Path-Based Statistical Static Timing Analysis for Large Integrated Circuits in a Weak Correlation Approximation

Dmytro Mishagli, Eugene Koskin and Elena Blokhina
School of Electrical and Electronic Engineering, University College Dublin, Ireland

***Abstract*—This work is aimed at the development of a path-based approach to Statistical Static Timing Analysis. Timing Analysis is an absolutely essential step in the verification of Very Large Scale Integration (VLSI) designs. We propose a novel analytical methodology for the fast calculations of VLSI delay. The problem is stated in such a way that becomes equivalent to finding the maximum of a large set of correlated random variables (RVs). For this purpose, a corresponding extension of extreme value theory of weakly-correlated RVs is developed. Results of simulations showing a comparison of our approach with Monte Carlo simulations are presented. Possible applications, extensions of our methodology and future steps are discussed.**

***Index Terms*—Timing Graph, Statistical Timing Analysis (SSTA), Delay, Very Large Scale Integration (VLSI), Extreme Value Distribution**

I. Introduction

The timing verification of digital integrated circuits (ICs) is an absolutely essential step: scaling down analogue and digital ICs poses very strict restrictions on their operation, in particular, in terms of timing and delay. Thus, timing analysis (TA) tools are used in order to verify and optimise digital design before fabrication [1]–[5]. However, with the reduction of feature size, the loss of fabricated ICs only increases, since the effect of uncertainties on the performance of circuits increases dramatically.

The impact of uncertainties, such as process, voltage or temperature variations, has risen significantly [6], since the nominal values of physical parameters (transistor channel length, etc.) are of the same order of magnitude as their deviations from these values. Additionally, the development of low-power applications [7], [8], such as Internet of Things and cloud technologies, brings only new challenges to IC design. An approach to handle these challenges is to treat circuit variations from the very beginning within statistical frameworks. Thus, statistical static timing analysis (SSTA) has been developing as a promising tool to replace deterministic TA [9].

Within SSTA, all delays are treated as random variables (RVs) with corresponding probability density functions (p.d.f.). There are three main challenges of SSTA that have been studied with different degrees of success: (i) impact of correlations, (ii) non-analytical operations such as $\max$, and (iii) non-Gaussian distributions of variations. We address the interested reader to the review articles [10]–[13] for further details.

Recently, an approach within the block-based framework was proposed in [14]. The authors considered delay propagation through a single gate and developed an algorithm for representing skewed distributions via Gaussian kernel functions. Being accurate, the proposed algorithm may still be computationally expensive since it requires evaluation of every gate or block of gates in an IC. In this study, we propose an alternative to [14] within the path-based framework. The path-based approach operates with sets of paths within one circuit. This allows one to describe topological correlations (due to shared paths) in a simple manner.

The paper is organised as follows. In Section II, we outline the basics of path-based SSTA, make necessary statistical preliminaries and present a general statement of the problem. In Sections III and IV, we develop an asymptotic theory for extreme value statistics, assuming that correlations are weak. This theory yields the distribution of the maximal delay of an IC with topological correlations taken into account. Sections V and VI present a comparison of our approach with numerical simulations and discussion of the obtained results.

II. Statement of the Problem

A. Some Definitions

For clarity, let us first make the following definitions of the functions that are widely used in this paper.

The p.d.f. $\omega_0(\mathbf{r})$ of n uncorrelated normal variables with means μ_k and variances σ_k:

$$\omega_0(\mathbf{r}) \stackrel{\text{def}}{=} \frac{1}{(2\pi)^{n/2}} \prod_{k=1}^{n} \frac{1}{\sigma_k} \varphi\left(\frac{x_k - \mu_k}{\sigma_k}\right), \tag{1}$$

where $\varphi(x)$ reads

$$\varphi(x) \stackrel{\text{def}}{=} e^{-\frac{x^2}{2}}. \tag{2}$$

In the case of only one RV, $\omega_0(x) = \varphi\left(\frac{x-\mu}{\sigma}\right)/\sqrt{2\pi}\sigma$.

The cumulative distribution function (c.d.f.) $\Phi(x)$ of the standard normal distribution:

$$\Phi_0(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} \varphi(x')dx' = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{x'^2}{2}} dx'. \tag{3}$$

B. Static Timing Analysis

In timing analysis, an IC is considered as a timing graph $G(E, V)$ with a set of edges E (gates) and vertices V (interconnects), and the problem of TA then is to find the maximum delay for such a graph, or the longest path (see Fig. 1).

In path-based timing analysis, the maximal delay D of a circuit is equal to the maximum of delays along all paths:

$$D = \max(\tau_1, \tau_2, \ldots, \tau_n), \tag{4}$$

where $\tau_i = \sum_{j=1}^{m_i} \tau_i^{(j)}$ is an accumulated delay along a path i and m_i is a number of delays in that path.

C. Extreme Values Statistics

In SSTA, the delay D (as well as individual path delays τ_i) is a random variable. Therefore, the problem of delay determination is equivalent to the following well–posed problem in statistics: for a sequence of n RVs $X_1, X_2, \ldots, X_n$ with a p.d.f. $\omega(\mathbf{r})$, where $\mathbf{r}$ is the radius-vector in the n-dimensional space of RVs, find the distribution of the RV ζ such that

$$\zeta = \max(X_1, X_2, \ldots, X_n). \tag{5}$$

In the general case, for arbitrary distributed n correlated variables $X_1, \ldots, X_n$, the solution to problem (5) is unknown and finding it remains challenging. However, there are two known results that can simplify the problem:

(i) Sum of m independent RVs, each with a finite expected value μ_i and variance σ_i^2, converges to a normal distribution if $m \to \infty$ (Central Limit Theorem, CLT) [15].

(ii) RV (5), if $n \to \infty$ and the RVs X_i are independent and identically distributed, is distributed according to one of three asymptotic limit laws: Fréchet, Weibull and Gumbel distributions. The latter is of our interest to the present discussion, since it satisfies a desired asymptotics [16].

Since in modern VLSI the number of nodes in a graph are typically of order of 10^9 and greater, one can conclude that the number of delays m_i in a given path i is sufficiently large in order to satisfy the requirements of the CLT. Thus, without any loss of generality, we assume that all of the RVs X_i that describe the accumulated delay along the i^{th} path in a graph G, are Gaussian ones.

Also, the number n of paths can be extremely large, providing the asymptotic behaviour required in (ii). Thus, for the case of the independent RVs X_i, the solution to problem (5) is

$$\Psi_n(x) = \exp\left(-e^{-\frac{x-a_n}{b_n}}\right), \tag{6}$$

which is the c.d.f. of the Gumbel distribution, where

$$a_n = \Phi_0^{-1}\left(1 - \frac{1}{n}\right), \quad b_n = \frac{1}{n \cdot \omega_0(a_n)}. \tag{7}$$

The mean $\tilde{\mu}_n$ and variance $\tilde{\sigma}_n^2$ of the Gumbel distribution are also represented via the parameters a_n and b_n:

$$\tilde{\mu}_n = a_n + \gamma b_n, \quad \tilde{\sigma}_n^2 = \frac{\pi^2}{6} b_n^2, \tag{8}$$

where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant.

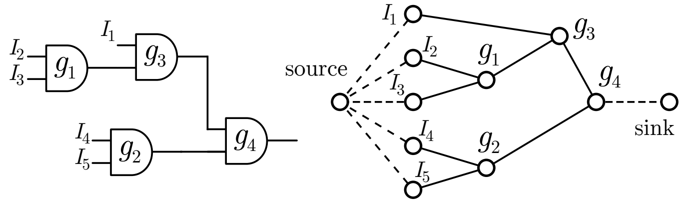

Fig. 1. An example combinational circuit with 5 inputs I_i and 4 gates g_j and its timing graph. There are $n = 5$ paths from source to sink in this example.

The illustration of the asymptotic solution (6) for the case of independent standard Gaussian RVs, $X_i \sim \mathcal{N}(0, 1)$, is shown in Fig. 2 for various n. One can observe *how* the distribution of ζ approaches to the Gumbel distribution (6), as the number n of terms increases.

D. Problem Formulation

In this study, we propose an approximate solution to problem (4) within the framework of path-based SSTA. Our aim is to take into account topological correlations in a graph. For this purpose, we consider the corresponding statistical problem (5).

We make the following assumptions: (i) all paths have a large number m_i of delays so that CLT holds, and the accumulated delay along the i^{th} path is described by a Gaussian RV; (ii) the number n of paths is large enough so that the maximum delay is given by the Gumbel law for the uncorrelated case.

The *aim of this work* is to find the distribution of the RV (5) in the case of the correlated RVs X_i assuming correlations are weak.

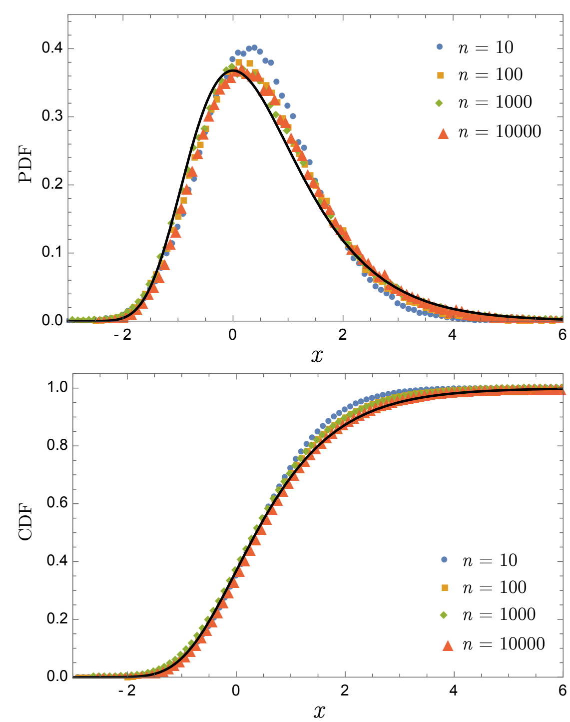

Fig. 2. The p.d.f. $\Psi'_n(x)$ and c.d.f. $\Psi_n(x)$ of the Gumbel distribution (black solid lines) versus numerical simulations of the RV (5) for different values of n. The samples are scaled so that $a_n = 0$ and $b_n = 1$.

III. Weak Correlation Approximation

In this section, we derive an expansion for the p.d.f. $\omega(\mathbf{r})$ of n weakly–correlated Gaussian RVs X_i. The p.d.f. $\omega(\mathbf{r})$ reads

$$\omega(\mathbf{r}) = \frac{1}{(2\pi)^n} \int\limits_{E_n} \chi(\mathbf{k}) e^{-i\mathbf{kr}} d\mathbf{k} \tag{9}$$

with the characteristic function[1]

$$\chi(\mathbf{k}) = \exp\left(i\mu_i k_i - \frac{1}{2}\Sigma_{ij} k_i k_j \right). \tag{10}$$

Here Σ_{ij} is the covariation matrix such that

$$\Sigma_{ij} = \delta_{ij}\sigma_i^2 + \varepsilon_{ij}, \quad \varepsilon_{ii} = 0. \tag{11}$$

The latter allows us to factorise the exponent in (10) as follows

$$\chi(\mathbf{k}) = \chi_0(\mathbf{k}) \cdot \exp\left(-\frac{1}{2}\varepsilon_{ij} k_i k_j \right),$$

where $\chi_0(\mathbf{k}) = \exp\left(i\mu_i k_i - \frac{1}{2}\sigma_i^2 k_i^2 \right)$ is the characteristic function of n uncorrelated normal RVs.

Let us assume that the non-diagonal entries of the covariation matrix are small by absolute value, and, therefore, ε_{ij} can be considered as perturbation:

$$\chi(\mathbf{k})\Big|_{\varepsilon_{ij}\to 0} = \chi_0(\mathbf{k}) + \varepsilon_{ij} \cdot \frac{\partial \chi(\mathbf{k})}{\partial \varepsilon_{ij}}\Big|_{\varepsilon_{ij}=0} + \ldots \tag{12}$$

We can express now the p.d.f. $\omega(\mathbf{r})$ via the non-correlated term ω_0 with the following remarkable property of the multivariate normal distribution:

$$\frac{\partial \chi}{\partial \varepsilon_{ij}}\Big|_{\varepsilon_{ij}=0} = \frac{1}{2} \frac{\partial^2 \chi_0}{\partial \mu_i \partial \mu_j}.$$

Indeed, limiting ourselves to the first order of smallness and substituting (12) into (9), we obtain the expansion for the p.d.f. of normal correlated random variables in the weak correlation case:

$$\omega(\mathbf{r}) \simeq \omega_0(\mathbf{r}) + \frac{\varepsilon_{ij}}{2} \frac{\partial^2 \omega_0(\mathbf{r})}{\partial \mu_i \partial \mu_j}. \tag{13a}$$

In principle, the second derivative can be written explicitly, but this is not necessary for the present discussion. Instead, we will proceed to differentiation by coordinates:

$$\omega(\mathbf{r}) \simeq \omega_0(\mathbf{r}) + \frac{\varepsilon_{ij}}{2} \frac{\partial^2 \omega_0(\mathbf{r})}{\partial x_i \partial x_j}. \tag{13b}$$

It is easy to see that expressions (13a) and (13b) are identical. At the same time, representation (13b) allows one to treat the mean values, μ_i and μ_j, as parameters, which simplifies the analysis below.

In the following Section, we shall consider the c.d.f. $\mathfrak{F}(x)$ for the RV (5), assuming the RVs X_i have the p.d.f. in the form of (13b).

[1] In this work, $\omega(\mathbf{r})$ and $\chi(\mathbf{k})$ correspond to the p.d.f. and characteristic function of *correlated* Gaussian RVs respectively; uncorrelated values are denoted by '0' subscript. We believe, this won't mislead the reader.

IV. Statistics of Weakly Correlated Extreme Values

The probability $\mathbb{P}$ that all of the RVs $X_1, X_2, \ldots, X_n$ are less than some given number z, $\mathbb{P}[\zeta < z]$, is

$$\mathbb{P}[\zeta < z] \equiv \mathfrak{F}(z) = \int\limits_{-\infty}^{z} \ldots \int\limits_{-\infty}^{z} \omega(\mathbf{r}) d\mathbf{r}, \tag{14}$$

where $d\mathbf{r} = dx_1 dx_2 \ldots dx_n$ is the elementary volume in the n-dimensional space of random variables X_i.

We have not imposed any restrictions on the mean values and standard deviations, μ_i and σ_i, while deriving expansion (13b). For the sake of simplicity, we assume that μ_i and σ_i are the same for all X_i. In such a case, RVs can be rescaled to be standard normal ones $X_i \sim \mathcal{N}(0, 1)$. The general case for arbitrary μ_i and σ_i will be reported elsewhere.

Substituting expansion (13b) into (14) and letting $n \to \infty$, we get two terms. The first term leads to the c.d.f. of the Gumbel distribution:

$$\int\limits_{-\infty}^{z} \ldots \int\limits_{-\infty}^{z} \omega_0(\mathbf{r}) d\mathbf{r} = \Phi(z)^n \underset{n\to\infty}{=} \Psi_n(z),$$

while the second one involves the integral of the form

$$I = \int\limits_{-\infty}^{z} \ldots \int\limits_{-\infty}^{z} \frac{\partial^2 \omega_0(\mathbf{r})}{\partial x_i \partial x_j} d\mathbf{r}.$$

The latter gives

$$I = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{z} \frac{\partial}{\partial x_i} e^{-\frac{x_i^2}{2}} dx_i \cdot \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{z} \frac{\partial}{\partial x_j} e^{-\frac{x_j^2}{2}} dx_j$$
$$\times \frac{1}{(2\pi)^{(n-2)/2}} \int\limits_{-\infty}^{z} \ldots \int\limits_{-\infty}^{z} \exp\left(-\frac{1}{2} \sum_{k\neq i,j} x_k^2 \right) \underbrace{dx_1 \ldots dx_n}_{\text{except } dx_i dx_j}$$
$$\underset{n\to\infty}{=} \frac{\varphi(z)^2}{2\pi} \Psi_{n-2}(z).$$

Finally, bringing all terms together and taking into account that $\Psi_{n-2}(x) \approx \Psi_n(x)$, we can write down the asymptotic expression for the c.d.f. of the maximum of n weakly–correlated Gaussian random variables in the first order approximation:

$$\mathfrak{F}(x) = \Psi_n(x) \left[1 + \frac{\varphi(x)^2}{4\pi} \sum_{\substack{i,j \\ i\neq j}} \varepsilon_{ij} + \ldots \right]. \tag{15}$$

Note that this formula is derived assuming the diagonal entries $\varepsilon_{ii} = 0$, and is valid for small correlations:

$$|\varepsilon_{ij}| \ll \delta_{ij}\sigma_i^2. \tag{16}$$

In the next Section, we perform a validation of the results obtained by comparing with Monte Carlo (MC) simulations and discuss possible applications as well as future steps.

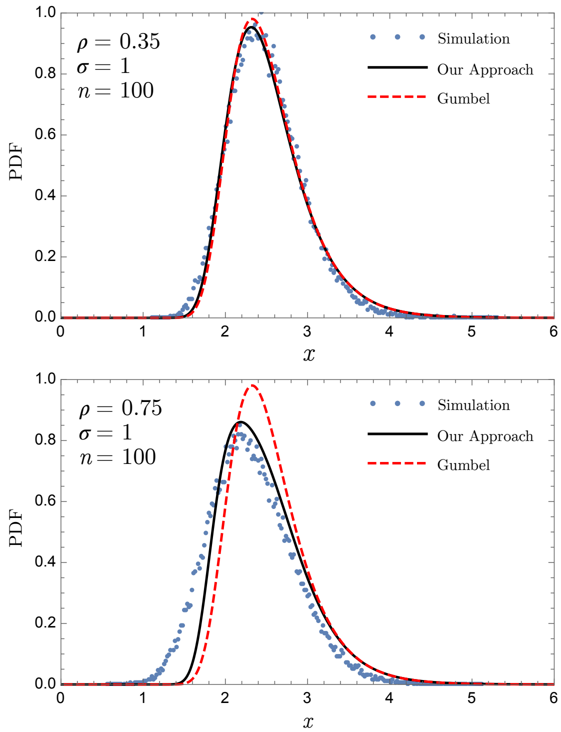

Fig. 3. The p.d.f. of the Gumbel distribution (dashed red) and given by (15) (solid black) together with the results of numerical simulations (blue dots) for different values of the correlation coefficient ρ.

V. Validation and Discussion

The reference method to verify VLSI design is MC simulations [2]. Standard way to generate correlated samples for MC simulations involves the Cholesky decomposition of the covariation matrix $\mathbf{\Sigma}$. This method, being a general approach, is applicable to any $\mathbf{\Sigma}$; however, it is still computationally expensive. For the sake of illustration of the results obtained, we choose the correlation between RVs X_i as follows:

$$\langle X_i X_j \rangle = \sigma^2 \rho^{|i-j|}, \quad \langle Y_i Y_j \rangle = \delta_{ij}, \quad \langle X_i Y_j \rangle = 0, \tag{17}$$

where $X_i \sim \mathcal{N}(0, \sigma)$ and $Y_i \sim \mathcal{N}(0, 1)$. In this case, random samples can be generated via the simple rule

$$X_{i+1} = \rho X_i + \sigma \sqrt{1 - \rho^2} Y_i, \tag{18}$$

where ρ is the correlation coefficient. A set of n samples has been generated according to rule (18) with the starting point X_0 chosen randomly from $\sim \mathcal{N}(0, \sigma)$, then (5) was computed. The procedure was repeated 10^4 times for each numerical experiment.

For a large number of RVs ($n \gtrsim 10^2$), the dependence of the p.d.f. of the RV ζ on the correlation coefficient ρ is studied. As one can see from Fig. 3, for a relatively small values of ρ, the deviation of both the perturbed and uncorrelated Gumbel distributions, (15) and (6) correspondingly, from simulations is negligible (error in mean is less than 2%). This implies that small correlations can be ignored without any loss of accuracy.

At the same time, the comparison in the strong–correlation regime, $\rho > 0.5$, shows that approximation (15), even being not applicable, is able to describe the correct trend.

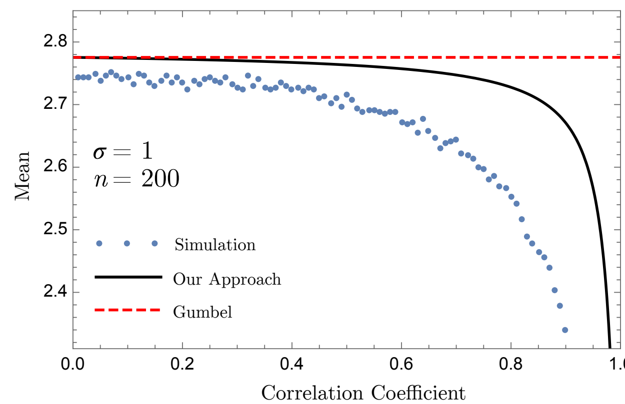

Fig. 4. The mean of RV (5) for $n = 200$ as a function of ρ. The black solid curve is the mean obtained from p.d.f. (15). The red dashed line is due to (8). Blue dots represent the results of the numerical simulation.

From the dependence of the mean of ζ on ρ shown in Fig. 4, one can see that the uncorrelated Gumbel distribution (6) leads to significant deviations from the numerical simulations for $\rho > 0.5$. Thus, (8) can be interpreted as the worst case delay of an IC. At the same time, weak–correlated formula (15) lies between that and the simulation dots. The deviation of (15) from simulations can be reduced by taking into account further terms in the expansion (12).

It should be pointed out that the mean is decreasing with the increase of the correlation coefficient, which is seen both from the simulation and the theory in Fig. 4. Recalling the meaning of the RV ζ in SSTA, we see that the total delay D in an IC is inversely proportional to the value of the correlation coefficient ρ, which can be used in the yield optimisation.

VI. Conclusion

In this paper, the problem of delay determination for VLSI is studied within the path-based approach to SSTA. Based on the CLT, RVs that describe the accumulated delays along paths are considered to have Gaussian distributions, and the asymptotic theory of the maximum of weakly–correlated Gaussian RVs is developed.

It is shown that in the case of large number of variables, the resulting delay distribution tends to the Gumbel law (6). The corresponding modification of the latter is derived by considering the non-diagonal entries of the covariance matrix as perturbations. The particular result (15) is obtained assuming all RVs X_i to be identically distributed. It should be emphasised that within our method, the procedure of delay determination is reduced to the computation of only one expression (15), that makes the methodology extremely fast.

The qualitative agreement between numerical simulations and equation (15) is taken by us as a strong evidence of the validity of our approach. The quantitative analysis of the problem, and the most general case of arbitrary distributed RVs X_i, will be the subject of a separate study. At the same time, the results obtained here are of interest on their own, and can be applied, *e.g.*, in the climate science to the problem of investigation of extreme value time series [17].

References

[1] S. H. Gerez, Ed., *Algorithms for VLSI Design Automation*. Wiley, 1998.

[2] S. Sapatnekar, *Timing*. Springer-Verlag, 2004.

[3] M. Orshansky, S. Nassif, and D. Boning, *Design for Manufacturability and Statistical Design: A Constructive Approach*, ser. Series on Integrated Circuits and Systems. Springer, 2008.

[4] J. Bhasker and R. Chadha, *Static Timing Analysis for Nanometer Designs. A Practical Approach*. Springer, 2009.

[5] L. Lavagno, I. L. Markov, G. Martin, and L. K. Scheffer, Eds., *Electronic Design Automation for IC Implementation, Circuit Design, and Process Technology*. CRC Press, 2016.

[6] M. Dietrich and J. Haase, Eds., *Process Variations and Probabilistic Integrated Circuit Design*. Springer, 2012.

[7] A. B. Kahng, "New game, new goal posts: A recent history of timing closure," in *Proc. DAC*, June 2015, pp. 1–6.

[8] D. Sengupta, V. Mishra, and S. S. Sapatnekar, "Invited: Optimizing device reliability effects at the intersection of physics, circuits, and architecture," in *Proc. DAC*, June 2016, pp. 1–6.

[9] A. Srivastava, D. Sylvester, and D. Blaauw, Eds., *Statistical Analysis and Optimization for VLSI: Timing and Power*. Springer, 2005.

[10] R. Ho, K. W. Mai, and M. A. Horowitz, "The future of wires," *Proceedings of the IEEE*, vol. 89, no. 4, pp. 490–504, Apr 2001.

[11] C. Visweswariah, "Death, taxes and failing chips," in *Proc. DAC*. IEEE, Jun. 2003, pp. 343–347.

[12] I. Nitta, T. Shibuya, and K. Homma, "Statistical static timing analysis technology," *Fujitsu Sci. Tech. J.*, vol. 43, no. 4, pp. 516–523, Oct. 2007.

[13] D. Blaauw, K. Chopra, A. Srivastava, and L. Scheffer, "Statistical timing anlaysis: From basic principles to state of the art," *IEEE Trans. Comput.–Aided Des. Integr. Circuits Syst.*, vol. 4, no. 8, 2008.

[14] J. Freeley, D. Mishagli, T. Brazil, and E. Blokhina, "Statistical simulations of delay propagation in large scale circuits using graph traversal and kernel function decomposition," in *Proc. SMACD*, Jul. 2018.

[15] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed., ser. Wiley Series in Probability and Mathematical Statistics, J. Wiley and Sons, Eds., 1968, vol. 1.

[16] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed., ser. Wiley Series in Probability and Statistics. John Wiley and Sons, 2003.

[17] M. Mudelsee, *Climate Time Series Analysis*, ser. Atmospheric and Oceanographic Sciences Library. Springer, 2010, no. 42.