Circuit size optimization with multiple sources of variation and position dependant correlation - art. no. 65210P

Article *in* Proceedings of SPIE - The International Society for Optical Engineering · March 2007

DOI: 10.1117/12.711794

CITATIONS
8

READS
31

4 authors, including:

Costas Spanos
University of California, Berkeley
314 PUBLICATIONS **4,465** CITATIONS

SEE PROFILE

Circuit Size Optimization with Multiple Sources of Variation and Position Dependant Correlation

Qian Ying Tang[1], Paul Friedberg, George Cheng, Costas J. Spanos
211 Cory Hall, Department of Electrical Engineering and Computer Sciences,
University of California, Berkeley, CA 94720

ABSTRACT

The growing impact of process variation on circuit performance requires statistical design approaches in which circuits are designed and optimized subject to an estimated variation. Previous work [1] has shown that by including extra margins in each of the gate delays and optimizing the gate sizes, the circuit delay variation can be reduced by half. Our work goes further by deploying extended models that include delay variations due to V_{th} and L_{eff}, as well as position dependant variation. Two types of models have been proposed to account for various variations: 1) a model that explicitly adds spatial correlation terms to the design objective; 2) a model that implicitly includes such effect through the use of a modified version of Pelgrom's model. These design models are used to size a 32-bit Ladner-Fischer adder and the circuit delay distributions are obtained from Monte Carlo simulations. The analysis shows that both types of models have a noticeable performance improvements over the model presented in [1]. In addition, the second model appears to be a more adequate method for modeling various variation components and has a better performance over the first model; the drawback is a more complicated object function.

Keywords: digital circuit sizing, spatial correlation, gate length variation, circuit performance, variability

1. INTRODUCTION

As device dimensions get smaller, the impact of process variation and the uncertainty introduced to circuit performance such as circuit delay or power consumption becomes a serious issue. From a design perspective, the general approach to combat such uncertainty is to identify sources of device parameter variation through test structure design or simulation, and then choose appropriate design methods to minimize the impact of the sources of variation identified. However, as circuits become complicated, a more systematic approach to circuit design is required. In particular, it will be beneficial to design an automated program that can adjust the size of each transistor in a circuit to achieve such a goal. As a result of process variation, a manufactured circuit may have a delay that differs significantly and randomly from the designed value. One way to characterize circuit sensitivity to such process variation is to evaluate the upper 95% quantile[2] of the distribution, understanding that reducing this quantile enhances robustness against process variation.

In a conventional deterministic circuit design where no process variation is taken into account, the critical path(s) of a circuit is fixed, and the design objective is to minimize the delay of the critical path(s). In a statistical design, however, any paths in the circuit may become a critical path due to the randomness introduced into the transistor delay through process variation. Therefore, in the latter case, the optimization of the circuit delay and its sensitivity to process variation has to take into account all possible paths in the circuit. In the work by Boyd [1], a geometric programming approach is used for this purpose. Denoting the gate size (width) for transistor i in a circuit as x_i, the objective of optimization is given as

$$\min\{\max_{p \in \{all\ \ circuit\ \ paths\ \}} \sum_{i \in p} [\overline{D}_i + k\sigma_i(D)]\}, \tag{1}$$

Subjects to the constraints

$$\sum_i x_i A_i \le A, \qquad x_i \ge 1, \tag{2}$$

[1] Tel: (510) 6429584; FAX: (510) 6422739; Email: qiantang@eecs.berkeley.edu
[2] The upper 95% quantile of a distribution of a random variable is defined as the value that is being exceeded by 5% of the population.

Design for Manufacturability through Design-Process Integration, edited by Alfred K. K. Wong,
Vivek K. Singh, Proc. of SPIE Vol. 6521, 65210P, (2007) · 0277-786X/07/$18 · doi: 10.1117/12.711794

where in equation (1), $\overline{D}_i$ denotes the nominal delay for transistor i, calculated from the standard resistor-capacitor (RC) delay model for transistors; k is a weighting factor that determines the relative importance of minimizing the variance of the delay as opposed to minimizing the nominal delay, and the value, typically around 2, that gives the best upper 95% quantile is used; $\sigma_i(D)$ denotes the standard deviation of the delay distribution of a transistor, and it is given by the following equation

$$\sigma_i(D) = \gamma(x_i^{-1/2})\overline{D}_i, \tag{3}$$

Derived from Pelgrom's model [2], which states that the standard deviation of the delay is inversely proportional to the square root of the area of the device. The proportionality factor γ is a model constant typically defined through measurements. Pelgrom's model is discussed in some detail in Section 3.1. In equation (2), A_i is the minimum area of transistor i, x_iA_i is the actual area of transistor i and A is a predefined maximum area allowed for the circuit under optimization; the summation is over all transistors in the circuit.

It is shown in [1] that compared with a deterministic design method, the geometric programming approach can reduce the circuit delay variation by half while improve the 95% quantile of circuit delay distribution by 6%. However, a number of limitations is noted in the design model presented in [1]: 1) only the effect of threshold voltage is taken into account in formulating the circuit delay variation while variation in other parameters, such as effective gate length, is ignored; and 2) only across wafer variation is considered in the formulation, while the with-in die and layout dependant variation is ignored. As a result, the model in [1] gives a too simplistic view of the entire problem. In this paper, new design models that capture all sources of variation including position dependant variation are presented. The performance of the new models is evaluated using a 32-bit Ladner-Fischer adder circuit [5], and results show noticeable improvements over the simple design model presented in [1].

2. SPATIAL CORRELATION BASED MODELLING

In a simplified view, the various sources of spatial process variation, e.g. the across wafer, within field and layout dependant variation, and the interactions between them can be lumped together and modeled by an empirically derived position dependant correlation term. Though this design method appears to be insufficient for a rigorous characterization of all process variation components, it does give some performance improvements over existing models in the statistical circuit sizing problem and provides some insight into the relative impacts of various components on circuit sensitivity to process variation. Furthermore, it is less computational intensive than the full model described in section 3.

2.1 Modeling of Spatial Correlation

Experiments showed that the spatial correlation between any two devices on a die and separated by some distances d can be modeled using the following piece-wise linear model [3]:

$$\rho_{ij} = \begin{cases} 1 - d_{ij} / X_L(1-\rho_B) & d_{ij} \le X_L \\ \rho_B & d_{ij} \ge X_L \end{cases}, \tag{4}$$

where ρ_{ij} represents the correlation coefficient between transistor i and j; d_{ij} denotes the separation distance transistor i and j; ρ_B is defined as the characteristic correlation baseline determined experimentally, and its value typically around 0.2; and X_L is the characteristic correlation length. The spatial correlation coefficient described by this model decreases linearly as the separation distance increases, and stays constant at ρ_B after the separation is larger than the characteristic correlation length.

2.2 Circuit Size Optimization

For a circuit consisting of n gates, the delay of each gate can be considered as a random variable having a Gaussian distribution with its own mean and variance. The circuit delay is the maximum (over all circuit paths) of the summation of each individual gate delays on a path, and it is, therefore also a random variable that can be approximated by a Gaussian-shaped distribution. The mean of the individual gate delay distribution can be calculated using the simple Resistance-Capacitance (RC) based delay model. The standard deviation (or the square root of the variance) of the gate delay of a transistor can be decoupled into two components: one arises from the threshold voltage (V_{th}) variation, and the other arises from the effective gate length (L_{eff}) variation. The contribution of both of these variation components to the gate delay variation can be derived from the Pelgrom's model [2] and the alpha-power law model [6]. In the ideal case

where the contributions from V_{th} and L_{eff} variations act additively, the standard deviation of the delay distribution of a particular transistor *i* can therefore be modeled[3] as:

$$\sigma_i(D) = \gamma_{Vth}(x_i^{-1/2})\overline{D}_i + \gamma_{Leff}(x_i^{-1/2})\overline{D}_i, \tag{5}$$

Where $\overline{D}_i$ is the mean of the gate delay for transistor *i*; γ_{Vth} and γ_{Left} is the model parameter for V_{th} caused variation and L_{eff} caused variation, respectively.

To find the optimal assignment for transistor widths that minimizes the 95% quantile of the circuit delay distribution, while takes into account of the different sources of process variation, the following objective function and constraints are used:

$$\text{Minimize: } \max_{p \in \{all\ circuit\ paths\}} \sum_{i \in p}[\overline{D}_i + w_1\sigma_i(D)] + w_2 \sum_{i,j \in p, i \neq j}[\rho_{ij}\sigma_i^2\sigma_j^2], \tag{6}$$

$$\text{Subject to: } \sum_i x_i A_i \leq A, \qquad x_i \geq 1. \tag{7}$$

where $\sigma_i(D)$ in equation (6) is calculated using equation (5) and ρ_{ij} is calculated using equation (4); the constants w_1 and, w_2 are the weighting factors that determine the relative importance of the individual transistor delay variation and the spatial correlation between transistor delays to the circuit sensitivity to process variation. For every path in the circuit, we sum together the following three terms: 1) the sum of the nominal delay of each transistor in the path, 2) the sum of the standard deviation of each transistor delay distribution weighted by w_1, and 3) the sum of all covariance terms between any two transistors in the path and weighted by w_2. The objective is to minimize the maximum of this total sum over all possible paths in a circuit. The constraints of the optimization problem are shown in equation (7), with the variables defined as in equation (2). The constraints simply state that all gate sizes have to be larger or equal than the minimum feature size and total circuit area must be smaller than some predefined value *A*. This optimization problem is solved using a geometric programming solver such as GGPLAB [9] or MOSEK [10].

2.3 Simulation Results and discussion

A 32-bit Ladner-Fisher adder is used as the test circuit to analyze the performance of the above formulation. To obtain the delay distribution of the circuit, a Monte Carlo analysis of 5000 samples is carried out after the circuit sizing is optimized using the optimization procedure described in section 2.2. For each Monte Carlo sample, the individual transistor delay is obtained by random sampling from the delay distribution characterized by a Gaussian distribution with mean equal to the nominal delay of a gate, standard deviation given by equation (5), and a correlation coefficient given in equation (4). The delay of the entire circuit is calculated by taking the maximum delay over all circuit paths. Tables 1 and 2 show the value of all parameters used in the simulation. A histogram that approximates the circuit delay distribution is obtained and shown in Figure 1. Table 3 lists the mean, standard deviation and upper 95% quantile of the circuit delay distribution for the various design methods.

Table 1. The various model parameters used in the Monte Carlo simulation.

Parameter name	Value
γ_{Vth}	0.15
γ_{Leff}	0.1
w_1	2
w_2	2.5
ρ_B	0.2
X_L	10mm
Separation distance between gates	100μm
A	15000 μm^2

[3] In reality, however, short-channel effects will result in some correlation between V_{th} variation and L_{eff} variation. A correlation term may be added to equation (5) for modeling such effect.

Table 2. The various gate types in the Ladner-Fisher Adder circuit and the gate parameters used in the simulation described in [1].

Gate Type	Input Capacitance (fF)	Intrinsic Capacitance (fF)	Resistance (kΩ)	Area (μm^2)
INV	3	3	0.48	3
NAND2	4	6	0.48	8
NOR2	5	6	0.48	10
AOI21	6	7	0.48	17
OAI21	6	7	0.48	16

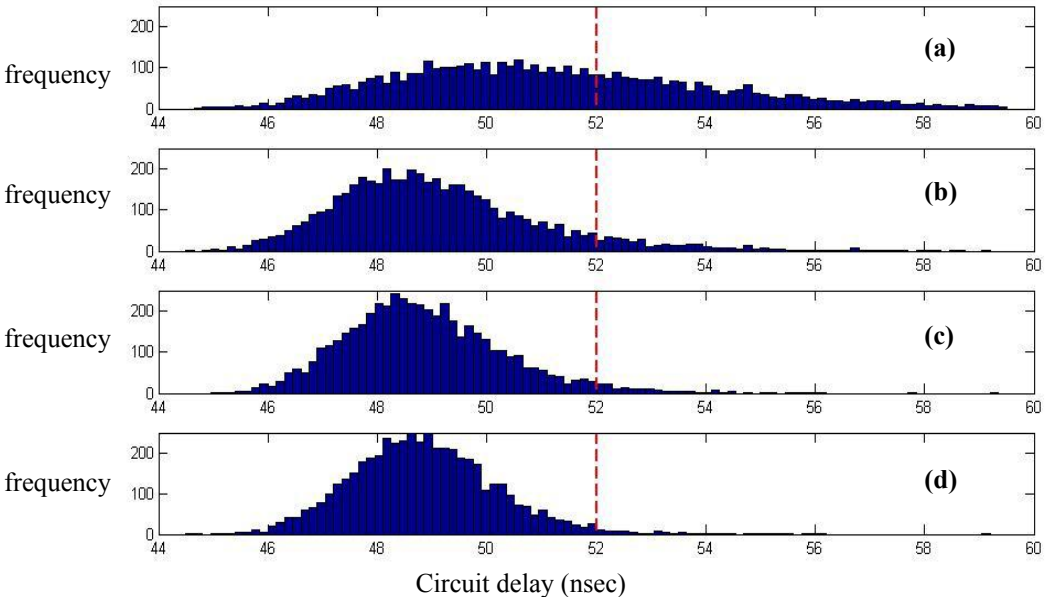

Figure 1. Histograms of the circuit delay and yield improvements with reference to a 52 nsec specification for different designs: (a) deterministic design, yield is 63%; (b) statistical design that considers only variation due to V_{th}, yield is 92%; (c) statistical design that considers variations due to both V_{th} and L_{eff} , yield is 96%; (d) statistical design that considers variations due to V_{th}, L_{eff} and position dependant correlation, yield is 98%.

Table 3. Comparison of different design methods.

Design Methods	Mean (nsec)	Std. Dev (nsec).	95% quantile (nsec)
(a) Deterministic Design	51.3	3.02	56.9
(b) Statistical Design, variation due to V_{th}	49.1	1.96	54.7
(c) Statistical Design, variation due to V_{th} and L_{eff}	48.9	1.52	52.6
(d) Statistical Design, variation due to V_{th} , L_{eff} and spatial correlation	48.9	1.36	51.5

Two major conclusions can be drawn from the simulation results: 1) the modified design model that includes the variation due to V_{th}, L_{eff} and spatial correlations has a superior performance than the model presented in [1] in terms of both reduced variation and upper quantile in circuit delay; 2) the performance improvements of the design model that includes spatial correlation (i.e. model (d)) is, on the other hand, marginal over the model that does not (i.e. model (c)). These conclusions suggest that the spatial correlation based model, even though it does lead to some degree of performance improvement, is nevertheless still inadequate for modeling the complex impact of process variation.

3. MODIFIED PELGROM'S MODEL AND APPLICATION TO CIRCUIT OPTIMIZATION

Elsewhere [11], it is shown that what is often referred to as "spatial correlation" is an artifact of un-modeled residuals after the decomposition of deterministic variation components. Therefore, a potentially better method for analyzing circuit performance and its response to process variation is to model the spatial or position dependant component implicitly rather than explicitly. The following sections discuss this modeling approach in more detail. In order to derive a correct representation of the distribution of the entire gate delay, the sources of variation are analyzed and modeled through a modified version of the Pelgrom's model [2]. Three main components are identified, namely across wafer variation, within field variation and layout dependant variation.

3.1 Pelgrom's Model

Pelgrom et al [2] derived a model for calculating the parameter mismatch between two rectangular devices of identical width and length and separated by a distance D_x. By decomposing the mismatch generation process into two components: 1) the random component or "white noise", and 2) a stochastic process with relatively long correlation distance, Pelgrom showed that for any parameter P of interest, the variance of the mismatch in P is given by

$$\sigma^2(\Delta P) = \frac{A_p^2}{WL} + S_p^2 D_x^2, \tag{8}$$

The fist term in equation (8) accounts for the white noise component, while the second term accounts for the systematic variation. The constant A_p is the area proportionality constant for parameter P, while S_p describes the variation of parameter P with spacing. This model suggests that a spatial correlation structure is inherent in the physical properties of semiconductor devices; specifically, devices situated closely together will be subject to a higher degree of correlation than devices separated by larger distances. The stochastic nature of the mismatch in parameter P comes from the fact that as the wafer is diced into chips, the exact position of the chip on the wafer is unknown. Thus the exact magnitude of the wafer-level variation for a chip is a random distribution rather than a deterministic function of position.

By a simple algebraic manipulation, as shown in equation (9), it is obvious that the so called "spatial correlation" between two devices is embedded in the mismatch model:

$$\begin{aligned}\sigma^2(\Delta P) &= \sigma^2[P(x_1,y_1) - P(x_2,y_2)] \\ &= \sigma^2[P(x_1,y_1)] + \sigma^2[P(x_2,y_2)] - 2Corr[P(x_1,y_1), P(x_2,y_2)]\end{aligned} \tag{9}$$

Where $P(x_1,y_1)$ and $P(x_2,y_2)$ is the parameter value for the two devices located at positions (x_1, y_1) and (x_2, y_2), respectively. By comparing the terms in equation (9) with those in equation (8), one can conclude that the first term $\frac{A_p^2}{WL}$ in Pelgrom's model corresponds to the variance of the parameter, while the second term $S_p^2 D_x^2$ corresponds to the so called "spatial correlation" between the distributions of the parameter of two devices. In the context of circuit sizing optimization, Pelgrom's model as represented in equation (8) can be used to model the variance of the gate delay. The major difference between this approach and that described in section 2 is that both terms in Pelgrom's model are used while in the modeling method of section 2, only the first term that captures the white noise component is used for deriving the variance of the gate delay.

A major assumption in deriving Pelgrom's model is that the only systematic variation that affects the parameter value is a parabolic shaped, wafer-level variation. It has been observed in [3], however, that variation introduced in manufacturing process also includes significant systematic variation patterns across each die, and a layout dependant variation component. Therefore, the second term of Pelgrom's model needs to be modified before it can be applied to the circuit size optimization problem.

3.2 Sources of variation

The first type of process variation is a deterministic and slowly varying component that varies across the entire wafer, often a parabolic shaped function centered at the middle of the wafer. Such variation can be a consequence of, for example, the loading effects in etching or deposition. For any parameter P (e.g. transistor width, gate length or threshold

voltage), the model described in section 3.1 gives an accurate representation of the variance of the mismatch in the parameter, $\sigma^2(\Delta P)$, between two rectangular devices due to wafer-level variation. Clearly, this type of variation introduces a random component to the parameter value for each device in a die. Figure 2(a) plots the parabolic shaped wafer-level variation vs. the radial distance r measured from the center of the wafer. Figure 2(b) shows a simulated plot of $\sigma^2(\Delta P)$ and Figure 2(c) shows the distribution of ΔP vs. the Euclidean separation distance between two devices.

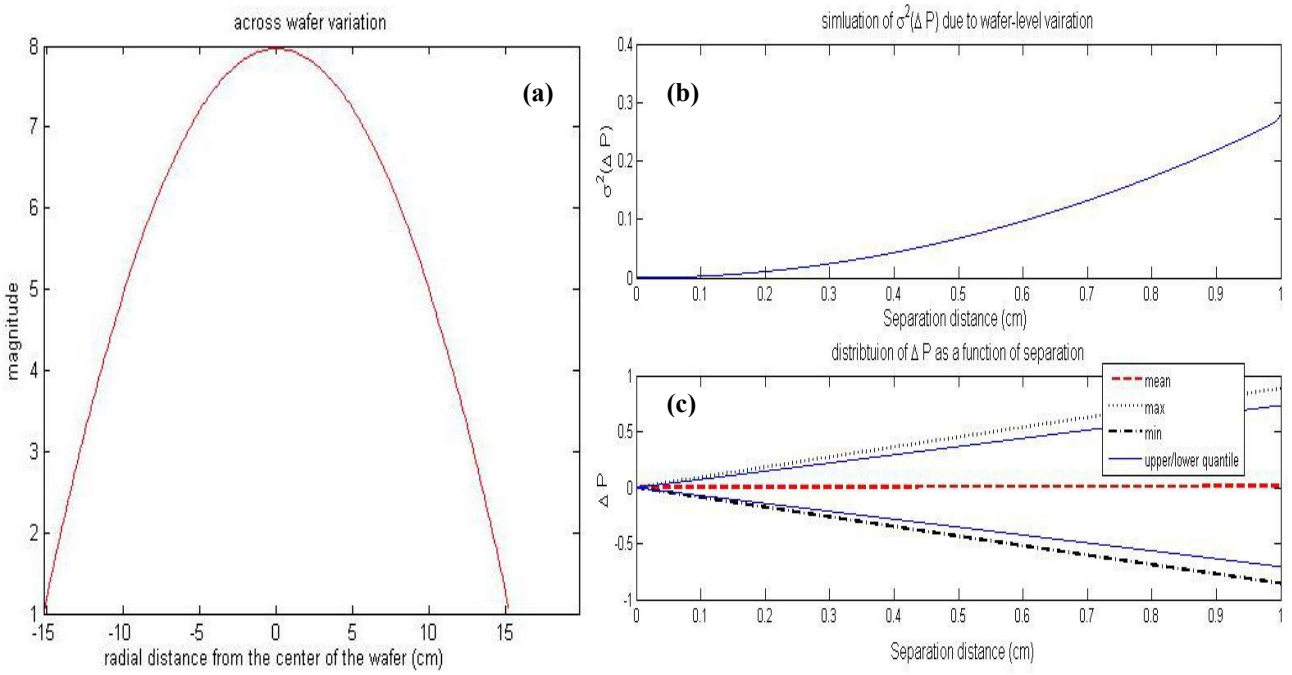

Figure 2. Simulation of the across wafer variation and the mismatch between two devices: (a) Parabolic shaped wafer-level variation vs. the radial distance r measured from the center of the wafer; (b) a simulated plot of $\sigma^2(\Delta P)$ vs. separation distance; (c) the distribution of ΔP vs. the separation distance between two devices.

The second type of process variation is a within field component that varies systematically and deterministically across a single die, but is identical for all dies in a wafer, mainly a consequence of mask errors and systematic variations in the exposure tool [3]. Such within field variation can also be approximated by a parabolic shaped function centered at the middle of a die as illustrated in figure 3(a). In terms of $\sigma^2(\Delta P)$, Pelgrom's model no longer gives an accurate predication of its relationship with device separation, D_x, due to this type of variation. It is found through simulation that instead of a quadratic dependence on D_x, $\sigma^2(\Delta P)$ exhibits a bell-shaped function with respect to D_x, where it is the smallest when the separation distance is either 0 or the size of the die, and largest when the separation distance is half of the die size, as illustrated in figure 3(b). Figure 3(c) shows the distribution of parameter mismatch, ΔP, as a function of device separation. Unlike the wafer-level variation, the within field variation is identical for all dies on a wafer. Therefore no stochastic component to the parameter, instead, a deterministic bias is introduced to the value of the parameter, and the magnitude of this bias depends on the position of the device on the die.

The manufactured width, length and threshold voltage of a device also depends on nearby structures. For example, the threshold voltage of a device can be modified due to the nearby STI structure or well extensions. Such layout or local density dependant effects can also be introduced during etching, lithography or CMP processing steps. Two approaches can be used to model the layout dependant effects: one is to identify all such sources of process variation and model each of them separately; the other is to make use of appropriately designed test structures and to extract the dependence of various parameters on layout or local density empirically. In the second approach, a moving window of fixed size and centered at the device of interest is often created, and the number of devices in this window is recorded to determine the local density. Then a simple curve fitting technique can be used to determine the relation between device parameter and the local density. Once the shape of the layout-dependant variation is determined, the quantity $\sigma^2(\Delta P)$ can be obtained through simulation. Figure 4 shows the simulated $\sigma^2(\Delta P)$ and the distribution of ΔP for an arbitrary layout-

dependant variation function. Again, since layout dependant variation is local and deterministic for a given layout on a die, such variation introduces only a deterministic bias to the device parameter.

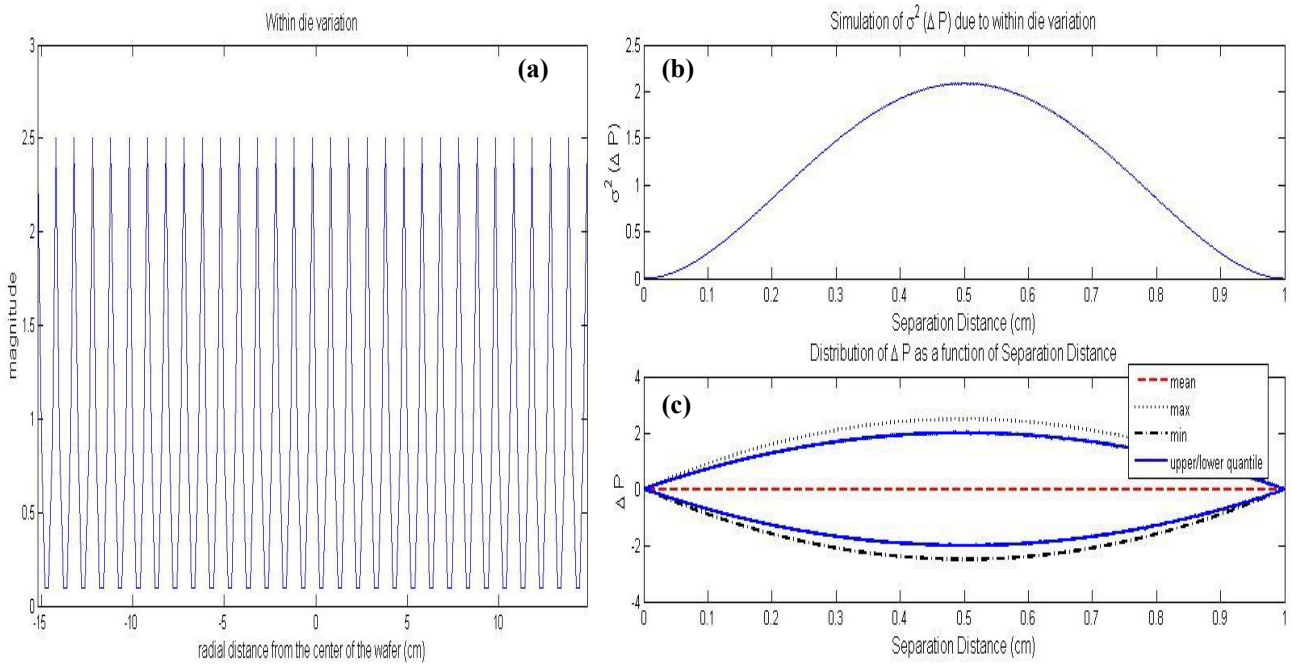

Figure 3. Simulation of the within die variation and the mismatch between two devices: (a) Parabolic shaped within die variation vs. the radial distance measured from the center of the wafer; (b) a simulated plot of $\sigma^2(\Delta P)$ vs. separation distance; (c) the distribution of ΔP vs. the separation distance between two devices

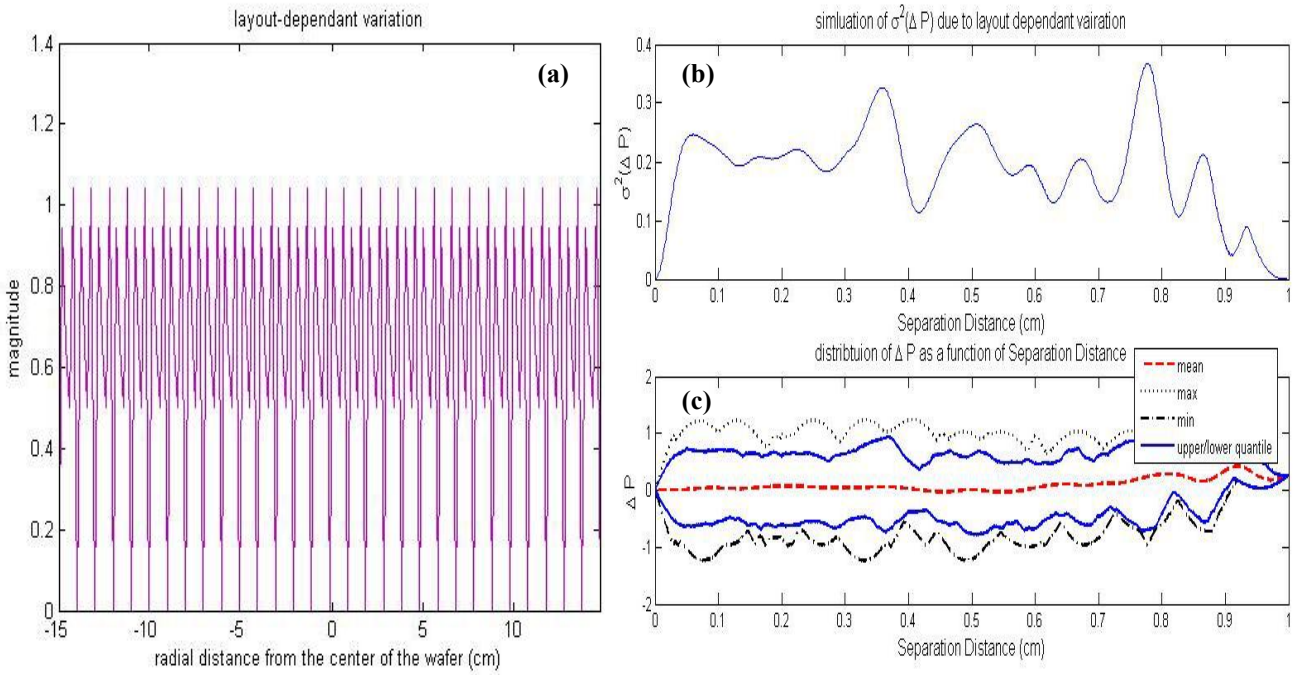

Figure 4. Simulation of the layout dependant variation and the mismatch between two devices: (a) Layout dependant variation pattern vs. the radial distance measured from the center of the wafer; (b) a simulated plot of $\sigma^2(\Delta P)$ vs. separation distance; (c) the distribution of ΔP vs. the separation distance between two devices

3.3 Formulation of the circuit sizing problem

For a particular device i located at position (x_i,y_i) on a die, the measured transistor width, length and threshold voltage is randomly distributed as a Gaussian distribution with some mean and variance. The mean and variance for the effective gate length for device i are given in equation (10) and (11), where L_i is the designed value of the gate length; $Density\_L(x_i,y_i,x_n,y_n)$ is the deviation of measured gate length from the designed value due to the effect of local density or circuit layout. The amount of the deviation depends both on the position of device i given by (x_i,y_i) and the position of its neighboring devices given by (x_n,y_n). $Field\_L(x_i,y_i)$ describes the contribution of the within field variation on gate length and it is a function of the device position.

$$\mathrm{E}(\mathrm{L}_{\mathrm{eff,i}}) = L_i + Density\_L(x_i, y_i, x_n, y_n) + Field\_L(x_i, y_i), \tag{10}$$

$$\mathrm{var}(\mathrm{L}_{\mathrm{eff,i}}) = \frac{A}{W_i L_i^2} + \sigma^2_{wafer,L}(x_i, y_i), \tag{11}$$

In equation (11), the first term accounts for the randomness introduced by white noise, as described in Pelgrom's model. The second term accounts for the randomness introduced by the wafer-level variation of the gate length, and it has the form shown in figure 2(b).

The mean and variance of the distributions of the transistor width and the threshold voltage of device i can be determined in a similar manner, and it is shown in equations (12)-(15).

$$\mathrm{E}(\mathrm{W}_{\mathrm{i}}) = W_i + Density\_W(x_i, y_i, x_n, y_n) + Field\_W(x_i, y_i), \tag{12}$$

$$\mathrm{var}(\mathrm{W}_{\mathrm{i}}) = \frac{A}{W_i^2 L_i} + \sigma^2_{wafer,w}(x_i, y_i), \tag{13}$$

$$\mathrm{E}(\mathrm{V}_{\mathrm{th,i}}) = V_{th,i} + Density\_V_{th}(x_i, y_i, x_n, y_n) + Field\_V_{th}(x_i, y_i), \tag{14}$$

$$\mathrm{var}(\mathrm{V}_{\mathrm{th,i}}) = \frac{A}{W_i L_i} + \sigma^2_{wafer,V_{th}}(x_i, y_i). \tag{15}$$

The delay of the transistor i can be expressed in terms of the random variables $\mathrm{W_i}$, $\mathrm{L_{eff,i}}$ and $\mathrm{V_{th,i}}$ as shown below:

$$\mathrm{D_i} = 0.69 R_i (C_i + C_L) \propto \frac{\mathrm{L_{eff,i}}}{\mathrm{W_i}} (Vdd - \mathrm{V_{th,i}})^{\alpha} (C_{\mathrm{int}} \mathrm{W_i} \mathrm{L_{eff,i}} + \sum_{f \in fanout} C_{\mathrm{int}} \mathrm{W_f} \mathrm{L_{eff,f}}), \tag{16}$$

Where R_i and C_i denotes the resistance and the capacitance of transistor i respectively; C_L denotes the loading capacitance of transistor i while C_{int} denotes the intrinsic capacitance.

The delay given in equation (16) is itself a random variable; its mean and variance can be determined by simply taking the expectation and variance of the above equation. Through simple algebraic manipulations, the mean can be explicitly evaluated, and the variance can be approximated, as follows:

$$\mathrm{E}(\mathrm{D_i}) = \frac{\mathrm{E}(\mathrm{L_{eff,i}})}{\mathrm{E}(\mathrm{W_i})} [Vdd - \mathrm{E}(\mathrm{V_{th,i}})]^{\alpha} [C_{\mathrm{int}} \mathrm{E}(\mathrm{W_i}) \mathrm{E}(\mathrm{L_{eff,i}}) + \sum_{f \in fanout} C_{\mathrm{int}} \mathrm{E}(\mathrm{W_f}) \mathrm{E}(\mathrm{L_{eff,f}})] \tag{17}$$

$$\begin{aligned} \mathrm{var}(\mathrm{D_i}) = & \left|\frac{\partial \mathrm{D_i}}{\partial \mathrm{L_{eff,i}}}\right|^2 \mathrm{var}(\mathrm{L_{eff,i}}) + \left|\frac{\partial \mathrm{D_i}}{\partial \mathrm{W_i}}\right|^2 \mathrm{var}(\mathrm{W_i}) + \left|\frac{\partial \mathrm{D_i}}{\partial \mathrm{V_{th,i}}}\right|^2 \mathrm{var}(\mathrm{V_{th,i}}) \\ & + \sum_{f \in fanout} \left[\left|\frac{\partial \mathrm{D_i}}{\partial \mathrm{L_{eff,f}}}\right|^2 \mathrm{var}(\mathrm{L_{eff,f}}) + \left|\frac{\partial \mathrm{D_i}}{\partial \mathrm{W_f}}\right|^2 \mathrm{var}(\mathrm{W_f}) \right] \end{aligned} \tag{18}$$

The circuit size optimization is formulated in the following form:

$$\begin{aligned} &\text{Objective:} \quad \min\{\max_{p \in \{all\ circuit\ paths\}} \sum_{i \in p} [\overline{D_i} + k \times \sigma(D_i)]\} \\ &\text{Constraints:} \quad \sum_i x_i A_i \le A,\ x_i \ge 1 \end{aligned} \tag{19}$$

where $\overline{D_i}$ is the mean value of the gate delay for transistor i given in equation (17); $\sigma(D_i)$ equals to $[Var(D_i)]^{1/2}$. The design variables in this optimization problem is the transistor gate width (or size) W_i's. The above problem can be solved using standard geometric programming solvers such as GGPLAB[9] or MOSEK[10].

4. SIMULATION RESULTS AND DISCUSSION

A circuit of 6-level gates each having a fan-out of 2, as illustrated in figure 5, is used to evaluate the performance of the optimization procedure. To obtain the delay distribution of the circuit, a Monte Carlo analysis of 5000 samples is carried out after the circuit sizing is optimized using the optimization procedure described in section 3.3. For each Monte Carlo sample, the individual transistor delay is obtained by random sampling from the gate delay distribution characterized by a Gaussian distribution with mean and variance given by equation (17) and (18). The distribution of the gate delay for each transistor is assumed to be independent of the others since the spatial correlation factor is already implicitly embedded into the spatial variance expression. The delay of the entire circuit is calculated by taking the maximum delay over all circuit paths. A histogram that approximates the circuit delay distribution is obtained and shown in Figure 6(b); the histogram obtained from deterministic design is plotted on the same figure**.** For comparison purpose, the circuit delay histograms obtained by using the optimization procedure described in section 2.2 and by using the corresponding deterministic design are plotted in figure 6(a). Table 4 lists the mean, standard deviation, upper 95% quantile and the percentage improvement in upper 95% quantile over the deterministic design of the circuit delay distribution for each design.

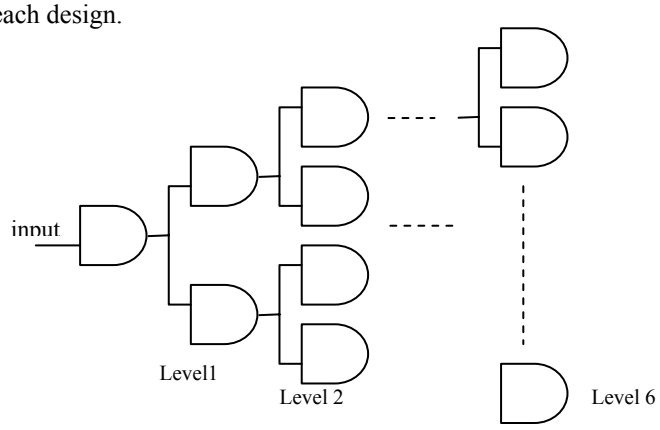

Figure 5. Schematic of the test circuit used in Monte Carlo simulation.

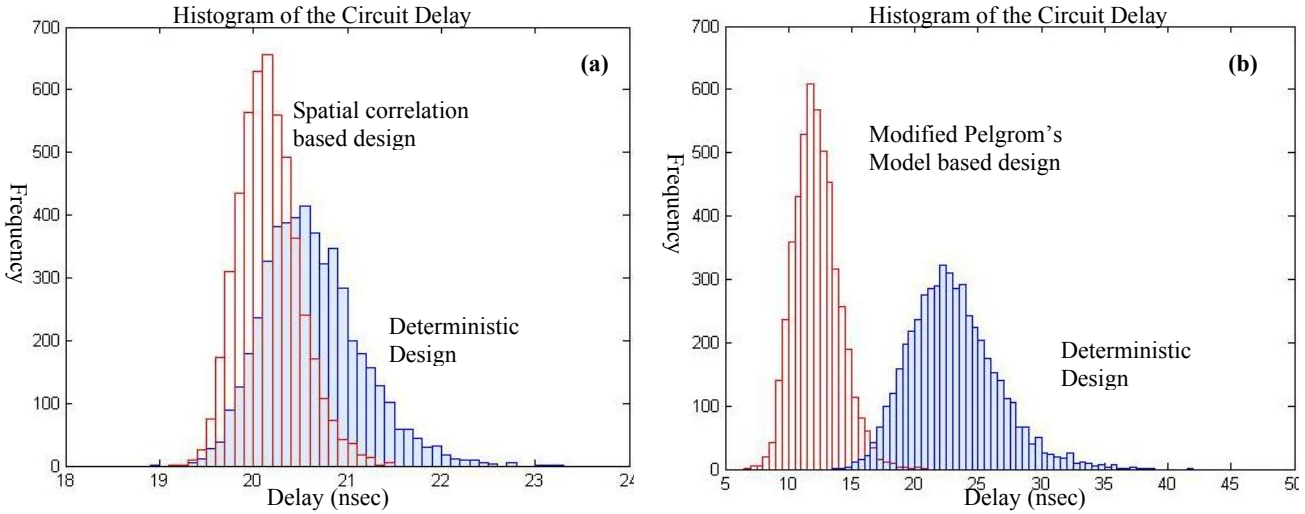

Figure 6. Histograms of the circuit delay for spatial correlation based design and modified Pelgrom's model based design, obtained from Monte Carlo simulations: (a) Comparison of the spatial correlation based design and the corresponding deterministic design; (b) Comparison of the modified Pelgrom's model based design and the corresponding deterministic design

Table 4. Comparison of different design methods.

Design	Mean **E(D)** (nsec)	Std. Dev (nsec)	Upper 95% quantile (nsec)	Percentage Improvement (%)
Spatial Corr. Based	20.2	0.32	20.7	4.2
Deterministic Design	20.6	0.52	21.6	
Pelgrom Based	12.2	1.77	15.4	47.3
Deterministic Design	23.0	3.48	29.2	

Figure 7 shows the comparison of the execution time between the spatial correlation based model and the modified Pelgrom's model based model. The execution time is measured on a 2.8GHz Pentium D processor computer, using sample circuits similar to that in figure 5, with different number of levels. The solver used is the MATLAB based GGPLAB solver[9]. Based on the simulation, the computation complexity of both design model is on the order of $O(x^{1.5})$ with respect to the number of gates in the circuit. However, the computation time for the Pelgrom based model is about five times larger the spatial correlation based model. By extrapolating to large circuit of, for example, 2,000 gates, the estimated computational time is calculated to be 1.5 hours and 7 hours for the two models respectively. Further reduction in execution time can be obtained by optimizing the geometric programming solver or by adopting a C/C++ based solver.

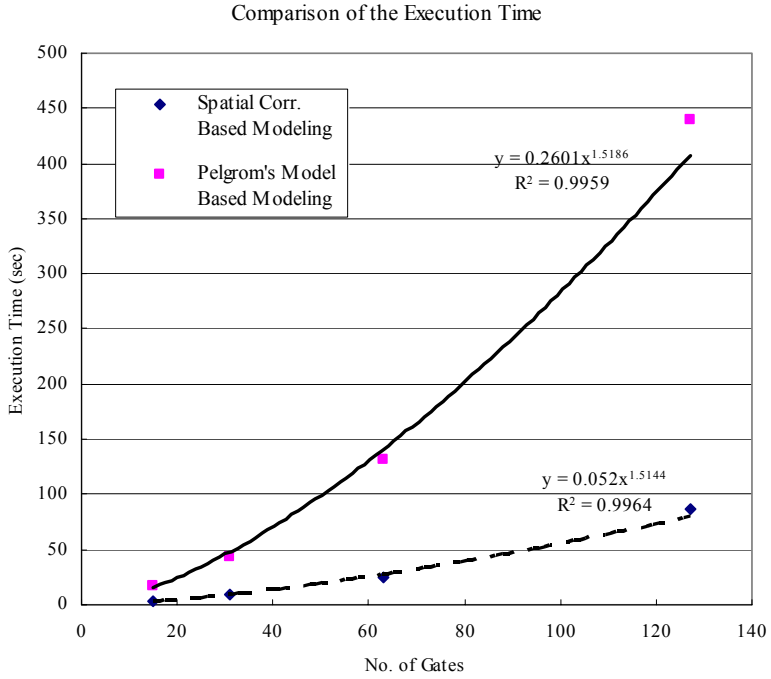

Figure 7. Comparison of the execution time between different design methods.

5. CONCLUSIONS

New design models that include different sources of variation and position dependant variation have been proposed for statistical circuit sizing problem. The position dependant variation can be modeled either explicitly, as an extra correlation term that depends on the separation distance, or implicitly by using a modified version of the Pelgrom's model in problem formulation. Monte Carlo analysis shows that both approaches have noticeable performance improvements over the existing design models, including both a reduced variance and reduced upper quantile of the

distribution of the circuit delay. The first approach affords a simpler objective function and thus requires less computational time than the second approach, but appears to be inadequate to fully model the position dependant variation and often has an inferior performance to the design approach that uses the modified Pelgrom's model.

ACKNOWLEDGEMENTS

This work is funded by SRC grant 1324.001.

REFERENCES

1. S. Boyd, S.-J. Kim, D. Patil, and M. Horowitz, "A Heuristic Method for Statistical Digital Circuit Sizing," *31st SPIE International Symposium on Microlithography,* February 2006.
2. M. Pelgrom, A. Duinmaijer, and A. Welbers, "Matching properties of MOS transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, no. 1, pp. 1433-1439, Oct. 1989.
3. P. Friedberg, Y. Cao, J. Cain, R. Wang, J. Rabaey, and C. Spanos, "Modeling Within-Die Gate Length Spatial Correlation for Process-Design Co-Optimization," *Design and Process Integration for Microelectronic Manufacturing III,* Proceedings of SPIE vol. 5675, Feb. 2005.
4. B. E. Stine, D. S. Boning, and J. E. Chung, "Analysis and decomposition of spatial variation in integrated circuit processes and devices," IEEE TSM, vol. 10, no. 1, Feb. 1997
5. S. Knowles, "A family of adders," in *Proceedings of 14th IEEE Symposium on Computer Arithmetic*, pp. 30–34, IEEE Computer Society Press, 1999.
6. T. Sakurai and A. Newton, "Alpha-power law MOSFET model and its application to CMOS inverter delay and other formulas," *IEEE Journal of Solid-State Circuits*, 25(2), pp. 584-593, 1990.
7. K. Okada, H. Onodera, and K. Tamaru, "Layout dependant matching analysis of CMOS circuits," *Analog integrated circuits and signal processing*, 25, pp. 309-318, 2000
8. S. Boyd, S.-J. Kim, L.Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," 2006. *Optimization and Engineering.* Available from www.stanford.edu/boyd/~gp_tutorial.html.
9. A. Mutapcic, K. Koh, S.-J. Kim, L. vandenberghe, and S. Boyd, *GGPLAB: A simple Matlab Toolbox for Geometric Programming,* 2005. Available from www.stanford.edu/~boyd/ggplab/.
10. MOSEK ApS, *The MOSEK Optimization Tools Version 2.5. User's Manual and Reference,* 2002. Available from www.mosek.com.
11. P. Friedberg, G. Cheng, Q. Y. Tang, and C. J. Spanos, "Modeling Micron-Scale Gate Length Variation and Spatial Correlation," to appear in *SPIE International Symposium on Microlithography*, 2007.