Models for Power Consumption and Power Grid Noise Due to Datapath Transition Activity *

Lijun Gao and Keshab K. Parhi
Department of Electrical and Computer Engineering
University of Minnesota
200 Union Street S.E., Minneapolis, MN 55455, US
{lgao, parhi}@ece.umn.edu

ABSTRACT
The average power consumption is proportional to the average value of transition activity, i.e., transition probability, and the variance of transition activity determines the strength of power grid noise. In this paper, for the first time, a simple accurate model for estimating the variance of transition activity was proposed, and the dual bit type (DBT) model for estimating the average transition activity was further developed. The model for estimating transition activity variance is based on linearly modeling the spatial correlation of bit-level transition activity, which leads to low computational complexity for computing the variance with very good estimation accuracy. The previous DBT model is made complete with the equation derived in this paper for computing the transition probability beyond the breakdown $BP_b$. In addition to DSP computational architecture and algorithm designs, the proposed simple models are of great significance for power grid noise decoupling and chip floor-planning designs.

1. INTRODUCTION
Over the past few years, the low-power design issues have received much attention mostly due to the emergence of mobile applications, where the average power consumption determines the lifetime of a battery. The average power consumption also plays an important role on the reliability and the packaging cost of high-performance ICs, such as microprocessors which consume large power in order to run at high clock frequency. As semi-conductor technology moves rapidly into deep sub-micron (DSM) stage, the design challenge for delivering reliable ICs becomes more severe due to very low supply voltage, shrinked wire width and much tighter noise margin constraints. Therefore, in DSM technology, the knowledge of the average power consumption alone is not sufficient for designing a reliable IC. Both the peak power consumption and the power grid noise have a heavy weight in design process. These two new factors can cause the malfunction of a digital circuit or the performance degradation of an analog circuit in a mixed signal IC chip, which prevents a system-on-chip (SOC) design from being implemented. From this viewpoint, the accurate knowledge and control of the average power consumption, the peak power consumption and the power grid noise are crucial for successful chip designs.

For CMOS circuits, there are three types of power consumption, i.e., static (or leakage), short circuit and dynamic (or switching). The switching power consumption arises when an output node makes a voltage transition from low to high, which is usually from 0V to power supply $V_{dd}$ [1], and can be represented as

$$P(n) = \alpha(n)C_L V_{dd}^2 f_{clk}$$

where $P(n)$ is the power consumption at clock cycle $n$; $\alpha(n)$ is the transition activity [2][3] at clock cycle $n$, whose value is 1 for transition (0 to 1) or 0 otherwise; $C_L$ is the capacitance load at the output. The average switching power is the average value of equation (1) with respect to $\alpha(n)$ and can be represented as

$$P = \alpha C_L V_{dd}^2 f_{clk}$$

where $\alpha = E[\alpha(n)]$, referred as transition probability in [2][3], and $\alpha(n)$ is assumed stationary. The values of capacitance load of circuits can be extracted using the method presented in [2][3]. To obtain the values of transition probability $\alpha$ for general logic circuits, it requires comprehensive simulations for input vector space. However, for DSP datapaths and components, the values of $\alpha$ can be more readily computed from signal statistical properties. In [2][3], the relationship between transition probability and signal statistics was provided as the dual bit type (DBT) model. In [4][5], more treatments of this issue were presented, where the bit-level lag-1 temporal correlation is modeled instead, and the transition probability is computed from the bit-level temporal correlation and the bit-level probability. However, in [4][5], a complex procedure using signal probability density function (pdf) is required to compute the bit-level probability. In spite of above aspects, none of these three approaches address the second order statistical properties of transition activity, which determine the magnitude of power grid noise.

In this paper, the DBT model is further developed and a for-
mula for computing the transition probability beyond break-
point $BP_3$, is derived. Thereafter, the proposed model for
computing power grid noise is presented. In this paper, all
discussions are limited to datapath situations because up to
40% of the total power may be dissipated in buses, drivers
and multiplexers [6]. For components, the transition activity
will be larger than that of datapaths due to glitching [5] and
this problem can be solved separately.

The remainder of this paper has four sections. In section 2,
the DBT model is summarized and the augmentation to it is
presented. In section 3, a linear model for characterizing the
spatial correlation of bit-level transition activity is proposed
to compute the variance of transition activity and thus
the power grid noise. In section 4, the information about in-
stantaneous and peak power consumptions are illustrated
using transition activity distribution plots. Section 5 is the
conclusion of this work.

2. AVERAGE TRANSITION ACTIVITY

The first part of this section serves as a summary of the
DBT model presented in [2][3]. The remaining part shows
the augmentation to the DBT model, which makes the DBT
model complete.

2.1 Preliminaries

For a stationary signal $x(n)$ with mean value of $\mu = E[x(n)]$,
its variance $\sigma^2$ is given by

$$\sigma^2 = E[(x(n) - \mu)^2] = E[x^2(n)] - \mu^2$$

and its lag-1 temporal correlation $\rho$ is defined as:

$$\rho = \frac{E[x(n) - \mu][x(n-1) - \mu]}{\sigma^2}$$

The square-root of variance, $\sigma$, is referred as standard devi-
ation of the signal. When the signal $x(n)$ is represented as a
$B$-bit binary vector $(x_{n-1}(n), \ldots, x_1(n), x_0(n))$, the quan-
tity $\alpha_i(n)$, given by

$$\alpha_i(n) = x_i(n-1)x_i(n),$$

is referred as bit-level transition activity of $i$-th bit, and the quantity $A(n)$, given by

$$A(n) = \sum_{i=0}^{B-1} \alpha_i(n),$$

is named as total word-level transition activity. If the sum-
mation in equation (6) is made on a fraction of adjacent
bits, then it is called partial word-level transition activity.
The mean values of equations (5) and (6) are referred as bit-
level transition probability and word-level transition proba-
bility, respectively. Since ARMA signal generation models
can generate stationary signals for widespread applications,
as in [4], seven ARMA signals are used in this paper to
generate signals and their definitions are shown in Table 1,
where $v(n)$ is a zero mean white Gaussian signal. For sim-
simplicity, the word-length is chosen as 16 bits throughout this
paper regardless of signal power values. In this paper, all
measurements are made on signals of sample length 40,000.

<table>
<thead>
<tr>
<th>Signal</th>
<th>$x(n)$</th>
<th>Lag-1 temporal correlation $\rho$</th>
<th>Std $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig1</td>
<td>$v(n) - 0.5x(n-1)$</td>
<td>-0.5</td>
<td>0.71</td>
</tr>
<tr>
<td>Sig2</td>
<td>$v(n)$</td>
<td>0</td>
<td>0.68</td>
</tr>
<tr>
<td>Sig3</td>
<td>$v(n) + 0.3x(n-1)$</td>
<td>0.3</td>
<td>0.87</td>
</tr>
<tr>
<td>Sig4</td>
<td>$v(n) + 0.5x(n-1)$</td>
<td>0.5</td>
<td>0.96</td>
</tr>
<tr>
<td>Sig5</td>
<td>$v(n) + 0.7x(n-1)$</td>
<td>0.7</td>
<td>0.91</td>
</tr>
<tr>
<td>Sig6</td>
<td>$v(n) + 0.4v(n-1) - 0.2v(n-2) + 0.05x(n-3) + 0.5x(n-1) + 0.3x(n-2) + 0.15x(n-3) + 0.05x(n-4) - 0.2x(n-5)$</td>
<td>0.9</td>
<td>245</td>
</tr>
<tr>
<td>Sig7</td>
<td>$v(n) + 0.99x(n-1)$</td>
<td>0.99</td>
<td>143</td>
</tr>
</tbody>
</table>

2.2 DBT model for transition probability

To describe the dual bit type (DBT) model in [2][3], two
plots of measured bit-level transition probability are shown
in Fig. 1 for two’s complement representation and Fig. 2 for
sign magnitude representation, respectively. In both figures,
dashed lines computed from DBT model are approximations
to the slope section of original curves.

![Measured bit-level transition probability for two's complement representation.](image)

The approximate linear model of bit-level transition proba-
bility for two’s complement representation in Fig. 1 is rep-
resented as

$$\alpha_i = \begin{cases} 0.25, & 0 \leq i \leq [BP_0] \\ 0.25 - \frac{i - [BP_0]}{BP_1 - BP_0}, & [BP_0] + 1 \leq i \leq [BP_0] + i - [BP_0] + [BP_1] \leq i \leq B - 1 \\ 0, & i = B - 1 \end{cases}$$

and the model for sign magnitude representation in Fig. 2 is
modeled as

$$\alpha_i = \begin{cases} 0.25, & 0 \leq i \leq [BP_0] \\ \frac{i - [BP_0]}{BP_1 - BP_0}, & [BP_0] + 1 \leq i \leq [BP_1] - 1 \leq i \leq B - 2 \\ 0, & i = B - 1 \end{cases}$$

The equations for computing $BP_0$ and $BP_1$ are derived in
[3]. However, they are empirically adjusted in this paper
to allow higher accuracy for larger range of $\rho$ values and
thus minorly different from original ones which also have the
same empirically derived parts. The empirically adjusted equations are shown as
\[ BR_0 = \log_2 \left[ \frac{\sigma(\sqrt{1 - \rho^2} + |\mu|/64)}{\sigma(\sqrt{1 - \rho^2} + |\mu|/3)} \right] \] (9)
\[ BR_1 = \log_2 \left[ \frac{|\mu| + 3\sigma(\sqrt{1 - \rho^2} + |\mu|/3)}{6\sigma(\sqrt{1 - \rho^2} + |\mu|/3)} \right] \] (10)

The models represented by equations (7)-(10) can provide good approximations for signals with small mean value or signals with the lag-1 correlation coefficient \( \rho > 0 \). For signals with both the lag-1 correlation coefficient \( \rho < 0 \) and the ratio of mean/standard deviation larger than 1, the equations (7) and (8) have to be modified appropriately, which is not addressed in this paper due to limited space.

Different from previous works, the breakpoints in this paper are computed as rational numbers instead of integer numbers, such that the approximate linear models have higher accuracy. In [2][3], the comparisons of the computed breakpoints versus the measured breakpoints are not provided. In this work, extensive computations and measurements were made, and the statistical errors of the estimated breakpoints over measured breakpoints are shown in Table 2. The data in Table 2 are obtained from 34 simulations, where \( \rho = -0.5 \sim 0.99 \) and \( \mu/\sigma = 0 \sim 1 \). The data in Table 2 shows that DBT model can provide good estimations of breakpoints \( BR_0 \) and \( BR_1 \). However, the above models are not complete because the equation for computing \( \alpha_{BP_1} \) was not provided in [2][3]. In the next subsection, this equation is derived.

Table 2: Statistical errors of computed v.s. measured breakpoints

<table>
<thead>
<tr>
<th>Two’s Complement</th>
<th>Sign Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>( BR_0 )</td>
<td>( BR_1 )</td>
</tr>
<tr>
<td>Mean of % Error</td>
<td>-2.44</td>
</tr>
<tr>
<td>Std. of % Error</td>
<td>2.88</td>
</tr>
</tbody>
</table>

2.3 Equation for computing \( \alpha_{BP_1} \)
The transition probability beyond \( BP_1 \) is the same as the transition probability of sign bit. If \( x(n) \) and \( x(n - 1) \) are signal values at time instance \( n \) and \( n - 1 \) respectively, the probability can be represented as
\[ \alpha_{BP_1} = \text{Prob}[x(n) < 0, x(n - 1) \geq 0] \]
\[ = \int_{-\infty}^{0} \int_{0}^{\infty} f_{x(n), x(n-1)}(x_1, x_2) dx_1 dx_2. \] (11)
Since all signals are assumed Gaussian, the joint pdf in equation (11) is Gaussian. Using equations (3) and (4), the covariance matrix of \( x(n) \) and \( x(n - 1) \) is obtained as
\[ K = \begin{pmatrix} \sigma^2 & \rho \sigma^2 \\ \rho \sigma^2 & \sigma^2 \end{pmatrix} \] (12)
and its inverse and determinant are
\[ K^{-1} = \frac{1}{\sigma^2(1 - \rho^2)} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \] (13)
and
\[ |K| = \sigma^4(1 - \rho^2), \] (14)
respectively. Then, the joint pdf equals
\[ f_{x(n), x(n-1)}(x_1, x_2) = \frac{1}{2\pi|K|^{1/2}} e^{\frac{1}{2}x^T K^{-1} x} \] (15)
where \( X = [(x_1 - \mu) \ (x_2 - \mu)]^T \). Combining equations (11) and (15), we have
\[ \alpha_{BP_1} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \int_{0}^{\infty} e^{-\frac{1}{2}u^2} erfc \left( \frac{-\mu/\sigma - \mu u}{\sqrt{2(1 - \rho^2)}} \right) du \] (16)
where \( erfc(x) \) is complementary error function, defined as \( erfc(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-t^2} dt \). The equation (16) and equations (7)-(10) are complete linear model for estimating bit-level and word-level transition probability. In Table 3, the computed \( \alpha_{BP_1} \) and \( A \) for signals listed in Table 1 are shown and estimation errors are also shown at bottom. Obtained from the same simulations for Table 2, the statistical errors of the computed \( \alpha_{BP_1} \) and total word-level transition probability \( A \) v.s. the measured values are shown in Table 4.

Table 3: Estimated \( \alpha_{BP_1} \) and \( A \) values

<table>
<thead>
<tr>
<th>Two’s Complement</th>
<th>Sign Magnitude</th>
<th>( \alpha_{BP_1} )</th>
<th>( A )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig1</td>
<td>0.333</td>
<td>4.24</td>
<td>0.333</td>
</tr>
<tr>
<td>Sig2</td>
<td>0.249</td>
<td>4.00</td>
<td>0.249</td>
</tr>
<tr>
<td>Sig3</td>
<td>0.202</td>
<td>3.81</td>
<td>0.202</td>
</tr>
<tr>
<td>Sig4</td>
<td>0.166</td>
<td>3.59</td>
<td>0.166</td>
</tr>
<tr>
<td>Sig5</td>
<td>0.128</td>
<td>3.25</td>
<td>0.128</td>
</tr>
<tr>
<td>Sig6</td>
<td>0.074</td>
<td>2.70</td>
<td>0.074</td>
</tr>
<tr>
<td>Sig7</td>
<td>0.023</td>
<td>1.82</td>
<td>0.023</td>
</tr>
<tr>
<td>Mean of % Error</td>
<td>-0.16</td>
<td>-0.28</td>
<td>-0.16</td>
</tr>
<tr>
<td>Std. of % Error</td>
<td>0.48</td>
<td>0.55</td>
<td>0.48</td>
</tr>
</tbody>
</table>

3. STANDARD DEVIATION OF TRANSITION ACTIVITY AND POWER GRID NOISE
As mentioned in section 1, the average power consumption can not provide enough information for power grid design because mean, standard deviation and maximum value of power supply current all are necessary data for determining the width of power supply wires and the values of decoupling
Table 4: Statistical errors of computed $\alpha_{BP}$ and $A$ v.s. measured values

<table>
<thead>
<tr>
<th>Two's Complement</th>
<th>Sign Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{BP}$</td>
<td>$A$</td>
</tr>
<tr>
<td>Mean of % Error</td>
<td>0.66</td>
</tr>
<tr>
<td>Std. of % Error</td>
<td>1.39</td>
</tr>
</tbody>
</table>

capacitors. Since fluctuations of supply voltage due to bit-level or word-level transition activity inject noise, it can lead to malfunction of digital circuits or performance degradation of analog circuits. In this section, the second order statistics of transition activity and the magnitude of power grid noise are computed with the proposed novel linear model.

3.1 Variance of transition activity

The mean value of bit-level transition activity is $\alpha_c = E[\alpha_c(n)]$ and the equation for computing its variance can be derived by using equation (5) as

$$\sigma^2_{\alpha_c(n)} = E[(\alpha_c(n) - \alpha_c)^2] = \alpha_c^2$$

For word-level transition activity, we only consider total word-level transition activity because partial word-level transition activity can be processed in the same way, and therefore, we use word-level transition activity to refer either total or partial word-level transition activity. The mean of total word-level transition activity is $\alpha = E[A(n)] = \sum_{i=0}^{B-1} \alpha_i$, and its variance can be computed using

$$\sigma^2_{\alpha(n)} = E[\alpha(n) - \alpha]^2 = E[A^2(n)] - \alpha^2.$$  \hspace{1cm} (18)

Since $\alpha^2$ is a known value from the computation in the last section, the problem remains to calculate $E[A^2(n)]$. Using equation (6), we have

$$E[A^2(n)] = E[\sum_{i=0}^{B-1} \sum_{j=0}^{B-1} a_i(n) a_j(n)]$$

$$= \sum_{i=0}^{B-1} \sum_{j=0}^{B-1} E[a_i(n)a_j(n)].$$  \hspace{1cm} (19)

Using definition of correlation coefficient, we define spatial correlation coefficient of bit-level transition activity as

$$\rho_{i,j} = E[\frac{\alpha_i(n) - \alpha_i\alpha_j}{\sqrt{\alpha_i - \alpha_i^2}\sqrt{\alpha_j - \alpha_j^2}}]$$

$$= E[\frac{\alpha_i(n)a_j(n) - \alpha_i\alpha_j}{\sqrt{\alpha_i - \alpha_i^2}\sqrt{\alpha_j - \alpha_j^2}}] \hspace{1cm} (20)$$

where $\alpha_i$ and $\alpha_j$ are known values computed by using DBT model. Then, we obtain following equation

$$E[\sum_{i=0}^{B-1} \sum_{j=0}^{B-1} E[a_i(n)a_j(n)]]$$

$$= \sum_{i=0}^{B-1} \sum_{j=0}^{B-1} [\rho_{i,j}\sqrt{\alpha_i - \alpha_i^2}\sqrt{\alpha_j - \alpha_j^2} + \alpha_i\alpha_j].$$  \hspace{1cm} (21)

Since $\alpha_i$ and $\alpha_j$ are known values, we only need to compute $\rho_{i,j}$.

3.2 Linear model for computing $\rho_{i,j}$

The spatial correlation coefficient matrix has a form of

$$(\rho_{i,j})_{B \times B} = \begin{bmatrix}
1 & \rho_{0,1} & \rho_{0,2} & \cdots & \rho_{0,B-1} \\
\rho_{1,0} & 1 & \rho_{1,2} & \cdots & \rho_{1,B-1} \\
\rho_{2,0} & \rho_{2,1} & 1 & \cdots & \rho_{2,B-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho_{B-1,0} & \rho_{B-1,1} & \rho_{B-1,2} & \cdots & 1
\end{bmatrix}$$  \hspace{1cm} (22)

and is diagonally symmetric. Therefore, only upper triangular part need to be computed. The computational model is based on the observations in Fig 3 and Fig 4, where each curve with different spatial distance corresponds to one of diagonal lines of the matrix $(\rho_{i,j})_{B \times B}$, i.e., spatial correlation coefficient of bit-level transition activity $\rho_{i,l+i}$ for $0 \leq i \leq B - l$. The bit-level transition probability is also plotted as a reference. Similar to bit-level transition probability, the spatial correlation of bit-level transition activity can be approximated with a linear model using $BP_0$ and $BP_1$ as well. The equations of this linear model are shown below in equation (23) for two's complement representation and equation (24) for sign magnitude representation, respectively.

$$\rho_{i,l+i} = \begin{cases}
0, & (0 \leq i \leq |BP_0 - 1|) \\
\frac{l}{BP_0 - BP_0 + 1}, & ([BP_0] \leq i \leq |BP_0 - 1|) \\
1, & ([BP_0] \leq i \leq B - 1)
\end{cases}$$  \hspace{1cm} (23)

$$\rho_{i,l+i} = \begin{cases}
0, & (0 \leq i \leq |BP_0 - 1|) \\
-0.05, & ([BP_0] \leq i \leq |BP_0 - 1|) \\
0, & ([BP_0] \leq i \leq B - 1)
\end{cases}$$  \hspace{1cm} (24)

where $l = 1, 2, \ldots, (B - 1)$ and $(i + l) \leq (B - 1)$.

3.3 Computation of standard deviation of word-level transition activity

Using the linear models presented in equations (23) and (24), and the computed mean values of transition activity in the
Spatial correlation of bit-level transition activity: sign magnitude representation.

In previous subsections, we presented that word-level transition activity $A(n)$ is a time-series which has mean and standard deviation values. In this subsection, the power spectrum of word-level transition activity is computed. Based on this computed spectrum, the spectrum of actual power supply current can be obtained by introducing a weighting function. This weighting function is the system response function of power supply current upon unit transition activity. Due to limited space, the latter issue is not included in this paper.

For discrete process $A(n)$, its autocorrelation function is defined as

$$R(k) = E[A(n)A(n-k)] = E[(A(n) - A)(A(n-k) - A)] + A^2$$

where $A = E[A(n)]$ and $R(0) = \sigma_A^2 + A^2$. The autocorrelation functions of word-level transition activity, which are obtained from simulations, are shown in Fig. 5 for two's complement representation and Fig. 6 for sign magnitude representation, respectively. For both representations, the word-level transition activity has very short correlation time and for simplicity, we can reasonably assume that $R(1) \approx A^2$. As a result, the correlation function $R(k)$ can be simplified as

$$R(k) = \sigma_A^2 \delta(k) + A^2$$

and its Fourier transformation is

$$A(e^{2\pi f}) = \sum_{k=\infty}^{k=-\infty} R(k)e^{-2\pi fk}$$

$$= \sigma_A^2 + A^2 \delta(f)$$

where $f$ is the normalized frequency and $A(e^{2\pi f})$ is the periodic function with a period of one, which corresponds to actual frequency of system clock rate. The equation (27) represents the power spectrum of word-level transition activity and its noise floor has a magnitude of $\sigma_A^2$. For more accurate computation of the noise floor, the value of $R(1)$ can be modeled as a function of $\rho$ and $\sigma_A^2$ by using polynomial curve fitting. In either case, the noise floor is dominated by the item of $\sigma_A^2$.

### 4. DISTRIBUTION OF WORD-LEVEL TRANSITION ACTIVITY

The information about instantaneous and peak power consumption due to word-level transition activity can be obtained from probability distributions of word-level transition activity, which is called probability mass function (pmf) for discrete random variables [7]. In Fig. 7 and Fig. 8, the pmfs, which are obtained from simulations, are shown for two's complement and sign magnitude representations, respectively. Since the sign bits in two's complement representation invoke high strength transition activity, its pmf tends to spread to the right end and thus leads to large mean and standard deviation values. Moreover, two's complement representation produces the strong peak transition activity that is almost as twice high as that by sign magnitude representation.

### 5. CONCLUSIONS
In this paper, a novel method is proposed to compute the variance of word-level transition activity and in turn the power grid noise due to this transition activity. The method is based on linearly modeling the spatial correlation of bit-level transition activity using breakpoints \( BP0 \) and \( BP1 \). In addition, the previous DBT model in \[2][3\] are further developed by deriving the equation to compute bit-level transition probability beyond the breakpoint \( BP1 \). As a result of all these efforts, both average power consumption and power grid noise can be computed at low computational complexity from statistical models of mean \( \mu \), variance \( \sigma^2 \) and lag-1 temporal correlation coefficient \( \rho \). The proposed models can be used towards power grid noise decoupling and chip floor-planning designs as well as DSP computational architecture and algorithm designs presented in \[1][5][6][7\].

6. REFERENCES


