A Note on the Relationship Between Signal Probability and Switching Activity*

Qing Wu, Massoud. Pedram

Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, CA 90089, USA
Tel: 213-740-4480
Fax: 213-740-7290
e-mail: qingwu@zugros.usc.edu

Xunwei Wu**

Department of Electronic Engineering
Hangzhou University
Hangzhou, Zhejiang 310028, CHINA
Tel: +86-571-8071224 (ext: 2281)
Fax: +86-571-8070107
e-mail: xwu@ms.fudan.sh.cn

Abstract—In current probability calculation algorithms for power estimation, switching activity E_{SW} of a node is calculated from its signal probability p by the following simple relation: $E_{SW} = 2p(1-p)$. It is generally understood that this simple relationship holds under the temporal independence assumption for the node. This paper however shows that the above equation also gives the expected value of the transition activity in any sequence that satisfies the given signal probability (averaged over all such sequences). Therefore, this equation can be used to calculate the switching activity under more general conditions than previously thought.

I. Signal Probability and Transition Probability

In the power estimation of CMOS circuits, the dominant term is the power required to charge or discharge the capacitance is given by

$$P = 0.5 C_L V_{DD}^2 f_{CLK} E_{SW} \tag{1}$$

where C_L is the physical capacitance at the output of the node, V_{DD} is the supply voltage, f_{CLK} is the clock frequency, E_{SW} (referred to as the average switching activity) is the average number of output transitions per clock cycle $1/f_{CLK}$ [1]. For a node in a circuit, C_L, V_{DD}, and f_{CLK} are usually given, but E_{SW} is dependent on the input pattern and the circuit structure. To estimate the power dissipation of a circuit with high accuracy, a large number of input signal patterns should be simulated and the average value of E_{SW} calculated.

E_{SW} may be interpreted as the probability of having a transition at the node in a clock cycle. It may be thus calculated by a probability propagation algorithm whereby the transition probability of an internal circuit node is calculated from the signal probabilities of the primary input variables [2]. The signal probability, which was first used to study circuit testability [3], can be represented as follows: There is a binary sequence x (of length l), where m bits are logic 1. Then the probability that the signal is measured as 1 in a random clock cycle is m/l, and we can denote $P(x = 1) = m/l$. Obviously, there are n bits ($n = l - m$) which are logic 0 in the sequence. Therefore, the probability that the signal is measured as 0 in a random clock cycle is n/l, and we can denote $P(x = 0) = n/l$. We have $P(x = 1) + P(x = 0) = 1$.

In Eq. (1) we need transition probability rather than the signal probability for calculating the power dissipation. Thus we need discuss changes in the signal value during a clock. We refer to this as the behavior of the signal. For a signal x in the circuit if we denote its logic values before and after a clock transition as $x(t)$ and $x(t')$, respectively, four value combinations may occur as shown in Table 1, where a special quaternary variable $\bar{x}$ denotes the behavior of the signal. Its four values are (0, α, β, 1), where α and β represent two kinds of transition behaviors and 0,1 represent two kinds of holding behaviors. (Note: although they have the same forms as signals 0 and 1, their meanings are different.)

TABLE 1

Four Behaviors of a Signal

$\bar{x}$	$x(t) \rightarrow x(t')$		Behavior
0	0	0	0-holding
α	0	1	α-transition
β	1	0	β-transition
1	1	1	1-holding

An l-bit signal sequence can be transformed as an (l - 1)bit behavior sequence. Probability of various behaviors in this sequence can be derived by dividing the number of occurrences of the behavior of interest into (l - 1). (If l is large enough, $l - 1 \approx l$). They can be denoted as $P(x^0)$, $P(x^\alpha)$, $P(x^\beta)$, $P(x^1)$. The behavior probabilities are related to the signal probability as follows

$$P(x = 0) = P(x^0) + P(x^\alpha), \tag{2}$$

$$P(x = 1) = P(x^1) + P(x^\beta). \tag{3}$$

Alternatively, we will have

* This work was supported by DARPA under contract #F33615-95-C-1627.

** Xunwei Wu is presently on leave at Univ. of Southern California..

ASP-DAC '97
0-89791-851-7$5.00 © 1997 IEEE

$$P(x = 0) = P(x^0) + P(x^\beta), \quad (4)$$

$$P(x = 1) = P(x^1) + P(x^\alpha). \quad (5)$$

From Eqs.(2) -(4) we have

$$P(x^0) + P(x^\alpha) + P(x^\beta) + P(x^1) = 1, \quad (6)$$

$$P(x^\alpha) = P(x^\beta). \quad (7)$$

Equation (6) is obvious. Equation (7) is reasonable since the number of rising and falling transitions must be equal. If we use $P(x^{1/2})$ to represent the probability of both transitions, i.e. E_{SW} in Eq.(1), then we have

$$P(x^{1/2}) = P(x^\alpha) + P(x^\beta) = 2P(x^\alpha) = 2P(x^\beta).$$

Note that the transition probabilities cannot be derived from Eqs.(2) - (5). Therefore, we have to find other relationships between signal probability and behavior probabilities if only the input signal probability is known.

II. Formula for Calculating Transition Probability from Signal Probability

Consider two consecutive clock cycles. Assuming that signal values in the two cycles are independent of one another, we can write:

$$P(x^0) = P(x=0) \cdot P(x=0) = [P(x=0)]^2 \quad (8)$$

$$P(x^1) = P(x=1) \cdot P(x=1) = [P(x=1)]^2 \quad (9)$$

$$P(x^{1/2}) = P(x^\alpha) + P(x^\beta) = 2P(x=1) \cdot P(x=0) \quad (10)$$

In the above equations Eq.(10) provides the transformation from signal probability into transition probability. Since the use of probabilities to estimate power was first proposed in [4], we consider the transformation formula in Eq.(10) is a development of this earlier work.

A stationary relationship between signal probability and transition probability is given by Eq.(10). It implies that as long as m bits are 1 in an l-bit sequence, the number of transitions will be given by Eq.(8). If the number of (rising or falling) unidirectional transitions is denoted by k_0, we can write

$$P(x^{1/2}) = \frac{2k_0}{l-1} = 2\left(\frac{m}{l}\right) \cdot \left(\frac{l-m}{l}\right)$$

If the length l is large enough, we will have

$$k_0 = \frac{m(l-m)}{l} \quad (11)$$

However, the stationary relationship is difficult to accept since eqn.(10) does not hold for an arbitrary sequence. For example, consider a 12-bit sequence with six 1's and six 0's. The signal probability will be $P(x = 1) = P(x = 0) = 0.5$. The three sequences shown in Fig.1 are examples of such a sequence. Their first and last bits have the same value, that is, the numbers of rising and falling transitions in each sequence are equal. According to Eq.(8) we should have $P(x^\alpha) = P(x^\beta) = 0.25$. So the number of rising and falling transitions is 6 if we take $l - 1 \approx l = 12$. However, only the sequence in Fig.1(a) has this many transitions. For this reason Eq.(10) is often criticized as being applicable only under the assumption of temporal independence.

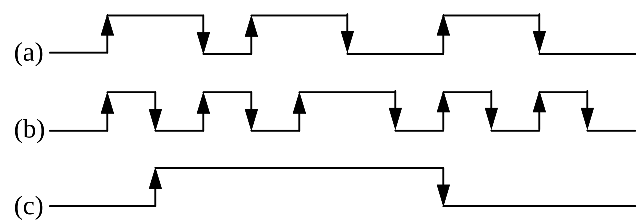

Fig. 1. Signal sequences ($l = 12$, $m = 6$)

In this paper we will show that Eq.(10) in fact gives the maximum likelihood estimate of the number of transitions in any random sequence of length l with m 1-bits.

III. Statistic of Sequences with Fixed Length and Signal Probability

There is a sequence x of length l , in which m bits are 1's and n bits are 0's, $m + n = l$. Therefore we have $P(x = 1) = m/l$ and $P(x = 0) = n/l$. For convenience we assume that the first and the last bits are the same, then in (l - 1) behaviors of the sequence, the number k of unidirectional (rising or falling) transitions are equal. Thus, the number of the transitions in the sequence is $2k$. Now we will answer the following questions:

(1) For the sequences with given l and m (n), how many k values can be observed, that is, what is the range, $k_{\min}$ and $k_{\max}$, for k?

(2) Given l and m (n), how many different sequences with exactly k transitions can be found?

If we denote this number by S_k we can calculate H, the total number of sequences with the given l and m (n), and k_{ave}, the average of unidirectional transitions, as

$$H = \sum_{k=k_{\min}}^{k_{\max}} S_k \quad (12)$$

$$k_{ave} = \frac{\sum_{k=k_{\min}}^{k_{\max}} k \cdot S_k}{H} \quad (13)$$

(3) Is there any difference between k_{ave} as calculated in Eq.(13) and k_0 as calculated in Eq.(11)?

For convenience we suppose the first bit and the last bit of the sequences are 0, as in Fig.1. Then we move all 1-cycles with their falling transitions together, and move all 0-cycles with their rising transitions together. For example, the sequence in Fig.1(a) can be decomposed into two parts shown as Fig.2(a) and (b). In Fig.2(a) we can at most arrange m falling transitions in the logic-1 zone, including the right most transition. However, in Fig.2(b) we can at most arrange (n - 1) rising transitions in the logic-0 zone. Therefore, number of the unidirectional transitions, k, can be from 0 to min(m, n-1). For example, for l = 100, m = 30, n = 70, k can take 1,2, …… 30.

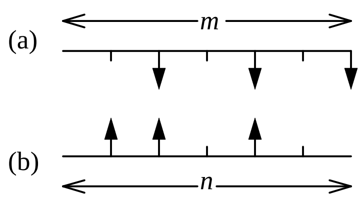

Fig. 2. Decomposition of signal sequence in Fig.1(a)

For the second question, we can analyze how many different sequences can be found for given l, m (n) and k by using the example in Fig.2. For Fig.2(a) this is the question about how many patterns can contain k-1 falling transitions at m - 1 positions in the logic-1 zone. (The last falling transition is fixed at the end of the zone.) The answer is $C_{m-1}^{k-1} = \frac{(m-1)!}{(m-k)!(k-1)!}$. On the other hand, for fig.2(b) the question is how many patterns we can contain k rising transitions at n - 1 positions in the logic-0 zone. The answer is $C_{n-1}^{k} = \frac{(n-1)!}{(n-k-1)!k!}$. Therefore, for the given l, m (n) and k we can find S_k different sequences:

$$S_k = C_{m-1}^{k-1} \cdot C_{n-1}^{k} \qquad (14)$$

Based on Eqs.(10),(11) we have

$$H = \sum_{k=1}^{\min(m,n-1)} S_k \quad \text{and} \quad k_{ave} = \frac{1}{H} \sum_{k=1}^{\min(m,n-1)} k \cdot S_k$$

Table 2 shows the sequence numbers, S_k, for l = 100, m = 30 and different k, k = 1,2, …… 30. The percentages in total sequences for l = 100, m = 30 are also given. The distribution curve is shown in Fig.3, where we find the peak appears at k = 21. Therefore we denote k_{peak} = 21. Based on Table 2, we can calculate the average number of unidirectional transitions as k_{ave} = 21.12 $\approx$ 21. Besides, for given l = 100, m = 30 we get k_0= 21 again! Note that we obtain k_{peak} and k_{ave} without any assumption about temporal independence.

TABLE 2

NUMBER OF SEQUENCES, S_K, FOR L = 100 AND M = 30

k	number of sequences S_k	S_k/H (%)
1	69	0.00 %
2	6.80e+4	0.00 %
3	2.13e+7	0.00 %
4	3.16e+9	0.00 %
5	2.67e+11	0.00 %
6	1.42e+13	0.00 %
7	5.12e+14	0.00 %
8	1.31e+16	0.00 %
9	2.43e+17	0.00 %
10	3.41e+18	0.00 %
11	3.65e+19	0.00 %
12	3.05e+20	0.00 %
13	2.01e+21	0.01 %
14	1.05e+22	0.07 %
15	4.40e+22	0.31 %
16	1.48e+23	1.04 %
17	4.05e+23	2.82 %
18	8.94e+23	6.24 %
19	1.60e+24	11.17 %
20	2.32e+24	16.16 %
21	2.70e+24	* 18.86 %
22	2.53e+24	17.63 %
23	1.88e+24	13.10 %
24	1.10e+24	7.64 %
25	4.93e+23	3.44 %
26	1.67e+23	1.16 %
27	4.09e+22	0.28 %
28	6.81e+21	0.05 %
29	6.88e+20	0.00 %
30	3.16e+19	0.00 %
Total	H = S S_k = 1.43e+25	100 %

Apart from m = 30, we also investigate m = 10, 20, 40, 50, 60, 70, 80, 90 and 100. For m = 100, the situation is similar to the case of m = 30, and the statistical results are listed in Table 3. Therefore, the transition probability derived from Eq.(8) has a new explanation: the calculated probability from given signal probability $P(x = 1)$ expresses not only the peak in the distribution of transitions, but also the average number of transitions. In fact we can replace $P(x^{1/2})$ in Eq.(8) by $P_{\text{ave}}(x^{1/2})$:

$$P_{ave}(x^{1/2}) = 2P(x=1)\cdot P(x=0)$$

Since the probability represents the average number of transitions in any sequence of length *l*, it can be reliably used to calculate the transition probability without any concern for presence or absence "temporal correlations". The calculated transition probability should be realized as the most probable and the average transition activity.

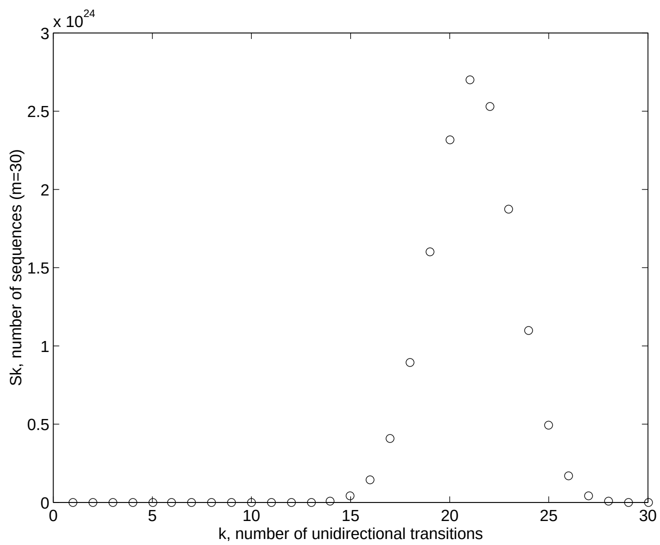

Fig. 3. Distribution curve of S_k

TABLE 3

AVERAGE UNIDIRECTIONAL TRANSITION NUMBER K_{AVE} AND PEAK VALUE K_{PEAK} OF SEQUENCE WITH L = 100 AND M = 10 $\longrightarrow$ 50

m	k_{ave}	k_{peak}	percentage of sequence number at k_{peak}
10	9.08	9	40.85 %
20	16.12	16	24.43 %
30	21.12	21	18.86 %
40	24.08	24	16.60 %
50	25.00	25	16.00 %
60	23.88	24	16.74 %
70	20.71	21	19.23 %
80	15.51	16	25.39 %
90	8.27	9	44.88 %

IV. DISCUSSIONS

To estimate the power dissipation in VLSI circuits with high accuracy, it is necessary to know the actual input pattern, which may be represented by detailed input sequences, by input signal probabilities, or by input transition probabilities. When some characteristic input sequences are given, the best way to estimate power is to input these sequences into a simulator and calculate the power by averaging results. If we extract the signal probability from the sequence and execute the probability algorithm to estimate power by using Eq.(10), the accuracy may be lost since Eq.(10) may not match the behavior of the given sequence. Therefore the deviation of results by using the SPICE simulator and the probabilistic simulator can be understood. However, as long as the loss is within a tolerable range, then the probability simulator may be used since it is significantly faster than actual simulation.

On the other hand, if the input signal probabilities are given we should take all sequences that satisfy the probability profile into account. Then we can use Eq.(10) to calculate the corresponding transition probability without being restrained by any temporal independence assumption. The calculated transition probability represents the average and the most probable value, which conforms to statistical principles. However, we should point out that it will be impossible to accurately compare the results with SPICE unless we simulate all sequences that have the signal probability profile and average them out.

Finally, if the input transition probabilities are given, we can bypass the signal probability propagation algorithm and use these transition probability directly.

REFERENCES

[1] M. Pedram, "Power minimization in IC Design: Principles and applications," *Transactions of ACM*, vol.1 , no. 1, pp.1-58, March, 1996.

[2] F. N. Najm, "A survey of power estimation techniques in VLSI circuits," *IEEE Transactions on VLSI Systems*, vol.2 , no.4, pp.446-455, 1995.

[3] K. P. Parker and E. I. McCluskey, "Probabilistic treatment of general combinational networks," *IEEE Transactions on Computers*, vol.24, no.6, pp.668-670, 1975.

[4] M. A. Cirit, "Estimating dynamic power constumptiuon of CMOS circuits," in *Proc. IEEE ICCAD*, pp.534-537, 1987.