A New Method for Constructing IP Level Power Model Based on Power Sensitivity

Heng-Liang Huang†, Jiing-Yuan Lin*, Wen-Zen Shen† and Jing-Yang Jou†

†Department of Electronic Engineering, National Chiao-Tung University, Hsin-Chu, Taiwan
e-mail: cooper@ee.nctu.edu.tw

*Global UniChip Corporation, Hsin-Chu, Taiwan
 e-mail: jylin@unichip.com.tw

Abstract — This paper proposes a nominal point selection method for IP (Intellectual Property) level power model based on power sensitivity. By analyzing the relationship between the dynamic power consumption of CMOS circuits and their input signal statistics, three nominal points are efficiently selected to construct a power model based on power sensitivity. Our experimental results on a number of benchmark circuits show the effectiveness of the proposed method. Estimation accuracy within 5.78% of transistor level simulations is achieved.

I. INTRODUCTION

Recently, IP (Intellectual Property) reusing has attracted lots of researchers’ attention. The urgent need of a feasible IP power model emerges from the rapid expansion of the portable electronics, in which the power dissipation has become an essential issue of circuit designs. In CMOS circuits, the major energy consumed is charging and discharging of capacitors, and is known as the dynamic power. Analyzing the switching activities of the nodes in a circuit is appropriate for calculating power consumption without simulating the circuit. However, the complexity of calculating the transition probability for each node grows exponentially with the size of a circuit. Especially after the concept of IP reusing is brought up, a power estimator that does not analyze every node in the IP block is strongly demanded.

Regression method had been used in [5][6] to construct a high level power model as a function of weighted input and output transitions. It is a good and fast estimation for randomly generated input signals, i.e., both the signal probabilities and the transition densities are around 0.5. However, cases with highly biased input probabilities could make the estimation of power far from accuracy.

In [2], the power consumption of a circuit is modeled as a function of the mean values of input signal probabilities, input transition densities and output transition densities. A power lookup table is built and indexed with those three values. Good accuracy can be obtained when the three values of the circuit under estimation are close to the pre-selected indices and the variances of the signal probability and the transition density are small. However, the output transition densities can not be calculated easily from the input statistics. Therefore, to explore the output density space requires simulating a tremendous amount of combinations of the input probabilities and the transition densities. It is very time consuming.

The concept of power sensitivity is first proposed in [7], in which power is modeled as the sum of weighted uncertainties of the input signal probabilities and the transition densities plus the power at a nominal point. The weights of those uncertainties are called the power sensitivities. A Statistical Technique to Estimate Power Sensitivities abbreviated to STEPS [7] is proposed to determine power sensitivities. Another Symbolic Technique to Obtain Power Sensitivities named STOPS [8] expresses the signal probability and the transition density of each node symbolically to determine the power sensitivities. The estimated power based on this kind of methods strongly depends on the selected nominal points. However, it is difficult to select the nominal points. The authors in [7][8] randomly selected some points but did not evaluate the accuracy of their models.

In this paper, an efficient nominal point selection method is proposed. We first select a nominal point that can achieve the minimum average error. According to the first nominal point, we select the second and the third nominal point. Based on our deduction, three nominal points are enough to construct an accurate power model. Experimental results fully demonstrate the accuracy of the power model constructed with our easy-to-use method.

This paper is organized as follows. In section II, we introduce some preliminary and review the power sensitivity model. Our nominal point selection algorithm is discussed in section III. Section IV shows the experimental results and some conclusions are given in section V.

II. PRELIMINARY

A. Signal Transition Activities

A digital signal at node \( x \), \( x(t) \), is either 1 or 0 if we neglect the rise/fall time and over/under shoots [1]. The expected value of a signal to be 1 in a clock cycle with a period \( T \) can be defined as

\[
p_x = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt ,
\]

where \( p_x \) is the equilibrium probability of \( x(t) \) or the probability of node \( x \) in short. Four types of transition that a signal can make between two consecutive periods are 00, 01, 10 and 11. The transition density of \( x(t) \) is defined as
\[
d_i = \lim_{T \to \infty} \frac{n_i(T)}{T},
\]
where \( n_i(T) \) is the number of 01 and 10 transitions of \( x(t) \) in the time interval \((-T/2, +T/2)\) [1]. Intuitively, we define the 01 and 10 transition probabilities as
\[
p_i^{01} = \lim_{T \to \infty} \frac{n_{01}(T)}{T},
\]
\[
p_i^{10} = \lim_{T \to \infty} \frac{n_{10}(T)}{T},
\]
where \( n_{01}(T) \) and \( n_{10}(T) \) are the number of 01 and 10 transitions of \( x(t) \) in the time interval \((-T/2, +T/2)\), respectively. Note that a digital circuit always makes a 10/01 transition at some time after a 01/10 transition. When \( T \) approaches infinity, the transition probabilities of 01 and 10 will converge to the same value. With the above definitions, four transition probabilities are derived and expressed as
\[
p_i^{00} = 1 - p_i - d_i / 2, \quad p_i^{01} = p_i^{10} = d_i / 2, \quad p_i = p_i - d_i / 2.
\]
Note that \( 0 \leq p_i \leq 1, \forall i, j \in \{0, 1\} \). A relationship between \( p_i \) and \( d_i \) is derived [2] as
\[
d_i / 2 \leq p_i \leq 1 - d_i / 2.
\]

B. Power Sensitivity [7]

For an \( n \)-input CMOS circuit, the logic value of the \( i \)th input node at time \( T \) is denoted as \( b_i^T \). The input pattern at time \( T \) can be expressed as the transpose of \([b_0^T b_1^T \ldots b_n^T]\). The input pattern transition of two consecutive patterns is denoted as \( V_i^{0} V_i^{1} \),
\[
V_i^{0} V_i^{1} = \begin{bmatrix} b_i^0 & b_i^1 \\ b_i^1 & b_i^0 \\ \vdots & \vdots \\ b_i^n & b_i^n \end{bmatrix}
\]
The average power consumption \( \text{Power}_{\text{avg}} \) of a CMOS circuit can be expressed as
\[
\text{Power}_{\text{avg}} = \sum_{V_i^{0} V_i^{1}} \text{Pr}(V_i^{0} V_i^{1}) \times \text{Power}(V_i^{0} V_i^{1}),
\]
where \( \text{Pr}(V_i^{0} V_i^{1}) \) is the probability of the input pattern transition and \( \text{Power}(V_i^{0} V_i^{1}) \) is the corresponding power consumption. Assume that the input signals are spatially independent,
\[
\text{Power}_{\text{avg}} = \sum_{V_i^{0} V_i^{1}} \prod_{i=0}^{n} p_i^{00} \times \text{Power}(V_i^{0} V_i^{1}),
\]
where \( p_i^{00} \) is the probability that the logic value of the \( i \)th node being \( b_i^0 \) in the beginning and \( b_i^0 \) at time \( T \). Note that power consumption is a function of both the probabilities and the transition densities of the input signals. Let \( S \) be a set of input statistics, and

\[
S = \begin{bmatrix} p_i^{00} & p_i^{10} & p_i^{10} & p_i^{11} \\ \vdots & \vdots & \vdots & \vdots \\ p_i^{00} & p_i^{10} & p_i^{10} & p_i^{11} \end{bmatrix}
\]
Apply Taylor's expansion and consider the first order approximation, we have
\[
\text{Power}_{\text{avg}}(S) = \text{Power}_{\text{avg}}(S_{\text{nom}}) + \sum_{V_i^{0} V_i^{1}} \frac{\partial \text{Power}_{\text{avg}}(S_{\text{nom}})}{\partial p_i^{00}} (S_{\text{nom}} - p_i^{00})
\]
where \( S_{\text{nom}} \) is a set of nominal input statistics,
\[
\Delta p_i^{00} = p_i^{00} - p_i^{00}_{\text{nom}} \quad \text{and}
\]
\[
S_{\text{nom}} = \begin{bmatrix} p_i^{00}_{\text{nom}} & p_i^{10}_{\text{nom}} & p_i^{10}_{\text{nom}} & p_i^{11}_{\text{nom}} \\ \vdots & \vdots & \vdots & \vdots \\ p_i^{00}_{\text{nom}} & p_i^{10}_{\text{nom}} & p_i^{10}_{\text{nom}} & p_i^{11}_{\text{nom}} \end{bmatrix}
\]
The partial derivative is the changing rate of the power consumption due to the change of one input transition probability. It is so called **power sensitivity** [7][8].

The partial derivatives can be estimated with a statistical method STEPS [7]. STEPS is an Monte Carlo based approach that simulates a circuit with randomly generated patterns to get samples of input power sensitivities until the mean value of those samples converges. Another method STOPS [8] for estimating power sensitivities requires topolological partitioning to reduce its enormous complexity.

In this paper, we adopt a straightforward method instead. The circuit is first simulated with the nominal input statistics \( S_{\text{nom}} \). Then, assign a small variation to the transition probabilities of an input node and proceed the simulation again to compute the changing rate of power corresponding to the variation. The steps are repeated for other input nodes until all partial derivatives are derived.

III. NOMINAL POINT SELECTION

A. Nominal 0

To construct a power model based on the power sensitivity, one must choose some nominal points first and calculate the power sensitivities according to the chosen nominal points. Without doubt, the more nominal points we choose the more accurate our model becomes. However, adding a nominal point requires \( 2 \times n + 1 \) more simulations to be carried out in a circuit with \( n \) inputs [7]. Furthermore, choosing nominal points arbitrarily makes the error range of the constructed power model unpredictable. It is thus important to choose the right nominal points to minimize the error of the power model with as few nominal points as possible.

This section focuses on finding a nominal input statistic to fit all kinds of \( S \) well. The average power is estimated by using equation (10), and the error can be expressed as
Error = Power\_avg (S) - \left[\text{Power\_avg (S\_nom)} + \sum_{i} \frac{\partial \text{Power\_avg}}{\partial p_{i}} (S\_nom) \times \Delta p_{i}^{\text{nom}}\right]. \tag{12}

Equation (12) can be approximated with the second order term, while neglecting the higher order terms. The average error will be

E[Error] \approx E \left[\sum_{i} \left(\frac{\partial^2 \text{Power\_avg}}{\partial p_i^2} (S\_nom) \times \Delta p_i^{\text{nom}} \times \Delta p_j^{\text{nom}}\right)\right]. \tag{13}

The second order partial derivatives can be expressed as

\frac{\partial^2 \text{Power\_avg}}{\partial p_i^2} (S\_nom) = \sum_{\forall i,j} \left(\prod_{\forall k \neq i,j} p_{k,\text{nom}}^{\text{nom}}\right) \times \text{Power}(V'^{i',j'}). \tag{14}

Let $p_{\max}$ be the maximum value of $p_{i,\text{nom}}^{\text{nom}}$. Power\_max is the maximum value of Power(V'i'j'). The right hand term (R.H.T.) in equation (14) can be expressed as

\text{R.H.T.} \leq 4^{-n} \times p_{\max}^{-n} \times \text{Power\_max}. \tag{15}

The average error is bounded by

\left|E[Error]\right| \leq \frac{1}{2^n} \times 4^{-n} \times p_{\max}^{-n} \times \text{Power\_max} \times \left[\sum_{\forall i,j} \left(\prod_{\forall k \neq i,j} p_{k,\text{nom}}^{\text{nom}}\right) \times (p_{i,j}^{\text{nom}} - p_{i,j}^{\text{nom}})\right]. \tag{16}

Since the transition probabilities of logic values at input nodes $i$ and $j$ are independent,

\left|E[Error]\right| \leq \frac{1}{2^n} \times 4^{-n} \times p_{\max}^{-n} \times \text{Power\_max} \times \sum_{\forall i,j} \left|\left(\prod_{\forall k \neq i,j} p_{k,\text{nom}}^{\text{nom}}\right) \times (p_{i,j}^{\text{nom}} - p_{i,j}^{\text{nom}})\right|. \tag{17}

From equation (17), the maximum estimation error can be minimized if we choose $p_{i,j}^{\text{nom}}$ to be equal to $E\left[ p_{i,j}^{\text{nom}} \right]$ for all $i$.

The following experiments are conducted to verify the above analysis. The input signals of all input nodes in a circuit are assumed mutually independent, and the transition probabilities are uniformly distributed. That means $d_i$ is uniformly distributed in $[0,1]$, and consequently, $p_i$ is uniformly distributed in $[d_i/2, 1-d_i/2]$. Both of them have the same mean value of 0.5. By equation (4), we choose $p_{i,\text{nom}} = d_{i,\text{nom}} = 0.5$ for all $i$. From the above analyses, we conclude that selecting the mean values of the input transition probabilities as the elements of the nominal input statistic, the minimum average error can be achieved. A nominal input statistics $S_{\text{nom}}$ is thus built with each element equals 0.25. This nominal point will be designated as nominal 0 ($N_0$) in this paper.

Five other nominal points are randomly constructed for comparison. Five hundred SIs are randomly generated following the distributions, $d_i \sim U[0.1]$ and $p_i \sim U[d_i/2, 1-d_i/2]$, where $U[a,b]$ is an uniform distribution between $a$ and $b$. One thousand input vectors corresponding to each $S$ are fed into powermill to simulate the exact power consumption. The power consumption for each $S$ is also evaluated by the first order power estimation with nominal point $N_0$ and five other randomly selected nominal points. The error of the power estimated with each nominal point is defined as the difference between the estimated power and the simulated power. The estimation errors are shown in Table I.

<table>
<thead>
<tr>
<th>Circuit</th>
<th>$N_0$</th>
<th>Rand0</th>
<th>Rand1</th>
<th>Rand2</th>
<th>Rand3</th>
<th>Rand4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cm138a</td>
<td>5.78%</td>
<td>12.63%</td>
<td>15.38%</td>
<td>11.07%</td>
<td>6.62%</td>
<td>8.08%</td>
</tr>
<tr>
<td>Cm150a</td>
<td>2.29%</td>
<td>4.47%</td>
<td>4.82%</td>
<td>4.07%</td>
<td>8.03%</td>
<td>3.12%</td>
</tr>
<tr>
<td>Cm151a</td>
<td>3.66%</td>
<td>7.83%</td>
<td>11.65%</td>
<td>5.38%</td>
<td>6.87%</td>
<td>10.57%</td>
</tr>
<tr>
<td>Cm152a</td>
<td>3.68%</td>
<td>41.35%</td>
<td>5.88%</td>
<td>8.46%</td>
<td>4.68%</td>
<td>6.82%</td>
</tr>
<tr>
<td>Cm162a</td>
<td>4.31%</td>
<td>28.5%</td>
<td>5.89%</td>
<td>7.47%</td>
<td>7.36%</td>
<td>6.30%</td>
</tr>
<tr>
<td>Cm163a</td>
<td>3.73%</td>
<td>29.92%</td>
<td>7.65%</td>
<td>7.22%</td>
<td>6.58%</td>
<td>10.63%</td>
</tr>
<tr>
<td>Cm42a</td>
<td>4.62%</td>
<td>22.08%</td>
<td>12.55%</td>
<td>5.97%</td>
<td>7.82%</td>
<td>5.50%</td>
</tr>
<tr>
<td>Cm62a</td>
<td>5.69%</td>
<td>24.59%</td>
<td>12.68%</td>
<td>9.70%</td>
<td>12.83%</td>
<td>12.76%</td>
</tr>
<tr>
<td>Cm85a</td>
<td>2.53%</td>
<td>16.72%</td>
<td>4.70%</td>
<td>4.11%</td>
<td>5.39%</td>
<td>3.87%</td>
</tr>
<tr>
<td>Cmb</td>
<td>1.82%</td>
<td>3.99%</td>
<td>3.22%</td>
<td>3.06%</td>
<td>2.75%</td>
<td>3.05%</td>
</tr>
<tr>
<td>Comp</td>
<td>2.90%</td>
<td>12.05%</td>
<td>4.57%</td>
<td>4.53%</td>
<td>5.15%</td>
<td>5.60%</td>
</tr>
<tr>
<td>Cu</td>
<td>2.62%</td>
<td>4.01%</td>
<td>5.74%</td>
<td>5.48%</td>
<td>4.65%</td>
<td>5.47%</td>
</tr>
<tr>
<td>Decod</td>
<td>4.89%</td>
<td>15.84%</td>
<td>14.07%</td>
<td>8.49%</td>
<td>9.69%</td>
<td>11.06%</td>
</tr>
<tr>
<td>F51m</td>
<td>2.45%</td>
<td>10.42%</td>
<td>4.56%</td>
<td>3.45%</td>
<td>4.32%</td>
<td>3.21%</td>
</tr>
<tr>
<td>Average</td>
<td>3.64%</td>
<td>16.73%</td>
<td>8.10%</td>
<td>6.20%</td>
<td>6.62%</td>
<td>6.87%</td>
</tr>
</tbody>
</table>

In Table I, we can see that the nominal point constructed with the mean values of input transition probabilities always gives the best estimation for all circuits under test.

B. More Nominal Points Selection

Although the average error of nominal 0 is minimized, the error is still unacceptable large in some cases. In order to improve the accuracy, some more nominal points should be selected. The cases, in which nominal 0 may produce large errors, are first examined. Take circuit cm138a as an example, the relationship between the power estimated with nominal 0 and the simulated one is shown in Fig. 1.

![Fig. 1. Simulated Power vs. Estimated Power at Nominal 0](image-url)
The solid line in Fig. 1 illustrates the ideal estimation of the simulated power. In Fig. 1, the estimated power consumption deviates from the simulated power when the later is away from the nominal power. From this observation, if we can choose some other nominal points around the mean of the larger and smaller halves of simulated power, we should be able to improve the performance. Consider equation (10),

let \( \text{Power}_{avg}(S_{nom}) = \frac{\partial \text{Power}}{\partial p^i_{nom}} (S_{nom}) \times p^i_{nom} + C \) then

\[
\text{Power}_{avg}(S) = \sum_{i=1}^{n} \frac{\partial \text{Power}}{\partial p^i_{nom}} (S_{nom}) \times p^i_{nom} + C.
\]  

(18)

The average power becomes a first order polynomial of input transition probabilities with \( 4n \) dimensions and can be reduced to \( 2n \) dimensions as a function of independent parameters \( p_i \) and \( d_i \), referred to equation (4).

\[
\text{Power}_{avg}(S) = \sum_{i=1}^{n} (a_i \times p_i + b_i \times d_i) + C',
\]

(19)

where the coefficients \( a_i, b_i \) and \( C' \) are

\[
a_i = \frac{\partial \text{Power}}{\partial p^i_{nom}} (S_{nom}) - \frac{\partial \text{Power}}{\partial p^i_{nom}} (S_{nom}),
\]

\[
b_i = \left( \frac{\partial \text{Power}}{\partial p^i_{nom}} (S_{nom}) + \frac{\partial \text{Power}}{\partial p^i_{nom}} (S_{nom}) \right) / 2,
\]

\[
C' = \sum_{i=1}^{n} \frac{\partial \text{Power}}{\partial p^i_{nom}} (S_{nom}) + C.
\]

(20)

When \( S_{nom} \) is given, \( a_i, b_i \) and \( C' \) are constant. Note that the valid values of \( (p_i, d_i) \) are constrained inside the triangle, referred to equation (5), as shown in Fig. 2. \( S_{nom} \) is first located, which is the input transition probability matrix corresponding to the maximum power. Since the optimal value of a linear equation with a limited feasible region happens only on the boundary points [10], only the three corners of the triangle with coordinates \((0,0), (1,0)\) and \((0.5,1)\) must be examined. The elements \( (p_{i, max}, d_{i, max}) \) of \( S_{max} \) can thus be characterized by equation (19) as one of \((0,0), (1,0), (0.5,1))\).

Similarly, \( S_{min} \), which is the input transition probability matrix corresponding to the minimum power, is also characterized.

By using the new input transition probability matrix \( S_{max} \) \( (S_{min}) \), together with the original nominal point \( S_{nom} \), a new nominal point is selected as the middle point between \( S_{nom} \) and \( S_{max} \) \( (S_{min}) \) and designated as \( N_{q\bar{q}} \) \( (N_{g\bar{g}}) \). The elements of the third quartile nominal point, \( N_{q\bar{q}} \), can be expressed as

\[
p_i\left(q_{q} \right) = \frac{(p_{i, max} + p_{i, min})}{2},
\]

\[
d_i\left(q_{q} \right) = \frac{(d_{i, max} + d_{i, min})}{2}.
\]

(21)

In the same manner, the first quartile nominal point, \( N_{g\bar{g}} \), can be obtained. Fig. 3 shows the experimental results of using only the nominal point \( N_{g\bar{g}} \). The estimation errors in the low power region are obviously reduced, while others are increased. On the other hand, Fig. 4 shows that less error is induced in the high power region when the power is estimated with only \( N_{g\bar{g}} \).

![Figure 2: Relationship between \( p_i \) and \( d_i \)](image)

Since the estimation error depends on the selected nominal point, a guideline for selecting the optimal nominal point is required for minimizing the estimation error. It is possible to build a power model constructed of these three nominal points and dynamically select the most suitable nominal point(s) to give power estimation. The details of constructing the 3-point power model will be discussed in the following section.

IV. EXPERIMENTAL RESULTS

With the three nominal points obtained in the previous
section, we can construct a 3-point model for a circuit. Considering an input statistic \( S \), let \( \text{Power}_{avg}(S|N_{0}), \text{Power}_{avg}(S|N_{1}), \text{and} \text{Power}_{avg}(S|N_{2}) \) be the estimated power done with the nominal points \( N_{0}, N_{1}, \) and \( N_{2} \), respectively. Let \( \text{Power}_{avg}(N_{0}), \text{Power}_{avg}(N_{1}), \) and \( \text{Power}_{avg}(N_{2}) \) be the nominal power of \( N_{0}, N_{1}, \) and \( N_{2} \), respectively. Due to the strong dependence of accuracy on the selected nominal point, we take the distance between the estimated power and the nominal power as parameters of interpolation while using the proposed 3-point model. The reported estimation power from our 3-point model is evaluated with the following equations:

\[
\begin{align*}
\text{dis}_{1} &= |\text{Power}_{avg}(S)_{|N_{1}} - \text{Power}_{avg}(N_{1})|, \\
\text{dis}_{2} &= |\text{Power}_{avg}(S)_{|N_{2}} - \text{Power}_{avg}(N_{2})|, \\
\text{Estimated Power} &= \frac{1}{\text{dis}_{1}} (\text{power}_{avg}(S)_{|N_{2}} \times \text{dis}_{2} + \text{power}_{avg}(S)_{|N_{1}} \times \text{dis}_{1}), \\
&\text{for dis}_{1} < 0, \\
&\text{dis}_{2} > 0. \\
\end{align*}
\]

From Table II, we can observe that 3-point model does improve the accuracy in most cases. Besides, even with 5 nominal points, the accuracy of the random method is still far behind the proposed 3-point model.

V. SUMMARY AND CONCLUSIONS

In this paper, an efficient nominal point selection method is proposed for power sensitivity based power model. The impact of nominal point selection on the accuracy of power sensitivity based model is also analyzed. Both the theoretical evaluations and the experimental results show that putting the nominal point on the mean of the input transition probabilities is an optimal choice. Furthermore, we propose a 3-point method to achieve even better performance. The proposed 3-point method can be used to build an accurate and efficient IP level power estimator. Further extension of the 3-point model to a 7-point model is possible if the model complexity is of reasonable size. The authors are currently focusing on putting this model into a system level power estimator.

REFERENCES