

國立交通大學

電機學院 IC 設計產業研發碩士班

碩士論文

考慮製程變異下應用 Bootstrap 信賴區間模擬空間相關性

BOOTSTRAP CONFIDENCE INTERVALS AS AN APPROACH to
MODEL WITHIN-DIE SPATIAL CORRELATION UNDER PROCESS
VARIATIONS

研究生：蘇炳熏

指導教授：李育民 教授

中華民國九十八年一月

考慮製程變異下應用 Bootstrap 信賴區間模擬空間相關性
BOOTSTRAP CONFIDENCE INTERVALS AS AN APPROACH to
MODEL WITHIN-DIE SPATIAL CORRELATION UNDER PROCESS
VARIATIONS

研究生：蘇炳熏

Student : Bing-Hsing Su

指導教授：李育民

Advisor : Yu-Min Lee



Industrial Technology R & D Master Program on
IC Design

January 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年一月

考慮製程變異下應用 Bootstrap 信賴區間模擬空間相關性

學生：蘇炳熏

指導教授：李育民

國立交通大學電機學院產業研發碩士班



摘要

隨著製程的技術進步及系統單晶片的到來，深次微米中的時序問題已越顯重要。傳統以 corner value 為基礎的時序分析將會導致預測的時序被過分低估。統計型靜態時序分析 (Statistical Static Timing Analysis - SSTA) 就是利用統計的方式去描述這些製程偏差, 把他們視為一些統計的隨機變數, 然後利用他們去預測時序並且得到更準確的結果。SSTA 使晶片設計者能得以將時間餘裕(timing margin)及良率(yield)做最佳化以提升晶片效能和可靠度。不同於其它論文之模型探討, 本文考量實際製程變異及晶圓應用的可行性, 提出一個實用且新穎的路徑學習重複取樣的方法論(path-based learning methodology with Balanced Bootstrap re-sampling)。此方法不須作任何timing model的假設, 只須從現有寶貴的晶圓量測資料中, 重覆作re-sampling learning的動作, 即可得到準確path delay的空間相關性

(spatial correlations)之推論。同時藉由建立信賴區間的方式，可得到path delay correlation 和 path distance兩者之間的關係和趨勢，由此推論path distance變化時，會有多少path delay的變化。此方法在晶圓廠有兩個方面可應用：(1)在先進晶圓製程階段，使用re-sampling 方法，可對有限的晶圓量測資料，快速建立近似的統計型時序模擬器。(2)在成熟晶圓製作階段，使用路徑學習方法，觀察測試晶片的量測資料，可取得統計型靜態時序分析(SSTA)之建模(modeling)。



BOOTSTRAP CONFIDENCE INTERVALS AS AN APPROACH to MODEL WITHIN-DIE SPATIAL CORRELATION UNDER PROCESS VARIATIONS

student : Bing-Hsing Su

Advisors : Dr. Yu-Min Lee

The logo of National Chiao Tung University is a circular emblem with a gear-like outer border. Inside the circle, there is a stylized figure holding a torch, and the year '1896' is inscribed at the bottom. The text 'Industrial Technology R & D Master Program of Electrical and Computer Engineering College National Chiao Tung University' is overlaid on the logo.

Industrial Technology R & D Master Program of
Electrical and Computer Engineering College
National Chiao Tung University

ABSTRACT

1896

With the advances to nanometer technologies and SOC, the process variation plays a more important in the future. Traditional corner value timing analysis becomes less effective and grossly conservative. Statistical timing models and simulation methods are required to capture these variation effects. The methodology of statistical timing analysis that characterizes time variables as statistical random variables offers a better approach for more accurate timing predictions. SSTA enable designer to setup and hold Timing Margins to optimize and improve the performance and reliability. The thesis considers real process variations and fabrication implementation designs a practical test chips and presents the implementation of a novel path-based learning methodology with Balanced Bootstrap re-sampling that accounts for process variations and their

spatial correlations. It doesn't need timing model hypothesis and make the accurate timing spatial correlation inference from fabrication measurement data. By constructing the confidence interval of the spatial correlation, we can get correlations and predictions for path delay and path distances. By this, we can know how many path distance changes will cause how many path delay correlation changes. It can be applied for two purposes. First, the bootstrap re-sampling can be used to produce a fast and approximated simulator for statistical timing simulations in the advanced production phase. Secondly, this path-based learning can be used as a vehicle to derive statistical static timing analysis (SSTA) based on observed measured data from the test chips in mature production phase.



誌 謝

這篇論文能夠順利地完成，首先要由衷地感謝我的指導教授 李育民博士。每當我遭遇困難疑惑之時，老師總像一盞明燈，持續的給予我方向，讓我能突破瓶頸，使得這篇論文能順利完成。

在實驗室裡，感謝培育學長寶貴的知識經驗傳承，正忠與國富在課業論文上的討論與加油打氣，以及實驗室學弟懷中、宗祐平日的關心與幫忙，使我紓解生活壓力，一路相伴與成長。

最後要深深地感謝我的老婆，負擔起家中大大小小的事，給我精神上最大的支持，以及兩個寶貝兒女的陪伴關心。僅在此將本論文獻給你們，共享這份喜悅與榮耀。

Contents

1 Introduction

1.1 Introduction	1
1.2 Motivation.....	2
1.3 Contribution.....	3

2 Preliminaries

2.1 Process Variation.....	4
2.2 Gate Length Variation.....	4
2.3 Process Variation Decomposition.....	4
2.4 Modeling of Spatial Correlation.....	5
2.5 Gate Length Spatial Correlation.....	6
2.6 Path-Base Spatial Correlation.....	8
2.7 The Path-Base Spatial Correlation Coefficient.....	9

3 Bootstrap Confidence Intervals for the Correlation Coefficient

3.1 Bootstrap Concept.....	10
3.2 Balanced Bootstrap.....	11
3.3 Bootstrap Confidence interval.....	12
3.3.1 The Percentile Method.....	12
3.3.2 The Bias-Corrected and Accelerated Method.....	13
3.3.3 The Normal Method.....	14
3.4 Bootstrap Confidence Intervals for the Correlation Coefficient.....	15

4 Experimental Results

4.1 Experiment Setup.....	16
4.2 Path Correlations with Multivariate Normal Gate Length Variations.....	17
4.3 Path Correlations with Multivariate Chi Square Gate Length Variations.....	21

5 Conclusions

List of Figures

Figure 1. Spatial Correlation Simulation.....	6
Figure 2. Inverter trend array.....	7
Figure 3. Path spatial correlation.....	9
Figure 4. Flowchart of Bootstrap percentile confidence interval.....	12
Figure 5. Flowchart of Bias-Corrected and Accelerated confidence interval.....	13
Figure 6. Flowchart of Bias-Corrected and Accelerated confidence interval.....	14
Figure 7. Flow of bootstrap of silicon data.....	15
Figure 8. Test Chip layout with multiplexor.....	16
Figure 9. Gate length histogram.....	17
Figure 10. A single inverter delay simulation with gate CD variations.....	17
Figure 11. Path delay correlation distributions of inverter trend 1, 2 and 3.....	18
Figure 12. The Histograms of correlation coefficients of Bootstrap of 5 by 5 inverter trend.....	18
Figure 13. Coverage rate for BCA, Percentile, Normal, Normal with filter and Fisher Z.....	19
Figure 14. Confidence interval bound for delta x distance.....	20
Figure 15. Confidence interval bound for delta x distance.....	20
Figure 16. The histogram of a single gate length variation with chi square distribution.....	21
Figure 17. The Histogram of correlation coefficients of Bootstrap of 1 and 2 inverter trend.....	21
Figure 18. Coverage rate for BCA, Percentile, normal, normal with filter and Fisher Z.....	22
Figure 19. Confidence interval bound for delta x distance.....	22
Figure 20. Confidence interval bound for delta Y distance.....	23

Chapter 1

Introduction

1.1 Introduction

In today's fabrication processes, there are many process parameters including doping, gate oxidation thickness, gate length, etc. Our research focus here, gate length variation is the dominant source of delay variation. This is due to the fact that gate length variation affects the device at different steps throughout the fabrication process, such as lithography dose variation, focus variation, etch rate variation, etc. These accumulated variations result in the electrical function far away from the target. For example, gate length variation can induce threshold voltage shift, current change and timing delay. The gate length has high correlation to its neighboring so the spatial correlation is our main research purpose. We give an innovated learning base SSTA modeling with bootstrap re-sampling to discuss spatially correlated timing delay caused by gate length variations.

As we know, most of the sources of gate length variation are systematic in nature, and it is possible to remove these systematic variations. By our modeling, we can do it well and give a good perdition for it.

Although many researchers have dealt with process variations, most of them have ignored spatial correlations by simply assuming zero correlation among devices on the chip. The difficulty in considering spatial correlations between parameters is that it can result in complicated path correlation structures which are hard to deal with. The authors in [1] compute path correlations on the basis of pair-wise gate delay covariance and use an analytic method to derive lower and upper bounds of circuit delay. The authors of [2] manipulate the complicated correlation structure with the Principal Component Analysis (PCA) technique to transform the sets of correlated parameters into sets of uncorrelated variables. The statistical timing computation is then performed with a PERTlike circuit graph traversal. The statistical timing analyzer in [3] imposes upper bounds and lower bounds on the delay correlations. These bounds can then be refined through learning the actual delay correlations from the path delay testing on silicon. Most of them have the strong hypothesis about path delay modeling and don't add process variation conditions into the modeling. Authors of [16], [20] propose a practical implementation to measure spatial correlation from test chips. It is good for huge silicon data but may be trouble in small data. In advanced technology, the process is not stable and silicon data are precious. How to use the limited silicon data to get the useful reference for spatial correlation modeling and how to validate the modeling with the silicon data is a hard topic

1.2 Motivation

SSTA is attractive because traditional worst-case corner timing analysis has become overly conservative due to the process variations. To take the advantage of SSTA technology in practice, the immediate challenge would be how to obtain an accurate and efficient statistical timing model. Many methodologies have been proposed [4], [5] to get a reliable timing model. There are many sources of pessimism in statistical timing model analysis and many of them are dependent on the method used for analysis. How to model and handle spatial correlations in SSTA is also an important research topic. Various spatial correlation models have been introduced. In [6], the authors use PCA to handle spatial correlations. They assume all delay distributions to be Gaussians, and approximate the MAX operation of 2 or more Gaussian distributions to be Gaussian as well. In [7], they propose a canonical first-order delay model and use an incremental block based timing analyzer to propagate arrival times. Considering process variations into SSTA modeling are more important when the feature size of MOS transistor scales down to nanometers. Under process variations, parameters such as the gate length, the gate width and the metal line height are random variables. Among so many types of variation parameters, spatial correlations are recognized to have significant impact on timing of design [9], [10], and the accuracy of SSTA analysis [11]. Our primary goal is to demonstrate that path-based learning is indeed feasible, and shows how it can be applied in test and diagnosis applications. Many papers describe the process variation modeling but lack to account for it into spatial correlations. A learning procedure from fabrication silicon data to study SSTA is practical and useful. We can design test chips to investigate electrical device parameters and use inverter trend to simulate path delays. After getting the precious silicon data, we use methodologies to model spatial correlation matrix and apply it into statistical static timing analysis (SSTA) modeling.

1.3 Contributions

In this work, we propose a novel path-based learning methodology with Balanced Bootstrap re-sampling that accounts for within-die process variations. By designing a test chip with the inverter trend array, we add gate length variations to the test chip and investigate the spatial correlation for the path delay. After running HSPICE to the test chip, we get the path delay and apply the Balanced Bootstrap confidence interval to estimate the spatial correlation. When the feature size of MOS transistor scales down to nanometers, the fabrication production cost is expensive and experiment silicon data is small. The Balanced Bootstrap can be applied for the small samples and doesn't need any timing model hypothesis like normal path delay distributions. We demonstrate the Bootstrap confidence interval has a good coverage rate for the true spatial path delay correlation. We also provide a prediction bound of the path spatial correlation for different distances



Chapter 2

Preliminaries

2.1 Process Variation

As the feature size of silicon is scaling down and the wafer size is getting larger, process variation is increasingly difficult to model and analysis. The variation is due to two factors: environmental and physical. Environmental factors occur at run-time and physical factors occur during the manufacturing of the circuit. Environmental process variation can decompose into lot to lot, wafer to wafer and die to die. Our research focuses on intra-die variation and the channel length (L_{eff}) variation is the emphasis. The reason is that MOSFETs are very critical to this type of variation and have large impacts on timing delay.

2.2 Gate Length Variation

For many process variation sources, gate length variation is the dominant source of the delay variation [8]. The gate length variation come throughout the fabrication process like focus variation, etch rate variation, scanner variation, etc. Gate length variations are pattern dependences. Close gates are higher similar length than far ones. In our thesis, we model gate length variation and put it into spatial correlations for path delays.

2.3 Process Variation Decomposition

We denote V as the measured gate length process parameter and model it as a random variable. Its overall variation can be decomposed into three distinct components: the inter-chip global variation X , the intra-chip spatial variation Y , and the purely uncorrelated random variation Z , i.e.,

$$V = v_0 + X + Y + Z; \quad (1)$$

where v_0 is the mean value of V , and X , Y and Z are random variables. How to extract the systematic variation has been studied in many papers [12], [13]. Our focus is on the zero-mean random variation components X , Y and Z . The inter-chip global variation X models the variation due to global variation effects. It is the same within the chip but different for different chips. The intra chip spatial variation Y models location-dependent variations within the chip. It is different at different locations within the same chip. The random variation Z models the purely uncorrelated random component that is not explainable by either the inter chip global variation X or intra-chip spatial variation Y . We assume

that X, Y and Z have mutually independent distributions. Hence, the variance of V can be given by

$$\sigma_V^2 = \sigma_X^2 + \sigma_Y^2 + \sigma_Z^2 \quad (2)$$

Where σ_X^2 , σ_Y^2 , and σ_Z^2 are the variances of X, Y, and Z respectively. The σ_V^2 is total variance.

2.4 Modeling of Spatial Correlation

In below 65nm technology, lithography of gate length CD has high correlated on neighboring patterns in the layout (proximity effect), the location in the layout (lens aberrations), and the density of features on the mask (flare). This implies that devices are more likely to have similar characteristics in close area than devices that are far apart. It means that correlation function $\rho(x_i, x_j, y_i, y_j)$ between any two points (x_i, y_i) and (x_j, y_j) depends only on the distance v between them, i.e.,

$$\rho(x_i, x_j, y_i, y_j) = \rho(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}) = \rho(v) \quad (3)$$

Suppose there are M chosen points on the chip and their joint spatial variation variable $Y = (Y_1, Y_2, \dots, Y_M)^T$ forms a multivariate Gaussian process with respect to their locations on the chip [19]. Any two points correlation function ρ can be calculated by (3) and hence the correlation matrix of spatial variation variable S is formed in (4):

$$\rho = \begin{bmatrix} 1 & \rho_{1,2} & \rho_{1,3} \cdots \rho_{1,M} \\ \rho_{2,1} & 1 & \rho_{2,3} \cdots \rho_{2,M} \\ \rho_{3,1} & \rho_{3,2} & 1 \cdots \rho_{3,M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{M,1} & \rho_{M,2} & \rho_{M,3} \cdots 1 \end{bmatrix} \quad (4)$$

For any two points V_1 and V_2 of measured gate length process parameters,

$$\begin{aligned} \text{cov}(V_i, V_j) &= \text{cov}(X + Y_i + Z, X + Y_j + Z) \\ &= E[(X + Y_i + Z) - E(X + Y_i + Z)][(X + Y_j + Z) - E(X + Y_j + Z)] \\ &= E[X + (Y_i - E(Y_i)) + Z][X + (Y_j - E(Y_j)) + Z] \\ &= \sigma_X^2 + \text{cov}(Y_i, Y_j) + \sigma_Z^2 \\ &= \sigma_X^2 + \rho(V_{i,j})\sigma_i\sigma_j + \sigma_Z^2 \end{aligned} \quad (5)$$

where $\rho(v_{i,j})$ is the spatial correlation coefficient between two locations that are $v_{i,j}$ distance apart. The same distance $v_{i,j}$ always corresponds to the same $\rho(v_{i,j})$, regardless of their locations.

Those M chosen gate length process parameters form the joint distribution variable $V = (V_1, V_2, \dots, V_M)^T$ and its mean and variance are

$$\mu_V = E(V) = V_0 = \begin{bmatrix} v_0 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ v_0 \end{bmatrix} \quad (6)$$

$$\begin{aligned} \sigma_V^2 &= \text{Var}(V) = E((V - E(V))(V - E(V))^T) \\ &= \begin{bmatrix} \text{cov}(V_1, V_2) & \text{cov}(V_1, V_2) & \text{cov}(V_1, V_3) & \dots & \text{cov}(V_1, V_M) \\ \text{cov}(V_2, V_1) & \text{cov}(V_2, V_2) & \text{cov}(V_2, V_3) & \dots & \text{cov}(V_2, V_M) \\ \text{cov}(V_3, V_1) & \text{cov}(V_3, V_2) & \text{cov}(V_3, V_3) & \dots & \text{cov}(V_3, V_M) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \text{cov}(V_M, V_1) & \text{cov}(V_M, V_2) & \text{cov}(V_M, V_3) & \cdot & \text{cov}(V_M, V_M) \end{bmatrix} \end{aligned} \quad (7)$$

The $\text{cov}(V_i, V_j)$ is defined in (5). From (1), we know that X, Y, Z are normal random variable so V is a multivariate Gaussian process with mean μ_V and σ_V^2 .

2.5 Gate Length Spatial Correlation

The final $\rho = \rho_x \times \rho_y$. The result is shown in Figure 1.

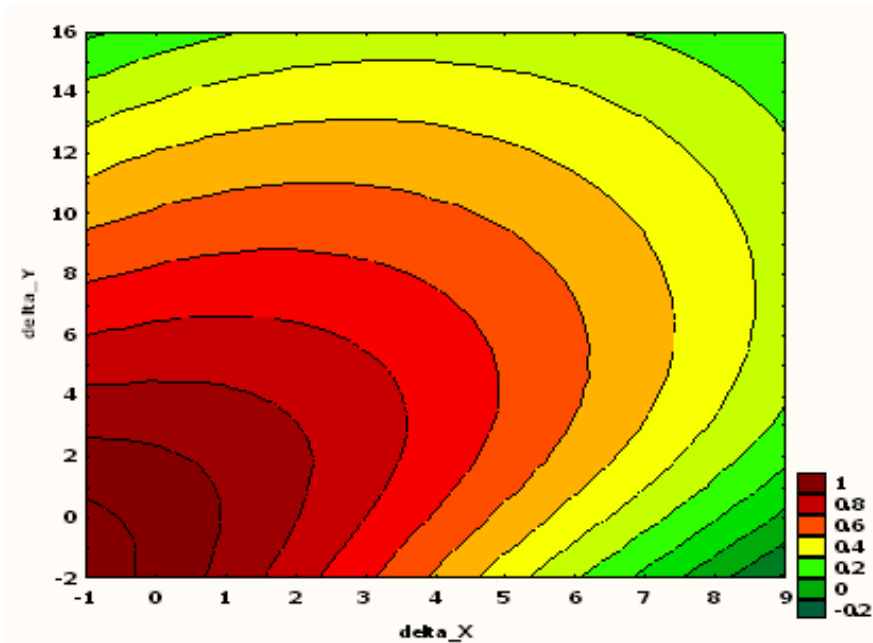


Figure 1. Spatial Correlation Simulation

There are other three commonly used spatial correlation function types: Linear, Exponential, and Gaussian

$$\begin{aligned}
 r(h) &= \max\{0, 1 - \alpha^2 \|h\|\} && \text{- Linear} \\
 r(h) &= \exp(-\alpha^2 \|h\|) && \text{- Exponential} \\
 r(h) &= \exp(-\alpha^2 \|h\|^2) && \text{- Gaussian}
 \end{aligned} \tag{10}$$

where h is (dx, dy) pair distance between two measure points and a is a fitting parameter.

In the exponential spatial correlation function, the distance can be separated into x distance dx and y distance dy . We can rewrite the equation to

$$r(h) = \exp(-\alpha^2 \|h\|) = \exp(-\alpha^2 \|dx\| - \alpha^2 \|dy\|) = \exp(-\alpha^2 \|dx\|) \bullet \exp(-\alpha^2 \|dy\|) \tag{11}$$

Let ρ_1 to be the correlation between samples separated by one unit, ρ_2 to be the correlation between samples separated by two units, and so on.

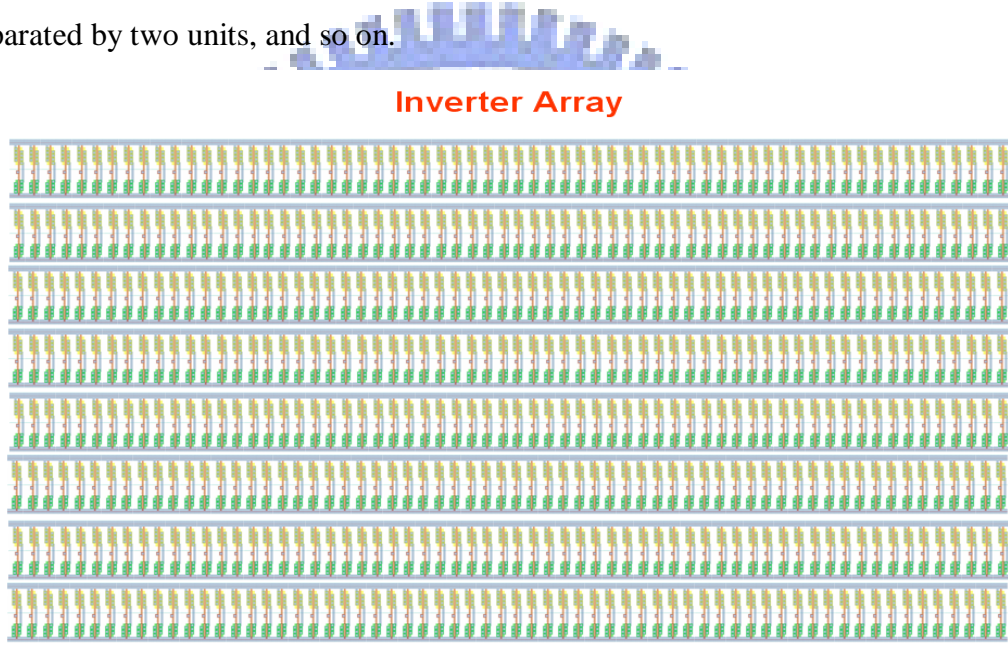


Figure 2. Inverter trend array

Suppose we have n by n inverter arrays as Figure 2. The distance between any two neighboring inverters have the same one unit distance. From this hypothesis, (10) can be expressed as

$$r(h) = \exp(-\alpha^2 \|dx\|) \bullet \exp(-\alpha^2 \|dy\|) = \rho_{dx} \bullet \rho_{dy} = \rho_{dx+dy} = \rho^{dx+dy} \tag{12}$$

where ρ is defined to $\rho_1 = \rho^1$ and $\rho_2 = \rho^2$. The gate length spatial correlation matrix for n by n inverter arrays is

$$\rho_{INV_{n*n}} = \rho(INV_{n*n}) = \rho \left(\begin{matrix} INV_1 \\ \cdot \\ \cdot \\ \cdot \\ INV_{n*n} \end{matrix} \right)$$

$$\begin{aligned}
&= \begin{bmatrix} \rho(\text{INV}_1, \text{INV}_1) & \rho(\text{INV}_1, \text{INV}_2) & \rho(\text{INV}_1, \text{INV}_3) & \dots & \rho(\text{INV}_1, \text{INV}_{n^*n}) \\ \rho(\text{INV}_2, \text{INV}_1) & \rho(\text{INV}_2, \text{INV}_2) & \rho(\text{INV}_2, \text{INV}_3) & \dots & \rho(\text{INV}_2, \text{INV}_{n^*n}) \\ \rho & \rho(\text{INV}_3, \text{INV}_1) & \rho(\text{INV}_3, \text{INV}_2) & \rho(\text{INV}_3, \text{INV}_3) & \dots & \rho(\text{INV}_3, \text{INV}_{n^*n}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho(\text{INV}_{n^*n}, \text{INV}_1) & \rho(\text{INV}_{n^*n}, \text{INV}_2) & \rho(\text{INV}_{n^*n}, \text{INV}_3) & \dots & \rho(\text{INV}_{n^*n}, \text{INV}_{n^*n}) \end{bmatrix} \\
&= \begin{bmatrix} 1 & \rho_{1+0} & \rho_{2+0} & \dots & \rho_{(n-1)+(n-1)} \\ \rho_{1+0} & 1 & \rho_{1+0} & \dots & \rho_{(n-2)+(n-1)} \\ \rho_{2+0} & \rho_{1+0} & 1 & \dots & \rho_{(n-3)+(n-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{(n-3)+(n-1)} & \rho_{(n-3)+(n-1)} & \rho_{(n-3)+(n-1)} & \dots & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{2n-2} \\ \rho & 1 & \rho & \dots & \rho^{2n-3} \\ \rho^2 & \rho & 1 & \dots & \rho^{2n-4} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho^{2n-2} & \rho^{2n-3} & \rho^{2n-4} & \dots & 1 \end{bmatrix} \tag{13}
\end{aligned}$$

From this analysis, we see that the correlation between Inverters closer together is weighted more strongly.

2.6 Path-Base Spatial Correlation

Given a n -stage path, the delay of each stage i can be characterized as a normal distribution $d_i \sim N(\mu_i, \sigma_i^2)$ for $1 \leq i \leq n$. The path delay P_d is the summation of these n normal distributions: $P_d \sim N(\mu, \sigma^2)$ where $\mu = \sum \mu_i$ and $\sigma^2 = \sum \sigma_i^2 + 2 \sum \rho_{ij} \sigma_i \sigma_j$. The ρ_{ij} denotes the correlation between the delay d_i and delay d_j , where $\rho_{ij} = 1$ for $i = j$ and ρ_{ij} for $i \neq j$. If $\rho_{ij} = 0$ for all $i \neq j$, i.e. delays are mutually independent. More generally speaking, given n delay random variables $d_1 \dots d_n$, their correlations can be specified with a symmetric positive semi-definite $n \times n$ correlation matrix $\rho = [\rho_{ij}]$.

Given two paths, P_x and P_y , and their gate delays, x_i and y_i , $i=1, \dots, n$. We can derive their correlation as below.

$$\begin{aligned}
\text{Corr}(P_x, P_y) &= \frac{\sum \sum \text{Cov}(x_i, y_j)}{\sigma_{P_x} \sigma_{P_y}} \\
\text{Cov}(x_i, y_j) &= \rho[x_i, y_j] \sigma_{x_i} \sigma_{y_j} \\
\sigma_{P_x} &= \sum \rho[x_i, x_j] \sigma_{x_i} \sigma_{x_j} \\
\sigma_{P_y} &= \sum \rho[y_i, y_j] \sigma_{y_i} \sigma_{y_j} \\
C_{P_x, P_y} &= \text{Corr}(P_x, P_y) = C(\rho, \sigma_x, \sigma_y)
\end{aligned} \tag{14}$$

If σ_x 's, σ_y 's are known, we have a function in terms of correlations $C(\rho)$. The Figure 3 illustrates the two delay path spatial correlation in the inverter array.

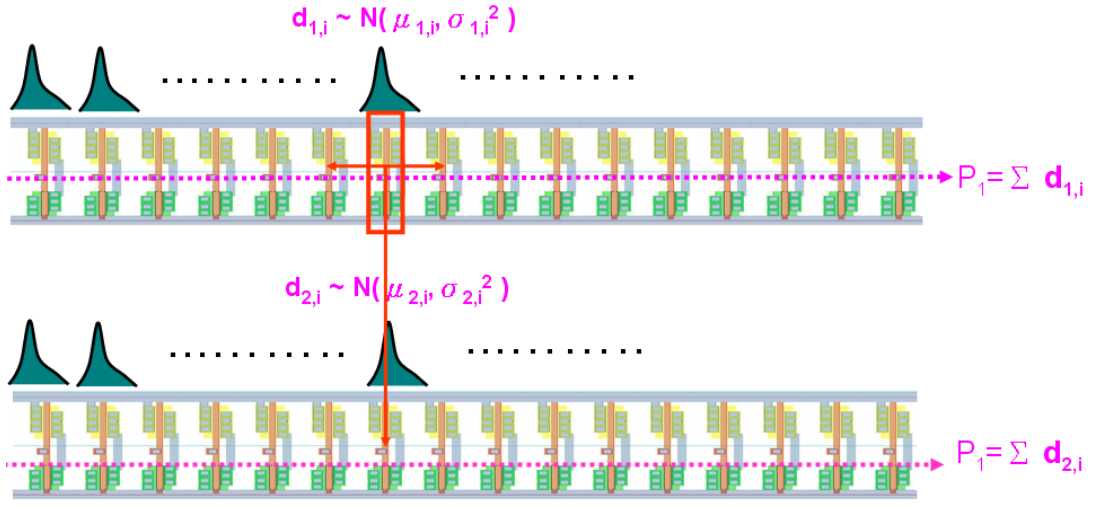


Figure 3. Path spatial correlation

2.7 The Path-Base Spatial Correlation Coefficient

In the path learning step, we select two paths P_a and P_b on the chip and measure their path delays. Suppose we get $D_a = [x_1, \dots, x_m]$ and $D_b = [y_1, \dots, y_m]$ on m silicon data. If path delays of P_a and P_b form the bivariate normal distribution, we can apply the Pearson's correlation coefficient as

$$r_{ab} = \frac{\sum_i (x_i y_i) - (\sum_i x_i)(\sum_i y_i) / t}{\sqrt{\sum_i (x_i^2) - (\sum_i x_i)^2 / t} \sqrt{\sum_i (y_i^2) - (\sum_i y_i)^2 / t}} \quad (15)$$

Each r_{ab} is a measured path correlation and r_{ab} is maximum likelihood estimation [17]. If all path delays form a multivariate normal distribution, we can apply this way to find out the spatial correlation matrix. One may want to know three things: (1) how confident is the Pearson's correlation coefficient? (2) How many samples the Pearson's correlation coefficient need? (3) How to do if the path delay distribution is not a normal distribution? We will discuss them in next chapter

Chapter 3

Bootstrap Confidence Intervals for the Correlation Coefficient

3.1 Bootstrap Concept

Bootstrap is a data based simulation method for statistical inference. The basic idea of bootstrap is to use the sample data to compute a statistic and to estimate its sampling distribution, without any model assumption. The usual confidence interval, based on Student's t distribution, is not appropriate due to the bias and skewness. The maximum likelihood estimators are dependent on an asymptotic distribution and may not be suitable for small samples. Most sampling distribution is unknown. The bootstrap uses the data and re-sampling to estimate that unknown sampling distribution.

Given a set of independent and identically distributed (iid) observations $X_i, i=1 \dots n$. A parameter is defined as the function of the values in the population $\theta = T(x)$, a statistic is defined as the function of the observations $\hat{\theta} = T(x)$, and the bootstrap estimates the sampling distribution $F_{\theta}(x)$ of that function. The data are used to estimate the unknown cumulative distribution function $F_{\theta}(x)$ of values in the population. Bootstrap samples are repeatedly drawn from the estimated population and give a set of bootstrap values $\hat{\theta}_i^* i=1, \dots, m$. The empirical distribution of those bootstrap values $\hat{F}_b(x)$ estimates the theoretical sampling distribution $F_{\theta}(x)$.

The bootstrap distribution $\hat{F}_b(x)$ is used to estimate bias, estimate a standard error (SE) and construct a confidence interval for the statistic of interest. The B_b is the bootstrap bias estimation, and s_b is the bootstrap SE from m bootstrap values:

$$B_b = \frac{\sum_{i=1}^B (\hat{\theta}_i^* - \hat{\theta})}{B} \quad (16)$$

$$s_b = \left[\frac{\sum_{i=1}^B (\hat{\theta}_i^* - \bar{\hat{\theta}}^*)}{B} \right]^{1/2} \quad (17)$$

The procedure of Bootstrap is

- (1) Given $x = (x_1, x_2, \dots, x_n)$ the (original-)sample.
- (2) Draw B independent samples $x_1^*, x_2^*, \dots, x_B^*$, all containing n data-values of x based on drawing with replacement
- (3) Calculate to each bootstrap-sample statistic of interest:

$$\hat{\theta}_i^* = T(x_i^*) \quad i= 1, 2, \dots, B$$

- (4) Calculate (10) and (11)

3.2 Balanced Bootstrap

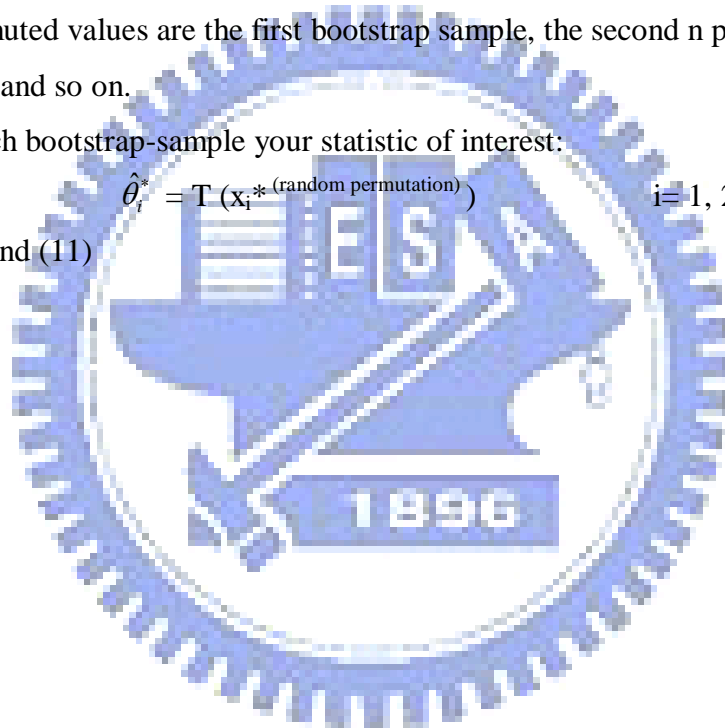
The balanced bootstrap is a selecting sampling method that can increase the precision of the bootstrap bias and SE. Under balanced bootstrap sampling, samples are generated in such a way that each original data point is present exactly B times in the entire collection of nB samples. Balanced bootstrap samples can be generated by constructing a population with B copies of each of the n observations, then randomly permuting that population. The first n permuted values are the first bootstrap sample, the second n permuted values are the second sample, and so on. The procedure of Balanced Bootstrap is

- (1) Given $x = (x_1, x_2, \dots, x_n)$ the (original-)sample.
- (2) Draw B independent samples $x_1^*, x_2^*, \dots, x_B^*$, all containing n data-values of x based on drawing with replacement. n copies this. Then, we have total nB samples. Randomly permute nB samples. The first n permuted values are the first bootstrap sample, the second n permuted values are the second sample, and so on.

- (3) Calculate to each bootstrap-sample your statistic of interest:

$$\hat{\theta}_i^* = T(x_i^* \text{ (random permutation)}) \quad i=1, 2, \dots, B$$

- (4) Calculate (10) and (11)



3.3 Bootstrap Confidence interval

Many methods have been proposed for constructing bootstrap confidence intervals [18]. We will discuss the most popular one: the so-called percentile method, bias-corrected and accelerated method and normal method

3.3.1 The Percentile Method

We can define a $1 - \alpha$ confidence interval for θ (For example, $\alpha = .05$ denotes 95% confidence.) The interval between the 2.5% and 97.5% percentiles of the bootstrap distribution of a statistic is a 95% bootstrap percentile confidence interval for the $\hat{F}_b(x)$. That is, we sort the B values of $\hat{\theta}_i^*$ and throw out lowest $(B+1)\alpha/2$ values, as well as the highest $(B+1)\alpha/2$ values. The flowchart is shown in the Figure 4.

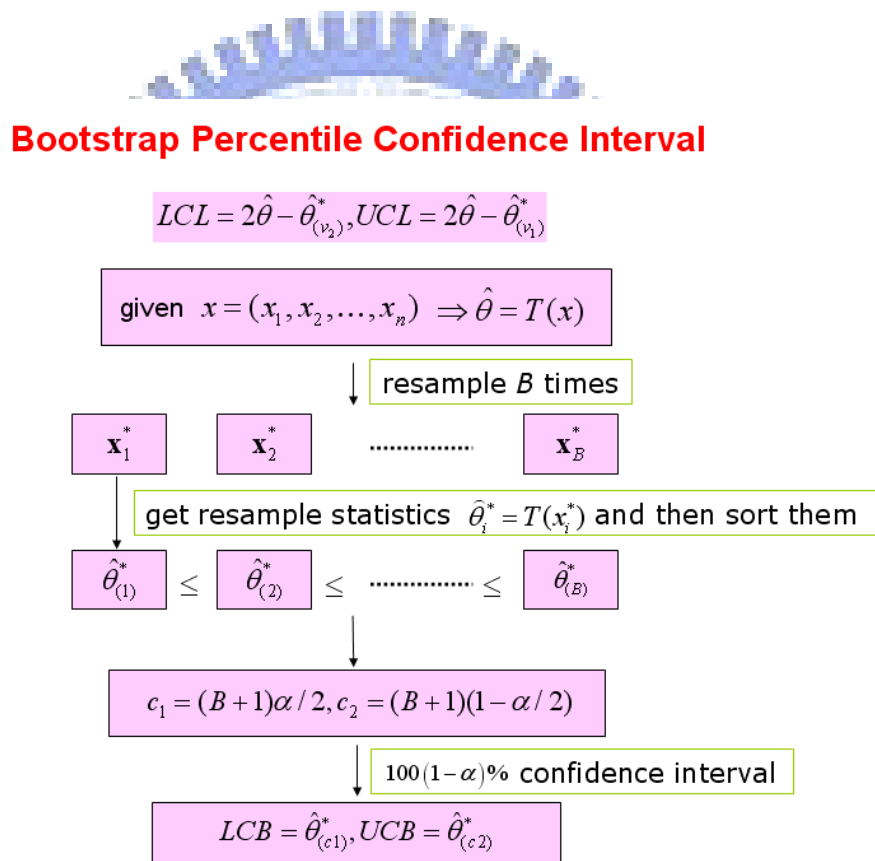


Figure 4. Flowchart of Bootstrap percentile confidence interval

3.3.2 The Bias-Corrected and Accelerated Method

The percentile bootstrap endpoints are easy to get and can work well, especially if the sampling distribution is symmetrical. The percentile bootstrap confidence intervals may have bad coverage when the sampling distribution is skewed. Efron (1987) proposed the bias-corrected and accelerated (BCA) bootstrap method to correct for the percentile bootstrap. The BCA method allows for the distribution to be with around an unknown constant (i.e., “biased”) and with non-constant variance (i.e., the variance tends to “accelerate” across values in the sampling distribution). This method adjusts the endpoints of the bootstrap distribution for bias and non-constant variance. The flowchart is shown in the Figure 5.

Bootstrap Bias-Corrected and Accelerated Confidence Interval

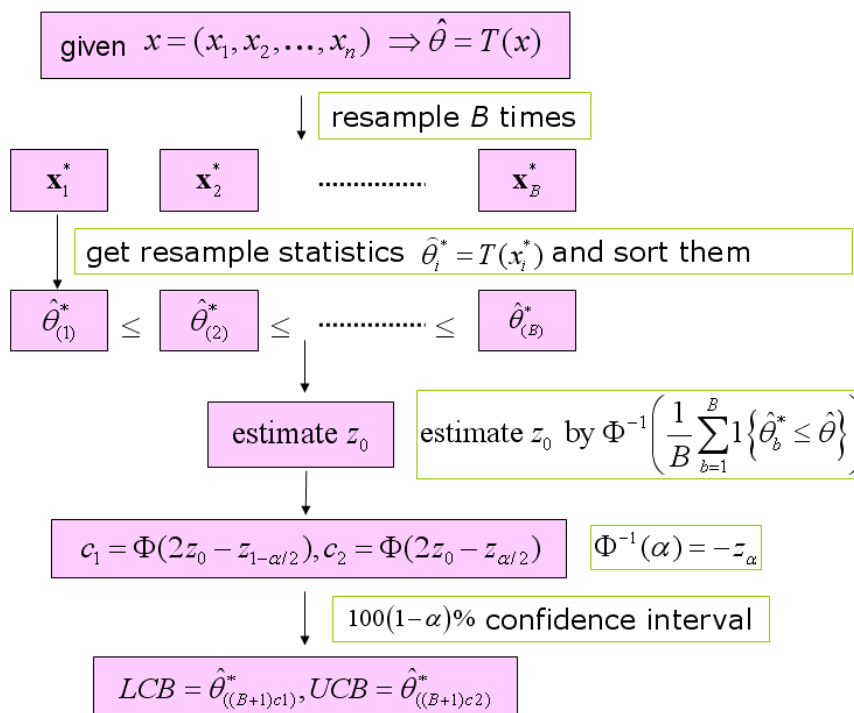


Figure 5. Flowchart of Bias-Corrected and Accelerated confidence interval

3.3.3 The Normal Method

When the distribution of $\hat{\theta}$ is asymptotically normal, we can get a confidence interval from the familiar normal distribution. The estimation $\hat{\theta}$ is the estimated from the original data and the estimate, $\hat{\theta}_i^*$, is from each bootstrap sample. The $1 - \alpha/2$ and $\alpha/2$ quantile confidence interval is

$$[\hat{\theta} - B_b - z_{(1-\alpha/2)}s_b, \hat{\theta} - B_b - z_{(\alpha/2)}s_b] \quad (18)$$

The B_b is the bootstrap bias estimation, and s_b is the bootstrap SE from (16) and (17). The flowchart is shown in the Figure 6.

Bootstrap Normal Confidence Interval

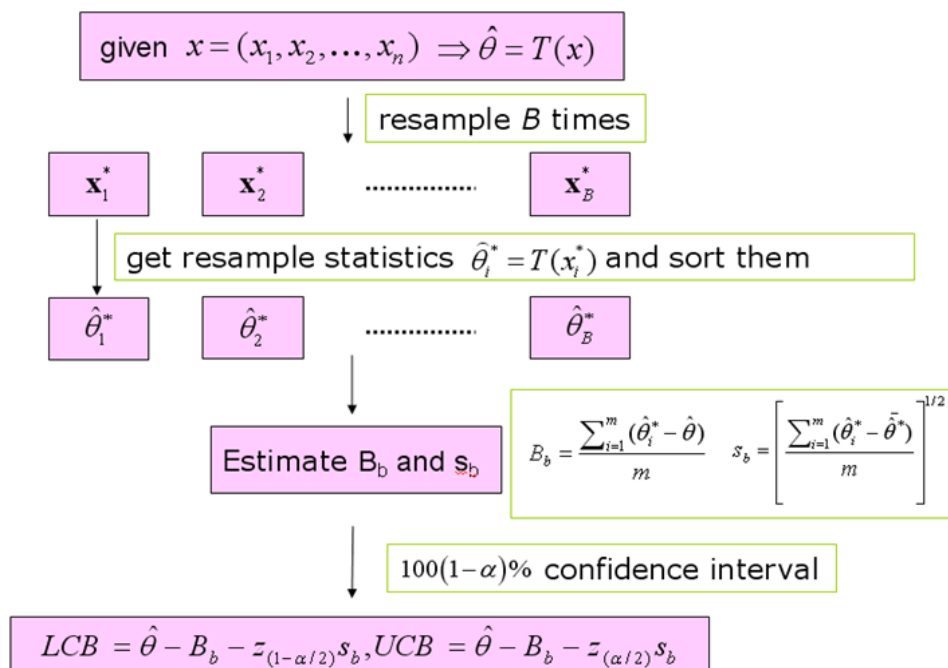


Figure 6. Flowchart of Bias-Corrected and Accelerated confidence interval

3.4 Bootstrap Confidence Intervals for the Correlation Coefficient

Pearson's correlation coefficient, r_{xy} , is often used to measure the association between two sampled (data size n) variables, x and y . For binormally distributed data, Fisher's (1921) transformation $z = 0.5 \log\left(\frac{1+r_{xy}}{1-r_{xy}}\right)$ yields z being asymptotically distributed with variance $1/(n-3)$. However, most cases are nonnormal data distributions and make normal-theory-based intervals questionable. The bootstrap is a good candidate to solve these problems. There are no distributionary assumptions because it uses the data to simulate the distribution. We can use it to find a confidence interval for the correlation coefficient of two path delays.

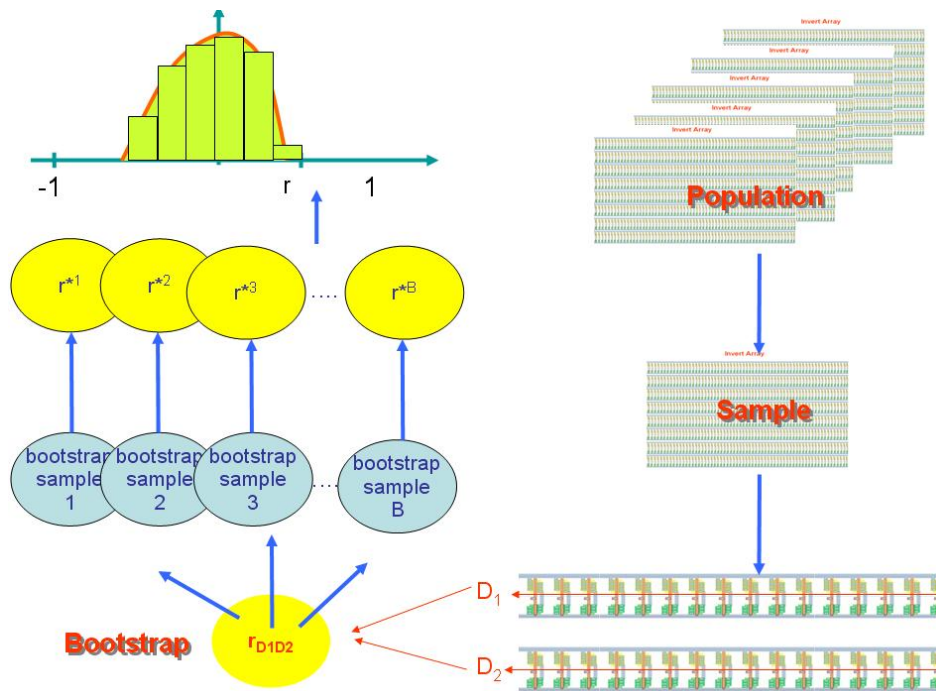


Figure 7. Flow of silicon data bootstrap

From Figure 7, we measure two path delays and use bootstrap re-sampling. We can get an approximation of the population distribution. If the path delay is approximated normal, we can use Bootstrap normal confidence interval to estimate spatial path delay correlations. If the path delay is skewed, we can use Bootstrap Bias-Corrected and Accelerated confidence interval to estimate delay correlations. If we know nothing about path delay, we can use Bootstrap percentile confidence interval to estimate delay correlations.

Chapter 4

Experimental Results

4.1 Experiment Setup

We design a novel test chip to run HSPICE to simulate path delay spatial correlations. Many papers propose the Ring Oscillator (RO) to find spatial process variation [14]. To get an efficient experiment result, we use the multiplexor with a 5 by 5 inverter trend in an inverter array. Each inverter trend has 900 inverters and the inverter trend array is 1640 um by 1640 um. The layout is shown in Figure 8. The inverter trend array can reduce process variations including photo impact, etch impact, CMP impact and micro loading impact. We simulate the 5 by 5 inverter trend delay with HSPICE simulations.

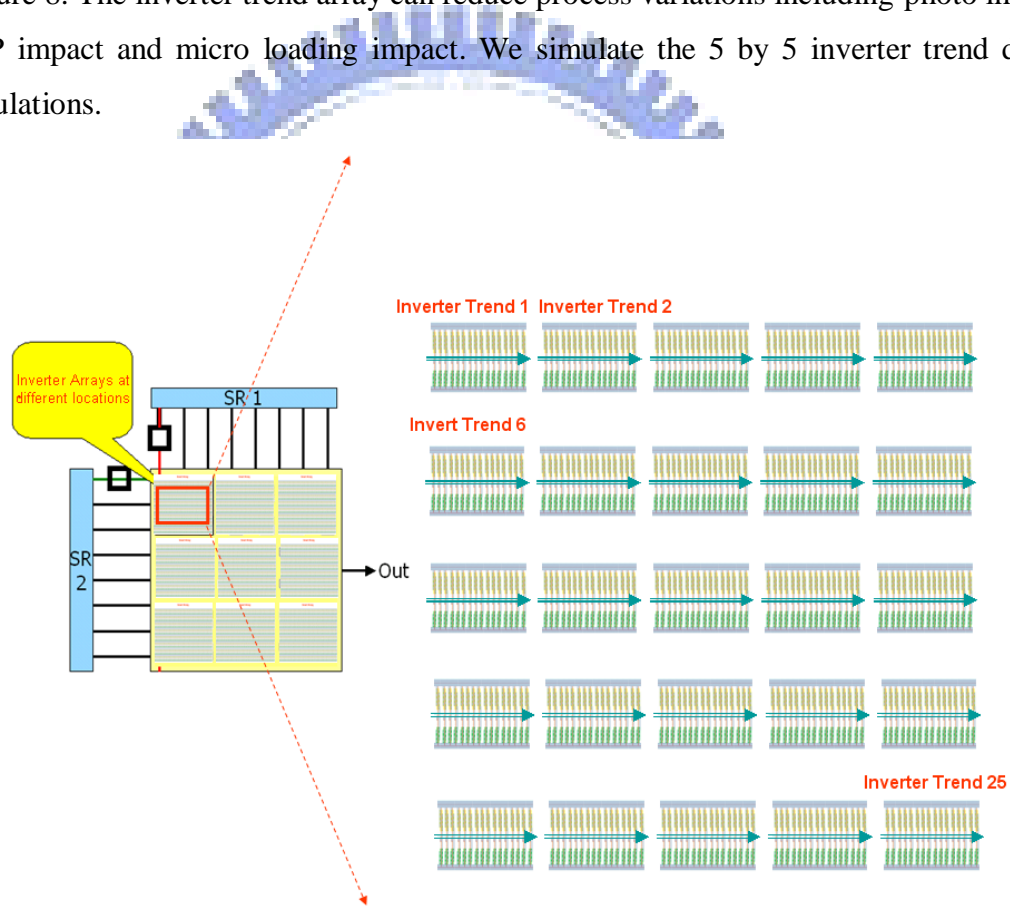


Figure 8. Test Chip layout with multiplexor

4.2 Spatial Path Correlations with Multivariate Normal Gate Length Variations

To detail understanding the correlation of CD variation and delays, we first simulate the single inverter delay with gate length CD variations. Suppose the gate length variation is normal and its histogram is shown in Figure 9. Its delay after Hspice simulation is shown in Figure 10 and is high correlated to CD variations

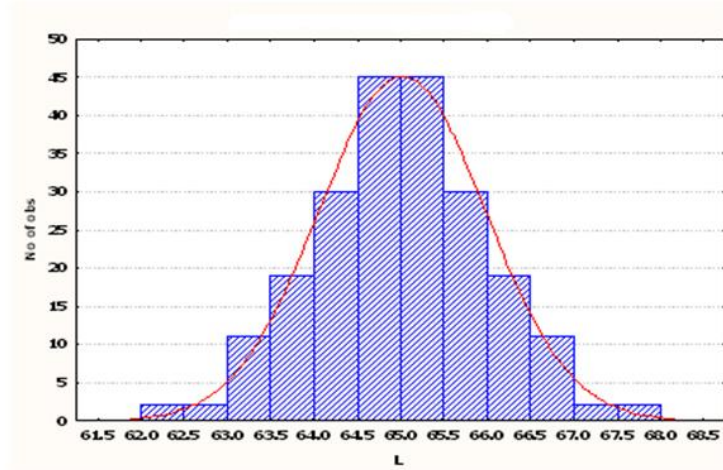


Figure 9. Gate length histogram

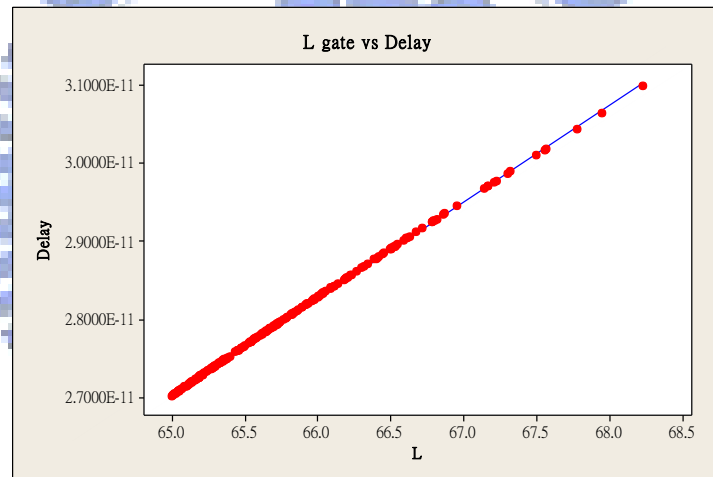


Figure 10. A single inverter delay simulation with gate CD variations

For the 5 by 5 inverter trend, the gate CD variation is multivariate normal distribution with correlation matrix

$$\begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{2(n-1)} \\ \rho & 1 & \rho & \dots & \rho^{(n-2)+(n-1)} \\ \rho^2 & \rho & 1 & \dots & \rho^{(n-3)+(n-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho^{2(n-1)} & \rho^{(n-2)+(n-1)} & \rho^{(n-3)+(n-1)} & \dots & 1 \end{bmatrix} \quad (19)$$

The correlation matrix is from (15) where $n=25$ & $\rho=0.98$.

To validate the accuracy of path base spatial correlation, we run Monte Carlo to generate 100000 patterns and set it as a golden set. The path delay distribution of the golden set is shown in Figure 11.

It shows inverter trend 1 and 2 has higher correlation than inverter trend of 1 and 3 due to the spatial distance. Close path distance has higher correlations.

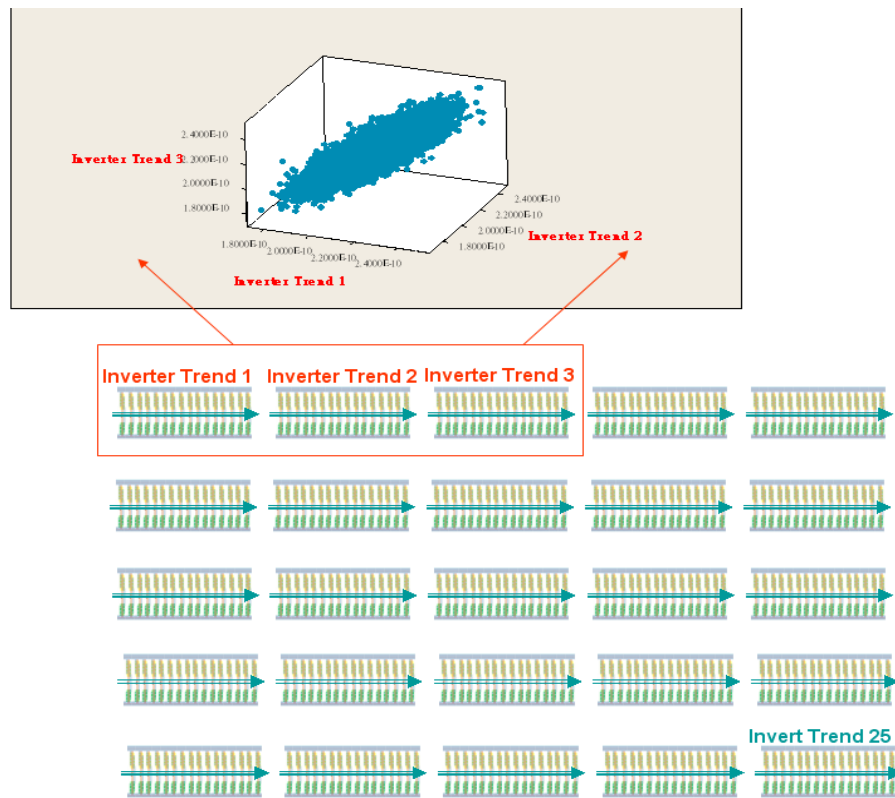


Figure 11. Path delay correlation distributions of inverter trend 1, 2 and 3

We select 100 samples from the golden set and run 2000 bootstrap re-sampling for the 5 by 5 inverter trend. The delay spatial correlation of this is in Figure 12.

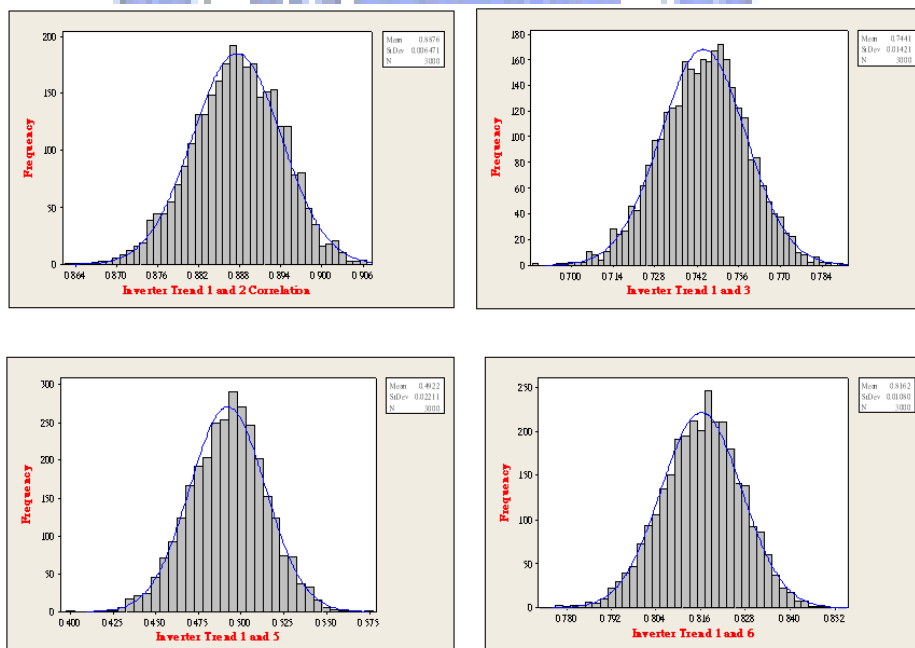


Figure 12. The Histograms of correlation coefficient of Bootstrap of 5 by 5 inverter trend

The histogram is approximated as normal distribution. Inverter trend 1 and 6 (in Figure 8) are most close and have the highest correlation distribution. The correlation of inverter trend 1 and 5 is lower due to the distance of them is far. The histogram variation of them is wide.

Confidence interval coverage rate is the probability that the confidence interval includes the true parameter from the population. If the coverage rate is the same as the stated size of the confidence interval, the intervals are accurate. We simulate 5 kinds of bootstrap confidence interval: Percentile method, Bias-Corrected and Accelerated method, Normal method, Normal method with outlier filter and Fisher Z. The normal with outlier filter is normal method excluding outlier with 3 standard variations.

The coverage rates are listed in Figure 13 with population size 100000, sample size 3000 and confidence $\alpha=95\%$

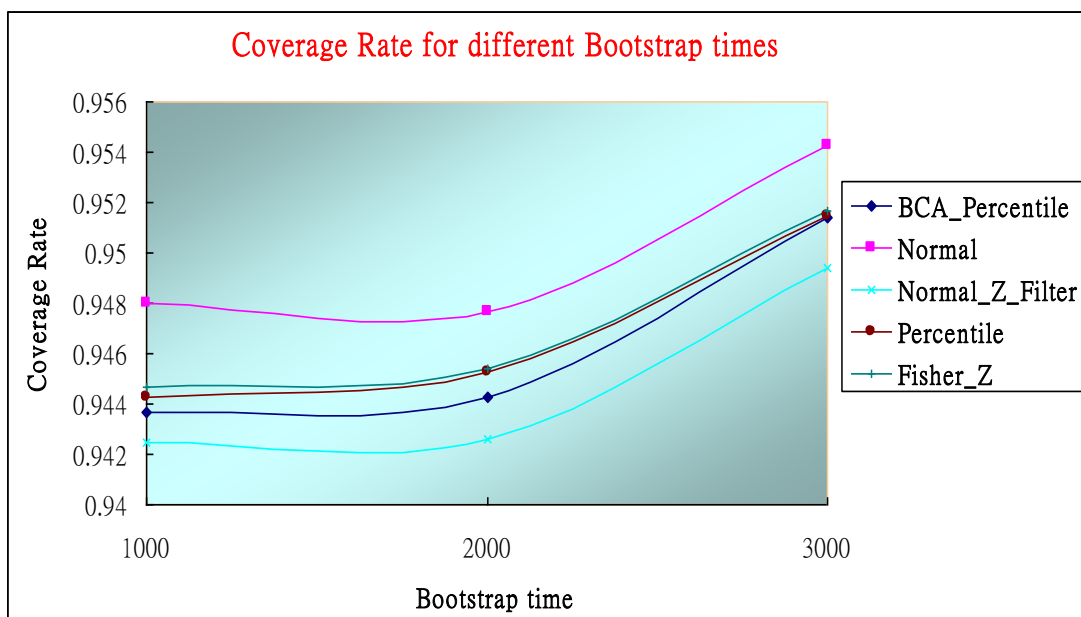


Figure 13. Coverage rate for BCA, Percentile, Normal, Normal with filter and Fisher Z

The coverage rate is around 95% under confidence $\alpha=95\%$. The Bootstrap normal confidence interval is better than other bootstrap ones due to normal distributions of gate length variations. The Fisher Z confidence interval has the same performance with Bootstrap methods. This is also because normal gate length variations. From Figure 12, we investigate the coverage rate is better as bootstrap times increase.

The correlations of path delays to delta x direction distance and delta y direction distance are listed in the Figure 14 and 15. The delta x distance is two inverter trend distance in the x direction and the delta y distance is two inverter trend distance in the y direction. The path delay correlation decrease as the path distance increase. When the trends are closer, the bound is smaller. This is because long distances have low correlations and cause more variations from them.

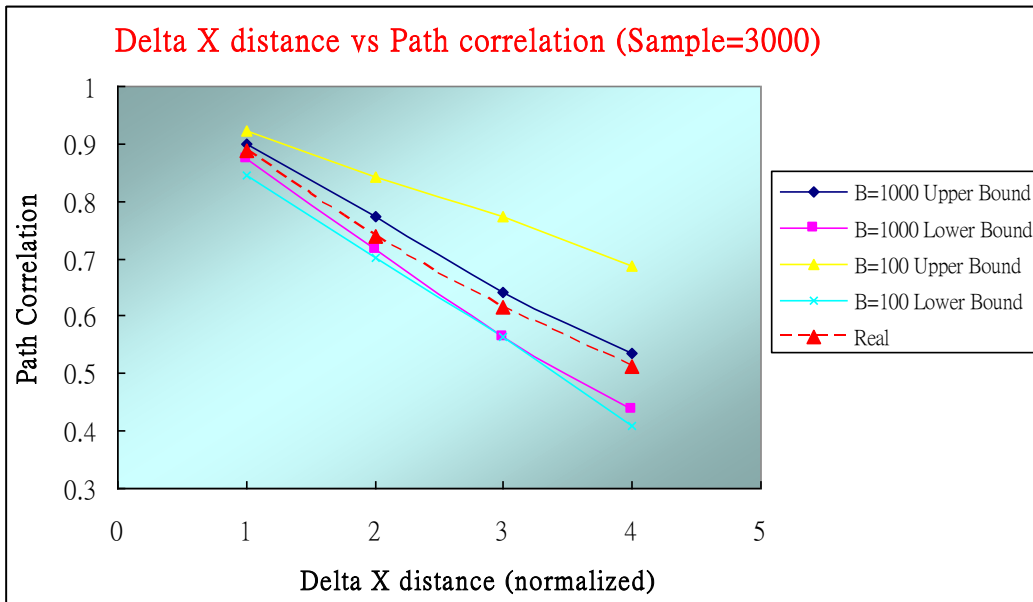


Figure 14. Confidence interval bound for delta x distance

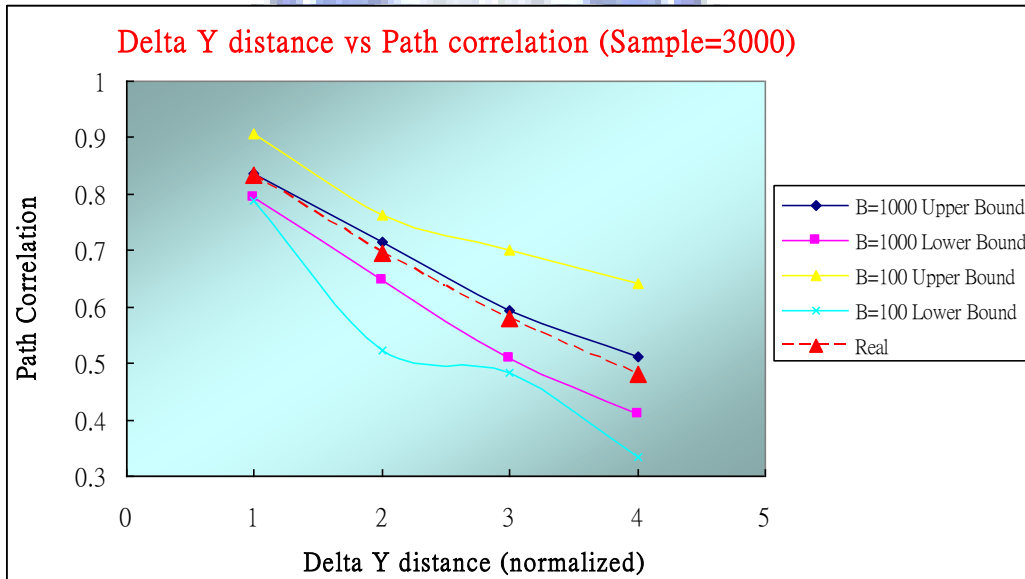


Figure 15. Confidence interval bound for delta Y distance

4.3 Correlations with Multivariate Chi Square Gate Length Variations

Most path delays are not normal distributed. We demonstrate a gate length variation with chi square distribution. The Figure 16 shows the histogram of a single gate length variation.

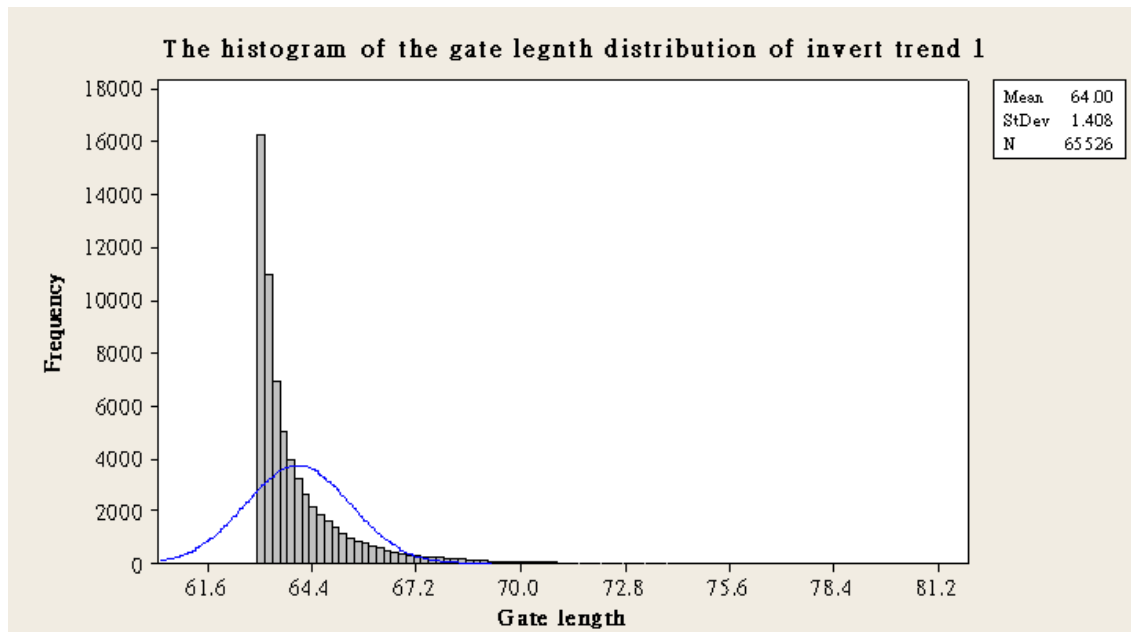


Figure 16. The histogram of a single gate length variation with chi square distribution

For the inverter trend array, multi gate length variations are multivariate chi square distribution with the same correlation matrix (19) and $\rho=0.98$. We use Bootstrap methodology to resample the correlation coefficient of the path delay. The result is shown in Figure17. The histogram shows the correlation coefficient distribution is non-normal.

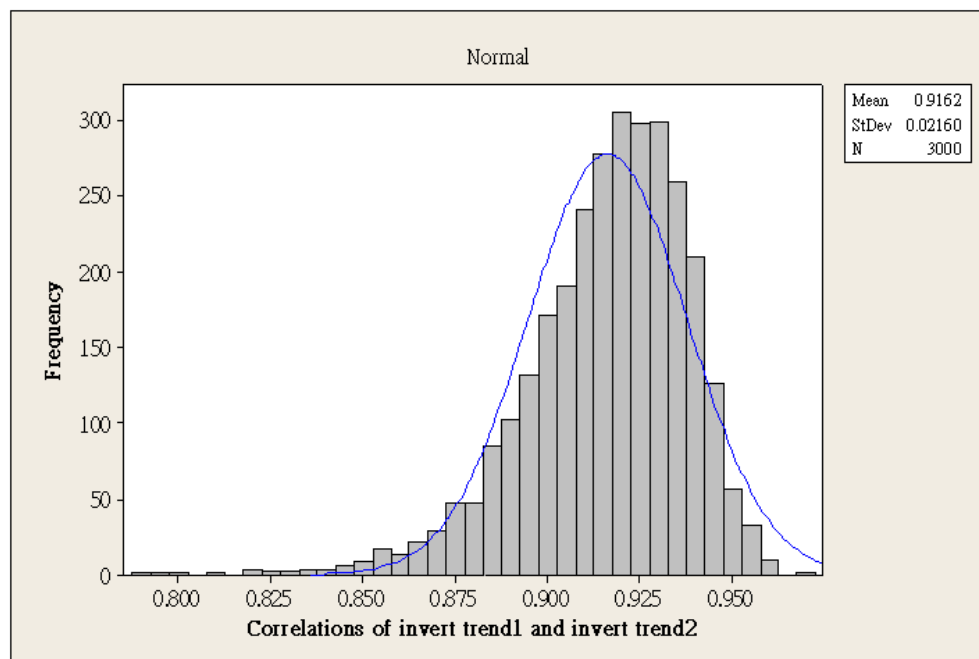


Figure 17. The histogram of the correlation coefficient of Bootstrap of 1 and 2 inverter trend

We simulate 5 kinds of Bootstrap confidence interval under the 95% confidence. The coverage rate is shown in Figure 18. The Bootstrap confidence intervals have 94.5% coverage rate but Fisher Z is lower about 93%. This is because the path delay is non-normal distribution and it is not suitable for Fisher Z confidence interval.

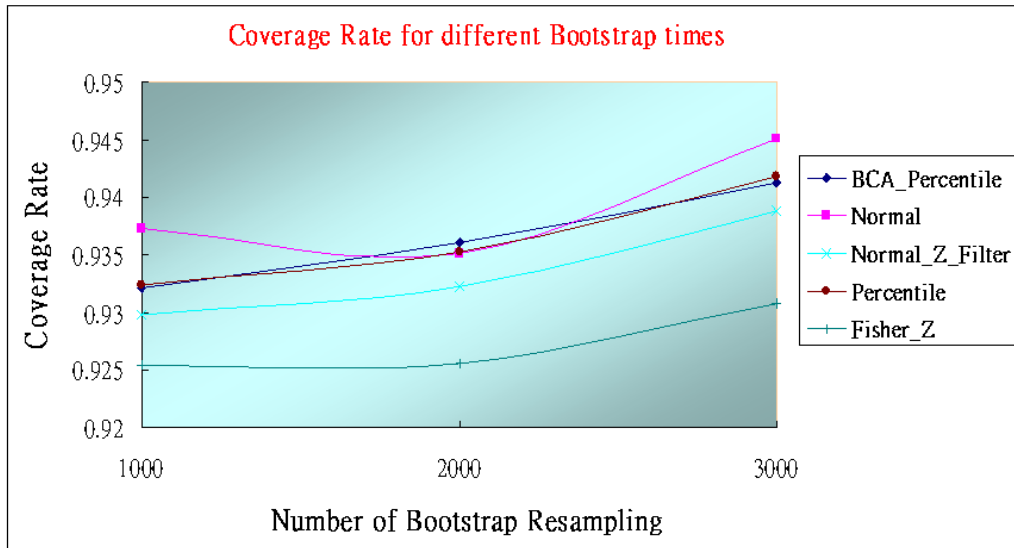


Figure 18. Coverage rate for BCA, Percentile, normal, normal with filter and Fisher Z

The correlations of path delays to delta x direction distances and delta y direction distances are listed in the Figure 19 and 20.

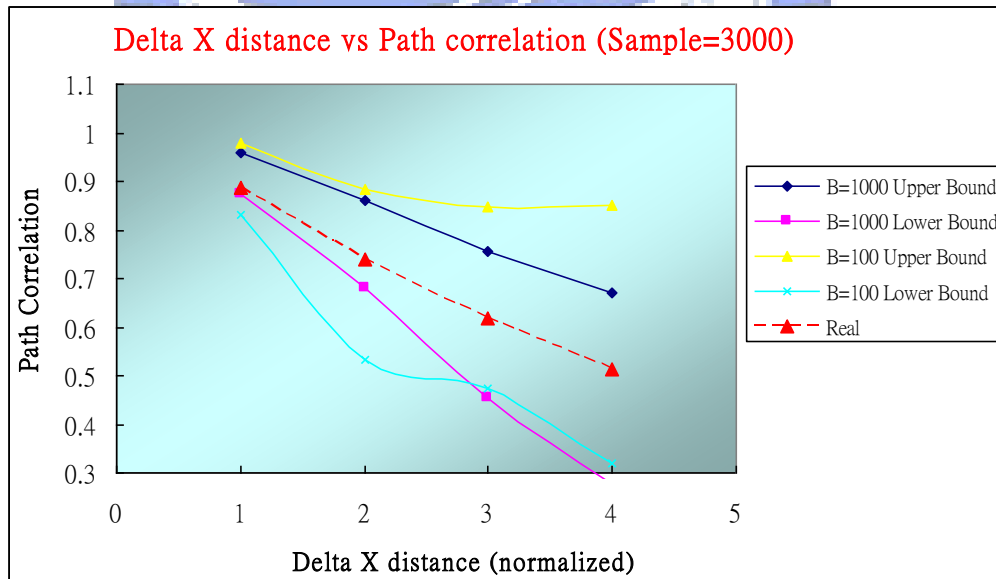


Figure 19. Confidence interval bound for delta x distance

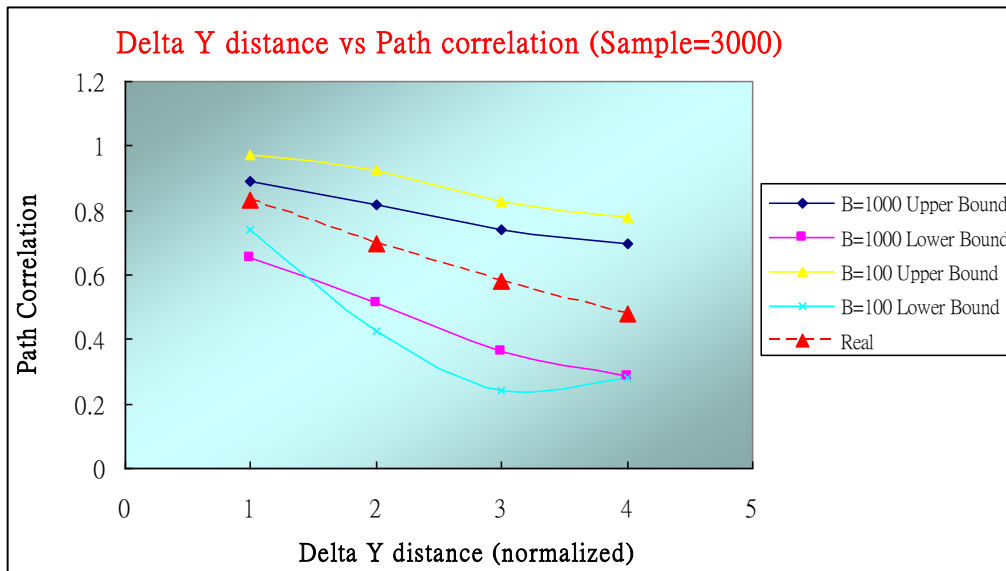
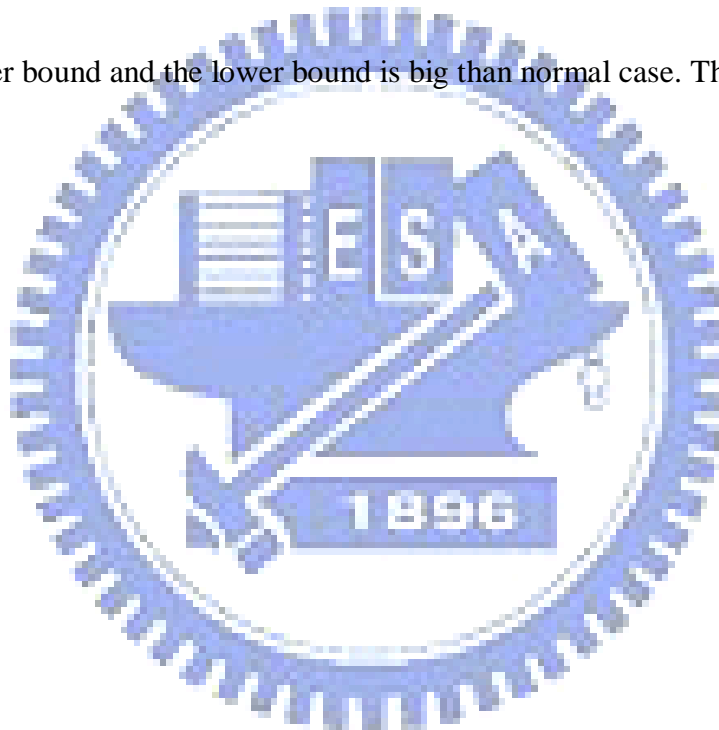


Figure 20. Confidence interval bound for delta Y distance

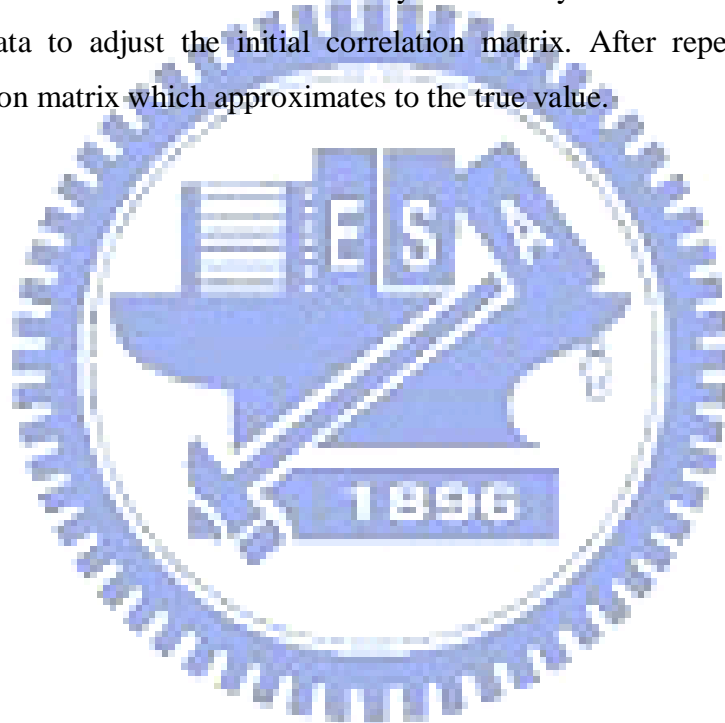
The gap of the upper bound and the lower bound is big than normal case. This is because the data are skewed seriously.



Chapter 5

Conclusion

In this thesis, we have applied bootstrap re-sampling algorithm and path-based learning methodology to estimate the spatial correlation of the path delay on the test chip. From this way, we can get a confidence interval for spatial correlation with a good coverage rate. It is also applied to non-normal correlation distribution. The plot of the path distance and the path delay correlation give us a predation way and give designer guidance for timing analysis. For the future work, we can select an initial correlation matrix which fall into Bootstap confidence intervals and put the initial correlation matrix to SSTA models. With the Bayesian theory or other learning methodologies, we use new silicon data to adjust the initial correlation matrix. After repeating this, we can get a converged correlation matrix which approximates to the true value.



Bibliography

- [1] M. O. and K. K., "A general probabilistic framework for worst case timing analysis," in Proceedings of the ACM/IEEE Design Automation Conference, New Orleans, Louisiana, USA, June 2002, pp.556–561.
- [2] H. C. and S. S., "Statistical timing analysis under spatial correlations," in IEEE transactions on computer-aided design of integrated circuits and systems, 2005, 24(9):1467-1482. Dd
- [3] B. N., "Refined statistical static timing analysis through learning spatial delay correlations," in DAC 2006: 149-154
- [4] A. A., V. Z., and D. T., "Statistical timing analysis using bounds and selective enumeration," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 22, pp. 1243.1260, Sept 2003.
- [5] A. A, D. B. and V. Z., "Statistical timing analysis for intra-die process variations with spatial correlations," in ICCAD, pp. 900.907, 2003.
- [6] L. Z., Y. H., C. C., "Statistical timing analysis with path reconvergence and spatial correlations," in the Design, Automation and Test in Europe, 2006. DATE '06.
- [7] C. V., K. R., K. K., S. G., and S. N., "First-order incremental block-based statistical timing analysis," in DAC, pp. 331.336, 2004.
- [8] P. F., W. C., and C. J., "Spatial Variability of Critical Dimensions," Proc. 22nd Int'l VLSI/ULSI Multilevel Interconnection Conf. (VMIC '05), Inst. Microelectronics Interconnection, 2005, pp. 539–546.
- [9] M. O and L. M., "Characterization of spatial intra field gate CD variability, its impact on circuit performance, and spatial mask-level correction," In IEEE Tran. on Semiconductor Manufacturing, vol 17, No 1, 2 - 11, 2004.
- [10] P. F, et. al, "Modeling Within-Die Spatial Correlation Effects for Process-Design Co-Optimization," Proc. of ISQED, 2005
- [11] V. K. and A. S., "A general framework for accurate statistical timing analysis considering correlations," In DAC, 2005.
- [12] P. G. and F. L., "Toward a Systematic-Variation Aware Timing Methodology", Proc. DAC, 2004
- [13] D. S., P. G., A. B., and J. Y., "Toward performance-driven reduction of the cost of RET-based lithography control", Proc. SPIE Conf. on Design and Process Integration for Microelectronic Manufacturing, Feb 2003.

- [14] J. C. and C. J., "Electrical linewidth metrology for systematic CD variation characterization and causal analysis," Proceedings of SPIE Int. Soc. Opt. Eng. 2003.
- [15] P. F., W. C., and C. J., "Spatial modeling of micron-scale gate length variation," in Data Analysis and Modeling for Process Control III
- [16] M. H., A. N., and C. S., "Ring Oscillator Sensitivity to Spatial Process Variation," Proc. 1st Int'l Workshop Statistical Metrology (IWSM 96), 1996
- [17] G. S. and R. B., "Statistical Inference", DUXBURY, 2002
- [18] B. E. and R. J., "An Introduction to the Bootstrap.", Chapman & Hall, New York
- [19] T. W., "An Introduction to Multivariate Statistical Analysis", JOHN WILEY & SONS52
- [20] J. D., D. K., S. L. , and J. K., "A unified statistical model for inter-die and intra-die process variation," in SISPAD, 2005

