the projection coefficients $a_{\zeta_i 1}$ and $a_{\zeta_i 2}$ for the different performance measures, clearly identifying the 3 different groups. Note that the correlation coefficient between two different performance measures is the scalar product of the corresponding vectors. The information displayed in Fig. is sufficient to determine the cardinality of the performance space, which can be graphically associated with the number and size of sectors shown in the graph.



$$\sum_{k=1}^{n} a_{\zeta ik}^2 = 1$$

Figure 2. Performances projection coefficients along the principal components.

Note that despite in some case we have a relatively large correlation coefficient between power and timing measures, the lack of overlap between the corresponding sectors give raise to independent clusters. Therefore in our case two different clusters are observed, one with all the timings ($G1 \cup G3$) and the other with all the power measurements (G2).

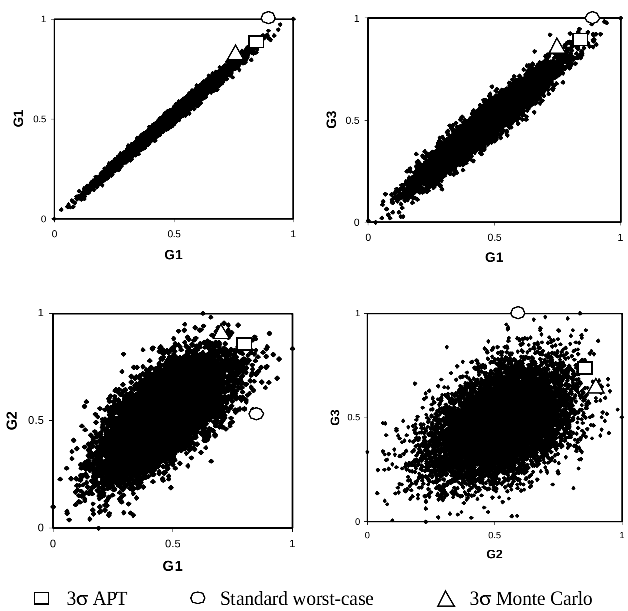


Figure 3. Scatter plots from 10,000 sample MC.

Based on a $1.35x10^{-3}$ probability of a worse performance to occur (i.e. a 3σ percentile point for normally distributed variables) two different SPICE level worst case models have been obtained by using the reverse modeling strategy illustrated in [2]

The scatter plots in Fig. 3 show the results of a 10,000 sample Monte-Carlo, comparing the predictions of the extracted worst case models (filled squares) with the marginal percentile point (filled triangles) and the standard worst-case predictions (dots), for the 3 different groups.

We note that, except for G2, the results of our method are always slightly more pessimistic than the desired marginal percentile point. This is however a necessary trade-off in order to get a common worst case model for all the timing (or power) performance measurements of a library.

5.Conclusions

A new methodology based on performance level PCA in order to determine the number and the value of realistic worst-case models for the performance of module library components has been presented in this work. By using performance clustering based on Principal Component subspace projections it is possible to determine the number and value of the worst-case corners characterized by a desired joint probability of a worse performance to occur. The application of the methodology to a standard cell library in a 0.25 μm CMOS technology, has demonstrated the usefulness and accuracy of the proposed method.

References

[1] D. Sylvester and K.Keutzer, “Getting to the bottom of Deep-Submicron Design” In *Proc. IEEE/ACM International Conference on Comp.-Aided Design*, pp. 203-211 Nov. 1998

[2] A.Nardi, A.Neviani, E. Zanoni and C. Guardiani, “The Impact of Unrealistic Worst Case Modeling on the Performance of VLSI Circuits in Deep Sub-Micron CMOS Technologies”, *In Proc. of IEEE International Workshop on Statistical Metrology*, pp. , Aug. 1998

[3] G.E.P. Box, N.R. Draper, “Empirical Model Building and Response Surfaces”, second edition, John Wiley &Sons, New York 1987

[4] K.V. Mardia, J.T. Kent and J.M. Bibby, “Multivariate Analysis” Academic Prss, San Diego (CA), 1979

[5] A. Dharchhoudhury and S.M. Kang, “An integrated Approach to Realistic Worst-Case Design Optimization of MOS Analog Circuits”, In *Proc. IEEE/ACM Design Automation Conference*, pp. 704-709, Jun. 1992

[6] H.E. Graeb, C.U. Wieser and K.J. Antreich, “Improved Methods for Worst-Case Analysis and Optimization Incorporating Tolerances”, In *Proc. IEEE/ACM Design Automation Conference*, pp.142-147, Jun. 1993

[7] J.C. Chen, C. Hu, Z. Liu and P. K. Ko, “Realistic Worst-Case Spice File Extraction Using Bsim3”, In Proc. IEEE Custom Integrated Circuits Conference, pp.375-378, May 1995

[8] A. Dal Fabbro, B. Franzini and C. Guardiani, “An Assigned Probability Technique to Derive Realistic Worst-Case Timing Models of Digital Standard-Cells”, In *Proc. IEEE/ACM Design Automation Conference*, pp.702-706, Jun. 1995

[9] “pdPca Reference Manual”, second edition, PDFSolutions, Inc., San Jose, Jan. 1998.

[10] R.M.D.A. Velghe, D.B.M. Klaassen, F.M. Klaassen, “MOS MODEL 9, level 902”, Unclassified Report NL-UR 003/94, June 1995.

[11] J.C. Zhang, M.A. Styblinski, “Yield and Variability Optimization of Integrated Circuits”, Kluwer Academic Publishers, Norwell (MA), 1995