A Yield Model for Integrated Circuits and Its Application to Statistical Timing Analysis

Farid N. Najm, *Fellow, IEEE*, Noel Menezes, *Member, IEEE*, and Imad A. Ferzli, *Student Member, IEEE*

Abstract—A model for process-induced parameter variations is proposed, combining die-to-die, within-die systematic, and within-die random variations. This model is put to use toward finding suitable timing margins and device file settings, to verify whether a circuit meets a desired timing yield. While this parameter model is cognizant of within-die correlations, it does not require specific variation models, layout information, or prior knowledge of intrachip covariance trends. The approach works with a "generic" critical path, leading to what is referred to as a "process-specific" statistical-timing-analysis technique that depends only on the process technology, transistor parameters, and circuit style. A key feature is that the variation model can be easily built from process data. The derived results are "full-chip," applicable with ease to circuits with millions of components. As such, this provides a way to do a statistical timing analysis without the need for detailed statistical analysis of every path in the design.

Index Terms—Correlations, die-to-die variations, generic critical path, parametric yield, principal component analysis, statistical timing analysis, timing margin, virtual corner, within-die variations.

I. INTRODUCTION

THE YIELD of integrated circuits can be thought of as the ratio of the number of "functional chips" to that of the total chips manufactured [1]. Traditionally, functionality in the context of yield estimation and optimization has been tied to contaminations in the manufacturing process and silicon defects which lead to dramatic circuit faults (e.g., consider the consequences of unwanted short circuits) [2]. This is referred to as functional yield or catastrophic-failure yield. Beyond catastrophic failures, yield is also reduced when a number of manufactured chips does not meet target design specifications. The most important example for digital integrated circuits is that some chips may fail to meet timing requirements in a given process and are sorted into "low-speed bins" [3], resulting in yield loss for the high-speed design. This gives rise to the concept of parametric yield, also known as circuit-limited yield. With technology scaling, circuits are more exposed to the adverse effects of process tolerances, and parametric yield becomes a crucial metric, especially in aggressive and high-speed designs. While the functional yield is primarily affected

by the process of manufacture and typically improves as the process matures [4], the parametric yield can be significantly enhanced at the design stage through design awareness of process variations and their impact on the overall circuit behavior. In this paper, we focus on the parametric part of the overall yield problem, with emphasis on timing margin budgeting, which directly relates to "timing yield" [5], [6].

As part of circuit-timing verification, one has to leave enough margin so that circuit-delay variations do not affect yield too adversely. This has been usually taken care of by using a set of worst case or corner-case files [7], [8] as part of timing verification, typically during static timing analysis (STA). Corner-case files specify the values of transistor parameters for various process corners, including the nominal and different extremes of device behavior. A standard practice for generating corners is to allow random variations of significant physical and electrical parameters across their $\pm 3\sigma$ window (where σ is the standard deviation of the variations) [3], and a circuit is deemed to have passed the timing test if it meets the performance constraints for all these corners.

This conventional corner-case approach is becoming less viable today, the main reason being the growing importance of within-die variations [6] in state-of-the-art technologies. For one thing, within-die variations appear as mismatches between devices or interconnect features on the same chip [5] and exhibit local layout-dependent trends. In consequence, an overwhelming number of corner cases would be needed if one were to adequately capture these variations; and since layout information is available only late in the design, checking for within-die variations becomes a final sign-off stage, impractical during the early phases [9]. In addition, the corner-case approach provides pass/fail outcomes, offering little or no quantitative feedback on the robustness of the design [8].

Statistical techniques offer an alternative. Statistical transistor modeling [8], [10] has been used for quite some time. Design centering [11], [12] methods have also been developed to tune the design to the manufacturing process. However, statistical methods have generally suffered from being too slow to be readily usable in a practical design methodology, aside from the inherent difficulty of gathering statistical data representative of various levels of process variations and extracting from this data information that can be effectively useful for the designer.

While die-to-die variations take the same value for all instances of a given feature on a chip and can be easily modeled, within-die variations have a statistical dependence on the location within a chip and are intrinsically more problematic. They are of two types: systematic and random. Systematic variations arise from the observed wafer-level variation trends

reflected on a given die [9], [13] or the spatial locations of some features on the die and their context in terms of the neighboring layout patterns. They account for the spatial correlations due to physical parameter variations across a chip. Random variations constitute the residual component of the total variations which, in essence, cannot be explained systematically [13] and are modeled by statistically independent Gaussian random variables (RVs) [13], [14]. Ideally, one would like to extract systematic variations and treat them as "deterministic" [15], [16]. This approach, however, necessitates detailed variation models and is hard to apply in the early design stages. A key contribution of this paper is to deal with statistical within-die variations, both systematic and random.

Recently, there has been a growing interest in tackling the timing-yield problem by employing statistical techniques as part of the circuit-timing analysis step. The aim is to extend traditional STA so that it takes into account statistical delay variations [5], [6], [16]–[22] leading to a statistical STA (SSTA). Given the difficulty of early estimation of systematic within-die variations, they have often been ignored; thus, within-die variations were accounted for on a random basis only [6], [13], [18], [21], [22], which is undesirable. In order to avoid making this assumption, one needs to express the within-die correlations with a model that can be easily built from process data. This point is key, and is hard to do—there are no published models, for instance, for how exactly the variations are correlated across the die as a function, say, of the distance between components. A model of correlations in terms of distance is mentioned and used in [23], but no details or data are given; it is not clear what shape the model should take for an arbitrary parameter nor how one would build it from process data. In [16], even though statistical within-die variations are not taken into account, a suggestion is made as to how one may include them and take care of correlation by enforcing correlation between neighboring features on a die. This theme was further developed recently where use was made of principal component analysis (PCA) [20] or a quad-tree partitioning [19] to express a regionwise spatial within-die correlation. Here, too, it is not clear how one would identify these regions and how the model would be built from process data. Finally, these methods depend on placement information, making them hard to use during circuit design and optimization.

The basic question that serves as premise to this paper is: Prior to design, what can be said regarding circuit characteristics? From the process/technology end, we assume that one is able to know device/interconnect-delay sensitivities and the extent of die-to-die and within-die correlations, without knowing how these correlations will affect different elements of the design once placement is done. Our analysis will account for this knowledge and will thus introduce a level of awareness of process variations, including their correlation, at this early stage of the design cycle. From the design side, we assume that one is able to make certain design decisions pertaining to "design style," such as the typical number of critical paths or their depth in a circuit. Then, given a target process technology, transistor models, and circuit design styles, the methodology presented in this paper can be applied to suggest preferred design styles (e.g., in terms of the number of high-speed

paths and their depths) or find preferred device characteristics (e.g., in terms of device delay sensitivities) which would forecast higher timing yields once the design is completed, considering that process variations are correlated but their correlations still unknown at the time of the analysis. The quantitative measure will come in the form of yield bounds and yield-margin curves that hold regardless of how the correlation structure of the process-induced variations in each physical parameter will look like once the placement is complete. Yield-margin curves can be used to quantify yield loss at a given timing margin. Our analysis will also result in a selection of "device file" settings lying within the extremes of device behavior with which to run deterministic STA, so that a given timing-yield requirement can be met. Our work, of which a preliminary version appeared in [24], straddles traditional STA and design space exploration techniques under process variations and links circuit-timing analysis as done today through the STA with process-induced parametric yield.

For example, while the "nominal" device file may call for a setting of $\Delta L = 0$ (for channel-length variations) and the "worst case" file may call for a setting of $\Delta L = +3\sigma_L$, our approach can be used to predict the value "δ," which we refer to as a virtual corner, such that if the setting of $\Delta L = \delta\sigma_L$ is used for all devices, and if the circuit timing is verified using deterministic STA, then the circuit will achieve the desired timing yield. Therefore, our approach preserves existing static timing methodology and only assumes the existence of statistical transistor models, which have been standard for some time.

We will perform our analysis using the concept of "generic critical path," in the style of [23], by examining the statistical properties of large ensembles of such paths. The validity of the generic-path model in the context of circuit-timing analysis was verified by a separate set of experiments described in the appendix of this paper. The generic paths will enable us to abstract most circuit-specific information and obtain estimates of yield loss considering a model of process variations including die-to-die and within-die with correlations. The model of generic paths is pertinent to circuit-timing analysis insofar as different circuits share fundamental characteristics which impact timing yield, and we think it is intuitively reasonable to use the generic-path model to capture the way logic paths contribute to timing-yield loss. It is well known, for instance, that the timing yield of highly optimized circuits is determined by a relatively large number of paths of comparable delay. This phenomenon of the "wall of paths" clustered close to the clock edge as a result of circuit optimization has been widely reported in the literature [25] and forms an intuitive basis for the generic-path model. Under this model, our analysis can serve as a theoretical framework for understanding the timing-yield implications of process variability.

The rest of this paper is organized as follows. Section II introduces our generic parameter model, combining die-to-die and systematic and random within-die (WDR) variations. In Section III, we build a model for circuit timing starting from the physical device and interconnect variations and show how circuit timing fits with our generic parameter model. Then, after a brief summary in Section IV, we gradually look at the impact of individual variation components on parametric yield. Section V

examines the situation where parameter variation is considered entirely die to die. Section VI adds within-die variations as purely random, and the full variation model is addressed in Section VII, with an application to circuit timing illustrated in Section VIII. When systematic within-die variations are excluded, our approach is to compute simple expressions to estimate parametric yield. When these variations are included, we derive readily usable expressions to bound the yield from above and below. Throughout this paper, we will derive the results first on a generic parameter, then illustrate them with a direct application to circuit timing. We will compare the necessary timing margins and case file settings in different situations by following a simple case where variation is progressively split between different components, and a specified timing yield is required. We conclude in Section IX.

II. PARAMETER MODEL

For a given circuit element or layout feature i, let its co-ordinates on the die be (x_i, y_i) and let $X(i)$ be a Gaussian RV that denotes the variation of a certain parameter of this element from its nominal (mean) value. Thus, for example, $X(i)$ may represent channel-length variations of transistor i. Let n be the number of occurrences of this parameter on a chip. For simplicity, we consider $X(i)$ to be zero mean, with the understanding that a mean shift can be easily incorporated into the model below.

We break down $X(i)$ into a die-to-die component X_{dd} and a within-die component X_{wd}, such that

$$X(i) = X_{\mathrm{dd}} + X_{\mathrm{wd}}(i), \qquad i = 1, \ldots, n. \tag{1}$$

Here, X_{dd} is a zero-mean Gaussian RV that takes the same value for all instances of an element on a given chip, irrespective of location, while $X_{\mathrm{wd}}(i)$ is a zero-mean Gaussian RV which can take different values for different instances of an element on the same die. We consider X_{dd} to be independent of $X_{\mathrm{wd}}(i)$, $\forall i$. This leads to the following relationship between the variances:

$$\sigma^2(i) = \sigma_{\mathrm{dd}}^2 + \sigma_{\mathrm{wd}}^2(i) \tag{2}$$

where $\sigma^2(i)$, σ_{dd}^2, and $\sigma_{\mathrm{wd}}^2(i)$ are the respective variances of $X(i)$, X_{dd}, and $X_{\mathrm{wd}}(i)$.

Within-die variation is further broken down into systematic and random components, respectively, as $X_{\mathrm{wds}}(i) = X_{\mathrm{wds}}(x_i, y_i)$ and $X_{\mathrm{wdr}}(i)$

$$X_{\mathrm{wd}}(i) = X_{\mathrm{wds}}(x_i, y_i) + X_{\mathrm{wdr}}(i), \qquad i = 1, \ldots, n \tag{3}$$

where each of these components is itself a zero-mean Gaussian RV. Systematic variations capture all location-dependent within-die correlations, so that $X_{\mathrm{wds}}(x_i, y_i)$ and $X_{\mathrm{wdr}}(j)$ are independent, $\forall i, j$, and $X_{\mathrm{wdr}}(i)$ and $X_{\mathrm{wdr}}(j)$ are independent, $\forall i \neq j$. The following relationship can then be written for within-die variance:

$$\sigma_{\mathrm{wd}}^2(i) = \sigma_{\mathrm{wds}}^2(x_i, y_i) + \sigma_{\mathrm{wdr}}^2(i) \tag{4}$$

where $\sigma_{\mathrm{wd}}^2(i)$, $\sigma_{\mathrm{wds}}^2(x_i, y_i)$, and $\sigma_{\mathrm{wdr}}^2(i)$ are the variances of $X_{\mathrm{wd}}(i)$, $X_{\mathrm{wds}}(x_i, y_i)$, and $X_{\mathrm{wdr}}(i)$, respectively.

The problem clearly lies in knowing and managing the joint distribution of the systematic variations. The covariance matrix itself, which may contain an information pertaining to millions of instances of a parameter on a given chip, is hard to estimate; and even when it is available, numerical operations can become extremely difficult. One way to express the systematic variations is to use the PCA [26] and write

$$X_{\mathrm{wds}}(i) = \sum_{j=1}^{p} a_{ij} Z_j \tag{5}$$

where Z_j are independent standard normal RVs (Gaussians with zero mean and unity variance), and $p \ll n$ is the order of the expansion. The RVs Z_j correspond to underlying independent unobservable factors. The value of p and the coefficients a_{ij} represent the extent of correlation across the die. For example, if $p = 1$, then the within-die spatial correlation coefficient is one: There is a perfect correlation; a single underlying RV Z_1 determines the value of the systematic component of X_{wd} all over the die. A $p > 1$ allows for less than perfect correlation.

We will adopt the PCA expansion (5) as our "correlation model" for the within-die component. At first glance, this model appears hard to use because it seems to depend on knowledge of the values of all the a_{ij} parameters. These coefficients depend on the correlation structure of physical parameters on the chip. In other words, a full PCA cannot be used to assess the impact of statistical variations on the design until after layout when the spatial locations of devices on chip are known. Even then, it is not clear how one would compute them from process data. A brute-force PCA over the millions of instances of a parameter on a chip would be impractical. However, without any explicit knowledge of the correlations, we will capitalize on the following property that relates the PCA coefficients to the parameter variances:

$$\sigma_{\mathrm{wds}}^2(x_i, y_i) = \sum_{j=1}^{p} a_{ij}^2. \tag{6}$$

Using this property, and without knowing the specifics of systematic variations, including die-level trends and layout information, we will obtain lower/upper bounds on the parametric yield that require only: 1) the order p of the PCA expansion, and 2) representative values (maximum, minimum, average) of the variance terms given previously.

III. TIMING MODEL

This section shows how a model for circuit timing can be built to conform with the generic model discussed in Section II and whose characteristics are obtainable from process files. Working toward the delay of a multistage generic critical path, we start with the physical variations of transistor threshold voltage and channel length. Physical variations lead to stage-delay variations, where a stage consists of a gate with fan-out interconnect, and we discuss interconnect-level variations. Finally, stage-delay variations are used to characterize path-delay variations as a generic parameter.

A. Device Variations

In deep submicrometer CMOS, process-induced delay variations are mainly due to variations in the MOSFET threshold voltage (V_t) and effective channel length (L_e). Due to short-channel effects, V_t may decrease with decreasing L_e (the V_t roll-off [3] effect), so that V_t and L_e are not independent variables. We assume that for small L_e variations, denoted by L, threshold voltage roll-off is linear with channel-length variation, and we express V_t as the sum of a linear term that depends on L_e ($K\Delta L_e$, where $K > 0$) and another term representing variations of V_t which are independent of L_e, denoted by V. With $E[\cdot]$ being the expected (or mean) value operator, we can express V_t and L_e of transistor i as follows:

$$L_e(i) = E[L_e(i)] + L(i) \tag{7}$$

and

$$V_t(i) = E[V_t(i)] + KL(i) + V(i). \tag{8}$$

We are interested in the RVs $L(i)$ and $V(i)$, which are assumed to be zero-mean Gaussian and independent of each other. We can thus break them up in the usual manner as

$$L(i) = L_{\text{dd}} + L_{\text{wds}}(x_i, y_i) + L_{\text{wdr}}(i)$$
$$V(i) = V_{\text{dd}} + V_{\text{wds}}(x_i, y_i) + V_{\text{wdr}}(i). \tag{9}$$

(Of course, here, V_{dd} is the die-to-die component of $V(i)$, and not the supply voltage.) The variances of $L(i)$ and $V(i)$, as well as the variance of their components, can be known from the technology files

$$\sigma_L^2(i) = \sigma_{\text{dd},L}^2 + \sigma_{\text{wds},L}^2(x_i, y_i) + \sigma_{\text{wdr},L}^2(i) \tag{10}$$
$$\sigma_V^2(i) = \sigma_{\text{dd},V}^2 + \sigma_{\text{wds},V}^2(x_i, y_i) + \sigma_{\text{wdr},V}^2(i). \tag{11}$$

B. Gate Delay

We assume that, for a given chip design in a given technology, one can define a "nominal" representative logic gate, with appropriate output loading and input slope. For reasons that will become clear, this gate should be typical of gates on critical paths in this technology. Due to the nonlinearity of the relationship between gate delay and transistor parameters, the mean value of gate delay does not necessarily coincide with its nominal value [the value corresponding to the case when $L(i) = 0$ and $V(i) = 0$]. Furthermore, the distribution of gate delay would not necessarily be Gaussian. Simple experiments with HSPICE, however, reveal that this nonlinearity is not strong, at least not in 0.13-μm CMOS. Therefore, we will ignore these complications and simply assume that gate delay is linearly dependent on transistor parameter variations, so that it becomes a Gaussian variable with mean equal to its nominal value. Another consequence of this linearity assumption is that, if $D(i)$ is the deviation of the delay of logic gate i from its mean (nominal) delay, then

$$D(i) = \alpha_0 L(i) + \beta(V(i) + KL(i)) \tag{12}$$

where α_0 and β are sensitivity parameters, with suitable units, that one can easily obtain from circuit simulation of a representative logic gate. Notice that, in general, $\alpha_0 > 0$ and $\beta > 0$. (For a specific industrial 0.13-μm process, we have found that for a minimum-sized inverter, $\alpha_0 \approx 0.857$ ps/nm and $\beta \approx 17.3$ ps/V.) Let $s_1 = \alpha_0 + K\beta$ and let $s_2 = \beta$. Then, we have

$$D(i) = s_1 L(i) + s_2 V(i). \tag{13}$$

Now, we define the following:

$$D_{\text{dd}} = s_1 L_{\text{dd}} + s_2 V_{\text{dd}}$$
$$D_{\text{wds}}(x_i, y_i) = s_1 L_{\text{wds}}(x_i, y_i) + s_2 V_{\text{wds}}(x_i, y_i)$$
$$D_{\text{wdr}}(i) = s_1 L_{\text{wdr}}(i) + s_2 V_{\text{wdr}}(i). \tag{14}$$

Clearly, D_{dd}, $D_{\text{wds}}(x_i, y_i)$, and $D_{\text{wdr}}(i)$ represent die-to-die, WDS, and WDR variations of gate delay, corresponding to the parametric model discussed in Section II. As a result, we can define the three components of the total gate-delay variance $\sigma_{\text{dd},D}^2$, $\sigma_{\text{wds},D}^2(x_i, y_i)$, and $\sigma_{\text{wdr},D}^2(i)$ and compute them from the variations of channel length and threshold voltage using (10), (11), and (14).

C. Interconnect Delay

Gate delay is one factor in total path delay, the other being interconnect. We extend the previous analysis to include interconnect-delay variations by considering each stage on a path to comprise a gate with corresponding fan-out interconnect. Let $\tau(i)$ be the delay of the interconnect line in stage i. Similar to a generic gate, we assume that each stage can be characterized by a nominal interconnect delay $\tau_0(i)$ that is the same for all stages, i.e., $\tau_0(i) = \tau_0$, $\forall i$. In a simple model, $\tau(i) = R(i)C(i)$, where $R(i)$ and $C(i)$ are, respectively, the resistance and capacitance of line i [27]. If ρ is the metal resistivity and ϵ_{ox} is the oxide permittivity, then

$$R(i) = \frac{\rho l(i)}{w(i)t(i)} \quad C(i) = \frac{\epsilon_{\text{ox}} w(i)l(i)}{T_{\text{ILD}}(i)}$$

where $l(i)$ and $w(i)$ are, respectively, the metal length and width, and $t(i)$ and $T_{\text{ILD}}(i)$ are, respectively, the thicknesses of the interconnect and the interlayer dielectric (ILD) oxide, in stage i. Then, we have

$$\tau(i) = \frac{\rho \epsilon_{\text{ox}} l(i)^2}{t(i)T_{\text{ILD}}(i)}. \tag{15}$$

Denote by t_0 and $T_{\text{ILD},0}$ the nominal values of metal and ILD thicknesses and let $t(i) = t_0 + \Delta t(i)$ and $T_{\text{ILD}}(i) = T_{\text{ILD},0} + \Delta T_{\text{ILD}}(i)$. Neglecting the second-order term $\Delta t(i)\Delta T_{\text{ILD}}(i)$, we write

$$\tau(i) \approx \frac{\rho \epsilon_{\text{ox}} l(i)^2}{t_0 T_{\text{ILD},0} + T_{\text{ILD},0}\Delta t(i) + t_0 \Delta T_{\text{ILD}}(i)}$$
$$= \frac{\tau_0}{1 + \frac{\Delta t(i)}{t_0} + \frac{\Delta T_{\text{ILD}}(i)}{T_{\text{ILD},0}}}. \tag{16}$$

Let $\Delta\tau(i) = \tau(i) - \tau_0$, then

$$\frac{\Delta\tau(i)}{\tau_0} = \frac{-\left(\frac{\Delta t(i)}{t_0} + \frac{\Delta T_{\mathrm{ILD}}(i)}{T_{\mathrm{ILD},0}}\right)}{1 + \left(\frac{\Delta t(i)}{t_0} + \frac{\Delta T_{\mathrm{ILD}}(i)}{T_{\mathrm{ILD},0}}\right)}. \tag{17}$$

Let $u = \Delta t(i)/t_0 + \Delta T_{\mathrm{ILD}}(i)/T_{\mathrm{ILD},0}$. Assuming small (relative) variations in ILD and metal thicknesses, $|u| \ll 1$, which leads to

$$\Delta\tau(i) \approx -\frac{\tau_0}{t_0}\Delta t(i) - \frac{\tau_0}{T_{\mathrm{ILD},0}}\Delta T_{\mathrm{ILD}}(i)$$
$$= -s_3\Delta t(i) - s_4\Delta T_{\mathrm{ILD}}(i). \tag{18}$$

Supposing zero-mean Gaussian deviations in ILD and metal thickness, we consider each to be a generic parameter as in Section II, with corresponding variation components $(T_{\mathrm{ILD},\mathrm{dd}}, T_{\mathrm{ILD},\mathrm{wds}}, T_{\mathrm{ILD},\mathrm{wdr}})$ and $(t_{\mathrm{dd}}, t_{\mathrm{wds}}, t_{\mathrm{wdr}})$ and their respective variance triplets $(\sigma^2_{T_{\mathrm{ILD}},\mathrm{dd}}, \sigma^2_{T_{\mathrm{ILD}},\mathrm{wds}}, \sigma^2_{T_{\mathrm{ILD}},\mathrm{wdr}})$ and $(\sigma^2_{t,\mathrm{dd}}, \sigma^2_{t,\mathrm{wds}}, \sigma^2_{t,\mathrm{wdr}})$.

While advanced interconnect variation models have been developed [28]–[30] and may depend on several physical parameters, we assume that in any case, it is possible to express the interconnect-delay variation as a linear summation of physical parameters P_i, in the form

$$\Delta\tau(i) \approx \sum \pm s_i P_i.$$

For simplicity, the analysis to follow includes only variations in ILD and metal thicknesses, as given in (18).

D. Stage Delay

Considering each stage to represent a gate with corresponding fan-out interconnect, we can now express stage-delay deviation $S(i)$ as follows:

$$S(i) = D(i) + \Delta\tau(i)$$
$$= [s_1 L_{\mathrm{dd}} + s_2 V_{\mathrm{dd}} - s_3 t_{\mathrm{dd}} - s_4 T_{\mathrm{ILD},\mathrm{dd}}]$$
$$+ [s_1 L_{\mathrm{wds}}(x_j, y_j) + s_2 V_{\mathrm{wds}}(x_j, y_j)$$
$$- s_3 t_{\mathrm{wds}}(x_j, y_j) - s_4 T_{\mathrm{ILD},\mathrm{wds}}(x_j, y_j)]$$
$$+ [s_1 L_{\mathrm{wdr}}(i) + s_2 V_{\mathrm{wdr}}(i) - s_3 t_{\mathrm{wdr}}(i) - s_4 T_{\mathrm{ILD},\mathrm{wdr}}(i)]. \tag{19}$$

The die-to-die, WDS, and WDR components of the variations in stage delay are readily identifiable from (19). Stage delay therefore conforms with the model for a generic parameter in Section II, in the sense that it has three variance components which can be computed from physical parameter variances, with the die-to-die component being the same for all stages and the within-die random components of two distinct stages being statistically independent.

E. Path Delay

Consider a path with a number N of logic stages, representative of a circuit's typical critical paths, which we will refer to as a generic critical path. Assume that path j comprises stages $1, \ldots, N$ and let $D_N(j)$ denote the deviation of the delay of path j from its mean (nominal) value. Then

$$D_N(j) = \sum_{i=1}^{N} S(i)$$
$$= NS_{\mathrm{dd}} + \sum_{i=1}^{N} S_{\mathrm{wds}}(x_i, y_i) + \sum_{i=1}^{N} S_{\mathrm{wdr}}(i) \tag{20}$$

where S_{dd}, $S_{\mathrm{wds}}(x_i, y_i)$, and $S_{\mathrm{wdr}}(i)$ are the die-to-die, WDS, and WDR variations in stage delay, derived in (19).

The gates on a path exist at various different locations. We make the simplifying assumption that as far as physical location on the die, for purposes of computing the within-die-systematic component, all gates (and stages) on path j share the same "nominal" coordinates (x_j, y_j), so that

$$D_N(j) = NS_{\mathrm{dd}} + NS_{\mathrm{wds}}(x_j, y_j) + \sum_{i=1}^{N} S_{\mathrm{wdr}}(i). \tag{21}$$

This approximation is motivated by the expectation that gates on a critical path should be nearby on the die, and differences between their position-dependent within-die-systematic variations should be minor. We note that this may no longer be true in latch-based designs that make use of a cycle stealing, and where critical paths may span large distances.

Using (19), we arrive at the following relationship between path-delay and physical parameter variations:

$$D_N(j) = [Ns_1 L_{\mathrm{dd}} + Ns_2 V_{\mathrm{dd}} - Ns_3 t_{\mathrm{dd}} - Ns_4 T_{\mathrm{ILD},\mathrm{dd}}]$$
$$+ [Ns_1 L_{\mathrm{wds}}(x_j, y_j) + Ns_2 V_{\mathrm{wds}}$$
$$- Ns_3 t_{\mathrm{wds}}(x_j, y_j) - Ns_4 T_{\mathrm{ILD},\mathrm{wds}}(x_j, y_j)]$$
$$+ \left[s_1 \sum_{i=1}^{N} L_{\mathrm{wdr}}(i) + s_2 \sum_{i=1}^{N} V_{\mathrm{wdr}}(i)\right.$$
$$\left. - s_3 \sum_{i=1}^{N} t_{\mathrm{wdr}}(i) - s_4 \sum_{i=1}^{N} T_{\mathrm{ILD},\mathrm{wdr}}(i)\right]. \tag{22}$$

We can readily identify the die-to-die, WDS, and WDR components of path delay from the previous expression. It can be seen that the systematic component is the same for all the considered generic critical paths. Assume that the considered paths $j = 1, 2, \ldots$ are disjoint, then the random components of path delay become strictly independent from one another and the structure of the generic parameter introduced in Section II is preserved. In the sequel, we will consider that path delay conforms with our generic parameter model, with three variance components related to physical parameters and obtained from process data, and such that its die-to-die component is the same for all instances (on the same die), its WDR components are independent for different instances, and its WDS components are

spatially correlated due to WDS physical parameter variations. For practical purposes, the assumption of disjoint paths is not restrictive and is needed to ensure that random components of path delay remain strictly uncorrelated. Recall that we are looking at generic paths which are representative of all the chip's critical paths, and the number of such disjoint paths is effectively large enough so that any statistics inferred on such a disjoint group reflect those of the critical paths collectively.

Based on the independence relations between the terms in (22), we now have expressions for the variance triplet of path delay, so that we can write

$$\sigma_{D_N}^2(j) = \sigma_{\mathrm{dd},D_N}^2 + \sigma_{\mathrm{wds},D_N}^2(x_j, y_j) + \sigma_{\mathrm{wdr},D_N}^2(j) \quad (23)$$

where

$$\sigma_{\mathrm{dd},D_N}^2$$
$$= N^2 s_1^2 \sigma_{\mathrm{dd},L}^2 + N^2 s_2^2 \sigma_{\mathrm{dd},V}^2 + N^2 s_3^2 \sigma_{\mathrm{dd},t}^2 + N^2 s_4^2 \sigma_{\mathrm{dd},T_{\mathrm{ILD}}}^2 \quad (24)$$

$$\sigma_{\mathrm{wds},D_N}^2(x_j, y_j)$$
$$= N^2 s_1^2 \sigma_{\mathrm{wds},L}^2(x_j, y_j) + N^2 s_2^2 \sigma_{\mathrm{wds},V}^2(x_j, y_j)$$
$$+ N^2 s_3^2 \sigma_{\mathrm{wds},t}^2(x_j, y_j) + N^2 s_4^2 \sigma_{\mathrm{wds},T_{\mathrm{ILD}}}^2(x_j, y_j) \quad (25)$$

$$\sigma_{\mathrm{wdr},D_N}^2(j)$$
$$= N s_1^2 \hat{\sigma}_{\mathrm{wdr},L}^2(j) + N s_2^2 \hat{\sigma}_{\mathrm{wdr},V}^2(j)$$
$$+ N s_3^2 \hat{\sigma}_{\mathrm{wdr},t}^2(j) + N s_4^2 \hat{\sigma}_{\mathrm{wdr},T_{\mathrm{ILD}}}^2(j) \quad (26)$$

where $\hat{\sigma}_{\mathrm{wdr},L}^2(j)$, $\hat{\sigma}_{\mathrm{wdr},V}^2(j)$, $\hat{\sigma}_{\mathrm{wdr},t}^2(j)$, and $\hat{\sigma}_{\mathrm{wdr},T_{\mathrm{ILD}}}^2(j)$ are the average variances of the within-die random variations of the corresponding physical parameters taken over path j. These path averages may be approximated by averages over the whole die, so that we write

$$\hat{\sigma}_{\mathrm{wdr},V}^2(j) \approx \hat{\sigma}_{\mathrm{wdr},V}^2$$
$$\hat{\sigma}_{\mathrm{wdr},L}^2(j) \approx \hat{\sigma}_{\mathrm{wdr},L}^2$$
$$\hat{\sigma}_{\mathrm{wdr},t}^2(j) \approx \hat{\sigma}_{\mathrm{wdr},D}^2$$
$$\hat{\sigma}_{\mathrm{wdr},T_{\mathrm{ILD}}}^2(j) \approx \hat{\sigma}_{\mathrm{wdr},T_{\mathrm{ILD}}}^2.$$

In turn, we can approximate the variance of delay on path j due to WDR parametric variations as a die-level average variance, in other words

$$\sigma_{\mathrm{wdr},D_N}^2(j) \approx \hat{\sigma}_{\mathrm{wdr},D_N}^2. \quad (27)$$

As usual, we define $\hat{\sigma}_{\mathrm{wdr},D_N}^2$ to be the average value of $\sigma_{\mathrm{wdr},D_N}^2(j)$ across the die. An important observation from (24)–(26) is that, unlike the two other components, WDR variance grows with N, not N^2. This gives a preliminary insight on the effect of random variations on path delay "averaging out," or becoming significantly less manifest than other variation components over paths with a large number of stages. We will return to this point in the subsequent analysis.

IV. PARAMETRIC AND TIMING YIELD

A key component of our approach is the application of the generic parameter yield model developed previously to path delay, and using that to compute chip timing yield. The reason we can do this is because, as shown, we can express the variance of path delay, for a large collection of critical generic paths, as a standard variance model with die-to-die, within-die systematic (WDS), and within-die random (WDR) components. With this, we can effectively talk about delay (be it of a gate, a stage, or a path) as being a "parameter."

Throughout this paper, we define the yield of a certain parameter X as the probability of satisfying a constraint on the maximum value of the parameter on a chip (with the understanding that the results to follow can be extended to cover minimum values and interval constraints)

$$Y(x) = \mathcal{P}\{X(i) \le x, \qquad i = 1, 2, \ldots, n\} \quad (28)$$

where n is the number of instances of that parameter on the die. We write $Y(x)$ to refer to the yield of a generic parameter and denote specifically by $\mathcal{Y}(d)$ a timing yield. In the following sections, we will build our way gradually to a full model of chip timing yield by first investigating the timing yield when only die-to-die variations are considered, then when only die-to-die and WDR variations are considered, before finally presenting the full model. This improves the clarity of the results, and also allows us to highlight interesting results that arise when only certain variation components are dominant.

V. DIE-TO-DIE YIELD

We start by considering only die-to-die variations.

A. Generic Parameter

We define the die-to-die yield of a generic parameter X as

$$Y_{\mathrm{dd}}(x) = \mathcal{P}\{X_{\mathrm{dd}} \le x\} = \Phi\left(\frac{x}{\sigma_{\mathrm{dd}}}\right) \quad (29)$$

where $\mathcal{P}\{\cdot\}$ is a probability, $\Phi(\cdot)$ is the cumulative distribution function (cdf) of a standard normal RV, and σ_{dd} is the variance of the die-to-die component of X. Since $X_{\mathrm{dd}}(i) = X_{\mathrm{dd}}, \forall i$, the die-to-die yield is governed by a single RV.

B. Timing Yield

The simple expression in (29) can be directly used to estimate the die-to-die timing yield $\mathcal{Y}_{\mathrm{dd}}(d)$, replacing σ_{dd} by σ_{dd,D_N}, obtained from (24). Fig. 1 plots the values of die-to-die timing yield versus the timing margin d, where d is normalized in numbers of standard deviations $\sigma = \sigma_{\mathrm{dd},D_N}$.

Now, suppose that a value $\mathcal{Y}$ is desired for the die-to-die timing yield. Then, we have

$$d = \mathcal{Y}_{\mathrm{dd}}^{-1}(\mathcal{Y})$$
$$= \Phi^{-1}(\mathcal{Y})\sigma_{\mathrm{dd},D_N}. \quad (30)$$

We can interpret (30) in two ways. Suppose that d_0 is the critical path delay when the nominal process files are used,
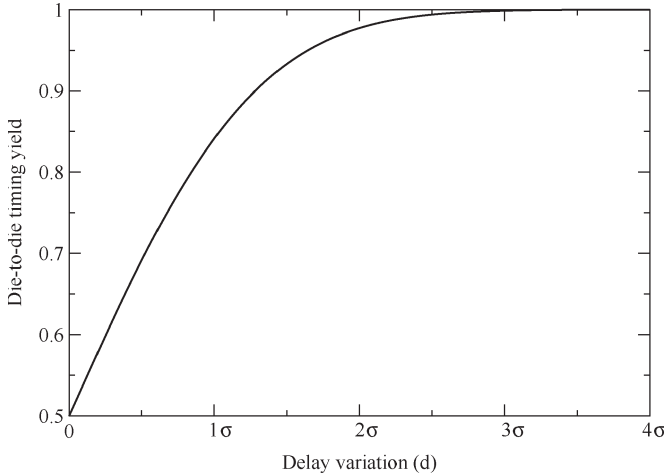
Fig. 1.   Timing yield when considering die-to-die variations alone.

i.e., excluding process variations. If the nominal timing margin (e.g., $T_{\text{clk}} - d_0$, where T_{clk} is the clock period) exceeds $\mathcal{Y}_{\text{dd}}^{-1}(\mathcal{Y})$, then the timing yield $\mathcal{Y}$ is met.

Alternatively, if a corner-case process file is set to produce a delay of $d_0 + \mathcal{Y}_{\text{dd}}^{-1}(\mathcal{Y})$ while still meeting the timing constraints, then the die-to-die yield $\mathcal{Y}$ is achieved. Therefore, we refer to plots such as Fig. 1 as yield-margin curves. For example, such a corner file can be built simply by allowing a delay deviation of $\mathcal{Y}_{\text{dd}}^{-1}(\mathcal{Y})/N$ at each of the N stages of the critical path. In turn, stage physical parameters can be modified so that stage delay deviates by this amount. $S(i)$ being the delay of stage i, we let

$$
\begin{aligned}
\frac{\mathcal{Y}_{\text{dd}}^{-1}(\mathcal{Y})}{N} &= S(i), \quad i = 1, \dots, n \\
&= s_1 L(i) + s_2 V(i) - s_3 \Delta t(i) - s_4 \Delta T_{\text{ILD}}(i) \\
&= \Phi^{-1}(\mathcal{Y}) \frac{\sigma_{\text{dd},D_N}}{N}
\end{aligned}
\tag{31}
$$

where the last equality follows from (30). This gives a range of possible settings of the physical parameters that achieve a timing deviation of $\mathcal{Y}_{\text{dd}}^{-1}(\mathcal{Y})$ over N-stage paths. For simplicity of illustration, if we impose on our corner-case files equal relative deflection (i.e., deviation from the mean or nominal value) of the physical parameters, i.e.,

$$
\frac{L(i)}{\sigma_{\text{dd},L}} = \frac{V(i)}{\sigma_{\text{dd},V}} = -\frac{\Delta t(i)}{\sigma_{\text{dd},t}} = -\frac{\Delta T_{\text{ILD}}(i)}{\sigma_{\text{dd},T_{\text{ILD}}}} = \delta, \quad \forall i \tag{32}
$$

then

$$
\delta = \frac{\mathcal{Y}_{\text{dd}}^{-1}(\mathcal{Y})/N}{s_1 \sigma_{\text{dd},L} + s_2 \sigma_{\text{dd},V} + s_3 \sigma_{\text{dd},t} + s_4 \sigma_{\text{dd},T_{\text{ILD}}}} \tag{33}
$$

is the value of the process file setting which, when applied to all the parameters, causes the right amount of path-delay deflection $\mathcal{Y}_{\text{dd}}^{-1}(\mathcal{Y})$ to be created so as to "test" the circuit for timing yield $\mathcal{Y}$.

It is of interest to note the following. Recall from (24) that the die-to-die path-delay variance is as follows:

$$
\sigma_{\text{dd},D_N}^2 = N^2 \left(s_1^2 \sigma_{\text{dd},L}^2 + s_2^2 \sigma_{\text{dd},V}^2 + s_3^2 \sigma_{\text{dd},t}^2 + s_4^2 \sigma_{\text{dd},T_{\text{ILD}}}^2 \right). \tag{34}
$$

Since the square of the sum of positive quantities is larger than the sum of their squares, then

$$
\sigma_{\text{dd},D_N} \leq N(s_1 \sigma_{\text{dd},L} + s_2 \sigma_{\text{dd},V} + s_3 \sigma_{\text{dd},t} + s_4 \sigma_{\text{dd},T_{\text{ILD}}}). \tag{35}
$$

From (30) and (33), it then follows that

$$
\delta \leq \Phi^{-1}(\mathcal{Y}). \tag{36}
$$

Thus, when equal deflections of the physical parameters are assumed, then δ, for a desired die-to-die yield $\mathcal{Y}$, is at most $\Phi^{-1}(\mathcal{Y})$, irrespective of the specific variance values.

In order to get a sense of typical values for δ, assume for simplicity that path-delay variance is equally divided among channel length, threshold voltage, dielectric thickness, and wire thickness, that is

$$
s_1^2 \sigma_{\text{dd},L}^2 = s_2^2 \sigma_{\text{dd},V}^2 = s_3^2 \sigma_{\text{dd},t}^2 = s_4^2 \sigma_{\text{dd},T_{\text{ILD}}}^2 = s^2
$$

and suppose a die-to-die yield of 95% is desired. From Fig. 1, we get that $\mathcal{Y}_{\text{dd}}^{-1}(95\%) = 1.6\sigma_{\text{dd},D_N}$. Then, each stage delay may be set to $\mathcal{Y}_{\text{dd}}^{-1}(95\%)/N = 1.6\sigma_{\text{dd},D_N}/N = 3.2s$. Imposing equal relative deflection in interconnect and device parameters, therefore using (33), yields

$$
\delta = \frac{3.2s}{4s} = 0.8. \tag{37}
$$

To sum up, δ defines a "corner-case" file for which the circuit should be tested for timing-constraint violations. If the circuit satisfies the timing constraints with this δ setting, then it would have a die-to-die timing yield of $\mathcal{Y}$. Observe that this file setting is not unique in the sense that any corner-case file which produces a delay deviation of $\mathcal{Y}_{\text{dd}}^{-1}(\mathcal{Y})$ around the nominal delay effectively serves the same purpose. Our particular derivation of the case file simply imposed equal stage timings and similar relative deviation in transistor physical parameters. We refer to these corners (defined by a δ value) that lie between device extremes and that could be used to design for a certain target yield as virtual corners.

VI. YIELD CONSIDERING DIE-TO-DIE AND WDR VARIATIONS

In this section, we consider only the die-to-die and WDR variations.

A. Generic Parameter

We start by considering a generic parameter X. Let $Y_{\text{dd,wdr}}(x)$ be the yield of X considering die-to-die and WDR variations of X. If n is the number of instances of X on the chip, then

$$
\begin{aligned}
Y_{\text{dd,wdr}}(x) &= \mathcal{P}\{X_{\text{dd}} + X_{\text{wdr}}(i) \leq x, \quad i = 1, \dots, n\} \\
&= \mathcal{P}\{X_{\text{wdr}}(i) \leq x - X_{\text{dd}}, \quad i = 1, \dots, n\}. \tag{38}
\end{aligned}
$$

We now recall a result from a basic probability theory that will be used repeatedly in this paper. Let $\mathcal{A}$ be an arbitrary event

and X be an RV with a probability density function (pdf) $f_X(\cdot)$. Then, we have (see [31, p. 85])

$$\mathcal{P}\{\mathcal{A}\} = \int_{-\infty}^{+\infty} \mathcal{P}\{\mathcal{A}|X = x\}f_X(x)dx. \qquad (39)$$

This result is simply an extension to the continuous case of the simple fact that $\mathcal{P}\{\mathcal{A}\} = \mathcal{P}\{\mathcal{A}|\mathcal{B}\} \cdot \mathcal{P}\{\mathcal{B}\} + \mathcal{P}\{\mathcal{A}|\overline{\mathcal{B}}\} \cdot \mathcal{P}\{\overline{\mathcal{B}}\}$, where $\mathcal{B}$ is another event. Applying (39) to (38) and denoting by $f_{X_{\mathrm{dd}}}(\cdot)$ the pdf of X_{dd}, give

$$Y_{\mathrm{dd,wdr}}(x)$$
$$= \int_{-\infty}^{+\infty} \mathcal{P}\{X_{\mathrm{wdr}}(i) \leq x - X_{\mathrm{dd}}, \ \forall i | X_{\mathrm{dd}} = z\} f_{X_{\mathrm{dd}}}(z)dz$$
$$= \int_{-\infty}^{+\infty} \mathcal{P}\{X_{\mathrm{wdr}}(i) \leq x - z, \ \forall i | X_{\mathrm{dd}} = z\} f_{X_{\mathrm{dd}}}(z)dz$$
$$= \int_{-\infty}^{+\infty} \mathcal{P}\{X_{\mathrm{wdr}}(i) \leq x - z, \ \forall i\} f_{X_{\mathrm{dd}}}(z)dz \qquad (40)$$

where the transition between the second and third equation in (40) is due to the statistical independence of X_{dd} and $X_{\mathrm{wdr}}(i)$, $\forall i$. If $\phi(\cdot)$ denotes the pdf of the standard normal distribution, then $f_{X_{\mathrm{dd}}}(z) = \phi(z/\sigma_{\mathrm{dd}})/\sigma_{\mathrm{dd}}$. Letting $z_0 = z/\sigma_{\mathrm{dd}}$ and using the statistical independence of $X_{\mathrm{wdr}}(i)$ and $X_{\mathrm{wdr}}(j)$ for $i \neq j$, we can write (40) as

$$Y_{\mathrm{dd,wdr}}(x) = \int_{-\infty}^{+\infty} \prod_{i=1}^{n} \Phi\left(\frac{x - z_0\sigma_{\mathrm{dd}}}{\sigma_{\mathrm{wdr}}(i)}\right) \phi(z_0)dz_0. \qquad (41)$$

We now introduce another basic result from the probability theory that will be used repeatedly in this paper, which is commonly referred to as the law of the unconscious statistician. If X is an RV with pdf $f_X(\cdot)$ and if $g(\cdot)$ is a function, then, (see [31, p. 106]) the mean of $g(X)$ can be written as

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f_X(x)dx. \qquad (42)$$

Based on this, and if we define Z_0 to be an independent[1] standard normal RV, then (41) can be written as

$$Y_{\mathrm{dd,wdr}}(x) = E\left[\prod_{i=1}^{n} \Phi\left(\frac{x - Z_0\sigma_{\mathrm{dd}}}{\sigma_{\mathrm{wdr}}(i)}\right)\right]. \qquad (43)$$

We have so far assumed the parameter variations to be Gaussian, which technically allows the variations to extend to $\pm\infty$, allowing for nonphysical situations. In reality, one would expect the process variations to be bounded by some upper and

lower bounds, as process tolerances fall within a certain range. If a device somewhere deviates by larger amounts, then chances are, there is a serious problem with that die, and that it would be lost due to other reasons, that is, other than timing yield. Therefore, when estimating a parametric yield, we truncate normal variations at $\pm k\sigma$, where k is an arbitrary positive number and σ is the standard deviation of the untruncated distribution. For simplicity of presentation, we apply the truncation only to the within-die random variations, since this is where the nonphysical tails of the normal distribution have a big effect on the yield estimates because the cdfs of the WDR variations are multiplied n times, as can be observed in (43).

We affect the truncation of the normal distribution by conditioning it over the interval $[-k\sigma, +k\sigma]$, leading to the so-called truncated normal distribution. Let $\Phi_k(\cdot)$ denote the cdf of the standard normal distribution truncated at $\pm k$, so that

$$\Phi_k(x) = \begin{cases} 0, & x < -k \\ \frac{\Phi(x)-\Phi(-k)}{\Phi(k)-\Phi(-k)}, & -k \leq x \leq k \\ 1, & x > k. \end{cases}$$

It can be easily verified that if X is a zero-mean Gaussian RV, with standard deviation σ and cdf $\Phi(x/\sigma)$, then conditioning X on being in the interval $[-k\sigma, +k\sigma]$ yields the cdf $\Phi_k(x/\sigma)$. We now write (43) as follows:

$$Y_{\mathrm{dd,wdr}}(x) = E\left[\prod_{i=1}^{n} \Phi_k\left(\frac{x - Z_0\sigma_{\mathrm{dd}}}{\sigma_{\mathrm{wdr}}(i)}\right)\right]. \qquad (44)$$

This can be easily computed by numerical integration using the integral form (41), with $\Phi_k(\cdot)$ used in place of $\Phi(\cdot)$. Fig. 2 shows the resulting yield-margin curves, where we have assumed for simplicity that $\sigma_{\mathrm{wdr}}(i)$ is constant across the die: $\sigma_{\mathrm{wdr}}(i) = \sigma_{\mathrm{wdr}}$, we have assumed that $\sigma_{\mathrm{dd}}^2 = \sigma_{\mathrm{wdr}}^2$, and we define $\sigma^2 = \sigma_{\mathrm{dd}}^2 + \sigma_{\mathrm{wdr}}^2$, so that $\sigma_{\mathrm{dd}} = \sigma_{\mathrm{wdr}} = \sigma/\sqrt{2}$. The values on the x axis are normalized in multiples of this σ. The figure includes yield plots for three different values of n, each for both cases of untruncated random variations and random variations truncated at $\pm 3\sigma_{\mathrm{wdr}}$. Effectively, truncating the distribution amounts to cutting off its "tail," leading to less pessimistic yield estimates, especially for large n: the $n = 1e + 06$ and $n = 1e + 08$ plots are indistinguishable. Comparing Fig. 1 with Fig. 2, we can readily contrast the level of yield loss when all the variance is taken to be die to die with that when the variance is split (equally, in our case) between die-to-die and WDR variations. For example, if the variation is attributed only to die-to-die effects, then we need to budget 1.6σ to meet a yield of 95%. Considering the WDR variations to be on a par with the die-to-die variations, the same yield level is reached only at 3.2σ (for very large n and taking truncated WDR variations), that is, the required margin doubles.

B. Timing Yield

Since we considered the path delay to conform to our generic parameter model (see Section III-E), then the results of Section IV-A can now be put to use in order to estimate the

---

[1]Throughout this paper, whenever an individual RV is described as "independent," this means that it is independent of all other RVs under consideration.
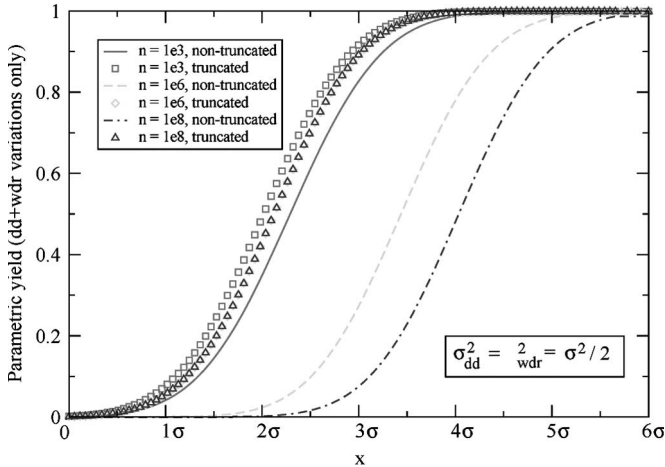
Fig. 2.   Comparison of parametric yield when using untruncated and truncated distributions.

Fig. 3.   Timing yield considering die-to-die and WDR variations.

timing yield $\mathcal{Y}_{\mathrm{dd,wdr}}(d)$. Applying (44) to derive an expression for timing yield, and using (27), we get

$$\mathcal{Y}_{\mathrm{dd,wdr}}(d) = E\left[\Phi_k^n\left(\frac{d - Z_0\sigma_{\mathrm{dd},D_N}}{\hat{\sigma}_{\mathrm{wdr},D_N}}\right)\right]. \qquad (45)$$

For simplicity, we assume that the variance of each physical parameter is equally split between die-to-die and WDR components

$$\sigma_{\mathrm{dd},L}^2 = \hat{\sigma}_{\mathrm{wdr},L}^2 = \sigma_L^2/2$$
$$\sigma_{\mathrm{dd},V}^2 = \hat{\sigma}_{\mathrm{wdr},V}^2 = \sigma_V^2/2$$
$$\sigma_{\mathrm{dd},t}^2 = \hat{\sigma}_{\mathrm{wdr},t}^2 = \sigma_t^2/2$$
$$\sigma_{\mathrm{dd},T_{\mathrm{ILD}}}^2 = \hat{\sigma}_{\mathrm{wdr},T_{\mathrm{ILD}}}^2 = \sigma_{T_{\mathrm{ILD}}}^2/2.$$

These and similar simplifying assumptions, made throughout this paper, are for illustration purposes only and do not affect the generality of the approach. From (24) and (26), it follows that $\sigma_{\mathrm{dd},D_N}^2 = N\sigma_{\mathrm{wdr},D_N}^2(j)$. Then, using (23) and (27), we can write

$$\sigma_{\mathrm{dd},D_N}^2 = N\hat{\sigma}_{\mathrm{wdr},D_N}^2 = \frac{N}{N+1}\sigma_{D_N}^2 \qquad (46)$$

where we have dropped the index j from $\sigma_{D_N}^2(j)$ for simplicity, and because we have used the die average $\hat{\sigma}_{\mathrm{wdr},D_N}^2$ in the equation. Plugging (46) into (45), we get

$$\mathcal{Y}_{\mathrm{dd,wdr}}(d) = E\left[\Phi_k^n\left(\sqrt{N+1}\frac{d}{\sigma_{D_N}} - \sqrt{N}Z_0\right)\right]. \qquad (47)$$

Fig. 3 plots $\mathcal{Y}_{\mathrm{dd,wdr}}(d)$ (with $k=3$) for different values of N and n. The yield-margin curves clearly show the impact of random variations on total path delay diminishing with the number of path stages.

We may follow the same approach as in Section V to compute the timing margins and derive corner-case files in order to check the circuit-timing yield. Once again, a timing margin of $\mathcal{Y}_{\mathrm{dd,wdr}}^{-1}(\mathcal{Y})$ is to be left when simulating the circuit with nominal files, to get a desired yield $\mathcal{Y}$. The value of $\mathcal{Y}_{\mathrm{dd,wdr}}^{-1}(\mathcal{Y})$
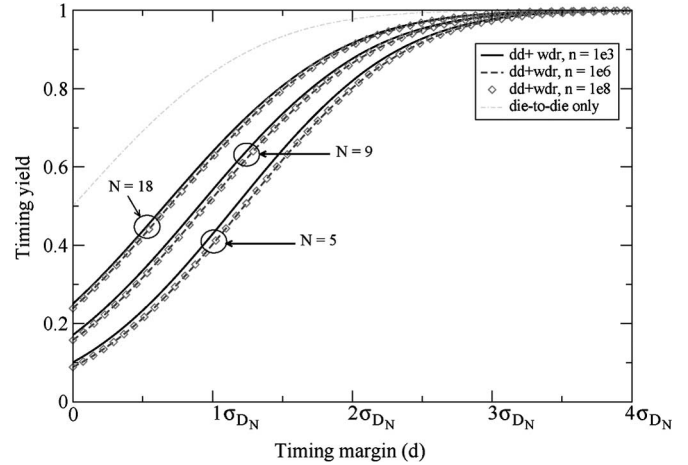
can be obtained from plots such as those in Fig. 2. For example, if $N = 9$ and $n = 1e + 08$, and a 95% yield is desired, then a margin of $2.5\sigma_{D_N}$ is to be budgeted (compared with $1.6\sigma_{D_N}$ when variations are die to die only). We can equally derive a corner-case file that checks for yield $\mathcal{Y}$. Following the discussion in Section V, we let each stage delay deviate by $\mathcal{Y}_{\mathrm{dd,wdr}}^{-1}(\mathcal{Y})/N$, so that

$$\frac{\mathcal{Y}_{\mathrm{dd,wdr}}^{-1}(\mathcal{Y})}{N} = s_1 L(i) + s_2 V(i) - s_3\Delta t(i) - s_4\Delta T_{\mathrm{ILD}}(i), \quad \forall i. \qquad (48)$$

We may further impose a proportional deflection of channel length, threshold voltage, wire thickness, and dielectric thickness by setting

$$\frac{L(i)}{\sigma_L} = \frac{V(i)}{\sigma_V} = \frac{-\Delta t(i)}{\sigma_t} = \frac{-\Delta T_{\mathrm{ILD}}(i)}{\sigma_{T_{\mathrm{ILD}}}} = \delta, \quad \forall i$$

leading to

$$\delta = \frac{\mathcal{Y}_{\mathrm{dd,wdr}}^{-1}(\mathcal{Y})/N}{s_1\sigma_L + s_2\sigma_V + s_3\sigma_t + s_4\sigma_{T_{\mathrm{ILD}}}}. \qquad (49)$$

As before, this value of δ specifies a virtual corner for the circuit device and interconnect that achieves a yield of $\mathcal{Y}$ (if only die-to-die and WDR variations are considered) when it meets the design timing constraints. To compare the new virtual corner with that in (37), we assume again that the device and interconnect parameters have the same effect on path delay

$$s_1^2\sigma_L^2 = s_2^2\sigma_V^2 = s_3^2\sigma_t^2 = s_4^2\sigma_{T_{\mathrm{ILD}}}^2 = s^2$$

so that $\sigma_{D_N} = \sqrt{2}Ns(1 + 1/N)^{1/2}$ and, with $N = 9$ as before, we arrive at

$$\delta = \frac{2.5\sqrt{2}Ns(1 + 1/N)^{1/2}}{4Ns} = 0.93 \qquad (50)$$

which is more than 16% larger than the value obtained in (37), where all variations were considered die to die.

VII. FULL PARAMETRIC YIELD MODEL

We now generalize the results to include all three components of the variations of the parameter X. The focus in this section is on a generic parameter, which conforms with the model of Section II. The application to timing yield will be done in the next section. Recall that the yield of X was defined as

$$Y(x) = \mathcal{P}\{X(i) \leq x, \quad i = 1, 2, \ldots, n\} \quad (51)$$

where $X(i) = X_{\mathrm{dd}} + X_{\mathrm{wds}}(x_i, y_i) + X_{\mathrm{wdr}}(i)$. Because $X_{\mathrm{dd}} \sim \mathcal{N}(0, \sigma_{\mathrm{dd}}^2)$ (this notation means that the RV is normally distributed with mean 0 and variance σ_{dd}^2) is independent of $X_{\mathrm{wds}}(x_i, y_i)$ and of $X_{\mathrm{wdr}}(i)$, $\forall i$, then $Z_0 = X_{\mathrm{dd}}/\sigma_{\mathrm{dd}}$ is an independent standard normal RV (mean 0, variance 1). With this notation, we write $Y(x)$ as

$$Y(x) = \mathcal{P}\{\sigma_{\mathrm{dd}} Z_0 + X_{\mathrm{wds}}(x_i, y_i) + X_{\mathrm{wdr}}(i) \leq x, \quad \forall i\}. \quad (52)$$

Applying (39) to (52), and because Z_0 is independent of all other RVs, then

$$Y(x) = \int_{-\infty}^{+\infty} \mathcal{P}\{X_{\mathrm{wdr}}(i) + X_{\mathrm{wds}}(x_i, y_i) \leq x - \sigma_{\mathrm{dd}} z_0, \forall i\} \phi(z_0) dz_0. \quad (53)$$

A. Yield Upper Bound

We start by stating the proof, in the multidimensional case, of a well-known lemma which will be useful in establishing an upper bound on parametric yield.

Lemma 1: Let $X_1, \ldots, X_n$ be nonnegative RVs. Then, $E[\prod_{i=1}^n X_i] \leq \prod_{i=1}^n (E[X_i^n])^{1/n}$.

Proof: Let $\mathbf{u} = [u_1, u_2, \ldots, u_n]$ be a real vector. If $u_i \geq 0$, then the function $f(\mathbf{u}) = \prod_{i=1}^n u_i^{1/n}$ is concave (see [32, p. 74]). Thus, $g(\mathbf{u}) = -f(\mathbf{u})$ is convex, and if $U_1, \ldots, U_n$ are nonnegative RVs, then by Jensen's inequality (see [32, p. 77]), we have that $E[g(\mathbf{u})] \geq g(E[\mathbf{u}])$ so that $E[\prod_{i=1}^n U_i^{1/n}] \leq \prod_{i=1}^n (E[U_i])^{1/n}$. The desired result follows by letting $X_i = U_i^{1/n}$. ∎

Let $f_{X_{\mathrm{wds}}}(\cdot)$ be the joint pdf of the n RVs $X_{\mathrm{wds}}(x_i, y_i)$. Using the independence of $X_{\mathrm{wdr}}(i)$, it can be shown that (53) leads to

$$Y(x) = \int_{-\infty}^{+\infty} \left[\underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{n} \prod_{i=1}^n \Phi_k \left(\frac{x - \sigma_{\mathrm{dd}} z_0 - x_i}{\sigma_{\mathrm{wdr}}(i)} \right) \right.$$
$$\left. \times f_{X_{\mathrm{wds}}}(x_1, \ldots, x_n) dx_1, \ldots, dx_n \right] \phi(z_0) dz_0$$
$$= E\left[\prod_{i=1}^n \Phi_k \left(\frac{x - \sigma_{\mathrm{dd}} Z_0 - X_{\mathrm{wds}}(x_i, y_i)}{\sigma_{\mathrm{wdr}}(i)} \right) \right] \quad (54)$$

where the second equality is due to (42).

Each $\Phi_k(\cdot)$ term in the above product can be considered a nonnegative RV in its own right. Let Z_1 be an independent standard normal RV. Applying Lemma 1, we obtain

$$Y(x) \leq \prod_{i=1}^n \left(E\left[\Phi_k^n \left(\frac{x - \sigma_{\mathrm{dd}} Z_0 - \sigma_{\mathrm{wds}}(x_i, y_i) Z_1}{\sigma_{\mathrm{wdr}}(i)} \right) \right] \right)^{1/n}. \quad (55)$$

The computation of this upper bound depends on the values of within-die variances for different locations on a die. When the values of $\sigma_{\mathrm{wds}}(x_i, y_i)$ and $\sigma_{\mathrm{wdr}}(i)$ are known, then the right-hand side of the inequality in (55) can be calculated. In a more realistic situation, however, the values of systematic and WDR variance components are not determined at all locations, but rather, we expect to know, from process data, representative values of these variances, such as their average, minimum, and maximum. In this case, one can show that the above upper bound can be adequately estimated via an integral expansion of (55) using only the average and extreme variance values [24].

As an interesting special case, assume there is a constant value of the variance $\sigma_{\mathrm{wdr}}^2(i)$ across the die, equal simply to σ_{wdr}^2, and similarly assume that $\sigma_{\mathrm{wds}}^2(x_i, y_i) = \sigma_{\mathrm{wds}}^2$, $\forall i$. If $Y_1(x)$ is the upper bound on $Y(x)$, given by (55), then it simplifies to

$$Y_1(x) = E\left[\Phi_k^n \left(\frac{x - \sigma_{\mathrm{dd}} Z_0 - \sigma_{\mathrm{wds}} Z_1}{\sigma_{\mathrm{wdr}}} \right) \right]. \quad (56)$$

This can be evaluated by numerical integration. To illustrate, consider the special case where $\sigma_{\mathrm{dd}}^2 = 2\sigma_{\mathrm{wds}}^2 = 2\sigma_{\mathrm{wdr}}^2$ and let $\sigma^2 = \sigma_{\mathrm{dd}}^2 + \sigma_{\mathrm{wds}}^2 + \sigma_{\mathrm{wdr}}^2$ be the total variance, then the yield upper bound is given by the simpler expression

$$Y_1(x) = E\left[\Phi_k^n \left(\frac{2x}{\sigma} - \sqrt{2} Z_0 - Z_1 \right) \right]. \quad (57)$$

The computation of $Y_1(x)$, involving a single double integration in this case, is fairly simple. For the illustrated case, this integration may be performed by affecting a change of variables: $u = \Phi(z_0)$ and $v = \Phi(z_1)$, so that

$$Y_1(x) = \int_0^1 \int_0^1 \Phi_k^n \left(\frac{2x}{\sigma} - \sqrt{2} \Phi^{-1}(u) - \Phi^{-1}(v) \right) du dv. \quad (58)$$

Fig. 4 plots this yield upper bound versus x, for $k = 3$. For comparison, we show the yield estimates when all the within-die variations are taken as random. The plots illustrate the potential for yield underestimation when within-die variations are considered to be entirely random. In our example, a yield of 95% may be reached at about 2.9σ (for $n = 1e + 08$) when accounting for systematic variations, while for the same deviation, random-only variations limit the yield to about 86%.
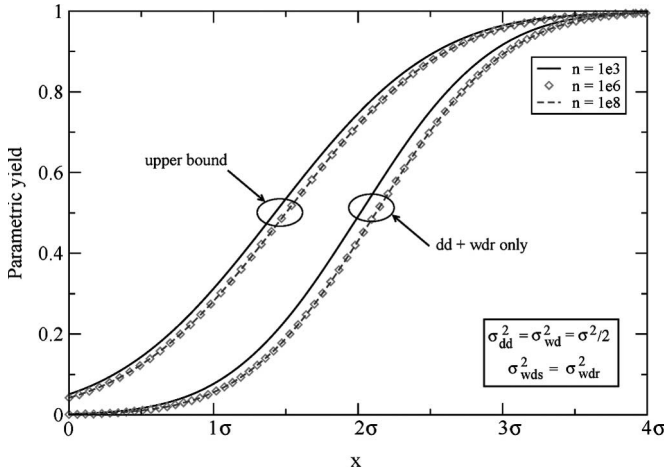
Fig. 4.   Upper bound on yield when considering die-to-die and both types of within-die variations.

B. Yield Lower Bounds

Yield lower bounds are more useful. Recalling the pth-order PCA expansion of $X_{\text{wds}}(x_i, y_i)$ given in (5), we start by defining the RV Q_p as follows:

$$Q_p = \begin{cases} Z_1 \sim \mathcal{N}(0,1) & p = 1 \\ \left(\sum_{j=1}^{p} Z_j^2 \right)^{1/2} \Rightarrow Q_p^2 \sim \chi_p^2 & p > 1 \end{cases}$$

where $\mathcal{N}(0,1)$ denotes the standard normal distribution and χ_p^2 denotes the chi-square distribution with p degrees of freedom [31]. For $p = 1$, we can write

$$X_{\text{wdr}}(i) + X_{\text{wds}}(x_i, y_i) = X_{\text{wdr}}(i) + a_{i1} Z_1$$
$$= X_{\text{wdr}}(i) + \sigma_{\text{wds}}(x_i, y_i) Q_1. \quad (59)$$

For $p > 1$, by Cauchy's inequality [31], we have

$$X_{\text{wdr}}(i) + X_{\text{wds}}(x_i, y_i)$$
$$= X_{\text{wdr}}(i) + \sum_{j=1}^{p} a_{ij} Z_j$$
$$\leq X_{\text{wdr}}(i) + \left(\sum_{j=1}^{p} a_{ij}^2 \right)^{1/2} \left(\sum_{j=1}^{p} Z_j^2 \right)^{1/2}$$
$$= X_{\text{wdr}}(i) + \sigma_{\text{wds}}(x_i, y_i) Q_p \quad (60)$$

where we used (6) to introduce $\sigma_{\text{wds}}(x_i, y_i)$.

Therefore, for $a \in \mathbb{R}$ and $i = 1, \ldots, n$, $X_{\text{wdr}}(i) + \sigma_{\text{wds}}(x_i, y_i) Q_p \leq a$ is a sufficient condition for $X_{\text{wdr}}(i) + \sum_{j=1}^{n} a_{ij} Z_j \leq a$, $\forall p$, and a necessary condition as well when $p = 1$. Define

$$Y_{\text{wd},0}(a) = \mathcal{P} \left\{ X_{\text{wdr}}(i) + \sigma_{\text{wds}}(x_i, y_i) Q_p \leq a, \quad \forall i \right\}.$$

Recalling (53) and using the PCA expansion of $X_{\text{wds}}(x_i, y_i)$, we can write

$$Y(x) = \int_{-\infty}^{+\infty} \mathcal{P} \left\{ X_{\text{wdr}}(i) + \sum_{j=1}^{p} a_{ij} Z_j \leq x - \sigma_{\text{dd}} z_0, \forall i \right\} \phi(z_0) dz_0$$
$$\geq \int_{-\infty}^{+\infty} Y_{\text{wd},0}(x - \sigma_{\text{dd}} z_0) \phi(z_0) dz_0 \quad (61)$$

where the inequality is obtained based on (60), and where the inequality reduces to an equality for $p = 1$.

Notice that $Y_{\text{wd},0}(a)$ can be written as

$$Y_{\text{wd},0}(a) = \mathcal{P} \left\{ Q_p \leq \min_{\forall i} \left(\frac{a - X_{\text{wdr}}(i)}{\sigma_{\text{wds}}(x_i, y_i)} \right) \right\}. \quad (62)$$

Consider each RV

$$M_i = \frac{a - X_{\text{wdr}}(i)}{\sigma_{\text{wds}}(x_i, y_i)}. \quad (63)$$

Each of these RVs is independent, and has a mean of $a / \sigma_{\text{wds}}(x_i, y_i)$ and a variance of $\sigma_{\text{wdr}}^2(i) / \sigma_{\text{wds}}^2(x_i, y_i)$. Let RV M denote the minimum

$$M = \min_{\forall i} \left(\frac{a - X_{\text{wdr}}(i)}{\sigma_{\text{wds}}(x_i, y_i)} \right) \quad (64)$$

and let $F_M(\cdot)$ and $f_M(\cdot)$, respectively, be the cdf and pdf of M. Considering that $X_{\text{wdr}}(i)$ is a truncated normal with the cdf $\Phi_k(\cdot)$ given earlier, it can easily be shown that

$$F_M(y) = 1 - \prod_{i=1}^{n} \Phi_k \left(\frac{a - \sigma_{\text{wds}}(x_i, y_i) y}{\sigma_{\text{wdr}}(i)} \right). \quad (65)$$

Let $f_{Q_p}(\cdot)$ denote the pdf of Q_p, then, making use of the fact that Q_p and M are independent, we can express $Y_{\text{wd},0}(a)$ as

$$Y_{\text{wd},0}(a) = P\{Q_p \leq M\}$$
$$= \int_{-\infty}^{+\infty} \int_{q}^{+\infty} f_M(y) f_{Q_p}(q) dy dq$$
$$= \int_{-\infty}^{+\infty} (1 - F_M(q)) f_{Q_p}(q) dq$$
$$= E[1 - F_M(Q_p)]$$
$$= E \left[\prod_{i=1}^{n} \Phi_k \left(\frac{a - \sigma_{\text{wds}}(x_i, y_i) Q_p}{\sigma_{\text{wdr}}(i)} \right) \right] \quad (66)$$

where we have made use, again, of the law of the unconscious statistician (42) in the fourth line. Applying (66) to (61) and making further use of (42) lead to

$$Y(x) \geq Y_0(x) = E\left[\prod_{i=1}^{n} \Phi_k\left(\frac{x - \sigma_{\mathrm{dd}}Z_0 - \sigma_{\mathrm{wds}}(x_i, y_i)Q_p}{\sigma_{\mathrm{wdr}}(i)}\right)\right] \tag{67}$$

where $Y_0(x)$ is the desired yield lower bound, which can be computed using the above expression. Notice that the inequality reduces to an equality for $p = 1$.

As was the case with the upper bound, the previous expression for the yield lower bound depends on the variances at different die locations, but a suitable integral expansion of (67) can be derived to estimate a yield lower bound when only the mean, maximum, and minimum values of within-die variances are known [24]. As an interesting special case, assume there is a constant value of the variance $\sigma_{\mathrm{wdr}}^2(i)$ across the die, equal simply to σ_{wdr}^2, and similarly assume that $\sigma_{\mathrm{wds}}^2(x_i, y_i) = \sigma_{\mathrm{wds}}^2$, $\forall i$, then the yield lower bound is given by the simpler expression

$$Y_0(x) = E\left[\Phi_k^n\left(\frac{x - \sigma_{\mathrm{dd}}Z_0 - \sigma_{\mathrm{wds}}Q_p}{\sigma_{\mathrm{wdr}}}\right)\right]. \tag{68}$$

This can be evaluated by numerical integration. To illustrate, we consider the same special case as in Section VII-A. Let $\sigma_{\mathrm{dd}}^2 = 2\sigma_{\mathrm{wds}}^2 = 2\sigma_{\mathrm{wdr}}^2$ and let $\sigma^2 = \sigma_{\mathrm{dd}}^2 + \sigma_{\mathrm{wds}}^2 + \sigma_{\mathrm{wdr}}^2$ be the total variance, then

$$Y_0(x) = E\left[\Phi_k^n\left(\frac{2x}{\sigma} - \sqrt{2}Z_0 - Q_p\right)\right]. \tag{69}$$

Again, computation of $Y_0(x)$ in this example comes down to a simple double integration, which may be done via a straightforward change of variables. For $p = 1$, let $u = \Phi(z_0)$ and $v = \Phi(q)$. Then, (69) leads to

$$Y_0(x) = \int_0^1\int_0^1 \Phi_k^n\left(\frac{2x}{\sigma} - \sqrt{2}\Phi^{-1}(u) - \Phi^{-1}(v)\right) du\,dv. \tag{70}$$

For $p > 1$, and denoting by $F_{\chi_p^2}(\cdot)$ the cdf of the χ^2 distribution with p degrees of freedom, let $u = \Phi(z_0)$ and $v = F_{\chi_p^2}(q)$. Then, the yield lower bound may be evaluated as

$$Y_0(x) = \int_0^1\int_0^1 \Phi_k^n\left(\frac{2x}{\sigma} - \sqrt{2}\Phi^{-1}(u) - \sqrt{F_{\chi_p^2}^{-1}(v)}\right) du\,dv. \tag{71}$$

An important observation is due: When $p = 1$, meaning that systematic variations are all perfectly correlated and can be expressed with only one principal component, then (57) and (69) are identical. This means that the lower bound (70) becomes an equality (rather than a bound on the yield), when $p = 1$.

Fig. 5 plots the yield lower bound curves for various values of n and p. For comparison, we also show yield plots when all the within-die variations are assumed random. Observe that
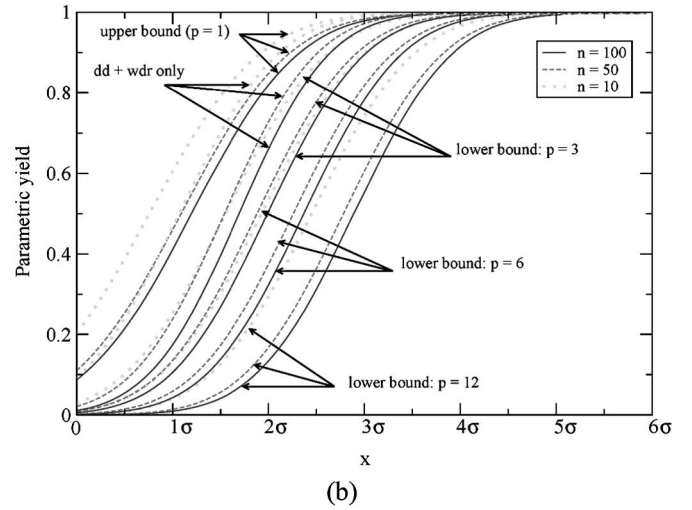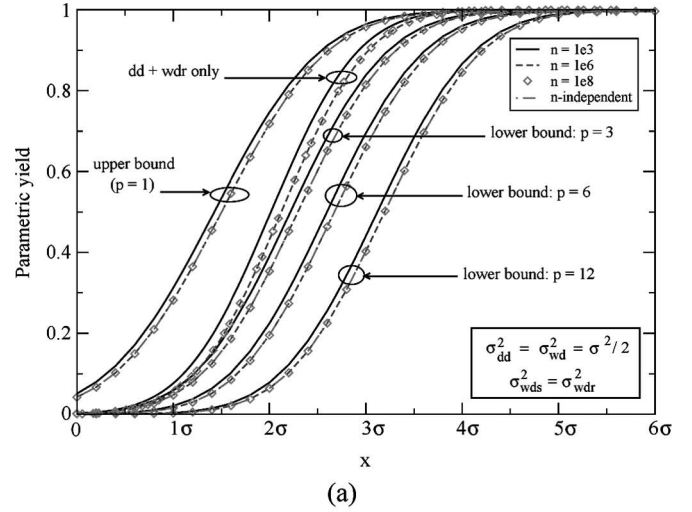

(a)

(b)

Fig. 5. Upper and lower yield bounds for (a) large n and (b) small n.

accounting for within-die variations as entirely random leads to yield estimates that are somewhere between the lower and upper bounds.

A crucial point in our results is that neither the upper nor the lower bound requires any prior knowledge of the layout-dependent within-die correlation. Instead, summary variance information and an estimate of a sufficient number of principal components to capture systematic parameter variations on a chip are all that is needed. While early knowledge of the detailed correlation functions is hard, estimating only the order p of the PCA is clearly much easier. Specifically, we can name three ways in which p can be estimated. First, based on knowledge of the process, it may be possible to simply identify a number of underlying independent factors that are responsible for the systematic variations, such as specific equipment or process steps. Second, one may associate each Z_j with a certain spatial location on the die, such as was done in [20]. Thus, if the chip area is partitioned into, say, four quadrants, and if one has some sense about distances over which the autocorrelation functions die down, one may be able to make an estimation of p. Third, and this may be the easiest approach, we can collect process data relative to some parameter from various locations across the die, and we use the PCA with different

candidate values of p, until we find one that gives an accurate decomposition. In any case, systematic variations generally exhibit smooth trends across a given die [14], [15], often in the form of a "slanted plane" [9], and we expect to be able to capture the bulk of correlations in systematic within-die variations with a small number p of principal components.

C. Yield Bounds Independent of n

It is noteworthy that we can further derive yield bounds which are independent of n. While the analysis to follow can be extended to the general case, we will, for clarity of presentation, limit the discussion to some special cases. From (68), let

$$V = \frac{x - \sigma_{\text{dd}} Z_0 - \sigma_{\text{wds}} Q_p}{\sigma_{\text{wdr}}}. \tag{72}$$

We start with a decomposition of the yield expression using conditional expected values, so that (68) gives

$$Y_0(x) = E\left[\Phi_k^n(V)|V \geq k\right] \times \mathcal{P}\{V \geq k\}$$

$$+ E\left[\Phi_k^n(V)|V < k\right] \times \mathcal{P}\{V < k\}. \tag{73}$$

Since $\Phi_k(x) = 1$ for $x \geq k$, then the first conditional expectation is 1. As for the second expectation, since $V < 1$, then $\Phi_k(V) < 1$ and $\Phi_k^n(V) \to 0$ for large n, which one would expect to have in an integrated circuit. Therefore, the second term can be neglected relative to the first, so that

$$Y_0(x) \approx \mathcal{P}\{\sigma_{\text{dd}} Z_0 + \sigma_{\text{wds}} Q_p \leq x - k\sigma_{\text{wdr}}\}. \tag{74}$$

The resulting (74) is most interesting. It gives the parametric yield (lower bound) for a large ensemble of parameters with a full statistical model, keeping in view the three variance components. To simplify the notation, if we let

$$U = \frac{x - k\sigma_{\text{wdr}} - \sigma_{\text{wds}} q}{\sigma_{\text{dd}}} \tag{75}$$

then, (74) can be simplified as follows:

$$Y_0(x) \approx \int_0^{+\infty} \int_{-\infty}^{U} \phi(z_0) f_{Q_p}(q) dz_0 dq$$

$$= \int_0^{+\infty} \Phi\left(\frac{x - k\sigma_{\text{wdr}} - \sigma_{\text{wds}} q}{\sigma_{\text{dd}}}\right) f_{Q_p}(q) dq$$

$$= E\left[\Phi\left(\frac{x - k\sigma_{\text{wdr}} - \sigma_{\text{wds}} Q_p}{\sigma_{\text{dd}}}\right)\right]. \tag{76}$$

Following the same steps, one can show that for large n, the yield upper bound given in (56) can be expressed as

$$Y_1(x) \approx E\left[\Phi\left(\frac{x - k\sigma_{\text{wdr}} - \sigma_{\text{wds}} Z_1}{\sigma_{\text{dd}}}\right)\right]. \tag{77}$$

These equations can be computed by numerical integration. To illustrate, we appeal once more to our illustrative special case. Let $\sigma_{\text{dd}}^2 = 2\sigma_{\text{wds}}^2 = 2\sigma_{\text{wdr}}^2$ and let $\sigma^2 = \sigma_{\text{dd}}^2 + \sigma_{\text{wds}}^2 + \sigma_{\text{wdr}}^2$ be the total variance, then the previous equation simplifies to

$$Y_0(x) \approx E\left[\Phi\left(\frac{\sqrt{2}x}{\sigma} - \frac{k}{\sqrt{2}} - \frac{Q_p}{\sqrt{2}}\right)\right] \tag{78}$$

and

$$Y_1(x) \approx E\left[\Phi\left(\frac{\sqrt{2}x}{\sigma} - \frac{k}{\sqrt{2}} - \frac{Z_1}{\sqrt{2}}\right)\right]. \tag{79}$$

Plots of these two results are shown in Fig. 5. As expected, these bounds are very tight; they are indistinguishable from the 1e6 and 1e8 curves in each group.

VIII. APPLICATION TO CIRCUIT TIMING

A. Yield-Margin Curves and Virtual Corners

We will now show how the results can be easily applied in the context of timing-yield estimation and timing margin budgeting. For each of the transistor channel length, transistor threshold voltage, wire thickness, and ILD thickness, we assume that, at all die locations, the die-to-die variance is equal to the within-die variance, and WDS and WDR variances are also equal. That is, $\forall i$, we let

$$\sigma_{\text{dd},L}^2 = 2\sigma_{\text{wds},L}^2(x_i, y_i) = 2\hat{\sigma}_{\text{wdr},L}^2 = \sigma_L^2/2$$

$$\sigma_{\text{dd},V}^2 = 2\sigma_{\text{wds},V}^2(x_i, y_i) = 2\hat{\sigma}_{\text{wdr},V}^2 = \sigma_V^2/2$$

$$\sigma_{\text{dd},t}^2 = 2\sigma_{\text{wds},t}^2(x_i, y_i) = 2\hat{\sigma}_{\text{wdr},t}^2 = \sigma_t^2/2$$

$$\sigma_{\text{dd},T_{\text{ILD}}}^2 = 2\sigma_{\text{wds},T_{\text{ILD}}}^2(x_i, y_i) = 2\hat{\sigma}_{\text{wdr},T_{\text{ILD}}}^2 = \sigma_{T_{\text{ILD}}}^2/2.$$

Applying the above with (24)–(26) gives the following relationship between the individual components of the path-delay variance, $\forall i$

$$\sigma_{\text{dd},D_N}^2 = 2\sigma_{\text{wds},D_N}^2(x_i, y_i) = 2N\hat{\sigma}_{\text{wdr},D_N}^2 = \frac{\sigma_{D_N}^2}{1.5 + \frac{0.5}{N}}.$$

Let $\mathcal{Y}_0(d)$ and $\mathcal{Y}_1(d)$ be the lower and upper bounds of timing yield, respectively. Plugging the variance relationships into (76) and (77) leads to

$$\mathcal{Y}_0(d) \approx E\left[\Phi\left(\frac{d}{\sigma_{D_N}}\sqrt{1.5 + \frac{0.5}{N}} - \frac{k}{\sqrt{2N}} - \frac{Q_p}{\sqrt{2}}\right)\right]$$

$$\mathcal{Y}_1(d) \approx E\left[\Phi\left(\frac{d}{\sigma_{D_N}}\sqrt{1.5 + \frac{0.5}{N}} - \frac{k}{\sqrt{2N}} - \frac{Z_1}{\sqrt{2}}\right)\right].$$

Plots of these bounds are shown in the yield-margin curves of Fig. 6, for various values of p, with $N = 9$ and $k = 3$.
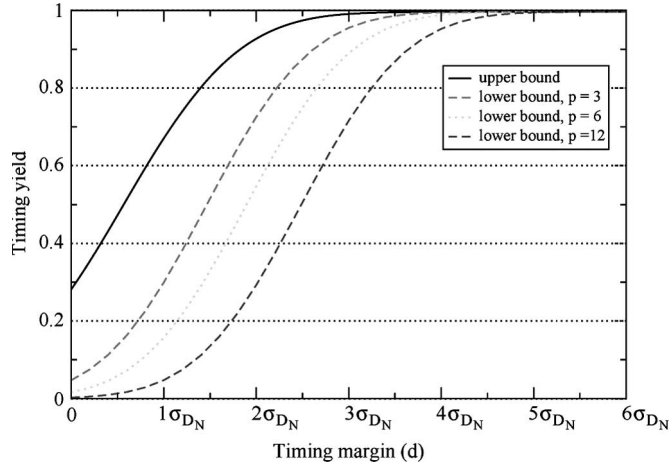
Fig. 6. Timing-yield bounds.

The same analysis as before can be done to compute the timing margins and corner file settings that check for yield, this time incorporating all parameter variations. Suppose that each of the four physical parameters requires in its representation three principal components. From (22), path delay will be expanded in terms of 12 principal components. Fig. 6 indicates that when the circuit is simulated at nominal process conditions, a timing margin of at most $4\sigma_{D_N}$ needs to be budgeted to achieve a yield of 95%. We can also easily derive an appropriate file setting by imposing a delay deviation of $4\sigma_{D_N}/N$ at each stage, so that

$$\frac{4\sigma_{D_N}}{N} = s_1 L(i) + s_2 V(i) - s_3 \Delta t(i) - s_4 \Delta T_{\text{ILD}}(i), \quad \forall i.$$

Now, let

$$\frac{L(i)}{\sigma_L} = \frac{V(i)}{\sigma_V} = \frac{-\Delta t(i)}{\sigma_{\text{t}}} = \frac{-\Delta T_{\text{ILD}}(i)}{\sigma_{T_{\text{ILD}}}} = \delta, \quad \forall i$$

where δ is our virtual corner in this case. This leads to

$$\delta = \frac{4\sigma_{D_N}/N}{s_1\sigma_L + s_2\sigma_V + s_3\sigma_{\text{t}} + s_4\sigma_{T_{\text{ILD}}}}.$$

To compare the resulting virtual corner with (37) and (50), we make the same assumption that all four physical parameters contribute equally to path-delay variance

$$s_1^2\sigma_L^2 = s_2^2\sigma_V^2 = s_3^2\sigma_{\text{t}}^2 = s_4^2\sigma_{T_{\text{ILD}}}^2 = s^2 \qquad (80)$$

so that (24)–(26) yield $\sigma_{D_N} = 2sN(0.75 + 0.25/N)^{1/2}$. We arrive at

$$\delta = \frac{4 \times 2s \times 9 \times (0.75 + 0.25/9)^{1/2}}{9 \times 4s} = 1.76$$

so that the circuit would need to be simulated (and its timing checked) with transistors' channel length and threshold-voltage set at their 1.76σ point, and interconnect and wire-thickness set at their -1.76σ point. Since s_1 and s_2, defined in (13), depend

on transistor sizing, the above provides a way in which δ can be controlled by circuit optimization and/or process tuning.

B. Monte Carlo (MC) Simulations

We compare the yield bounds derived above with the yield-margin curves obtained from MC simulations. The purpose of the MC experiments is to check the validity of the proposed yield bounds under arbitrary correlations of the physical parameters on chip, since the bounds are meant to provide conservative estimates of yield for any correlations.

The experimental setup was as follows: We first generate a number of identical generic paths, each of a certain depth. Both the number of paths and their depths are user specified. The stages in each path are all similar and made of a generic gate with generic fan-out interconnect. The sensitivities of gate delay to channel length and threshold voltage are derived from a 0.13-μm technology and are given as inputs in these experiments (ignoring threshold voltage roll-off). The user also specifies a nominal RC delay for the generic gate fan-out, which is used to compute the sensitivities of interconnect delay to metal and ILD thickness as per (18), given a nominal thickness of ILD (Metal1—Substrate) and interconnect (Metal1). The user also defines the standard deviation of the channel length, and (80) is used to compute the standard deviation of the other three physical parameters. The user also specifies the number of principal components for each of L_{e}, V_{t}, Δt, and ΔT_{ILD} and the fraction of total variability due to each of the die-to-die, WDS, and within-die random components in each of these parameters.

MC experiments were performed by generating random co-efficients for the PCA expansion (a_{ij}) of every parameter in every gate, such that the sum of squares of each coefficient corresponds to the parameter variance in every instance. This assignment of PCA coefficients is repeated `NumberOfCurves` times, each corresponding to a different correlation structure for each parameter on the chip. For every one of these random coefficient assignments, we sample independent realizations of RVs corresponding to die-to-die, within-die random, and standard normal RVs corresponding to the PCA expansion for every parameter. This sampling is such that a global die-to-die variable is sampled for every parameter type (channel length, threshold voltage, ILD, and metal thickness) and assigned to every gate or interconnect instance on the circuit; the PCA-derived standard normal RVs (whose number corresponds to the PCA order expansion for every parameter) are sampled independently for every parameter and assigned to all parameter instances on the circuit; finally, the random components are sampled independently for every parameter and every instance of the circuit. Thus, the parameter model presented in this paper is reproduced with its within-die correlation based on PCA expansions. For every one of the `NumberOfCurves` co-efficient assignments, independent RV sampling is repeated `MonteCarloSamples` times, each leading to one realization of stage, path, and circuit delay. A Monte-Carlo-based realization of a yield-margin curve is deduced for the given correlation structure when `MonteCarloSamples` delay realizations are obtained. Altogether, `NumberOfCurves` realizations of
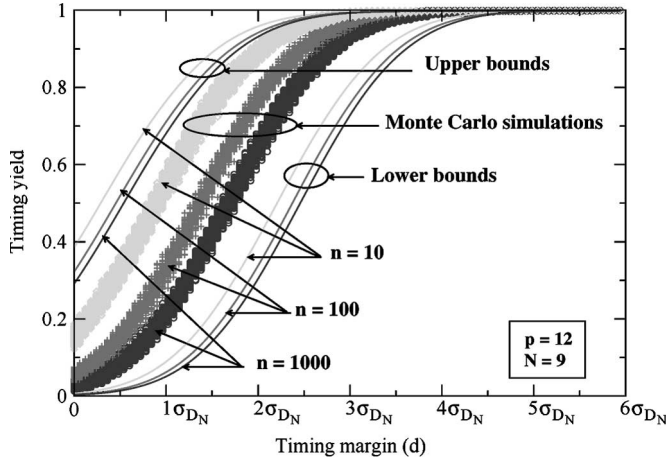
Fig. 7. Yield bounds and MC curves for various number of paths.

Fig. 8. Yield bounds and MC curves for various PCA orders.

yield-margin curves are obtained, each corresponding to a particular (randomly generated) correlation structure. The following summarizes the MC procedure.

-**Repeat** `NumberOfCurves` times:
   -For each instance i of L, V, t, and T_{ILD}, generate random PCA coefficients $a_{\mathrm{param},ij}$, where param $= L, V, t, T_{\mathrm{ILD}}$.
   -**Repeat** `MonteCarloSamples` times:
     -Sample independently L_{dd}, V_{dd}, t_{dd}, and $T_{\mathrm{ILD,dd}}$.
     -Sample independent standard normal RVs $Z_{\mathrm{param},i}$ for the wds variations of each parameter type.
     -For each gate and interconnect i, sample independently $L_{\mathrm{wdr}}(i)$, $V_{\mathrm{wdr}}(i)$, $t_{\mathrm{wdr}}(i)$, and $T_{\mathrm{ILD,wdr}}(i)$.
     -Compute the realization of each parameter at each instance i:
$$L(i) = L_{\mathrm{dd}} + \sum a_{L,ij}Z_{L,i} + L_{\mathrm{wdr}}(i)$$
$$V(i) = V_{\mathrm{dd}} + \sum a_{V_t,ij}Z_{V_t,i} + V_{\mathrm{wdr}}(i)$$
$$t(i) = t_{\mathrm{dd}} + \sum a_{t,ij}Z_{t,i} + t_{\mathrm{wdr}}(i)$$
$$T_{\mathrm{ILD}}(i) = T_{\mathrm{ILD,dd}} + \sum a_{T_{\mathrm{ILD}},ij}Z_{T_{\mathrm{ILD}},i} + T_{\mathrm{ILD,wdr}}(i)$$
     -Compute a realization of the delay of all stages i:
$$S(i) = s_1 L(i) + s_2 V(i) + s_3 t(i) + s_4 T_{\mathrm{ILD}}(i)$$
     -Compute a realization of the delay of all paths j as $D_N(j) = \sum S(i)$, for all stages i lying on path j.
     -A realization of circuit delay is: $\max D_N(j)$.
     -end **Repeat**
   -Rank circuit-delay realizations and deduce a realization of a yield-margin curve.
-end **Repeat**.

The MC experiments were carried out by writing C programs, with various results illustrated in Figs. 7–9, comparing the derived upper and lower bounds with MC simulations when varying the number of paths n, the number of principal components p of path delay and path depth N. The maximum number of paths for which it was practical to run MC simulations was 1000. The figures were obtained with `NumberOfCurves` $= 300$ and `MonteCarloSamples` $= 2000$. In all cases, the bounds were found to hold for all correlations and all timing margins, and to mirror the results of MC simulations when parameters are varied. We note the overlap of the MC curves for various p and various N, shown in Figs. 8 and 9. From the plots, we
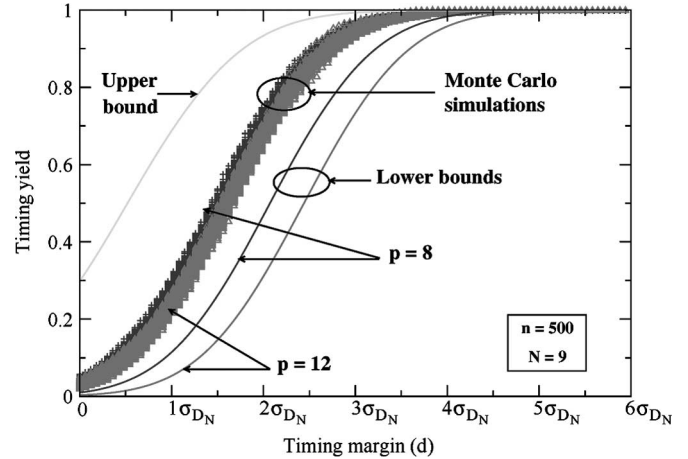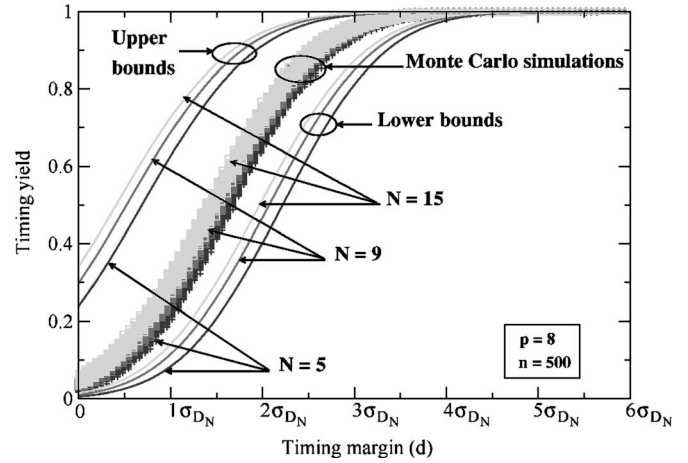

Fig. 9. Yield bounds and MC curves for various path depths.

observe that the lower bounds tend to be closer to the MC curves for larger numbers of paths and smaller PCA orders, and upper bounds tend to be closer to the MC curves for smaller numbers of paths. In the other cases, the MC curves fall around the center of the interval defined by the lower and upper bounds. Tables I–III illustrate the results of the corresponding figures (respectively Figs. 7–9), emphasizing timing margins for selected yield levels. The tables compare the timing margins obtained by the lower bound curve and the upper bound curve and shown in the third and fourth columns of the tables (remark that the timing margin corresponding to the upper bound curve is smaller than that corresponding to the lower bound curve and vice versa), with the spread in the timing margin obtained from our MC simulations, which we refer to as the "Monte Carlo range," and which is shown as the last column in each table.

IX. CONCLUSION

We proposed a generic parameter model for process-induced variation and applied it in statistical timing analysis. The model is a "full-chip" model, in that, it can be applied with ease to large chips, before layout. This is achieved by using a measure of yield based on use of a generic critical path concept and capturing the statistics of a large collection of such paths with

TABLE I
COMPARISON OF MC RESULT WITH THE DERIVED BOUNDS FOR
DIFFERENT VALUES OF THE NUMBER OF CRITICAL PATHS n

Yield (%)	n	Upper Bound (in units of σ)	Lower Bound (in units of σ)	Monte Carlo Range (in units of σ)
50	10	0.29	2.22	0.81–0.99
	100	0.45	2.32	1.29–1.53
	1000	0.54	2.46	1.59–1.77
70	10	0.79	2.70	1.23–1.47
	100	0.99	2.85	1.71–1.95
	1000	1.05	2.94	2.01–2.25
85	10	1.32	3.15	1.65–1.95
	100	1.47	3.33	2.13–2.37
	1000	1.56	3.39	2.49–2.67
90	10	1.55	3.36	1.89–2.19
	100	1.72	3.54	2.31–2.67
	1000	1.80	3.62	2.67–2.91
95	10	1.92	3.70	2.19–2.55
	100	2.08	3.87	2.67–3.03
	1000	2.16	3.94	2.91–3.27
99	10	2.64	4.34	2.67–3.39
	100	2.79	4.50	3.15–3.75
	1000	2.87	4.59	3.45–3.99
99.5	10	2.85	4.64	2.85–3.75
	100	3.09	4.79	3.33–4.29
	1000	3.18	4.87	3.63–4.35

TABLE II
COMPARISON OF MC RESULTS WITH THE DERIVED BOUNDS FOR
DIFFERENT VALUES OF THE PCA ORDER p

Yield (%)	p	Upper Bound (in units of σ)	Lower Bound (in units of σ)	Monte Carlo Range (in units of σ)
50	8	0.52	2.07	1.41–1.59
	12		2.45	1.53–1.71
70	8	1.05	2.55	1.83–2.07
	12		2.89	1.95–2.13
85	8	1.54	3.01	2.25–2.55
	12		3.39	2.31–2.61
90	8	1.80	3.24	2.49–2.79
	12		3.60	2.49–2.85
95	8	2.14	3.57	2.79–3.09
	12		3.93	2.73–3.21
99	8	2.86	4.23	3.33–3.81
	12		4.59	3.39–3.93
99.5	8	3.17	4.47	3.51–4.35
	12		4.86	3.57–4.23

TABLE III
COMPARISON OF MC RESULTS WITH THE DERIVED BOUNDS FOR
DIFFERENT VALUES OF PATH DEPTH N

Yield (%)	N	Upper Bound (in units of σ)	Lower Bound (in units of σ)	Monte Carlo Range (in units of σ)
50	5	0.70	2.22	1.47–1.71
	9	0.52	2.07	1.41–1.59
	15	0.42	1.98	1.35–1.53
70	5	1.21	2.70	1.95–2.13
	9	1.05	2.55	1.83–2.07
	15	0.93	2.46	1.83–2.01
85	5	1.71	3.15	2.37–2.61
	9	1.54	3.01	2.25–2.55
	15	1.44	2.91	2.25–2.49
90	5	1.98	3.36	2.61–2.85
	9	1.80	3.24	2.49–2.79
	15	1.68	3.14	2.43–2.73
95	5	2.28	3.71	2.85–3.15
	9	2.14	3.57	2.79–3.09
	15	2.04	3.48	2.73–3.09
99	5	3.00	4.32	3.39–3.99
	9	2.86	4.23	3.33–3.81
	15	2.75	4.11	3.21–3.81
99.5	5	3.30	4.60	3.57–4.29
	9	3.17	4.47	3.51–4.35
	15	3.05	4.40	3.45–4.17

The MC capability in this framework works by first generating the systematic variation profiles (or spatial maps) for a set of artificial dies. These profiles are generated using the PCA for the process parameter that exhibits spatial correlation (channel length). The mathematically generated population of dies models the parameter variations that would be seen after manufacturing. Through prior experimentation, we have determined that 800 dies (MC samples) provide an acceptable accuracy for SSTA purposes. Depending on the location of the cell within each die sample, a WDS is calculated for each cell. Hence, a cell that falls within a high-L_e region of the die would be slower than nominal. The WDR offsets are generated using a Gaussian random number generator. STA is then carried out for each of the generated dies using cell delays that are modified depending on the offsets. The stage delay (driver input to receiver input) was calculated as

$$\text{delay} = \text{nominal\_delay}(\text{input slope}, \text{cload})$$
$$+ \text{wds\_offset} \times \text{sensitivity}_{\text{wds}}(\text{input slope}, \text{cload})$$
$$+ \text{wdr\_offset} \times \text{sensitivity}_{\text{wdr}}(\text{input slope}, \text{cload}).$$

In the equation, $\text{sensitivity}_{\text{wds}}$ and $\text{sensitivity}_{\text{wdr}}$ are obtained from the variation-aware library models. Path delays are measured as the difference in arrival times between a capture flop data pin and a launch flop clock pin (over one-cycle paths). Path delay for the longest path is stored for each of the dies and data from the STA runs for all dies was used to generate a distribution.

For the generic-path-based MC analysis, the generic path was created out of an alternating sequence of inverters and NAND2s with a total delay equal to the cycle time T of the design. Each die sample was divided into a grid with the dimension of each grid square set to the distance at which the systematic

a model of within-die correlations based on PCA. This results in a methodology, whereby one can select the right setting of the transistor parameters to be used in simulation or in traditional timing analysis in order to verify the performance while guaranteeing a certain desired yield.

APPENDIX

We conducted a series of experiments to substantiate the validity of the generic-path model in the context of circuit-timing analysis. The accuracy of the generic-path model was verified by comparing the timing results from a standard direct STA-based MC analysis with a generic-path-based MC analysis.

Our standard MC experiments are carried out using the SSTA capabilities in an industrial STA framework. This includes a comprehensive library characterization process to model variation for the significant process parameters affecting delay.

TABLE IV
DESIGN 1: 19 264 CELLS

Metric	Percent. error
Mean	2.40
Sigma	-9.86
90% yield	2.07
95% yield	2.05
99% yield	1.57

TABLE V
DESIGN 2: 131 897 CELLS

Metric	Percent. error
Mean	2.26
Sigma	-10.17
90% yield	1.71
95% yield	1.43
99% yield	0.54

correlation would drop to 0.98. That is, all devices within each grid square are highly correlated. An identical generic path is assigned to each grid square. Essentially, this generic path represents all critical paths within the grid square. For each die sample (or spatial profile), a delay analysis is carried out for the set of generic paths with the variation-aware delay for the path calculated as shown above. The maximum path delay determines the critical delay for that die sample. The overall distribution is obtained by combining the results from all die samples.

Tables IV and V compare the maximum path-delay distribution obtained from the STA-based MC analysis to that obtained from the generic-path-based MC run. These results were obtained for two industrial designs at two successive technology nodes. In these tables, the percentage error is defined as $100 * (Generic\_Path—Direct\_MC)/Direct\_MC$. These tables corroborate the validity of the generic-path model and its usefulness in an actual design cycle.

ACKNOWLEDGMENT

REFERENCES

[1] S. W. Director, W. Maly, and A. J. Strojwas, *VLSI Design Manufacturing for Yield Enhancement*. Boston, MA: Kluwer, 1990.

[2] W. Moore, W. Maly, and A. Strojwas, Eds., *Yield Modelling and Defect Tolerance in VLSI*, Bristol, U.K.: Adam Hilger, 1987.

[3] K. Bernstein, K. M. Carrig, C. M. Durham, P. R. Hansen, D. Hogenmiller, E. J. Nowak, and N. J. Rohrer, *High Speed CMOS Design Styles*. Boston, MA: Kluwer, 1999.

[4] J. B. Khare and W. Maly, *From Contamination to Defects, Faults, and Yield Loss*. Boston, MA: Kluwer, 1996.

[5] M. Eisele, J. Berthold, D. Schmitt-Landsiedel, and R. Mahnkopf, "The impact of intra-die device parameter variations on path delays and on the design for yield of low voltage digital circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 5, no. 4, pp. 360–368, Dec. 1997.

[6] A. Gattiker, S. Nassif, R. Dinakar, and C. Long, "Timing yield estimation from static timing analysis," in *Proc. IEEE Int. Symp. Quality Electron. Des.*, San Jose, CA, Mar. 26–28, 2001, pp. 437–442.

[7] S. R. Nassif, *Statistical Worst-Case Analysis for Integrated Circuits Statistical Approach to VLSI*, ser. Advances in CAD for VLSI, vol. 8,

[8] K. Singhal and V. Visvanathan, "Statistical device models for worst case files and electrical test data," *IEEE Trans. Semicond. Manuf.*, vol. 12, no. 4, pp. 470–484, Nov. 1999.

[9] D. Boning and S. Nassif, "Models of process variations in device and interconnect," in *Design of High-Performance Microprocessor Circuits*, A. Chandrakasan, W. J. Bowhill, and F. Fox, Eds. New York: IEEE Press, 2001, ch. 6.

[10] L. Mizrukhin, J. Huey, and S. Mehta, "Prediction of product yield distributions from wafer parametric measurements of CMOS circuits," *IEEE Trans. Semicond. Manuf.*, vol. 5, no. 2, pp. 88–93, May 1992.

[11] K. K. Low and S. W. Director, "A new methodology for the design centering of IC fabrication processes," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 10, no. 7, pp. 895–903, Jul. 1991.

[12] R. W. Dutton and A. J. Strojwas, "Perspectives on technology and technology-driven CAD," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 19, no. 12, pp. 1544–1560, Dec. 2000.

[13] S. R. Nassif, "Within-chip variability analysis," in *Proc. IEEE Int. Electron. Devices Meeting*, San Francisco, CA, Dec. 6–9, 1998, pp. 283–286.

[14] B. E. Stine, D. S. Boning, and J. E. Chung, "Analysis and decomposition of spatial variation in integrated circuit processes and devices," *IEEE Trans. Semicond. Manuf.*, vol. 10, no. 1, pp. 24–41, Feb. 1997.

[15] V. Mehrotra, S. L. Sam, D. Boning, A. Chandrakasan, R. Vallishayee, and S. Nassif, "A methodology for modeling the effects of systematic within-die interconnect and device variations on circuit performance," in *Proc. ACM/IEEE Des. Autom. Conf.*, Los Angeles, CA, Jun. 5–9, 2000, pp. 172–175.

[16] J. A. G. Jess, K. Kalafala, W. R. Naidu, R. H. J. M. Otten, and C. Visweswariah, "Statistical timing for parametric yield prediction of digital integrated circuits," in *Proc. ACM/IEEE Des. Autom. Conf.*, Anaheim, CA, Jun. 2–6, 2003, pp. 932–937.

[17] C. Visweswariah, "Death, taxes and failing chips," in *Proc. ACM/IEEE Des. Autom. Conf.*, Anaheim, CA, Jun. 2–6, 2003, pp. 343–347.

[18] A. Agarwal, D. Blaauw, V. Zolotov, and S. Vrudhula, "Computation and refinement of statistical bounds on circuit delay," in *Proc. ACM/IEEE Des. Autom. Conf.*, Anaheim, CA, Jun. 2–6, 2003, pp. 348–353.

[19] A. Agarwal, D. Blaauw, and V. Zolotov, "Statistical timing analysis for intra-die process variations with spatial correlations," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Des.*, San Jose, CA, Nov. 9–13, 2003, pp. 900–907.

[20] H. Chang and S. S. Sapatnekar, "Statistical timing analysis considering spatial correlations using a single PERT-like traversal," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Des.*, San Jose, CA, Nov. 9–13, 2003, pp. 621–625.

[21] A. Devgan and C. Kashyap, "Block-based static timing analysis with uncertainty," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Des.*, San Jose, CA, Nov. 9–13, 2003, pp. 607–614.

[22] S. Bhardwaj, S. B. Vrudhula, and D. Blaauw, "Tau: Timing analysis under uncertainty," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Des.*, San Jose, CA, Nov. 9–13, 2003, pp. 615–620.

[23] K. A. Bowman and J. D. Meindl, "Impact of within-die parameter fluctuations on future maximum clock frequency distributions," in *Proc. IEEE Custom Integr. Circuits Conf.*, San Diego, CA, May 6–9, 2001, pp. 229–232.

[24] F. N. Najm and N. Menezes, "Statistical timing analysis based on a timing yield model," in *Proc. ACM/IEEE Des. Autom. Conf.*, San Diego, CA, Jun. 7–11, 2004, pp. 460–465.

[25] X. Bai, C. Visweswariah, P. N. Strenski, and D. J. Hathaway, "Uncertainty-aware circuit optimization," in *Proc. Des. Autom. Conf.*, New Orleans, LA, Jun. 10–14, 2002, pp. 58–63.

[26] M. S. Srivastava, *Methods of Multivariate Statistics*. New York: Wiley-Interscience, 2002.

[27] J. M. Rabaey, *Digital Integrated Circuits: A Design Perspective*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[28] Y. Liu, L. T. Pileggi, and A. J. Strojwas, "Model order-reduction of RC(L) interconnect including variational analysis," in *Proc. Des. Autom. Conf.*, New Orleans, LA, Jun. 21–25, 1999, pp. 201–206.

[29] E. Acar, S. Nassif, Y. Liu, and L. T. Pileggi, "Time-domain simulation of variational interconnect models," in *Proc. IEEE Int. Symp. Quality Electron. Des.*, San Jose, CA, Mar. 18–21, 2002, pp. 419–424.

[30] E. Malavasi, S. Zanella, M. Cao, J. Uschersohn, M. Misheloff, and C. Guardiani, "Impact analysis of process variability on clock skew," in *Proc. IEEE Int. Symp. Quality Electron. Des.*, San Jose, CA, Mar. 18–21, 2002, pp. 129–132.

S. W. Director and W. Maly, Eds. Amsterdam, The Netherlands: North-Holland, 1994.

[31] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge Univ. Press, 2004. [Online]. Available: http://www.stanford.edu/~boyd/cvxbook.html

Farid N. Najm (S'85–M'89–SM'96–F'03) received the B.E. degree in electrical engineering from the American University of Beirut (AUB), Beirut, Lebanon, in 1983 and the M.S. and Ph.D. degrees in electrical and computer engineering (ECE) from the University of Illinois at Urbana-Champaign (UIUC), in 1986 and 1989, respectively.

From 1989 to 1992, he worked with Texas Instruments, Dallas. Then, he joined the ECE Department at UIUC as an Assistant Professor and eventually became an Associate Professor in 1997. In 1999, he joined the ECE Department at the University of Toronto, Toronto, ON, Canada, where he is currently a Professor and a Vice Chair of ECE. He has coauthored the text *Failure Mechanisms in Semiconductor Devices*, 2nd ed. (Wiley, 1997). His research is on computer-aided design (CAD) for integrated circuits, with an emphasis on circuit level issues related to power dissipation, timing, and reliability.

Dr. Najm is an Associate Editor for the IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS. He received the IEEE Transactions on CAD Best Paper Award in 1992, the NSF Research Initiation Award in 1993, the NSF CAREER Award in 1996, and was an Associate Editor for the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS from 1997 to 2002. He served as a General Chairman for the 1999 International Symposium on Low-Power Electronics and Design (ISLPED-99) and as a Technical Program Cochairman for ISLPED-98. He has also served on the technical committees of ICCAD, DAC, CICC, ISQED, and ISLPED.

Noel Menezes (S'93–M'95) received the B.Eng. degree from Maharaja Sayajirao University, Baroda, India, and the M.S. and Ph.D. degrees from the University of Texas, Austin.

He is currently the Manager with the Core CAD research group at Intel's Strategic CAD Labs in Hillsboro, OR. In the past, he has worked on interconnect optimization algorithms, clock tree synthesis, gate-delay engines, statistical static timing analysis, and inductance extraction. Results of his research have been applied in the design of various IBM and Intel microprocessors. He is also a Coinventor of the popular Thevenin gate-driver model for resistive interconnects which is used in most industry timing analyzers. He has been a member of the technical program committee for most major CAD conferences.

Imad A. Ferzli (S'00) received the B.E. degree in computer and communications engineering from the American University of Beirut, Beirut, Lebanon, in 2001, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2004. He is currently working toward the Ph.D. degree at the University of Toronto.

He interned with Magma Design Automation, Eindhoven, The Netherlands, in the summer of 2006. His research interests include CAD development for design for manufacturability and process variations, power grid verification and design, timing analysis, and yield modeling and prediction.

Mr. Ferzli is a Student Member of the Semiconductor Research Corporation.