

A New Statistical Max Operation for Propagating Skewness in Statistical Timing Analysis

Kaviraj Chopra

Bo Zhai

David Blaauw

Dennis Sylvester

Department of EECS
University of Michigan
Ann Arbor, MI

ABSTRACT

Statistical static timing analysis (SSTA) is emerging as a solution for predicting the timing characteristics of digital circuits under process variability. For computing the statistical max of two arrival time probability distributions, existing analytical SSTA approaches use the results given by Clark in [8]. These analytical results are exact when the two operand arrival time distributions have jointly Gaussian distributions. Due to the nonlinear max operation, arrival time distributions are typically skewed. Furthermore, nonlinear dependence of gate delays and non-gaussian process parameters also make the arrival time distributions asymmetric. Therefore, for computing the max accurately, a new approach is required that accounts for the inherent skewness in arrival time distributions. In this work, we present analytical solution for computing the statistical max operation.¹ First, the skewness in arrival time distribution is modeled by matching its first three moments to a so-called skewed normal distribution. Then by extending Clark's work to handle skewed normal distributions we derive analytical expressions for computing the moments of the max. We then show using initial simulations results that using a skewness based max operation has a significant potential to improve the accuracy of the statistical max operation in SSTA while retaining its computational efficiency.

1. INTRODUCTION

Process control precision is worsening with continuous process scaling due to smaller dimensions, smaller number of doping atoms and aggressive lithographic techniques. This results in an increase in process parameters fluctuations, that causes variations in electrical characteristics of transistors and interconnects. These variations in electrical characteristics of circuit components affect timing and result in chip operating frequency variation. Traditionally corner based static timing analysis have been used to guard against yield loss resulting from these variations; however, with increasing number of sources of variation, corner based methods are becoming overly pessimistic and computationally expensive.

¹This work was supported in part by the National Science Foundation (NSF), under Grant CCR-0205227.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICCAD'06, November 5-9, 2006, San Jose, CA

Copyright 2006 ACM 1-59593-389-1/06/0011 ...\$5.00.

An alternative approach, namely, statistical static timing analysis (SSTA) has emerged as a possible solution for statistically quantifying the variability in timing performance. Existing SSTA approaches can be broadly classified into block based SSTA [13, 4, 1, 9, 6, 17] and path based SSTA [14, 2]. A path-based SSTA requires enumeration of an exponential number of paths, therefore, block-based SSTA is considered to be a more efficient technique. Among these, the analytical methods, presented in [3, 6, 17], appeal to be the more promising approaches for a computationally efficient implementation of SSTA. In [3], the author introduced a linear time analytical SSTA algorithm assuming uncorrelated normal random variables for delay distribution. Using a first order parametric delay model, a method for handling correlations in global sources of variation due to both spatial correlation and path re-convergence was presented in [6, 17]. Their SSTA algorithm included a PERT-like topological traversal of a circuit graph, where at each node the maximum arrival time distribution is computed in terms of the parametric delay model. For propagating arrival time distributions, one needs to compute the sum and the maximum of two arrival time at each node in the circuit graph. The computation of the sum function is relatively simple; however, the statistical max of two correlated arrival time variables is non-trivial.

The max operation in existing SSTA approaches is invariably based on analytical results given in [8]. Clark derived analytical expressions for finding the moments of the max of two correlated normal random variables and an expression for computing the correlation of the resulting max with any other jointly normal variable. The Clark's max results are exact when the two operand random variables have a bivariate normal distribution. However, the result of the max of two normal variables is typically a positively skewed non normal distribution. Skewness is a statistical parameter used to describe asymmetry in a random variables probability distribution. A probability distribution is said to have positive(negative) skewness if it has a long tail in the positive(negative) direction (see Figure 1). Both the above mentioned analytical approaches [6, 17], use these expressions for computing the moments of statistical max of two arrival time random variables. Unfortunately in SSTA, the asymmetric non-normal arrival time distributions resulting from the max operation performed at one node are inputs to the max operation which is needed to be performed at a downstream node. Additionally, variations in interconnect and few process parameters also have asymmetric non-normal distributions [19]. However, existing analytical SSTA approaches have to approximate the non-normal arrival time distribution with a normal distribution for applying Clark's max. The error of this approximation increases when the difference of the mean relative to the standard deviation decreases and it becomes maximum when two means are equal [8]. For a typical design, there can be several thousand critical paths and the

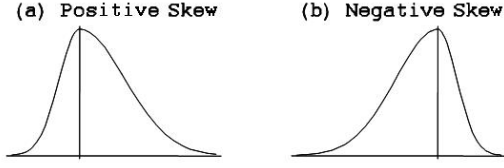


Figure 1: Examples for Asymmetric PDFs

means of their output arrival time distributions and arrival time distributions at common internal nodes will be closely aligned with each other. Therefore, in such a case Clark's max based SSTA methods may result in inadequate accuracy, in particular, for power optimized designs having a large number of nodes with zero or small slack. Recently, SSTA algorithms using higher order nonlinear parametric delay models with non gaussian distributions were proposed in [7, 12, 18, 19]. However, for computing the max operation, these approaches either use numerical techniques and/or employ the Clark's max requiring normal approximation. A conditional max based heuristic analytical method was presented in [19] where the max operations is postponed until the two arrival time distributions are skewed.

In this work, we extend Clark's max approach and give analytical results for computing the approximate maximum of a set of asymmetric random variables. The problem of computing the max of a finite set of random variables has been well studied. Several approaches derived Clark's results using different methods [11, 5]. In our method, given the first three moments of any asymmetric distribution, we give analytical expressions to map it to a skewed normal (explained later) representation having same moments. We then derive analytical results for computing the moments of the max of two correlated skewed normal distributions assuming a bivariate skew normal distribution. The derivation is similar in spirit to Clark's approach, although it is more general since we can compute the moments for a bivariate skewed normal random variables.

The rest of the paper is organized as follows. In Section 2 we explain the skewed normal distribution and give analytical expressions for computing the parameters of a skewed normal distribution from the mean, variance and skewness of arrival time distribution. A bivariate skewed normal distribution and the derivation for the proposed max operation are given in Section 3. In Section 4, we give numerical results illustrating the efficacy of the proposed max operation. Section 5 concludes the paper.

2. MODELING SKEWNESS

Arrival time distributions and circuit delay distributions are typically skewed, due to the nonlinear max operation and nonlinear dependence of delay on process parameters. We need an analytical representation that is flexible enough to capture the skewness in asymmetric arrival time distributions and at the same time be of the functional form which allows analytical derivation of the maximum operation. After studying several skewed representations, in [10], we found a general method for introducing skewness into any unimodal symmetric distribution. Their basic idea is to simply introduce inverse scale factors in the left and the right half real lines around the mean. Let $f(x)$ be the normal distribution with mean μ and variance σ given by

$$f(x) = \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right), \text{ where } \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Using the method presented in [10], a skewed normal distribution $f_\gamma(x)$ can be computed from the normal distribution $f(x)$, by scaling its left half and right half by factors γ and its inverse $1/\gamma$, respectively. This gives us the skewed normal distribution,

$$f_\gamma(x) = \frac{2}{\sigma(\gamma + 1/\gamma)} \left\{ \phi\left(\frac{(x-\mu)\gamma}{\sigma}\right) I_{(-\infty, \mu]}(x) + \phi\left(\frac{x-\mu}{\gamma\sigma}\right) I_{(\mu, \infty)}(x) \right\},$$

where, $I_{(-\infty, \mu]}(x)$ and $I_{(\mu, \infty)}(x)$ are the Indicator functions:

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise.} \end{cases}$$

For a skewed normal distribution, we can observe that scaling variable x corresponds to an inverse scaling of the standard deviation σ around its mean. Therefore, $f_\gamma(x)$ can be alternatively written as

$$f_\gamma(x) = \frac{2}{\sigma_l + \sigma_r} \left\{ \phi\left(\frac{(x-\mu)}{\sigma_l}\right) I_{(-\infty, \mu]}(x) + \phi\left(\frac{x-\mu}{\sigma_r}\right) I_{(\mu, \infty)}(x) \right\},$$

where,

$$\sigma_l = \frac{\sigma}{\gamma} \quad \text{and} \quad \sigma_r = \sigma\gamma.$$

Note that the resulting skewed distribution $f_\gamma(x)$ has a functional form similar to the original non-skewed distribution $f(x)$. If the skewness parameter γ is greater(less) than unity then $f_\gamma(x)$ is positively(negatively) skewed. For $\gamma = 1$ we get back the original symmetric normal distribution. Furthermore, $f_\gamma(x)$ is both continuous and differentiable and is completely defined by only three parameters μ , σ and γ . These were the key appealing properties that motivated us to use this representation for deriving the proposed max operation.

Existing SSTA approaches model and propagate only the mean and variance of the arrival time distribution. For improving the accuracy of SSTA algorithm, in addition to the mean and variance, we wish to propagate the skewness in asymmetric arrival time distributions. In such an SSTA framework, the input parameters of the max operation will include mean, variance and skewness of the two input arrival time distributions and their correlation. We first want to map the arrival time distribution characterized by its mean, variance and skewness to a skewed normal distribution $f_\gamma(x)$. Let μ_γ , σ_γ and Sk_γ be the given mean, standard deviation and skewness of a skewed arrival time distribution and μ , σ and γ are the three parameters that define the desired skewed normal distribution $f_\gamma(x)$. For finding $f_\gamma(x)$, we express the mean, variance and skewness of the skewed normal distribution as function of its parameters μ , σ and γ and then match these to the μ_γ , σ_γ and Sk_γ of a skewed arrival time distribution to solve for μ , σ and γ . The analytical expressions for mean μ_γ , variance σ_γ^2 and skewness Sk_γ of $f_\gamma(x)$ derived in terms of its parameters (μ , σ and γ) are given as follows:

$$\mu_\gamma = \mu + \sqrt{\frac{2}{\pi}} \left(\gamma - \frac{1}{\gamma} \right) \sigma \quad (1)$$

$$\sigma_\gamma^2 = \frac{(\pi\gamma^4 - 2\gamma^4 - \pi\gamma^2 + 4\gamma^2 + \pi - 2) \sigma^2}{\pi\gamma^2} \quad (2)$$

The skewness of distribution defined by the ratio of the third centered moment and cubed standard deviation is given by

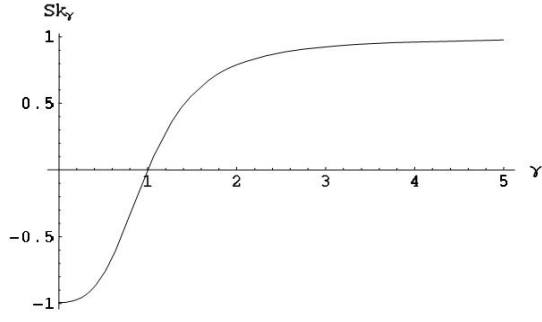


Figure 2: The γ parameter of $f_\gamma(x)$ vs. Skewness Sk_γ

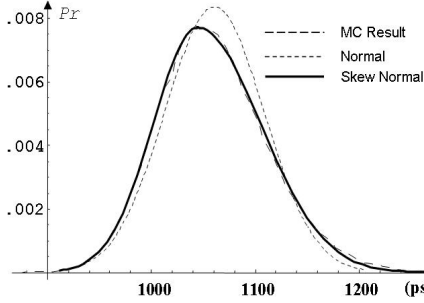


Figure 3: Comparison between Skewed Normal distribution and Normal distribution for a typical Monte Carlo based Arrival Time distribution.

$$Sk_\gamma = \frac{\sqrt{2} (1 - \gamma^2) (\pi (\gamma^4 - 3\gamma^2 + 1) - 4(\gamma^2 - 1)^2)}{(\pi (\gamma^4 - \gamma^2 + 1) - 2(\gamma^2 - 1)^2)^{\frac{3}{2}}} \quad (3)$$

Fortunately, the skewness Sk_γ (Eq. 3) is only a function of γ and is independent of the other two parameters μ and σ . A plot of this function is given in Figure 2, where it can be seen that skewness Sk_γ is a well behaved function and it monotonically increases with γ . Therefore, for a given Sk_γ , one can efficiently compute γ either using pre-computed look-up tables or using numerical methods with very fast convergence. Using γ , σ_γ and μ_γ we can analytically solve equations 2 and 1 for parameters σ and μ , respectively. Thus given mean, variance and skewness of an arrival time distribution we can easily map it to a skewed normal distribution. In Figure 3, we show plots of a typical skewed arrival time distribution approximated by a skewed normal distribution and normal distribution. It is evident that compared to existing normal approximations, skewed normal is a much better representation that can accurately capture the inherent skewness in arrival time distributions.

3. SKEWED NORMAL MAX OPERATION

Based on the skewed normal representation explained in the previous section, we now present the skewed normal max operation. For analytically expressing the max function of two correlated arrival time random variables X and Y , we need to know their joint probability distribution function. In [8], the author uses the following bivariate normal distribution for the two operand random variables.

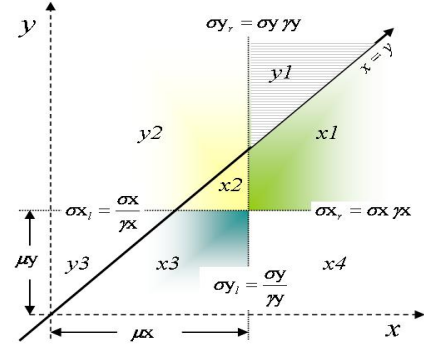


Figure 4: Standard deviations of a bivariate Skewed Normal distribution and seven regions of integration for $\mu_x > \mu_y$

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \phi\left(\frac{x - \mu_x}{\sigma_x}, \frac{y - \mu_y}{\sigma_y}\right)$$

where, $\phi(x, y) = \frac{1}{\sqrt{1 - \rho^2}} e^{-\frac{x^2 - 2\rho xy + y^2}{2(1 - \rho^2)}}$

Recall that bivariate normal representation being symmetric will introduce errors in the computation of the recursive max operation for SSTA purposes. Therefore, similar to the univariate skewed normal presented in the previous section, we add two inverse scale parameters γ_x and γ_y for random variables X and Y around their respective means μ_x and μ_y for introducing skewness in the bivariate distribution.

$$f_\gamma(x, y) = \frac{1}{\Gamma\sigma_x\sigma_y} \left(\begin{aligned} &\phi\left(\frac{x - \mu_x}{\sigma_{x_l}}, \frac{y - \mu_y}{\sigma_{y_l}}\right) I_{(-\infty, \mu_x)}(x) I_{(-\infty, \mu_y)}(y) \\ &+ \phi\left(\frac{x - \mu_x}{\sigma_{x_l}}, \frac{y - \mu_y}{\sigma_{y_r}}\right) I_{(-\infty, \mu_x)}(x) I_{[\mu_y, \infty)}(y) \\ &+ \phi\left(\frac{x - \mu_x}{\sigma_{x_r}}, \frac{y - \mu_y}{\sigma_{y_l}}\right) I_{[\mu_x, \infty)}(x) I_{(-\infty, \mu_y)}(y) \\ &+ \phi\left(\frac{x - \mu_x}{\sigma_{x_r}}, \frac{y - \mu_y}{\sigma_{y_r}}\right) I_{[\mu_x, \infty)}(x) I_{[\mu_y, \infty)}(y) \end{aligned} \right)$$

$$\text{where } \Gamma = \frac{\pi}{2} \left(\gamma_x + \frac{1}{\gamma_x} \right) \left(\gamma_y + \frac{1}{\gamma_y} \right) + \left(\gamma_x - \frac{1}{\gamma_x} \right) \left(\gamma_y - \frac{1}{\gamma_y} \right) \tan^{-1} \left(\frac{\rho}{\sqrt{1 - \rho^2}} \right);$$

$$\sigma_{x_l} = \frac{\sigma_x}{\gamma_x}; \quad \sigma_{y_l} = \frac{\sigma_y}{\gamma_y}; \quad \sigma_{x_r} = \sigma_x \gamma_x; \quad \text{and } \sigma_{y_r} = \sigma_y \gamma_y.$$

Due to the correlation ρ , the normalizing constant term Γ differs from the univariate case. Figure 4 graphically illustrates how the two indicator functions partition the X, Y plane into 4 quadrants having different standard deviations around the mean vector (μ_x, μ_y) . Strictly speaking the arrival time distributions may not necessarily have a bi-variate skew normals; however, introducing additional skewness parameters allows us more degrees of freedom in comparison with [8]. Furthermore, in the absence of skewness

the bi-variate skew-normal representation reduces to the bivariate normal representation and therefore in this case we get the exact same results as [8]. For this bivariate skew normal distribution we now derive results for computing the moments of the max of X and Y based on the on the original derivation given in [8]. Let $v(i)$ be the i^{th} moment of $\max(X, Y)$ given by

$$\begin{aligned} v(i) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\max(x, y))^i f_{\gamma}(x, y) dy dx \\ &= \oint_{(x, y) \in X > Y} x^i f_{\gamma}(x, y) d(x, y) \\ &+ \oint_{(x, y) \in X \leq Y} y^i f_{\gamma}(x, y) d(x, y) \end{aligned}$$

As shown in Figure 4 the region $X > Y$ gets further partitioned into 4 sub-regions x_1, x_2, x_3 and x_4 where the sub-script denotes the standard deviation quadrant and likewise, region $X \leq Y$ gets partitioned into sub-regions y_1, y_2 and y_3 . Therefore, we can write the i^{th} moment of $\max(X, Y)$ as

$$v(i) = \sum_{j=1}^3 v_{y,j}(i) + \sum_{j=1}^4 v_{x,j}(i)$$

where, $v_{x,j}(i)$ and $v_{y,j}(i)$ are the i^{th} moment of $\max(x, y)$ in the j^{th} quadrant. The complete derivation of $v(i)$ over all the seven regions is repetitive and tedious. Therefore, in this paper we will present the key steps encountered while deriving the expression for moments of sub-region y_1 . The i^{th} moment of max, for $Y > X$ in the 1st quadrant, is given as follows:

$$v_{y,1}(i) = \frac{1}{\Gamma \sigma_x \sigma_y} \int_{\mu_x}^{\infty} \int_x^{\infty} y^i \phi \left(\frac{x - \mu_x}{\sigma_x}, \frac{y - \mu_y}{\sigma_y} \right) dy dx$$

Using the Leibnitz rule, we compute the partial derivative of $v_{y,1}(i)$ with respect to μ_x :

$$\begin{aligned} \frac{\partial v_{y,1}(i)}{\partial \mu_x} &= \frac{1}{\Gamma \sigma_x \sigma_y} \int_{\mu_x}^{\infty} \int_x^{\infty} y^i \frac{\partial}{\partial \mu_x} \phi \left(\frac{x - \mu_x}{\sigma_x}, \frac{y - \mu_y}{\sigma_y} \right) dy dx \\ &- \frac{1}{\Gamma \sigma_x \sigma_y \sqrt{1 - \rho^2}} \int_{\mu_x}^{\infty} y^i e^{-\frac{(y - \mu_y)^2}{2\sigma_y^2}} dy \end{aligned}$$

We first change order of integration variables in the inner integral,

$$\begin{aligned} \frac{\partial v_{y,1}(i)}{\partial \mu_x} &= \frac{1}{\Gamma \sigma_x \sigma_y \sqrt{1 - \rho^2}} \int_{\mu_x}^{\infty} y^i e^{-\frac{(y - \mu_y)^2}{2\sigma_y^2}} \\ &\int_{\mu_x}^y \frac{\partial e^{-\frac{(x - \mu_x - \rho(y - \mu_y))^2}{2(1 - \rho^2)}}}{\partial \mu_x} dx dy \\ &- \frac{\gamma_1}{\Gamma \sigma_x \sigma_y \sqrt{1 - \rho^2}} \int_{\mu_x}^{\infty} y^i e^{-\frac{(y - \mu_y)^2}{2(1 - \rho^2)}} dy \end{aligned}$$

Now, the inner integral of the first term in the above expression is in an integrable form. We evaluate this integral and an additional term due to the integration cancels out the second term and gives us the following simplified result.

$$\frac{\partial v_{y,1}(i)}{\partial \mu_x} = -\frac{1}{\Gamma \sigma_x \sigma_y \sqrt{1 - \rho^2}} \int_{\mu_x}^{\infty} y^i e^{-\frac{(y - \mu_y)^2}{2\sigma_y^2}} e^{-\frac{(y - \mu_x - \rho(y - \mu_y))^2}{2(1 - \rho^2)}} dy$$

Similar to [8], we first make the substitution $y = \frac{(\sigma_x \sigma_y \sqrt{1 - \rho^2})n}{a} + \mu_y + \frac{(\mu_x - \mu_y) \sigma_y (\sigma_y - \sigma_x \rho)}{a^2}$ and then $\mu_x = \mu_y + am$.

$$\begin{aligned} \frac{\partial v_{y,1}(i)}{\partial m} &= -\frac{1}{\Gamma} e^{-\frac{m^2}{2}} \int_{\frac{m(\sigma_x - \sigma_y \rho)}{\sigma_y \sqrt{1 - \rho^2}}}^{\infty} e^{-\frac{n^2}{2}} \\ &\left(\mu_y + \frac{n \sigma_x \sigma_y \sqrt{1 - \rho^2}}{a} + \frac{m \sigma_y (\sigma_y - \sigma_x \rho)}{a} \right)^i dn \end{aligned}$$

where,

$$a^2 = \sigma_x^2 + \sigma_y - 2\rho \sigma_x \sigma_y$$

Now note that for $m = \infty$, the random variable $X \gg Y$ and therefore, at $m = \infty$ all moments $v_{y,1}(i) = 0$. Using this observation one can express $v_{y,1}(i)$ as follows:

$$v_{y,1}(i) = \frac{1}{\Gamma} \int_{\alpha}^{\infty} e^{-\frac{m^2}{2}} \int_{k_3 m}^{\infty} (\mu_y + k_1 n + k_2 m)^i e^{-\frac{n^2}{2}} dn dm \quad (4)$$

where,

$$\begin{aligned} k_1 &= \frac{\sigma_x \sigma_y \sqrt{1 - \rho^2}}{a}, & k_2 &= \frac{\sigma_y (\sigma_y - \sigma_x \rho)}{a}, \\ k_3 &= \frac{\sigma_x - \sigma_y \rho}{\sigma_y \sqrt{1 - \rho^2}} & \text{and } \alpha &= \frac{(\mu_x - \mu_y)}{a}. \end{aligned}$$

For a given positive integer value of i , the above integral can be expressed in terms of well known special functions. For example the first moment can be written as

$$\begin{aligned} v_{y,1}(1) &= \frac{\sqrt{\pi}}{\Gamma \sqrt{2}} \left(\frac{(k_1 - k_2 k_3)}{\sqrt{k_3^2 + 1}} \operatorname{erfc} \left(\frac{\alpha \sqrt{k_3^2 + 1}}{\sqrt{2}} \right) \right. \\ &\left. + e^{-\frac{\alpha^2}{2}} k_2 \operatorname{erfc} \left(\frac{k_3 \alpha}{\sqrt{2}} \right) \right) + \frac{\mu_y}{\Gamma} T(\alpha, k_3 \alpha) \end{aligned}$$

where, $\operatorname{erfc}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ is the complementary error function and $T(x, y)$ is the Owen's T-function [15], given by

$$T(x, y) = \frac{1}{\pi} \int_0^y \frac{e^{-\frac{1}{2}x^2(1+t^2)}}{1+t^2} dt.$$

The special functions $\operatorname{erfc}(x)$ and $T(x, y)$ are commonly encountered while integrating univariate and bivariate normal distributions, respectively. Precise numerical tables or accurate closed form analytical approximations exist for both $\operatorname{erfc}(x)$ and $T(x, y)$. Efficient analytical solution for computing $T(x, y)$ is given in [16]. Thus similar to [8], the moments of the max can be found in a computationally efficient manner. Likewise, higher moments can also be found by evaluating the integral given in Equation 4 at higher values of i . Using similar manipulations, the moments of max in all seven regions can be computed. Note that the case when $\mu_x = \mu_y$ we will only have 6 regions instead of 7 as $(\mu_x = \mu_y)$ will lie on the line $x = y$. For such a case, $T(x, y)$ can be analytically integrated by changing the variables to polar coordinates.

3.1 Applying SN Max to SSTA

In this sub-section, we will briefly discuss how the proposed skew normal max results can be applied to block based SSTA. First we review how Clark's max results are currently being used for computing $\max(X, Y)$ of two arrival times X and Y . As mentioned earlier, X and Y are expressed in terms of a canonical form [6, 17]. First, the variances and the covariance of X and Y are computed from the canonical form. Then by substituting the mean, variance and correlation values into the analytical results given in [8], the statistics of the max distribution, namely, tightness probability (i.e. probability of one input being greater than the other [17]), the mean and the variance of their maximum are computed. Thereafter, the approximate distribution of the $\max(X, Y)$ in canonical form is determined by taking a linear combination of the two input arrival times weighted by their tightness probability [6, 7, 19]. Finally, the mean and variance of the resulting canonical form distribution are matched with the analytically computed mean and variance of the max computed from Clark's max results.

The focus of the proposed skew normal max operation is to take into account skewness of X and Y in addition to mean and variance of the arrival time distribution. Non-linear canonical timing models (for example, the quadratic timing model) proposed in previous approaches [7, 12, 18, 19], are inherently skewed, and if X and Y are expressed in terms of such non-linear canonical forms, then the skewness of X and Y can also be analytically found in addition to the variance and covariance. Recall from the discussion in Section 2, that the parameters of skew normal approximations of X_γ and Y_γ can be efficiently found from their respective mean, variance and skewness. Assuming that X_γ and Y_γ have a joint skew normal distribution given in the previous section, the approximate moments of the $\max(X, Y)$ can be computed from the analytical results derived in the previous section. The tightness probability, mean, variance and skewness of the maximum can be found from these moments of the max. Thus, given the statistics of two correlated arrival time distribution we can compute the statistics of their maximum and their respective tightness probability needed for the statistical maximum algorithm in SSTA. Therefore, these results can be applied to existing SSTA approaches that rely on [8] for computing the approximate $\max(X, Y)$ in canonical form.

4. NUMERICAL RESULTS

In this section, we will present a comparison between the proposed Skew normal max results and Clark's max results. Our goal is to show the usefulness of the proposed max operation in an SSTA framework. Therefore, to emulate the actual use of these results in true statistical max algorithm, we generated a test suite consisting of skewed arrival time distributions by running 50,000 Monte Carlo(MC) simulations on a toy circuit that mimics the behavior of a real circuit. In this setup, the arrival times at the primary inputs were assumed to have a correlated multivariate normal distribution. The the relative mean alignment, the ratio of the variance and the correlation were swept with in a reasonable range for generating skewed distributions at the internal nodes. For each max operation performed at an internal node, the two input operand arrival time distributions and their resulting maximum arrival time distributions were logged during MC simulations. The statistics of the operand arrival times were used as an input to the proposed max function implemented in C++. We implemented the analytical results presented in [16] for evaluating $T(X, Y)$. For comparison purposes, we also implemented the 5-parameter Clark's max results. The error in the result of both the proposed max operation and Clark's max operation was computed relative to the MC simu-

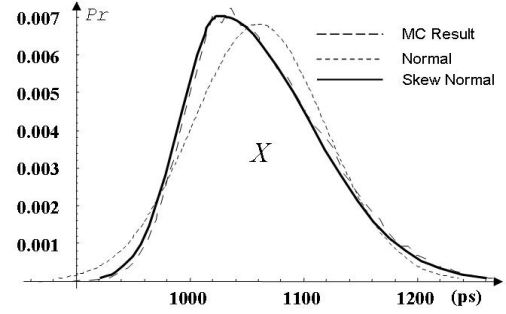


Figure 5: Example: Input X PDF.

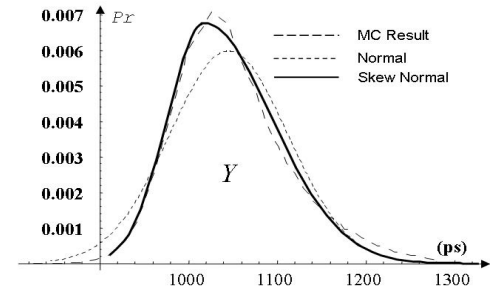


Figure 6: Example: Input Y PDF.

lation results of the output arrival distribution for each test case.

Now for every test case, we computed the statistical parameters of the two input arrival time distributions, namely, μ_{X_γ} , σ_{X_γ} , Sk_{X_γ} , μ_{Y_γ} , σ_{Y_γ} , Sk_{Y_γ} and ρ . These 7 statistical parameters were the input to the proposed skewed normal function. Using the moment matching method presented in Section 2, we first find the parameters of skew normal distribution and then using the analytical max results derived in Section 3, we compute the output moments of $\max(X, Y)$. An example illustrating the efficacy of the max operation is given in Figure 5, 6 and 7. Given the statistics of $X(1060.55, 58.56, 0.56)$, $Y(1045.53, 66.73, 0.80)$ and their correlation, the parameters of skewed normal probability distribution function are first computed. It can be seen from these figures that the skewed normal distribution accurately represents the MC generated skewed arrival time distribution as compared to the symmetric normal for both the inputs. Consequently, as shown in Figure 7 a skewness based treatment of the input arrival time distribution gives a $\max(X, Y)$ distribution that accurately matches the MC simulation results.

We found that the error in the standard deviation of the max operation based on a normal assumption increases significantly with increase in skewness of the two input arrival time distributions. This is illustrated in Figure 8 where, we show a plot of percentage error in computing the standard deviation of the $\max(X, Y)$ as a function of the skewness in X , Sk . It is evident from this plot that the proposed skewed normal max operation can significantly improve the accuracy of existing SSTA approaches.

Furthermore, as mentioned in [8], the error of the max operation also increases when the difference between μ_{X_γ} and μ_{Y_γ} decreases. We observed a similar trend in our simulation results. In Figure 9 we present error plots of percentage error in standard deviation of output arrival time as a function of $\frac{\mu_{X_\gamma} - \mu_{Y_\gamma}}{a}$. It is clear from this

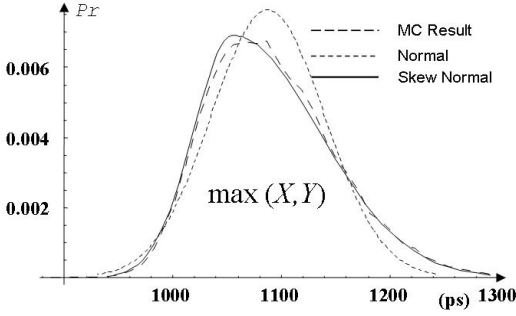


Figure 7: Example: Result $\max(X, Y)$ PDF.

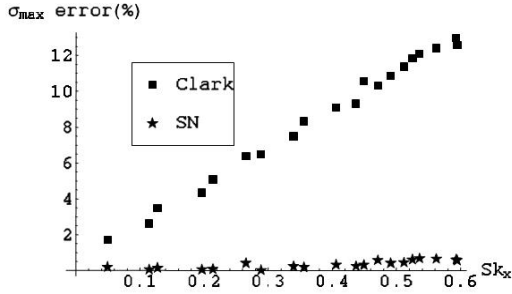


Figure 8: Comparison of standard deviation σ_{max} error (%) as a function of input arrival time skewness Sk_x

plot that the proposed method exhibits much better robustness to difference in the mean of input arrival time distribution.

Recall from the discussion in Section 3.1, that while computing the correlated coefficients of $Z = \max(X, Y)$, a linear sum of the two operand coefficients weighted by their respective tightness probability is computed. In Figure 10, we show a comparison of error for this step between the Skew normal max and Clark's Max operation. It is evident from this plot that the error in computing the correlation coefficient using both the max operations is very similar. This result illustrates the fact that correlated propagation of canonical forms can be achieved using tightness probabilities computed from Skew Normal max results.

In addition to the above results, for evaluating the impact of the proposed skew normal max operation on the benchmark circuits we also compared the proposed max operation with Clark's max

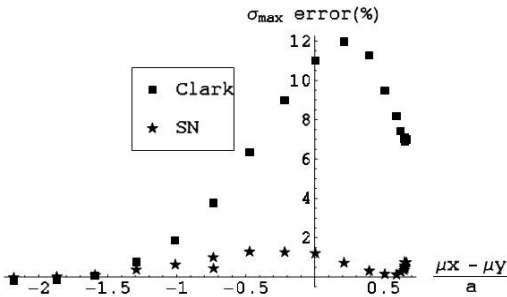


Figure 9: Comparison of standard deviation σ_{max} error (%) as a function of $\frac{\mu_X - \mu_Y}{a}$

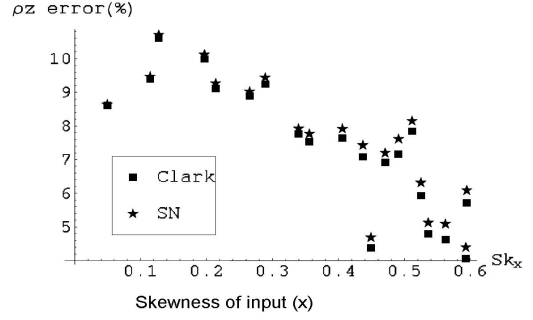


Figure 10: Comparison of error in correlation coefficient calculation as a function of input arrival time skewness Sk_x

by comparing them with MC simulations. The benchmark circuits were synthesized using an industrial 0.13μ technology and placed using Cadence Silicon Ensemble. The $3\sigma/mean$ of 20% was considered for channel length and gate length-independent threshold voltage variations. All variation in V_{th0} was assumed to be random (due to random dopant effects), whereas half the variation in channel length was considered to be correlated. The gates in the library were characterized for delay using SPICE simulations for different values of channel length and V_{th0} , which were fit to a linear canonical form. A grid based spatial correlation model similar to the one proposed in [6] was used. In the absence of real correlation data, the correlation coefficient among different squares on the grid was assumed to be inversely proportional to the distance between the centers of their grids. MC simulations were performed by generating correlated and random Gaussian random variables for the process parameters. and arrival time distributions were logged for each max operation computed in the circuit. Both the max results were tested against these MC generated statistics for each max operation computed in the circuit.

Similar to the previous research [6, 17], the maximum error in mean computed from both Clark and SN Max was found to be less than 1%; however, as shown in Table 1, for standard deviation as expected the accuracy of Skew Normal Max was found to be better than the Clark's Max. The maximum and average error in standard deviation computed using both Clark and the proposed Skew Normal max over all max operations performed in the circuit is listed in Table 1. Note that in most of the cases the proposed skew normal max operation has better accuracy in standard deviation than the Clark's max approach. Interestingly, although the maximum error in Clark's standard deviation is on the high side but the average error is comparable to the Skew Normal max. This result suggests that it for an efficient implementation in an SSTA algorithm one can selectively use the proposed skew normal max operation where necessary. On the other hand, due to the linear canonical form and normal distributions of process parameters our current results only modeled the skewness introduced due to the nonlinear max operation. Therefore, we believe for more realistic nonlinear delay models and non gaussian process parameters, due to additional inherent skewness further improvement over Clark's max can be achieved.

Additional investigation of the sources of error in the proposed max operation revealed that most of the maximum error cases for the Skew Normal max occurred when the correlation was relatively high (typically > 0.9). Because of the systematic nature of this error we believe it will be possible to reduce it further. However, this requires a more detailed study which is the subject of our ongoing research. Nevertheless, to the best of our knowledge, this

is the first work that analytically addresses the problem of computing the maximum of non-normal distributions in SSTA. We believe that the proposed work appeals to be a promising new direction for improving the accuracy of max operation in SSTA algorithms.

Table 1: Comparison of Maximum and average error of standard deviation between Clark’s max and SN Max over all the max operations

Circuit	#Gates	Max. % Error in SD		Avg. Error in SD	
		Clark	Skew Normal	Clark	Skew Normal
c432	257	2.800	1.153	0.054	0.051
c499	545	1.334	1.685	0.067	0.066
c880	501	2.345	-0.780	0.054	0.066
c1908	604	2.588	2.090	0.078	0.070
c2670	781	1.209	0.688	0.048	0.047
c3540	1164	3.489	2.222	0.064	0.068
c5315	1693	6.736	2.589	0.064	0.063
c7552	2153	3.694	3.063	0.085	0.078
i2	193	0.776	0.561	0.044	0.041
i4	265	1.072	-0.773	0.156	0.137
i5	424	1.834	0.912	0.145	0.091
i6	462	0.977	-0.651	0.090	0.066
i7	770	1.086	0.879	0.099	0.087
i8	1014	1.795	1.151	0.072	0.059
i10	2483	2.583	1.486	0.050	0.049

5. CONCLUSION

In this work we present an analytical approach that extends Clark’s max results to skew normal distributions for computing the statistical maximum of two skewed arrival time distributions. An efficient method is presented to approximate the arrival time distribution using skew normal representation. This done by matching the mean, the variance and the skewness of arrival time distributions to that of the skew normal approximation. Using this method we then derived analytical results for computing the approximate moments of the maximum of the arrival time distribution assuming their joint PDF as a bi-variate skew normal distribution. From these moments the tightness probability, mean, variance and skewness of the maximum can be computed and therefore the presented results can be applied in existing SSTA algorithms that work on Clark’s results. Our numerical results show that the proposed max operation can improve the accuracy of existing SSTA approaches. Furthermore, the skewness based proposed max function can be used to augment existing SSTA framework to propagate three moments.

6. REFERENCES

- [1] Aseem Agarwal, David Blaauw, and Vladimir Zolotov. Statistical timing analysis using bounds and selective enumeration. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 22(9):1243–1260, Sept 2003.
- [2] Aseem Agarwal, David Blaauw, Vladimir Zolotov, Savithri Sundareswaran, Min Zhou, Kaushik Gala, and Rajendran Panda. Statistical delay computation considering spatial correlations. In *ACM/IEEE Asia-Pacific Design Automation Conference (ASP-DAC)*, 2003.
- [3] MRCM Berkelaar. Statistical delay calculation, linear time method. In *Workshop TAU*, 1997.
- [4] Sarvesh Bhardwaj, Sarma Vrudhula, and David Blaauw. τ AU: Timing Analysis under Uncertainty. In *IEEE International Conference on Computer Aided Design*, 2003.
- [5] M. Cain. The moment-generating function of the minimum of bivariate normal random variables. 48(2):124–125, 1994.
- [6] Hongliang Chang and Sachin Sapatnekar. Statistical timing analysis considering spatial correlations using a single pert-like traversal. In *IEEE International Conference on Computer Aided Design*, 2003.
- [7] Hongliang Chang, Vladimir Zolotov, Sambasivan Narayan, and Chandu Visweswariah. Parameterized block-based statistical timing analysis with non-gaussian parameters, nonlinear delay functions. In *DAC ’05: Proceedings of the 42nd annual conference on Design automation*, 2005.
- [8] C.E. Clark. The greatest of a finite set of random variables. 9:85–91, 1961.
- [9] Anirudh Devgan and Chandramouli Kashyap. Block-based static timing analysis with uncertainty. In *IEEE International Conference on Computer Aided Design*, 2003.
- [10] C. Fernandez and M.F.J. Steel. On bayesian modelling of fat tails and skewness. pages 359–371, 1998.
- [11] O. Kella. On the distribution of the maximum of bivariate normal random variables with general means and variances. 15(11):3265–3276, 1986.
- [12] Vishal Khandelwal and Ankur Srivastava. A general framework for accurate statistical timing analysis considering correlations. In *DAC ’05: Proceedings of the 42nd annual conference on Design automation*, pages 89–94, New York, NY, USA, 2005. ACM Press.
- [13] Jing-Jia Liou, Angela Krstic, Li-C. Wang, and Kwang-Ting Cheng. False-path-aware statistical timing analysis and efficient path selection for delay testing and timing validation. In *DAC ’02: Proceedings of the 39th conference on Design automation*, pages 566–569, New York, NY, USA, 2002. ACM Press.
- [14] Michael Orshansky and Kurt Keutzer. A general probabilistic framework for worst case timing analysis. In *DAC ’02: Proceedings of the 39th conference on Design automation*, pages 556–561, New York, NY, USA, 2002. ACM Press.
- [15] D. B. Owen. Tables for computing bivariate normal probabilities. 27:1075–1090, 1956.
- [16] Mike Patefield. Fast and accurate calculation of Owen’s T function. *Statistical Software*, 5(5):1–25, 2000.
- [17] C. Visweswariah, K. Ravindran, K. Kalafala, S. G. Walker, and S. Narayan. First-order incremental block-based statistical timing analysis. In *DAC ’04: Proceedings of the 41st annual conference on Design automation*, pages 331–336, New York, NY, USA, 2004. ACM Press.
- [18] Yaping Zhan, Andrzej J. Strojwas, Xin Li, Lawrence T. Pileggi, David Newmark, and Mahesh Sharma. Correlation-aware statistical timing analysis with non-gaussian delay distributions. In *DAC ’05: Proceedings of the 42nd annual conference on Design automation*, pages 77–82, New York, NY, USA, 2005. ACM Press.
- [19] Lizheng Zhang, Weijen Chen, Yuheng Hu, John A. Gubner, and Charlie Chung-Ping Chen. Correlation-preserved non-gaussian statistical timing analysis with quadratic timing model. In *DAC ’05: Proceedings of the 42nd annual conference on Design automation*, pages 83–88, New York, NY, USA, 2005. ACM Press.