國立臺灣大學電機資訊學院電子工程學研究所

碩士論文

Graduate Institute of Electronics Engineering

College of Electrical Engineering & Computer Science

National Taiwan University

Master thesis

考慮信號轉換時間之統計靜態時序分析

Slew-Aware Statistical Static Timing Analysis

劉繼蔚

Liu Chi-Wei

指導教授：陳中平 博士

Advisor: Chen,Charlie Chung-Ping, Ph.D.

中華民國 97 年 7 月

July, 2008

摘要

近年的研究趨勢中，統計分析已成為一廣泛受到注意與重視之主題。在本作品中, 藉由靜態時序分析此一應用，我們提供了另一種基於數學推論，可據以進行統計分析的觀點。以實驗模型為基本,利用統計靜態時序分析中的積分法，我們所提出的方法在其實驗模型滿足以下假設時可證明其正確性：(1)其模型滿足數學上之well-defined 的性質；(2)其實驗模型所定義之自變數為相互獨立；(3)所定義之自變數可分成二組無交集且無遺漏之分割，令之為 A_1 和 A_2，並且存在一映成函數，其定義域為待測之統計特性與 A_2 之聯集，而其值域為 A_1。

Abstract

Statistical Analysis draws much research attention in recent years. In this work, with the static timing analysis as target application, a mathematical analysis is made to provide another viewpoint of its statistical result. Starting from the experiment model, a statistical analysis approach based on the integration method is provided and proven to be exact with respect to the model under these requirements for the model: (1) the model is well-defined; (2) the model is based on mutually independent variables; (3) there is at least a bi-partition of independent variables, says A_1 and A_2, such that there's an onto function from the union of A_2 and properties to A_1.

Table of Content

List of Figures

Chapter 1

Introduction

Systematic yield model for process-induced uncertainty remains a challenge since its firstly identified as a challenge by International Technology Roadmap for Semiconductors (ITRS) in 2001[1]. Among all research topics involved, referring to the viewpoint from EETimes, statistical static timing analysis (SSTA) draws extensive discussion to be used for verification of the designs manufactured at 90 nm or below since DAC'05[2]. However, even SSTA itself does still not yet acquire well-acknowledged industrial success. In this work, a model-based statistical analysis approach is proposed. This approach would be proved to be exact with respect to the model under these requirements for the model: (1) the model is well-defined; (2) the model is based on mutually independent variables; (3) there is at least a bi-partition of independent variables, says A_1 and A_2, such that there's an onto function from the union of A_2 and properties to A_1. With this approach, some issues of recent path-based and block-based SSTA methodology are discussed.

Chapter 2

Preliminary and Related Work

The framework of statistical timing evaluation proposed in this work is based on the deterministic timing model. When talking about the deterministic model, it can be traced back to the previous work about static timing analysis. With the shrinking of the feature size, SSTA emerges. Two main branches, path-based SSTA and block-based SSTA are then described.

2.1. Static timing analysis (STA)

Static timing analysis (STA) is a widely-used method for performance evaluation in electronic design automation. In this section, no detailed or tedious concepts would be introduced. A sketch is made based on the idea proposed in [3, 4] by R.B. Hitchcock et

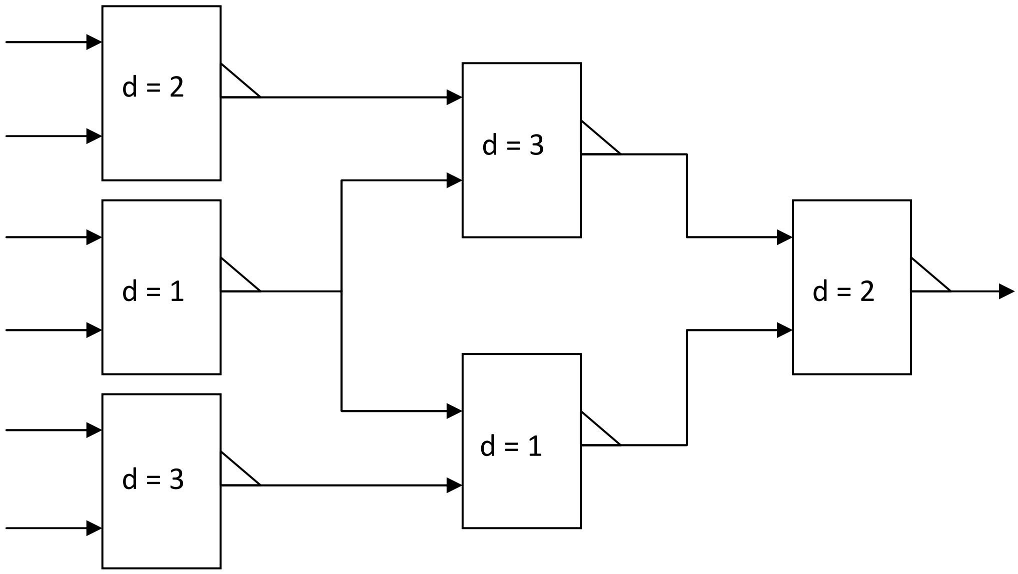


Fig. 2-1 A Sample Circuit

al. It's assumed that the delay is contributed by the gate. This assumption is still valid if interconnect is treated as a special kind of gate. An example is given as Fig. 2-1.

And it's obvious that the time-delay as an event could be modeled as an activity network. The arrival time (AT) could then be computed with the method by [5] and shown as Fig. 2-2. It's convenient to find that from this methodology, if defined AT(g) as the arrival time at the output of the specific gate g, $FAN_{in}(g)$ as the set of input cells of the gate, and $d(g,g_i)$ as the gate delay of the gate with respect some input signal from gate g_i:

$$AT(g) = \max\{d(g,g_i) + AT(g_i) \mid g_i \in FAN_{in}(g)\} \tag{2-1}$$

It's very important to clearly point out the two basic operations in this type of timing analysis: add operator and max operator. The add operator reflects the fact that the

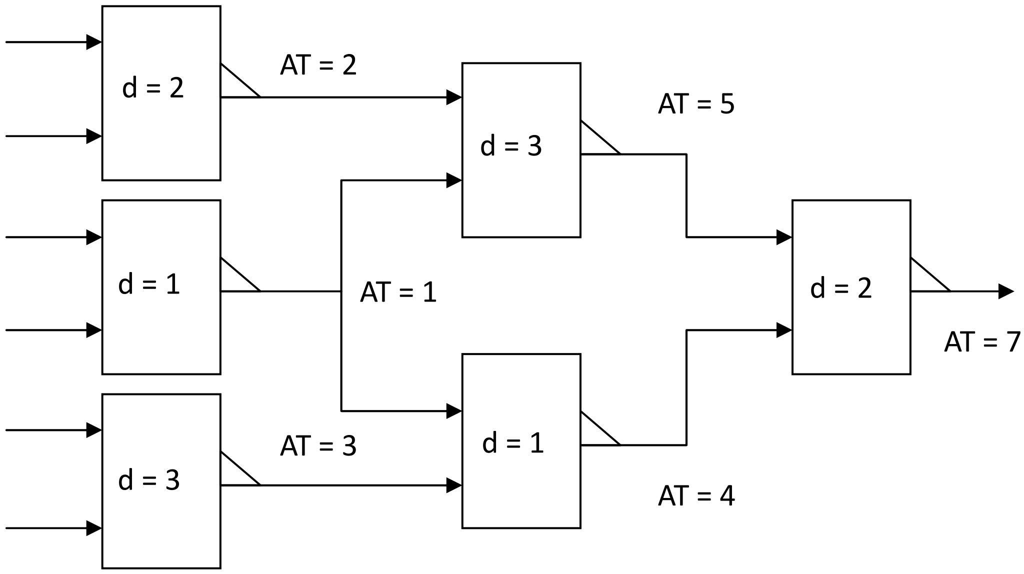


Fig. 2-2 STA Result

arrival time is the summation of sensitized gate delay. And the max operator is related to the concern of critical delay.

2.2. Statistical static timing analysis (SSTA)

This topic is not recently emerged one. SSTA could be traced back to some works over ten years such as [6]. The main difference is that the concerned delay or arrival time is no longer a deterministic value, but described with a distribution instead. For example, the sample circuit in Fig.2-1, now is assumed with the behavior illustrated as Fig.2-3.

Assuming a simplest but impractical property that every delay distribution and every possible summation of the delay distributions is independent, arrival time could be

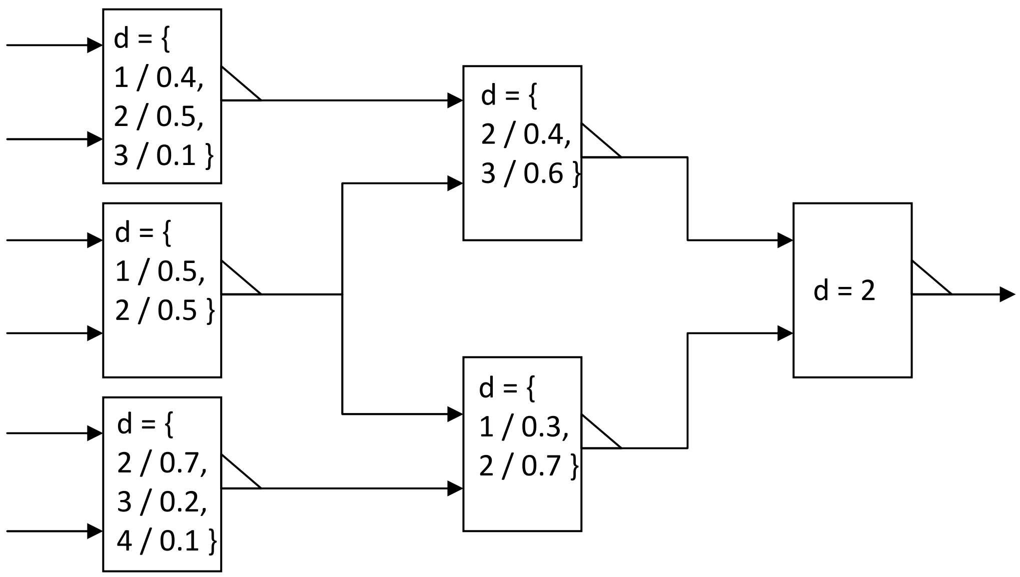


Fig. 2-3 Circuit with delay described in a distribution

found with the method mentioned in [7] as Fig. 2-4. The main idea in [7] is that the max operator for two random variables could be computed with the cumulative distribution function of one variable and the probability distribution function of the other variable. It directly copes with the distribution. Since both the add operation and max operation are defined, the SSTA goes almost the same as STA in equation 2-1. It must be noted that the result in Fig.2-4 is based on impractical assumptions. In general cases, the delays are correlated. Recalling to the cause of the distribution, [8] illustrate that we can relate the variation of timing properties to the variation of some design parameters. In this sense, there are works describes the delay as various model such as first order canonical model in [9] as equation 2-2:

$$d = a_0 + \sum_{i=1}^{n} a_i \Delta X_i + a_{n+1} \Delta R_a \tag{2-2}$$

,where a_0 is the nominal value, ΔX_i are random variables representing the global variations, and ΔR_a is another random variable referring to the uncorrelated variation. In [10], it extends the uncorrelated term in equation 2-2 to vector of local variance. And in [11] it gives another viewpoint of equation 2-2 from Taylor expansion. As another example, in [12], it provides the quadratic timing model as equation 2-3:

$$D = m_0 + \alpha R + \sum_i \beta_i G_i + \sum_{i,j} \Gamma_{ij} G_i G_j \qquad (2\text{-}3)$$

This suggests a series of approaches that with well-defined timing model and two basic operators, add and max, SSTA could be operated as STA. For the timing models mentioned above, add operator is a linear combination of the operands with respect to

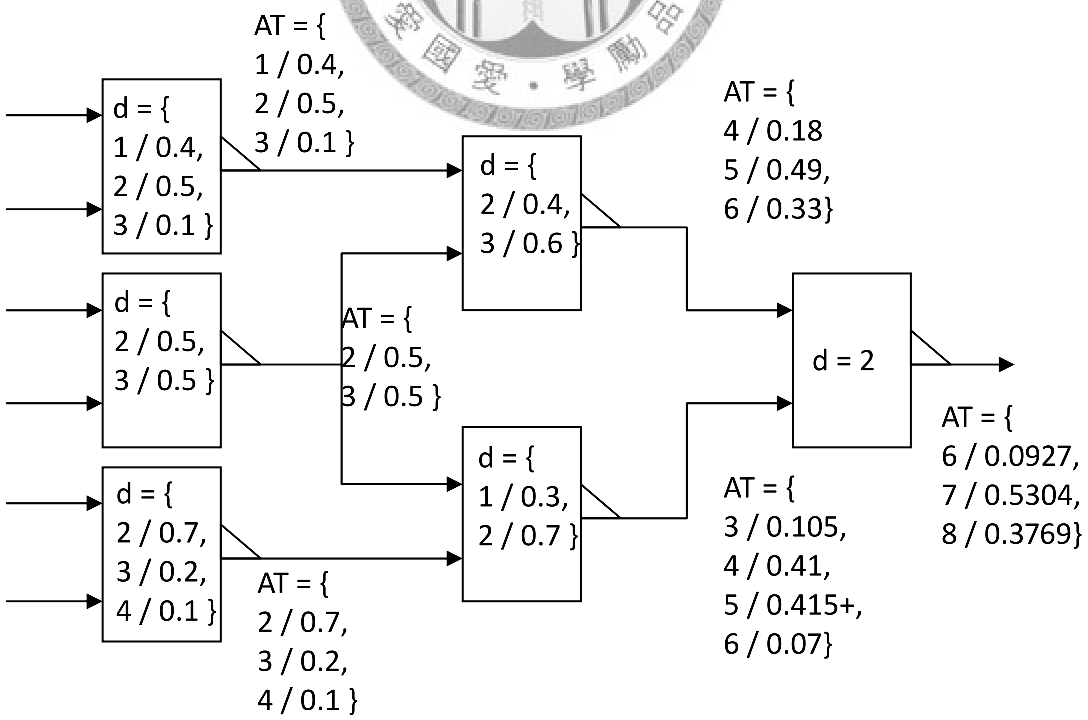


Fig. 2-4 SSTA Result

the coefficients of the timing model. However, max operator is not the case due to its non-linearity. The strategy to the use of max operates creates two branches which are not mutually exclusive: block-based SSTA and path-based SSTA.

2.2.1. Block-based SSTA

The term "block-based" means that the delay would be resolved, i.e. max operator is applied, at some internal block before further computation. In the extreme case, the SSTA is operated as conventional STA in the sense of equation 2-1. As a result, in a block-based SSTA, the main task is to find a relationship between the resolved coefficients of the timing model and timing models of the operands. The most common method is to assume every variation is modeled as a Gaussian random variable. With this assumption of Gaussian random variables, Clark's approximation [13] which is a linear approximation would be used for this max operator. There is other solution not based on the Clark's approximation such as [14], which uses curve-fitting to find the resulted coefficients to the results of the max operation. In [15], it proposes a conditional max approximation which uses a pre-computed skewness to determine the linearity of the max operator. Block-based SSTA in the documents is typically expected to have better performance in runtime.

2.2.2. Path-based SSTA

Path-based SSTA goes in another track. If reviewing the equation 2-1, it's possible in the equation that keeps the max operator unresolved. At the sink node, arrival time from various signal propagation path could be collected. Taking a max operator to this collection, the distribution of critical delay will then be found. The most arguable point is that path-based SSTA might require the enumeration of a great amount of paths. In [16] it suggests that the information of criticality could be used to skip non-critical path, and the methods in [17] and [18] are adopted in that work. Path-based typically takes the advantage of better accuracy. This comes from two sources: one is from the less uses

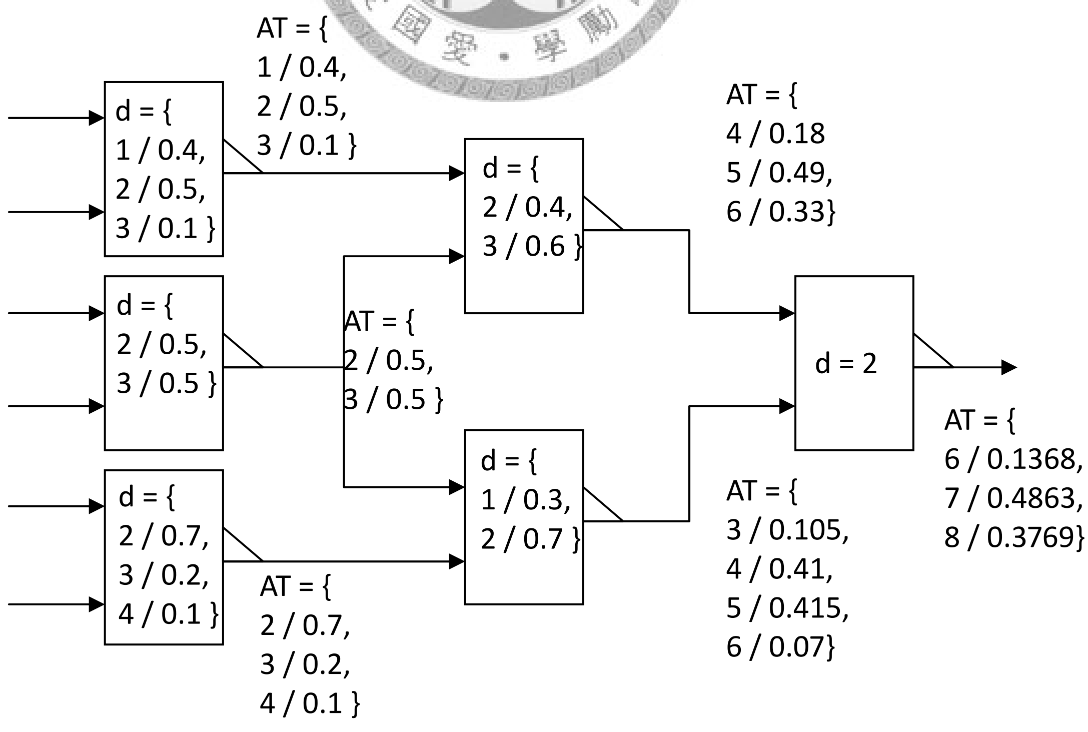


Fig. 2-5 SSTA considering reconvergence path

of max operations; the other is that the path-based strategy facilitates tracking the correlation. The Fig.2-5 is an example illustrating that if the structural correlation due to the reconvergence path is taken into consideration. With the path-based SSTA, it's much easier to cancel the effect of common path and re-calculate the correlation from path to path since the information of the paths is kept. It doesn't mean that path-based SSTA is an exact engine. As the example of Fig. 2-5 suggesting, the accuracy still relies on the well-extracted correlation from path to path.

2.3. Slope Propagation

The impact of timing with respect to signal transition time is well pointed out in [19] and [20] with STA. Although it's not directly followed additive effects with the parametric variation sources, the impact does hold. Worse than that, this impact of signal transition would not be strictly a deterministic value. It would be a distribution as the delay time between gates.

2.4. Monte Carlo method and SSTA

Monte Carlo method is widely used in SSTA, usually for the validation of proposed SSTA methodology. However, there are works such as proposed in [21]. It directly lists every function of the delay and output transition time in canonical form and then rearranges it as a large sparse matrix. With this sparse matrix, it extracts the statistical result with Monte Carlo method. From this study, we find that the evaluated timing performance of the design is bound if the parametric timing model is given. This stands as the basis of our work.

Chapter 3

Proposed Method

To give our method an introduction is that it starts from the deterministic model. The term “model” refers to a set of well-defined variables and functions, and by knowing the practical parameters, one can use the functions to predict any property of the design that provided by the model. In the successive sections, firstly we’ll start from deterministic model where all parameters are treated as some particular values. And then the deterministic model will be extended to statistical one by knowing that every sampling to the statistical space would result in a set of deterministic values. And since each sampling is a set of deterministic values, it would not violate the deterministic model. From section 3.1 to section 3.3, the model is separated into two parts: in section 3.1 and section 3.2, the model is used to relate the intermediate signals or parameters about the relationship between the signals; in section 3.3, the model is further used to relate those signals to the concerned properties. In section 3.4 all the pieces above are meshed up and give a formal methodology to gain statistical distribution of the concerned property from a deterministic model. At the final section of this chapter, 3.5, some examples are given to illustrate how the method proposed in this work is used in application.

3.1. Model, Response, and general overview

It's very obvious that as a model is bound, the response or behavior expected by the model is then fixed for any particular design.

Take Fig.3-1 as an example. The target component is an inverter. To determine the behavior of the inverter, we may run simulations based on some extracted behavior or take measurement to a real element. By given its input signals and estimating its output response under various specification of the gate, the result can finally be summarized as a characterized library. Fig. 3-2 illustrates a possible result in the form of a table.

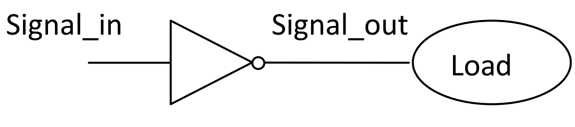


Fig.3-1 An Inverter with load

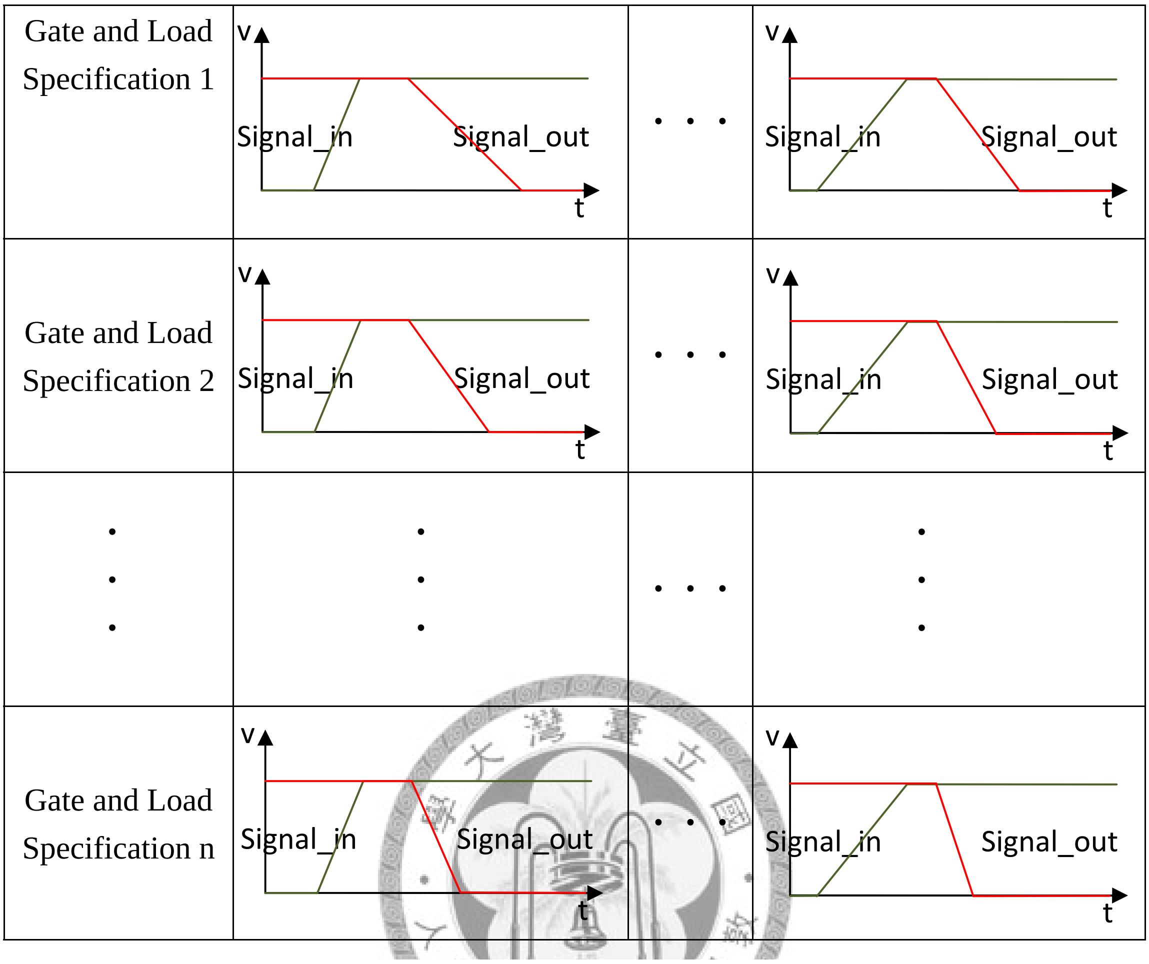
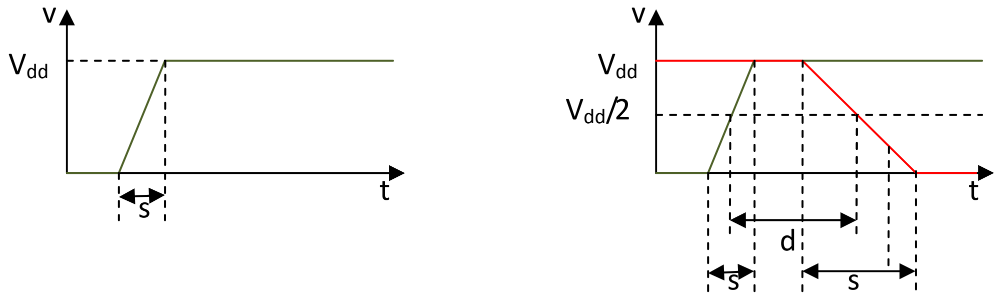


Fig.3-2 Characterized Library

Assuming that all signals are ramp-shaped with known V_{dd}, it follows that every signal can be described with a single variable referring to the slope as in Fig.3-3(a). Considering a pair of stimulus and response as in Fig.3-3(b), then another variable describing the delay between input and output is required if this delay is concerned. Letting that a vector X containing four variables is used to determine the specification of the gate itself and the load it's connected to, then the output signal, now described with a particular slew, could be fit with a pre-guessed function as a model. So is the



(a) Single signal (b) Stimulus-Response Pair

Fig.3-3 Signal Representation

delay d. This relationship could be written as following equations that:

$$\begin{cases} s_{out} = f_{s_{out}}(\overrightarrow{X}, s_{in}) \\ d = f_d(\overrightarrow{X}, s_{in}) \end{cases}, \overrightarrow{X} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T \quad (3\text{-}1)$$

It's very important that in this modeled relationship, all the behaviors to this gate have been explicitly determined if all the required parameters are known, no matter as a deterministic value, or as a set of values with a probability distribution.

For instance, this always holds true that: if vector X and s_{in} is known, such as $\overrightarrow{X} = \overrightarrow{X_0} = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}^T$ and $s_{in} = 3.15$, then

$$\begin{cases} s_{out} = f_{s_{out}}(\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}^T, 3.15) \\ d = f_d(\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}^T, 3.15) \end{cases} \quad (3\text{-}2)$$

It's worth of noting that this claim about model never assumes the correctness of the model. The only requirement is that the model itself is "well-defined", that is, every property derived by this model should be consistent. But even a pair of s_{out} and d is computed in the model, it doesn't mean the same value will be estimated in practical usage. This must be remarkably claimed here that "**models take all the responsibility for its self-consistency and the consistency between the expected behavior and the practical response.**"

In this sense, everything based on a particular model is known if that model is clearly given. In this work, the method to analyze the behavior is illustrated. A widely-used first-order canonical model is adopted as an example.

3.2. Models and Cascading of Functions

Without loss of generality, it's assumed that the model has already been given. For the successive sections in this chapter, the signals are discussed as the slew-based model in section 3.1. Now we can describe the behaviors of the design by the conjunction of the functions. For example:

As Fig.3-4, following the relationship as Eq. 3-1 assumed in section 3.1, it can be written:

$$\begin{cases} s_2 = f_{s,INV1}(\overrightarrow{X_{INV1}}, s_1) \\ d_2 = f_{d,INV1}(\overrightarrow{X_{INV1}}, s_1) \\ s_3 = f_{s,INV2}(\overrightarrow{X_{INV2}}, s_2) \\ d_3 = f_{d,INV2}(\overrightarrow{X_{INV2}}, s_2) \end{cases} \tag{3-3}$$

Cascading those functions, i.e., replacing the intermediate responses, s_2 in this case, with the respective function, it results in Eq. 3-4.

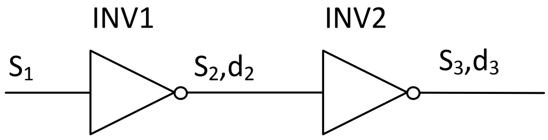


Fig.3-4 Connected Gates

$$\begin{cases} s_2 = f_{s,INV1}(\overrightarrow{X_{INV1}}, s_1) \\ d_2 = f_{d,INV1}(\overrightarrow{X_{INV1}}, s_1) \\ s_3 = f_{s,INV2}(\overrightarrow{X_{INV2}}, f_{s,INV1}(\overrightarrow{X_{INV1}}, s_1)) \\ d_3 = f_{d,INV2}(\overrightarrow{X_{INV2}}, f_{s,INV1}(\overrightarrow{X_{INV1}}, s_1)) \end{cases} \tag{3-4}$$

Considering that all vectors X_{INV} could be concatenated as a new vector containing all the variables required, the set of equations 3-4 could be re-written with this new vector as:

$$\begin{cases} s_2 = \overline{f_{s,INV1}}(\overrightarrow{X_{spec}}) \\ d_2 = \overline{f_{d,INV1}}(\overrightarrow{X_{spec}}) \\ s_3 = \overline{f_{s,INV2}}(\overrightarrow{X_{spec}}) \\ d_3 = \overline{f_{d,INV2}}(\overrightarrow{X_{spec}}) \end{cases}, \overrightarrow{X_{spec}} = \begin{bmatrix} \overrightarrow{X_{INV1}} \\ \overrightarrow{X_{INV2}} \\ s_1 \end{bmatrix} \tag{3-5}$$

It's worth noting for this simplified symbolic representation that the cardinality of the vector X_{spec} may not be equal to the sum of the cardinality of the vector X_{INV1} and X_{INV2} plus one because there may be repeated variables and only one copy is kept in the concatenated X_{spec}. And finally, the set of equations 3-5 can be written as:

$$\vec{Y} = \begin{bmatrix} s_2 \\ d_2 \\ s_3 \\ d_3 \end{bmatrix} = \begin{bmatrix} \overline{f_{s,INV1}}(\overrightarrow{X_{spec}}) \\ \overline{f_{d,INV1}}(\overrightarrow{X_{spec}}) \\ \overline{f_{s,INV2}}(\overrightarrow{X_{spec}}) \\ \overline{f_{d,INV2}}(\overrightarrow{X_{spec}}) \end{bmatrix} = \vec{f}(\overrightarrow{X_{spec}}) \tag{3-6}$$

3.3. Models and Particular Property

Considering the case in Fig.3-4, if some property, such as the arrival time(AT) at terminal of INV_2 is concerned, this property, could be calculated with:

$$AT_{INV2} = d_2 + d_3 \quad (3\text{-}7)$$

The equation 3-7 could imply a particular sense if the equations 3-6 are taken into consideration together, that is:

$$\begin{aligned} AT_{INV2} &= d_2 + d_3 \\ &= 0 \cdot s_2 + 1 \cdot d_2 + 0 \cdot s_3 + 1 \cdot d_3 \\ &= 0 \cdot \overline{f_{s,INV1}}(\overrightarrow{X_{spec}}) + 1 \cdot \overline{f_{d,INV1}}(\overrightarrow{X_{spec}}) + 0 \cdot \overline{f_{s,INV2}}(\overrightarrow{X_{spec}}) + 1 \cdot \overline{f_{d,INV2}}(\overrightarrow{X_{spec}}) \\ &= \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix} \cdot \vec{f}(\overrightarrow{X_{spec}}) \end{aligned} \quad (3\text{-}8)$$

Although Eq. 3-8 is in the form of linear combination, not every property could be written as a linear combination of the set of the functions in Eq. 3-6. For example, one might find that in order to improve the accuracy, there must be some cubic correction term with respect to s_3 as Eq. (3-9).

$$AT_{INV2} = 0 \cdot \overline{f_{s,INV1}}(\overrightarrow{X_{spec}}) + 1 \cdot \overline{f_{d,INV1}}(\overrightarrow{X_{spec}}) + \alpha(\overline{f_{s,INV2}}(\overrightarrow{X_{spec}}))^3 + 1 \cdot \overline{f_{d,INV2}}(\overrightarrow{X_{spec}}) \quad (3\text{-}9)$$

Or as another example, it’s found that the model require a correction term with respect to s_2 and s_3 if both s_2 and s_3 are larger than some threshold such as equation 3-10.

$$AT_{INV2} = \begin{cases} [\alpha \quad 1 \quad \beta \quad 1] \cdot \vec{f}(\overrightarrow{X_{spec}}), H(s_2 - s_{th})H(s_3 - s_{th}) > 0 \\ [0 \quad 1 \quad 0 \quad 1] \cdot \vec{f}(\overrightarrow{X_{spec}}), otherwise \end{cases}, H(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \quad (3\text{-}10)$$

As a consequence, it's preferred to represent all the cases together with a function representation. These function representation could be further re-arranged as a composite function. For the cases of AT in above, it may then look like:

$$AT_{INV2} = g(\vec{f}(\overrightarrow{X_{spec}})) = g \circ \vec{f}(\overrightarrow{X_{spec}}) \quad (3\text{-}11)$$

The form of composite function gives a great insight the property AT is a function of X_{spec}. Carefully recalling the reasoning about equation 3-11, there's almost no limitation to the left-hand side of the equation 3-11. That is, for any property variable P which is predictable in the model, the model should contain a special function g such that:

$$P = g \circ \vec{f}(\vec{X}) \quad (3\text{-}12)$$

The suffix 'spec' is omitted in equation 3-12 for visualized simplicity. The equation 3-12 could be extends by jointly listing several properties with each respective function g such that:

$$\vec{P} = \vec{g}(\vec{f}(\vec{X})) = \vec{g} \circ \vec{f}(\vec{X}) \quad (3\text{-}13)$$

The equation 3-13 should be treated as a part of the requirement to the property "well-defined" mentioned in section 3.1 when the properties of the design is taken as a part of the model. And it must be reminded that the function g, and the composite function are both not restricted to any type. That is, it may be very complex, discontinuous, or even just a list of relations, while equation 3-13 still holds valid.

3.4. From deterministic model to statistical result

Without loss of generality, assuming that the models of the signal and each functional element are given, every design is then a cascading of functional element such as logic gates and connected with intermediate signals. Now we can describe any behavior by the conjunction of the functions as section 3.2. Now it's assumed here that every design discussed in successive parts of this work is capable of evaluation through Monte Carlo method. This sometimes is achieved by properly selecting a set of X_{spec} or applying principal component analysis to find a new X'_{spec} such that the variables are mutually independent. With this property, all functions can be re-written as functions of independent variables. As a consequence, a big function system could be found with respect to these independent variables. It's better to make a remark that all previous works requiring evaluation based on Monte Carlo method inevitably demands this assumption. This provides a good reason for this work to hold the assumption.

Following the equations 3-6, it's very straight-forward that whatever model it is, it could be finally written in the form:

$$\vec{Y} = \vec{f}(\vec{X}) \tag{3-14}$$

And in section 3.3, we conclude that that for any property P in equation 3-12:

$P = g \circ \vec{f}(\vec{X})$

Since all Xs are mutually independent, this theorem holds:

Theorem 1:

If $P = g \circ \vec{f}(\vec{X})$, where $\vec{X} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T$ and each (x_i, x_j) pair is mutually independent, then the distribution of property P, say $p_P(P)$:

$$p_P(P) = \sum_{g \circ \vec{f}(\vec{X}) = P} p_{\vec{X}}(\vec{X}) = \sum_{g \circ \vec{f}(\vec{X}) = P} \left(\prod_i p_{x_i}(x_i) \right), \vec{X} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T \qquad (3\text{-}15)$$

Proof:

Since the summation counts events that are mutually exclusive, by the addition principle, the first equality mark is true. Then since each x_i is mutually independent, according to the multiplication principle the second equality mark is true. As a result, the theorem is true.

If in the theorem 1, starting from equation 3-13 rather than equation 3-12, we get another similar theorem:

Theorem 2:

If $\vec{P} = \vec{g} \circ \vec{f}(\vec{X})$, where $\vec{X} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T$ and each (x_i, x_j) pair is mutually independent, then the joint distribution of property vector P, say $p_{\vec{P}}(\vec{P})$:

$$p_{\vec{P}}(\vec{P}) = \sum_{g \circ \vec{f}(\vec{X}) = \vec{P}} p_{\vec{X}}(\vec{X}) = \sum_{g \circ \vec{f}(\vec{X}) = \vec{P}} \left(\prod_i p_{x_i}(x_i) \right), \vec{X} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T \qquad (3\text{-}16)$$

Proof:

Similarly, the first equality holds for addition principle and multiplication principle for the second one. And consequently, the theorem is true.

Now considering how the solution set is found. For simplicity, we take the case of single property as example. For the constraint that $\left\{\vec{X}\middle| g\circ\vec{f}(\vec{X})=P\right\}$, apparently, $g\circ\vec{f}(\vec{X})=P$ is the only limitation for $\vec{X}$ to be satisfied.

Assume this property holds:

$$P = g\circ\vec{f}(\vec{X}) = h_1(x_1) + h_2(\vec{X'}) \qquad (3\text{-}17),$$

where $\vec{X'} = [0 \quad I_{|\{X\}|-1}]\vec{X}$. This may not be true for all cases. Especially the function $g\circ f$ may be very complicated and no variable is separable. However, for usual cases of artificial models, it's not a rare case to have linear terms. If h_1 is properly selected, we have its respective X_1 being selected without loss of generality. Rearranging equation 3-17, we get:

$$h_1(x_1) = h_2(\vec{X'}) - P \qquad (3\text{-}18)$$

It's assumed here that h_1 is invertible. This is not always true. However, if linear term as mentioned above is selected as h_1, this assumption holds. Then from equation 3-18, we get:

$$x_1 = h_1^{-1}(h_2(\vec{X'}) - P) \qquad (3\text{-}19)$$

In the equation 3-19, noted that every x_i besides x_1 is free, the equation 3-15 in theorem 1 then becomes:

$$p_P(P) = \int_{x_2=-\infty}^{+\infty} \cdots \int_{x_n=-\infty}^{+\infty} p_{x_1}(h_1^{-1}(h_2(\overrightarrow{X'}) - P)) \prod_{i=2}^{n} (p_{x_i}(x_i) dx_i) \tag{3-20}$$

And in the case of property vector, it's more complicated.

Starting from equation 3-13, it's known that $\vec{P} = \vec{g} \circ \vec{f}(\vec{X})$. Assume m properties are considered, as equation 3-17, we may get:

$$\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_m \end{bmatrix} = \begin{bmatrix} h_{11}(x_1) \\ h_{21}(x_2) \\ \vdots \\ h_{m1}(x_m) \end{bmatrix} + \begin{bmatrix} h_{12}(\overrightarrow{X''}) \\ h_{22}(\overrightarrow{X''}) \\ \vdots \\ h_{m2}(\overrightarrow{X''}) \end{bmatrix}, \overrightarrow{X''} = \begin{bmatrix} x_{m+1} & \cdots & x_n \end{bmatrix}^T \tag{3-21}$$

Similarly rearranging equation 3-21 and assuming that every h_1 is invertible, it follows that:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} h_{11}^{-1}(h_{12}(\overrightarrow{X''}) - P_1) \\ h_{21}^{-1}(h_{22}(\overrightarrow{X''}) - P_2) \\ \vdots \\ h_{m1}^{-1}(h_{m2}(\overrightarrow{X''}) - P_m) \end{bmatrix} \tag{3-22}$$

With equation 3-22, the equation 3-16 becomes:

$$p_{\vec{P}}(\vec{P}) = \int_{x_{m+1}=-\infty}^{+\infty} \cdots \int_{x_n=-\infty}^{+\infty} \prod_{i=1}^{m} p_{x_i}(h_{i1}^{-1}(h_{i2}(\overrightarrow{X''}) - P_i)) \prod_{i=m+1}^{n} (p_{x_i}(x_i) dx_i) \tag{3-23}$$

Carefully reviewing the reasoning progress to derive the equation, it's apparently that the linear condition in equation 3-17 is not necessary. For successive reasoning to hold valid, the only condition it required is that it exist some relationship that some variable, e.g. x_1, is separable such that there's an invertible function h_1 where:

$$h_1(x_1) = H(\overrightarrow{X''}, P) \tag{3-24}$$

For example, another possible operation is multiplication. If we have:

$$P = h_1(x_1) \cdot h_2(\overrightarrow{X'}) \tag{3-25}$$

Then, similarly,

$$x_1 = h_1^{-1}\left(\frac{P}{h_2(\overrightarrow{X'})}\right) \tag{3-26}$$

And finally:

$$p_P(P) = \int_{x_2=-\infty}^{+\infty} \cdots \int_{x_n=-\infty}^{+\infty} p_{x_1}\left(h_1^{-1}\left(\frac{P}{h_2(\overrightarrow{X'})}\right)\right) \prod_{i=2}^{n} (p_{x_i}(x_i) dx_i) \tag{3-27}$$

For the case of the property vector, it becomes :

$$\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_m \end{bmatrix} = \begin{bmatrix} h_{11}(x_1) \cdot h_{12}(\overrightarrow{X''}) \\ h_{21}(x_1) \cdot h_{22}(\overrightarrow{X''}) \\ \vdots \\ h_{m1}(x_m) \cdot h_{m2}(\overrightarrow{X''}) \end{bmatrix} \tag{3-28}$$

Rearranging with the inversion of h_{i1}:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} h_{11}^{-1}\left(\frac{P_1}{h_{12}(\overrightarrow{X''})}\right) \\ h_{21}^{-1}\left(\frac{P_2}{h_{22}(\overrightarrow{X''})}\right) \\ \vdots \\ h_{m1}^{-1}\left(\frac{P_m}{h_{m2}(\overrightarrow{X''})}\right) \end{bmatrix} \tag{3-29}$$

And finally the joint PDF is found:

$$p_{\vec{P}}(\vec{P}) = \int_{x_{m+1}=-\infty}^{+\infty} \cdots \int_{x_n=-\infty}^{+\infty} \prod_{i=1}^{m} p_{x_i}(h_{i1}^{-1}(\frac{P_i}{h_{i2}(\overrightarrow{X''})})) \prod_{i=m+1}^{n} (p_{x_i}(x_i)dx_i) \quad (3\text{-}30)$$

In the examples listed above, we conclude the two lemmas below.

Lemma 1:

If $P = g \circ \vec{f}(\vec{X})$, the probability distribution function of P would be in the form:

$$p_P(P) = \int_{x_{k+1}=-\infty}^{+\infty} \cdots \int_{x_n=-\infty}^{+\infty} p_{\overrightarrow{X_1}}(R(\overrightarrow{X_2}, P)) \prod_{i=k+1}^{n} (p_{x_i}(x_i)dx_i) \quad (3\text{-}31)$$

if and only if there is an onto function R from $\begin{bmatrix} \overrightarrow{X_2} \\ P \end{bmatrix}$ to $\overrightarrow{X_1}$, where $\overrightarrow{X_1} = [x_1 \quad \cdots \quad x_k]^T$, $\overrightarrow{X_2} = [x_{k+1} \quad \cdots \quad x_n]^T$, and $1 \le k \le n, k \in N$.

Proof:

(if-part)

If there's an onto function R as claimed, by theorem 1, equation 3-31 is true..

(only if-part)

For equation 3-31 to hold true, R must be at least a function to be used as the argument of the probability function. And then the only problem is the onto relation. Considering the equation 3-31 which is a special case of equation 3-15, since theorem 1 relies on all cases enumeration to support the equality of equation 3-15, all possible $\overrightarrow{X_1}$ must be considered for equation 3-31 to hold true. Therefore, for each $\overrightarrow{X_1}$, there is some $\begin{bmatrix} \overrightarrow{X_2} \\ P \end{bmatrix}$ such that $R(\overrightarrow{X_2}, P) = \overrightarrow{X_1}$. If not, there is a special $\begin{bmatrix} \overrightarrow{X_1} \\ \overrightarrow{X_2} \end{bmatrix}$ such that $g \circ \vec{f}(\begin{bmatrix} \overrightarrow{X_1} \\ \overrightarrow{X_2} \end{bmatrix})$ is undefined which is a contradiction to the well-defined property. By definition, R is an onto function from $\begin{bmatrix} \overrightarrow{X_2} \\ P \end{bmatrix}$ to $\overrightarrow{X_1}$.

Lemma 2:

If $\vec{P} = \vec{g} \circ \vec{f}(\vec{X})$, the joint probability distribution function of $\vec{P}$ would be in the form:

$$p_{\vec{P}}(\vec{P}) = \int_{x_{k+1}=-\infty}^{+\infty} \cdots \int_{x_n=-\infty}^{+\infty} p_{\overrightarrow{X_1}}(R(\overrightarrow{X_2}, \vec{P})) \prod_{i=k+1}^{n} (p_{x_i}(x_i) dx_i) \tag{3-32}$$

if and only if there is an onto relation R from $\begin{bmatrix} \overrightarrow{X_2} \\ \vec{P} \end{bmatrix}$ to $\overrightarrow{X_1}$, where $\overrightarrow{X_1} = [x_1 \quad \cdots \quad x_k]^T$, $\overrightarrow{X_2} = [x_{k+1} \quad \cdots \quad x_n]^T$, and $1 \le k \le n, k \in N$.

Proof:

(if-part)

If there’s an onto relation R as claimed, by theorem 2, equation 3-32 is true..

(only if-part)

For equation 3-31 to hold true, R must be at least a function to be used as the argument of the probability function. And then the only problem is the onto relation. Considering the equation 3-32 which is a special case of equation 3-16, since theorem 2 relies on all cases enumeration to support the equality of equation 3-16, all possible $\overrightarrow{X_1}$ must be considered for equation 3-32 to hold true. Therefore, for each $\overrightarrow{X_1}$, there is some $\begin{bmatrix} \overrightarrow{X_2} \\ \vec{P} \end{bmatrix}$ such that $R(\overrightarrow{X_2}, \vec{P}) = \overrightarrow{X_1}$. If not, there is a special $\begin{bmatrix} \overrightarrow{X_1} \\ \overrightarrow{X_2} \end{bmatrix}$ such that $g \circ \vec{f}(\begin{bmatrix} \overrightarrow{X_1} \\ \overrightarrow{X_2} \end{bmatrix})$ is undefined which is a contradiction to the well-defined property. By definition, R is an onto function from $\begin{bmatrix} \overrightarrow{X_2} \\ \vec{P} \end{bmatrix}$ to $\overrightarrow{X_1}$.

It's obvious that the examples to separate the variables by additive inverse or by multiplicative inverse are both special cases of above lemmas.

3.5. Examples and Simulation Results

In this section, some simple example would be given to give more illustration about how to compute the distribution analytically.

Example 1: Given a gate model as Fig. 3-4, by given that:

$$\begin{aligned} \Delta D &= \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 \\ \Delta S &= \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \end{aligned} \quad (3\text{-}33),$$

where the nominal value is ignored and only the difference variables are modeled. Assuming that every X is mutually independent standard Gaussian random variable, i.e. N(0, 1), calculate the joint distribution of the delta delay and delta slew at the output terminal.

Fig. 3-5 A single Gate

Sol:

From the system 3-33, we can write that:

$$\begin{cases}\Delta D = \begin{bmatrix}\alpha_1 & \alpha_2 & \alpha_3\end{bmatrix}\begin{bmatrix}X_1 & X_2 & X_3\end{bmatrix}^T \\ \Delta S = \begin{bmatrix}\beta_1 & \beta_2 & \beta_3\end{bmatrix}\begin{bmatrix}X_1 & X_2 & X_3\end{bmatrix}^T\end{cases} \Rightarrow \begin{bmatrix}\alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3\end{bmatrix}\begin{bmatrix}X_1 & X_2 & X_3\end{bmatrix}^T = \begin{bmatrix}\Delta D \\ \Delta S\end{bmatrix}$$

Its augmented matrix then is:

$$\left[\begin{array}{ccc:c}\alpha_1 & \alpha_2 & \alpha_3 & \Delta D \\ \beta_1 & \beta_2 & \beta_3 & \Delta S\end{array}\right] \tag{3-34}$$

Without loss of generality, the reduced echelon form of 3-34 is:

$$\left[\begin{array}{ccc:c}1 & 0 & \dfrac{\alpha_3\beta_2 - \alpha_2\beta_3}{\beta_2\alpha_1 - \beta_1\alpha_2} & \dfrac{\Delta D\beta_2 - \Delta S\alpha_2}{\beta_2\alpha_1 - \beta_1\alpha_2} \\ 0 & 1 & \dfrac{\alpha_1\beta_3 - \alpha_3\beta_1}{\beta_2\alpha_1 - \beta_1\alpha_2} & \dfrac{\Delta S\alpha_1 - \Delta D\beta_1}{\beta_2\alpha_1 - \beta_1\alpha_2}\end{array}\right] \tag{3-35}$$

From 3-35, it's followed that:

$$\begin{bmatrix}X_1 \\ X_2\end{bmatrix} = \begin{bmatrix}\dfrac{\Delta D\beta_2 - \Delta S\alpha_2}{\beta_2\alpha_1 - \beta_1\alpha_2} - \dfrac{\alpha_3\beta_2 - \alpha_2\beta_3}{\beta_2\alpha_1 - \beta_1\alpha_2}X_3 \\ \dfrac{\Delta S\alpha_1 - \Delta D\beta_1}{\beta_2\alpha_1 - \beta_1\alpha_2} - \dfrac{\alpha_1\beta_3 - \alpha_3\beta_1}{\beta_2\alpha_1 - \beta_1\alpha_2}X_3\end{bmatrix} \tag{3-36}$$

By Theorem 2, and from 3-36 the joint distribution is:

$$p_{\Delta D,\Delta S}(\Delta D, \Delta S) = \int_{-\infty}^{\infty} p_{X_1}(X_1)p_{X_2}(X_2)p_{X_3}(X_3)dX_3 \tag{3-37}$$

Replacing the X_1, X_2 in Eq. 3-37 with 3-36, and introducing the probability distribution function of the standard Gaussian random variable, the integration would then be:

$$p_{\Delta D,\Delta S}(\Delta D,\Delta S)=\frac{1}{2\pi}\frac{(\beta_2\alpha_1-\beta_1\alpha_2)e^{-\frac{\sum_{i=1}^{3}(\Delta S\alpha_i-\Delta D\beta_i)^2}{2((\alpha_3\beta_2-\alpha_2\beta_3)^2+(\alpha_1\beta_3-\alpha_3\beta_1)^2+(\alpha_1\beta_2-\alpha_2\beta_1)^2)}}}{\sqrt{(\alpha_3\beta_2-\alpha_2\beta_3)^2+(\alpha_1\beta_3-\alpha_3\beta_1)^2+(\alpha_1\beta_2-\alpha_2\beta_1)^2}} \tag{3-38}$$

We can validate this result by given random instances and comparing to the Monte Carlo method. For example, one instance might be:

$$\begin{cases}\Delta D=\begin{bmatrix}-0.4326 & 0.1253 & -1.1465\end{bmatrix}\begin{bmatrix}X_1 & X_2 & X_3\end{bmatrix}^T \\ \Delta S=\begin{bmatrix}-1.1656 & 0.2877 & 1.1909\end{bmatrix}\begin{bmatrix}X_1 & X_2 & X_3\end{bmatrix}^T\end{cases} \tag{3-39}$$

The Fig. 3-6 shows the result. The LHS figure is from the equation 3-38 and the RHS is from the Monte Carlo method with two million samples. The upper figure illustrates the respective joint distribution and the lower figure is the contour.

Example 2: Let everything invariant but given the gate model as

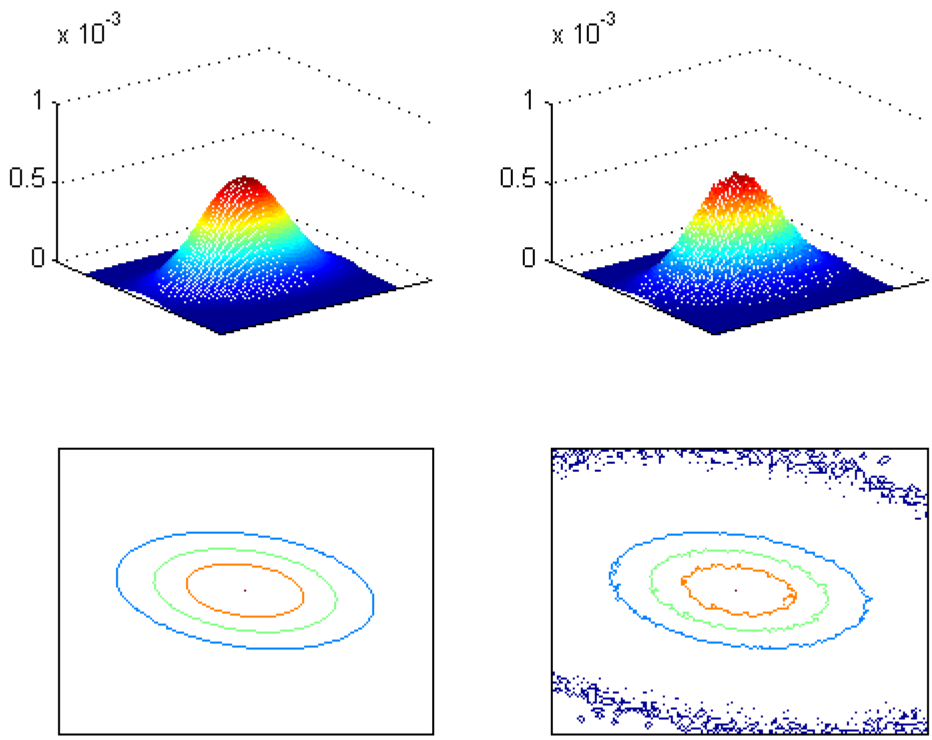


Fig. 3-6 Simulation result of example 1

$$\Delta D = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5$$
$$\Delta S = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 \tag{3-40}$$

Repeat example 1.

Sol:

Similarly, from 3-40, calculating its reduced echelon form of its augmented matrix, then it can be found:

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \dfrac{\Delta D\beta_2 - \Delta S\alpha_2}{\beta_2\alpha_1 - \beta_1\alpha_2} \\ \dfrac{\Delta S\alpha_1 - \Delta D\beta_1}{\beta_2\alpha_1 - \beta_1\alpha_2} \end{bmatrix} - \begin{bmatrix} \dfrac{\alpha_3\beta_2 - \alpha_2\beta_3}{\beta_2\alpha_1 - \beta_1\alpha_2} & \dfrac{\alpha_4\beta_2 - \alpha_2\beta_4}{\beta_2\alpha_1 - \beta_1\alpha_2} & \dfrac{\alpha_5\beta_2 - \alpha_2\beta_5}{\beta_2\alpha_1 - \beta_1\alpha_2} \\ \dfrac{\beta_3\alpha_1 - \beta_1\alpha_3}{\beta_2\alpha_1 - \beta_1\alpha_2} & \dfrac{\beta_4\alpha_1 - \beta_1\alpha_4}{\beta_2\alpha_1 - \beta_1\alpha_2} & \dfrac{\beta_5\alpha_1 - \beta_1\alpha}{\beta_2\alpha_1 - \beta_1\alpha_2} \end{bmatrix} \begin{bmatrix} X_3 \\ X_4 \\ X_5 \end{bmatrix} \tag{3-41}$$

The symbolic result similar to Eq. 3-38 is very tedious, only the random instance and its joint distribution would be listed:

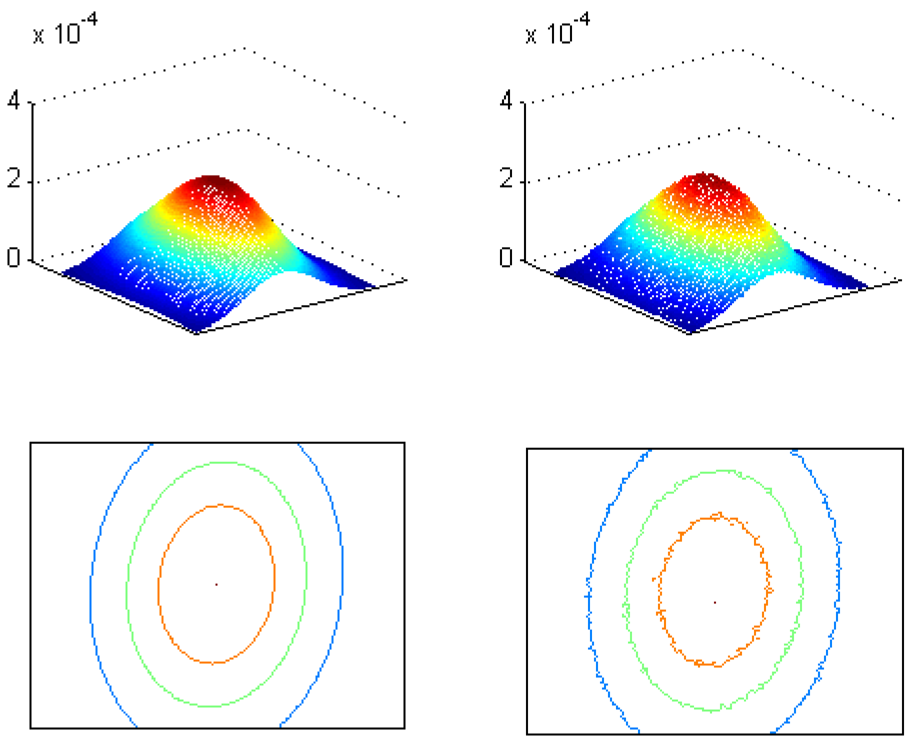


Fig. 3-7 Simulation result of example 2

$$\Delta D = -0.5419X_1 - 1.2991X_2 + 1.0187X_3 - 3.1138X_4 + 0.9024X_5$$
$$\Delta S = 1.2769X_1 - 1.4422X_2 - 0.1041X_3 - 0.0600X_4 - 0.7245X_5 \quad (3\text{-}42)$$

Its joint probability is:

$$p_{\Delta D,\Delta S}(\Delta D, \Delta S) = 0.0514 \times e^{-0.0372\Delta D^2 + 0.0107\Delta D\Delta S - 0.1186\Delta S^2} \quad (3\text{-}43)$$

Running the simulation with fifteen million Monte Carlo samples, Fig. 3-7 is found.

Example 3: As Fig. 3-8, now the two gates are concatenated.

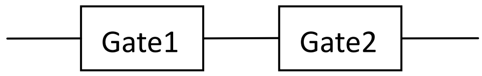


Fig. 3-8 Concatenated Gates

By given that:

$$\Delta D_1 = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3$$
$$\Delta D_2 = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \quad (3\text{-}44)$$
$$\Delta AT = \Delta D_1 + \Delta D_2$$

The delta AT is the variance affected by the variance sources. Letting every other assumption the same as the first example, calculate the distribution of delta AT.

Sol:

From 3-44, it directly follows:

$$\Delta AT = \sum_{i=1}^{3} (\alpha_i + \beta_i) X_i = \sum_{i=1}^{3} \gamma_i X_i$$

(3-45)

According to Theorem 1, we can find

$$p(\Delta AT) = Ke^{-\frac{\Delta AT^2}{2(\gamma_1^2 + \gamma_2^2 + \gamma_3^2)}} \quad (3\text{-}46)$$

K is a scalar which could be resolved with the law of the total probability that:

$$p(\Delta AT) = \frac{1}{\sqrt{2\pi(\gamma_1^2 + \gamma_2^2 + \gamma_3^2)}} e^{-\frac{\Delta AT^2}{2(\gamma_1^2+\gamma_2^2+\gamma_3^2)}} \tag{3-47}$$

. There's another viewpoint from Eq. 3-45. Since every X is mutually independent

Gaussian, the delta AT is consequently another Gaussian, where:

$$\begin{cases} \mu_{\Delta AT} = \sum_{i=1}^{3} \gamma_i \mu_{X_i} = 0 \\ \sigma_{\Delta AT}^2 = \sum_{i=1}^{3} \gamma_i^2 \sigma_{X_i}^2 = \sum_{i=1}^{3} \gamma_i^2 \end{cases} \tag{3-48}$$

From 3-48, by the definition of normal distribution:

$$p(\Delta AT) = \frac{1}{\sqrt{2\pi(\gamma_1^2 + \gamma_2^2 + \gamma_3^2)}} e^{-\frac{\Delta AT^2}{2(\gamma_1^2+\gamma_2^2+\gamma_3^2)}} \tag{3-49}$$

The consistency of the Eq. 3-47 and Eq. 3-49 is nothing wonder.

Example 4: As the previous example 3, however, the model is then given as:

$$\begin{aligned} \Delta D_1 &= \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 \\ \Delta S_1 &= \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \\ \Delta D_2 &= \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \gamma_4 \Delta S_1 \\ \Delta AT &= \Delta D_1 + \Delta D_2 \end{aligned} \tag{3-50}$$

Repeat example 3.

Sol:

Similarly, from 3-50:

$$\Delta AT = \sum_{i=1}^{3} (\alpha_i + \gamma_i + \gamma_4 \beta_i) X_i = \sum_{i=1}^{3} \tau_i X_i \tag{3-51}$$

Thus, from Theorem 1,

$$p(\Delta AT) = K e^{-\frac{\Delta AT^2}{2(\tau_1^2+\tau_2^2+\tau_3^2)}} \tag{3-52}$$

Determine K by the law of the total probability,

$$p(\Delta AT) = \frac{1}{\sqrt{2\pi(\tau_1^2 + \tau_2^2 + \tau_3^2)}} e^{-\frac{\Delta AT^2}{2(\tau_1^2+\tau_2^2+\tau_3^2)}} \tag{3-53}$$

And similarly, it can be verified by direct computation with Gaussian random variables:

$$\begin{cases} \mu_{\Delta AT} = \sum_{i=1}^{3} \tau_i \mu_{X_i} = 0 \\ \sigma_{\Delta AT}^2 = \sum_{i=1}^{3} \tau_i^2 \sigma_{X_i}^2 = \sum_{i=1}^{3} \tau_i^2 \end{cases} \tag{3-54}$$

The distribution from direct computation with Gaussian random variables is:

$$p(\Delta AT) = \frac{1}{\sqrt{2\pi(\tau_1^2 + \tau_2^2 + \tau_3^2)}} e^{-\frac{\Delta AT^2}{2(\tau_1^2+\tau_2^2+\tau_3^2)}} \tag{3-55}$$

Example 5: As Fig. 3-9, considering the gate with more than single input.

Fig. 3-9 A gate with two inputs

Given the model as:

$$\begin{cases} AT = D_g + \max\{AT_a, AT_b\} \\ D_g = \alpha_g X + c_g \begin{bmatrix} S_a \\ S_b \end{bmatrix} + \mu_{D_g} \\ AT_a = \alpha_a X + \mu_{AT_a} \\ AT_b = \alpha_b X + \mu_{AT_b} \\ S_a = \beta_a X + \mu_{S_a} \\ S_b = \beta_b X + \mu_{S_b} \end{cases} \tag{3-56}$$

The D_g is the gate delay. AT_a is the arrival time of one input, and AT_b is another one. The X is a vector representing the all possible X_i as in previous examples. μ is the mean value. Try calculating the distribution of AT.

Sol:

From the system 3-56,

$$\begin{cases} AT = (\alpha_g + c_g \begin{bmatrix} \beta_a \\ \beta_b \end{bmatrix} + \alpha_a)X + (\mu_{D_g} + \mu_{AT_a} + c_{gS_a}\mu_{S_a} + c_{gS_b}\mu_{S_b}) \\ AT_a = \alpha_a X + \mu_{AT_a} & ,\max\{AT_a, AT_b\} = AT_a \\ AT_b = \alpha_b X + \mu_{AT_b} \end{cases}$$

$$\begin{cases} AT = (\alpha_g + c_g \begin{bmatrix} \beta_a \\ \beta_b \end{bmatrix} + \alpha_b)X + (\mu_{D_g} + \mu_{AT_b} + c_{gS_a}\mu_{S_a} + c_{gS_b}\mu_{S_b}) \\ AT_a = \alpha_a X + \mu_{AT_a} & ,\max\{AT_a, AT_b\} = AT_b \\ AT_b = \alpha_b X + \mu_{AT_b} \end{cases}$$

In order to relate this system with translational mean value to the previous ones centered at zero, rearrange the mean and redefine the variable such as:

$$\begin{cases} AT' = AT - \mu_{AT_1} = (\alpha_g + c_g \begin{bmatrix} \beta_a \\ \beta_b \end{bmatrix} + \alpha_a)X \\ AT_a' = AT_a - \mu_{AT_a} = \alpha_a X & ,\max\{AT_a, AT_b\} = AT_a \\ AT_b' = AT_b - \mu_{AT_b} = \alpha_b X \end{cases}$$
$$\begin{cases} AT' = AT - \mu_{AT_2} = (\alpha_g + c_g \begin{bmatrix} \beta_a \\ \beta_b \end{bmatrix} + \alpha_b)X \\ AT_a' = AT_a - \mu_{AT_a} = \alpha_a X & ,\max\{AT_a, AT_b\} = AT_b \\ AT_b' = AT_b - \mu_{AT_b} = \alpha_b X \end{cases} \quad (3\text{-}57)$$

Further, it can be simplified by define a difference variable:

$$\Delta d = AT_a - AT_b \quad (3\text{-}58)$$

We can rewrite 3-57 with 3-58 as:

$$
\begin{cases} AT_1' = AT - \mu_{AT_1} = (\alpha_g + c_g \begin{bmatrix} \beta_a \\ \beta_b \end{bmatrix} + \alpha_a) X \\ \Delta d' = \Delta d - (\mu_{AT_a} - \mu_{AT_b}) = (\alpha_a - \alpha_b) X \end{cases}, \Delta d' \geq -(\mu_{AT_a} - \mu_{AT_b})
$$
$$
\begin{cases} AT_2' = AT - \mu_{AT_2} = (\alpha_g + c_g \begin{bmatrix} \beta_a \\ \beta_b \end{bmatrix} + \alpha_b) X \\ \Delta d' = \Delta d - (\mu_{AT_a} - \mu_{AT_b}) = (\alpha_a - \alpha_b) X \end{cases}, \Delta d' < -(\mu_{AT_a} - \mu_{AT_b}) \tag{3-59}
$$

System 3-59 suggests that AT’ has two different functions controlled by $\Delta d'$. We can use lemma 2 to calculate the joint distribution. And then by the condition of mutually exclusive, by the additive principle:

$$
p(AT') = \int_{-(\mu_{AT_a}-\mu_{AT_b})}^{\infty} p(AT_1', \Delta d) d\Delta d + \int_{-\infty}^{-(\mu_{AT_a}-\mu_{AT_b})} p(AT_2', \Delta d) d\Delta d
$$

(3-60)

Validate Eq. 3-60 with random instances such as:

$$
\begin{bmatrix} \alpha_g \\ \alpha_a \\ \alpha_b \\ \beta_a \\ \beta_b \end{bmatrix} = \begin{bmatrix} 1.7286 & -0.2249 & -5.7929 & -0.1382 & 1.5832 & -10.3104 \\ -1.1533 & 1.3209 & 2.2992 & 7.5368 & 3.1228 & 1.4017 \\ -1.0996 & 7.6782 & -4.8447 & 2.6592 & -2.5813 & -5.3724 \\ 1.1005 & -3.2224 & 1.6291 & -3.1907 & 3.4132 & 3.6074 \\ 1.6450 & -2.3941 & 0.6852 & -2.9263 & -0.5577 & 0.8148 \end{bmatrix}, \tag{3-61}
$$

$$
\begin{bmatrix} \mu_{D_g} \\ \mu_{AT_a} \\ \mu_{AT_b} \\ \mu_{S_a} \\ \mu_{S_b} \end{bmatrix} = \begin{bmatrix} 2.3608 \\ 1.9679 \\ 0.0001 \\ 0.3937 \\ 1.4846 \end{bmatrix} ,
$$

(3-62)

$$
\begin{bmatrix} C_{gS_a} \\ C_{gS_b} \end{bmatrix} = \begin{bmatrix} 3.5797 \\ 1.2131 \end{bmatrix} \tag{3-63}
$$

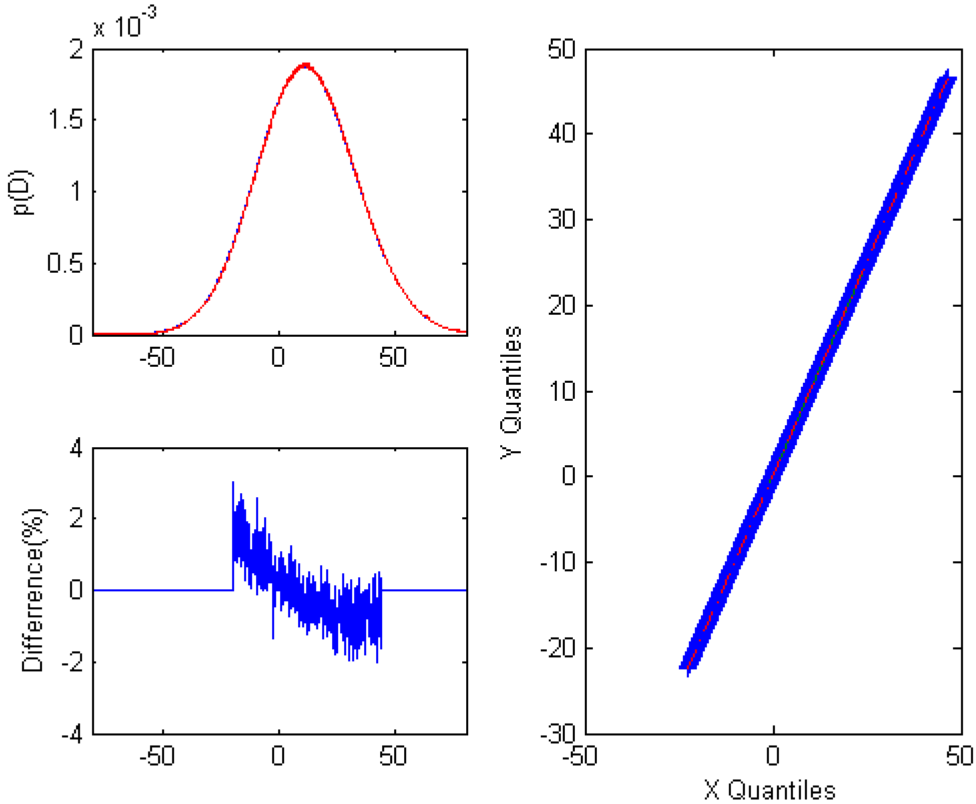


Fig. 3-10 Simulation result of example 5

The simulation result is as Fig.3-10. In Fig. 3-10, the upper subplot of LHS contains both the distributions: from Eq. 3-60 and from Monte Carlo method. The lower subplot of LHS illustrates the difference, between $\mu \pm \sigma$. The RHS is the Q-Q plot which identifies the regularity of the two distributions.

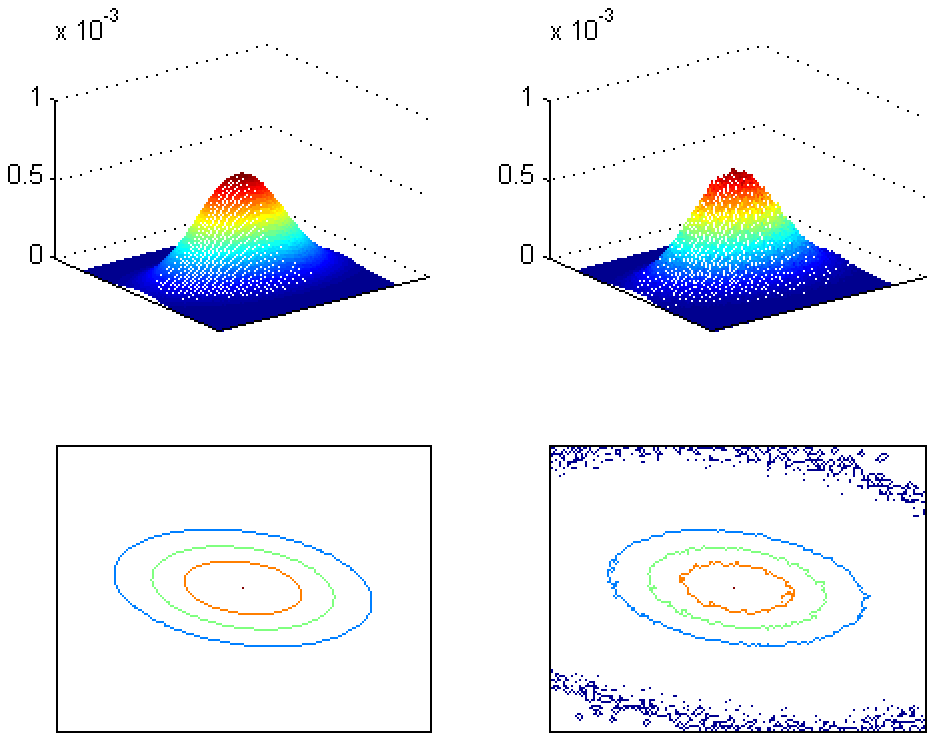
In the last example, it's noted that we can separate the system in 3-59 to two types of functions: one is to relate the properties to the parametric variables, and the other is to relate the parametric variables to the variables controlling previous functions. It must be noted clearly that although the generalized function form looks simple in our method, its practical use might be tedious in the integration. Such as in the last example, the definition of property function may not be invariant. This not only affects the integration where the joint distribution is extracted, but the final property distribution would be affected as well when integration is used to find the marginal probability.

Chapter 4

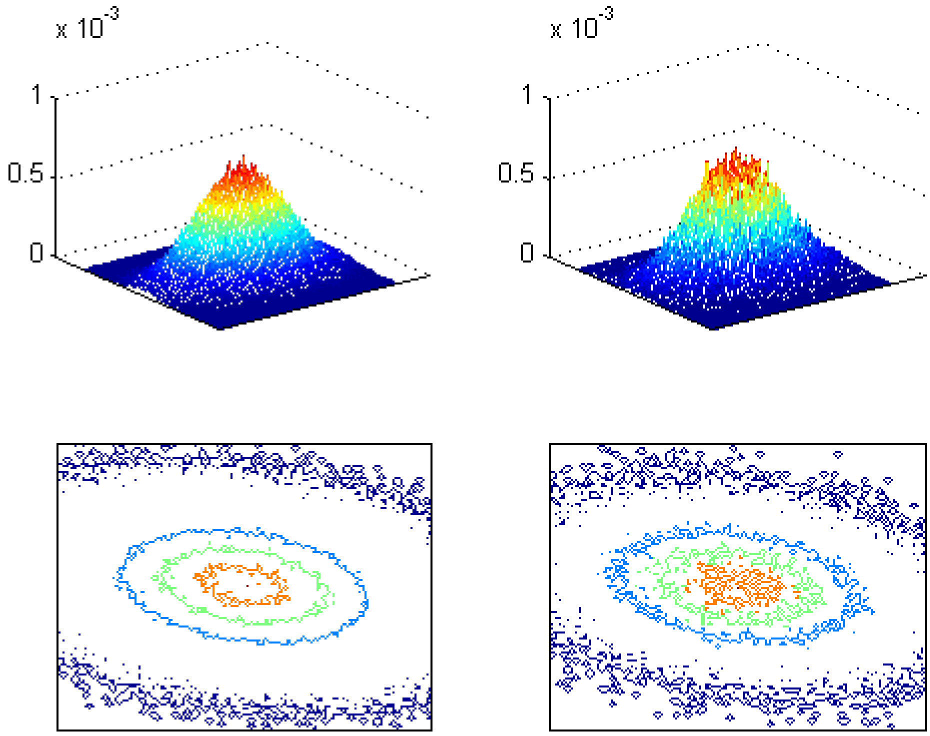
Discussion

The first question we would be interested in is the validation of our method. The examples in the section 3.5 provide some confidence. We shall compare those results with Monte Carlo especially changing the numbers of the samples. Four sets of subplots are listed in Fig. 4-1. Two of them are the same as what has been shown in section 3.5. The others are based on Monte Carlo with half million and one-tenth million samples respectively.

Similar listing would be found in Fig. 4-2, where originally fifteen million samples are used. The comparative simulations are based on five million and one million samples. An obvious trend is that the required samples significantly increasing with the

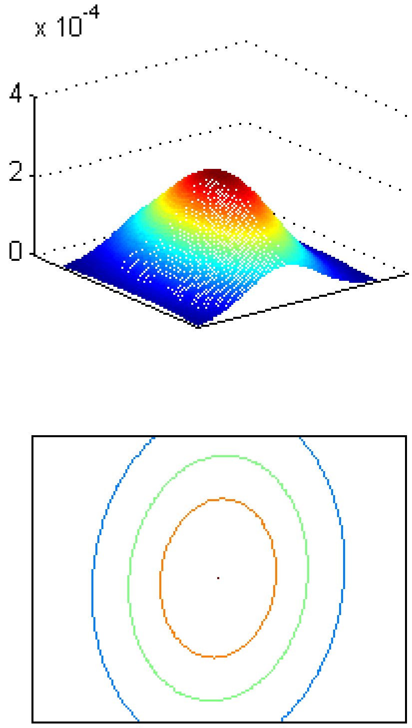


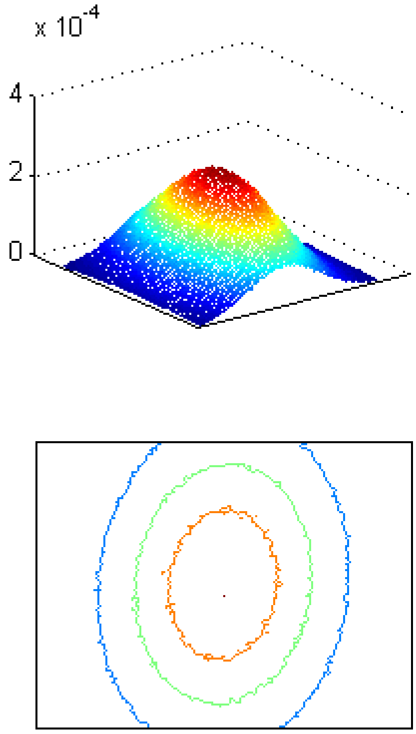
(a) Proposed Method (b) MC with 2,000,000 pts



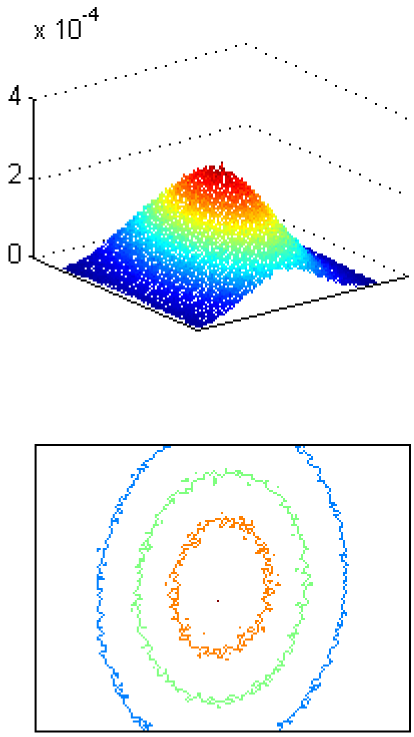
(c) MC with 500,000 pts (d) MC with 100,000 pts

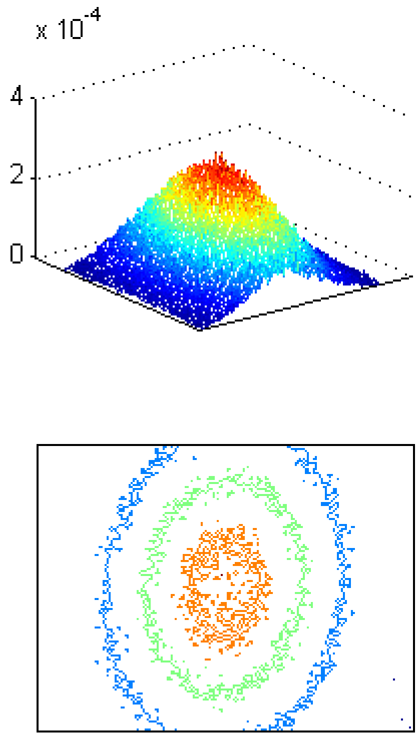
Fig. 4-1 More simulation results of example 1





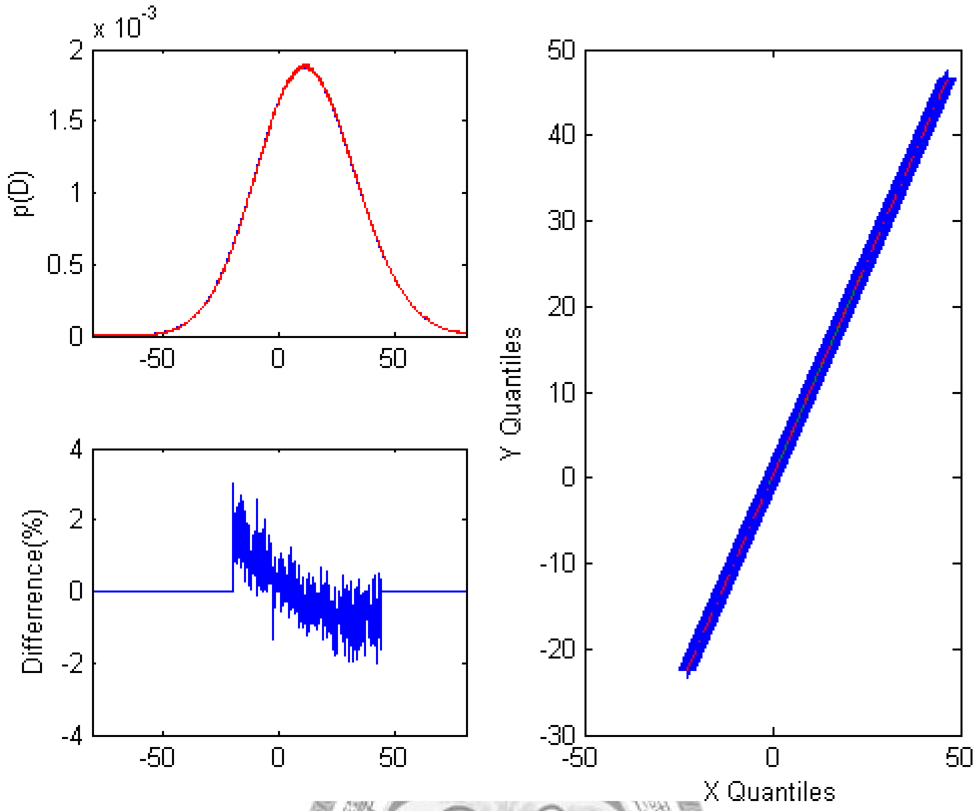
(a) Proposed Method (b) MC with 15,000,000 pts



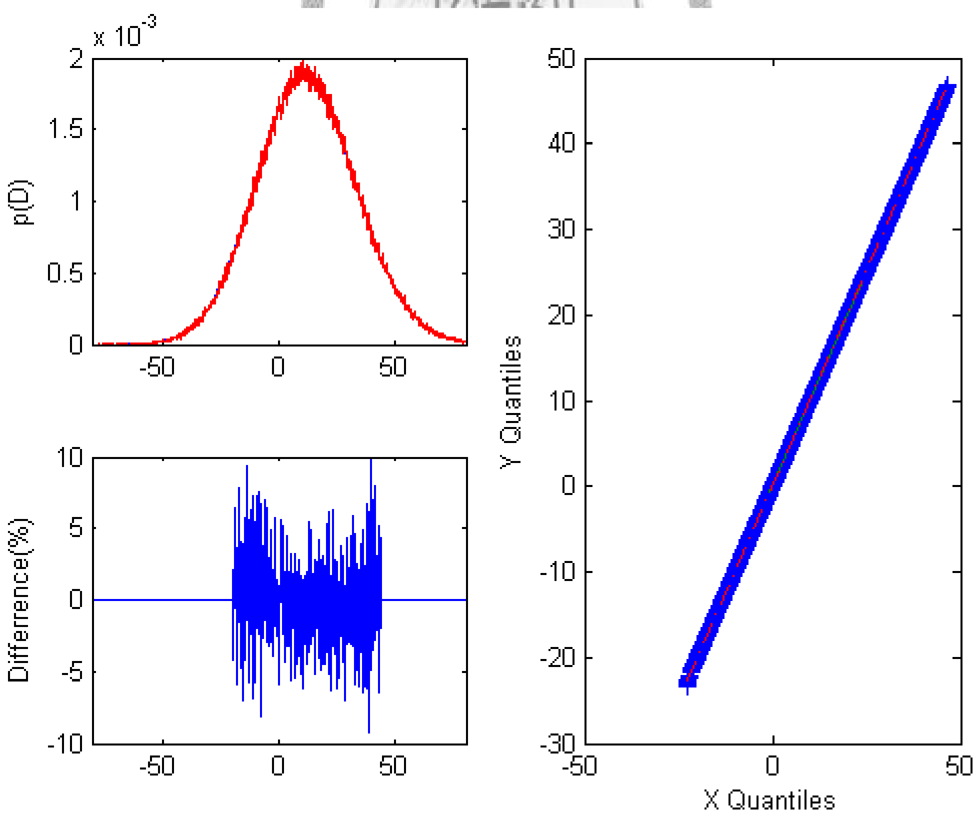


(c) MC with 1,000,000 pts (d) MC with 5,000,000 pts

Fig. 4-2 More simulations of example 2



(a) Proposed method v.s.MC with 30,000,000 pts



(b) Proposed method v.s.MC with 1,000,000 pts

Fig. 4-3 More simulations of example 5

number of the variance variables taken into consideration.

Fig. 4-3 is the result of the example 5 originally with thirty million samples. And the comparative case uses one million samples. It's very important to find that for the Monte Carlo method, the improvement rate with the increasing samples might be much worse than linear.

From our method, we can look back to the path-based and block-based SSTA. For the path-based SSTA, it's not hard to find similar track within the example 2. Similar to example 2, all paths could be enumerated in our method. But it's very important that the path-based would take all paths into consideration. As Fig. 4-4 we give an example.

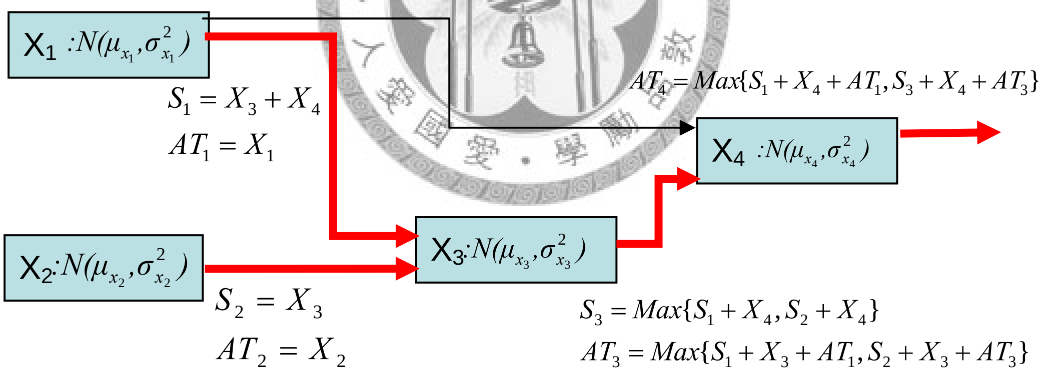


Fig. 4-4 A netlist for example

If only the red paths are taken, comparing the result to the result with all paths, from Monte Carlo method, we get Fig. 4-5.

It's important to know the trend that the tail of the distribution would not be caught. This observation could be found with theorem 1 and theorem 2 since the probability based on the additive parts in the proof might be partially truncated if not all paths are

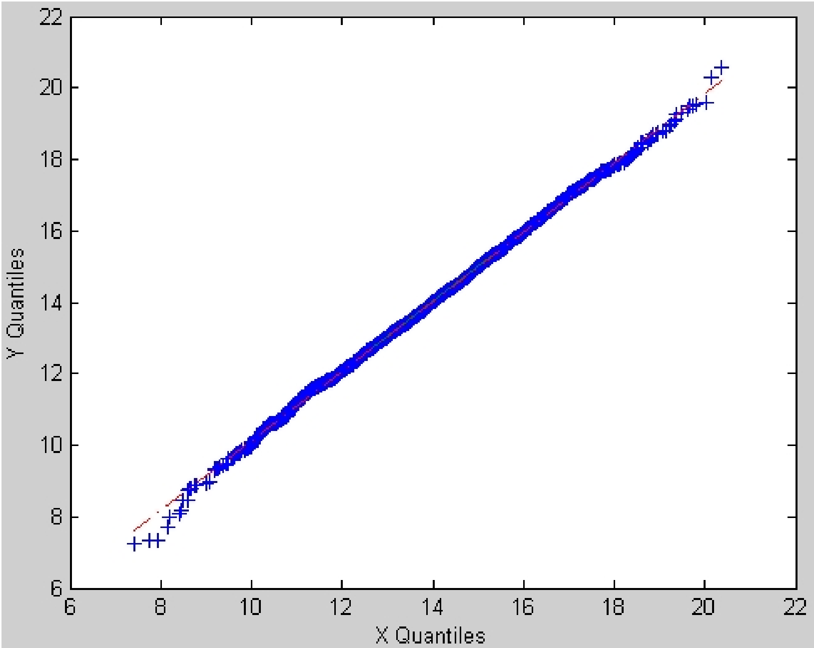


Fig. 4-5 Q-Q Plot for all paths MC v.s. M.C. with critical paths

taken into consideration. It may suggest a weighted summation or weighted average is required as a correction based on the effects from the non-critical paths.

As for the block-based SSTA, from example 5, we know that the function would split because the nonlinear max operation makes the function translation diverge. This would be far more complicated if the mapping function take more physical effects into consideration. Traditional block-based SSTA doesn't elegantly solve this problem and leave it a main error source as claimed in [14]. The curve-fitting method in [14] in some sense is the effort finding a mean function $p(\overline{AT},\Delta d)$ instead of the $p(AT_1^{'},\Delta d)$ and $p(AT_2^{'},\Delta d)$ in Eq. 3-60. The continuous result may suggest the existence of this mean function. However, it's out of the scope of our current work.

Chapter 5

Conclusion

Statistical analysis is a growing topic in recent IC industry. In this work, an analytical analysis is provided to give another viewpoint of the statistical analysis. Examples in SSTA are given to illustrate how this method is applied, and random instances are given as validation. By the theorems and lemma given in this work, we provide the sufficient and necessity condition of the mathematical exactness. The might engineering tractability is the goal of our future work. And finally it’s expected to be a much more powerful statistical analysis framework with this method.

Bibliography

[1] "ITRS Papers and Presentations", [webpage] (2008, Jan. 14), Available at HTTP: http://www.itrs.net/

[2] M. Celik, "Statistical static timing analysis: A view from the future" [online article] (2007,Jul. 27), Available at HTTP: *http://www.eetimes.com/showArticle.jhtml?articleID=201201507*

[3] R.B. Hitchcock, Sr., G.L. Smith, and D.D. Cheng, "Timing analysis of computer hardware," *IBM Journal of Research and Development*, pp. 100-105, January 1982.

[4] R.B. Hitchcock, "Timing verification and the timing analysis program," in *Proc. DAC*, 1982, pp. 594-604.

[5] E. Horowitz, S. Sahni, and D. Mehta, "Activity Networks" in *Fundamental of*

Data Structures in C++, W.H. Freeman, NY, 1995, pp.378-393.

[6] H.-F. Jyu, S. Malik, S. Devadas, K.W. Keutzer, “Statistical timing analysis of combinational logic circuits,” *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol.1, pp.126-137, 1993.

[7] A. Devgan and C. Kashyap, “Block-based static timing analysis with uncertainty,” *ICCAD 2003*, pp.607-614.

[8] V. Mehrotra, “Modeling the effects of systematic process variation on circuit performance,” PhD Dissertation, MIT.

[9] C. Visweswariah, K. Ravindran, and K. Kalafala, “First-order parameterized block-based statistical timing analysis,” *TAU’04*, Feb 2004.

[10] L. Zhang, W. Chen, Y. Hu, and C.C.-P. Chen,”Statistical timing analysis with extended pseudo-canonical timing model,” *DATE’05*, 952-957.

[11] H. Chang and S. S. Sapatnekar, “Statistical timing analysis considering spatial correlations using a single PERT-like traversal,” *ICCAD 2003*, pp.621-625.

[12] L. Zhang, W. Chen, Y. Hu, J. A. Gubner, C.C. Chen, “Correlation-perserved non-gaussian statistical timing analysis with quadratic timing model,” in *Proc. DAC*, 2005, pp.83-88..

[13] C. E. Clark, “The greatest of a finite set of random variables,” *Operations*

Research, pp. 145-162, 1961.

[14] S. Bhardwaj, P. Ghanta, and S. Vrudhula, “A framework for statistical timing analysis using non-linear delay and slew model,” *ICCAD 2006*, pp.225-230.

[15] L. Zhang, W. Chen, Y. Hu, and C. C. –P. Chen, “Statistical static timing analysis with conditional linear MAX/MIN approximation and extended canonical timing model,” *Computer-Aided Design of Integrated Circuit and Systems, IEEE Transactions on*, vol.25, pp.1183-1191, 2006.

[16] W.-S. Wang and M. Orshansky, “Path-based statistical timing analysis handling arbitrary delay correlations: theory and implementation,” *Computer-Aided Design of Integrated Circuit and Systems, IEEE Transactions on*, vol.25, pp.2976-2988, 2006.

[17] S. Yen, D. Du, and S. Ghanta, “Efficient algorithms for extracting the k most critical paths in timing analysis,” in *Proc. DAC*, 1989, pp.649-652.

[18] Y.-C. Ju and T. Saleh,” Incremental techniques for the identification of statically sensitizable critical paths,” in *Proc. DAC*, 1991, pp.541-546.

[19] A.I. Kayssi, K.A. Sakallah, and T.N. Mudge, “The impact of signal transition time on path delay computation,” *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol.40, pp.302-309, 1993.

[20] D. Blaauw, V. Zolotov, and S. Sundareswaran, ”Slope propagation in static timing analysis,” *Computer-Aided Design of Integrated Circuit and Systems, IEEE Transactions on*, vol.21, pp.1180-1195, 2002.

[21] A. Ramalingam, G.–J. Nam, A.K. Singh, M.Orshansky, S.R. Nassif, and D.Z. Pan, “An accurate sparse matrix based framework for statistical static timing analysis,” *ICCAD 2006*, pp.231-236.