

Analytical High-level Power Model for LUT-based Components

Ruzica Jevtic and Carlos Carreras

Dpto. de Ingeniería Electrónica, E.T.S.I. Telecomunicación, Universidad Politécnica de Madrid, Ciudad Universitaria s/n, Madrid, Spain
{ruzica,carreras}@die.upm.es

Abstract. This paper presents an extended high-level model for logic power estimation of multipliers and adders implemented in FPGAs in the presence of glitching and correlation. The model is based on an analytical computation of the switching activity produced in the component and the FPGA implementation details of the component structure. It is extended to consider operands of different word-lengths, both zero-mean and non-zero mean signals, and the glitching produced inside the component, taking into account the sign nature of the autocorrelation coefficients of the components' inputs. The number of simulations needed for the model characterization is extremely small and can be reduced to only two. As the final power model is analytical, it is capable of providing power estimates in milliseconds. The results show that the mean relative error is within 10% of low-level power estimates given by the XPower tool.

1 Introduction

Due to their low cost and ability for reconfiguration, FPGAs have become an ideal solution for various embedded designs. As they are aimed for implementation of many different designs, a large number of the routing and logic resources of the FPGA architecture is inefficiently used. This in turn, causes a significant increase in their power consumption and creates the need for FPGA power optimization.

The algorithmic complexity of recent and upcoming DSP applications is growing fast, leading to severe time penalties when power optimization is carried out at low levels of abstraction. Thus, it is necessary to use power optimization techniques at the earliest possible time, which results in development of new high-level models and methods.

Most high-level power estimation techniques use multivariable regression over a large number of accurate power estimations at the transistor level, obtained through simulations for different input statistics, in order to represent power consumption in the form of an equation with variable parameters depending on the input and output signal statistics. However, at high levels of abstraction the design architecture has not been defined yet. Therefore, a new set of simulations is required for the module characterization each time the module's parameters

change. There are some high-level power models that try to avoid this problem by constructing models that are parameterizable in terms of different combinations of input's word-lengths [4, 8]. In these cases, the dependency of the models on different input signal statistics is expressed in terms of coefficient values stored in a table, obtained again through extensive simulations aimed to cover a large number of different signal statistics. Most of the existing approaches use interpolating techniques in order to estimate power for signal statistics different from those used in the simulations.

We present the methodology which overcomes the above mentioned problems. It supports power consumption estimation of multipliers and adders implemented in FPGAs considering both zero-mean and non-zero mean signals and any autocorrelation coefficient value. The resulting power models are parameterizable in terms of input signal statistics and input word-lengths without a need of constructing a large table of coefficients. Furthermore, the glitching produced inside the component is also modelled, taking into account the sign nature of the autocorrelation coefficients of the components' inputs. The approach adopted in this paper extends the basic concepts presented in [5] where only zero-mean signals with positive autocorrelation coefficients were considered.

The paper is organized as follows. Section 2 highlights the previous work done in the area of high-level power estimation. Section 3 presents some preliminaries of the work used for estimating signal transition activity from the word-level statistics obtained from non-zero mean signals. In Section 4, we present an improved power model, with special attention on modelling the glitching effects when the zero-mean constraint is relaxed. Experimental results are given in Section 5. We conclude this paper in Section 6.

2 Related Work

High-level estimation models can be divided into three groups according to the characterization of the input data set. The first one [2, 8], is based on n chosen bit-level signal statistics such as transition probability of input bits, input spatial correlation between the bits inside the same input word, input signal probability, etc. All these statistics appear in the model as average values obtained from Data Flow Graph simulations and are introduced into a power equation as variables. Coefficients standing by the variables are obtained through extensive simulations and are listed into an n -dimension array where each dimension corresponds to one of the bit-level statistics. The main flaw of this model is that it can be applied only to a specific component with a fixed word-length, so a new set of simulations is needed for model characterization each time this parameter changes.

The second power estimation group is based on power macromodels constructed by using the spatio-temporal correlation defined as Hamming distance (ie. the number of bit transitions between two consecutive input vectors). Beside this parameter, many models also use Signal distance, which describes the number of input bits that are fixed to logic one in two consecutive input vectors. The models are based on a very large number of low-level simulations, although cer-

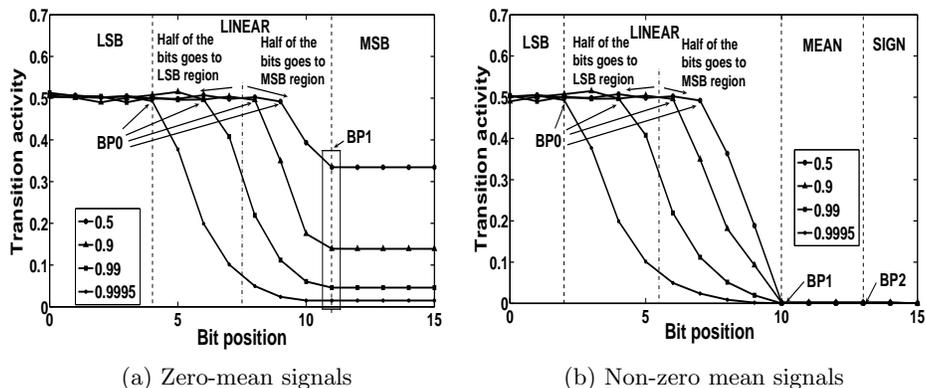


Fig. 1: Bit transition activity vs. bit position in a word for different autocorrelations

tain modifications permit a trade-off between the accuracy and the time needed for model characterization [3, 4].

The third power estimation group considers word-level signal statistics such as variance, mean and autocorrelation coefficient. The work in [1], presents characterized equations for arithmetic components implemented in FPGAs, based on a table of coefficients for different signal statistics. As the only variable introduced in the equations is the component size, this methodology is not capable of producing estimates for components with operands of different word-lengths. In [6], the authors first noted that a signal word can be divided into three regions according to its word-level signal statistics: LSB uncorrelated bits, correlated bits belonging to the so-called linear region and MSB bits. Based on the activity analysis of signal-word division, they build black-box model of the capacitance switched in each activity region of the module obtained through extensive simulations. This signal word division has also been used in the work presented here, but instead of measuring the activity of the bits in each region, it has been estimated from word-level statistics and adapted to non-zero mean signals as it is explained in the next section.

3 Signal Model

In this work, we consider gaussian signals with mean μ , variance σ^2 and autocorrelation coefficient ρ . As already mentioned, our approach is based on a signal-word division into 3 regions (see Fig. 1a): LSB data bits with switching activity of 0.5 as they behave as uncorrelated bits (in Fig. 1a it is marked for $\rho = 0.9995$), MSB data bits with switching activity that strongly depends on word-level signal statistics, and a linear region which can be approximated by a linear interpolation of the previous two regions [6]. As presented in [5], the expression for the breakpoint $BP0$ which separates the LSB region from the linear

region is:

$$BP0 = \lceil \log_2(\sqrt{1 - \rho^2} \cdot \sigma_x) \rceil \quad (1)$$

The expression for the breakpoint $BP1$, which separates the linear region and the MSB region is taken from [6] and represents the number of bits in a signal that are needed to cover the signal variation around its mean:

$$BP1 = \lceil \log_2(3\sigma) \rceil \quad (2)$$

For a zero-mean signal these two breakpoints are sufficient in order to account for the contribution of each of the regions to the total switching activity of the component. However, in the case of non-zero mean signals with a gaussian distribution, the previous signal model does not account for some important effects. In the extended model for non-zero mean signals, the region beyond breakpoint $BP1$ is transformed into two subregions with mean and sign bits (see Fig. 1b). The activity of these regions is zero, regardless of the autocorrelation coefficient, but the values of these bits depend on the value and the sign of the mean respectively, and as such, have a great impact on power consumption. Fig. 2 represents a regional decomposition of the array multiplier according to its input activity regions. It can be seen that, the mean bits of one operand equal to '1' will cause the switching activity propagation of the other operand on the corresponding outputs of the AND gates and thus, also on the outputs of the component's full-adder cells. On the other hand, if these bits are '0's, the switching activity at the outputs of all basic elements in this region will be zero, regardless of the switching activity of the other operand. Hence, it is clear that the values of the mean bits contribute significantly to the component's total power, although they do not exhibit any switching activity.

The new breakpoint is obtained as follows. The maximum value of non-zero mean gaussian signals is $\mu + 3\sigma$ and $m = \log_2(\mu + 3\sigma)$ bits are needed for its binary representation. As previously mentioned, $\log_2(3\sigma)$ bits are needed for the representation of the signal variation around the mean. Thus, there will be $x = \log_2(\mu + 3\sigma) - \log_2(3\sigma)$ bits that are not changing in a data word and whose values correspond to the x upper bits of the mean. If a signal word has N bits, then this region is followed by $N - m$ bits, all taking value '0' or '1' depending on the sign of the mean. These are the bits that form the sign region. As a consequence, the MSB region of the signal-word when considering non-zero mean signals, is composed only of half of the bits belonging to the linear region. Hence, the third breakpoint $BP2$, which separates the mean region from the sign region, is calculated as follows:

$$BP2 = \lceil \log_2(\mu + 3\sigma) \rceil \quad (3)$$

The switching activities of the bits in each region are computed as in [7]:

$$t_i = 2 \cdot p_i \cdot (1 - p_i) \cdot (1 - \rho_i) \quad (4)$$

where p_i is the bit probability and ρ_i is the bit-level autocorrelation coefficient, which can be approximated by ρ for the MSB bits, has a value 0 for the LSB bits and a value 1 for the mean and sign bits.

output of an AND gate respectively, all of them located in a corresponding row i and column j . The method used for the switching activity computation at the output of these elements is based on the switching activities of the input bits and the probabilities of these bits being '0' or '1'. This method is explained in detail in [5]. For non-zero gaussian signals, the bit probabilities depend on the signal statistics as explained in Section 3. Once the constant a is computed, formula (6) can be used for power estimation of any other component size and signal statistics. In order to obtain a power estimate, it is only necessary to re-compute the total switching activity for the new input parameters.

4.1 An Improved Glitching Model

The glitching occurs due to the different signal delays entering the same logic component. Its amount is directly proportional to the transition activities of the inputs. Thus, we consider that the most significant amount of glitching produced inside the component is generated by the most active regions of its inputs.

When considering zero-mean gaussian signals and only positive autocorrelation coefficients, LSB input regions exhibit the highest switching activity. This can be deduced directly from the equation (4). In this context, glitching has been modelled in [5] as the sum of average equivalent glitching produced by each cell belonging to the LSBx-LSBy region of the component (see Fig. 2). However, for negative autocorrelation coefficients, it is clear that MSB regions exhibit a higher switching activity than all neighbouring regions. Therefore, a new glitching model has to be developed to account for the contribution of these regions as well. This is achieved by introducing a new factor into an expression for glitching that depends on the value of the autocorrelation coefficient, as follows.

Assuming that the MSB bits have an equal probability of being '0' or '1', the expression for their switching activity according to (4) becomes:

$$t_i = 0.5 \cdot (1 - \rho) \quad (8)$$

As the LSB bits have a switching activity of 0.5, the relationship between the switching activity of these two regions can be expressed as a coefficient $l = 1 - \rho$. At the same time, this is the relationship we expect between the average glitching produced in the MSB and LSB part, as the amount of glitching is proportional to the transition activity of the input bits. Hence, the extended glitching model which represents a sum of glitching in the four component's regions (LSBx-LSBy, LSBx-MSBy, MSBx-LSBy and MSBx-MSBy in Fig. 2) is expressed as:

$$G = k \cdot \sum_{i=1}^4 (1 - \rho_{1i}) \cdot (1 - \rho_{2i}) \cdot FA_i = k \cdot G' \quad (9)$$

where G is the amount of glitching, k is an empirically derived constant which represents the average glitching at the output of one LUT in the LSBx-LSBy part of the component, ρ_{1i} and ρ_{2i} are the bit-level autocorrelation coefficients of the LSB/MSB regions of inputs (depending on the particular region of the

component), and FA_i is the number of full-adder cells in the corresponding component's region. The final model for estimating the power consumption in the presence of glitching and autocorrelation is given as follows:

$$P = b \cdot (SW + k \cdot G') \quad (10)$$

Two low-level power measurements for different component sizes using the same ρ are sufficient in order to determine coefficients b and k . As the factors SW and G' are known and P is measured, the coefficients can be easily obtained. However, in order to increase the accuracy of the model, we use a multivariable regression approach with more than two measurements for obtaining these two coefficients. The number of measurements is still significantly smaller than any other existing high-level approach for building power macro-modules. It is clear that the model is parameterizable in terms of the operands word-lengths and the input signal statistics.

5 Experimental results

We split the model evaluation into two sets of experiments. In the first set we evaluate the accuracy of the power estimation model for non-zero mean signals presented in (10), considering both components with inputs of the same size and components where the input bit-widths differ. In the second part, we focus on evaluating the power model for zero-mean signals against the estimates obtained by the power model described in [5]. Both sets of experiments have been performed on multipliers and adders implemented as Xilinx IP Cores in Virtex II devices. All the estimated values have been compared to low level power estimates provided by the Xilinx tool XPower [9]. The signals used as input stimuli had Gaussian distributions with means equal to 0, 10 and 125 respectively. We have chosen these values for the mean in order to see the difference between the power values obtained for signals with many and a few '1's in their mean.

We have used 16x16 and 32x32 multipliers and signals with zero-mean gaussian distributions and autocorrelation coefficients of 0, 0.9, -0.9 and -0.99 for our characterization set. Multiple regression over relative error was performed for obtaining the constants b and k .

Fig. 3 presents the estimation errors for multipliers and adders when operands have the same word-lengths. Input word-lengths varied between 8 and 40 bits, and autocorrelation coefficients varied between -0.9995 and 0.9995. It can be noted that a similar error performance is obtained for all adders when signals with mean 10 and mean 125 are applied to its inputs. This is a direct consequence of a feature of the operation *add*. As the mean bits do not change their value, when added, they will give the same result at the outputs of the full-adder cells, and hence, there will be no switching activity in this part of the adder. This result has also been confirmed by the identical XPower values.

Next, Fig. 4 shows the errors obtained for multipliers and adders with different operand sizes. The experimental set includes three autocorrelation coefficients of -0.99, 0 and 0.99. Again, the adder errors for mean values of 10 and

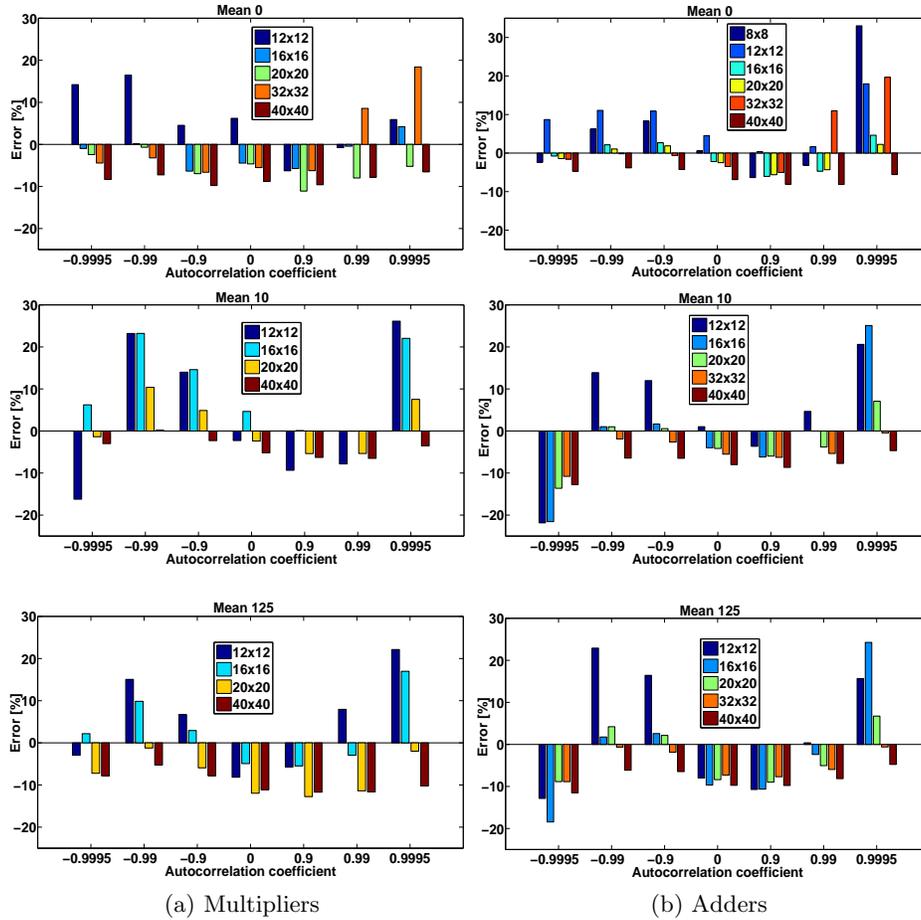


Fig. 3: Error performance for multipliers and adders with operands of the same size

125 were almost the same, so we have plotted only one of them (Fig. 4e). Another important observation from this figure is that the power model for adders clearly underestimates when non-zero mean signals are applied to its different-size inputs. This is due to the fact that the glitching model does not consider the amount of glitching produced in the part of the adder where the input bits of one operand belong to the mean region, while the input bits of the other operand belong to the LSB or MSB region. In reality, there will be glitching in this part of the adder due to the different arrival times of the carry bit and the input bits of the other operand. However, the glitching model presented here, does not take this into account. Overall, the mean relative error for multipliers is 8.34% and for adders 11.14%. It can be seen that the models are capable of giving quite accurate results over a wide range of component sizes, signal autocorrelation coefficients and mean values.

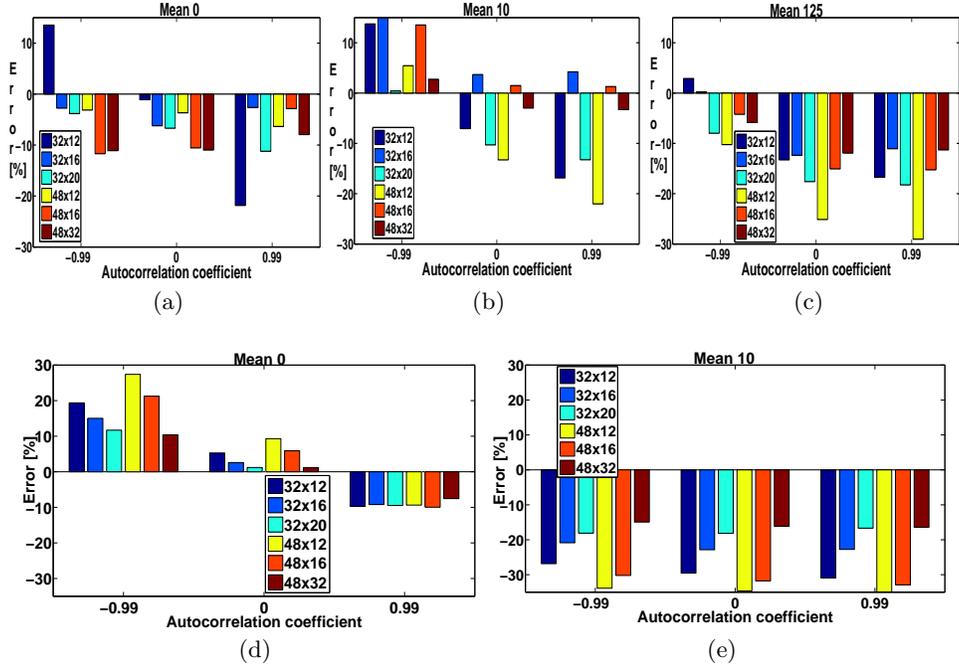


Fig. 4: Error performance for multipliers - a), b), c) and adders - d), e) with different size operands for various signal statistics

The second set of experiments evaluates adders and multipliers with input word-lengths between 8 and 40 bits. The autocorrelation coefficient values are only positive and vary between 0 and 0.9995, to allow comparison with the model in [5] which does not support negative autocorrelation coefficients.

For an easier comparison, we give the absolute value of the errors in Fig. 5. It can be seen that the accuracy of the new model does not worsen when positive autocorrelation coefficients are considered, while providing power estimates for a significantly larger set of input signal parameters.

6 Conclusion

We have presented a high-level power estimation model for DSP components implemented as IP cores in FPGAs. The model is parameterized in terms of the input signal statistics and the operands' word-lengths. It accounts for the different power behaviour observed when considering zero-mean and non-zero mean input signals. The number of simulations needed for model characterization is extremely reduced compared to other high level power models. The results show that the model is accurate within 10% over a wide range of input signal statistics and bit-widths.

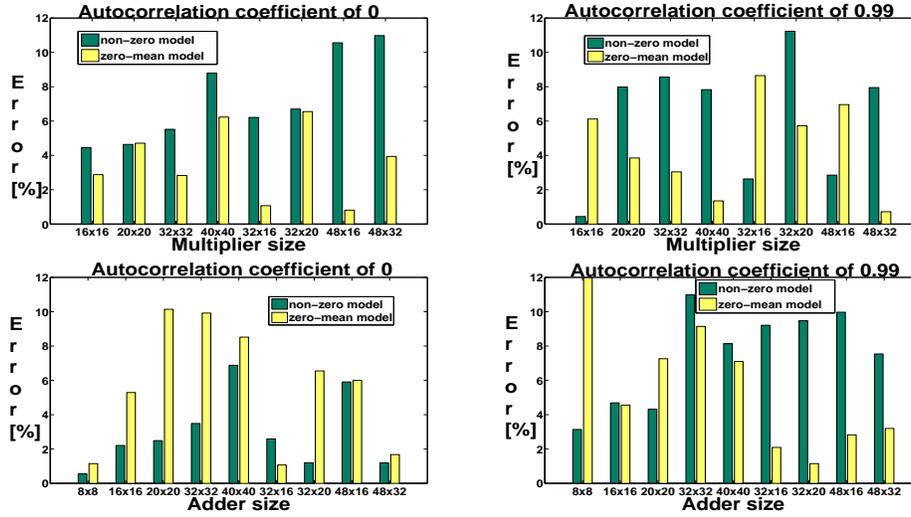


Fig. 5: Error performance for multipliers and adders for various signal statistics considering two approaches: the model presented here and the model presented in [5]

Acknowledgment

This work was supported in part by the Spanish Ministry of Education and Science under project TEC2006-13067-C03-03/TCM.

References

1. J. A. Clarke, A. A. Gaffar, G. A. Constantinides, and P. Y.K. Cheung, *Fast word-level power models for synthesis of FPGA-based arithmetic*, Proc. ISCAS (2006) 1299-1302
2. S. Gupta, and F. N. Najm, *Power Modeling for High Level Power Estimation*, IEEE Trans. VLSI Syst., vol. 8, (Feb. 2000) 18-29
3. D. Helms, E. Schmidt, A. Schulz, A. Stammermann, and W. Nebel, *An Improved Power Macro-Model for Arithmetic Datapath Components*, PATMOS 2002, LNCS 2451, (2002) 16-24
4. G. Jochens, L. Kruse, E. Schmidt, and W. Nebel, *A New Parameterizable Power Macro-Model for Datapath Components*, DATE '99, (1999) 29-36
5. R. Jevtic, C. Carreras and G. Caffarena, *Switching Activity Models for Power Estimation in FPGA Multipliers*, ARC'07, LNCS, vol. 4419, (March 2007) 201-213
6. P. Landman, and J. Rabaey, *Architectural Power Analysis: The dual bit type method*, IEEE Trans. On VLSI Systems, vol. 3, no. 2, (1995) 173-187
7. S. Ramprasad, N. R. Shanbhag, and I. N. Hajj, *Analytical Estimation of Signal Transition Activity from Word-Level Statistics*, IEEE Trans. On Computer-Aided Design of Integrated Circuits and Systems, vol. 16, no. 7, (July 1997) 718-733
8. L. Shang, and N. K. Jha, *High-level Power Modeling of CPLDs and FPGAs*, in Proc. of the Int. Conf. on Comp. Design. IEEE Computer Society, (2001) 46-53
9. Xilinx Inc. www.xilinx.com