Efficient Estimation of Signal Transition Activity in MAC Architectures

Alberto García, Lukusa D. Kabulepa, Manfred Glesner
Institute of Microelectronic Systems
Darmstadt University of Technology
Karlstrasse 15, Darmstadt, D-64283 Germany
{agarcia, kabulepa, glesner}@mes.tu-darmstadt.de

ABSTRACT

Because of the increasing demand of portable digital systems, it is of great interest to extend the existing high-level power estimation techniques to handle architectures with non linear components, as they appear in relevant practical applications. In this paper we focus on the estimation of the transition activity in MAC structures implementing FIR filters. Based on a divide and conquer approach, an accurate yet efficient estimation procedure is developed. The technique has been evaluated for different synthetic and real data sets. In all cases, our results depict only very slight discrepancies with respect to precise bit level simulations.

Categories and Subject Descriptors

B.8.2 [**Hardware**]: Performance Analysis & Design Aids

General Terms

Design Performance

Keywords

Low power, power estimation, transition activity, MAC

1. INTRODUCTION

The increasing demand on portable applications, and the possibility of integrating in a single chip high performance (and power consuming) digital systems, have shown the importance of including power estimation and optimization techniques in the ICs design flow. Since the impact of these techniques is specially significant at higher levels of abstraction [3], tools working at the architectural and system level are highly desired.

Different approaches ([9, 8, 10, 7]) have been developed to estimate the power consumption in DSP architectures based on the analysis of the number of transitions in a signal after its word level statistics.

In [2] the power consumption for common DSP macro-blocks was related to the total number of transitions at its outputs. Parameterizable analytical expressions were developed to describe these dependencies. Therefore, given an architecture composed by the interconnections of these blocks, it is possible to calculate the power consumption by estimating the transition activity of the different interconnecting signals.

The analysis of the bit transition profile in different common DSP signals, led Landman [6] to the development of the Dual Bit Type (DBT) model. In this model, the transition activity is described by a piecewise linear function, based on three parameters: the LSB breakpoint (BP_0), the MSB breakpoint (BP_1) and the transition activity in the MSB. This analysis was centered in Gaussian distributed data, but in some applications, as in filtering with MAC architectures, other distributions can appear. In [5] a technique for estimating the transition activity under general probability distributions (PDF) is presented, but its application field is restricted to uncorrelated data.

The DBT model has been used in the pioneer work of [8] to estimate the transition activity at the nodes of a MAC architecture. The basic idea of the formulation consists on calculating a modified BP_0 that takes into account the effect of the FIR filter coefficients. Mathematically:

$$BP_0 = \log_2 \sigma_{eff} + \log_2 \left[\sqrt{1-\rho^2} + \frac{|\rho|}{8} \right] \quad (1)$$

where σ_{eff} is calculated after the probability distribution at the output of the MAC multiplier, $f(x)$, solving the non-linear equation:

$$f(\sigma_{eff}) = f(0)\, e^{-1/2} \quad (2)$$

The main disadvantage of this approach is that it requires the computation of a PDF and the solution of a non-linear equation. Additionally, since the correlation of the FIR filter coefficients is not taken into account while calculating σ_{eff}, this method suffers of some lack of accuracy. As a motivation, in fig. 1 the transition activity at the output of the MAC multiplier is represented together with the estimation provided by the aforementioned estimation procedure. We simulated two 20-taps filters with impulse response $h_1 = [a, a, \ldots, a, b, \ldots, b]$ and $h_2 = [a, b, a, b, \ldots, a, b]$ for $a = 3217$ and $b = -219$. Although the σ_{eff} is equal in both cases, the different degree of correlation of the filter coefficients has a significant influence in the transition activity that is not foreseen. A third filter, h_3, similar to h_1 but with $h_3[0] = 3$ was also simulated. The small variation in the filter response produces an exaggerated change in σ_{eff} that does not reflect the real behavior of the MAC.

The rest of the paper is organized as follows: in the next section the estimation problem is formulated and an approximation

ISLPED'02, August 12-14, 2002, Monterey, California, USA.

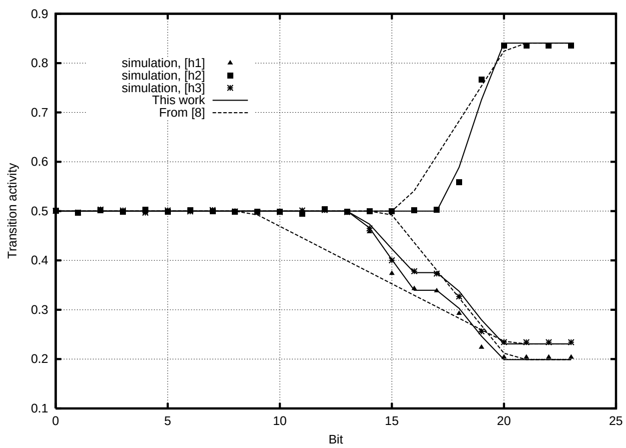

Figure 1: Comparison of the bit level transition activity estimated by [8] and this work.

procedure based on a decomposition approach is presented. This basic formulation is used in section 4 to estimate the transition activity for a complete filter. The resulting approximation technique is validated in section 5 with a set of digital FIR filters. Finally, we present some conclusions.

2. DATA MODEL AND HARDWARE ARCHITECTURE

In a wide spectrum of DSP applications, the incoming data signal can be assumed to have a Gaussian distribution with mean μ, variance σ^2 and correlation coefficient ρ [9]. We consider that the signal is stationary, or at least that its statistical parameters change so slowly that we can divide the signal in quasi-stationary segments. In order to simplify the incoming notation, let us define:

$$\sigma_n \stackrel{\text{def}}{=} \frac{\sigma}{2^B}$$
$$t_m \stackrel{\text{def}}{=} \frac{\arccos(\rho)}{\pi}$$
$$dt \stackrel{\text{def}}{=} 0.5 - t_m = \frac{\arcsin(\rho)}{\pi}$$

where σ_n measure the relative value of σ with respect to the representation range, t_m corresponds to the transition activity in the most significant bit (MSB) of the bus (see [1]), and dt describes the difference between the transition activity in the least and most significant bit.

Let h_i be the impulse response of a k-tap FIR filter. A common architecture for implementing it (see fig. 2) consists of a multiply accumulate (MAC) unit cyclically fed with h_i and the input signal. We assume that both the multiplier and the adder are registered, and therefore no glitches are presented at the output nodes[1].Then, the sequence of values produced at the output of the multiplier (node 3 of fig. 2) are as shown in the fig. 2.

Since the transition activity at nodes 1,2 and 4 has been efficiently estimated in [8], in the present work we will focus on node 3.

[1]This assumption does not decrease the validity of the method because in practice registers are used in order to increase the data throughput rate

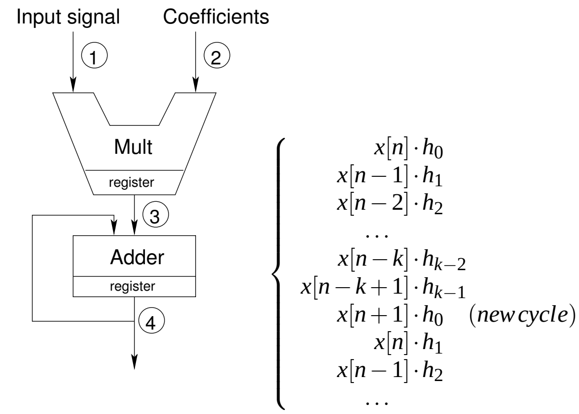

Figure 2: MAC architecture (left) and data sequence at the output of the multiplier (right).

3. PROBLEM DECOMPOSITION.

Let us denote with $T_h(\sigma,\rho,h_a,h_b)$ the transition activity induced after cyclically multiplying a Gaussian signal (with parameters σ and ρ) by the two factors h_a and h_b (see fig. 3). The total transition activity at the node 3 of the MAC can be expressed after this function by:

$$T = \frac{1}{k}\left\{T_h(\sigma,\rho_k,h_{k-1},h_0) + \sum_{i=0}^{k-2} T_h(\sigma,\rho,h_i,h_{i+1})\right\} \qquad (3)$$

where the first summand takes into account that at the beginning of a new filter cycle (i.e. signal from $x[n-k+1]\cdot h_{k-1}$ to $x[n+1]\cdot h_0$), the input data jump k positions and therefore the lag-k correlation, ρ_k, should be used.

Furthermore, we can observe that multiplying a signal with standard deviation σ by the factors $\{h_a,\ h_b\}$ is equivalent to multiply a signal of standard deviation $\sigma\cdot|h_a|$ by $\{1,\ h_b/h_a\}$. Then,

$$T_h(\sigma,\rho,h_a,h_b) = T_h(\sigma\cdot|h_a|,\rho,1,h_b/h_a)$$

If the ratio $|h_b/h_a|$ is significantly smaller than 1, the region where the bits are uncorrelated is determined by $|h_a|$ and practically independent of $|h_b|$ (see fig. 3). Respectively, if $|h_b| >> |h_a|$ the breakpoint is determined by $|h_b|$. Thus with high accuracy we can consider T_h independent of $|h_b/h_a|$ when this ratio is smaller than 0.5.

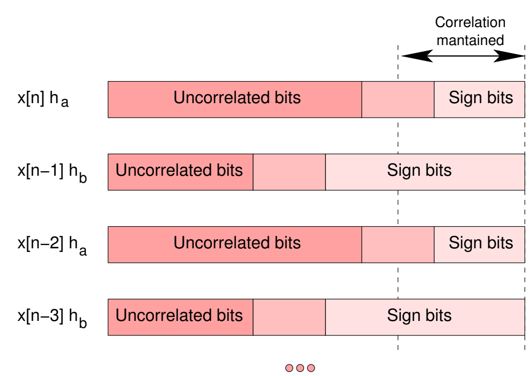

Figure 3: Correlation regions after multipling a Gaussian signal by $\{h_a, h_b\}$.

With respect to the transition activity in the MSB, t_h, we find that it is equal to t_m when the factors h_a, h_b have the same signs, but $t_h = 1 - t_m$ in the other case (i.e. if the sign of the coefficients changes, the sign of the multiplication result toggles only when the sign of $x[n]$ does not vary.) Similar to [4] we use a polynomial in dt to model effect of data correlation in the total transition activity. We get:

$$T_h \approx \frac{B}{2} + dt \left\{ \log_2(2.3\sigma_n h_M) - 7|dt|^3 \right\} dt_h \tag{4}$$

where $h_M = \max(h_a, h_b)$ takes into account the variation in the breakpoints and $dt_h = dt \cdot \text{sign}(h_a \cdot h_b)$ accounts for the effects in the MSB.

When h_a and h_b are close, the total transition activity can be estimated similarly to [4] by:

$$T_h \approx \frac{B}{2} + dt \left\{ \log_2(3.2\sigma_n h_M) - \sqrt{\frac{|dt|}{2t}} \right\} dt_h \tag{5}$$

A wide set of simulations have been carried out to determine the accuracy of the present approach. Fig. 4 compares the measured and estimated total transition activity for 7 synthetic signals with $\sigma = 183$, ρ ranging from 0.2 till 0.99 and 16 bits of word width. The x-axis shows in logarithmic scale the ratio between the two filter coefficients. Around zero, a decrease in the transition activity (model by eq. 5) appears. Outside this region, the transition activity is linear as predicted by eq. 4. Only for highly correlated signals a minor oscillation around -1 and 1 is observed.

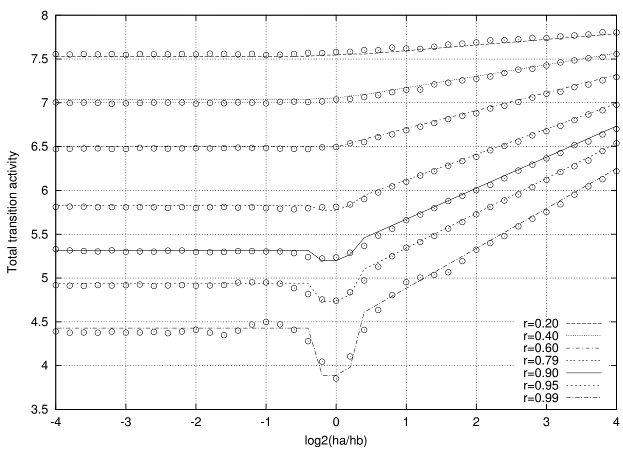

Figure 4: Measured (dots) and estimated (lines) activity versus filter coefficients ratio for signals with different correlation factor.

In most cases, the main interest is in the total transition activity and not in the activity at each single bit. Nevertheless, if required, the breakpoints can be estimated after T_h and dt (in fact, it is the reverse process used in [1] to estimate T_h after the breakpoints). Mathematically, $BP_0 = B_m - \Delta B_m/2$ and $BP_1 = B_m + \Delta B_m/2$ where:

$$B_m = \frac{T_h - B \cdot t_m}{0.5 - t_m} \tag{6}$$

For the separation between the breakpoints, ΔB_m, we found that a good approximation is given by $\log_2 3 + 20 \cdot dt^3$. This approach has been used in fig. 1 to estimate the bit level transiton activity induced by each filter coefficent pair. A very close match between measurements and estimations is observed.

4. TOTAL TRANSITION ACTIVITY ESTIMATION

To calculate the total transition activity at the output of the MAC multiplier, equations 3, 4 and 5 can be used. Nevertheless, in practice, the lag-k correlation of the signal is very small and its effects can be neglected. Further on, the decrease of the transition activity when $h_a \approx h_b$ has an overall small effect on the total activity. Under this simplifications we obtain:

$$T \approx \frac{B}{2} + dt \left\{ \left(\log_2(2.3\sigma_n) - 7|dt|^3 \right) \Delta D_\rho + \Delta B \right\} \tag{7}$$

where the two parameters:

$$\Delta B = \frac{1}{k} \sum_{i=0}^{k-2} \text{sign}(h_i h_{i+1}) \cdot \log_2 \max(h_i, h_{i+1}) \tag{8}$$

$$\Delta D_\rho = \frac{1}{k} \sum_{i=0}^{k-2} \text{sign}(h_i h_{i+1}) \tag{9}$$

determine completely the switching characteristics of the filter.

5. EXPERIMENTAL RESULTS

To probe the accuracy of the presented formulation, the bit transition activity at the output of the MAC multiplier is simulated and compared respectively with the estimations provided by [8] and the present work (eq. 7). The experiment was conducted for three synthetic sets of data with correlation factor 0.87, 0.92 and 0.97, standard deviation 91.4 and a representation width of 12 bits. The data was generated with a lag-one auto regressive model as in [2]. The analyzed FIR structures are 41-tap band pass filters with a band width of $\pi/10$ radians and a normalized (to half the sample rate) central frequency varying from 0.15 to 0.85. A hamming window is used for the FIR design and the coefficients are quantized also with 12 bits. In order to avoid the discretization effects in the LSBs, the first 4 bits are discharged.

As it can be observed in table 1, the error induced by our estimation approach is very small, around one order of magnitude smaller than in previously reported approaches. The errors became smaller for filters with a central frequency close to 0.5 because in this case the output tends to be uncorrelated. When the correlation increases, the transition activity in the most and least significant bit differs more significantly and thus small inexactitudes in the estimation have higher impact on the total accuracy. Notice that the extremely high accuracy obtained in this experiments is due to the fact that the data has an exact Gaussian distribution. In practice, higher errors should be expected with real data sets.

In order to evaluate the proposed estimation procedure with signals from practical applications, a set of experiments is performed using input streams from a band pass sigma delta modulator. The data stream corresponds to the output of the in-phase channel of a simplified OFDM receiver with similar parameters as a HiperLAN-II system. The MAC structure implements the final low-pass filter before the signal decimation by 16. A variety of filters with similar specifications but designed with different windows and lengths are analyzed. The decrease of the transition activity with respect to a random noise signal was simulated and compared with our estimation (see fig. 5). In contrast to [8], a very good track of the small variations in the activity induced by the different filter coefficients can be observed.

f_0	ρ	Trans. activity Total	Trans. activity 4 to 23	error (%) From [8]	error (%) Eq. 7-9
0.15	0.87	10.1	8.3	-6.0	0.1
	0.92	9.8	8.0	-6.4	0.5
	0.97	9.5	7.7	-7.2	0.7
0.25	0.87	10.4	8.8	-4.2	0.3
	0.92	10.2	8.7	-4.4	0.5
	0.97	10.0	8.4	-4.8	0.7
0.35	0.87	11.2	9.4	-2.9	0.1
	0.92	11.1	9.3	-3.0	0.2
	0.97	11.0	9.2	-3.3	0.2
0.45	0.87	11.8	9.9	-1.2	0.1
	0.92	11.8	9.9	-1.2	0.1
	0.97	11.7	9.9	-1.2	0.3
0.55	0.87	11.9	10.1	1.2	-0.0
	0.92	11.9	10.1	1.4	0.0
	0.97	11.9	10.0	1.5	-0.1
0.65	0.87	12.4	10.6	2.7	0.0
	0.92	12.4	10.6	2.9	-0.0
	0.97	12.5	10.7	3.2	-0.0
0.75	0.87	12.7	11.1	3.8	0.0
	0.92	12.9	11.3	3.9	-0.1
	0.97	13.0	11.5	4.2	-0.2
0.85	0.87	13.5	11.7	4.5	-0.0
	0.92	13.8	12.0	4.8	-0.3
	0.97	14.0	12.2	5.1	-0.3

Table 1: Error in the estimation of the transition activity for different band pass filters.

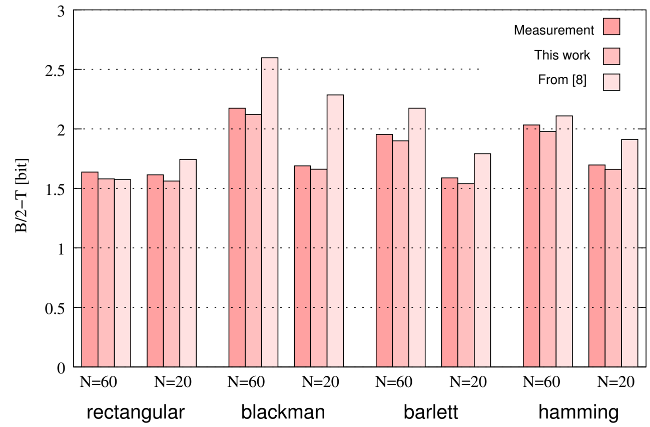

Figure 5: Difference in the transition activity with respect to random noise for a set of low pass FIR filters designed with different windows.

6. CONCLUSIONS

In this paper, an efficient technique for estimating the node transition activity in MAC architectures implementing FIR filters has been proposed. The emphasis has been placed on the estimation of the activity at the output of the MAC multiplier, since at this node, previously proposed approaches required complex and thus slow computation methods. Our method does not only overtake this limitation, but also results in more accurate estimations. Additionally, the switching characteristics of the filter have been related to two parameters, ΔB and ΔD_ρ (eq. 8 and 9 respectively) that can be used as metrics for comparing the power consumption of different FIR filters.

Different experiments with both synthetic and real data sets have been reported to prove the validness of the approach.

Since the implementation of a n-tap FIR filter with a MAC architecture is based on the sharing of the multipliers and adders inside each tap, this work represents a first step towards the estimation of the transition activity in shared architectures.

7. REFERENCES

[1] S. Bobba, I. N. Hajj, and N. R. Shanbhag. Analytical expressions for average bit statistics of signal lines in DSP architectures. In *Proc. of the IEEE ISCAS 1998*, pages 33 – 36, 1998.

[2] S. Bobba, I. N. Hajj, and N. R. Shanbhag. Analytical expressions for power dissipation of macro-blocks in DSP architectures. In *12th Int. conf. on VLSI design*, pages 33 – 36, January 1999.

[3] A. P. Chandrakasan and R. W. Broderson. *Low Power Digital CMOS Design*. Kluwer Academic Publishers, 1995.

[4] A. García, L. D. Kabulepa, and M. Glesner. Estimation of power consumption in encoded data buses. In *Design, Automation and Test in Europe*, March 2002.

[5] A. García, L. D. Kabulepa, and M. Glesner. Transition activity estimation for generic data distributions. In *IEEE Int. Symp. on Circuits and Systems*, May 2002.

[6] P. Landman and J. Rabaey. Architectural power analysis: the dual bit type method. In *IEEE Trans. on VLSI Systems*, pages 173 – 187, June 1995.

[7] M. Lundberg, K. Muhammad, K. Roy, and S. K. Wilson. High-level modeling of switching activity with application to low-power DSP system synthesis. In *Proc. of the Int. Conf. on Acoustic, Speech and Signal Processing*, volume 4, pages 1877–1880. IEEE, 1999.

[8] S. Nikolaidis, E. Karaolis, and E. D. Kyriakis-Bitzaros. Estimation of the transition activity in MAC architectures implementing FIR filters. *IEEE Trans. on CAD*, pages 164 – 179, January 2000.

[9] S. Ramprasad, N. R. Shanbhag, and I. N. Hajj. Analytical estimation of signal transition activity from word level statistics. *IEEE Trans. on CAD*, pages 718 – 733, July 1997.

[10] J. H. Satyanarayana and K. K. Parhi. Theoretical analysis of word-level switching activity in the presence of glitching and correlation. *IEEE Trans. on VLSI Systems*, pages 148 – 159, April 2000.