41st Southeastern Symposium on System Theory
University of Tennessee Space Institute
Tullahoma, TN, USA, March 15-17, 2009

Computing Bounds on Dynamic Power Using Fast Zero-Delay Logic Simulation

Jins D. Alexander and Vishwani D. Agrawal

Auburn University, Department of Electrical and Computer Engineering, Auburn, AL 36849, USA

alexaji@auburn.edu, *vagrawal@eng.auburn.edu*

Abstract

To consider process variation, we model gates with given lower and upper bounds on delays. For given input vectors, we first find logic transitions using zero-delay simulation. Our algorithms then determine the ambiguity (transient) interval, and maximum and minimum numbers for possible transitions. Computing these for all gates requires a linear-time analysis of each vector-pair. Weighting with node capacitances estimates lower and upper bounds on dynamic power. Results compare favorably with power analysis using Monte Carlo simulation, which requires significantly more computing resources. Bounded variations for node capacitances and leakage, not used in this work, are suggested for future investigation.

Keywords: *Dynamic power, digital circuits, process variation, bounded delay simulation, power estimation.*

I. Introduction

The dynamic power consumed by a digital CMOS circuit depends on logic and glitch transitions, the latter being a function of delays in the circuit. We often take fixed (nominal or worst-case) delay values for design and analysis. This is not correct for two reasons. First, the delay of a gate changes depending on signal states, device temperature, power supply fluctuations, and interconnect coupling noise. Second, and this is relevant to today's nanoscale technologies, there are wider process variations.

A. Motivation

To deal with the variability, significant advances have been made in logic simulation, timing analysis and delay testing areas with the use of the *bounded delay* [22] model. The delay of a gate is thus modeled as a range, typically known as (min, max). In the present work, we adopt the bounded delay model to develop a dynamic power analysis methodology. Although we have only focused on the delay variations, other parameters such as capacitance and leakage current may also be similarly treated in the future. The analysis may also be extended to include leakage power.

II. Background

Computationally efficient power estimation techniques often use selected random vectors or propagate signal probabilities [18]. A Monte Carlo approach involves applying randomly generated vectors and monitoring power until a desired confidence level is obtained [6]. Several approaches have been proposed for finding the peak power. Statistical and deterministic approaches [6], [23] have been used as opposed to traditional simulation-based methods to speed up estimation. They include generating test vectors that can cause higher power consumption [24]. In this method the authors assign transitions to gate nodes based on fan-outs and justify them by backtracking to primary input. Another technique includes developing state diagram to depict state to state transitions for estimating the energy consumption [21]. Genetic algorithm (GA) has been used to generate patterns that can simulate a high power consumption based on the fitness property of current vector [12], [15]. This is similar to the weighted activity function based approach of estimating patterns that maximize power consumption [10].

Some of the above methods use a zero-delay model and hence do not consider glitches. Others use fixed delay approaches and do not consider process variation. Effect of process variation on power has been discussed [9], [17] mainly for optimization and reduction of leakage power. A possible way to deal with process variation is to use a Monte Carlo approach [23]. Here, we take the variations between individual dies into account and consider delays to be random variables. We then run Monte Carlo simulations to get a high quality simulated maximum power sample

978-1-4244-3325-4/09/$25.00 ©2009 IEEE

from which by statistical techniques we estimate the mean and variance of maximum power. Due to the number of simulations required to have a reasonably accurate estimate, this technique is quite time consuming.

Process variation has been considered in logic-level simulation, critical path delay and timing analysis through delay modeling of gates. Both statistical [3], [14] and bounded delay models [4], [5], [7] have been used in this respect.

We use the zero-delay logic simulation, which is the fastest mode for simulating a given set of vectors. It, however, determines only the steady-state logic transitions and completely ignores the glitches. This result, therefore, gives an absolute lower bound on the dynamic power [19]. In reality, the number of glitches at a gate output depends upon all transitions (including glitches) arriving at the inputs of the gate, the arrival times of those transitions, and the inertial delay of the gate. The arrival times of transitions at the gate inputs depend upon the delays of other gates on the paths from primary inputs. Assuming that process variation makes all delays random variables, power estimation adds another dimension to the Monte Carlo method. Thus, each vector-pair should be resimulated a large number of times, every time with a different sample of delays for all gates. Our novel analysis completely eliminates such repetitions. Instead, gate delays are modeled as minimum and maximum bounds and the zero-delay result is analytically augmented to determine the minimum and maximum number of transitions each signal can have [2]. We obtain very similar results as would be obtained by Monte Carlo simulation but with significantly reduced computation.

III. Defining Signal Delays

In the bounded delay model, each gate is assigned the lower and upper bounds for delays, also called the *min-max* delay specification [4], [5], [7], [8], [11], [16], [20], [22].

To represent a signal in the bounded delay model, we use the term *ambiguity interval* for the duration of uncertainty where we cannot deterministically tell exactly when the signal transitions within that interval. The ambiguity intervals are shown in Figure 1 and are described as:

- EA is the earliest arrival time for signal.
- LS is the latest stabilization time for signal.
- IV is the initial value of signal.

Fig. 1: Ambiguity intervals of a rising signal (a) and a signal with steady logic state (b), respectively.

- FV is the final value of signal.

Our analysis follows the events that occur at gate outputs in a circuit after a change, typically on application of a new vector, occurs at primary inputs. We determine the ambiguity (transient) interval for the signal at the output of a gate from (1) bounded delay of the gate and (2) steady-state signal values, and the corresponding *initial values (IV)* and *final values (FV)*. We will assume that the primary inputs change at deterministic times synchronized with a clock. In general, however, input change ambiguities can be specified and treated in a similar manner as described here.

All times defined here are with reference to the start of the clock period. Because all gates are assumed to have delays within their respective specified (min, max) bounds, a typical gate inputs go through transient intervals before settling to some final values (FV). In the bounded delay formulation we do not precisely know how many transitions each signal makes. The ambiguity interval is defined as the interval from the earliest arrival time (EA) to the latest stabilization time (LS).

For clarity of discussion, we will use the example of an AND gate with four inputs as shown in Figure 2. Borrowing from the literature [4], we define $EAdv$ as the earliest arrival time of a signal that causes the input of the gate to change from controlling value (e.g., 0 for AND gate) to non-controlling value (1 for AND gate). Also, $LSdv$ is the latest stabilization time of an input signal changing from dominant value to non-dominant value. Similarly, $EAsv$ and $LSsv$ are the earliest arrival time and the latest stabilization time, respectively, for an input signal changing from non-dominant to dominant values. It should be noted that for $EA = \infty$ and $LS = -\infty$, the output is defined as having no ambiguity region or is in steady state condition. The following result, given here without elaboration, determines the output ambiguity interval.

The ambiguity interval (EA, LS) for the output signal

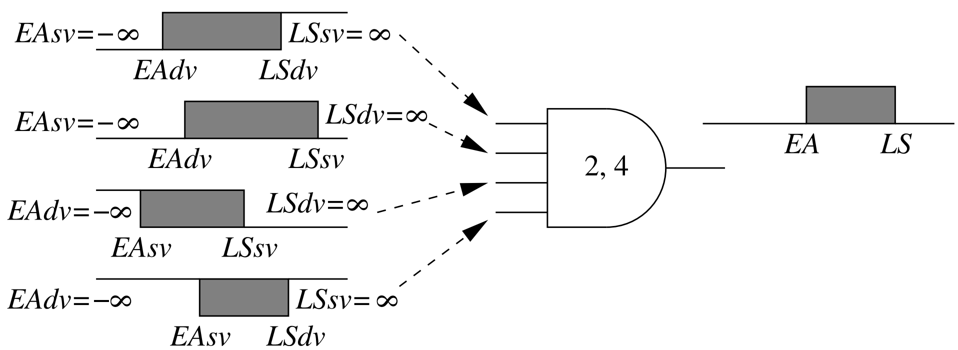

Fig. 2: A four-input AND gate with delay bounds (2, 4). Shaded regions are ambiguity intervals.

of a logic gate is determined from the ambiguity intervals of input signals, their pre-transition and post-transition steady-state values, and minimum and maximum gate delays, as follows [2]:

For all gate inputs i,

$$\begin{aligned} E1 &= maximum\{EAdv(i)\} && (1)\\ E2 &= minimum\{EAsv(i)\} && (2)\\ L1 &= minimum\{LSsv(i)\} && (3)\\ L2 &= maximum\{LSdv(i)\} && (4)\\ EA' &= maximum\{E1, E2\} && (5)\\ LS' &= minimum\{L1, L2\} && (6) \end{aligned}$$

Then

$$(EA, LS) = \begin{cases} \{EA' + mindel, LS' + maxdel\} \\ \qquad\qquad \text{if } (LS' - EA') \geq maxdel \\ \\ \{\infty, -\infty\} \qquad \text{if } (LS' - EA') < mindel \end{cases}$$

where the inertial delay of the gate is bounded as $(mindel,\ maxdel)$.

In general, the output of a gate can have multiple ambiguity regions separated by deterministic signal values as we will demonstrate in the following discussion. In that case, each ambiguity region as well as each deterministic interval will be affected by the inertial filtering caused by *mindel*. For simplicity, we can take a pessimistic view by combining all possible ambiguity regions into one.

Once the ambiguity interval (EA, LS) is determined, the steady-state values allow a straightforward conversion to the detailed signal specification. For the example of Figure 2, at the output of the AND gate, $EA = EAdv$ and $LS = LSsv$. Also notice when IV takes dominant (non-dominant) value, $EAsv$ $(EAdv)$ $= -\infty$. Similarly, when FV takes dominant (non-dominant) value, $LSdv$ $(LSsv) = \infty$.

Assuming that primary inputs (PIs) are glitch-free, the minimum number of transitions at a PI can be either 0 or 1. Consider a gate for which the ambiguity intervals and the minimum and maximum numbers of transitions, *mintran* and *maxtran*, respectively, for fan-ins are known. We can then estimate the value of *maxtran* at the output. We consider: (1) causality - an output transition must be caused by an input transition, and (2) filtering - gate inertia can filter out transitions that are closer in time than the gate delay. In the absence of detailed information, we assume that the transitions at a fan-in are evenly spaced within the ambiguity interval.

Agrawal et al. [1] have derived two upper bounds for the number of events possible at the output of a gate. However, in their derivation, neither ambiguity intervals nor the initial (IV) and final (FV) output values were considered. In this work, we use an improved bound [2].

The maximum number of transitions is defined as the minimum of the two upper bounds [2]:

$$maxtran \quad = \quad minimum(Nd, N) \qquad (7)$$

where *Nd* is the maximum number of transitions permitted by the gate inertial delay and *N* is the sum of all transitions present at the gate input.

We calculate Nd as the largest number of transitions that can be accommodated in the ambiguity interval given by the gate delay bounds and the output IV and FV values. We consider the filtering of glitches by gate inertia. Note that most transitions can be accommodated if they are evenly spaced over the output ambiguity interval with a spacing just equal to the inertial delay. We consider the following cases:

1) If the output has a static hazard, then we allow an

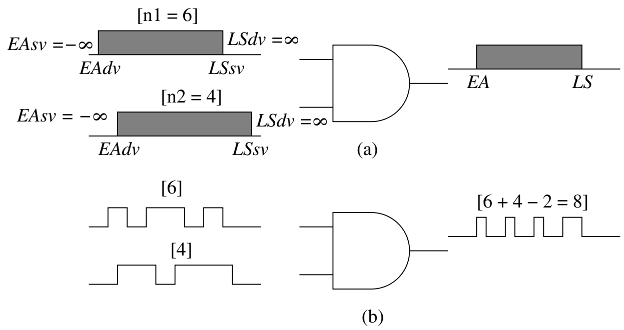

Fig. 3: Effect of modification factor k on the second upper bound.

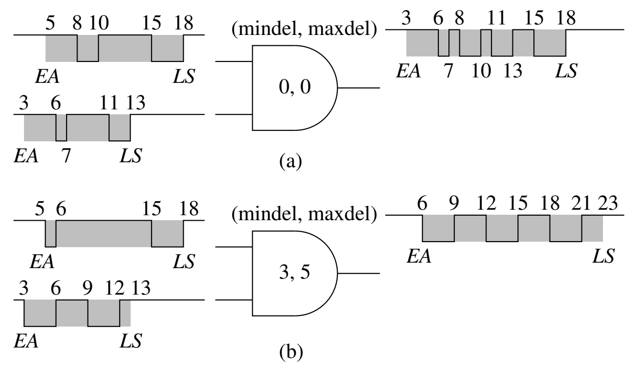

Fig. 4: Filtering of transitions in a two-input AND gate.

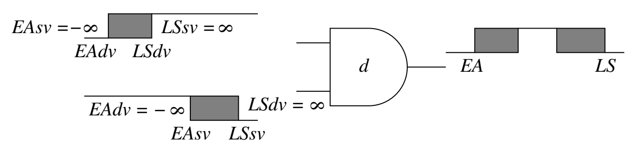

Fig. 5: Estimating lower bound on output transitions of a 2-input AND gate.

even number of transitions determined by $tpd \times (2n-1) \leq LS - EA$, and the number of transitions is given by $2n$, where tpd is the gate delay given by the minimum delay bound, n is the number of hazards that can possibly be accommodated in the ambiguity interval, LS is the latest stabilization time and EA is the earliest arrival time for the output signal, as given earlier.

2) Similarly, for an output signal with a dynamic hazard we would get an odd number of transitions determined by $tpd \times 2n \leq LS - EA$, and the number of transitions is given by $2n - 1$.

In our analysis, we have taken $maxdel$ in the estimation of Nd.

Next, to find the second upper bound (N), we modify the sum $Nsum$ of all input transitions as:

$$N \quad = \quad Nsum - k \tag{8}$$

where k = 0, 1, or 2 for a 2-input gate and is determined by the ambiguity regions and the IV and FV values of inputs. The procedure is explained by the example of Figure 3. In (a), for the given input transitions $n1$ = 6 and $n2$ = 4, the output cannot have the total sum of input transitions. This is because, when we consider the sum, in the case of a two signals going from a controlling value to a non-controlling value, only one of the two transitions should be counted. Thus. we see that a correction factor k is required, which would give us a $maxtran = 8$ for the example in (a). The example in (b) is a possible deterministic signal representation of the same.

The two upper bounds are illustrated in Figure 4. In (a) the gate is assumed to have zero delay. Thus, based upon the above discussion we get the maximum number of transitions from the first upper bound, Nd = ∞ and the second upper bound gives us N = 8. The minimum of the two gives us the final $maxtran$ = 8 value. However when we consider the gate delay bounds (3, 5) as in (b) we find that 8 transitions can never occur within the output ambiguity bounds of (6, 23). Using the above principles we get Nd = 6 and N = 8. In fact, the maximum possible $maxtran$ is only 6.

A rather pessimistic lower bound on the minimum number of transitions, *mintran*, can be found from the steady-state values at the gate output. This bound is 0 or 1 depending upon whether the output values before and after transients, i.e., IV and FV, are same or different. This has been used as the condition for minimum glitch power design [1]. However, when there are split ambiguity regions, we can obtain a tighter lower bound. Without discussing the details, which can be found elsewhere [2], we give an example.

In Figure 5, there are at least two essential signal changes that must occur within the output ambiguity interval. Thus, there will always be a hazard in the output as long as:

$$(EAsv - LSdv) \quad \geq \quad maxdel \tag{9}$$

where $maxdel$ is the maximum delay of the gate producing the transient. In this case $mintran$ is not 0, but is 2.

TABLE I: Per vector energy consumption in picojoule in benchmark circuits for 1000 random vectors by Monte Carlo simulation of 1000 sample circuits and bounded delay analysis. (*Sun UltraSparc 10 with 4GB Shared Memory.)

Circuit name	Monte Carlo simulation - *picojoule per vector*				Bounded delay analysis - *picojoule per vector*			
	Minimum	Maximum	Average	CPU s	Minimum	Maximum	Average	CPU s*
c880	1.086	10.847	4.340	298.26	1.080	11.140	4.240	0.34
c1355	3.606	13.577	7.310	423.69	3.600	20.150	10.928	0.59
c1908	4.870	29.470	15.580	840.85	4.590	57.050	17.750	0.69
c2670	8.470	51.190	24.390	1452.24	8.390	59.010	23.200	1.09
c3540	6.036	66.660	30.770	1810.18	5.970	96.180	35.100	1.39
c5315	29.810	91.100	56.41	3435.53	23.030	113.200	55.610	2.14
c6288	45.360	194.860	129.700	20944.53	11.840	406.340	153.710	2.60
c7552	35.050	146.120	82.790	5834.87	29.470	196.310	82.180	3.34

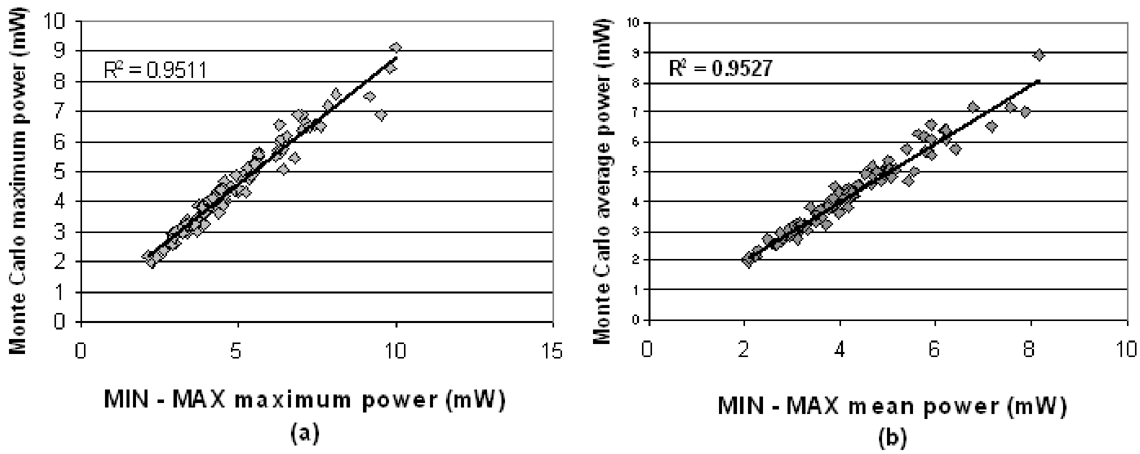

Fig. 6: Monte Carlo simulation versus bounded delay analysis for c880. Each point represents one vector-pair. One hundred sample circuits with $nominal \pm 20\%$ delay variation were simulated and for each vector-pair (a) maximum and (b) average power was determined.

IV. Results

Our results consist of power analysis of ISCAS85 benchmark circuits for 1000 random vectors. The circuits were implemented using TSMC025 2.5V CMOS library. Process variation was modeled by assuming 15% intra-die and 5% inter-die variations [13]. For illustrative purposes, standard size gate delay of 10ps and wire-load delay models were used to determine the nominal gate delays from which bounds (mindel, maxdel) were obtained by assuming a $\pm 20\%$ variation. It should be noted that any kind of variation model is suitable for our method. Node capacitances for the circuits were obtained from the Spice modeling files.

Table I gives two sets of data. The first set (columns 2-5) is from a Monte Carlo event-driven simulation. These results are for 1000 circuit samples. Each gate in a sample circuit was assigned delay using a random number uniformly distributed in its (mindel, maxdel) range. From the simulation of 1000 random vectors, we have listed the energy consumption (pJ) for two vectors consuming the least and the most energy, and the average for all 1000 vectors. We observe that the bounded delay analysis always gives lower minimum energy and higher maximum energy. This is expected. As we simulate more circuit samples, the Monte Carlo numbers drop in the minimum column and increase in the maximum column. Notably, to take the variability into account, an event-driven simulator is essentially used in the Monte Carlo analysis, which takes much more computing resources. The CPU times in Table I are for a Sun UltraSparc 10 with 4GB shared memory system.

We next conducted another Monte Carlo analysis using the event-driven simulator. This time, 100 sample circuits were generated for c880 by inserting random delays. Each sample circuit was simulated by the event-driven simulator for the same set of 100 random vectors. For each vector-pair, we obtained the minimum, average and maximum power. The maximum and average numbers are shown in Figure 6 against the corresponding vector-pair results from the bounded delay analysis. The values of R^2 shown on the regression graphs are *coefficients of determination* from Microsoft Excel, whose ideal fit value is 1.0. This shows a close correlation between our zero-delay simulation and bounded delay analysis with the traditional Monte Carlo procedures using the event-driven simulation.

V. Conclusion

The motivation of this research is to develop an efficient power estimation method with consideration of process variations. Our present target is dynamic power. We have used a bounded delay model and new results for bounds on gate transitions. This analysis has a linear-time complexity in the number of gates and is an efficient alternative to the Monte Carlo analysis. Presently, we can include leakage based on signal states that are obtained from inherent zero-delay simulation in this method. Our expectation for the future is to consider process variation in leakage as well. Besides, node capacitances that are considered to be fixed here, can also have process-dependent variation. We hope to investigate that in the future.

References

[1] V. D. Agrawal, M. L. Bushnell, G. Parthasarathy, and R. Ramadoss, "Digital Circuit Design for Minimum Transient Energy and a Linear Programming Method," in *Proc. 12th International Conf. VLSI Design*, Jan. 1999, pp. 434–439.

[2] J. D. Alexander, "Simulation Based Power Estimation for Digital CMOS Technologies," Master's thesis, Auburn University, Dept. of ECE, Dec. 2008.

[3] J. W. Bierbauer, J. A. Eiseman, F. A. Fazal, and J. J. Kulikowski, "System Simulation with MIDAS," *AT&T Tech. J.*, vol. 70, no. 1, pp. 36–51, Jan. 1991.

[4] S. Bose and V. D. Agrawal, "Delay Test Quality Evaluation Using Bounded Gate Delays," in *Proc. 25th IEEE VLSI Test Symp.*, 2007, pp. 23–28.

[5] S. Bose, H. Grimes, and V. D. Agrawal, "Delay Fault Simulation With Bounded Gate Delay Model," in *Proc. International Test Conf.*, 2007. Paper 26.3.

[6] R. Burch, F. N. Najm, P. Yang, and T. N. Trick, "A Monte Carlo Approach for Power Estimation," *IEEE Trans. VLSI Systems*, vol. 1, no. 1, pp. 63–71, Mar. 1993.

[7] S. Chakraborty and D. L. Dill, "More Accurate Polynomial-Time Min-Max Timing Simulation," in *Proc. 3rd International Symp. Advanced Research in Asynchronous Circuits and Systems*, Apr. 1997, pp. 112–123.

[8] S. Chakraborty, D. L. Dill, and K. Y. Yun, "Min-Max Timing Analysis and an Application to Asynchronous Circuits," *Proc. IEEE*, vol. 87, no. 2, pp. 332–346, Feb. 1999.

[9] H. Chang and S. S. Sapatnekar, "Full-Chip Analysis of Leakage Power Under Process Variations, Including Spatial Correlations," in *Proc. 42nd Design Automation Conf.*, June 2005, pp. 523–526.

[10] S. Devadas, K. Keutzer, and J. White, "Estimation of Power Dissipation in CMOS Combinational Circuits using Boolean Function Manipulation," *IEEE Trans. CAD*, vol. 11, no. 3, pp. 373–383, Mar. 1992.

[11] S. Hassoun, "Critical path analysis using a dynamically bounded delay model," in *Proc. 37th Design Automation Conf.*, 2000, pp. 260–265.

[12] M. S. Hsiao, E. M. Rudnick, and J. H. Patel, "An Estimator for Peak Sustainable Power of VLSI Circuits," in *Proc. International Symp. Low Power Electronics and Design*, 1997, pp. 178–183.

[13] S. P. Khatri, R. K. Brayton, and A. L. Sangiovanni-Vincentelli, *Cross-talk Noise Immune VLSI Design using Regular Layout Fabrics*. Boston: Kluwer Academic Publishers, 2001.

[14] K. N. Lalgudi, D. Bhattacharya, and P. Agrawal, "Architecture of a Min-Max Simulator on MARS," in *Proc. 6th International Conf. VLSI Design*, Jan. 1993, pp. 246–249.

[15] Y.-M. Liang, K.-T. Cheng, and A. Krstic, "Estimation of Maximum Power and Instantaneous Current Using a Genetic Algorithm," in *Proc. lEEE Custom Integrated Circuits Conf.*, May 1997, pp. 135–138.

[16] M. Linderman and M. Leeser, "Simulation of Digital Circuits in the Presence of Uncertainty," in *Proc. International Conf. Computer-Aided Design*, 1994, pp. 248–251.

[17] S. Mukhopadhyay and K. Roy, "Modeling and Estimation of Total Leakage Current in Nano-Scaled CMOS Devices Considering the Effect of Parameter Variation," in *Proc. International Symp. Low Power Electronics and Design*, Aug. 2003, pp. 172–175.

[18] F. N. Najm, "A Survey of Power Estimation Techniques in VLSI Circuits," *IEEE Trans. VLSI Systems*, vol. 2, no. 4, pp. 446–455, Dec. 1994.

[19] K. Roy and S. C. Prasad, *Low-Power CMOS VLSI Circuit Design*. Springer, 2000.

[20] S. Roy, P. P. Chakrabarti, and P. Dasgupta, "Bounded Delay Timing Analysis Using Boolean Satisfiability," in *Proc. 20th International Conf. VLSI Design*, Jan. 2007, pp. 295–302.

[21] J. H. Satyanarayana, K. K. Parhi, L. Song, and Y.-N. Chang, "Systematic Analysis of Bounds on Power Consumption in Pipelined and Non-Pipelined Multipliers," in *Proc. IEEE International Conf. Computer Design*, Oct. 1996, pp. 492–499.

[22] C. J. Seger, "A Bounded Delay Race Model," in *Proc. of the IEEE International Conf. Computer Aided Design*, Nov. 1989, pp. 130–133.

[23] C.-Y. Wang, T.-L. Chou, and K. Roy, "Maximum Power Estimation for CMOS Circuits Under Arbitrary Delay Model," in *Proc. IEEE International Symp. Circuits and Systems*, May 1996, pp. 763–766.

[24] C.-Y. Wang, K. Roy, and T.-L. Chou, "Maximum Power Estimation for Sequential Circuits Using a Test Generation Based Techniques," in *Proc. IEEE Custom Integrated Circuits Conf.*, Apr. 1996, pp. 229–232.