Unit – II

Power Estimation

Activity Factor

The activity factor is a powerful and easy-to-use technique for reducing power. If a circuit can be turned off entirely, the activity factor and dynamic power go to zero.

Blocks are typically turned off by stopping the clock this is called *clock gating*.

When a block is on, the activity factor is 1 for clocks and substantially lower for nodes in logic circuits.

The activity factor of a logic gate can be estimated by calculating the switching probability. Glitches can increase the activity factor.

Clock Gating

Clock gating ANDs a clock signal with an enable to turn off the clock to idle blocks. It is highly effective because the clock has such a high activity factor,

Gating the clock to the input registers of a block prevents the registers from switching and thus stops all the activity in the downstream combinational logic.

Clock gating can be employed on any enabled register. Sometimes the logic to compute the enable signal is easy; for example, a floating-point unit can be turned off when no floating-point instructions are being issued.

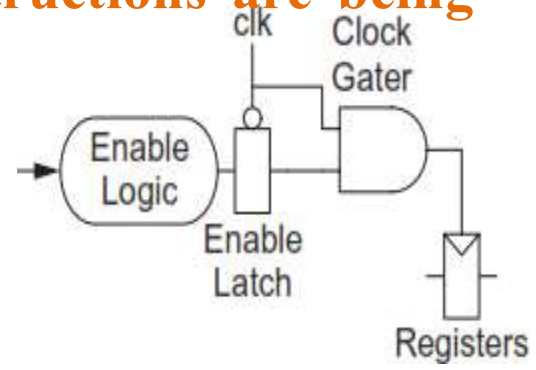


FIG: Clock gating

Switching Activity

The switching activity is comprised of the following two parameters:

(a) Static probability

(b) Transition rate

Static probability(for a given net) refers to the expected state of the signal.

For example, a static probability value of 0.2 implies that the signal is at logic-1 for 20% of the time (and logic-0 for 80% of time).

A 50% duty cycle for a clock signal implies that the static probability of the clock signal is 0.5 (or the clock is logic-0 for 50% of time and logic-1 for 50% of time).

The transition rate is the number of transitions per unit time. The transition rate is also referred to as toggle rate.

For periodic signals such as clocks where the frequency of the signal is specified, the transition rate is twice the frequency of the signal (since there are two transitions—rising and falling—within each cycle).

The power analysis utilizes the switching activity (static probability and transition rate) for each signal in the design.

The probability that pins *CK and Q are at 1 is 50%.* *However, the toggle* rate for pin *CK is 8 toggles in 40 ns* (200 million transitions per second). The toggle rate for pin *Q is 4 toggles in 40 ns* (100 million transitions per second).

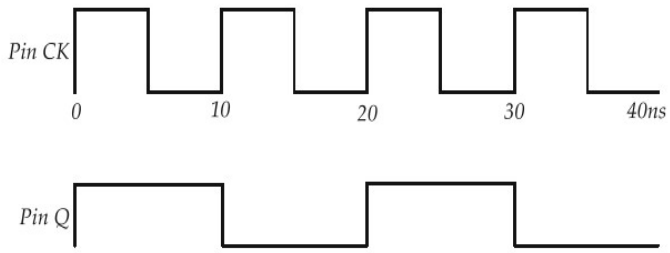


Fig: Example waveforms with same static probability but different transition rates

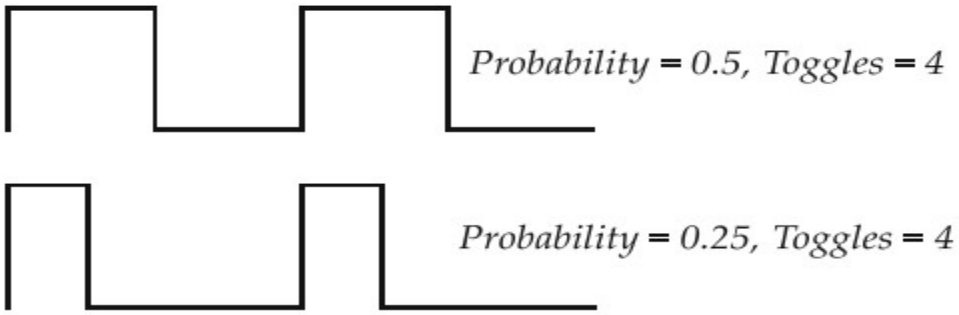


Fig: Example waveforms with same transition rates but different static probability

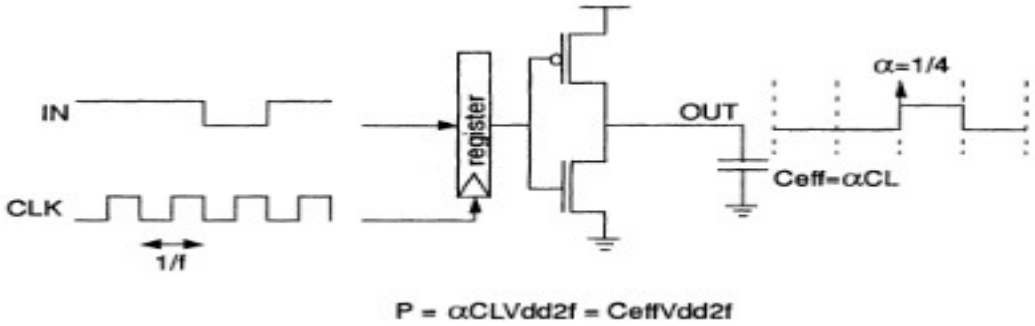
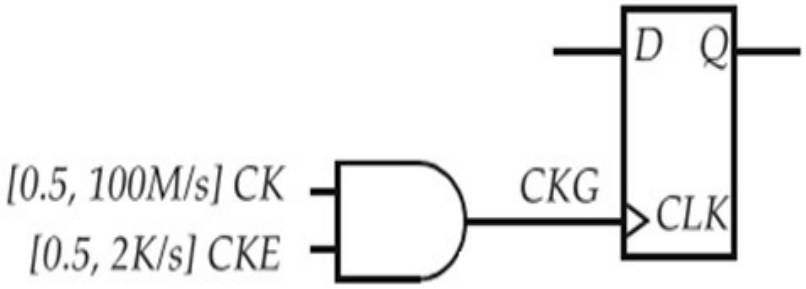


Fig: Interpretation of switching activity in synchronous systems



0.5, 100M/s : It is a signal with time period of 10ns and 50% duty cycle (1/10ns=100M/s)
0.5, 2k/s : It is a signal with time period of 500us and 50% duty cycle (1/500us=2k/s)

Fig: Example of reduced toggle rate at the output of an *and* gate

If *Pi* is the probability that node *i is 1. Pi'*= 1 – *Pi* is the probability that node *i is 0.*

αi the activity factor of node *i,* $\alpha_i = \bar{P}_i P_i$

Table lists the output probabilities of various gates as a function of their input probabilities

Gate	P_Y
AND2	$P_A P_B$
AND3	$P_A P_B P_C$
OR2	$1 - \bar{P}_A \bar{P}_B$
NAND2	$1 - P_A P_B$
NOR2	$\bar{P}_A \bar{P}_B$
XOR2	$P_A \bar{P}_B + \bar{P}_A P_B$

TABLE : Switching probabilities

Probabilistic Power Analysis

In the simulation approach of power analysis, The basic principle of the approach is to copy the circuit behavior over time, based on a predictable power model. In logic simulation, all signal events are recorded precisely by the simulation software. The events are then tabulated and averaged over time to obtain the actual average power dissipation of the circuit.

The basic model of probabilistic power analysis is different from simulation-based approaches. A logic signal is viewed as a random zero-one process with certain statistical characteristics. Here we derive several numerical statistical characteristics of the signal. The power dissipation of the circuit is then derived from the statistical quantities.

The primary reason for applying probabilistic analysis is computation efficiency. The number of statistical quantities to be analyzed is generally orders of magnitude smaller than the number of events to be processed compared to the simulation approach.

Typically, only a few statistical quantities need to be computed at a given node of the circuit as opposed to thousands of events during simulation.

The biggest drawback of the probabilistic approach is the loss in accuracy. In some circuits probabilistic analysis results in large error. Though the probabilistic power estimation technique is a crucial tool in today's digital VLSI design.

The application of probabilistic power analysis techniques has mainly been developed for gate-level abstraction and above.

Random Logic Signals

The modeling of zero-one logic signals is essential to understand probabilistic analysis. By taking few essential statistical parameters of a signal, we can construct a very compact description of the signal and analyze its effect on a circuit.

Characterization of Logic Signals

Logic signal consists of a waveform with zero-one voltage levels. The most precise way to describe a logic signal is to record all transitions of the signal at the exact times the transitions occur. Consider the logic signal in Figure.

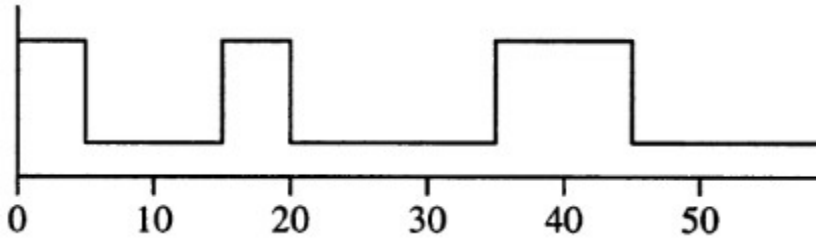


FIG: A logic signal and its characterization.

To represent the signal, we write down the initial state of the signal (state 1) and the time value when each transition occurs (5, 15, 20, 35, 45). Such description of the logic waveform Allows us to analyze the signals for various purposes.

For example, if we wish to compute the frequency of the signal, we count how many times the signal changes state and divide the number by the observation period.

This exact characterization of the signal gives us the full details of the signal history, allowing precise reconstruction of the signal.

For some purposes, the exact characterization of the signal is results in too much computation resource.

For example, if we only wish to know the frequency of the signal, there is no need to know the initial state and the exact switching times; the number of switches should be sufficient.

Consider the two signals shown in Figure. Although they appear to be different, the number of transitions is identical, which means that the frequencies of both signals are the same.

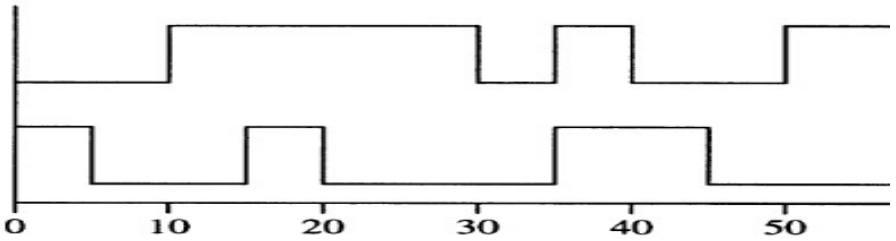


FIG: Two different logic signals with identical frequency.

If we wish to apply the $P = CV^2 f$ *equation, then only the frequency of the signal is necessary for the computation. The exact times at which the transitions occur are not important for our computation*. We need to know the number of transitions per unit time. By describing only the frequency of the signal, we can reduce the computation requirements. The disadvantage of this type of signal characterization is that the exact instant at which each individual transition occurs is not known.

By describing only certain *characteristic of the signal* (frequency of the signal) allows us to characterize a large number of different signals into a single class.

For example, there are infinite number of logic signals with frequency 1 MHz. But for the purpose of computing power dissipation using $P=CV^2f$ *equa*tion, all such logic signals result in identical power dissipation.

Many quantities such as signal probabilities, correlations, probability density functions, etc have been used for this purpose.

In this chapter, we will discuss the application of such signal representation method in power estimation and analysis of digital VLSI.

Probability and Frequency

The switching frequency f is a very important characteristic in the analysis of a digital signal for *power dissipation analysis* ($P = CV^2 f$).

The switching frequency *f of a digital signal is defined as half the number of transitions per unit* time, i.e.,

$$f = \frac{N(T)}{2T}$$

where *N(T) is the number of logic transitions observed in the period T.*

In the continuous random signal model, the observation period is often not specified. Under this situation, it is understood that the observation period is *infinite* ***i.e.*** the period is defined from time zero to *T with the value of T* approaches infinity

$$f = \lim_{T \to \infty} \frac{N(T)}{2T}$$

Static Probability and Frequency

The *static probability of a digital signal is the ratio of the time it spends in logic 1(t1)* to the total observation time *to+t1* *(duty cycle) expressed in a probability value between zero and* one, i.e.,

$$p = \frac{t_1}{t_0 + t_1}$$

It states how often the signal stays in logic 1 as opposed to logic 0. The probability that the signal is at logic 0 is 1-*p*.

The static probability and the frequency of a digital signal are related. If the static probability is zero or one, the frequency of the signal has to be zero(p(1)=P(0) =0) because if the signal makes a transition, the ratio of logic 1 to logic 0 has to be strictly between zero and one.

Consider a discrete random signal with static probability p. ***We assume that the signal* is *memoryless*, *meaning the current state of the logic signal is independent of its his*tory.**

The probability that the state is logic 1 is $p^1 = P$ ***and the probability that it is logic 0 is*** $p^0 = (1 - p)$. ***Sup*pose that the state is logic 1, the *conditional probability that the next state is also logic* 1 is** $p^{11} = P$ ***and the conditional probability that the next state is logic 0 is*** $p^{10} = (1-p)$.

***This is a direct consequence of our memoryless assumption. Simi*larly, we can compute** p^{01}, p^{00} **and summarize the results as follows:**

$$p^0 = p^{00} = p^{10} = (1-p)$$

$$p^1 = p^{01} = p^{11} = p$$

The probability *T that a transition occurs at a clock boundary is the probability of a* zero-to-one transition p^{01} plus the probability of a one-to-zero transition p^{10}

$$T = T^{01} + T^{10} = p^0 p^{01} + p^1 p^{10} = 2p(1-p)$$

The expected frequency *f is half the transition probability and we have* $f=p(1-p)$

The above equation relates the static probability *p of a memory less random logic sig*nal to its expected frequency *f* .

Figure *plots the frequency as a function of the* static probability.

Note that the frequency is zero at $p = 0$, $p = 1$ *and that the maxi*mum expected frequency of 0.25 is achieved when $p = 0.5$.

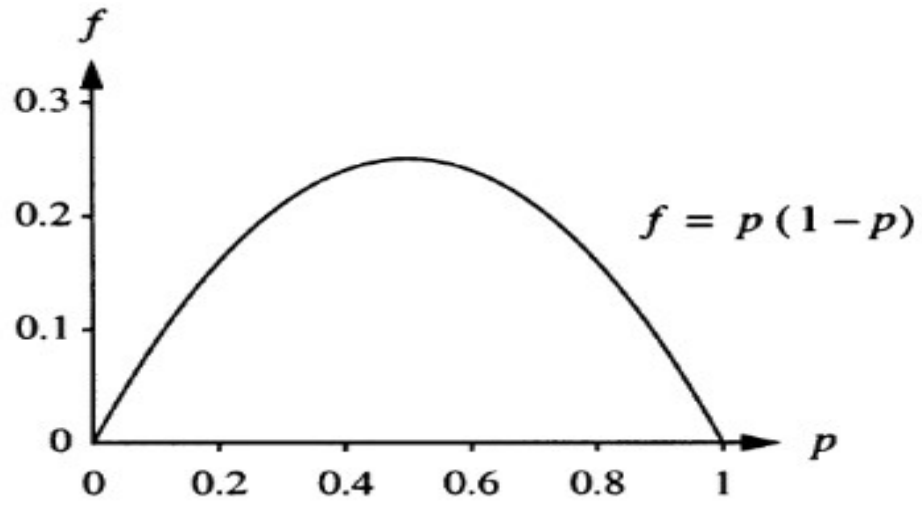


FIG: Expected frequency and static probability of discrete random signals.

The static probability is the only independent variable that characterizes the signal. All other properties of the signal, such as the frequency, can be derived from the static probability.

Probabilistic Power Analysis Techniques

In this section the general power analysis model for logic signals expressed in the random zero-one process will be discussed.

The basic framework for probabilistic power analysis is shown in Figure .

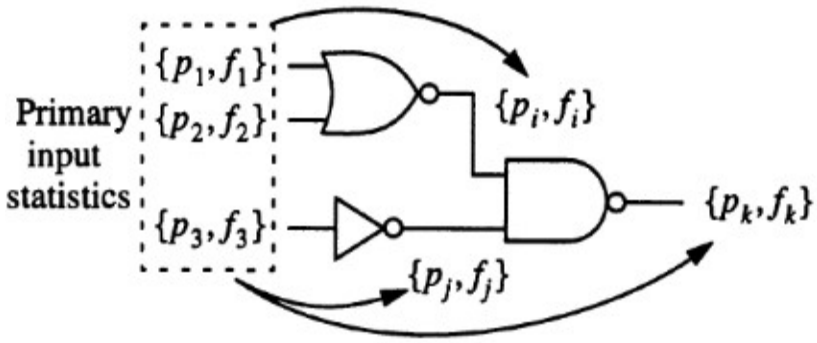


FIG: Propagation of statistical quantities probabilistic power analysis.

The statistical quantities at the primary inputs *{P1, f1} , {P2, f2} , {P3, f3} are specified. The statis*tical quantities at the internal nodes and outputs,

{Pi, fi} , {Pj, fj} , {Pk, fk} are derived from those of the primary inputs.

The power dissipation of the circuit is then computed from the statistical quantities at each node of the circuit and will be discussed later.

Propagation of Static Probability in Logic Circuits

WKT the static probability of a signal is the probability that the signal is at logic1 at any given time. To understand propagation of static probability through a logic gate, Consider a two-input AND gate as shown in Figure. If the static probabilities of the inputs are *P1 and P2 respectively and the two signals are statistically uncorrelated*, *the output static probability is P1P2* because the AND-gate sends out a logic 1 if and only if its inputs are at logic 1.

FIG: Static probability propagation of a two-input AND gate.

To derive the general formula for the propagation of static probability through an arbitrary Boolean function.

Let y = *f(x1, ... , xn) be an n-input Boolean function.*

Applying Shannon's decomposition with respect to *xi we have*

$$y = x_i f_{x_i} + \overline{x_i} f_{\overline{x_i}}$$

in which *fxi (fxi')is the new Boolean function obtained by setting xi = 1 (xi'= 0) in f{(x1,... , xn).*

Let the static probabilities of the input variables be P(x1),……,P(xn).

Since the two sum terms in the decomposition cannot be at logic 1 simultaneously, they are mutually exclusive. We can simply add their probabilities

$$P(y) = P(x_i f_{x_i}) + P(\overline{x_i} f_{\overline{x_i}}) = P(x_i)P(f_{x_i}) + P(\overline{x_i})P(f_{\overline{x_i}})$$

The probabilities *P(fxi) and P(fxi',) are computed from the recursive application of Shannon's* decomposition.

At the end of the recursion, *P(y) will be* expressed as an arithmetic function of the input probabilities *P(xi).* *(WKT P(xi') = 1 - P(x)).*

Shannon's Expansion

- **Divide a switching function into smaller functions**
- **Pick a variable x, partition the switching function into two cases: x=1 and x=0**

$$f(x,y,z,\ldots)= xf(x=1,y,z,\ldots) + x'f(x=0,y,z,\ldots)$$

Example

$$f(x)=xf(1)+x'f(0)$$

$$f(x,y)=xf(1,y)+x'f(0,y)$$

Transition Density Signal Model

Transition Density is the number of toggles per unit time. Consider periodic logic signals shown in Figure. The static probabilities and transition densities of the signals are given in the figure. Note that the top and the middle signals have identical static probabilities but different transition densities. The middle and the bottom signals have identical static probabilities and transition densities and are indistinguishable under this formulation.

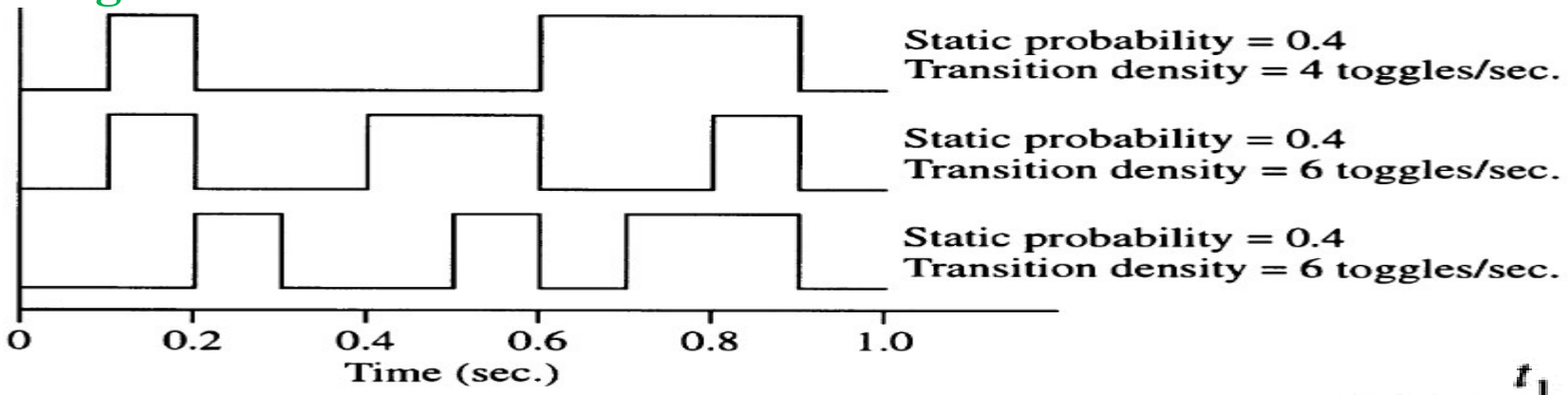


FIG: Static probabilities and transition densities of logic signals.

$$p = \frac{t_1}{t_0 + t_1}$$

WKT: The *static probability of a digital signal is the ratio of the time it spends in logic 1(t1)* to the total observation time *to+t1*

Propagation of Transition Density

For given a Boolean function y = *f(x1, ... , xn), the static probability P(x) and* transition density *D(x) of the input variables, we want to find the static probability P(y) and transition density D(y) of the output signal y.*

Here we use a *zero gate-delay model in which the output changes* with the input instantaneously(i.e. assumed that gate delay is zero).

The output *y makes a transition because an input xi* switches and the states of the other inputs happen to be such that the transition propagates to the output *y*.

WKT Shannon's decomposition equation

$$y = x_i f_{x_i} + \overline{x}_i f_{\overline{x}_i}$$

According to the Shannon's decomposition equation, when *xi = 1, the output of y is fxi, Similarly when xi = 0, the output is fxi'.*

If a 1-to-0 or 0-to-l transition in xi were to trigger a logic change in *y, the Boolean functions fxi and fxi' must have different values.*

i.e.

1. $f_{x_i} = 1, f_{\bar{x}_i} = 0$; or
2. $f_{x_i} = 0, f_{\bar{x}_i} = 1$.

In other words, the exclusive-OR of the two functions has to be 1,

i.e.

$$f_{x_i} \oplus f_{\bar{x}_i} = 1$$

The exclusive-OR of the two functions is called the *Boolean difference of y with* respect to *xi, denoted as*

$$\frac{dy}{dx_i} = f_{x_i} \oplus f_{\bar{x}_i}$$

The notation *dy/ dxi is* another Boolean function obtained from *f(x1, x2, ….xn)*

From the above discussion, it is clear that an input transition at *xi propagates to the* output y if and only if *dy / dxi*= 1.

Let ***P(dy/dxi)*** ***be the static probability that the*** **Boolean function** *dy/dxi* **evaluates to logic 1, and let** ***D(xi)*** **be the transition density of** ***xi,***

Because of the uncorrelated inputs assumption, the output will have transition **density of**

$$P\left(\frac{dy}{dx_i}\right) D(x_i)$$

The total transition density of the output y **for all input signals x*i*** **as**

$$D(y) = \sum_{i=1}^{n} P\left(\frac{dy}{dx_i}\right) D(x_i)$$

Since ***dy/ dxi*** **is just another Boolean function, we can compute** ***P(dy/ dxi) given*** *P(xj)'s* ***using***

$$P(y) = P(x_i f_{x_i}) + P(\overline{x_i} f_{\overline{x_i}}) = P(x_i)P(f_{x_i}) + P(\overline{x_i})P(f_{\overline{x_i}})$$

Gate Level Power Analysis Using Transition Density

The algorithm (steps) for Gate Level Power Analysis Using Transition Density for combinational circuits are:

1. **For each internal node *y of the circuit, find the Boolean function of the node with* respect to the primary inputs.**
2. **Find the transition density *D(y) of each node y = f(x1, ……xn) using Equation*** $$D(y) = \sum_{i=1}^{n} P\left(\frac{dy}{dx_i}\right) D(x_i)$$
3. **Compute the total power with the formula**

$$P = \sum_{\text{all nodes } y} 0.5 C_y V^2 D(y).$$

In step 3, 0.5*D(y)* *is the frequency and Cy is the capacitance of the node y.*

***The power* dissipation is dependent on**

1. **The input probabilities;**
2. **Transition densities; and**
3. **The Boolean function of each node with respect to the primary inputs.**

The main disadvantage of the transition density technique is

i) The accuracy is less , An inherent assumption of the analysis is that the logic gates of the circuits have zero delay.

ii) The signal glitches and spurious transitions are not properly modeled.

iii) The transition density model works well only on combinational circuits. For sequential cells it is difficult to model

Monte Carlo Simulation

Monte Carlo methods are a class of computational algorithms that rely(based) on repeated random sampling to obtain numerical results".

i.e Monte Carlo algorithms are used for introducing random variations within the given limits to explore the corner cases of any problem.

The Monte Carlo algorithms are used over a wide variety of applications including Risk Analysis, Finances, Statistics, Physics, and Electronic Designs.

Monte Carlo analysis is based on statistical distributions. On each simulation run, it calculates every parameter randomly according to a statistical distribution model.

With this analysis, we can see in which region the circuit will work most of the time.

The normal distribution is one particular bell shaped curve; it has two parameters (mean and variance) and plays a huge role in many areas of statistics, particularly in connection with the central limit theorem.

Random sampling, which actually refers to a whole huge group of sampling methods.

Ex: A Train comes to a station every day at 8:00am and some time it's late by few minutes and some day early. If we record all the train timing for a long duration of time say a year or two you will get a normal distribution centered at 8:00am.

Everyday if we record some random train arrival for some duration and you do it for a year. The distribution will be random distribution.

All simulation-based power analysis systems require a set of *simulation vectors* *at the primary inputs of the circuit to trigger* the circuit activities. To obtain the power dissipation of the circuit, the switching activity information is collected and applied to the appropriate power models after simulation.

The simulation vectors applied to the circuit have a substantial impact on the final power number reported because power dissipation of a digital circuit heavily depends on its switching activities.

Each simulation vector causes some energy to be dissipated and the total power dissipation is derived by summing up the energy of each vector and dividing over the simulation time.

If we simulate a circuit with few vectors, we would expect that the power dissipation result obtained is not truthful because the vector length is too short. Most part of the circuit is probably not exercised enough to obtain the actual toggling activities.

On the other hand, If we simulate the circuit for millions and millions of vectors to obtain a very accurate measure of the power dissipation. But it is wasting of computer resources to simulate that many vectors.

How do we know that we have simulated enough vector length? All these lead to an important question of the *stopping Criteria* of the simulation: when do we stop simulation so that the result is accurate enough for our purpose?

Statistical Estimation of Mean

The stopping criteria of simulation-based power analysis were first studied by Burch ,Trick, and et.all.

To formulate the problem, we define a basic *sample period T* in which a single power dissipation value is observed. For example, *T may* be several vectors or several clock cycles.

After a particular simulation period *Ti, the* power dissipation *Pi of the circuit during the period Ti is computed.*

As such, we obtained a series of *power samples* $P_0, P_1, \ldots, P_N,$

The estimated power dissipation P of the circuit under simulation is given by the average value of the samples.

$$P = \frac{(P_0 + P_1 + \ldots + P_N)}{N} \quad \text{----------1}$$

This is the mean estimation problem in statistics

The power samples *Pi are random variables following some unknown probability* density function. The distribution of *Pi depends on the circuit, simulation vectors and* the sample interval.

Let μ be mean and σ^2 be the variance of *Pi respectively.*

According to the *central limit theorem in statistics, the sample mean P* approaches the normal distribution for large *N regardless of the distribution of Pi, For* theoretical development, let us assume that the samples *Pi have normal distribution*.

Basic statistical theory states that the average of normally distributed random variables also has normal distribution. The mean of *P is exactly* μ *and its variance is*

$$\sigma_P^2 = \frac{\sigma^2}{N}$$

Where σ^2 be the variance of *Pi* , σ_p^2 is the *variance from* mean of *P*

As we increase the sample size *N*, *the variance* σ_p^2 *decreases so that we obtain more* accurate measures of the true mean μ. The normal distribution curve for *P is shown* in Figure

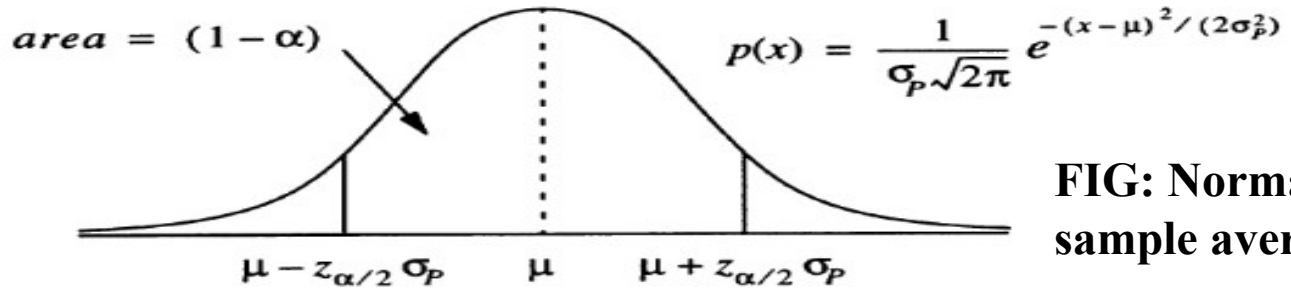


FIG: Normal distribution curve of sample average *P*.

Let ε be the maximum error tolerance, *typically with values less* than 10%. Given ε , *the probability that P is within the* ε *error range of* the true mean μ, In other words the probability for the condition $0 \leq |P - \mu|/\mu \leq \varepsilon$. *If this probability is high, we will trust the estimate P; otherwise,* we must increase the sample size *N to gain more confidence. We can obtain this* probability by integrating the normal distribution curve *p(x) in Figure*.

Generally the **probability** is more conveniently expressed by a **confidence variable α**. The *confidence level is defined as 100 (1 - α.)%. A confidence level of 100%* (α = 0) means that *P is absolutely within the error tolerance of ε. Typically, the* confidence level is set to more than 90% to be meaningful, i.e., $\alpha \leq 0.1$.

The relationships among ε, α *and N can be obtained by defining a variable* $Z_{\alpha/2}$ ($Z_{\alpha/2}$ *is typically obtained from a mathematical table known as the z-dis*tribution function)*such that the area between* $\mu - Z_{\alpha/2}\sigma_p$ *and* $\mu + Z_{\alpha/2}\sigma_p$ under the normal distribution curve *p(x) is (1 - α). From Figure, this is also the confidence level of the* condition $|P - \mu| \leq Z_{\alpha/2}\,\sigma_p$.

To ensure the condition $|P - \mu|/\mu \leq \varepsilon$, *we require that*

$$\frac{|P-\mu|}{\mu} \leq \frac{z_{\alpha/2}\,\sigma_P}{\mu} \leq \varepsilon$$

Substituting the equation $\sigma_p^2 = \sigma^2/N$ into the above inequality we get

$$\frac{z_{\alpha/2}\,\sigma}{\mu\sqrt{N}} \leq \varepsilon$$

Since α and ε *are fixed constants prescribed by the experimenter, the equation can be* rewritten as

$$N \geq \left(\frac{z_{\alpha/2}\, \sigma}{\varepsilon\, \mu} \right)^2 \quad \text{----------x}$$

Equation x tells us the minimum number of samples *N to be taken for* the error of *P is within the tolerance of ε. In other* words, if we make a hundred measurements of *P with N samples, only* 100 (1-α)% of the measurements will have error more than ε.

The value of $Z_{\alpha/2}$ *is typically obtained from a mathematical table known as the z-dis*tribution function. The following table shows some values of (1 - α) and $Z_{\alpha/2}$.

TABLE 2.1 Table of z-distribution.

$(1-\alpha)$	$z_{\alpha/2}$
0.9	1.65
0.95	1.96
0.99	2.58
0.998	3.00
1.0	∞

EX. Prob1: The standard deviation of the power samples measured from a circuit has been observed to have ±20% fluctuation from the mean. How many samples are required so that we are 99% confidence that the error of sample mean is within ±5%?

We have $\sigma/\mu = 0.2$, $\varepsilon=0.05$, $(1-\alpha)=0.99$, Therefore $Z_{\alpha/2}$ from the table for $(1-\alpha)=0.99$ is 2.58

$$N = \left(\frac{z_{\alpha/2}\,\sigma}{\varepsilon\,\mu}\right)^2$$

$N=(0.2\text{x}2.58/0.05)^2$ =107 Samples

TABLE 2.1 Table of z-distribution.

$(1-\alpha)$	$z_{\alpha/2}$
0.9	1.65
0.95	1.96
0.99	2.58
0.998	3.00
1.0	∞

Information theory-based approach (Signal Entropy)

Entropy theory has been successfully used in communication systems to analyze the information contents. In entropy analysis, the signals in a logic circuits are treated as a collection of random signals.

The entropy or randomness of the signals is related to the average switching activities of the circuit.

Basics of Entropy

Entropy is a measure of the randomness carried by a set of discrete events observed* over time(average information content). Suppose a system has *m possible events {S1, S2 ••• ,Sm} where each* event *Si occurs with probability Pi and

P1 + P2 + ... + Pm = 1.

The information content (Ci) carried by an event *Si* is given by

$$C_i = \log_2 \frac{1}{p_i}$$

Since $0 \leq Pi \leq 1$, *the logarithm term is non-negative and we have $Ci \geq 0$.*

***The entropy of the* system is the average information content of the *system and this is the weighted sum of the information* contents of Ci.**

***The entropy of the* system is given by**

$$H = \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i}$$

In logic signal analysis, the events correspond to the word-level values of the signals.

In an n-bit logic signal, there are $m = 2^n$ distinct word-level values in the range of $0 \leq i \leq 2^n-1$

Power Estimation Using Entropy

Consider a combinational logic circuit with m-bit input X and n-bit output *Y* . An entropy based power estimation method for the combinational logic is shown in figure

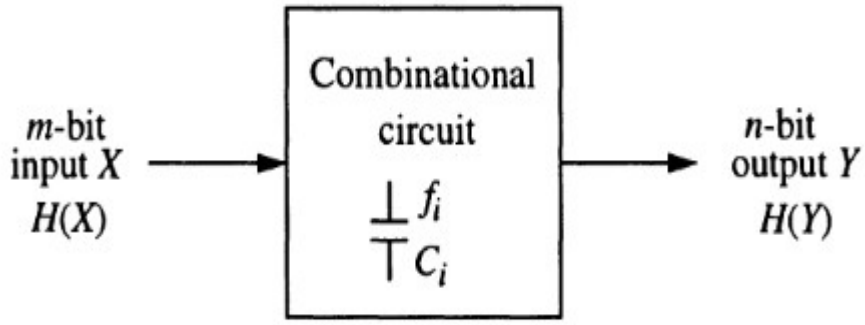


FIG: Power estimation of combinational logic using entropy analysis.

Assuming constant *Vdd, let fi be the switching frequency of a node capacitance Ci* in the circuit and let *N be the total number of nodes, then the* power dissipation of the circuit is

$$P = \sum_{i=0}^{N} C_i V_{dd}^2 f_i$$

If we assume that the node frequency *fi = F is constant for all nodes i*, *we can write*

$$P \propto F \sum_{i=0}^{N} C_i = FA \qquad \text{----------1}$$

where *A is the sum of node capacitance, proportional to the area or size of the circuit.*

Let *H(X) and H(y) be the entropy of the input and output signals.* From the study of the complexity of Boolean function, a relation has been empirically observed between the average implementation area *A and the output entropy H(Y). For large value of n,*

$$A \propto \frac{2^n}{n} H(Y) \qquad \text{----------2}$$

The average frequency *F* *of the circuit is related to the entropy of the signals*

$$F \approx \frac{2}{3\,(m+n)}\,[H(X)+H(Y)] \qquad \text{----------3}$$

Substituting Equations 2 and 3 in eqn 1 gives

$$P \propto F \sum_{i=0}^{N} C_i = FA$$

$$P \propto \frac{2^{n+1}}{3n\,(m+n)}\,H(Y)\,[H(X)+H(Y)] \qquad \text{----------4}$$

The proportionality constant depends on the operation voltage and node capacitance. Equation 4 expresses the power dissipation of a combination circuit with respect to its input *m* and output *n* bit sizes, *and entropy measures H(x), H(y).*

The entropy measures *H(x) and H(y) are typically obtained by monitoring the sig*nals X and *Y during a high-level simulation of the circuit.* In general, *H(Y) ≤ H(X) because the outputs of a logic gate make fewer transitions than its* inputs.

Equation 4 is applicable for large combinational logic circuits with high degree of randomness. The power estimation equation typically fails when the circuits exhibit structural regularity (multipliers, counters and decoders) or highly correlated signals (voice, image).

The Entropy analysis of a sequential circuit is shown in Figure. A high-level simulation of the circuit is performed and the input and output entropy measures of the combinational logic are observed.

The entropy power estimation method for combinational logic is applied and added to the power dissipation of the sequential elements, which can be easily estimated from the input and output activities of their signals.

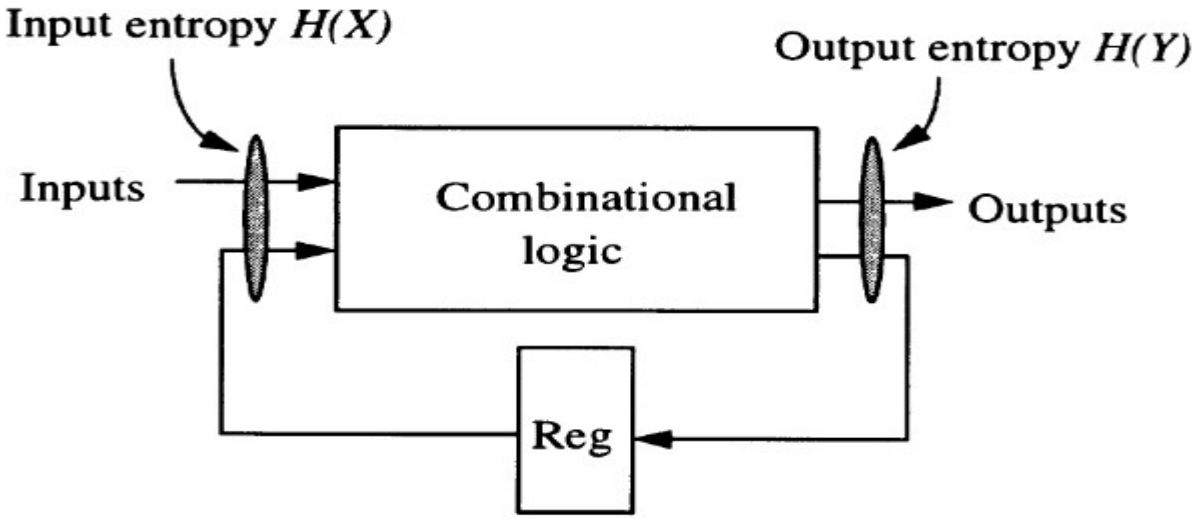


FIG: Entropy analysis of a sequential circuit.