

# Early Analysis and Budgeting of Margins and Corners Using Two-Sided Analytical Yield Models

Khaled R. Heloue, *Student Member, IEEE*, and Farid N. Najm, *Fellow, IEEE*

**Abstract**—Manufacturing process variations lead to variability in circuit delay and, if not accounted for, can cause excessive timing yield loss. The familiar traditional approaches to timing verification, such as the use of process corners and predefined timing margins, cannot readily handle within-die variations. Recently, statistical static timing analysis (SSTA) has been proposed as a way to deal with variability. Although many powerful techniques have been proposed, the fact that SSTA requires a significant change of methodology has delayed its wide adoption. In this paper, we propose a framework whereby the familiar concepts of corners and margins, which are generally meaningful at the transistor or cell level, are elevated to the chip level in order to handle within-die variations. This is achieved by using high-level models, such as the *generic path model* or the *generic circuit model* with different *classes of paths*, to represent the behavior of typical designs. These models allow us to determine “yield-specific” margins (setup and hold margins) and virtual corners, which, if applied during standard (deterministic) timing analysis, would guarantee the desired yield. Our framework can be used at an early stage of circuit design and is consistent with traditional timing verification methodology.

**Index Terms**—Classes of paths, early analysis, generic path, process variations, setup and hold margins, statistical timing analysis, timing yield, virtual corners.

## I. INTRODUCTION

PROCESS variations have an impact on circuit delay and can cause timing yield loss if circuit delay fails to meet the performance constraints. Traditionally, process variations have been taken care of in various ways. In microprocessors, it is typical to check circuit timing with *nominal* transistor files, and to specify some predefined *timing margin* which should be left as slack between the nominal delays and the timing thresholds, in order to account for process variations. In ASICs, the practice is to typically design circuits by making sure that the chip passes the timing requirements at all *process corners*, including nominal, worst, and best cases of device behavior. If these settings are too pessimistic, then designers are forced to waste time and effort optimizing a circuit using design conditions that are too stringent.

Manuscript received October 29, 2007; revised February 15, 2008 and May 6, 2008. Current version published September 19, 2008. This work was supported in part by Intel Corporation and in part by Altera Corporation. This paper was recommended by Associate Editor D. Sylvester.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: khaled@eecg.utoronto.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCAD.2008.2003291

There has been considerable discussion in the literature that the traditional methods of using process corners or using a timing margin are breaking down. For one, in microprocessors, where nominal process files are used and a timing margin needs to be left as slack, there are no easy ways to decide what margin should be used, particularly when one wants to account for within-die variations, which have become important in modern technology. In addition, if a given margin is used, there is no indication as to what the resulting timing yield would be. On the other hand, for ASICs, the number of corners is increasing, making it very expensive to explore all corners, and anyway, this traditional corner analysis approach cannot handle *within-die* statistical variations [1].

Statistical techniques offer one alternative approach to dealing with variability. Due to the increased importance of within-die variations, there has been an increased interest in tackling the timing yield problem by employing statistical techniques as part of the circuit timing analysis step [2]–[7]. The aim is to include statistical delay variations as an extension to traditional STA leading to statistical static timing analysis (SSTA). In early works [1], [3], [8], [9], within-die variations were assumed to be totally uncorrelated. This assumption is not true in practice; however, it is usually hard to express the correlations between within-die parameter variations with a model built from process data. Different attempts to model correlations have been proposed. In [4], principal component analysis (PCA) was used to decorrelate variations on a set of independent random variables (RVs). In [10], a quad-tree partitioning is used to express a regionwise spatial correlation among within-die variations. In [6], correlation is taken care of using a canonical model, where each variation is expressed in terms of global sources of variations. In most of the published works, variations are assumed to have a Gaussian (normal) distribution, and linear (first-order) delay models are used. Recently, however, several authors have investigated the case of non-Gaussian distributions and/or nonlinear delay models [11], [12]. Overall, SSTA has significant computational cost, and some have suggested [7] that some SSTA techniques unnecessarily complicate the design flow, and the point was made that it may be useful to explore early and simple SSTA techniques that would maintain the current timing verification methodology.

With regard to the design flow, most proposed SSTA techniques are applicable to fully specified and placed designs and require advanced correlation models built from extensive process data. Thus, these methods are “final sign-off” tools and are not usable during the circuit design phase. Also, requiring detailed process data can be problematic. For one, it may not be available during the early design phase of a new chip, which is

often being manufactured in a new process. In addition, there is no agreement yet on a simple and universal model of within-die correlation. One possibility is to assume that the correlation dies down with on-die distance according to some polynomial function. However, it is shown in [13] that distance is not the key determinant of correlation but that correlation depends on whether the distance in question is along a vertical or horizontal direction. There is also evidence that correlation depends on location within the reticle field, and not purely on distance. Finally, SSTA techniques require a change in the traditional timing verification methodology, which explains why these techniques are not yet widely adopted. Given all these reasons, it seems worthwhile to explore other options.

In this paper, we address the aforementioned challenges by concentrating on the following aspects:

- 1) focusing on *early analysis*, possibly in the absence of a complete circuit instance, which allows one to take corrective actions early in the design phase;
- 2) operating in the absence of within-die correlation information, for the reasons stated earlier, namely that correlation information can either be unknown for a new process or inaccurate;
- 3) keeping the current timing verification methodology which uses corners and timing margins in place, and extending it to handle within-die variations. This requires elevating the concepts of corners and margins to the circuit (i.e., chip) level whereby virtual corners and timing margins are deduced from circuit performance.

In order to enable early analysis, and given the absence of complete details of the circuit implementation, it is necessary to construct high-level models that capture circuit characteristics. A model was used in [14] in which a *generic critical path* model was employed as a way to capture the dependence of circuit worst-case delays in a given technology, in the absence of a specific circuit instance. In this model, a circuit is assumed to consist of a large number of identical paths, whose delay is representative of critical path delays in the target design. Each of these generic paths consists of a specific number of identical stages (a stage is a logic gate and the interconnect at its output). This model has been successfully employed in the industry, and was recently extended in [15], where it is shown that the statistics of a large number of such generic paths match very well the statistics of critical path delays in the actual circuit. In this paper, we make use of the generic path model, and extend it by proposing a flexible *generic circuit* model based on multiple *classes* of generic paths, with different depth. Such a representation can model virtually any circuit by simple discretization of its path delays into specified classes. We use these models to represent and analyze typical designs in a target technology.

As stated earlier, we will assume that within-die variations have some *unknown* correlation on the die. This does not mean that we ignore correlation. On the contrary, we prove that, by taking correlation to its extremes, we get useful (conservative) bounds on the timing yield. Assuming that variations are positively correlated (this is further discussed hereafter), we show

that the conservative case is achieved when the systematic<sup>1</sup> within-die variations of gate delay are assumed to be totally correlated ( $\rho = 1$ ) along a path, and the systematic within-die variations of path delay are assumed to be uncorrelated ( $\rho = 0$ ) across a block (i.e., a collection of paths).

The analysis hereafter will lead to closed-form expressions for the timing yield. Furthermore, the results will be two sided, i.e., will take into account yield loss due to both setup and hold time violations. The yield expressions allow one to extract “yield-specific” timing margins that should be used on top of the nominal maximum and minimum circuit delay, during deterministic timing analysis, in order to guarantee the desired yield. We refer to these margins as setup and hold timing margins. We also show how to extract virtual corners at the chip level, as opposed to traditional process corners which are extracted at the transistor or gate level. Since these virtual corners are based on *circuit performance*, they do take within-die statistical variations into account. Finally, we propose a method that allows “controlled” budgeting of yield loss between setup and hold violations. Our framework is therefore consistent with traditional timing verification methodology. It can be applied at an early stage of the design flow to predict which corners and margins to use in order to achieve a target yield. Early versions of this work have appeared in [16].

The rest of this paper is organized as follows. Sections II and III describe the generic path model and the mathematical framework pertaining to the formulation of the timing yield expressions. In Section IV, we show how we determine the yield-specific margins and virtual corners that should be used to guarantee a desired target yield. Sections V and VI describe the generic circuit model and the mathematical framework leading to the selection of setup and hold margins. Margin budgeting, as well as the control over yield loss, is presented in Section VII. We conclude in Section VIII.

## II. BACKGROUND: THE GENERIC PATH MODEL

In this section, we first present the parameter model which captures process variations, and then show how we construct the generic path model. The purpose is to provide some background based on the material from [15] and [16].

### A. Parameter Model

For a given circuit element or layout feature  $i$ , let  $X(i)$  be a zero-mean 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  $i$ . It is standard practice [17] to break up the variations into *die-to-die* and *within-die* components as follows:

$$X(i) = X_{dd} + X_{wd}(i). \quad (1)$$

<sup>1</sup>The adjective “systematic” has unfortunately acquired two different meanings in the literature. To some, systematic variations are those variations that are repeatable, nonrandom, and deterministically dependent on some design or layout feature. To others [14], systematic variations are variations that are random but are not independent, i.e., what one may refer to as *correlated* variations. There are ways to reconcile the two views, but in any case, this paper assumes the second viewpoint: Systematic variations are those portions of the total variations which are random and correlated.

The die-to-die component  $X_{dd}$  is an *independent*<sup>2</sup> zero-mean Gaussian RV that takes the same value for all instances of this element on a given die, irrespective of location. The within-die component  $X_{wd}(i)$  is a zero-mean Gaussian which can take different values for different instances of that element on the same die. This leads to the following relationship between the variances:

$$\sigma^2(i) = \sigma_{dd}^2 + \sigma_{wd}^2(i). \quad (2)$$

Then, the within-die component is further broken down into two components, namely, a *systematic* or somewhat correlated component and a “random” component

$$X_{wd}(i) = X_{wds}(i) + X_{wdr}(i) \quad (3)$$

where, for each  $i$ , the random component  $X_{wdr}(i)$  is an independent zero-mean Gaussian. A similar relationship follows for the variances

$$\sigma_{wd}^2(i) = \sigma_{wds}^2(i) + \sigma_{wdr}^2(i). \quad (4)$$

The within-die systematic component is expressed in the following way:

$$X_{wds}(i) = \sigma_{wds}(i)Z_{wds}(i) \quad (5)$$

where  $Z_{wds}(i)$ 's are the *correlated standard normal* RVs (mean zero and variance one).

Hence, our model for parameter variation  $X(i)$  consists of an independent zero-mean die-to-die component  $X_{dd}$  with variance  $\sigma_{dd}^2$ , a correlated systematic within-die component  $X_{wds}(i)$  with variance  $\sigma_{wds}^2(i)$ , and an independent random within-die component  $X_{wdr}(i)$  with variance  $\sigma_{wdr}^2(i)$ . In previous work, PCA is used to model within-die correlation by expressing variations in terms of global variables. In this paper, however, we do not require the use of any specific correlation model. Instead, we express the within-die systematic components as positively correlated RVs and derive bounds on the timing yield which hold irrespective of the correlation model used. We show next how gate delay variations and path delay variations can be modeled using the same parameter model defined previously.

## B. Gate Delay

In general, there is a nonlinear relationship between gate delay and transistor parameters. HSPICE simulations in an industrial 90-nm CMOS library reveal, however, that this non-linearity is not strong. Therefore, we will simply assume that gate delay is linearly dependent on the process and, hence, is Gaussian with mean that is equal to its nominal value.

For all the transistors within one logic gate, we assume that their channel length variations are captured with a *single* RV  $L(i)$ , and their threshold voltage variations are captured with a *single* RV  $V(i)$ . We also assume that  $L(i)$  and  $V(i)$  have been “decorrelated” using PCA [18], or any other transform,

<sup>2</sup>Throughout this paper, whenever an individual RV is described as “independent,” this means that it is independent of all other RVs under consideration.

so that they are independent. For each gate, we can extract sensitivities to the different varying process parameters using circuit simulation. Hence, if  $D(i)$  is the deviation of the delay of logic gate  $i$  from its mean (nominal) delay, we have

$$D(i) = \alpha L(i) + \beta V(i) \quad (6)$$

where  $\alpha$  and  $\beta$  are sensitivity parameters, with suitable units, that one can easily obtain from circuit simulation of a representative logic gate. As a result, we can express the statistical variations in delay of gate  $i$  as

$$D(i) = D_{dd} + D_{wds}(i) + D_{wdr}(i) \quad (7)$$

so that

$$\begin{aligned} D_{dd} &= \alpha L_{dd} + \beta V_{dd} \\ D_{wds}(i) &= \alpha L_{wds}(i) + \beta V_{wds}(i) \\ D_{wdr}(i) &= \alpha L_{wdr}(i) + \beta V_{wdr}(i) \\ \sigma_{dd,D}^2 &= \alpha^2 \sigma_{dd,L}^2 + \beta^2 \sigma_{dd,V}^2 \\ \sigma_{wds,D}^2(i) &= \alpha^2 \sigma_{wds,L}^2(i) + \beta^2 \sigma_{wds,V}^2(i) \\ \sigma_{wdr,D}^2(i) &= \alpha^2 \sigma_{wdr,L}^2(i) + \beta^2 \sigma_{wdr,V}^2(i). \end{aligned} \quad (8)$$

These equations provide a way in which the statistical model of gate delay (i.e., its three variances) can be computed from the underlying statistical model of transistor parameters.

## C. Generic Path Delay

Consider a generic path of  $N$  logic stages (a stage is a logic gate and the interconnect at its output). We will focus only on gate delays and on transistor  $L$  and  $V_t$  variations. The methodology can be easily applied when more device parameters are of interest or when interconnect parameter variations are to be included as well. Let  $D_N(j)$  denote the deviation of the delay of path  $j$  from its mean (nominal) value

$$D_N(j) = \sum_{i=1}^N D(i) = ND_{dd} + \sum_{i=1}^N D_{wds}(i) + \sum_{i=1}^N D_{wdr}(i). \quad (10)$$

Recall that we are working under the assumption that within-die systematic correlation is unknown, and thus, the systematic path delay component  $\sum_{i=1}^N D_{wds}(i)$  cannot be explicitly resolved as is. However, because we will be using upper and lower bounds in the yield analysis section where we study the statistics of a large collection of generic paths, we also use bounds here when deriving the delay model for these paths. In Appendix A, we prove that a lower (upper) bound on the distribution of a sum of Gaussian RVs with unknown correlation is achieved when these RVs are assumed to be totally correlated (independent). Applying this result to the systematic within-die component of path delay leads to the following equation for the variances:

$$\sigma_{D_N}^2(j) = \sigma_{dd,D_N}^2 + \sigma_{wds,D_N}^2(j) + \sigma_{wdr,D_N}^2(j) \quad (11)$$

where

$$\begin{aligned}\sigma_{\text{dd},D_N}^2 &= N^2\sigma_{\text{dd},D}^2 \\ \sigma_{\text{wds},D_N}^2(j) &= N^2\sigma_{\text{wds},D}^2(j), \quad \text{for lower bound} \\ \sigma_{\text{wds},D_N}^2(j) &= N\hat{\sigma}_{\text{wds},D}^2(j), \quad \text{for upper bound} \\ \sigma_{\text{wdr},D_N}^2(j) &= \sum_{i=1}^N \sigma_{\text{wdr},D}^2(i) = N\hat{\sigma}_{\text{wdr},D}^2(j)\end{aligned}$$

and where  $\hat{\sigma}_{\text{wdr},D}^2(j)$  and  $\hat{\sigma}_{\text{wds},D}^2(j)$  are the average values of  $\sigma_{\text{wdr},D}^2(i)$  and  $\sigma_{\text{wds},D}^2(i)$  over all  $N$  gates on this path. With this, we have a full statistical model of path delay, so that we can treat it as a “parameter” similarly to  $X(i)$ .

### III. YIELD ANALYSIS

With the random parameter model given previously, we now define the *parametric Max yield* and the *parametric Min yield* for parameter  $X$  as

$$Y_{\text{max}}(x) = \mathcal{P}\{X(i) \leq x, \quad i = 1, 2, \dots, n\} \quad (12)$$

$$Y_{\text{min}}(x) = \mathcal{P}\{X(i) > -x, \quad i = 1, 2, \dots, n\} \quad (13)$$

where  $n$  is the number of instances of this parameter on chip. Here,  $X(i)$  is a *generic* parameter that may represent process variations (channel length or threshold voltage), or path delay variations as we have shown in Section II-C, in which case the parametric yield becomes a *timing* yield. In other words, when  $X(i)$  represents path delay variation, the parametric Max yield is the probability that *all* path delay variations are less than  $x$ , and the parametric Min yield is the probability that all path delay variations are greater than  $-x$ . Note that  $x$  is generally positive since, in the Max yield case, we are interested in  $x$  values that are greater than the mean of  $X$  (i.e., zero), while in the Min yield case, we are interested in values that are smaller.

Thus, the material in this section, although focused on parametric yield, will actually be directly useful for computing timing yield. It is noteworthy that both Max and Min yields are of equal importance. Chips do fail because the maximum (minimum) circuit delay is greater (smaller) than a performance constraint determined by the setup (hold) constraints.

#### A. Slepian's Inequality

In this section, we will find bounds on the two parametric yields that were defined earlier. Using the following theorem from multivariate normal probability [19], we will prove that these bounds are a direct consequence of extreme cases of correlation between the within-die systematic components across the die.

1) *Slepian's Inequalities*: Let  $V = (\{V(i)\}, i = 1, \dots, n)$  and  $W = (\{W(i)\}, i = 1, \dots, n)$  be two random vectors of size  $n$ , with both of them being multivariately distributed with zero mean and covariance matrices  $\Sigma = \{\sigma_{ij}\}$  and  $\Gamma = \{\gamma_{ij}\}$ , respectively. Also, let  $\sigma_{ii} = \gamma_{ii}$  for  $i = 1, \dots, n$ . If  $\sigma_{ij} \geq \gamma_{ij}$  for all  $i \neq j$ , in symbols  $\Sigma \geq \Gamma$ , then

$$\mathcal{P}\{V(i) \leq a_i, \forall i\} \geq \mathcal{P}\{W(i) \leq a_i, \forall i\} \quad (14)$$

$$\mathcal{P}\{V(i) > a_i, \forall i\} \geq \mathcal{P}\{W(i) > a_i, \forall i\}. \quad (15)$$

Note that  $\sigma_{ii} = \sigma_i^2$  is the variance of  $V(i)$  and  $\sigma_{ij}$  is the covariance of  $V(i)$  and  $V(j)$ .

The aforesaid result is quite interesting. It states that, given two random vectors  $V$  and  $W$  having the same variances (i.e.,  $\sigma_{ii} = \gamma_{ii}$ ), if one of them is more correlated than the other (i.e.,  $\sigma_{ij} \geq \gamma_{ij}$  for all  $i \neq j$ ), then (14) and (15) hold.

A direct consequence of Slepian's inequalities is to be able to find upper and lower bounds on both Max and Min parametric yields. As presented in Section II-A, the parameter vector  $X$  is the following:

$$X(i) = X_{\text{dd}} + X_{\text{wds}}(i) + X_{\text{wdr}}(i) \quad (16)$$

where  $X_{\text{wds}}(i)$ 's are arbitrarily correlated across the die. To overcome the lack of information about within-die systematic correlation, which is typically hard to get, we will introduce two RVs that represent extreme cases of within-die systematic correlation: the case of independence and the case of total correlation.

Let  $X^{(0)}$  be a random vector whose elements have the same marginal distributions as those of  $X$  but with the property that its within-die systematic components are independent. Also, let  $X^{(1)}$  be a random vector whose elements have the same marginal distributions as those of  $X$  but with the property that its within-die systematic components are positively totally correlated. In other words,  $X$ ,  $X^{(0)}$ , and  $X^{(1)}$  have the same individual variances but only differ in the extent of correlation of the within-die systematic components, i.e., they differ in their covariances.

Now, suppose that  $\Sigma$ ,  $\Sigma^{(0)}$ , and  $\Sigma^{(1)}$  are the covariance matrices of parameter vectors  $X$ ,  $X^{(0)}$ , and  $X^{(1)}$ , respectively. These matrices have the same diagonal elements (representing variances) but differ by their off-diagonal elements (representing covariances). Remember that  $X$  is a zero-mean random vector, and its covariance matrix  $\Sigma = \{\sigma_{ij}\}$  will be simplified to the following, for  $i \neq j$ :

$$\begin{aligned}\sigma_{ij} &= E[X(i) \cdot X(j)] \\ &= \sigma_{\text{dd}}^2 + E[X_{\text{wds}}(i) \cdot X_{\text{wds}}(j)]\end{aligned} \quad (17)$$

where  $\sigma_{ij}$  is the off-diagonal element of the covariance matrix  $\Sigma$ . The second part of the aforementioned equation is the covariance  $\Sigma_{\text{wds}} = \{\sigma_{\text{wds}ij}\}$  of the within-die systematic component. Therefore, we can write

$$\sigma_{ij} = \sigma_{\text{dd}}^2 + \sigma_{\text{wds}ij}. \quad (18)$$

Notice here that the covariance of  $X^{(0)}$  and  $X^{(1)}$  will have the same form as the aforementioned equation and only differ by the value of  $\sigma_{\text{wds}ij}$ : For  $X^{(0)}$ ,  $\sigma_{\text{wds}ij} = 0$  since the within-die systematic components are independent (correlation coefficient equals zero). For  $X^{(1)}$ ,  $\sigma_{\text{wds}ij}$  is maximum, since maximum covariance occurs when the systematic RVs are positively totally correlated (correlation coefficient equals +1).

If we assume that the within-die systematic components are positively correlated (i.e., their covariances are always positive

$\sigma_{\text{wds}_{ij}} \geq 0$ ), then the correlation coefficients are between the two extremes of zero and one, so that

$$\Sigma^{(0)} \leq \Sigma \leq \Sigma^{(1)}. \quad (19)$$

Some comments are in order with respect to this positivity assumption. The assumption of positive correlation is practical for many sources of variability. A physical variation that slows down a transistor, a gate, or a path is likely to have the same effect on another that lies nearby. If the other device/gate/path is far away, then it would probably be independent anyway. In this paper, the RVs that will be assumed to be positively correlated are the gate delays in one case and the path delays in another case. Thus, we are *not* assuming that all process variables are positively correlated. By saying that two gates (or two paths) are positively correlated, we mean that when process variations cause the delay of one of them to increase, then the delay of the other does not decrease. Notice, however, that both gate and path delays are typically functions of *several* process variables. Therefore, this assumption is not as strong as requiring that corresponding sensitivities to the same process parameters be always of the same sign. To be sure, most commonly used process parameters, such as  $L$  and  $V_{\text{th}}$ , have similar effects on gate and path delay, e.g., an increase in channel length will almost always slow down a logic gate; such process variables lead to positive correlation. However, the fact that gate and path delay variations are an *aggregate* effect of dependence on a number of process variables is the best justification for our assumption; even if one or two pathological process variables, in advanced technology for instance, are such that they cause opposite effects, they are unlikely to dominate to such an extent so as to lead to overall negative correlation. Finally, in the case of path-to-path correlation, the fact that many paths share common subpaths would also contribute to overall positive correlation between them. In any case, given this assumption, then combining (19) with Slepian's inequalities, we can write

$$\begin{aligned} \mathcal{P}\{X^{(0)}(i) \leq x, \forall i\} &\leq Y_{\text{max}}(x) \leq \mathcal{P}\{X^{(1)}(i) \leq x, \forall i\} \\ \mathcal{P}\{X^{(0)}(i) > -x, \forall i\} &\leq Y_{\text{min}}(x) \leq \mathcal{P}\{X^{(1)}(i) > -x, \forall i\} \end{aligned} \quad (20)$$

where  $Y_{\text{max}}(x)$  and  $Y_{\text{min}}(x)$  are the parametric Max and Min yields defined earlier. The aforesaid equations are fundamental, because they capture the lower and upper bounds on the Max and Min yields when the within-die correlation is unknown or uncertain. In the following sections, we will derive expressions for the lower and upper bounds given in (20).

### B. Lower Bounds

Consistent with the previous section, we will now find the expressions for the Max and Min yield lower bounds by starting with an assumption that  $X_{\text{wds}}(i)$ 's are independent. Let  $Y_{\text{max}}^{(\text{lb})}(x)$  and  $Y_{\text{min}}^{(\text{lb})}(x)$  be the Max and Min yield lower bounds, respectively. For the lower bound analysis, the within-die systematic and random components  $X_{\text{wds}}(i)$  and  $X_{\text{wdr}}(i)$  are both independent RVs. Therefore, we can replace them both

by an independent within-die component  $X_{\text{wd}}(i) = X_{\text{wds}}(i) + X_{\text{wdr}}(i)$ . The following are the yield lower bound expressions:

$$Y_{\text{max}}^{(\text{lb})}(x) = \mathcal{P}\{X_{\text{dd}} + X_{\text{wd}}(i) \leq x, \forall i\} \quad (21)$$

$$Y_{\text{min}}^{(\text{lb})}(x) = \mathcal{P}\{X_{\text{dd}} + X_{\text{wd}}(i) > -x, \forall i\}. \quad (22)$$

Since  $X_{\text{dd}}$  is an independent zero-mean Gaussian with variance  $\sigma_{\text{dd}}^2$ ,  $Z_0 = X_{\text{dd}}/\sigma_{\text{dd}}$  is an independent standard normal RV (mean zero and variance one), and the expressions for the yield lower bounds can be expanded as

$$Y_{\text{max}}^{(\text{lb})}(x) = \mathcal{P}\{\sigma_{\text{dd}}Z_0 + X_{\text{wd}}(i) \leq x, \forall i\} \quad (23)$$

$$Y_{\text{min}}^{(\text{lb})}(x) = \mathcal{P}\{\sigma_{\text{dd}}Z_0 + X_{\text{wd}}(i) > -x, \forall i\}. \quad (24)$$

We now recall a result from 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)  $f(x)$ . Then (see [20, p. 85]), we have

$$\mathcal{P}\{\mathcal{A}\} = \int_{-\infty}^{+\infty} \mathcal{P}\{\mathcal{A}|X = x\}f(x)dx. \quad (25)$$

This result is an extension to the continuous case of the simple fact that  $\mathcal{P}\{\mathcal{A}\} = \mathcal{P}\{\mathcal{A}|\mathcal{B}\} \cdot \mathcal{P}\{\mathcal{B}\} + \mathcal{P}\{\mathcal{A}|\bar{\mathcal{B}}\} \cdot \mathcal{P}\{\bar{\mathcal{B}}\}$ , where  $\mathcal{B}$  is another event. Applying (25) to (23) and (24) and denoting by  $\phi(\cdot)$  the pdf of the standard normal distribution give

$$\begin{aligned} Y_{\text{max}}^{(\text{lb})}(x) &= \int_{-\infty}^{+\infty} \mathcal{P}\{X_{\text{wd}}(i) \leq x - \sigma_{\text{dd}}z, \forall i\} \phi(z)dz \\ Y_{\text{min}}^{(\text{lb})}(x) &= \int_{-\infty}^{+\infty} \mathcal{P}\{X_{\text{wd}}(i) > -x - \sigma_{\text{dd}}z, \forall i\} \phi(z)dz. \end{aligned} \quad (26)$$

Since  $X_{\text{wd}}(i)$ 's are independent, then we can express their joint probability as a product

$$\begin{aligned} Y_{\text{max}}^{(\text{lb})}(x) &= \int_{-\infty}^{+\infty} \prod_{i=1}^n \mathcal{P}\{X_{\text{wd}}(i) \leq x - \sigma_{\text{dd}}z\} \phi(z)dz \\ Y_{\text{min}}^{(\text{lb})}(x) &= \int_{-\infty}^{+\infty} \prod_{i=1}^n \mathcal{P}\{X_{\text{wd}}(i) > -x - \sigma_{\text{dd}}z\} \phi(z)dz. \end{aligned} \quad (27)$$

Now, replacing the integrals by the mean or expected value operator  $E[\cdot]$  and using the fact that, for a normal RV  $W$  with zero mean,  $\mathcal{P}\{W > -w\} = \mathcal{P}\{W \leq w\}$ , we get

$$Y_{\text{max}}^{(\text{lb})}(x) = E\left[\prod_{i=1}^n \Phi\left(\frac{x - \sigma_{\text{dd}}Z_0}{\sigma_{\text{wd}}(i)}\right)\right] \quad (28)$$

$$Y_{\text{min}}^{(\text{lb})}(x) = E\left[\prod_{i=1}^n \Phi\left(\frac{x + \sigma_{\text{dd}}Z_0}{\sigma_{\text{wd}}(i)}\right)\right] \quad (29)$$

where  $\sigma_{\text{wd}}^2(i) = \sigma_{\text{wds}}^2(i) + \sigma_{\text{wdr}}^2(i)$  is the variance of  $X_{\text{wd}}(i)$  and  $\Phi(\cdot)$  is the cumulative distribution function (cdf) of the standard normal.

Equations (28) and (29) represent the parametric Max and Min yield lower bounds. Note the dependence on the number of parameter instances  $n$ . Later in this paper, we will show that as  $n$  goes to infinity, the aforementioned equations will take other forms which are independent of  $n$ .

### C. Upper Bound

Similarly, we will now find expressions for upper bounds on the parametric Max and Min yields. Recall from Section III-A that when  $X_{\text{wds}}(i)$ 's are positively totally correlated, one gets an upper bound on the parametric yields. The following equations represent the upper bounds on the parametric yield, where  $X_{\text{wds}}(i)$ 's are totally correlated:

$$Y_{\text{max}}^{(\text{ub})}(x) = \mathcal{P} \{ X_{\text{wdr}}(i) + X_{\text{wds}}(i) + \sigma_{\text{dd}}Z_0 \leq x, \forall i \} \quad (30)$$

$$Y_{\text{min}}^{(\text{ub})}(x) = \mathcal{P} \{ X_{\text{wdr}}(i) + X_{\text{wds}}(i) + \sigma_{\text{dd}}Z_0 > -x, \forall i \}. \quad (31)$$

Total correlation of the within-die systematic components means that there is a unique RV that models all the systematic variations across the die. Hence, the model for  $X_{\text{wds}}(i)$  will be the following:

$$X_{\text{wds}}(i) = \sigma_{\text{wds}}(i)Z_1 \quad (32)$$

where  $Z_1$  is an independent standard normal. Making use of (25) twice on variables  $Z_0$  and  $Z_1$ , and noting that  $X_{\text{wdr}}(i)$ 's are independent with variance  $\sigma_{\text{wdr}}^2(i)$ , we get the following expressions:

$$Y_{\text{max}}^{(\text{ub})}(x) = E \left[ \prod_{i=1}^n \Phi \left( \frac{x - \sigma_{\text{dd}}Z_0 - \sigma_{\text{wds}}(i)Z_1}{\sigma_{\text{wdr}}(i)} \right) \right] \quad (33)$$

$$Y_{\text{min}}^{(\text{ub})}(x) = E \left[ \prod_{i=1}^n \Phi \left( \frac{x + \sigma_{\text{dd}}Z_0 + \sigma_{\text{wds}}(i)Z_1}{\sigma_{\text{wdr}}(i)} \right) \right]. \quad (34)$$

The same analysis used in the lower bound case is applied here to get the aforesaid equations. For the sake of conciseness, several steps were omitted.

As a result of Sections III-B and III-C, we have an upper and a lower bound on each of the parametric Max and Min yields. With a simple change of variables, as will be illustrated in the next section, these bounds can be computed by numerical integration. If a Max or Min yield of, for example, better than 90% is desired, then one can set the lower bound  $Y_{\text{max}}^{(\text{lb})}(x)$  or  $Y_{\text{min}}^{(\text{lb})}(x)$  to 0.9 and work backward to get the value of the threshold  $x$ . The upper and lower bounds are the results of extreme cases of within-die systematic correlations. Therefore, when combined together, they become very useful in estimating the yield when the correlations are uncertain or unknown.

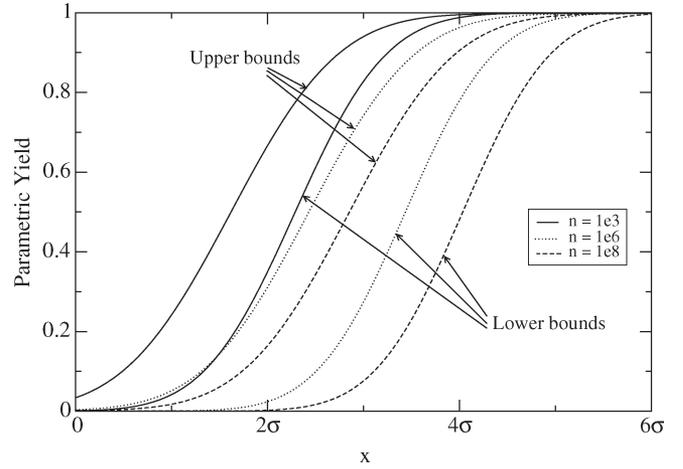


Fig. 1. Parametric max yield bounds.

### D. Illustration

As an illustration, we will assume all variances to be the same across the die. This means that  $\sigma_{\text{wds}}^2(i) = \sigma_{\text{wds}}^2$  and  $\sigma_{\text{wdr}}^2(i) = \sigma_{\text{wdr}}^2$ . Then, we can rewrite (28), (29), (33), and (34) in the following way:

$$Y_{\text{max}}^{(\text{lb})}(x) = E \left[ \Phi^n \left( \frac{x - \sigma_{\text{dd}}Z_0}{\sigma_{\text{wd}}} \right) \right] \quad (35)$$

$$Y_{\text{min}}^{(\text{lb})}(x) = E \left[ \Phi^n \left( \frac{x + \sigma_{\text{dd}}Z_0}{\sigma_{\text{wd}}} \right) \right] \quad (36)$$

$$Y_{\text{max}}^{(\text{ub})}(x) = E \left[ \Phi^n \left( \frac{x - \sigma_{\text{dd}}Z_0 - \sigma_{\text{wds}}Z_1}{\sigma_{\text{wdr}}} \right) \right] \quad (37)$$

$$Y_{\text{min}}^{(\text{ub})}(x) = E \left[ \Phi^n \left( \frac{x + \sigma_{\text{dd}}Z_0 + \sigma_{\text{wds}}Z_1}{\sigma_{\text{wdr}}} \right) \right] \quad (38)$$

where  $\sigma_{\text{wd}}$  was defined earlier to be  $\sqrt{\sigma_{\text{wds}}^2 + \sigma_{\text{wdr}}^2}$ .

In order to compute the aforementioned equations, we use the definition of the expected value operator (as an integral), and with a change of variables of  $u = \Phi(z_0)$  and  $v = \Phi(z_1)$ , we get

$$Y_{\text{max}}^{(\text{lb})}(x) = \int_0^1 \Phi^n \left( \frac{x - \sigma_{\text{dd}}\Phi^{-1}(u)}{\sigma_{\text{wd}}} \right) du$$

$$Y_{\text{min}}^{(\text{lb})}(x) = \int_0^1 \Phi^n \left( \frac{x + \sigma_{\text{dd}}\Phi^{-1}(u)}{\sigma_{\text{wd}}} \right) du$$

$$Y_{\text{max}}^{(\text{ub})}(x) = \int_0^1 \int_0^1 \Phi^n \left( \frac{x - \sigma_{\text{dd}}\Phi^{-1}(u) - \sigma_{\text{wds}}\Phi^{-1}(v)}{\sigma_{\text{wdr}}} \right) dudv$$

$$Y_{\text{min}}^{(\text{ub})}(x) = \int_0^1 \int_0^1 \Phi^n \left( \frac{x + \sigma_{\text{dd}}\Phi^{-1}(u) + \sigma_{\text{wds}}\Phi^{-1}(v)}{\sigma_{\text{wdr}}} \right) dudv. \quad (39)$$

The plots of the bounds on the parametric Max yield for different values of  $n$  are shown in Fig. 1, where we have assumed that  $\sigma_{\text{dd}}^2 = 0.5\sigma^2$  and  $\sigma_{\text{wds}}^2 = \sigma_{\text{wdr}}^2 = 0.25\sigma^2$ , with

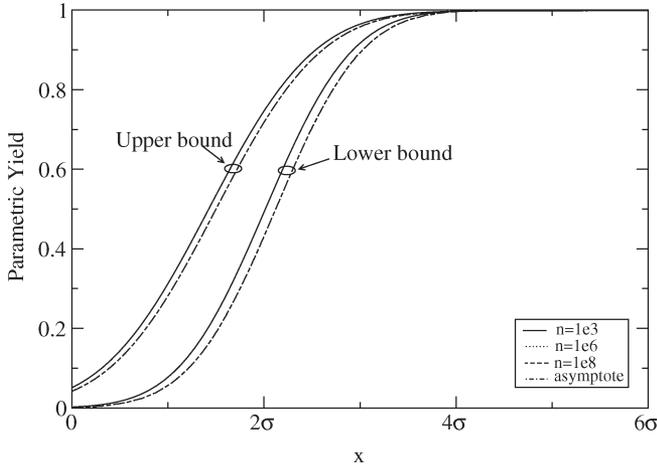


Fig. 2. Parametric max yield bounds for  $k = 3$ .

$\sigma^2$  being the total variance. Notice that the yield decreases for larger  $n$ , as expected. Also note that as  $x$  increases, it is more likely that the max of  $X(i)$  is less than  $x$ .

#### E. Bounded Variations

Notice that, in the previous expressions for the yield, the yield decreases for larger  $n$ . One would somewhat expect this, but it is surprising to note that the yield approaches zero as  $n$  goes to infinity for any combination of values of the three variances. This may also be seen in the plots of Fig. 1. This is somewhat nonphysical and arises due to the fact that we have assumed that the distribution of  $X_{\text{wdr}}(i)$  is normal; recall that the normal distribution extends to  $\pm\infty$  in both directions. In reality, one would expect process variations (and, consequently, delay variations) to be bounded. If a device somewhere deviates by large amounts, like  $6\sigma$  or  $7\sigma$ , then chances are that there is a serious problem with that die and that it would be lost due to other reasons other than timing yield. Therefore, it is a good idea to limit the spread of the cdf of  $X_{\text{wdr}}(i)$  to some multiple of  $\sigma$  in order to avoid these nonphysical effects at large  $n$ . In this section, therefore, we will use a *truncated normal* distribution for  $X_{\text{wdr}}(i)$ . For clarity of presentation, we will restrict the analysis to the illustrative special case introduced in Section III-D where all variances are the same across the die. The analysis can be extended to the general case. Suppose, therefore, that  $X_{\text{wdr}}(i)$  is bounded by  $\pm k\sigma$ , and let  $\Phi_t(x)$  represent the cdf of the truncated standard normal, which is zero for  $x \leq -k$  and one for  $x \geq k$ .

We can plug  $\Phi_t(\cdot)$  instead of  $\Phi(\cdot)$  (for  $X_{\text{wdr}}(i)$ ) into the aforementioned equations and plot the resulting Max and Min yield integrals, as shown in Fig. 2 for the Max yield. In this case, the yield loss at higher  $n$  values is limited so that the 1e6 and 1e8 plots in each group are indistinguishable. This is to be expected, because the “tail” of the distribution has been cut off, and it is primarily the tail that causes the yield loss at very large  $n$ .

When working with a truncated normal, it is noteworthy that we can derive asymptotes on the parametric yield curves that are independent of  $n$ . This means that as  $n$  tends to infinity, the parametric yield curves will be equal to these asymptotes. The

TABLE I  
REQUIRED SETTINGS FOR A YIELD LOWER BOUND

Yield Lower bound	Setup or Max Yield	Hold or Min Yield
Path delays	Sum over gates $\rho_{\text{wds}} = 1$	Sum over gates $\rho_{\text{wds}} = 1$
Block delay	Max over paths $\rho_{\text{wds}} = 0$	Min over path $\rho_{\text{wds}} = 0$

derivations are not shown, for space limitation, but lead to the following results, where  $Y_0(x) = \lim_{n \rightarrow \infty} Y(x)$ :

$$Y_{0\text{max}}^{(\text{lb})}(x) = \Phi\left(\frac{x - k\sigma_{\text{wd}}}{\sigma_{\text{dd}}}\right) \quad (40)$$

$$Y_{0\text{min}}^{(\text{lb})}(x) = \Phi\left(\frac{x - k\sigma_{\text{wd}}}{\sigma_{\text{dd}}}\right) \quad (41)$$

$$Y_{0\text{max}}^{(\text{ub})}(x) = E\left[\Phi\left(\frac{x - \sigma_{\text{dd}}Z_0 - k\sigma_{\text{wdr}}}{\sigma_{\text{wds}}}\right)\right] \quad (42)$$

$$Y_{0\text{min}}^{(\text{ub})}(x) = E\left[\Phi\left(\frac{x + \sigma_{\text{dd}}Z_0 - k\sigma_{\text{wdr}}}{\sigma_{\text{wds}}}\right)\right] \quad (43)$$

where, as before,  $Z_0 = X_{\text{dd}}/\sigma_{\text{dd}}$  is an independent standard normal RV. Equations (40) and (41), particularly when applied to timing yield as we will do later in this paper, exhibit similar observations as was made in [21]. Namely, the within-die variations determine the mean of the yield, while the die-to-die variations determine the spread of the yield. Fig. 2 shows the plot of these asymptotes for the Max yield case. These asymptotes are very tight and indistinguishable on the plot from the 1e6 and 1e8 curves.

In practice, one is quite interested in the yield *lower bound* expressions. Table I summarizes the conditions required to get a lower bound on the Max (Setup time) yield and the Min (Hold time) yield. It summarizes the results of this section, along with Section II-C, to show that, in the case of path delay, where gate delays are the “parameters,” in order to produce a lower bound on the yield, one must set the within-die systematic variations of gate delays to be totally correlated ( $\rho_{\text{wds}} = +1$ ). Looking at block delay, where delay is the maximum/minimum of a large number of paths, with path delays as the “parameters,” and in order to produce a lower bound on both Setup (Max) and Hold (Min) yields, one must set the within-die systematic variations of path delays to be independent ( $\rho_{\text{wds}} = 0$ ).

#### IV. APPLICATION: TIMING MARGINS AND VIRTUAL CORNERS

In the previous section, we have presented the mathematical framework leading to closed-form expressions for the parametric yield. We have also noted that when  $X(i)$ 's represent path delay variations, the parametric yield is essentially a timing yield, which is of more interest. As stated in Section I, two *dual* concepts are important for designers, namely, the concept of a *timing margin* to account for variations and/or the concept of a *corner* or device file setting, with which to run timing analysis. In this section, we show how we use the timing yield expressions in order to select, for a desired target yield, the corresponding timing margin and *virtual corner* that guarantee

that yield. Note that the corner is labeled as virtual because it is extracted from a design's performance, unlike traditional corners which are extracted by looking at transistors' or gates' performances.

Let  $N_1$  be the number of stages (gates) on a path that would be representative of long critical paths in a given design. We also consider that the design contains a large number of such *disjoint* (nonintersecting) critical paths of  $N_1$  gates. If one is interested in the maximum circuit delay, then the Max timing yield expression can be used

$$Y_{\max}(\tau_1) = \mathcal{P}\{D_{N_1}(j) \leq \tau_1, \forall j\} \geq Y_{0_{\max}}^{(lb)}(\tau_1) \quad (44)$$

where  $D_{N_1}(j)$  is the path delay variation described in Section II-C and  $Y_{0_{\max}}^{(lb)}(\cdot)$  is the lower bound expression for the yield found in Section III-E, with path delay being considered as the "parameter." Since the paths being considered are disjoint, then any correlations between their delays are only due to correlations in the process variations, and not to the sharing of circuit component. If  $\mathcal{Y}$  is the desired target Max yield, then the techniques of Section III effectively provide the inverse function to compute the margin  $\tau_1$  for any desired  $\mathcal{Y}$

$$\tau_1 = Y_{0_{\max}}^{(lb)-1}(\mathcal{Y}). \quad (45)$$

This  $\tau_1$  is the timing margin of an  $N_1$ -gate path for the desired specified Max yield  $\mathcal{Y}$ . Therefore, in order to get the desired yield, the circuit should be designed to "pass" the timing constraints when  $D_{N_1}(j) = \tau_1$  for all  $j$ . Now, to get the virtual corner, we need to move from the margin to the process (transistors). We use the same technique of Najm and Menezes [5], where the margin  $\tau_1$  is equally distributed first among the gates along the path, i.e.,  $D(i) = \tau_1/N_1$ , and then, using the gate delay model, we relate the transistor corner with the gate margin

$$\delta = \frac{\tau_1/N_1}{\alpha\sigma_L(i) + \beta\sigma_V(i)}. \quad (46)$$

This  $\delta$  effectively defines the device setting for which the circuit should be tested for timing constraint violations so as to guarantee that the timing Max yield is at least  $\mathcal{Y}$ . Similar analysis can be used for the Min yield case to verify the minimum circuit delay.

We now illustrate this with a numerical example. Assume that the variances of channel length and threshold voltage follow the following ratio:

$$(\sigma_{dd,L}^2, \sigma_{wds,L}^2, \sigma_{wdr,L}^2) = (0.5\sigma_L^2, 0.25\sigma_L^2, 0.25\sigma_L^2) \quad \forall i \quad (47)$$

$$(\sigma_{dd,V}^2, \sigma_{wds,V}^2, \sigma_{wdr,V}^2) = (0.5\sigma_V^2, 0.25\sigma_V^2, 0.25\sigma_V^2) \quad \forall i. \quad (48)$$

At the gate level,  $\sigma_D^2 = (\alpha^2\sigma_L^2 + \beta^2\sigma_V^2)$  and

$$(\sigma_{dd,D}^2, \sigma_{wds,D}^2, \sigma_{wdr,D}^2) = (0.5\sigma_D^2, 0.25\sigma_D^2, 0.25\sigma_D^2) \quad \forall i. \quad (49)$$

In addition, at the path level, we have

$$\sigma_{dd,D_{N_1}}^2 = 0.5N_1^2\sigma_D^2 \quad (50)$$

$$\sigma_{wds,D_{N_1}}^2 = 0.25N_1^2\sigma_D^2, \quad \text{for lower bound} \quad (51)$$

$$= 0.25N_1\sigma_D^2, \quad \text{for upper bound} \quad (52)$$

$$\sigma_{wdr,D_{N_1}}^2 = 0.25N_1\sigma_D^2. \quad (53)$$

The last two equations are notable for the absence of the square factor in  $N_1$ , since they result from averaging of variations due to independence. Let  $\sigma_{D_{N_1}}^2$  be the total path variance. If we want a Max yield that is greater than 95%, then we use the expression in (45) with  $\mathcal{Y} = 0.95$  and truncation factor  $k = 3$  to get the timing margin which is, in this case,  $\tau_1 \approx 3.1\sigma_{D_{N_1}}$ .

Then, using (46), we get

$$\delta = \left( \frac{3.1\sigma_{D_{N_1}}/N_1}{\alpha\sigma_L + \beta\sigma_V} \right). \quad (54)$$

Note that  $\sigma_{D_{N_1}} = N_1\sigma_D\sqrt{0.75 + 0.25N_1^{-1}}$  resulting from adding the path variances in the lower bound case, and assuming that  $N_1 = 9$ , we get

$$\delta = 2.73 \frac{\sqrt{\alpha^2\sigma_L^2 + \beta^2\sigma_V^2}}{\alpha\sigma_L + \beta\sigma_V}. \quad (55)$$

If  $r = (\alpha\sigma_L)/(\beta\sigma_V)$ , then

$$\delta = 2.73 \frac{\sqrt{1+r^2}}{1+r}. \quad (56)$$

If, for example,  $r = 1$ , then  $\delta \approx 1.93$ , so that the circuit would need to be simulated (and its timing checked), with all its transistors' channel lengths and threshold voltage being set at their  $+1.93\sigma$  points. Notice that, since  $\alpha$  and  $\beta$  depend on transistor sizing, (56) provides a way in which  $\delta$  can be controlled by circuit optimization and/or process tuning.

Note that  $\delta \approx 1.93$  is obtained from the lower bound which corresponds to the worst possible setting of correlation as discussed earlier. We have also used the upper bound on the yield (corresponding to best setting of correlation) to compute an optimistic virtual corner at a 95% yield, and it was found to be  $\delta \approx 1.43$ .

## V. GENERIC CIRCUIT MODEL

The generic path model for circuits is useful to study setup margins (Max yield) or hold margins (Min yield) separately. However, in order to study the interactions among long and short paths, and the tradeoffs between setup and hold margins, a more detailed model is required. In this section, we provide such a model, which we refer to as the *generic circuit* model, which involves the specification of different *classes* of paths, as an extension of the generic path model.

According to this model, a generic circuit is a collection of a certain number  $M$  of classes of paths. Each class  $j = 1, 2, \dots, M$  consists of  $n_j$  identical generic paths of *depth*  $N_j$ . The depth is the number of stages along a path as defined earlier; a stage is a logic gate or cell along with its fan-out interconnect network. In a logic circuit, if a logic path is

TABLE II  
GENERIC CIRCUIT WITH FIVE CLASSES OF PATHS

$j$	1	2	3	4	5
$N_j$	4	5	6	8	9
$D_{N_j}^{(nom)}$	8 (ns)	10 (ns)	11.1 (ns)	12 (ns)	13.5 (ns)
$\gamma_j$	30%	10%	10%	20%	30%
$n_j$	3000	1000	1000	2000	3000

well optimized, then every stage is loaded by a total output capacitance which is three to four times as big as its own input capacitance, hence the standard FO4 configuration, so that the stages on a path have approximately the same delay.

We use the same assumption as before that all stages have the same nominal delay and the same sensitivity to process variables ( $\alpha$  and  $\beta$ ). Thus, all stages are identical, and all paths within the same class are identical; paths in different classes have different depth.

The classes can be viewed as the result of discretization of path delays into “bins” of paths with equal or similar delay. We further assume the existence of a “circuit depth histogram” which provides the fraction of the total number of paths that each class constitutes. We express these fractions as probabilities

$$\mathcal{P}\{N = N_j\} = \gamma_j, \quad j = 1, 2, \dots, M \quad (57)$$

where  $N$  is a discrete RV that represents path depth,  $N_j$  is the depth of paths in class  $j$ , and  $\gamma_j$  is the fraction of total paths that have a depth of  $N_j$ . If  $n$  is the total number of paths in a circuit, then the number of paths  $n_j$  in class  $j$  is

$$n_j = \gamma_j \times n \quad \sum_{j=1}^M \gamma_j = 1. \quad (58)$$

An example is given in Table II, which, for a test circuit  $\mathcal{A}$  with five classes, lists the values of  $N_j$  and  $\gamma_j$  for every class. Also shown is the corresponding path count  $n_j$  in every class based on an assumed total path count of 10 000. The table also shows the nominal path delay  $D_{N_j}^{(nom)}$  in every class. Based on such a circuit model, we will provide the mathematical framework that leads to predicting timing margins that need to be left to guarantee a desired timing yield. We will also validate our margins with Monte Carlo simulations. We consider the generic circuit to be an intuitive and useful model for logic circuits in the same way as the generic path has been found to be useful [14], [15]. Recall that timing yield is the probability that circuit delay is within the specified timing constraints. These constraints can be upper limits on the maximum circuit delay, lower limits on the minimum circuit delay, or both in what is known as interval or two-sided constraints. Yield loss is incurred if circuit timing falls outside the constraints. Note that constraints on the maximum circuit delay are also known as setup time constraints, and constraints on the minimum delay are known as hold time constraints. In the next section, we define and express the setup yield and the hold yield in closed forms. Then, we combine the two analyses and present a margin budgeting scheme which allows to make a tradeoff between yield loss due to setup and hold violations.

## VI. SETUP AND HOLD YIELD ANALYSIS

Setup yield  $Y_S$  can be written in the following informal way:

$$Y_S = \mathcal{P}\{\text{All path delays are less than max constraint}\}. \quad (59)$$

Recall that, in our generic circuit representation with multiple classes, each class consists of identical generic paths whose delay variation is modeled using the generic parameter model. Let  $D_{N_j}(i)$  be the path delay variation of path  $i$  in class  $j$ . Then, we can express path delay variation using the parameter model in the following way:

$$D_{N_j}(i) = D_{N_j}^{(dd)} + D_{N_j}^{(wds)}(i) + D_{N_j}^{(wdr)}(i) \quad (60)$$

$$= D_{N_j}^{(dd)} + D_{N_j}^{(wd)}(i) \quad (61)$$

where  $j = 1, \dots, M$  is the class index,  $i = 1, \dots, n_j$  is the path index in class  $j$ , and the right-hand side is simply the three components of variability. Note that, in the second line, we have combined the systematic and random components into one within-die component of variability whose variance is equal to the sum of the variances of the systematic and random components. Now, combining the aforesaid model for path delay with (59) gives the formal expression for setup yield

$$Y_S = \mathcal{P} \left\{ \bigcap_{j=1}^M \left[ \bigcap_{i=1}^{n_j} (D_{N_j}(i) \leq \tau_j) \right] \right\} \quad (62)$$

where the  $\bigcap$  operator is used to indicate that we are interested in the probability of all these *joint* events.

In other words, the setup yield is the probability that, *over all* classes  $M$ , all path delay variations in each class are less than a timing margin  $\tau_j$  (specific to each class  $j$ ). This margin is the amount of “padding” or slack, which should be kept between the maximum delay constraint and the nominal path delay in class  $j$ , to account for delay variations. Using the previous expression for the setup yield, we will later show how we can predict a unique setup margin  $\tau_s$  that should be left as slack on the maximum nominal path delay in order to guarantee a desired setup yield for the circuit under consideration.

Starting from (62), and noting that the second intersection defines a statistical max operation, we can write the following:

$$Y_S = \mathcal{P} \left\{ \bigcap_{j=1}^M \left[ \max_{i=1}^{n_j} (D_{N_j}(i)) \leq \tau_j \right] \right\}. \quad (63)$$

In Section III-A, it was shown, using *Slepian's inequality*, that if the systematic within-die delay variations are assumed to be uncorrelated, then this would lead to a lower bound on the statistical max and min operations. We will use again this result here to write the following *lower bound* on the setup yield  $Y_S$ :

$$Y_S \geq \mathcal{P} \left\{ \bigcap_{j=1}^M \left[ \max_{i=1}^{n_j} (D_{N_j}^{(dd)} + D_{N_j}^{(wd)}(i)) \leq \tau_j \right] \right\} \quad (64)$$

where  $D_{N_j}^{(wd)}(i)$ 's are assumed to be uncorrelated for different  $i$ 's (hence independent because it is Gaussian).

### A. Asymptotic Convergence

Let  $Z_j$  be an RV that is equal to the maximum delay variation for class  $j$ . This is shown in the following:

$$Z_j = \max_{i=1}^{n_j} \left( D_{N_j}^{(dd)} + D_{N_j}^{(wd)}(i) \right) \quad (65)$$

$$= D_{N_j}^{(dd)} + \max_{i=1}^{n_j} \left( D_{N_j}^{(wd)}(i) \right) \quad (66)$$

$$= D_{N_j}^{(dd)} + W_j \quad (67)$$

where  $W_j = \max_{i=1}^{n_j} (D_{N_j}^{(wd)}(i))$ . Note that because the die-to-die component is the same for all  $i$ 's, then we can take it outside the max operation. We will now prove that as  $n_j \rightarrow \infty$ , the variance of  $W_j$  will go to zero; therefore, we can approximate  $W_j$  by a constant  $\mu_j$  that is equal to its 50th percentile  $W_{j,50\%}$ .

Let  $Y$  be the maximum of  $n$  independent identically distributed standard normal RVs. In studying the asymptotic properties of such an RV, it has been shown in [22] that  $V = l_n(Y - l_n)$  has a unique limiting distribution that is independent of  $n$ , where  $l_n$  is a coefficient (not an RV) that is proportional to  $\sqrt{2 \log n}$ . Since  $V$  has a distribution that is independent of  $n$ , its variance  $s^2$  is independent of  $n$ . Therefore

$$s^2 = \text{Var}[l_n(Y - l_n)] = l_n^2 \text{Var}[Y] \quad (68)$$

where  $\text{Var}[\cdot]$  denotes variance and

$$\text{Var}[Y] = \frac{s^2}{l_n^2} \rightarrow 0, \quad \text{as } n \rightarrow \infty \quad (69)$$

which is true because  $l_n^2$  is proportional to  $\log n$ .

This result can be applied to  $W_j$ , which is the maximum of  $n_j$  independent zero-mean Gaussian RV  $D_{N_j}^{(wd)}(i)$ , after scaling each by its standard deviation to transform them to standard normal RVs. Since the variance of  $W_j$  goes to zero for large  $n_j$ , we can approximate it by its 50th percentile  $W_{j,50\%}$  which we denote by  $\mu_j$ . Now, let  $\hat{Z}_j$  be an RV such that

$$\hat{Z}_j = D_{N_j}^{(dd)} + \mu_j. \quad (70)$$

Then,  $Z_j$  converges to  $\hat{Z}_j$  for large  $n_j$ . Note that  $\mu_j$  is given by

$$\mu_j = W_{j,50\%} = \Phi^{-1} \left( (0.5)^{\frac{1}{n_j}} \right) \times \sigma_{N_j}^{(wd)} \quad (71)$$

where  $\sigma_{N_j}^{(wd)}$  is the standard deviation of  $D_{N_j}^{(wd)}$ .

Fig. 3 shows the plots of the distributions of  $Z_j$  (solid) and  $\hat{Z}_j$  (dashed) for increasing values of  $n_j = 10, 100, 1e3, 1e6, 1e7,$  and  $1e9$ . The approximation is very tight for  $n_j$  that is greater than 100, and the two become indistinguishable for larger  $n_j$ . Since we anticipate each class of paths to contain hundreds, if not thousands, of paths, then the approximation is very good, and  $\hat{Z}_j$  will be used in lieu of  $Z_j$  in the analysis hereafter. Also note that the use of truncated normals accelerates the convergence and guarantees that  $\mu_j$  in (71) remains bounded.

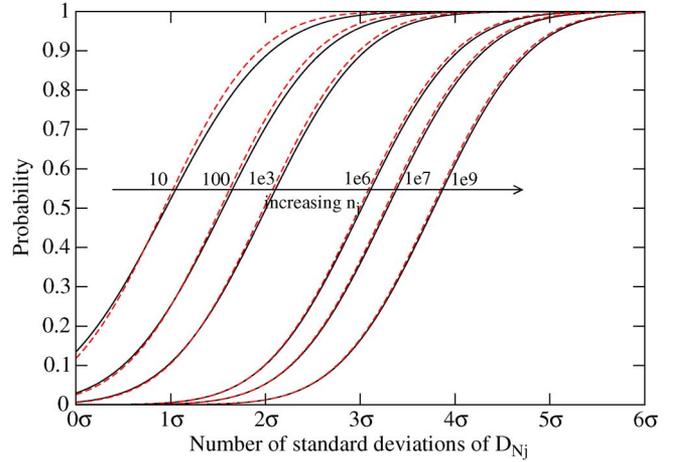


Fig. 3. Approximate and true distributions.

### B. Setup Yield Expression

Starting from (64), and replacing the expression inside the brackets by  $\hat{Z}_j$ , we get the following expression for the setup yield:

$$Y_S \geq \mathcal{P} \left\{ \bigcap_{j=1}^M \left( \hat{Z}_j \leq \tau_j \right) \right\} \quad (72)$$

$$= \mathcal{P} \left\{ \bigcap_{j=1}^M \left( D_{N_j}^{(dd)} \leq \tau_j - \mu_j \right) \right\}. \quad (73)$$

Recall that the die-to-die component of variation affects all the circuits in the same way, i.e., the variation is governed by a single RV

$$D_{N_j}^{(dd)} = \sigma_{N_j}^{(dd)} Z_0 \quad (74)$$

where  $\sigma_{N_j}^{(dd)}$  is the standard deviation of the die-to-die component and  $Z_0$  is a standard normal RV. Note that  $\sigma_{N_j}^{(dd)}$  depends on  $j$  and is different for different classes. Combining (73) and (74) gives the following expression for the setup yield:

$$Y_S \geq \mathcal{P} \left\{ \bigcap_{j=1}^M \left( Z_0 \leq \frac{\tau_j - \mu_j}{\sigma_{N_j}^{(dd)}} \right) \right\} \quad (75)$$

$$= \mathcal{P} \left\{ Z_0 \leq \min_{j=1}^M \left( \frac{\tau_j - \mu_j}{\sigma_{N_j}^{(dd)}} \right) \right\} \quad (76)$$

$$= \Phi \left( \min_{j=1}^M \left( \frac{\tau_j - \mu_j}{\sigma_{N_j}^{(dd)}} \right) \right) = Y_o \quad (77)$$

where  $Y_o$  is the desired lower bound on the setup yield  $Y_S$ .

### C. Setup Margin $\tau_s$

The aforementioned equation for the setup yield  $Y_S$  is very important because it will allow us to budget each margin  $\tau_j$  and to finally assign a unique setup margin  $\tau_s$  for the circuit,

which guarantees a desired setup yield as follows. Let  $a_{\min}$  be such that

$$a_{\min} = \min_{j=1}^M \left( \frac{\tau_j - \mu_j}{\sigma_{N_j}^{(dd)}} \right). \quad (78)$$

Recall that  $\mu_j$  and  $\sigma_{N_j}^{(dd)}$  are determined and fixed for each class, but  $\tau_j$  is to be determined. If a setup yield of, for example,  $Y_S = \mathcal{Y}$  is required, then we can set the lower bound  $Y_o$  in (77) to  $\mathcal{Y}$ , and get a value for  $a_{\min}$  as follows

$$a_{\min} = \Phi^{-1}(\mathcal{Y}). \quad (79)$$

We can then use this value of  $a_{\min}$  to budget the yield between the different classes. In order to reduce the design effort for class  $j$ , one would like the margin  $\tau_j$  to be small, and so we require

$$\frac{\tau_j - \mu_j}{\sigma_{N_j}^{(dd)}} = a_{\min}, \quad j = 1, \dots, M \quad (80)$$

where, effectively, we require the different class contributions to the yield to be equal. In this way, we have “relaxed” the yield requirements for each class while ensuring that the overall desired yield is met. As mentioned, this is important, since it has the benefit of reducing the design effort required to meet the timing and yield constraints.

This gives us the following expression for each class margin:

$$\tau_j = \sigma_{N_j}^{(dd)} \times a_{\min} + \mu_j, \quad j = 1, \dots, M. \quad (81)$$

Recall that the margin  $\tau_j$  represents how much designers should leave as slack between the nominal path delay of class  $j$  and the maximum delay constraint (setup constraint). Let  $D_{N_j}^{(nom)}$  be the nominal path delay of the generic paths of class  $j$ . Then, we can define a unique setup margin  $\tau_s$  to be padded on top of the maximum nominal path delay for all classes in order to guarantee that the desired setup yield is attained. This is expressed in the following way:

$$\tau_s = \max_{j=1}^M \left( D_{N_j}^{(nom)} + \tau_j \right) - \max_{j=1}^M \left( D_{N_j}^{(nom)} \right) \quad (82)$$

which may be explained with the help of Fig. 4. For each class with the nominal path delay of  $D_{N_j}^{(nom)}$  (hashed bars), we get the class margins  $\tau_j$  (solid bars) using (81). We then apply (82), subtracting the maximum nominal delay from the maximum delay with the added margin  $\tau_j$  to get the setup margin  $\tau_s$ . Therefore, we are providing a unique margin to be allowed on the nominal maximum circuit delay in order to guarantee the required yield. Note that no other path delays exceed the point defined by the maximum nominal delay with the added setup margin  $\tau_s$ .

#### D. Monte Carlo Validation

In this section, we validate our approach with Monte Carlo analysis based on an industrial 90-nm CMOS library. For this

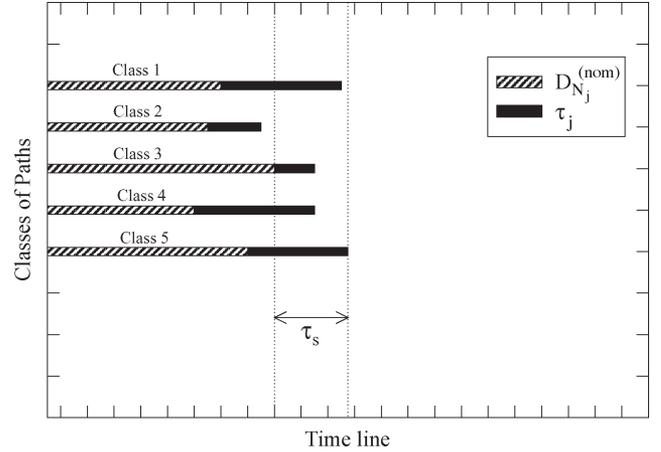


Fig. 4. Setup margin computation.

matter, we use circuit  $\mathcal{A}$  that was defined in Table II. This circuit has five classes of paths with different depths, nominal delay values, and numbers of paths. We assume that two process parameters are varying (channel length and threshold voltage), to which we determine gate delay sensitivities for every class of path. In our experiment, it was observed that the variation in process parameters resulted in 20%–27%  $3\sigma$  gate delay variation. Also, we have assumed a breakdown of (50%, 25%, and 25%) total parameter variance into die-to-die, systematic within-die, and random within-die variances, respectively, as was done earlier. Applying the analysis described in Section V, we can get the path delay model for each class of paths.

We assume that the desired setup yield is  $\mathcal{Y} = 99.73\%$ . Applying our analysis on circuit  $\mathcal{A}$ , without truncation of parameter variations, we get a setup margin  $\tau_s = 4.85$  ns that corresponds to 36% of the maximum nominal path delay of 13.5 ns found in Table II. Recall that this margin is valid for arbitrary within-die correlation since it was derived from the yield lower bound. It is therefore a conservative timing margin; this means that, for a given correlation setting, the true setup margin is expected to be less than the margin that we have predicted. We verify this statement through Monte Carlo simulations for different within-die correlations. To this end, we have considered a  $k \times k$  grid on which we place all the 10 000 paths of circuit  $\mathcal{A}$ . We also use a number ( $p$ ) of RVs to model the within-die correlations between paths. If  $p$  is large, within-die correlation between paths is weak. If  $p$  is small, correlation is strong. We have run Monte Carlo for 10 000 samples and generated the distribution of maximum path delay. From this distribution, we determined the 99.73% delay percentile and subtracted from it the maximum nominal path delay to get the setup margin. We have run this experiment for different values of  $p$  to emulate situations ranging from high correlation to almost independence. Monte Carlo analysis predicted setup margins ranging from 3.54 ns (for strong correlations) to 4.54 ns (for weak correlations) as opposed to our 4.85-ns predicted margin. Fig. 5 shows the predicted margins for the experiments that we performed. Note that the leftmost Monte Carlo margin corresponds to the case of strong correlations, and the rightmost Monte Carlo margin corresponds to the case of weak correlations.

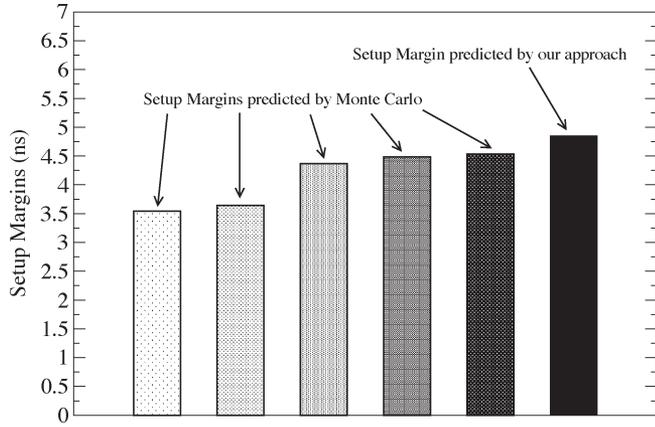


Fig. 5. Monte Carlo setup margins for 99.7% setup yield.

### E. Hold Yield

In the previous section, we presented in detail the analysis that led to the final expression for the setup yield  $Y_S$  and the corresponding setup margin  $\tau_s$ . In this section, we will briefly repeat the same analysis for the hold yield. Recall that the hold yield can be informally defined in the following way:

$$Y_H = \mathcal{P}\{\text{All path delays are greater than min constraint}\}. \quad (83)$$

Using the same analysis as before, we can rewrite the hold yield to be

$$Y_H = \mathcal{P} \left\{ \bigcap_{j=1}^M \left[ \bigcap_{i=1}^{n_j} (D_{N_j}(i) \geq -\tau'_j) \right] \right\} \quad (84)$$

where  $\tau'_j$  is the margin (positive) that should be left between the minimum constraint and the nominal path delay for class  $j$ . The aforementioned equation can be transformed in the same way as previously done

$$Y_H = \mathcal{P} \left\{ \bigcap_{j=1}^M \left[ \min_{i=1}^{n_j} (D_{N_j}(i)) \geq -\tau'_j \right] \right\} \quad (85)$$

$$= \mathcal{P} \left\{ \bigcap_{j=1}^M \left[ D_{N_j}^{(dd)} + \min_{i=1}^{n_j} (D_{N_j}^{(wd)}(i)) \geq -\tau'_j \right] \right\}. \quad (86)$$

Again, using Slepian's inequality, the assumption that  $D_{N_j}^{(wd)}(i)$ 's are independent will give us a lower bound on the hold yield. This allows us again to approximate the minimum of  $n_j$  independent Gaussian RVs by a constant  $\mu'_j$  that is equal to  $U_{j,50\%}$ , the 50th percentile of the distribution of  $U_j = \min_{i=1}^{n_j} (D_{N_j}^{(wd)}(i))$  [22]

$$\mu'_j = U_{j,50\%} = \Phi^{-1} \left( 1 - (0.5)^{\frac{1}{n_j}} \right) \times \sigma_{N_j}^{(wd)}. \quad (87)$$

The final expression for the hold yield is

$$Y_H \geq 1 - \mathcal{P} \left\{ Z_0 \leq \max_{j=1}^M \left( \frac{-\tau'_j - \mu'_j}{\sigma_{N_j}^{(dd)}} \right) \right\} \quad (88)$$

$$= 1 - \Phi \left( \max_{j=1}^M \left( \frac{-\tau'_j - \mu'_j}{\sigma_{N_j}^{(dd)}} \right) \right). \quad (89)$$

In the same way as before, we will choose a  $b_{\max}$  to guarantee a desired hold yield of, for example,  $Y_H \geq Y_o = \mathcal{Y}$

$$b_{\max} = \max_{j=1}^M \left( \frac{-\tau'_j - \mu'_j}{\sigma_{N_j}^{(dd)}} \right) = \Phi^{-1} (1 - \mathcal{Y}). \quad (90)$$

Then, we can budget  $\tau'_j$  in such a way that

$$\frac{-\tau'_j - \mu'_j}{\sigma_{N_j}^{(dd)}} = b_{\max}, \quad j = 1, \dots, M. \quad (91)$$

This gives us the margins that should be left between the minimum delay constraint (hold constraint) and the nominal path delay of every class of paths

$$\tau'_j = -\sigma_{N_j}^{(dd)} \times b_{\max} - \mu'_j, \quad j = 1, \dots, M. \quad (92)$$

Note here that  $\tau'_j$  is positive because  $b_{\max}$  and  $\mu'_j$  are both negative.

Finally, we combine the aforesaid margins with the values of nominal path delay  $D_{N_j}^{(nom)}$  to get a unique hold margin  $\tau_h$  for the circuit. This margin should be left as slack between the minimum constraint and the minimum nominal path delay over all classes

$$\tau_h = \min_{j=1}^M (D_{N_j}^{(nom)}) - \min_{j=1}^M (D_{N_j}^{(nom)} - \tau'_j). \quad (93)$$

Our hold analysis was validated using the same strategy presented in Section VI-D, and have found that Monte Carlo simulations predicted hold margins ranging from 1.56 ns to 2.05 (for various correlations) as opposed to 2.22 ns predicted by our approach.

## VII. APPLICATION: MARGIN BUDGETING

In the previous section, we presented a way to verify that a desired setup yield or hold yield is met. The method is simple, in that it provides a setup margin  $\tau_s$ , and a hold margin  $\tau_h$ , which should be allowed on the nominal maximum and minimum path delays, respectively, to guarantee the desired yields. We now combine the two analyses together, and give an expression for the total yield  $Y_T$ , which captures both setup and hold yield loss. This provides control over the amount of yield loss from either setup or hold violations. This means that if hold yield is dear, one can pick the margins in such a way

that *all* timing yield losses are due to setup yield loss and vice versa. The total yield is

$$Y_T = \mathcal{P} \left\{ \bigcap_{j=1}^M \left[ \left( D_{N_j}^{(dd)} + \min_{i=1}^{n_j} \left( D_{N_j}^{(wd)}(i) \right) \geq -\tau'_j \right) \cap \left( D_{N_j}^{(dd)} + \max_{i=1}^{n_j} \left( D_{N_j}^{(wd)}(i) \right) \leq \tau_j \right) \right] \right\}. \quad (94)$$

Again, a lower bound is guaranteed when the within-die variations are assumed to be uncorrelated. Then, by applying the approximations that were presented earlier, mainly replacing the minimum and maximum operations by their constant estimators gives the final expression for the total yield

$$Y_T \geq \mathcal{P} \left\{ \bigcap_{j=1}^M \left[ -\tau'_j - \mu'_j \leq D_{N_j}^{(dd)} \leq \tau_j - \mu_j \right] \right\} \quad (95)$$

$$= \mathcal{P} \{ b_{\max} \leq Z_0 \leq a_{\min} \} \quad (96)$$

$$= \Phi(a_{\min}) - \Phi(b_{\max}) \quad (97)$$

where  $\mu_j$ ,  $\mu'_j$ ,  $Z_0$ ,  $b_{\max}$ , and  $a_{\min}$  are as defined before. Note that we need to assign values for  $a_{\min}$  and  $b_{\max}$  in order to guarantee a desired total yield  $Y_T$ . Once we have chosen their values, the same previous analysis applies, i.e., find  $\tau_j$  and  $\tau'_j$  from  $a_{\min}$  and  $b_{\max}$  using (81) and (92), respectively, and then determine the setup margin  $\tau_s$  and the hold margin  $\tau_h$  using (82) and (93), respectively.

Looking at (97), it is easy to see that if a total yield of, for instance,  $Y_T \geq 95\%$  is desired, then different solutions  $(a_{\min}, b_{\max})$  exist. For each solution, the amount of yield loss from setup and hold violations differs, but the total yield loss is the same. Let  $L_T$ ,  $L_S$ , and  $L_H$  be the total, setup, and hold yield losses, respectively. Then, we can write

$$L_T = 1 - Y_T = 1 - \Phi(a_{\min}) + \Phi(b_{\max}) \quad (98)$$

$$L_S = 1 - Y_S = 1 - \Phi(a_{\min}) \quad (99)$$

$$L_H = 1 - Y_H = \Phi(b_{\max}). \quad (100)$$

Note that the total yield loss equals the sum of yield loss from setup and hold. We now define the “yield loss budget”  $\xi$  in  $[0, 1]$  to be the fraction of total yield loss induced by hold violations. This means that

$$L_H = \xi L_T \quad (101)$$

$$L_S = f(1 - \xi) L_T. \quad (102)$$

Combining the aforementioned four equations enables us to get a unique  $(a_{\min}, b_{\max})$  and, hence, a unique  $(\tau_h, \tau_s)$  solution for a specific yield loss budget  $\xi$

$$b_{\max} = \Phi^{-1}(\xi L_T) \quad (103)$$

$$a_{\min} = \Phi^{-1}(1 - (1 - \xi)L_T). \quad (104)$$

We have tested the previous analysis on circuit  $\mathcal{A}$  shown in Table II. For a total yield of  $Y_T = 95\%$  and a yield loss budget of  $\xi = 0.8$ , our results show that, without truncating the parameter distributions, the resulting margins are  $\tau_s = 4.44$  ns (33% of the maximum nominal delay of 13.5 ns) and  $\tau_h =$

TABLE III  
SETUP AND HOLD MARGINS FOR DIFFERENT TEST CIRCUITS

	Yield = 95%				Yield = 98%			
	$k = 3$		$k = \infty$		$k = 3$		$k = \infty$	
	$\tau_s$	$\tau_h$	$\tau_s$	$\tau_h$	$\tau_s$	$\tau_h$	$\tau_s$	$\tau_h$
$\mathcal{A}$	30%	21%	33%	23%	32%	22%	35%	25%
$\mathcal{B}$	23%	20%	26%	24%	25%	22%	28%	24%
$\mathcal{C}$	25%	20%	28%	23%	27%	23%	30%	25%

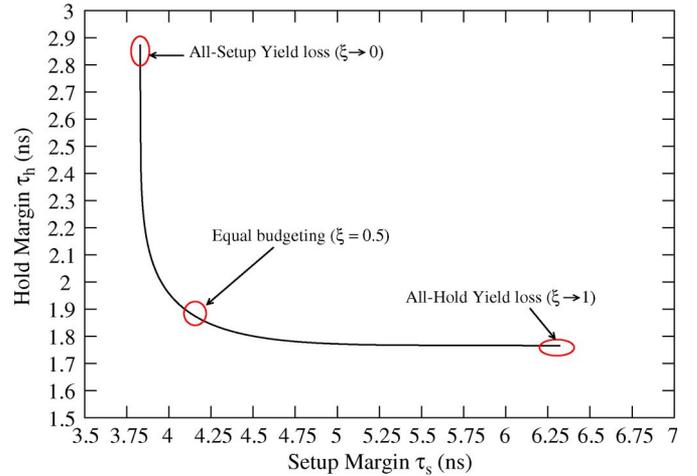


Fig. 6. Yield loss budgeting: Sweep over  $\xi \in (0, 1)$ ,  $Y_T = 95\%$ .

1.81 ns (23% of the minimum nominal delay of 8 ns). We have performed additional tests for different test circuits with varying number of classes of paths; the margins were derived at two yield targets, namely, 95% and 98%, and for two cases, namely, without truncation ( $k = \infty$ ) and with truncation ( $k = 3$ ), and a yield loss budget of  $\xi = 0.8$ . Our results are recorded in Table III, where setup and hold margins are normalized to the nominal maximum and minimum circuit delay, respectively.

Fig. 6 shows an interesting plot of  $\tau_h$  versus  $\tau_s$ , for circuit  $\mathcal{A}$  at a desired yield of 95%, which was obtained by sweeping the yield loss budget  $\xi$  between  $0^+$  and  $1^-$ . As expected, when the setup margin is increased (more slack), causing setup yield loss to decrease, we can reduce the hold margin and incur more hold yield loss while maintaining the same total yield. This gives designers more flexibility in their choice of margins.

## VIII. CONCLUSION

In this paper, an early technique for timing analysis was proposed, one that maintains the traditional methodology of using timing margins and corners. Our technique is based on a timing yield model which allows one to *determine* and *budget* the setup and hold margins required for a proposed chip design, as well as to elevate the concept of corners to the chip level by carefully selecting a “virtual corner” based on the desired circuit yield metric. The technique is valid irrespective of (and for any possible setting of) the within-die correlations, which allows one to handle situations of unknown or uncertain correlations. The process requires up-front specification of the target process technology and the circuit style, through a *generic circuit description* in terms of one or more *classes of generic paths*. As a result, the proposed technique provides setup and hold timing margins which, if observed for the longest and shortest nominal path delays during subsequent circuit design,

would allow the chip to meet the timing yield target. These margins are not unique, but may be traded off against each other, so that one may allow more setup yield loss than hold yield loss, or vice versa, by simply changing a single parameter.

#### APPENDIX BOUNDS ON SUM OF GAUSSIAN RVs

In this section, we will find bounds on the distribution of the sum of  $N$  normally distributed RV  $X(i)$ 's with unknown correlation. These bounds will turn out useful when we build the generic path variational delay model. Essentially, we prove that if the unknown correlation ranges from zero (independence) to one (total correlation), then a lower bound on the distribution of the sum is achieved when  $X(i)$ 's are assumed to be totally correlated, and an upper bound is achieved when  $X(i)$ 's are assumed to be independent.

Let  $S = \sum_{i=1}^N X(i)$ , where  $S$  is a zero-mean normal RV with some variance; call it  $\sigma_s^2$ . The distribution of  $S$  is as follows:

$$\mathcal{P}\{S \leq x\} = \Phi(x/\sigma_s) \quad (105)$$

where  $\Phi(\cdot)$  is the standard normal cdf. Since  $\Phi(\cdot)$  is increasing, then  $\mathcal{P}\{S \leq x\}$  is decreasing in  $\sigma_s$  (for  $x \geq 0$ , which is the range of interest if one is looking to achieve high yield). The variance of  $S$  is a function of the variances of  $X(i)$ 's and the unknown correlation coefficient  $\rho_{ij}$

$$\sigma_s^2 = \sum_{i=1}^N \sigma_{X_i}^2 + 2 \sum_{i \neq j} \rho_{ij} \sigma_{X_i} \sigma_{X_j}. \quad (106)$$

Therefore, it is easy to see that a lower (or upper) bound on the distribution of  $S$  is achieved when  $\sigma_s$  is maximized (or minimized), i.e., when  $\rho_{ij}$  are set to one (or zero).

#### REFERENCES

- [1] 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.
- [2] J. Jess, K. Kalafala, W. Naidu, R. Otten, and C. Visweswariah, "Statistical timing for parametric yield prediction of digital integrated circuits," in *Proc. DAC*, Anaheim, CA, Jun. 2–6, 2003, pp. 932–937.
- [3] A. Agarwal, D. Blaauw, V. Zolotov, and S. Vrudhula, "Computation and refinement of statistical bounds on circuit delay," in *Proc. Des. Autom. Conf.*, Anaheim, CA, Jun. 2–6, 2003, pp. 348–353.
- [4] H. Chang and S. S. Sapatnekar, "Statistical timing analysis considering spatial correlations using a single PERT-like traversal," in *Proc. ICCAD*, San Jose, CA, Nov. 9–13, 2003, pp. 621–625.
- [5] F. N. Najm and N. Menezes, "Statistical timing analysis based on a timing yield model," in *Proc. Des. Autom. Conf.*, San Diego, CA, Jun. 7–11, 2004, pp. 460–465.
- [6] C. Visweswariah, K. Ravindran, K. Kalafala, S. Walker, and S. Narayan, "First-order incremental block-based statistical timing analysis," in *Proc. DAC*, San Diego, CA, Jun. 7–11, 2004, pp. 331–336.
- [7] C. Amin, N. Menezes, K. Killpack, F. Dartu, U. Choudhury, N. Hakim, and Y. Ismail, "Statistical static timing analysis: How simple can we get?" in *Proc. Des. Autom. Conf.*, Jun. 13–17, 2005, pp. 652–657.
- [8] 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.
- [9] S. Bhardwaj, S. B. K. Vrudhula, and D. Blaauw, " $\tau$ -tau: Timing analysis under uncertainty," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Des.*, San Jose, CA, Nov. 9–13, 2003, pp. 615–620.
- [10] 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.*, Nov. 9–13, 2003, pp. 900–907.
- [11] Y. Zhan, A. J. Strojwas, X. Li, and L. T. Pileggi, "Correlation-aware statistical timing analysis with non-Gaussian delay distributions," in *Proc. DAC*, Anaheim, CA, Jun. 13–17, 2005, pp. 77–82.
- [12] H. Chang, V. Zolotov, S. Narayan, and C. Visweswariah, "Parameterized block-based statistical timing analysis with non-Gaussian parameters, nonlinear delay functions," in *Proc. DAC*, Anaheim, CA, Jun. 13–17, 2004, pp. 71–76.
- [13] B. Cline, K. Chopra, D. Blaauw, and Y. Cao, "Analysis and modeling of CD variation for statistical static timing," in *Proc. ICCAD*, San Jose, CA, Nov. 5–9, 2006, pp. 60–66.
- [14] K. A. Bowman, S. G. Duvall, and J. D. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration," *IEEE J. Solid-State Circuits*, vol. 37, no. 2, pp. 183–190, Feb. 2002.
- [15] F. N. Najm, N. Menezes, and I. A. Ferzli, "A yield model for integrated circuits and its application to statistical timing analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 26, no. 3, pp. 574–591, Mar. 2007.
- [16] K. R. Heloue and F. N. Najm, "Statistical timing analysis with two-sided constraints," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Des.*, San Jose, CA, Nov. 6–10, 2005, pp. 829–836.
- [17] S. G. Duvall, "Statistical circuit modeling and optimization," in *Proc. Int. Workshop Statistical Metrol.*, Jun. 2000, pp. 56–63.
- [18] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.
- [19] Y. L. Tong, *The Multivariate Normal Distribution*. New York: Springer-Verlag, 1990.
- [20] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.
- [21] 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.*, 2001, pp. 229–232.
- [22] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed. New York: Wiley-Interscience, 2003.



**Khaled R. Heloue** (S'01) received the B.E. degree in computer and communications engineering (with high distinction) from the American University of Beirut, Beirut, Lebanon, in 2003 and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2005, where he is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering.

His research includes computer-aided design (CAD) for integrated circuits, with an emphasis on timing analysis under variability, yield modeling and estimation, design for manufacturability, and full-chip leakage estimation. He interned with Intel's Strategic CAD Labs, Hillsboro, OR, in the summer of 2008.



**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, Beirut, Lebanon, in 1983 and the Ph.D. degree in electrical and computer engineering (ECE) from the University of Illinois, Urbana (UIUC), in 1989.

From 1989 to 1992, he was with Texas Instruments, Dallas. He was then with the ECE Department, UIUC, as an Assistant Professor and became an Associate Professor in 1997. Since 1999, he has been with the ECE Department, University of Toronto, Toronto, ON, Canada, where he is currently a Professor. His research includes computer-aided design (CAD) for very large scale integration (VLSI), with an emphasis on circuit-level issues related to power, timing, variability, and reliability.

Dr. Najm received the IEEE TRANSACTIONS ON CAD Best Paper Award in 1992, the NSF Research Initiation Award in 1993, and the NSF CAREER Award in 1996. He was an Associate Editor for the IEEE TRANSACTIONS ON VLSI SYSTEMS from 1997 to 2002 and is currently an Associate Editor for the IEEE TRANSACTIONS ON CAD OF INTEGRATED CIRCUITS AND SYSTEMS.