Non-Gaussian Statistical Parameter Modeling for SSTA with Confidence Interval Analysis

Lizheng Zhang*, Jun Shao†, Charlie Chung-Ping Chen‡
*‡ ECE Department, University of Wisconsin, Madison, WI53706, USA
† Department of Statistics, University of Wisconsin, Madison, WI53706, USA
Email:{*lizhengz@cae.wisc.edu, †shao@stat.wisc.edu, ‡chen@engr.wisc.edu}

***Abstract*— Most of the existing statistical static timing analysis (SSTA) algorithms assume that the process parameters of have been given with 100% confidence level or zero errors and are preferable Gaussian distributions. These assumptions are actually quite questionable and require careful attention.**

In this paper, we aim at providing solid statistical analysis methods to analyze the measurement data on testing chips and extract the statistical distribution, either Gaussian or non-Gaussian which could be used in advanced SSTA algorithms for confidence interval or error bound information.

Two contributions are achieved by this paper. First, we develop a moment matching based quadratic function modeling method to fit the first three moments of given measurement data in plain form which may not follow Gaussian distributions. Second, we provide a systematic way to analyze the confident intervals on our modeling strategies. The confidence intervals analysis gives the solid guidelines for testing chip data collections. Extensive experimental results demonstrate the accuracy of our algorithm.

I. Introduction

As the technology feature size goes below the deep-sub-wavelength domain, the variations of manufacturing parameters are getting more and more significant and greatly impacts the yield [1], [2]. Facing this issue, the classical corner-based timing analysis requires significantly change to reduce the pessimism induced by the ignorance of correlations between parameters. For this reason, statistical static timing analysis (SSTA) has been proposed and extensively studied to provide a systematic timing analysis methodology considering these parameters' complicated statistical behavior and their impacts on timing analysis. [3]–[5] and [6].

Generally speaking, most of the existing SSTA algorithms assume that the statistical characteristics of these process parameter variation have been calibrated *perfectly* at the foundry side (almost as a black box). SSTA algorithms then assume zero confidence interval or zero errors on these given parameters and perform SSTA confidently. In reality, these assumptions are actually questionable and require careful revisions.

First, due to the limited budget and human resources, there could often be only a few test chips or circuits could be manufactured and measured to estimate the data distribution. Therefore, confidence interval or error is finite and has to be carefully analyzed to properly reflect the fidelity of the data sources. Using only finite measurement data samples, it is mandatory to consider confidence interval and error bounds for formal statistical analysis. This effect has been most likely ignored in the literatures.

Second, due to the nature of semiconductor process, even the independent parameter such as gate length and interconnect width/thickness may not follow Gaussian distributions, not to mention the derived parameters such as gate and interconnect delay. Therefore forcing any distribution, especially those with non-zero skewness to be a Gaussian distribution can fundamentally introduce large errors. To deal with the non-gaussian and non-linear issues, several researchers have proposed to use either numerical integration or quadratic function of Gaussian random variables as the basic modeling function [7]–[9]. Since numerical integrations require significant computational efforts, quadratic function approach may be more efficient in runtime-wise. However, those quadratic functions are assumed perfectly given so far. It is not clear how to fit the real measurement data using these quadratic models.

In this paper, we simultaneously provide solutions for these two problems. First, we develop a moment matching based method to fit the quadratic function model of which the first three moments, mean, variance and skewness, matches the given primitive measurement data. Second, we provide a systematic way to analyze the confident interval or error for our modeling strategies. With such, the quality of the modeling process can be analytically evaluated. The confident interval analysis also gives solid guidelines for designing testing chips and making measurement on these chips. Extensive experimental results demonstrate the accuracy of our algorithms.

The rest of the papers are organized as follows. In section II, we present the basics of process variation and measurement process. Section III introduces the moment-matching based quadratic parameter modeling and section V is the confidential interval analysis for the quadratic modeling. At last, in the section V, experiment results of the quadratic fitting and confidence interval analysis are provided.

II. Parameter Variation and Measurement

IC timing parameter variations will cause device and circuit to deviate from their designed value. Classical worst case timing analysis produces timing predictions that are often too pessimistic and grossly conservative. On the other hand, statistical timing analysis (STA) that characterizes timing delays as statistical random variables offers a better approach for more accurate and realistic timing prediction.

A. Quadratic Timing Model

However, to realize the full benefit of STA, one must address a challenging issue that gate/wire delays in a circuit could be correlated since two delays might be affected by the same parameter variation such as voltage supply uncertainties, gate channel length variations, wire geometry variations,...,etc. In [4], [5], [10] the delay D is explicitly related with these parameter variations Y_i by the *canonical timing model*:

$$D = \mu + \alpha R + \sum_i \beta_i Y_i \tag{1}$$

where R accounts the cumulative effect of all variation sources other than considered parameter variations.

The canonical timing model (1) provides an elegant way to deal with the correlations([4]). Unfortunately, the nonlinear relationship between the gate/wire delay and the parameter variations can not be accurately reflected by the *linear* canonical timing model. To mitigate this deficiency, in [7], a *quadratic timing model* is proposed to augment the linear canonical timing model with second order terms:

$$D = m + \alpha R + \sum_i \beta_i Y_i + \sum_{i,j} \Gamma_{ij} Y_i Y_j \tag{2}$$

where Γ_{ij} are quadratic coefficients and m is a constant term which may be different from the mean value of the delay timing variable.

Nevertheless, both canonical and quadratic time model assume parameter variations to be Gaussian distribution which is not always true in reality. In cases when a parameter variation, Y, is not Gaussian, we would like to express it as a quadratic function of a quadratic function of some other Gaussian random variable X and conduct the timing analysis using the method proposed in [7]:

$$Y = aX^2 + bX + c \tag{3}$$

To be a fact, additional truncation may be needed then we substitute the quadratic parameter expression into the quadratic delay model.

B. Parameter Variation Components

All circuit timing parameters are physically measurable quantities and so that their variations can be represented as a set of measurement data of the actual values of the parameter in many manufactured circuits.

Roughly speaking, there are four components of parameter variations in current manufacturing process: (1) the variations between different lots of wafers, *lot-to-lot variations* Y_l, (2)the variations between different wafers in the same lot, *wafer-to-wafer variations* Y_w, (3)the variations between different chips in the same wafer, *chip-to-chip variations* Y_c and (4)the variations between different gates in the same chip, *gate variations* Y_g. The overall parameter variation is then the linear superposition of all these four variation components:

$$Y(l, w, c, g) = Y_l + Y_w + Y_c + Y_g \tag{4}$$

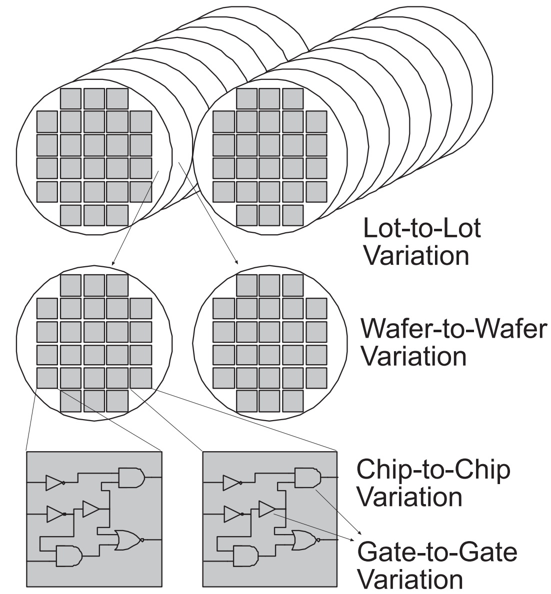


Fig. 1: Variation Sources

where l, w, c, g are indices of the lot, wafer, chip and gate repsectively. So to study the overall parameter variation, it is equivalent to study each of its four components individually assuming these four components are orthogonal to each other. Since all four variation components of a parameter variation can be treated similarly, we will use the *gate-to-gate variation* as an example.

C. Measurement Data Matrix

Of course, we are not able to measure the value of Y_g directly. Instead, we can only measure the value of $Y(l, w, c, g)$ which is a comprehensive effect of all four variation components. In order to characterize the gate-to-gate variation Y_g, we partition the four indices l, w, c, g into two group indices: one group index is the gate index g only and the other group index is the combination of the rest three indices and noted as(l, w, c). With these two group indices, the measurement data set $\{Y_{lwcg}\}$ can be organized as a *measurement data matrix* as:

$$\boldsymbol{M}_Y = \begin{pmatrix} Y_{1111} & Y_{1121} & \cdots & Y_{1211} & \cdots & Y_{2111} & \cdots \\ Y_{1112} & Y_{1122} & \cdots & Y_{1212} & \cdots & Y_{2112} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \end{pmatrix}$$

in which, each row i is a measurement data set for gate i's variation for parameter Y at all chips, wafers and lots. Clearly, the total number of columns will be the same as the total chips manufactured and total number of rows will be the same as the gate number in each chip.

For example, if we manufacture 4 four-gate chips in a single lot of two wafers. Each wafer has two chips on it. For the timing parameter of effective gate channel length L_{eff}, we will get 16 measurement data as: $\{L_{l,w,c,g}\} = \{L_{1111}, L_{1112}, L_{1113}, L_{1114}, L_{1121}, L_{1122}, L_{1123}, L_{1124}, L_{1211}, L_{1212}, L_{1213}, L_{1214}, L_{1221}, L_{1222}, L_{1223}, L_{1224}\}$. If we are interested in the gate-to-gate variation, then we will arrange these 16 measurement data as a 4×4 measurement data matrix as:

$$\boldsymbol{M}_L = \begin{pmatrix} L_{1111} & L_{1121} & L_{1121} & L_{1221} \\ L_{1112} & L_{1122} & L_{1122} & L_{1222} \\ L_{1113} & L_{1123} & L_{1123} & L_{1223} \\ L_{1114} & L_{1124} & L_{1124} & L_{1224} \end{pmatrix}$$

where the first row is the measurement data set for L_{eff} at gate 1, and the second row is that at gate 2, so on so forth.

III. Quadratic Parameter Fitting

From N manufactured chips with G gates on each chip, we can get a $G \times N$ measurement data matrix for a specific gate's timing parameter such as effective channel length L_{eff}. From such a measurement data matrix, we can use statistics methods to estimate the distribution characteristics of the timing parameter at each gate, i.e. mean, variance, skewness etc. Moreover, the parameter variation is not independent between different gates and the covariance between them can then be estimated using the corresponding rows in the measurement data matrix.

A. Moment Estimation

From N manufactured chips, if $Y_1, Y_2, ..., Y_N$ is one row in the measurement data matrix for a timing parameter at the gate Y, then the arithmetic average of these measurement data

$$\hat{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \tag{5}$$

will be the unbiased estimator for the parameter mean at gate Y.

Similarly, unbiased estimators for the variance $\hat{S}^2$, and the skewness $\hat{K}^3$ of the timing parameter at the gate Y using the above data set are:

$$\hat{S}_Y^2 = \frac{1}{N-1}\sum_{j=1}^{N}\left(Y_j - \hat{Y}\right)^2 \tag{6}$$

$$\hat{K}_Y^3 = \frac{N}{(N-1)(N-2)}\sum_{j=1}^{N}\left(Y_j - \hat{Y}\right)^3 \tag{7}$$

If $Z_1, Z_2, ..., Z_N$ is another row in the measurement data matrix as the value of the considered timing parameter at another gate Z, the estimator for the covariance between the parameters at gate Z and Y is:

$$\hat{C}_{ZY} = \frac{1}{N-1}\sum_{j=1}^{N}\left(Z_j Y_j - \hat{Z}\hat{Y}\right) \tag{8}$$

B. Quadratic Fitting

In order to translate the estimated distribution characteristic of the timing parameter, represented by its mean, variance, skewness and covariance, into the sensitivity form needed in the statistical timing analysis, we will have to represent the timing parameter as a function of another computational parameter which is usually assumed to be a standard Gaussian random variable.

If the timing parameter approximately follows Gaussian distribution, the estimated skewness of the parameter will then be close to zero and the timing parameter will then be a linear function of the computational parameter. But if the parameter has significant skewness estimation, it is then be significantly non-Gaussian. We treat such case by setting the timing parameter as a quadratic function of the corresponding computational parameter, as shown in the fitting equation (4), and matching all first three moments between the timing parameter and the function of the computational parameter.

Theorem 1: Assuming random variable $Y = aX^2 + bX + c$ where X is a Gaussian random variable with $\mu_X = 0$ and $\sigma_X = 1$, then:

$$\mu_Y = E\{Y\} = a + c \tag{9}$$

$$\sigma_Y^2 = E\{(Y - E\{Y\})^2\} = 2a^2 + b^2 \tag{10}$$

$$\kappa_Y^3 = E\{(Y - E\{Y\})^3\} = 8a^3 + 6ab^2 \tag{11}$$

Proof: For the Gaussian random variable X, its moments will be $E\{X\} = \mu_X = 0$, $E\{X^2\} = \sigma_X^2 = 1$, $E\{X^3\} = 0$, $E\{X^4\} = 3$, $E\{X^5\} = 0$ and $E\{X^6\} = 15$. Using these results, the moments of Y can be evaluated as:

$$\begin{aligned} & E\{Y\} = E\{aX^2 + bX + c\} \\ = \; & aE\{X^2\} + bE\{X\} + c = a + c \\ & E\{(Y - E\{Y\})^2\} = E\{(aX^2 + bX - a)^2\} \\ = \; & a^2 - 2abE\{X\} - (2a^2 - b^2)E\{X^2\} + \\ & 2abE\{X^3\} + a^2E\{X^4\} \\ = \; & 2a^2 + b^2 \\ & E\{(Y - E\{Y\})^3\} = E\{(aX^2 + bX - a)^3\} \\ = \; & -a^3 + 3a^2bE\{X\} + 3a(a^2 - b^2)E\{X^2\} - \\ & b(6a^2 - b^2)E\{X^3\} - 3a(a^2 - b^2)E\{X^4\} + \\ & 3a^2bE\{X^5\} + a^3E\{X^6\} \\ = \; & 8a^3 + 6ab^2 \end{aligned}$$

■

So knowing the left hand side moments by making estimations on the measurement data for timing parameter Y, we are able to solve the parameters of a, b and c and express the timing parameter of Y as a quadratic function of computational parameter of X.

Equations (10) and (11) in theorem 1 are nonlinear equations, so it is not always possible to get real solution for a, b and c from them.

Theorem 2: Equations (9), (10) and (11) in theorem 1 will have real solution for a, b and c if and only if Y's skewness

$$|\kappa_Y| \leq \sqrt{2}\sigma_Y \tag{12}$$

Proof: We prove the necessity first.

Substitute the equation (10) into (11), we will get a cubic equation about a:

$$f(a) = 4a^3 - 6\sigma_Y^2 a + \kappa_Y^3 = 0 \tag{13}$$

For equation (10) to have real solution for b, it is necessary and sufficient to have

$$\sigma_Y^2 \geq 2a^2 \quad or \quad |a| \leq \frac{\sigma_Y}{\sqrt{2}}$$

Notice that for any a in this range

$$\frac{df(a)}{da} = 12a^2 - 6\sigma_Y^2 \leq 0$$

So we must have

$$f(-\frac{\sigma_Y}{\sqrt{2}}) \geq 0 \text{ and } f(\frac{\sigma_Y}{\sqrt{2}}) \leq 0$$

which is equivalent to $|\kappa_Y| \leq \sqrt{2}\sigma_Y$ with some simple additional derivation.

The sufficiency is simple since all proof steps in above are reversible. ∎

So exact moment matching can only be achieved when the underneath distribution is not heavily skewed ($|\kappa_Y| \leq \sqrt{2}\sigma_Y$). This restriction, however, can be satisfied in most realistic cases for timing parameter estimation. In cases with heavily skewed parameter distribution, approximation has to be made for moment matching equations. For example, we can minimize the skewness matching error while matching mean and variance exactly.

Theorem 3: If equations (9), (10) and (11) in theorem 1 have real solutions for a, b and c, then a will be one of the following three values whichever is real and in the range of $|a| \leq \sigma_Y/\sqrt{2}$:

$$a_1 = -\frac{2\sigma_Y^2 + \Delta^{2/3}}{2\Delta^{1/3}} \quad (14)$$

$$a_2 = \frac{2(1+i\sqrt{3})\sigma_Y^2 + (1-i\sqrt{3})\Delta^{2/3}}{4\Delta^{1/3}} \quad (15)$$

$$a_3 = \frac{2(1-i\sqrt{3})\sigma_Y^2 + (1+i\sqrt{3})\Delta^{2/3}}{4\Delta^{1/3}} \quad (16)$$

where $\Delta = \kappa_Y^3 + i\sqrt{8\sigma_Y^6 - \kappa_Y^6}$ is a complex number. Solutions for b and c cab be found as:

$$b = \sqrt{\sigma_Y^2 - 2a^2} \quad and \quad c = \mu_Y - a \quad (17)$$

Proof: From the proof of theorem 2, we know that if there are solutions for a, b, c, then $|a| \leq \sigma_Y/\sqrt{2}$ and a must be a solution of the cubic equation (13). All there solutions of equation (13) are listed as equations (14) to (16). Also, from the proof of theorem 2, there will be only one out of these three solutions located in the range of $|a| \leq \sigma_Y/\sqrt{2}$.

As long as the solution of a is identified, substitute it into equations (9) and (10), we then compute b and c as equations in (17). ∎

C. Covariance Translation

By translating the timing parameters into computational random variables, the original covariances between timing parameter at different gates has to be translated into the covariances between computational parameter at those gates too. This is achieved by using the following theorem:

Theorem 4: Assuming $Y_1 = a_1X_1^2 + b_1X_1 + c_1$ and $Y_2 = a_2X_2^2 + b_2X_2 + c_2$ are timing parameters in gate 1 and 2 and X_1 and X_2 are computatonal Gaussian random variables at gate 1 and 2. $\mu_{X1} = \mu_{X2} = 0$, $\sigma_{X1} = \sigma_{X2} = 1$, and $\rho_X = cov(X_1, X_2)$. Then:

$$cov(Y_1, Y_2) = \rho_X b_1 b_2 + 2\rho_X^2 a_1 a_2 \quad (18)$$

Proof: Although random variables X_1 and X_2 are correlated, it can be verified that random variables

$$\phi_1 = \frac{X_1 + X_2}{\sqrt{2(1+\rho_X)}} \quad and \quad \phi_2 = \frac{X_1 - X_2}{\sqrt{2(1-\rho_X)}}$$

are independent Gaussian random variables with $\mu_{\phi1} = \mu_{\phi2} = 0$, $\sigma_{\phi1}^2 = \sigma_{\phi2}^2 = 1$, $E\{\phi_1^3\} = E\{\phi_2^3\} = 0$, and $E\{\phi_1^4\} = E\{\phi_2^4\} = 3$. Since

$$\begin{cases} X_1 = \sqrt{\frac{1+\rho_X}{2}}\phi_1 + \sqrt{\frac{1-\rho_X}{2}}\phi_2 \\ X_2 = \sqrt{\frac{1+\rho_X}{2}}\phi_1 - \sqrt{\frac{1-\rho_X}{2}}\phi_2 \end{cases}$$

then the covariance between Y_1 and Y_2 can be evaluated as:

$$\begin{aligned} & cov(Y_1, Y_2) = E\{Y_1Y_2\} - E\{Y_1\}E\{Y_2\} \\ = \;& -a_1a_2 - a_2c_1 - a_1c_2 + \frac{\rho_X^2 - 1}{4}a_1a_2E\{\phi_1^2\phi_2^2\} \\ & \frac{1+\rho_X}{2}(b_1b_2 + a_2c_1 + a_1c_2)E\{\phi_1^2\} + \\ & \frac{1-\rho_X}{2}(a_2c_1 + a_1c_2 - b_1b_2)E\{\phi_2^2\} + \\ & \frac{(1+\rho_X)^2}{4}a_1a_2E\{\phi_1^4\} + \frac{(1-\rho_X)^2}{4}a_1a_2E\{\phi_2^4\} \\ = \;& \rho_X b_1b_2 + 2\rho_X^2 a_1a_2 \end{aligned}$$

where the covariance results for ϕ_1 and ϕ_2, $E\{\phi_1\phi_2\} = 0$ and $E\{\phi_1^2\phi_2\} = E\{\phi_1\phi_2^2\} = 0$ are used. ∎

So knowing the covariance between Y_1 and Y_2, $cov(Y_1, Y_2)$ as we have estimated from the measurement data, we are able to solve the correlation coefficient between X_1 and X_2 which can be used in statistical timing analysis.

IV. Confidence Interval of Quadratic Fitting

From N manufactured chips, we can get an estimation for quadratic coefficients $\hat{a}, \hat{b}, \hat{c}$ and correlation coefficient $\hat{\rho}_X$ using the moment matching method introduced above if the timing parameter $Y = aX^2 + bX + c$ is quadratically dependent on the Gaussian computational parameter X. But we also would like to have some idea of the accuracy for this type of estimation. Since it is hard to directly derive the estimators for coefficients a, b, c, we will instead use a statistical *Jackknife* method based on the following theorem:

Theorem 5: With a $G \times N$ measurement data matrix for a timing parameter from N manufactured chips, let θ be some statistical parameter of interest such as the quadratic fitting coefficient a, b, c for the timing parameter at a gate or the correlation coefficient ρ_X between the computational parameters at two gates. Let $\hat{\theta}$ and $\hat{\theta}_i$, respectively, be estimators of parameter θ computed from the complete measurement data matrix and a reduced measurement data matrix obtained by omitting the i^{th} column in the complete measurement data matrix. We can then construct a *pseudovalue set* $P_1, P_2, ..., P_N$ with $P_i = N\hat{\theta} - (N-1)\hat{\theta}_i$ and an unbiased estimator for θ will be:

$$\hat{\theta} = \hat{P} = \frac{1}{N}\sum_{i=1}^{N} P_i \tag{19}$$

and the end points for the confidence interval of the estimation, at the confidence level of α, are:

$$\hat{\theta} \pm \frac{t_{N-1,\alpha/2}}{\sqrt{N}}\sqrt{\frac{\sum_{i=1}^{N}(P_i - \hat{P})^2}{N-1}} \tag{20}$$

where $t_{N-1,\alpha/2}$ is the $(1-\alpha/2)$ quantile of the Student's **t** distribution with $N-1$ degree of freedom

Proof: See [11]. ■

A. Example of ρ_X Estimation

For example, if we manufacture 4 chips $\{1, 2, 3, 4\}$ and each chip has gate A and B, then for the effective gate channel length L_{eff}, we will have the following measurement data matrix:

$$\boldsymbol{M}_L = \begin{pmatrix} L_{1A} & L_{2A} & L_{3A} & L_{4A} \\ L_{1B} & L_{2B} & L_{3B} & L_{4B} \end{pmatrix}$$

If we are going to use the quadratic form of standard Gaussian computational parameters X_A and X_B to fit the distribution of L_A and L_B respectively, the correlation coefficient ρ_X between X_A and X_B can then be estimated using theorem 4 and the confidence interval of ρ_X can be estimated using the following steps of *Jackknife* method:

1) Estimate $cov(L_A, L_B)$ using equation (8) and the complete measurement data matrix of $\boldsymbol{M}_L$
2) Compute $\hat{\rho}_X$ from theorem 4
3) Construct 4 reduced measurement data matrices as:

$$\begin{aligned} \boldsymbol{M}_L^1 &= \begin{pmatrix} L_{2A} & L_{3A} & L_{4A} \\ L_{2B} & L_{3B} & L_{4B} \end{pmatrix} \\ \boldsymbol{M}_L^2 &= \begin{pmatrix} L_{1A} & L_{3A} & L_{4A} \\ L_{1B} & L_{3B} & L_{4B} \end{pmatrix} \\ \boldsymbol{M}_L^3 &= \begin{pmatrix} L_{1A} & L_{2A} & L_{4A} \\ L_{1B} & L_{2B} & L_{4B} \end{pmatrix} \\ \boldsymbol{M}_L^4 &= \begin{pmatrix} L_{1A} & L_{2A} & L_{3A} \\ L_{1B} & L_{2B} & L_{3B} \end{pmatrix} \end{aligned}$$

4) Repeat step 1) and 2) for each reduced measurement data matrix, we will get four estimations $\hat{\rho}_X^1$,$\hat{\rho}_X^2$,$\hat{\rho}_X^3$ and $\hat{\rho}_X^4$
5) Construct 4 pseudovalues as: $P_1 = 4\hat{\rho}_X - 3\hat{\rho}_X^1$, $P_2 = 4\hat{\rho}_X - 3\hat{\rho}_X^2$, $P_3 = 4\hat{\rho}_X - 3\hat{\rho}_X^3$, $P_4 = 4\hat{\rho}_X - 3\hat{\rho}_X^4$
6) The confidence interval of ρ_X estimation is then equal to that of the mean estimation for pseudovalue set $\{P_1, P_2, P_3, P_4\}$ at the same confidence interval. For example, if the confidence level is 5%, then:

$$\rho_X = \hat{\rho}_X \pm \frac{1.59}{\sqrt{3}}\sqrt{\sum_{i=1}^{4}(P_i - \hat{\rho}_X)^2}$$

B. Error Reduction

So using theorems 5, we are able to estimate both expected values and confidence intervals for all interesting quadratic fitting parameters for a timing parameter. To gauge the accuracy of such type of estimation, we proposed here to use the following ratio:

Definition 1: The accuracy of the statistical mean estimation using theorem 5 is represented by the *estimation error* of ϵ_θ, defined as the ratio between the estimated confidence interval range and the expectation value as:

$$\epsilon_\theta = \frac{t_{N-1,\alpha/2}}{\hat{\theta}\sqrt{N}}\sqrt{\frac{\sum_{i=1}^{N}(P_i - \hat{P})^2}{N-1}}$$

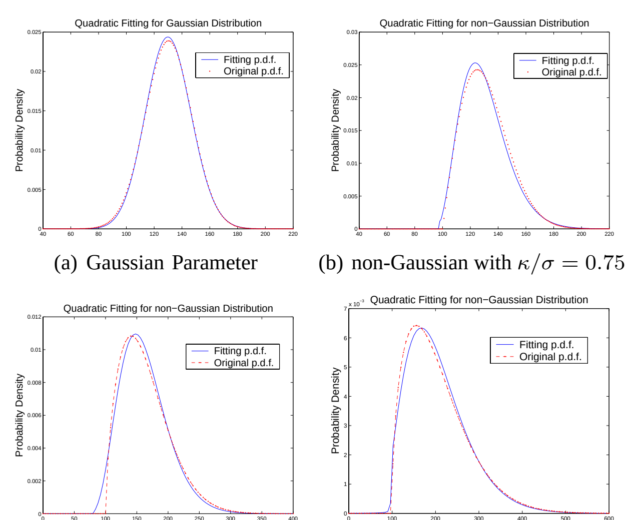
According to theorem 5, the size of the peudovalues will be the same as the original measurement data size and so that the estimation error will reduce as $1/\sqrt{N}$. Clearly, if we want to get one more digit of accurate estimation, we need roughly $100\times$ more chips being measured.

V. Experiment and Discussions

Gaussian and non-Gaussian timing parameters may have significantly different statistical behavior during our estimation and moment fitting. Since there are many type of non-Gaussian distributions, we use Weibull distribution as an example for experiment purpose.

A. Quadratic Fitting with Moment Matching

It is clear that the moment matching between the timing parameter Y and its quadratic fitting form $aX^2 + bX + c$ based on Y's N measurement data $Y_1, Y_2, ..., Y_N$ will have minimum error if the timing parameter Y follows Gaussian distribution since Gaussian random variable can always be expressed as a linear transformation of a standard Gaussian random variable. Such accuracy is clearly shown in figure 2(a).



(a) Gaussian Parameter (b) non-Gaussian with $\kappa/\sigma = 0.75$

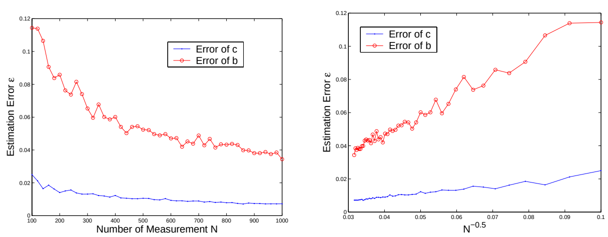
(c) non-Gaussian with $\kappa/\sigma = 0.90$ (d) non-Gaussian with $\kappa/\sigma = 1.00$

Fig. 2: $p.d.f.$s of quadratic form fitting based on $N = 1000$ measurement data for Gaussian and non-Gaussian parameter. κ/σ is the skewness coefficient of the parameter.

But for non-Gaussian timing parameter Y, there is no guarantee that it can be precisely expressed as a quadratic function of a standard Gaussian random variable and so that there will be some intrinsic error for our quadratic fitting $Y = aX^2 + bX + c$. But since we will match all three moments when we do the fitting. The fitting accuracy is still very reasonable as shown in figures 2(b), 2(c) and 2(d) where non-Gaussian distributions with different skewness coefficients κ/σ are shown.

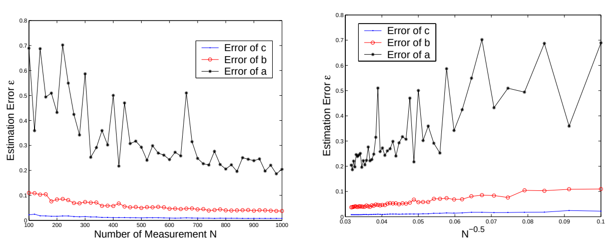
B. Estimation Convergence

Apparently, the size of the measurement data will be with most interest not only because it will directly affect the estimation accuracy but more importantly, because it directly reflect the cost to get the estimation: more measurement will result in higher estimation accuracy but on the other hand, more measurement means manufacturing more chips which is expensive. So the data size will actually be the result of trade-off between accuracy and cost.



(a) Estimation error ϵ v.s. N (b) Estimation error ϵ v.s. $1/\sqrt{N}$

Fig. 3: Estimation error ϵ for Gaussian distribution at different measurement data size N

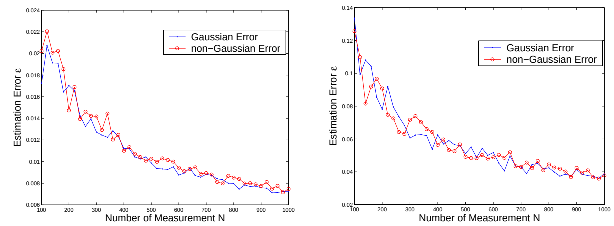


(a) Estimation error ϵ v.s. N (b) Estimation error ϵ v.s. $1/\sqrt{N}$

Fig. 4: Estimation error ϵ for non-Gaussian distribution at different measurement data size N

Figure 3 and 4 demonstrate the estimation error reduction at the rate of $1/\sqrt{N}$ when data size N increases for Gaussian and non-Gaussian timing parameters.

From figure 5, the estimation error seems not very sensitive to the type of distributions of the timing parameter. The estimation done for non-Gaussian timing parameter will roughly has the same size of confidence interval as that done for the Gaussian timing parameter if the measurement data size is the same.



(a) Error for coefficient c (b) Error for coefficient b

Fig. 5: Estimation error ϵ for Gaussian and non-Gaussian distribution at different measurement data size N

VI. Acknowledgement

This work was partially funded by TSMC, UMC, Faraday, SpringSoft, National Science Foundation under grants CCR-0093309 & CCR-0204468 and National Science Council of Taiwan, R.O.C. under grant NSC 92-2218-E-002-030. Also great thanks to professor Barry D. Van Veen for the great discussions.

References

[1] S. Nassif, "Within-chip variability analysis," *Electron Devices Meeting, 1998. IEDM '98 Technical Digest., International*, pp. 283 – 286, Dec 1998.

[2] S. R. Nassif, "Modeling and analysis of manufacturing variations," *CICC*, pp. 223–228, 2001.

[3] J. Le, X. Li, and L. Pileggi, "Stac: statistical timing analysis with correlation," *Design Automation Conference, 2004. Proceedings. 41st*, pp. 343 – 348, June 2004.

[4] H. Chang and S. S. Sapatnekar, "Statistical timing analysis considering spatial correlations using a single pert-like traversal," *ICCAD'03*, pp. 621–625, Nov 2003.

[5] C. Visweswariah, K. Ravindran, and K. Kalafala, "First-order parameterized block-based statistical timing analysis," *TAU'04*, Feb 2004.

[6] author suppressed for anonymous review, "Statistical timing analysis with extended pseudo-canonical timing model," *DATE'05*, March 2005.

[7] L. Zhang, W. Chen, Y. Hu, J. A. Gubner, and C. C. Chen, "Correlation-preserved non-gaussian statistical timing analysis with quadratic timing model," *Design Automation Conference, 2005. Proceedings. 42nd*, pp. 83–88, 2005.

[8] Y. Zhan, A. Strojwas, X. Li, L. T. Pileggi, D.Newmark, and M. Sharma, "Correlation-aware statistical timing analysis with non-gaussian delay distributions," *Design Automation Conference, 2005. Proceedings. 42nd*, pp. 77 – 82, 2005.

[9] H. Chang, V. Zolotov, S. Narayan, and C. Visweswariah, "Parameterized block-based statistical timing analysis with non-gaussian parameters, nonlinear delay functions," *Design Automation Conference, 2005. Proceedings. 42nd*, pp. 71–76, 2005.

[10] A. Agarwal, D. Blaauw, and V. Zolotov, "Statistical timing analysis for intra-die process variations with spatial correlations," *Computer Aided Design, 2003 International Conference on. ICCAD-2003*, pp. 900 – 907, Nov 2003.

[11] K. M. Wolter, *Introduction to Variance Estimation*. New York, Berlin, Heidelberg, Tokyo: Springer-Verlag, 1985.