Statistical Estimation of Average Power Dissipation in CMOS VLSI Circuits Using Nonparametric Techniques *

Li-Pen Yuan†, Chin-Chi Teng‡, and Sung-Mo Kang†

†Dept. of ECE and Coordinated Science Lab.
U. of Illinois at Urbana-Champaign
1308 W. Main St., Urbana, IL 61801

‡EPIC Design Technology, Inc.
310 N. Mary Ave.
Sunnyvale, CA 94086-4111

Abstract

In this paper, we present a new statistical technique for estimation of average power dissipation in digital circuits. Present statistical techniques estimate the average power based on the assumption that the power distribution can be characterized by a preassumed function. Large error can incur when the assumption is not met. To overcome this problem, we propose a nonparametric technique in which no distribution function needs to be assumed. A set of distribution-independent upper and lower bounds of the average power are developed using the properties derived from the order statistics. A stopping criterion is designed based on the bounds for a desired percentage error with a specified confidence level. Since it does not resort to assuming any particular distribution function, the technique can be applied to all the circuits irrespective of their power distributions. Comparison is made against the present statistical technique based on the central limit theorem. Experimental results show that the proposed technique is much more accurate and robust, yet the efficiency characteristic of statistical techniques is still preserved.

I. Introduction

For state-of-the-art VLSI technology, power analysis and power optimization have become crucial design concerns and have received much attention from DA community. The importance of accurate power analysis is twofold. First, since the battery life of portable equipment and several reliability problems are directly related to power dissipation, accurate power analysis is essential. Second, the quality of the synthesized circuit optimized for low power strongly depends on the accuracy of cost function (power) evaluation. For these reasons, power estimation has become the focus of research efforts in recent years.

Among the approaches proposed in the past to tackle the power estimation problem, statistical technique is an attractive choice because of its accuracy, efficiency and simplicity. For the average power consumption of the whole circuit, it usually requires only a few hundred input vectors to generate an accurate estimate. Since the only information needed to draw a statistical conclusion is the consumed power, to implement a statistical technique, a variable delay circuit simulator can be modified easily to monitor the power and all the signal correlations are implicitly taken into consideration.

Various statistical approaches have been proposed to address the whole circuit average power [1] and individual node activity estimation problem [2][3]. This paper only discusses the first problem. In [1], the average power of a circuit with m gate output nodes over a time interval $(-T/2, +T/2]$ is modeled as a random variable $\boldsymbol{P}_T$ and is expresses as:

$$\boldsymbol{P}_T = \frac{V_{DD}^2}{2} \sum_{i=1}^{m} C_i \frac{\boldsymbol{n}_i(T)}{T}, \tag{1}$$

where C_i is the load capacitance at node i, random variable $\boldsymbol{n}_i$ is the number of transitions occurred at node i, and V_{DD} is the power supply voltage. The average power of the circuit can be expressed as the expected value of $\boldsymbol{P}_T$. By assuming that $\boldsymbol{P}_T$ is normally distributed over any T, for a confidence level $1-\alpha$, the sample average η_T and sample standard deviation s_T of N different $\boldsymbol{P}_T$ samples obey the following relation:

$$|\boldsymbol{P}_T - \eta_T| < \frac{t_{\alpha/2} s_T}{\sqrt{N}}, \tag{2}$$

where $t_{\alpha/2}$ is obtained from the t distribution with $(N-1)$ degrees of freedom. By dividing both sides of (2) by η_T, the absolute error relative to η_T is bounded. For a desired percentage error ϵ, the simulation is continued until the following criterion is satisfied:

$$\frac{t_{\alpha/2} s_T}{\eta_T \sqrt{N}} < \epsilon. \tag{3}$$

The major assumption made in deriving (3) is that $\boldsymbol{P}_T$ over any T is normally distributed. It was assumed [1] that for most of the circuits, the distribution of $\boldsymbol{P}_T$ is at least approximately normal. However, it has been observed that the assumption is not valid for several benchmark circuits. Since the causes for the nonnormal distribution of $\boldsymbol{P}_T$ are not clearly understood, this assumption will inevitably discourage the use of statistical technique as a reliable average power estimator. To overcome this problem, in this paper we propose a new statistical approach for average power estimation. We call this technique *nonparametric*, because conclusion

*This research was supported by Joint Services Electronics Program (N00014-96-J-1270) and Semiconductor Research Corp. (SRC96DP109).

ISLPED 1996 Monterey CA USA
0-7803-3571-8/96/$5.00 © 1996

can be drawn for a statistical property (mean) without assuming any particular distribution function or parameter. With use of this technique, the average power dissipation can be estimated with high accuracy and robustness by only analyzing the power sample data with distribution-independent statistics. Owing to this feature, the proposed technique can be applied to circuits with arbitrary power distributions.

The rest of the paper is organized as follows. In Section II we show how to express the power consumption of a circuit as a random variable so that statistical techniques can be applied. In Section III, first we use the properties derived from the distribution-independent order statistics to construct a confidence band of the cumulative distribution function (*cdf*). Next, we show how to develop upper and lower bounds of the average power using the concept of confidence band. Based on the distribution-independent bounds, we design a stopping criterion to terminate the random simulation when the bounds satisfy the user-specified accuracy and confidence level. We implemented the proposed technique and tested it on a set of benchmark circuits. The results are reported in Section IV with discussions and comparisons with those obtained from the approach in [1], followed by the concluding remarks in Section V.

II. Problem Formulation

Consider a digital circuit with n primary inputs. An *input pattern* of the circuit is a vector composed of 0's and 1's to assign a value to each primary input. Input pattern can be treated as a random variable $\boldsymbol{V}$ and generated by an input generation machine which can take into account spatial and/or temporal correlations among the primary inputs, such as those suggested in [4] and [5].

In static CMOS circuits, a gate dissipates power only when the gate output switches due to logic state transition at primary inputs. Here the power due to leakage currents is ignored. Without loss of generality, we assume that the logic state of primary inputs switch at the same time or remain unchanged. The power dissipation of the circuit with m gate output nodes is a function of the two consecutive input vectors $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ and can be expressed as:

$$\boldsymbol{P} = \frac{V_{DD}^2}{2T} \sum_{i=1}^{m} C_i \boldsymbol{n}_i(\boldsymbol{V}_1, \boldsymbol{V}_2), \qquad (4)$$

where the symbols have the same meanings as those in (1) except T. In (1), T is a time interval with arbitrary length; in (4) T is set such that during one unit of T any primary input can only switch once. If the circuit is embedded in a synchronous environment, T is the clock cycle time. Since $\boldsymbol{P}$ is a function of $\boldsymbol{n}_i$, $i = 1, \cdots, m$, it is also a random variable and possesses a distribution function. It should be noted that since the number of transitions at each gate output can only take nonnegative integers, $\boldsymbol{n}_i$ has a discrete distribution. According to (4), so does $\boldsymbol{P}$. For practical circuits, the sample space of $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ is usually large enough so that the difference between two adjacent observable values of $\boldsymbol{P}$ is small. Therefore, we can assume that $\boldsymbol{P}$ has a continuous distribution function [6].

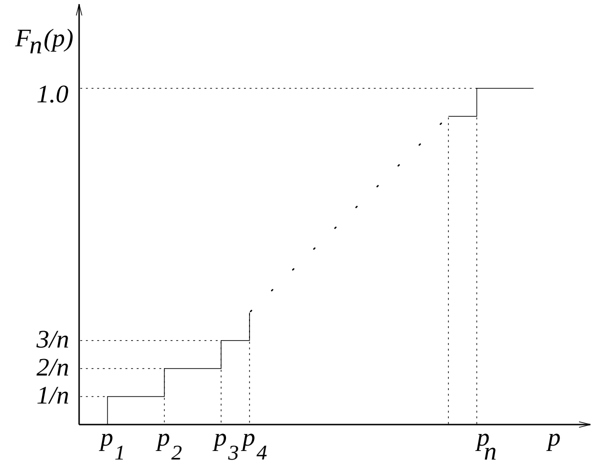

Figure 1: Construction of sample *cdf* $F_n(p)$.

III. Estimation of Average Power

A. Order Statistics

For a given circuit, let $F(p)$ be the *cdf* of $\boldsymbol{P}$. Suppose that the random variables $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_n$ form a random sample of $F(p)$, where $\boldsymbol{P}_1 < \boldsymbol{P}_2 < \ldots < \boldsymbol{P}_n$, and let $p_1, \ldots, p_n$ denote the observed values of $\boldsymbol{P}_1 \ldots, \boldsymbol{P}_n$. A sample *cdf* $F_n(p)$ is constructed from the values $p_1, \ldots, p_n$ such that for any p $(0 < p < \infty)$ the value of $F_n(p)$ represents the proportion of observed values in the sample which is less than or equal to p, as depicted in Fig. 1. $F_n(p)$ can be regarded as the *cdf* of a discrete distribution which assigns probability $1/n$ to each of the n values $p_1, \ldots, p_n$. For the i th order statistic $\boldsymbol{P}_i$, we define a random variable $\boldsymbol{Z}_i = F(\boldsymbol{P}_i)$. It is noteworthy that the distribution of $\boldsymbol{Z}_i$ is independent of that of $\boldsymbol{P}_i$. The pdf $g_i(z)$ of $\boldsymbol{Z}_i$ is [7]

$$g_i(z) = \frac{n!}{(i-1)!(n-i)!} z^{i-1}(1-z)^{n-i}. \qquad (5)$$

After some algebraic manipulation, we can express the *cdf* $G_i(z)$ of $\boldsymbol{Z}_i$ recursively:

$$G_i(z) = G_{i-1}(z) - \frac{n!}{(i-1)!(n-i+1)!}(1-z)^{n-i+1} z^{i-1}. \qquad (6)$$

Using (6), for $\boldsymbol{Z}_i$ we can find its $1-\alpha$ confidence interval $[z_i^{min}, z_i^{max}]$ simply by solving the equations $G_i(z) = \alpha/2$ and $G_i(z) = 1-\alpha/2$. Let $\boldsymbol{P}_j$, $\boldsymbol{P}_{j+1}$ be the j th and $j+1$ th smallest values in a sample of size n and $[z_j^{min}, z_j^{max}]$ $[z_{j+1}^{min}, z_{j+1}^{max}]$ be the corresponding $1-\alpha$ confidence interval of $\boldsymbol{Z}_j = F(\boldsymbol{P}_j)$ and $\boldsymbol{Z}_{j+1} = F(\boldsymbol{P}_{j+1})$, respectively. For random variable $\hat{\boldsymbol{P}}_j$, $\boldsymbol{P}_j < \hat{\boldsymbol{P}}_j < \boldsymbol{P}_{j+1}$, the following relation always holds because F is a non-decreasing function:

$$F(\boldsymbol{P}_j) \le F(\hat{\boldsymbol{P}}_j) \le F(\boldsymbol{P}_{j+1}), \qquad (7)$$

Since $\Pr(F(\boldsymbol{P}_j) \ge z_j^{min}) = 1-\alpha/2$ and $\Pr(F(\boldsymbol{P}_{j+1}) \le z_{j+1}^{max}) = 1-\alpha/2$, the following $1-\alpha$ confidence interval is always valid for $\hat{\boldsymbol{P}}_j$:

$$\Pr(z_j^{min} \le F(\hat{\boldsymbol{P}}_j) \le z_{j+1}^{max}) \ge 1-\alpha. \qquad (8)$$

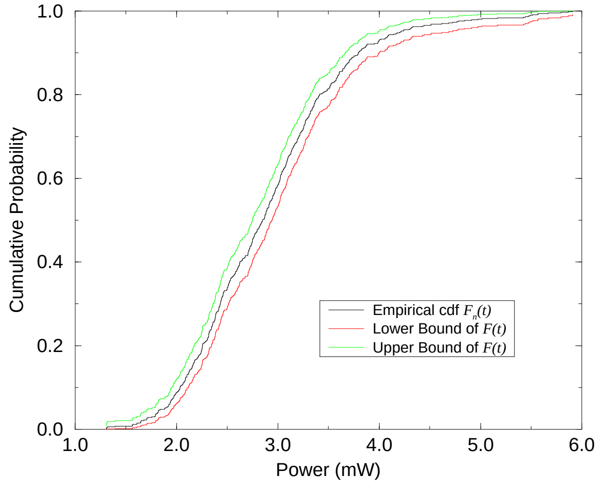

Figure 2: The empirical *cdf* and 99% confidence band of $F(p)$ of circuit C880 when sample size is 672.

(8) means that the confidence of $F(\hat{\boldsymbol{P}}_j)$ between z_j^{min} and z_{j+1}^{max} is *at least* $1-\alpha$. Given the observed values of the order statistics $p_1, \ldots, p_n$, let $p_0 = 0$ and $p_{n+1} = \infty$. For arbitrary p, $0 < p < \infty$, we can always find some index j, $j = 0, \ldots, n$, such that $p_j < p < p_{j+1}$. From (8), for random variable $\hat{\boldsymbol{P}}_j$, $p_j < \hat{\boldsymbol{P}}_j < p_{j+1}$, the $1-\alpha$ confidence interval of $F(\hat{\boldsymbol{P}}_j)$ is $[z_j^{min}, z_{j+1}^{max}]$. Therefore, a stairwise $1-\alpha$ confidence band of the unknown *cdf* $F(p)$ can be constructed by connecting the endpoints of every individual confidence interval, and can be expressed as:

$$\Pr(B_L(p) \leq F(p) \leq B_U(p)) \geq 1-\alpha, \tag{9}$$

where $B_L(p) = z_i^{min}$, $B_U(p) = z_{i+1}^{max}$, $p_i < p < p_{i+1}$, $i = 1, \ldots, n$. Fig. 2 illustrates the empirical *cdf* $F_n(t)$ as well as $B_L(p)$ and $B_U(p)$ of a benchmark circuit C880 when n is 672.

B. Bounds of Average Power and Stopping Criterion

The $1-\alpha$ confidence band (9) is distribution-independent and is the key result for finding the bounds of the average power and designing the stopping criterion to terminate the random simulation. To begin with, recall that the average power μ_p of a circuit is the mean of $\boldsymbol{P}$:

$$\mu_p = E[\boldsymbol{P}] = \int_0^\infty p f(p) dp, \tag{10}$$

where $f(p) = dF(p)/dp$. Let $u = F(p)$, by substituting $f(p)dp$ by du and p by $F^{-1}(u)$, the integration in (10) can be performed on the domain of variable u as

$$\mu_p = \int_0^1 F^{-1}(u) du. \tag{11}$$

By referring to Fig. 3, we can see that μ_p is just the area between $F^{-1}(u)$ and u axis and it is bounded

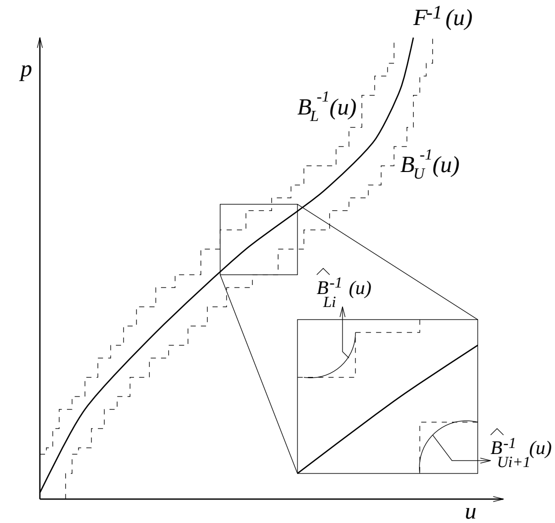

Figure 3: Lower and upper bound of average power.

from above by the area between $B_L^{-1}(u)$ and u, and from below by the area between $B_U^{-1}(u)$ and u, respectively. However, care must be taken here since both $B_L(p)$ and $B_U(p)$ are non-decreasing stairwise functions whose inverse functions do not exist. To make this statement mathematically correct, we argue that we can connect any two adjacent points z_i^{max} and z_{i+1}^{max} by some smooth function $\hat{B}_{Ui}(p)$ whose inverse function does exist, as shown in Fig. 3. To keep the bounds as tight as possible, we can make $\hat{B}_{Ui}(p)$ arbitrarily close to $B_U(p)$ and still let its inverse function exist. Suppose k is the parameter controlling the convexity of $\hat{B}_{Ui}(p)$, then $B_U(p)$ is the limiting function of $\hat{B}_{Ui}(p)$ as k approaches infinity:

$$\lim_{k\to\infty} \hat{B}_{Ui}(p, k) = B_U(p). \tag{12}$$

Following the similar procedure, we can make another invertible function $\hat{B}_L(p)$ to approximate $B_L(p)$. Note that since the confidence band encompassed by $B_U(p)$ and $B_L(p)$ are still embraced by $\hat{B}_U(p)$ and $\hat{B}_L(p)$, there is no loss in confidence level by making this approximation.

Since $\hat{B}_U(p)$ and $\hat{B}_L(p)$ are invertible, we can now bound μ_p by:

$$\int_0^1 \hat{B}_U^{-1}(u) du \leq \mu_p \leq \int_0^1 \hat{B}_L^{-1}(u) du, \tag{13}$$

Let $\mu_{pL} = \int_0^1 \hat{B}_U^{-1}(u)du$, $\mu_{pU} = \int_0^1 \hat{B}_L^{-1}(u)du$, and $\bar{\mu}_p$ be the sample mean of a sample of size n, (13) can be recast as

$$\frac{\mu_{pL} - \bar{\mu}_p}{\bar{\mu}_p} \leq \frac{\mu_p - \bar{\mu}_p}{\bar{\mu}_p} \leq \frac{\mu_{pU} - \bar{\mu}_p}{\bar{\mu}_p}. \tag{14}$$

As the simulation proceeds, more power data are collected and used to construct the sample *cdf* $F_n(p)$. By solving $G_i(z) = \alpha/2$ and $G_i(z) = 1-\alpha/2$ for

Circuit Name	No. Inputs	No. Outputs	No. Gates
C432	7	36	200
C499	41	32	439
C880	60	26	337
C1355	41	32	439
C1908	25	33	437
C2670	233	140	727
C3540	50	22	944
C5315	178	123	1200
C7552	207	108	1794
dalu	75	16	712
apex6	135	99	687
x3	135	99	646
vda	17	39	606
k2	45	43	1015
frg2	143	139	784
pair	173	137	1268
t481	16	1	586
rot	135	107	644
i8	133	81	910
i9	88	63	369
i10	257	224	2110

Table 1: Statistics of ISCAS85 and MCNC91 benchmark circuits.

$i = 1, \ldots, n$, we can see that the calculated $1-\alpha$ confidence band $\hat{B}_L(p)$ and $\hat{B}_U(p)$ will become increasingly closer towards the the real *cdf* $F(p)$. As a result, the calculated bounds of average power $\bar{\mu}_{pL}$ and $\bar{\mu}_{pU}$ will approach $\bar{\mu}_p$. For a desired percentage error ϵ and confidence level $1-\alpha$ specified up-front by the user, the power simulation can be stopped when the following criterion is satisfied:

$$\max\left[\frac{\bar{\mu}_p - \mu_{pL}}{\bar{\mu}_p}, \frac{\mu_{pU} - \bar{\mu}_p}{\bar{\mu}_p}\right] \leq \epsilon. \tag{15}$$

The first sample size n to satisfy (15) is defined as the *convergent sample size*. By (15), we can bound $|\mu_p - \bar{\mu}_p|/\bar{\mu}_p$ to the specified percentage error and guarantee the obtained sample mean $\bar{\mu}_p$ is close enough to μ_p. Since the derivation of (15) only employs the properties of the order statistics and requires no assumption on the distribution of $\boldsymbol{P}$, the stopping criterion is thus *distribution-independent* and can be applied to any type of circuits.

IV. Experimental Results and Discussion

The proposed average power estimation technique has been implemented as a distribution-independent power estimator (DIPE) and applied to a set of benchmark circuits to estimate the average power dissipation. The statistics of the test circuits are tabulated in Table 1. Table 2 shows the simulation results for the set of benchmark circuits. The circuits are assumed to operate at a clock frequency of 20MHz with 5V power supply. For each circuit, every primary input is assumed to be independent of one another and has a signal probability 0.5. The applied input patterns can thus be generated via a random number generator. The maximum error allowed was specified as 5% with 0.99 confidence. In Table 2, SIM is the average power obtained by calculating the sample mean of a power sample of size 1 million, and represents the best estimate we can get. LB(UB) is the lower(upper) bound of the average power, and $\bar{\mu}_p$ is the sample average power. LB, UB, and $\bar{\mu}_p$ are all obtained using the convergent sample size listed in the last column. For all the circuits, the technique produced very accurate estimate of the average power. Another distinguished property of the technique is that it is *dimensionally-independent* [1], i.e., the convergent sample size for a specified accuracy is independent of the circuit size. Thus, this technique is useful for average power estimation of very large circuits.

Circuit Name	SIM (mW)	LB (mW)	UB (mW)	$\bar{\mu}_p$ (mW)	Sample Size
C432	1.646	1.577	1.732	1.650	1248
C499	5.845	5.561	6.102	5.846	288
C880	2.907	2.777	3.055	2.911	672
C1355	5.843	5.550	6.036	5.841	288
C1908	5.357	5.100	5.629	5.366	576
C2670	7.331	6.981	7.654	7.334	384
C3540	15.275	14.531	16.009	15.278	672
C5315	21.357	20.313	22.149	21.317	320
C7552	33.309	31.727	34.919	33.362	672
dalu	4.737	4.514	4.976	4.740	1056
apex6	3.906	3.712	4.095	3.905	704
x3	4.251	4.047	4.442	4.254	448
vda	1.893	1.807	1.990	1.896	1248
k2	2.867	2.733	3.016	2.873	832
frg2	4.221	4.005	4.416	4.207	768
pair	9.465	9.018	9.469	9.469	320
t481	1.878	1.788	1.970	1.877	1216
rot	3.845	3.647	4.007	3.837	320
i8	6.704	6.391	7.037	6.702	1664
i9	5.857	5.584	6.157	5.866	1472
i10	22.313	21.257	23.371	22.340	544

Table 2: Power estimation results using DIPE.

Although the convergent sample size is not a function of circuit size, it shows certain correlation with the standard deviation σ of $\boldsymbol{P}$. If we plot the mean convergent sample size over 1000 simulation runs with σ normalized by the mean μ of $\boldsymbol{P}$, as shown in Fig. 4, it grows linearly as σ/μ increases for all the test circuits. The reason for this observation is that every power value in the sample is weighted differently when calculating the bounds of the average power. For simplification, we assume that $\hat{B}_{\{U,L\}}(p)$ are close enough to $B_{\{U,L\}}(p)$ so that the bounds can be obtained by calculating the areas between $B^{-1}_{\{U,L\}}(u)$ and the u axis:

$$\begin{aligned} \bar{\mu}_{pL} &= \sum_{i=2}^{n-1} p_{i-1}(B_U(i) - B_U(i-1)), \\ \bar{\mu}_{pU} &= \sum_{i=1}^{n-1} p_i(B_L(i+1) - B_L(i)). \end{aligned} \tag{16}$$

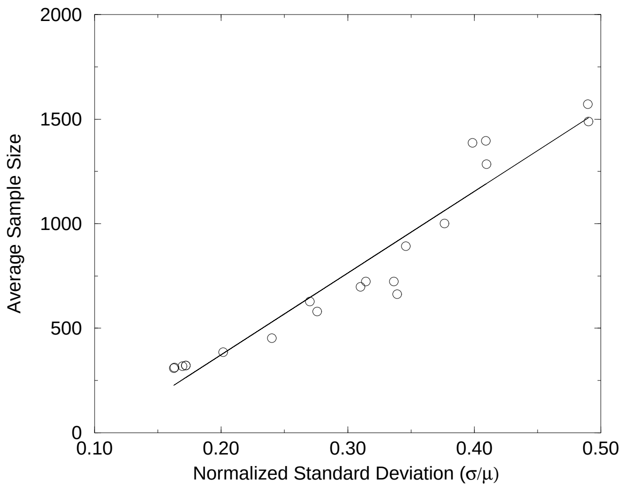

Figure 4: Correlation between average convergent sample size and normalized standard deviation.

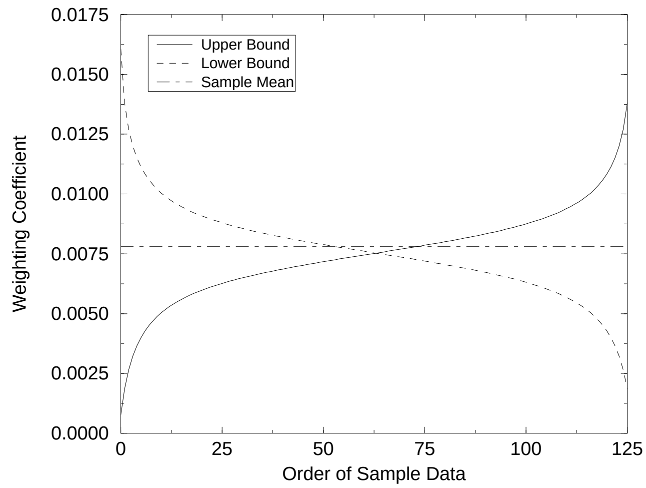

Figure 5: Weighting coefficients for the calculation of sample average power and the power bounds. Sample size n is 128.

Since $(B_U(i) - B_U(i-1))$ and $(B_L(i+1) - B_L(i))$ vary along with i, the bounds can be viewed as the nonlinearly weighted sum of the sample data, while the sample mean $\bar{\mu}_p$ is obtained by summing the sample data weighted by the same coefficient $1/n$. The variation of weighting coefficients as a function of order i is plotted in Fig. 5 when sample size is 128. It can be clearly seen that the smaller power values are more heavily weighted in estimating $\bar{\mu}_{pL}$ than $\bar{\mu}_{pU}$, while the larger power values are more heavily weighted in estimating $\bar{\mu}_{pU}$ than $\bar{\mu}_{pL}$. The sample data ordered in between have approximately equal weighting coefficients. Because sample data are unequally weighted, when evaluating the sample mean and the bounds, the standard deviation plays an important role in deciding the sample size for the bounds to converge. To understand this, suppose that at some point of the simulation, the current sample size is s and q new sample data are collected from the simulation which are smaller than the current sample mean. Let w_u, w_{av}, w_l denote the weighting coefficients of $B_U(p)$, $F_n(p)$, and $B_L(p)$, respectively. Since $w_u < w_{av} < w_l$, compared to the old bounds, the new upper bound will be closer to the new sample mean while the new lower bound will be farther from the new sample mean. The situation when the collected power data are larger than the current sample mean can be discussed similarly. On the other hand, if the newly added data are close to the current sample mean, they will be approximately equally weighted in estimating all three values, and push the data of extreme values away from the middle of the empirical *cdf* to let them be more lightly weighted. Consequently, both new bounds are closer to the new sample mean.

Based on the above observations, we can understand why the convergent sample size is proportional to the normalized standard deviation σ/μ of $\boldsymbol{P}$. If σ/μ is large, relatively the sample data tend to spread over a wider value range so that it is more difficult for $\bar{\mu}_{pL}$ and $\bar{\mu}_{pU}$ to converge. If σ/μ is small, the sample data tend to focus on a limited value band and convergence will be faster.

For comparison, we implemented the power estimation technique McPower [1] which is based on the central limit theorem and applied it to the same set of test circuits. The accuracy specification is the same as Table 2. In our implementation, an average power sample is obtained by taking the mean of a sample of size 32, instead of 50 as used in [1]. The difference in sample size, however, should not interfere the comparison because 1) it is explicitly assumed in [1] that the sample mean is normally distributed for any sample size; and 2) a rule-of-thumb sample size for the sample mean obtained from any distribution to be at least approximately normal is about 30 [8]; hence 32 is used as an appropriate sample size. By using the same accuracy specification, we compared the performance of the two techniques. The comparison results collected from 1000 simulation runs are listed in Table 3. In this table, Min, Max, and Avg represent the minimum, maximum, and average sample size used during the 1000 simulation runs, respectively; Err shows the percentage of the runs violating the accuracy specification. Since the confidence level is specified as 0.99, for 1000 runs the error percentage is at most 1% if the assumption that the sample mean is normally distributed is valid. As shown in Table 3, the error percentage of McPower is more than 1% for 9 of 21 test circuits, implying that the assumption is not generally valid. On the contrary, for all the simulation runs conducted for all the benchmark circuits, no error is detected from the the results generated by DIPE. It should be noted that the estimation results reported here are based on the assumption that the primary input signals are spatially and the input patterns are temporally uncorrelated. If such correlations exist, which usually happens to practical circuits, it is possible that the power distribution will further deviate from normal. In such cases, the estimation error percentage of McPower is expected to be even higher.

The robustness and accuracy of DIPE come from requiring larger sample size for the average power bounds to satisfy the accuracy specification and to stop the

Circuit	McPower				DIPE			
Name	Min	Max	Avg	Err	Min	Max	Avg	Err
C432	64	768	376	1.3	992	1696	1386	0.0
C499	64	320	168	0.0	256	352	311	0.0
C880	64	512	265	0.8	480	960	628	0.0
C1355	64	352	170	0.0	256	352	311	0.0
C1908	64	384	202	0.2	480	704	580	0.0
C2670	64	448	195	0.0	352	448	386	0.0
C3540	64	608	304	0.5	608	800	698	0.0
C5315	64	320	175	0.1	288	352	320	0.0
C7552	64	768	380	1.7	608	704	662	0.0
dalu	64	832	367	2.6	800	1120	1008	0.0
apex6	64	736	366	1.4	672	800	725	0.0
x3	64	512	230	0.5	416	512	453	0.0
vda	64	928	464	2.6	1120	1664	1395	0.0
k2	64	704	354	2.4	704	1056	893	0.0
frg2	64	736	319	0.9	576	896	725	0.0
pair	64	320	170	0.3	288	352	317	0.0
t481	64	864	455	2.2	1056	1472	1284	0.0
rot	64	352	167	0.0	256	384	323	0.0
i8	64	928	654	1.4	1344	1728	1571	0.0
i9	64	928	604	2.3	1312	1632	1487	0.0
i10	64	480	247	0.9	480	608	537	0.0

Table 3: Performance comparison of DIPE and McPower from the statistics of 1000 simulation runs.

simulation. It is a natural consequence of distribution-independent statistical approach, since it implicitly takes into account all possible variations of distribution functions and cannot take advantage of properties belonging to any particular one of them. As a result, if the sample mean power does have a normal distribution, the estimation technique based on the properties of normal distribution will produce correct estimate with a smaller sample size. On the other hand, if the sample mean power is not normally distributed, such technique would not be able to achieve the desired accuracy and confidence. When applied to practical circuits, nevertheless, DIPE seems to be more favorable because the power distribution of a circuit is usually not available at the time of average power estimation. Consequently it will be hard to justify the validity of any assumption on the functional form of power distribution. Since our approach is distribution-independent and can still produce an accurate power estimate with comparably low computational cost, it is suitable to be used as a reliable power estimator.

It should be noted, though, that according to the central limit theorem, no matter what the distribution of $\boldsymbol{P}$ is, as the sample size approaches infinity, the limiting distribution of the sample mean is normal. In such case, (3) can be used for average power estimation. However, the issue about how to decide a "large enough" sample size so that (3) can be used still remains unsolved. It is also expected that such sample size will be circuit dependent. For example, from Table 3, for circuits like C499 and C1355, a sample size of 32 is large enough for the central limit theorem to hold, while it is not the case for circuits *dalu* and *vda*. As there is no way to decide the sample size beforehand, our approach provides an appealing alternative that user does not have to have any knowledge about the circuit to obtain the correct average power estimate.

V. Conclusion

We have proposed a novel statistical technique for estimation of average power dissipation of digital circuits. By using the properties derived from the order statistics, we can construct a $1 - \alpha$ confidence band of the unknown *cdf* $F(p)$. An upper and lower bound of the average power can be obtained from the confidence band and used to design a stopping criterion to terminate the simulation. Since it is distribution-independent, the technique can be applied to any type of circuits. Experimental results show that it is much more robust than other statistical techniques and yet the computational cost is still very low.

REFERENCES

[1] R. Burch, F. N. Najm, P. Yang, and T. Trick, "A Monte Carlo approach for power estimation," *IEEE Trans. VLSI Systems,* vol. 1, no. 1, pp. 63-71, 1993.

[2] M. Xakellis and F. Najm, "Statistical estimation of the switching activity in digital circuits," *31st Design Automation Conf.*, pp. 728-733, 1994.

[3] T.-L. Chou and K. Roy, "Statistical estimation of sequential circuit activity," *IEEE/ACM Int. Conf. Computer-Aided Design,* pp. 34-37, 1995.

[4] R. Marculescu, D. Marculescu, and M. Pedram, "Efficient power estimation for highly correlated input streams," *32nd Design Automation Conf.*, pp. 628-634, 1995.

[5] J. Monteiro and S. Devadas, "Techniques for the power estimation of sequential logic circuits under user-specified input sequences and programs," *Int. Symp. Low Power Design,* pp. 33-38, 1995.

[6] A. Hill, C.-C. Teng, and S.-M. Kang, "Simulation based maximum power estimation," *1996 Int. Symp. Circuit and Systems.*

[7] R. Hogg and A. Craig, *Introduction to mathematical statistics,* 5th ed. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[8] R. A. Johnson, *Miller and Freund's probability and statistics for engineers,* 5th ed. Englewood Cliffs, NJ: Prentice Hall, 1994.