A NEW STATISTICAL METHOD FOR MAXIMUM POWER ESTIMATION IN CMOS VLSI CIRCUITS

N. E. EVMORFOPOULOS and J. N. AVARITSIOTIS*

*Department of Electrical and Computer Engineering, National
Technical University of Athens, 9, Iroon Polytechniou St.,
Zographou, Athens, 15773, Greece*

(Received 11 October 1999; In final form 18 November 1999)

A method for maximum power estimation in CMOS VLSI circuits is proposed. The method is based on extreme value theory and allows for the calculation of the upper end point of the probability distribution which is followed by the instantaneous power drawn from the supply bus. The main features of the method are the relatively small and circuit-independent subset of input patterns required for accurate prediction of maximum power and its simulative nature which ensures that no over-simplifying assumptions are made. Application of the proposed method to eight distributions, which come close to the behavior of power consumption in VLSI circuits, proved its superior capabilities with respect to existing methods.

Keywords: Maximum power; reliability; CMOS VLSI circuits; statistics; extremes

1. INTRODUCTION

Excessive instantaneous power dissipation in CMOS VLSI circuits is a potential source of many reliability and performance problems such as physical failures due to overheating or electromigration, functional failures caused by voltage drop (or IR-drop) on power rails, reduced noise margins and switching speed degradation. The issues regarding reliability are becoming even more important as we move into deep-submicron ICs which exhibit decreased feature sizes and

*Corresponding author. e-mail: abari@cs.ntua.gr

power supply levels. Consequently, there is an increasing need for tools and methods that perform fast and accurate estimation of the circuit's maximum power requirements during the design phase.

Previous work on this field has focused on the creation and exploitation of deterministic or stochastic information (such as switching activity and signal correlations) about the circuit's inputs, without explicitly simulating the circuit. While most related techniques exhibit good timing figures, they make some simplifying assumptions that often result in loss of accuracy. In addition, their speed and algorithmic efficiency generally degrade with large or complex circuit structures. Statistical methods might offer a promising alternative but they have not seen a widespread application as yet, due to the inadequacy of existing statistical models to deal with extreme situations such as in maximum power analysis. Even though a step towards the right direction was made in [1], the reported method overlooked some important facts about extremes.

This paper presents a new statistical method for maximum power estimation which is based on extreme value theory and more precisely on the properties of extreme order statistics of the probability distribution of the instantaneous power drawn from the power supply bus. The proposed method has proved extremely fast in its operation, as only a small subset of input patterns is required to obtain a reasonably accurate estimate of maximum power. Furthermore, the number of units of this subset is constant and independent of the circuit size or complexity, whereas most competing techniques impose a computational load that grows exponentially with circuit size and may prove prohibitively expensive to apply to large circuits. On the other hand, there are no over-simplifying assumptions to hinder the necessary accuracy, since the statistical inference is performed on the outcome of highly-accurate simulation of the circuit under typical input vectors.

2. THEORETICAL BACKGROUND

2.1. Order Statistics and Extreme Order Statistics

DEFINITION 2.1 Let $X_1, X_2, \dots, X_n$ be a sample of size n from a given population with cumulative distribution function (cdf) $F(x)$.

The units X_r are drawn in random so that they form independent and identically distributed (iid) random variables with cdf equal to $F(x)$. If they are subsequently rearranged in an increasing order of magnitude, $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$, then the r -th term $X_{r:n}$ of this new sequence is called the r -th order statistic of the sample. In particular, the terms $X_{1:n} = \min(X_1, X_2, \dots, X_n)$ and $X_{n:n} = \max(X_1, X_2, \dots, X_n)$ are called extreme order statistics (minima and maxima order statistics respectively) or simply extremes.

THEOREM 2.1 [2] *The cdf $F_{X_{r:n}}(x)$ of the r -th order statistic of a sample drawn from a population with cdf $F(x)$, is given by:*

$$F_{X_{r:n}}(x) = r \binom{n}{r} \int_0^{F(x)} u^{r-1} (1-u)^{n-r} du = I_{F(x)}(r, n-r+1) \quad (2.1)$$

where $I_p(a, b)$ is the incomplete beta function. By setting $r = n$ we obtain the cdf of the maxima:

$$F_{X_{n:n}}(x) = F^n(x) \quad (2.2)$$

The above expressions require the knowledge of the parent cdf. However, in the frequently encountered case where the probability distribution of a physical quantity is unknown, they prove quite insufficient. Therefore we have to turn to the asymptotic distributions of extremes (if any) that emerge in the limit $n \rightarrow +\infty$ of the sample size.

2.2. Asymptotic Distributions of Maxima Order Statistics

DEFINITION 2.2 A cdf $F(x)$ is said to lie in the domain of attraction for maxima of cdf $H(x)$ when it satisfies:

$$\lim_{n \rightarrow +\infty} F^n(a_n + b_n x) = H(x) \quad (2.3)$$

for given normalizing sequences $\{a_n\}$ and $\{b_n > 0\}$.

THEOREM 2.2 [3] *The only three possible types of non-degenerated (i.e., not asymptotically constant) distributions $H(x)$ satisfying condition*

(2.3) are:

$$H_{1,\beta}(x) = \begin{cases} \exp(-x^{-\beta}) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}, \beta > 0 \quad \text{Frechet} \quad (2.4)$$

$$H_{2,\beta}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ \exp(-(-x)^\beta) & \text{if } x < 0 \end{cases}, \beta > 0 \quad \text{Weibull} \quad (2.5)$$

$$H_3(x) = \exp(-\exp(-x)) \quad -\infty < x < +\infty \quad \text{Gumbel} \quad (2.6)$$

THEOREM 2.3 [3] *Let the upper end point of cdf $F(x)$ be defined as $\omega(F) = \sup\{x: F(x) < 1\}$. It represents the maximum value that the associated random variable can acquire and becomes $\omega(F) = F^{-1}(1)$ if the random variable is bounded or $\omega(F) = +\infty$ in the opposite case. Then, cdf $F(x)$ lies in the domain of attraction for maxima of:*

- (i) $H_{1,\beta}(x)$ if and only if $\omega(F) = +\infty$ and $\lim_{t \rightarrow +\infty} (1 - F(tx)/(1 - F(t))) = x^{-\beta}, x > 0$.
- (ii) $H_{2,\beta}(x)$ if and only if $\omega(F) < +\infty$ and $\lim_{t \rightarrow +\infty} (1 - F(\omega(F) - (1/tx))/(1 - F(\omega(F) - (1/t)))) = x^{-\beta}, x < 0$, or by substituting $t \leftrightarrow (1/t) > 0$ and $x \leftrightarrow -(1/x) < 0$, $\lim_{t \rightarrow 0^+} (1 - F(\omega(F) + xt)/(1 - F(\omega(F) - t))) = (-x)^\beta, x < 0$.
- (iii) $H_3(x)$ if and only if $\lim_{n \rightarrow +\infty} n[1 - F(F^{-1}(1 - (1/n)) + x(F^{-1}(1 - (1/ne)) - F^{-1}(1 - (1/n))))] = \exp(-x)$, where the inverse function $F^{-1}(a)$ corresponds to the 100a quantile (or percentile) of $F(x)$.

By introducing parameters μ and $\sigma > 0$, we obtain the following generalized parametric expressions for the three limit distributions (2.4)–(2.6):

$$H_1(x; \mu, \sigma, \beta) = \exp\left[-\left(\frac{x - \mu}{\sigma}\right)^{-\beta}\right] \quad \text{Frechet family} \quad (2.7)$$

$$H_2(x; \mu, \sigma, \beta) = \exp\left[-\left(-\frac{x - \mu}{\sigma}\right)^\beta\right] \quad \text{Weibull family} \quad (2.8)$$

$$H_3(x; \mu, \sigma) = \exp\left[-\exp\left(-\frac{x - \mu}{\sigma}\right)\right] \quad \text{Gumbel family} \quad (2.9)$$

Parameters μ , σ and β are referred to as location (or threshold), scale and shape parameters respectively.

2.3. Properties of Asymptotic Distributions

This section contains some important facts about the three aforementioned families of asymptotic distributions. Theorem 2.3 implies that a parent distribution with a finite upper end point ($\omega(F) < +\infty$) cannot lie in a Frechet-type domain of attraction, whereas a distribution with an infinite upper end point ($\omega(F) = +\infty$) cannot belong to a Weibull-type domain of attraction. However, these observations are not enough to guarantee a Frechet or Weibull-type domain of attraction for a particular distribution. In fact the Gumbel limit distribution, even though it is not bounded by itself, asymptotically models many parent distributions with either finite or infinite end points. The upper end point of the Weibull family (2.8) is equal to the location parameter μ , while both Frechet and Gumbel families have infinite upper end points.

The domain of attraction of a cdf $F(x)$ is determined by the rate with which its right tail approaches the value of one as $x \rightarrow \omega(F)$. Depending on this rate we can distinguish between three families of distributions [4], namely the Cauchy-Pareto family, the Beta family (a special case of which is the well-known uniform distribution) and the Exponential family (which includes most of the common statistical distributions such as the Normal and Gamma ones). The Cauchy-Pareto family is characterized by an infinite upper end point and a very slow rate of convergence to one at its tail and is therefore asymptotically modeled by a Frechet limit distribution. The tail of a distribution of the Exponential family converges to one at least as fast as that of the exponential distribution with cdf $F(x) = 1 - \exp(-(x - a_2)/a)$. Both unbounded and bounded cdfs belong to this family and are all modeled by a Gumbel limit distribution at their extremes. Finally, members of the Beta family of distributions, which asymptotically approach the Weibull limit distribution, exhibit the faster (approximately linear) convergence towards one and are all bounded by finite upper limits.

2.4. Tail Equivalence

DEFINITION 2.3 Two cdfs $F(x)$ and $G(x)$ are referred to as right tail equivalent if and only if

$$\omega(F) = \omega(G) \text{ and } \lim_{x \rightarrow \omega(F)} \frac{1 - F(x)}{1 - G(x)} = 1.$$

THEOREM 2.4 [5] *Two right tail equivalent cdfs $F(x)$ and $G(x)$ belong to the same domain of attraction for maxima and have the same normalizing sequences a_n and b_n i.e.,*

$$\lim_{n \rightarrow +\infty} F^n(a_n + b_n x) = \lim_{n \rightarrow +\infty} G^n(a_n + b_n x) = H(x).$$

The practical implication of the above theorem when dealing with extreme value problems, is that a common cdf $F(x)$ may be fitted to the right tail of an unknown cdf $G(x)$ which is observed to follow the same limit distribution as $F(x)$. This will be the basis of the maximum power estimation method proposed throughout this paper.

2.5. Probability Plots

The basic concept of a probability plot of a given parametric family $F(x; \theta)$ of distributions with vector parameter θ , is to modify the random variable X and probability P axes scales in such way that the plot against X of any cdf belonging to that family becomes a straight line [6]. This allows one to decide whether or not a set of ordered data follows a probability distribution drawn from that particular family. More specifically, if $F(x; \theta)$ can be written as $y = F(x; \theta) = h^{-1}(ag(x) + b)$, then the transformation $\xi = g(x)$, $\eta = h(y)$ changes $F(x; \theta)$ into a family of straight lines $\eta = a\xi + b$.

Such a transformation can be applied to the parametric families (2.7)–(2.9) by taking logarithms twice:

$$-\log(-\log(y)) = \beta \log(x - \mu) - \beta \log \sigma \quad \text{Frechet probability plot equation} \quad (2.10)$$

$$-\log(-\log(y)) = -\beta \log(\mu - x) + \beta \log \sigma \quad \text{Weibull probability plot equation} \quad (2.11)$$

$$-\log(-\log(y)) = \frac{1}{\sigma} x - \frac{\mu}{\sigma} \quad \text{Gumbel probability plot equation} \quad (2.12)$$

The above equations provide the required x -axis transformations $\xi = g(x) = \log(x - \mu)$, $\xi = g(x) = -\log(\mu - x)$ and $\xi = g(x) = x$ for Frechet, Weibull and Gumbel probability plots respectively, as well as the y -axis transformation $\eta = h(y) = -\log(-\log(y))$ which is identical in all three cases. We notice that the x -axis of the Gumbel plot undergoes no change whereas the other two plots turn to logarithmic *abscissa* scales. Moreover, Weibull and Frechet plots require knowledge of the location parameter μ for the computation of the transformed x -axis scale.

Gumbel probability plot is the standard starting point when the domain of attraction of the data set is unknown. The primary reason for this is that the Gumbel family of distributions occupies a central position between Weibull and Frechet families and the shape of data on a Gumbel plot allows us to determine the type of limit distribution (further details on this subject are cited in the next section). Another reason is the ignorance of the location parameter μ for the Weibull and Frechet cdfs. In that case we should first lay the data on a Gumbel plot and, after having decided in favor of a Weibull or Frechet cdf, we proceed by trial and error with μ on the corresponding plot until a straight line is obtained.

2.6. Selection of Limit Distribution

In this section we show that the domain of attraction of a set of data can be invoked directly from the corresponding Gumbel probability plot and specifically from the curvature it exhibits.

The transformations of Frechet and Weibull families (2.7), (2.8) on a Gumbel probability plot give $\eta = -\log(((\xi - \mu)/\sigma)^{-\beta})$ and $\eta = -\log((-(\xi - \mu)/\sigma)^\beta)$ respectively. By differentiating twice with respect to ξ , the above equations yield $(d\eta/d\xi) = (-\beta/(x - \mu)^2)$ and $(d\eta/d\xi) = (\beta/(x - \mu)^2)$. Bearing in mind that $\beta > 0$, we deduce that Frechet distributions appear as concave curves ($d\eta/d\xi < 0$) whereas Weibull distributions appear as convex curves ($d\eta/d\xi > 0$) on a Gumbel probability plot (obviously Gumbel distributions appear normally as straight lines).

The curvature of the plot can be quantified as the quotient of the slopes of two straight lines fitted on two successive intervals of the particular plot segment we are interested in [2]. More precisely, we

compute the quantity $S = (S_{n_1, n_2} / S_{n_3, n_4})$, where

$$S_{n_i, n_j} = \left(\left((n_j - n_i + 1) \sum_{k=n_i}^{n_j} \left(-x_k \log \left(-\log \left(\frac{k-1/2}{n} \right) \right) \right) \right. \right. \\ \left. \left. - \left[\sum_{n_i}^{n_j} x_k \right] \left[\sum_{k=n_i}^{n_j} \left(-\log \left(-\log \left(\frac{k-1/2}{n} \right) \right) \right) \right] \right) \right) / \\ \left((n_j - n_i + 1) \sum_{k=n_i}^{n_j} \left(-\log \left(-\log \left(\frac{k-1/2}{n} \right) \right) \right)^2 \right. \\ \left. - \left[\sum_{k=n_i}^{n_j} \left(-\log \left(-\log \left(\frac{k-1/2}{n} \right) \right) \right) \right]^2 \right)$$

is the slope of the least-squares straight line fit on the order statistics x_k with index k between n_i and n_j .

If the value of S , calculated on a Gumbel probability plot, is well above 1 or well below 1, we may decide in favor of a Weibull or Frechet-type domain of attraction respectively. Significance levels of this test are also provided in the literature [2], though not by analytical calculation but from extensive Monte Carlo simulations (over 5000 units) that were applied on a Gumbel parent.

2.7. Parameter Estimation

The next step after determining the domain of attraction of a parent distribution, is to estimate parameters μ , σ and (possibly) β which are found in expressions (2.7)–(2.9).

If the domain of attraction turns out to be of Gumbel-type, we must first perform a least-squares straight line fit to the Gumbel probability plot and then estimate parameters μ and σ by matching the line equation coefficients to the ones in (2.12).

If the domain of attraction is of Weibull or Frechet-type, we must create the corresponding probability plot for various values of μ until we obtain the best straight line in terms of the minimum fit error. After μ is determined, we may proceed with the estimation of parameters σ and β by matching the line equation to (2.10) or (2.11), just as in the Gumbel case. However, for the particular problem we are dealing

with, we do not need those parameters, since (as it will become clear in Section 3) we are only interested in the location parameter μ .

3. MAXIMUM POWER ESTIMATION PROCEDURE

3.1. Problem Formulation

The term “maximum power” refers to the worst-case instantaneous power (or equivalently current) that is drawn from the supply bus. In order to render the estimation problem precise, we have to consider the sources of power dissipation in digital CMOS circuits, which are [8] $P_{\text{switching}}$, $P_{\text{short-circuit}}$ and P_{leakage} . The first two quantities are the dynamic components of total power since they only occur at transitions between logic states, whereas the third one is the static component as it is a constant source of dissipation. Because of its static nature, leakage power may be regarded as an offset to the total instantaneous power, while at the same time, it is two to three orders of magnitude smaller than the other two components. It is therefore clear that the worst-case power conditions arise strictly on transition intervals and that any corresponding analysis should be restricted inside a clock cycle which marks a transition between two consecutive binary vectors. Because of this, maximum power estimation is also referred to as cycle-accurate power estimation.

Within this context, estimation of maximum power is a fundamentally different problem than that of average power estimation. The former is related to the maximum value, among all possible vector pairs, of the instantaneous power consumed within a clock cycle, whereas the latter involves the mean power of a sequence of input patterns applied over a large number of clock cycles that cover an extended period of time. From a statistical point of view, the probability distribution of average power exhibits smoothing properties over time and approximates a normal (Gaussian) distribution as the number of clock cycles tends to infinity. On the other hand, instantaneous power is considered within short periods of time and therefore its probability distribution is completely undetermined. In addition, a normal distribution model may be suitable for the center of the distribution but not for the tail which always follows an

asymptotic extreme value model. Consequently, conventional probabilistic and statistical techniques [9, 10], though have proved successful for average power estimation, are not adequate to handle maximum power and a radically different approach should be adopted which makes use of extreme value theory.

Instantaneous power dissipation in a CMOS circuit depends upon the specific pair of binary vectors (v_1, v_2) that define a transition on a particular clock edge. If the number of primary inputs is n , then such a vector consists of n bits and obviously the population of input vector pairs contains $2^n \cdot 2^n = 4^n$ units. We will restrict our analysis to combinational circuits and furthermore, we will assume that all primary inputs execute their transitions simultaneously on the same clock edge. However, the generalization to a sequential circuit is straightforward, if we expand the “previous” input vector v_1 to encompass the state bits (or outputs) of all sequential elements (flip-flops) of the circuit. In that case, if the number of state bits is m , then the input population will contain $2^{n+m} \cdot 2^n = 4^n \cdot 2^m$ units. We should also remark that the input population may not always include all possible vector pairs but only selected ones that correspond to typical circuit operation or input switching activity (constrained problem).

3.2. Procedure Description

Since we have defined an input population of vector pairs, we may regard the instantaneous power consumption of a particular circuit as a random variable with a certain probability distribution. The required maximum power will then be represented by the upper end point $\omega(F)$ of this distribution. The proposed method closely follows the concepts presented in Section 2. However, in order to apply them we must assume a continuous probability distribution of power consumption and an input population of infinite size. Both assumptions are most frequently justified in practice for circuits of reasonable size with a moderate number of primary inputs (8 or more).

A flow chart of the estimation procedure is depicted in Figure 1. The main point is to acquire a number of distinct samples, form an additional sample of maxima and graphically transfer it to a Gumbel probability plot in order to determine the domain of attraction and ultimately the upper end point of the probability distribution. The

FIGURE 1 Flow chart of the proposed method for maximum power estimation.

procedure begins with the generation of a number of vector pairs in random way or under certain constraints. This is equivalent to random sampling out of the unconstrained or constrained population of input vector pairs. The circuit due for maximum power analysis is entered in a transistor-level or gate-level simulator such as SPICE or PowerMill. Each generated vector pair is fed as input for transient analysis to the simulator and the peak power during transition time is recorded. The set of power values obtained for all generated vector pairs forms one of the samples to be used hereafter and the procedure is repeated for the rest of them.

The choice of sample size is critical at this point, in order to ensure the validity of the asymptotic models for the minimum possible number of units. Experiments performed on various parent distributions have shown that the rate of convergence to the corresponding domains of attraction is reasonably fast. Indeed, most distributions already approach their limit distribution as soon as the sample reaches a size of about 30 to 50 units, depending on the type of distribution, the problem size and the quality of the selected units. The worst case is encountered in the case of normal distribution, which takes about 100 units to approach its Gumbel-type limit distribution. Thus, the sample size is set to $n = 100$. However one may further improve the quality of the final estimate by increasing the sample size. The choice of the number of samples, or equivalently the size of the sample of maxima, is less critical and a total of $m = 30$ is more than adequate in most cases. On the whole, the total number of units required to obtain an estimate of the maximum power is 3000, which is significantly smaller (by a factor of 2X to 10X) than the majority of the existing methods.

The sample of maxima of power values is subsequently laid on a Gumbel probability plot and the domain of attraction is determined with the aid of the curvature method presented in Section 2.6. According to Monte Carlo tests reported in [2], the critical value of the curvature S for a sample size of 100 is $S = 2.588$, with a significance level of 95%. The indices of the order statistics that define the intervals encountered in the calculation of S are chosen as $n_1 = 1$, $n_2 = n_3 = m/2$ and $n_4 = m$.

Due to physical limitations, it is obvious that the probability distribution of power consumption for a particular circuit is upper bounded (*i.e.*, it has a finite upper end point) and therefore it can only

lie in the domain of attraction of Gumbel or Weibull families (Theorem 2.3). Consequently the proposed method has considered both types of asymptotic models.

If the limit distribution turns out to be of Gumbel-type, we may infer (due to the discussion in Section 2.3) that the parent distribution belongs to the Exponential family regarding its rate of convergence to one as $x \rightarrow \omega(F)$. The tail equivalence principle (Theorem 2.4) allows us then to approximate the right tail of the unknown parent distribution with a function of the form:

$$F(x) = a_1 - \exp\left(-\frac{x - a_2}{a}\right), \quad a_1 > 1 \quad (3.1)$$

If a_1 is close to 1 then $F(x)$ is a probability distribution with the same rate of convergence to one (and consequently with the same Gumbel-type domain of attraction) as the standard exponential distribution, but which has a finite upper end point equal to:

$$\omega(F) = F^{-1}(1) = a_2 - a \log(a_1 - 1) \quad (3.2)$$

Estimates of parameters a_2 and a (location and scale respectively) may be derived with the same approach cited in Section 2.7 for the Gumbel parameter estimation, *i.e.*, by fitting a straight line on a probability plot and by subsequently matching the coefficients of the line and plot equations. Care though has to be taken in that the type of probability plot is now exponential and also that the sample does not necessarily follow the distribution for which the plot is designed, which means that the fit should only include the right tail of the sample. A good choice for the number of points comprising the tail is $2\sqrt{n}$. Parameters a_2 and a are estimated for all 30 samples of power and mean values are taken out of them. The extra parameter a_1 in (3.2) cannot be estimated from the exponential probability plot since the latter is designed for the standard exponential distribution $F(x) = 1 - \exp(-(x - a_2)/a)$. However, a_1 is associated with the Gumbel location parameter μ of the sample of maxima, through the relation (see Appendix):

$$a_1 = \exp\left(-\frac{\mu - a_2}{a}\right) + 1 - \frac{1}{n} \quad (3.3)$$

For the estimation of parameter μ we refer again to Section 2.7 and the Gumbel probability plot of maxima that we have already created. When (3.3) is substituted in (3.2), we obtain the following estimate of the upper end point for the Gumbel case:

$$\omega(F) = a_2 - a \log \left[\exp \left(-\frac{\mu - a_2}{a} \right) - \frac{1}{n} \right] \quad (3.4)$$

The case of a Weibull limit distribution is considerably easier than the Gumbel one, since the upper end point is equal to the location parameter μ which can be estimated as the value to give the best straight line in terms of the minimum fit error on the Weibull plot (Section 2.7).

To be more specific, we calculate the sum of squared errors between the data and the corresponding coordinates of the straight line fit as a function of μ , and then employ a function minimization algorithm such as the Nelder–Mead simplex method [7].

4. EXPERIMENTAL RESULTS

Experiments have been conducted on two sets of distributions which are assumed to come close to the behavior of instantaneous power consumption in VLSI circuits. Distributions that belong to the first set asymptotically follow the Gumbel model, whereas members of the second set are associated with the Weibull extreme model. Each set embodies four distributions with range of values 15.00–49.54 mW, 15.00–222.69 mW, 15.00–533.08 mW and 500.00–707.23 mW respectively. The difference between these two sets is obviously their speed of ascent from the lower end point towards the upper end point. Experimental results for the first and the second set of distributions are shown in Tables I and II respectively. Both tables display the average, maximum and minimum values of the estimated upper end point and the relative percentage error over 100 runs of the procedure for each distribution. The intermediate graphs created for a single run of the estimation procedure applied on the third distribution of both sets are depicted in Figures 2–4 (Gumbel set) and Figures 5–7 (Weibull set).

TABLE I Results for distributions with Gumbel-type domain of attraction

<i>Actual upper end point</i>	<i>Estimated value</i>			<i>Relative error (%)</i>		
	<i>Average</i>	<i>Max</i>	<i>Min</i>	<i>Average</i>	<i>Max</i>	<i>Min</i>
49.54	50.93	53.83	48.66	2.90	8.66	0.07
222.69	229.07	244.05	219.03	2.95	9.59	0.11
533.08	548.82	590.01	516.99	3.36	10.68	0.01
707.23	716.17	737.33	700.79	1.32	4.26	0.01

TABLE II Results for distributions with Weibull-type domain of attraction

<i>Actual upper end point</i>	<i>Estimated value</i>			<i>Relative error (%)</i>		
	<i>Average</i>	<i>Max</i>	<i>Min</i>	<i>Average</i>	<i>Max</i>	<i>Min</i>
49.54	49.54	49.57	49.50	0.02	0.09	0
222.69	222.67	222.93	222.35	0.03	0.15	0
533.08	532.98	533.42	531.07	0.03	0.38	0
707.23	707.21	707.50	706.76	0.01	0.07	0

FIGURE 2 Approximate probability distribution of Gumbel-type domain of attraction and exponential fit on tail data.

Two important observations can be made by closer examination of the above tables. The first one is that the method appears to be much more precise when dealing with distributions of Weibull-type

FIGURE 3 Exponential probability plot (with equation $-\log(1 - y) = x/a - a_2/a$) and least-squares straight line fit on tail data.

FIGURE 4 Gumbel probability plot of Gumbel-type domain of attraction and least-squares straight line fit.

FIGURE 5 Approximate probability distribution of Weibull-type domain of attraction.

FIGURE 6 Gumbel probability plot of Weibull-type domain of attraction.

FIGURE 7 Weibull probability plot of Weibull-type domain of attraction and best least-squares straight line fit.

domain of attraction. This is not a direct result of our method but stems from the fact that, for random and equal-sized samples, a distribution with a fast rate of ascent will exhibit more sample points belonging to its right tail than a distribution with a slower rate of ascent (as is obvious from Figs. 2 and 5). The second observation is that estimates of the upper end point in the Gumbel case appear slightly biased towards larger values than the actual ones on average. This can be attributed to the fact that estimation of parameters a_2 and a of (3.1) is performed through a probability plot which is designed for the standard exponential distribution. However, this distribution has an infinite upper end point, which is a cause for a small positive bias on the estimated value of a finite one.

5. CONCLUSIONS

Fast and accurate estimation of maximum instantaneous power may both allow to spot reliability problems early in the design phase as well as to perform efficient design and optimization of the power

distribution network. A statistical approach based on extreme value theory has been proposed to address these issues. The method's strongest benefit is its speed, which is mainly due to the fact that a small subset of input vectors captures all the necessary information for extreme value analysis. In addition, the simulative nature of the method renders it a very attractive alternative solution, as its integration within the design flow of digital CMOS VLSI circuits is straightforward.

References

- [1] Qiu, Q., Wu, Q. and Pedram, M. (1998). "Maximum power estimation using the limiting distributions of extreme order statistics", *ACM/IEEE Design Automation Conference*.
- [2] Castillo, E. (1988). *Extreme Value Theory in Engineering*, Academic Press.
- [3] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics 2nd edn.*, Krieger.
- [4] Reiss, R.-D. and Thomas, M. (1997). *Statistical Analysis of Extreme Values*, Birkhauser Verlag.
- [5] Resnick, S. (1971). "Tail equivalence and its applications", *J. Applied Probability*, **8**, 136–156.
- [6] Hahn, G. and Shapiro, S. (1994). *Statistical Models in Engineering*, Wiley.
- [7] Nelder, J. and Mead, R. (1965). "A simplex method for function minimization", *Computer J.*, **7**, 308–313.
- [8] Weste, N. and Eshragian, K. (1993). *Principles of CMOS VLSI Design: A Systems Perspective 2nd edn.*, Addison-Wesley.
- [9] Burch, R., Najm, F., Yang, P. and Trick, T., "A Monte Carlo approach for power estimation", *IEEE Trans. VLSI Systems*, **1**, 63–71, Mar., 1993.
- [10] Stamoulis, G. and Hajj, I. (1993). "Improved techniques for probabilistic simulation including signal correlation effects", *ACM/IEEE Design Automation Conference*.

APPENDIX

Proof of Equation (3.3)

Theorems 2.2 and 2.3 allow, for large enough n , substitution of $F^n(a_n + b_n x)$ by $H(x)$ or, if normalizing constants a_n and b_n are passed to the right hand side, substitution of $F^n(x)$ by $H((x - a_n)/b_n)$. A comparison of the latter with (2.7)–(2.9) will reveal that $a_n = \mu$ for any sample of maxima. In particular for the Gumbel case, $a_n = F^{-1}(1 - (1/n))$ [3]. By taking $F(x) = a_1 - \exp(-((x - a_2)/a))$, we deduce that $F(a_n) = a_1 - \exp(-((a_n - a_2)/a)) = 1 - (1/n)$ or $a_1 = \exp(-((a_n - a_2)/a)) + 1 - (1/n) = \exp(-((\mu - a_2)/a)) + 1 - (1/n)$.