

An Experimental Methodology for the Estimation of Spatially Correlated Parametric Yield in Thin Film Devices

Edwin T. Carlen and Carlos H. Mastrangelo

Center for Integrated Sensors and Circuits

Department of Electrical Engineering and Computer Science, University of Michigan

Rm. 2405, EECS Bldg., Ann Arbor, MI 48109-2122, USA

(313)-763-7162, FAX:(313)-763-9324

Abstract— In this paper we present an experimental methodology for parametric yield estimation that accounts for spatial correlations between features of the same device at specific wafer locations. Each device feature is representative of a device parameter that must fit with a specific tolerance box and may be influenced by several steps of the manufacturing process. If the process flow is known and each of its steps is characterized in a spatially correlated manner, the feature pointwise probability density functions (PDFs) can be accurately reconstructed from the processing step pointwise PDFs. This method thus permits the estimation of pointwise device yield more accurately than the common multilevel (run, wafer, die) averaging approach. Because spatially correlated phenomena is subject to both random and systematic non uniformities, the pointwise step (PDFs) are determined by a decomposition process that separates the systematic and random error components. The systematic PDFs are determined from interpolation functions representing the spatial variations across the entire wafer lot, and the random PDFs are approximated using a combination of principle component analysis and factor analysis with a few uncorrelated random variables valid for the entire lot.

I. INTRODUCTION

Over the past two decades there has been much work in yield estimation, modeling, and design centering for semiconductor manufacturing [1]. In thin film devices, point defects determine the catastrophic yield loss while the uniformity variations determine the “out-of-spec” or parametric yield loss [2]. While many elaborate mathematical models, techniques, and tools [3, 4] have been developed for parametric yield estimation, most of these models are largely based on untested assumptions in the parameter distributions with very little hard supporting data [5]. For example, the most widely known statistical process simulator, FABRICS [3], uses monte carlo (MC) methods coupled with random disturbance generators to model the lot, wafer, and die level fluctuations. These disturbance models are based on assumptions of uniform average variances [6] for the entire lot and experimentally “tuned” inter-level average covariances. One of the primary difficulties with this approach is that the spatial information is lost by the averaging process which may easily lead to incorrect yield estimates.

For example, consider the wafer shown in Fig. 1(a)

where areas A and B have different device features that fall within the yield acceptability region. If areas A and

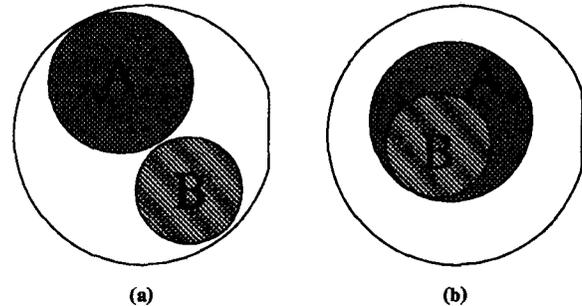


Fig. 1. Representation of acceptable yield areas for different components of a device on a wafer. Even though the average yield in both diagrams is finite and constant, the configuration on the left (a) has zero yield and (b) nonzero yield.

B are not overlapping, the actual parametric yield is zero, but the average parametric yield is not. Therefore only through careful construction of spatially correlated (SC) PDF models an accurate yield prediction is made.

This paper outlines a methodology for the estimation of SC pointwise parametric yield from a description of the device process and its SC process step PDFs. This methodology is used to generate figures of merit for synthetic process flows generated by the process compiler MISTIC [7]. Section II outlines the basic device representation and an unifying vector representation of its basic features. Section III summarizes the technique for construction of SC pointwise PDFs for deposition and etching processes. Section IV describes the connection between the process flow, the step SC PDFs and the ultimate pointwise SC device parametric yield.

II. BASIC DEVICE FEATURE REPRESENTATION

In order to estimate the device yield, it is necessary to define a device representation that captures all of its characteristic features. The representation adopted here first partitions the device onto a series of distinct zones. Each zone is treated as a one dimensional stack of layers where all points within each layer parallel to the wafer surface are considered equivalent because they are subject to the same processing conditions. Consider the MOSFET shown in Fig. 2 where the device is partitioned into seven distinct zones expressed as device vector $\vec{d} = \{z_1, z_2, z_3, \dots, z_7, z_9\}^T$. In this device, even though

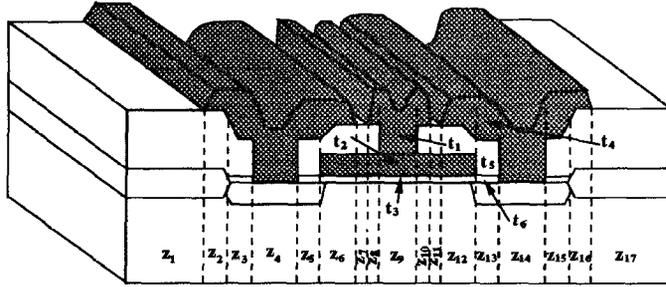


Fig. 2. MOSFET depicting the partitioning of the device into zones for parametric yield estimation.

the gate oxide spans many zones, each zone containing the gate oxide must be considered separately. The characteristic device features within each zone are identified with a vector of parameters. In this paper we focus on the vector of layer thicknesses \vec{t} which is directly affected by the processing steps and ultimately influences the device characteristics. In this example, the layer thickness vector in zone z_9 is $z_9 = \{t_1, t_2, t_3\}$ where t_1 , t_2 , and t_3 are the aluminum, polysilicon gate, and gate oxide thicknesses in the flat areas of z_9 , respectively. This vector was subject to different processing steps than that for zone $z_{13} = \{t_4, t_5, t_6\}$, and so on.

In order to simplify the problem further, we assume that devices are infinitesimally small. This assumption is consistent with the observation that the scale where spatial uniformity changes substantially is much larger than the device size. Therefore, the parametric yield of an entire device is associated with a particular point x in the wafer lot. With the aid of this simplifying representation the connection of individual process steps with the device features is clear. Each layer in a zone is either grown, etched, or unaffected by individual process steps. If each process has a rate and time, assuming linearity, each thickness in each layer at point x is determined from

$$t_i(x) = D_i(x)\tau_{i_0} + \sum_{j=1}^N \alpha_{ij} \frac{Q_j(x)\tau_j}{S_{ij}}, \quad (1)$$

where $D_i(x)$ is the deposition rate of the original material i , τ_{i_0} is the deposition time, $Q_j(x)$ is the etch rate of the j^{th} etchant on t_j , S_{ij} , its selectivity, α_{ij} is a weight coefficient, and N is the number of etching steps affecting t_i . Each deposition \hat{D} and etching rate \hat{Q} is a random variable containing both systematic and random variations. Eq. (1) can therefore be used to find the SC PDF for the device vector \vec{d} if the process step SC PDFs are known.

This estimation methodology is used in the process compiler MISTIC. MISTIC is a process compiler which uses a database of experimentally characterized materials and processing steps and a device cross section to generate process flows as shown in Fig. 3. The compiler reads the device representation and selects the appropriate process steps to construct it. Each of these process steps contains a model for its SC rate PDF stored in the database. This information is used in Eq. (1) to estimate the device SC PDF and its parametric yield. Integration of the device

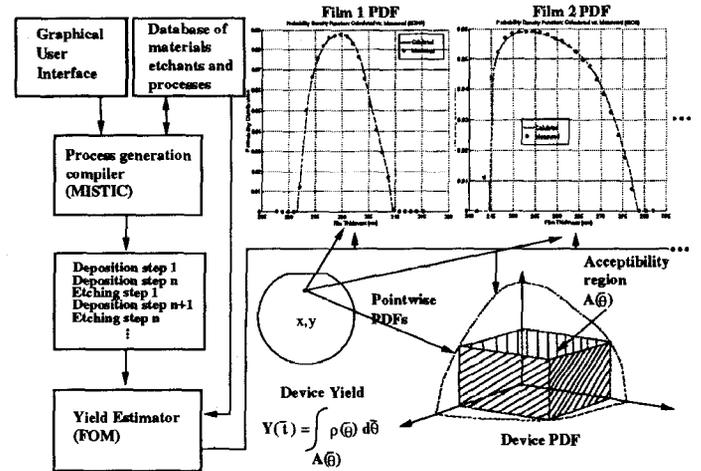


Fig. 3. Organization of the parametric yield estimator within the MISTIC process compilation environment

PDF over the specified acceptability region results in the pointwise parametric yield estimation shown on the right side of Fig. 3. Section III presents the formulation of models for deposition and etching rate PDFs and Section IV shows how to combine them.

III. STATISTICAL CHARACTERIZATION OF THIN FILM PROCESSING STEPS

The effect of a process step on a device is influenced by both deterministic and random factors. Deterministic errors are caused by systematic factors inherently present in the processing equipment such as temperature and pressure gradients that cause predictable non uniformities on the sample. Another important source of systematic error is the loading effect where local rates are affected by the balance between the diffusion of fresh reactants and the depletion of species. In many process steps, the systematic spatial variations are much larger than the random parts; hence they must be included for accurate parametric yield estimation [2].

Random fluctuations in the process step rates are caused by several factors that change randomly both during and between runs. For example, in LPCVD reactors, the deposition rate is affected by two types of random fluctuations. Rate fluctuations are introduced through errors in the control settings such as pressure, temperature, and gas flow rates. These errors are associated with identifiable parts of the equipment which remain uncontrolled or out of spec. Random rate fluctuations are also introduced through the presence of random processes such as gas turbulence and other statistical phenomena. In many VLSI processes, the dominant source of random fluctuations is the former hence their random behavior can be modeled well with just a few random variates.

As part of this study, we have performed a statistical characterization of LPCVD and reactive growth steps. Approximately 35,000 thickness measurements for LPCVD oxide, nitride, polysilicon, and thermal oxide thin films were recorded representing a total of 40 furnace runs. Each run consisted of a lot of 25, 100 mm wafers. Within

each lot, 875 measurements were performed with 35 locations per wafer. From these measurements a reduced model for the SC rate PDF was extracted. A good fitting model for the deposition and growth rates \hat{D} at each point within the lot volume that includes both systematic and random rate components is

$$\hat{D}(x) = D_s(x)(1 + \phi), \quad (2)$$

where $\hat{D}(x)$ is the estimated deposition rate, $D_s(x)$ is the systematic rate variation, ϕ is a zero mean random rate fluctuation, and x is the physical location of the point in the lot. Eq. (2) tells us that the rate variations are primarily influenced by the systematic factor $D_s(x)$ while the randomness is mostly determined by an “average” rate fluctuation. This approximation is justified by the large amount of correlation observed between the points in the lot. This simplifying expression is obtained using a combination of principle component and factor analysis [8, 9]. A refined version of Eq. (2) includes both an “average” and a small “local” fluctuation. Experimentally, the best fit was obtained using

$$\phi = w\phi_l + (1 - w)\phi_w(j) \quad (3)$$

where ϕ_l and ϕ_w are equal variance lot and wafer level independent random variables with $w \approx 0.95$, and $\phi_w(j)$ depends on the wafer number j . Fig. 4 depicts a scatter

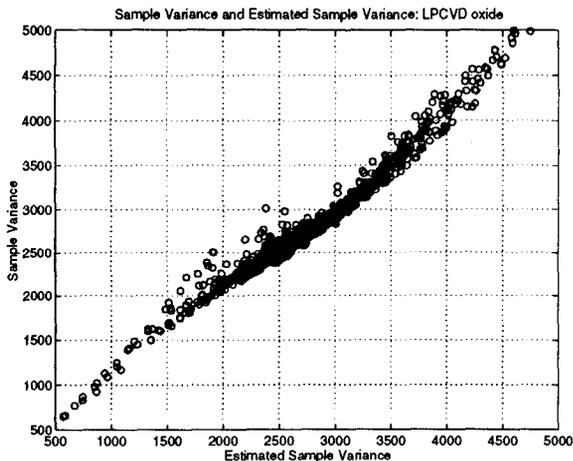


Fig. 4. Scatter plot depicting the closeness of the fit between sample variance calculated from ten LPCVD furnace runs and estimated variance calculated with Eq. (2) and Gaussian random variates. One lot contained 875 measurement locations with 35 locations on each wafer

plot showing the experimental variance and that predicted by the model of Eq. (2). The experimental variance was calculated from 10 LPCVD oxide furnace runs with 875 measurement locations within a single lot. The estimated variance was generated from Eq. (2) using the two added gaussian random number generators of Eq. (3). As evident in Fig. 4, Eq. (2) reproduces the variance satisfactorily while the position dependent randomness has been almost completely retained. The average relative error

between all calculated and estimated variances is approximately 5% for all of the thin films mentioned.

In Eq. (2), the systematic variation term $D_s(x)$ represents the repeatable thickness variations across the wafer surface and lot. Fig. 5(a) depicts an example of the thin film thickness variation across the surface of a LPCVD oxide wafer and Fig. 5(b) depicts an example of the thin film thickness variation along the length of the boat for a single LPCVD oxide furnace run. The systematic distri-

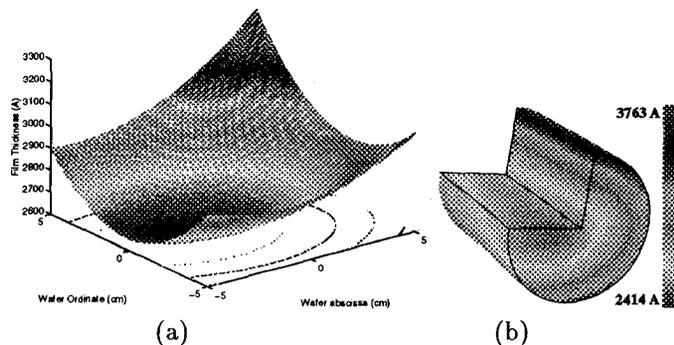


Fig. 5. Thickness variations for LPCVD oxide (a) wafer level (b) lot level

bution function has been calculated analytically using 6-term quadratic fit functions where all contours have been characterized as ellipses or hyperbolas [8]. This analytical model for the systematic variation provides a rapid calculation scheme roughly 100 times faster than the MC method. Fig. 6 shows the comparison between the ana-

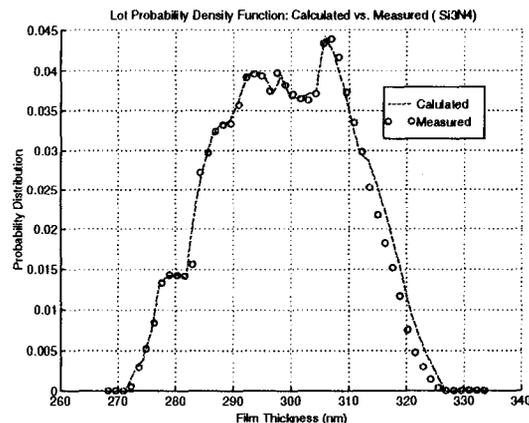


Fig. 6. Comparison between analytically calculated and measured (MC generated) lot level PDFs for a single LPCVD furnace run of nitride

lytically calculated lot level PDF and the MC estimated PDF for a single furnace run of LPCVD nitride showing an excellent agreement of the analytical approximations. We are currently in the process of developing statistical process models for etching steps.

IV. PARAMETRIC YIELD CALCULATIONS

The sections above outline the construction of the device vector \vec{d} and describe how individual process steps

affect it. In the most general sense, each thickness component in the m -dimensional device vector is related to the process steps by the following matrix relationship

$$\vec{d} = P \vec{R}, \quad (4)$$

where \vec{R} is the vector containing all n process step rates, and P is an $m \times n$ process matrix. If the process affects the device vector linearly, the process matrix has the form

$$P = M \alpha S^{-1} \tau, \quad (5)$$

where M is a boolean masking matrix, α is the weight coefficient matrix, S^{-1} is the inverse selectivity matrix, and τ is the $n \times n$ diagonal process time matrix. The mask matrix M specifies if any interaction between the device vector components and processing steps exists. The pointwise parametric yield is hence calculated as [5]

$$Y(\vec{t}, x) = \int_{-\infty}^{\infty} \Phi(\vec{t}, B) \rho(\vec{t}, x) d\vec{t}, \quad (6)$$

where $\rho(\vec{t}, x)$ is the pointwise multivariate SC PDF of the device thickness vector, and $\Phi(\vec{t}, B)$ is an indicator function dependent on the tolerance box $B = \{\vec{t}_\beta, \vec{t}_\alpha\}^T$

$$\Phi(\vec{t}, B) = \begin{cases} 1, & \vec{t}_\alpha \leq \vec{t} \leq \vec{t}_\beta \\ 0, & \text{elsewhere.} \end{cases} \quad (7)$$

The evaluation of the yield integral can be simplified if the indicator function is eliminated. Under the justified assumption of a gaussian multivariate distribution, the parametric yield is

$$Y(\vec{t}, x) = \int_{\Omega_{\vec{t}_\beta, \vec{t}_\alpha}} \frac{|\Sigma_t|^{-\frac{1}{2}}}{(2\pi)^{\frac{m}{2}}} \exp\left(-\frac{1}{2} \vec{t}^T \Sigma_t^{-1} \vec{t}\right) d\vec{t}, \quad (8)$$

where Σ_t is the covariance matrix of \vec{t} and $\{\vec{t}_\alpha, \vec{t}_\beta\}$ are the lower and upper thickness boundaries, respectively. In order to perform the integration though, Σ_t must be determined. Since the device vector depends linearly on the process rate vector \vec{R} , then

$$\Sigma_t = P \Sigma_R P^T. \quad (9)$$

In general the calculation of Eq. (8) is not easily performed for high dimensions; however, since not all device components in \vec{d} are influenced by all processing steps because of causality and blocking considerations, M and Σ_t are in most cases relatively sparse. Therefore Eq. (8) can be partitioned into a product of decoupled integrations with smaller dimensions

$$Y(\vec{t}) = \prod_{i=1}^q Y_i(\vec{t}_i), \quad (10)$$

where

$$Y_i(\vec{t}_i) = \int_{t_{\beta_{i_1}}}^{t_{\alpha_{i_1}}} \cdots \int_{t_{\beta_{i_v}}}^{t_{\alpha_{i_v}}} \frac{|\Sigma_i|^{-\frac{1}{2}}}{(2\pi)^{\frac{v}{2}}} \exp\left(-\frac{1}{2} \vec{t}_i^T \Sigma_i^{-1} \vec{t}_i\right) d\vec{t}_{i_v}, \quad (11)$$

and where v is the dimension of the i^{th} subcovariance matrix, and \vec{t}_i is a subset vector of the device thickness vector \vec{t} . The q subcovariance matrices are determined from analyses of the masking matrix.

The evaluation of Eq. (11) can be performed with the assumption that the yield calculations are probability evaluations of polyhedrals under a reduced dimension multivariate gaussian distribution. The calculation of these probabilities are determined numerically over a v -dimensional polyhedra using the methods cited in [10].

SUMMARY

An experimental methodology for the estimation of spatially correlated parametric yield has been presented. The method uses approximate models for the SC PDFs of a vector specifying the device features. The SC PDFs for the vector are obtained using experimentally determined SC PDFs for individual process steps and knowledge of the process flow. The yield calculation is performed numerically using a correlated multivariate gaussian distribution obtained from the experimental data over a specified tolerance polyhedra.

ACKNOWLEDGMENT

We thank D. K. Jones of the UM Solid-State Laboratory for performing the LPCVD furnace runs. We also thank Dr. S. B. Crary for assistance with the I-OPT DOE software and many helpful discussions. This project was partially supported by the National Science Foundation under grant ECS-9309229.

REFERENCES

- [1] D. Moore and H. Walker, *Yield Simulation for Integrated Circuits*. Boston: Kluwer, 1987.
- [2] B. E. Stine, D. S. Boning, and J. E. Chung, "Analysis and decomposition of spatial variation in integrated circuit processes and devices," *IEEE Transactions on Semiconductor Manufacturing*, vol. 10, no. 1, February 1997.
- [3] W. Maly and A. J. Strojwas, "Statistical simulation of the IC manufacturing process," *IEEE Trans. on Computer-Aided Design*, vol. CAD-1, no. 7, pp. 120-131, July 1982.
- [4] P. Feldman and S. Director, "Integrated circuit quality optimization using surface integrals," *IEEE Trans. on Computer-Aided Design*, vol. 12, no. 12, pp. 1868-1879, December 1993.
- [5] J. C. Zhang and M. A. Styblinski, *Yield and Variability Optimization of Integrated Circuits*. Boston: Kluwer, 1995.
- [6] C. J. Spanos and S. W. Director, "Parameter extraction for statistical IC process characterization," *IEEE Trans. on Computer-Aided Design*, vol. CAD-5, no. 1, January 1986.
- [7] M. Hasanuzzaman and C. H. Mastrangelo, "Process compilation of thin film microdevices," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, no. 7, July 1996.
- [8] E. Carlen and C. Mastrangelo, "Statistical multi-lot characterization of spatial thickness variations in lpcvd oxide, nitride, polysilicon, and thermal oxide films," *To be presented at the 1997 SPIE Conference, Austin, TX, October 1-2, 1997*.
- [9] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall, 1982.
- [10] Z. Drezner, "Computation of the multivariate normal integral," *ACM Transactions on Mathematical Software*, vol. 18, no. 4, pp. 470-480, Dec. 1992.