Switching Activity Reduction of MAC-Based FIR Filters with Correlated Input Data

Oscar Gustafson[1], Saeeid Tahmasbi Oskuii[2],
Kenny Johansson[1], and Per Gunnar Kjeldsberg[2]

[1] Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden, Email: {oscarg, kennyj}@isy.liu.se
[2] Department of Electronics and Telecommunications, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway, Email: {saeeid.oskuii, per.gunnar.kjeldsberg}@iet.ntnu.no

Abstract. In this work we consider coefficient reordering for low power realization of FIR filters on fixed-point multiply-accumulate (MAC) based architectures, such as DSP processors. Compared to previous work we consider the input data correlation in the ordering optimization. For this we model the input data using the dual bit type approach. Results show that compared with just optimizing the number of switches between coefficients, the proposed method works better when the input data is correlated, which can be assumed for most applications.

Key words: FIR filter, MAC, dual bit type, switching activity, coefficient reordering

1 Introduction

Energy consumption is becoming the major cost measure when implementing integrated circuits. This trend is motivated both by the benefit of increased battery life for portable products as well as reducing cooling problems. Many of these systems include a digital signal processing (DSP) subsystem which performs a convolution or a sum-of-product computation. These computations are often performed using a, possibly embedded, programmable DSP processor.

The probably most common form of convolution algorithms is the finite-length impulse response (FIR) filter. The output of an N:th-order FIR filter is computed as

$$y(n) = \sum_{i=0}^{N} h(i)x(n-i) \tag{1}$$

where the filter coefficients, $h(n)$, determine the frequency response of the filter. The transfer function of the FIR filter is

$$H(z) = \sum_{i=0}^{N} h(i)z^{-i} \tag{2}$$

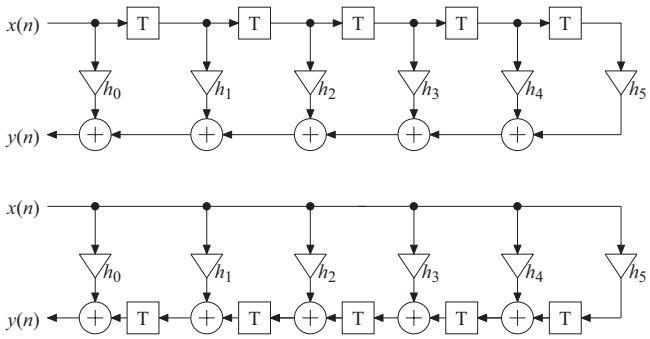

Fig. 1. (above) Direct form and (below) transposed direct form fifth-order FIR filter.

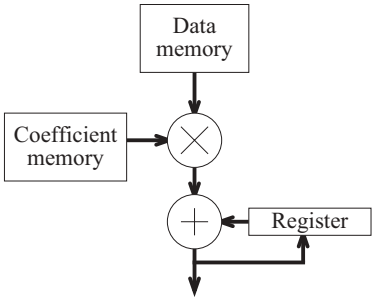

Fig. 2. Multiply-accumulate (MAC) architecture suitable for realizing direct form FIR filters.

The two most common filter structures for realizing the transfer function in (2) are the direct form and the transposed direct form structures depicted in Fig. 1. As can be seen from Fig. 1 the basic arithmetic operation is a multiplication followed by an addition. This is usually called a multiply-accumulate (MAC) operation and is commonly supported in programmable DSP processors [1]. If a direct form FIR filter is realized the input data is stored in one memory, while the coefficients are stored in another memory. Then each output is computed by performing $N + 1$ MAC operations. An abstracted suitable architecture is shown in Fig. 2. It is also possible to use a similar architecture when implementing FIR filters in application specific integrated circuits (ASICs) and field-programmable gate arrays (FPGAs) [2, 3]. Many FPGAs have dedicated general fixed-point multipliers and some even have a complete fixed-point MAC as a dedicated building block.

For integrated circuits implemented in CMOS technology the sources of power dissipation can be classified as dynamic, short circuit, and leakage power. Even though the impact of leakage power increases with decreasing feature size, the main source of power dissipation for many integrated circuits is still the dy-

namic power. The dynamic power consumption for a CMOS circuit is expressed as

$$P_{\text{dynamic}} = \alpha C f V_{DD}^2 \tag{3}$$

where α is the switching activity, C is the capacitance, f is the frequency, and V_{DD} is the power supply voltage.

In this work we focus on reducing the switching activity in fixed-point MAC-based realizations of direct form FIR filters. The same ideas can be applied to other convolutions and sum-of-product computation, but for clarity we will only discuss FIR filters here. We focus on reducing the number of switches on the inputs of the multiplier. It should be noted that this will also decrease the number of switches on the buses connecting the memories to the multiplier.

Previous work considering switching activity reduction of FIR filters on MAC-based architectures can be divided in two class. For the first class it is assumed that the MAC operations are performed in increasing (or decreasing) order and the approaches optimize the coefficient values such that the number of switches between adjacent coefficients is small [4, 5]. In [4] a heuristic optimization method is proposed, while in [5] an optimal method based on mixed integer linear programming is presented. The second class aims at reordering the computations such that the number of switches between succeeding coefficients are small [4, 6]. In [4] a framework that both optimizes the coefficients and the order was proposed. However, the optimization and reordering are not performed simultaneously. Furthermore, the work in [4] neglects the fact that the input data is correlated. Input data correlation is treated in [6] by determining a lookup table based on simulation. This lookup table will grow rapidly with increased filter length, and, hence, only short filters and convolutions are considered in [6].

There are also works where several output samples are computed interleaved with, typically, more than one accumulator, leading to reduced switching activity [7, 8].

In this work we characterize the input data using the dual bit type method [9] and derive equations for computing the correlation between samples more than one sample period apart. This is used to formulate a Hamiltonian path (or traveling salesman, TSP) problem that is solved to find the best ordering of the computations. While the focus of this work is on FIR filters, similar techniques can be applied for other applications based on subsequent MAC-operations.

In the next section we review the issues related to correlated input data and derive the correlation equations for the input data. Then, in Section 3 the proposed optimization approach is presented. This approach is extended to include possible negation of coefficients in Section 3.2. In Section 4 results are presented that highlight the importance of the contribution. Finally, in Section 5 some concluding remarks are given.

2 Correlated Input Data

Signals in real world applications can in many cases be approximated as a Gaussian stochastic variable. This leads to that their binary representations have

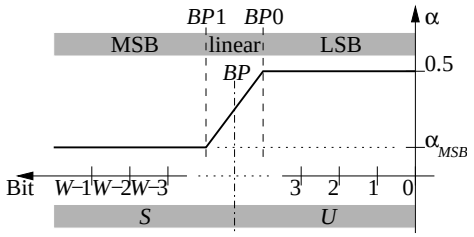

Fig. 3. Illustration of the dual bit type properties.

different switching probabilities for different bit positions. However, certain properties for different regions can be observed [9–11]. In this work we focus on two's complement representation, but a similar derivation can be performed using, e.g., sign-magnitude representation.

The dual bit type (DBT) method [9] is based on the fact that the binary representation of most real world signals can be divided into a few regions, where the bits of each region have a well defined switching activity. In [9] the three regions LSB, linear, and MSB was defined as illustrated in Fig. 3. Because of the linear approximation of the middle region, it was stated that two bit types is sufficient. In the LSB region the switching probability is 1/2, which corresponds to random switching. Hence, the bits are divided into a uniform white-noise (UWN) region, U, and a sign region, S, as shown in Fig. 3.

The word-level statistics, i.e., mean, μ, variance, σ^2, and correlation, ρ, of the signal are used to determine the breakpoints and the switching activities. The correlation for a signal is computed as

$$\rho = \frac{\mu_\Delta + \mu^2}{\sigma^2} \tag{4}$$

where μ_Δ is the average value of the signal multiplied by the signal delayed one sample. Typically, we have $\mu = 0$, which gives that the probability of a bit in the two's complement representation being one, p, is 1/2.

In [9] the break points of the regions are defined as

$$BP0 = \log_2 \sigma + \log_2(\sqrt{1-\rho^2} + \frac{|\rho|}{8}) \tag{5}$$

$$BP1 = \log_2(|\mu| + 3\sigma) \tag{6}$$

$$BP = \frac{BP0 + BP1}{2} \tag{7}$$

With a data wordlength of W bits the number of bits in the region S is

$$W_S = W - BP - 1 \tag{8}$$

If p_Q is the probability that a single-bit signal, Q, is one and the temporal correlation of Q is ρ_Q, then the switching activity, a_Q, of Q is defined as [10]

$$\alpha_Q = 2p_Q(1-p_Q)(1-\rho_Q) \tag{9}$$

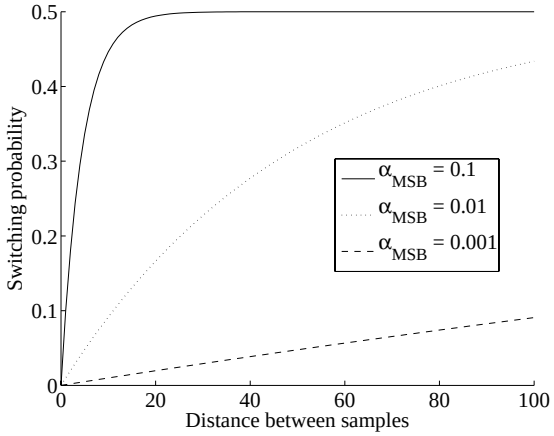

Fig. 4. Resulting switching activity for the MSB region, S, after reordering.

The probability for a one is assumed to be $1/2$ for all bits in the twos-complement representation of a signal, as the mean value is 0. Furthermore, it was stated in [10] that the temporal correlation for bits in the MSB region is close to the word-level correlation, ρ. Hence, the switching activity in the MSB region can be computed from as

$$\alpha_{MSB} = \frac{(1-\rho)}{2} \tag{10}$$

Now, when the filter coefficients are reordered, the switching probability for adjacent data words is changed. Let $\alpha_{m,D}$ denote the switching probability between two bits at position, m, with time index i and $i+D$. We have

$$\alpha_{m,D} = \begin{cases} 0 & D = 0 \\ \begin{cases} 1/2 & m \in S \\ \alpha_{MSB} & m \in U \end{cases} & D = 1 \\ (1-\alpha_{m,1})\alpha_{m,D-1} + \alpha_{m,1}(1-\alpha_{m,D-1}) & D \geq 2 \end{cases} \tag{11}$$

In Fig. 4 the effect of reordering on the switching probability is shown for some initial switching probabilities, α_{MSB}. From this it can be seen that the switching probability increases monotonically toward $1/2$ with increasing distance between samples. Hence, while reordering may decrease the switching probability of the coefficients, it will increase the switching probability of the input data sent to the multiplier.

3 Proposed Approach

Let the coefficient $h(i)$ be represented using a B-bit two's complement representation as

$$h(i) = -b_{i,B-1} + \sum_{k=0}^{B-2} b_{i,k} 2^{-(B-1-k)} \tag{12}$$

where $b_{i,j} \in \{0, 1\}$. Hence, the number of switches when changing the coefficient from $h(i)$ to $h(j)$ (or vice versa) is

$$c_{h,i\to j} = c_{h,j\to i} = \sum_{k=0}^{B-1} b_{i,k} \oplus b_{j,k} \tag{13}$$

This is the Hamming distance measure used for coefficients in [4–6].

For the input data it is not possible to explicitly compute the number of switches. Instead the switching probability from (11) is used to obtain

$$c_{x,i\to j} = c_{x,j\to i} = \sum_{k=0}^{W-1} \alpha_{k,|i-j|} \tag{14}$$

The total transition cost for selecting $h(j)$ as the input coefficient after $h(i)$ is now

$$c_{tot,i\to j} = c_{tot,j\to i} = c_{h,i\to j} + c_{x,i\to j} \tag{15}$$

By forming a fully connected weighted graph with $N+1$ nodes, where each node correspond to a filter coefficient, $h(i)$, and each edge weight is obtained from (15) the ordering problem becomes a problem of finding a path visiting all nodes in the graph once with minimum weight. This problem is known as a symmetric traveling salesman problem (TSP). The TSP-problem is NP-hard, but it is in general possible to solve rather large instances in reasonable time. We have used GLPK [12] to solve problems with about 100 coefficients to optimality in tens of seconds.

3.1 Multiplier and Bus Power Consumption

It should be noted that switches at different inputs to the multiplier affects the power consumption differently [13]. Hence, if it is possible to characterize the used multiplier it is also possible to weight the $c_{h,i\to j}$ and $c_{x,i\to j}$ terms. However, the results in [13] also indicate that the variation is not large. Our own simulations also show that when all inputs are randomly distributed with a switching and one probability of 0.5 except for one which is known to switch every cycle, the variation in power consumption is insignificant. Hence, this aspect is not included in the results.

For the cases that we will implement a custom multiplier and not use an existing one in a DSP, an FPGA, or a macro library, it is worth noticing that it is possible to optimize the power consumption of the multiplier based on the expected switching probability [14].

For buses the traditional power model has been to count the number of switches. However, for deep sub-micron technology the interwire capacitances is dominating over the wire-to-ground capacitances [15]. Hence, one should possibly include these in the cost function as well. In general it is hard to determine the exact bus structure, especially for DSPs and FPGAs, and, hence, in the results section we only consider the number of switches.

3.2 Selective Negation

In [4] it was proposed that if the MAC operation is able to conditionally subtract the output of the multiplier, it is possible to negate some of the coefficients to reduce the switching even further. However, they did not provide any solution as how to decide which coefficients to be negated. In this work we evaluate these ideas by considering the case that we change all coefficients to positive values and selectively subtract the results that corresponds to negative coefficients. This could in general be solved using a modified TSP formulation known as the equality generalized TSP (E-GTSP) [16].

4 Results

To illustrate the results of the proposed design technique we will consider three FIR filters of varying lengths. These will be optimized according to the methodology in Section 3 using different data distributions. All FIR filters are designed using the Remez exchange algorithm for the given specifications. For simplicity we assign the same weights for the passband and the stopband ripples. The filter coefficients are scaled by a power of two such that the magnitude of the largest coefficients is represented using all available bits, i.e., $0.5 \leq \max(|h(i)|) < 1$. Finally, the coefficients are rounded to the used wordlength. It should be noted that the used wordlength are enough even for harder specifications [17]. Hence, it would be possible to design filters with shorter wordlengths for most designs. However, this aspect is not considered here, but the rounded coefficients are used to demonstrate the properties of the proposed reordering methodology.

4.1 Design 1

For the first design the passband and stopband edges are at 0.2π rad and 0.3π rad, respectively. For this design we aim at a general purpose DSP with a 24×24-bit multiplier. With a filter order of 65 we obtain the results shown in Table 1, where the switching activity denotes the total number of switches at the bus and multiplier inputs for a complete FIR filter computation. From the results it can be seen that savings in switching activity between 1.5% and 7.2% are obtained taking the correlation of the input data into account.

4.2 Design 2

For the second design, we will consider implementation in an FPGA which includes a general 18×18-bits multiplier. Again we use an FIR filter designed using the Remez algorithm with identical maximum passband and stopband ripples. For the passband and stopband edges we select 0.6π rad and 0.8π rad, respectively. The filter coefficients are scaled as in the previous design. To obtain reasonable stopband attenuation we select a filter order of 40.

As the statistical properties of the input signal are estimated it is of interest to know how a change in the actual input characteristics affects the switching

Data characteristics	Natural order[3]	Optimized for coefficient Hamming distance [4]	Optimized using data characteristics	Reduction
Random	1472.0	1020.0	1020.0	-
$W_S = 4, \alpha_{MSB} = 0.1$	1366.4	1012.0	996.6	1.5%
$W_S = 4, \alpha_{MSB} = 0.01$	1342.6	943.5	929.2	1.5%
$W_S = 8, \alpha_{MSB} = 0.1$	1260.8	1004.0	964.6	3.9%
$W_S = 8, \alpha_{MSB} = 0.01$	1213.3	867.0	837.7	3.4%
$W_S = 12, \alpha_{MSB} = 0.1$	1155.2	996.1	924.0	7.2%
$W_S = 12, \alpha_{MSB} = 0.01$	1083.9	790.5	744.3	5.8%

Table 1. Total switching activity of the data and coefficient values for Design 1.

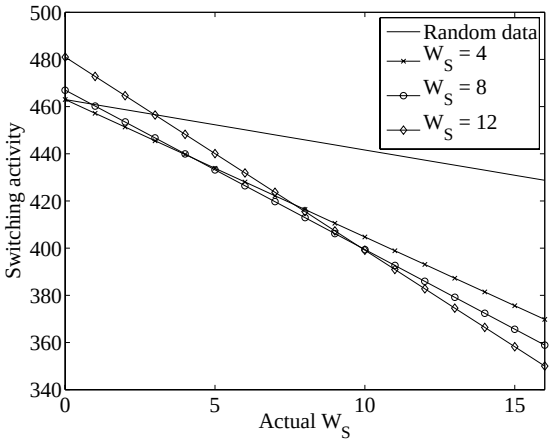

Fig. 5. Resulting estimated switching activity for different coefficient orders with respect to the actual W_S. $\alpha_{MSB} = 0.1$

activity. For this we have considered the case where $\alpha_{MSB} = 0.1$ and varied W_S for four different designs, each optimized for coefficient Hamming distance [4] (corresponding to $W_S = 0$) and $W_S = 4, 8$, and 12. The results are shown in Fig. 5, where it can be seen that the designs that consider the input data correlation in most cases result in less switching activity compared to not considering (as in [4]) independent of the actual value of W_S. It is only for small W_S:s that the orderings designed for large W_S are worse compared with the ordering designed for Hamming distance (corresponding to $W_S = 0$). Hence, the proposed design methodology will reduce the switching activity as long as the estimated parameters are reasonably close to the actual parameters of the input data.

4.3 Design 3

In the third design, again we consider implementation in an FPGA. Now we consider the effect of making all signs positive. This can easily be realized by

Data characteristics	Original sign		Positive sign		Reduction	
	Natural	Optimized	Natural	Optimized	Natural	Optimized
Random	1890.0	1152.0	1604.0	1128.0	15.1%	2.1%
$W_S = 6, \alpha_{MSB} = 0.1$	1635.6	1111.1	1349.6	1068.1	17.5%	3.9%
$W_S = 6, \alpha_{MSB} = 0.01$	1578.4	963.8	1292.4	929.9	18.1%	3.5%
$W_S = 10, \alpha_{MSB} = 0.1$	1470.0	1062.0	1184.0	1000.5	19.5%	5.8%
$W_S = 10, \alpha_{MSB} = 0.01$	1374.9	829.2	1088.9	779.2	20.8%	6.0%

Table 2. Total switching activity of the data and coefficient values for Design 3.

replacing the adder in Fig. 2 with an adder/subtracter. To control this an additional bit is required in the coefficient memory. It should be noted that using this approach is similar to using the sign-magnitude number representation. This sign bit should possibly be included in the switching activity analysis. From a bus point of view it should be included while from a multiplier point of view it should not, as it does not affect the power consumption of the multiplier. We choose to not include it in the cost function in this design study. For this design we consider a 105:th-order FIR filter with passband and stopband edges at 0.5π rad and 0.53π rad, respectively.

The results are shown in Table. 2 for the cases where the original sign is used and when all coefficients are transformed into having positive sign. It is clear that while significant savings can be obtained by using absolute valued coefficients when realizing the multiplications in their natural order, the advantage decreases when coefficient reordering is considered. The reason for the larger savings using natural ordering is that the value of the coefficients vary for the impulse response leading to many switches for the sign bits. For the optimized version this is already considered through the optimization. One would expect that by using the E-GTSP formulation the switching activity can be reduced even further.

5 Conclusion

In this work we have proposed an approach to low-power realization of FIR filters on MAC-based architectures when the input data correlation is considered. The reordering of computations to reduce the switching activity now also depends on the input data correlation, which is represented using the dual bit type method. Furthermore, we proposed how to form a problem when we consider the possibility to negate coefficients to reduce the switching activity further. The proposed approach provide a more accurate modeling compared to [4] which did not consider input data correlation. The results show that as long as we have correlated input data and the dual bit type parameter estimation is reasonably correct we obtain lower switching activity using the proposed methodology compared to [4]. Compared to [6] the proposed method can handle arbitrary large FIR filters, while the modeling in [6] depended on simulations making it complex to obtain results for long filters due to the large look-up tables required. For this work the corresponding results are easily computed from the presented equa-

tions. However, this requires that the input data is characterized for use of the dual bit type method.

References

1. P. Lapsley, J. Bier, A. Shoham, E. A. Lee, *DSP Processor Fundamentals: Architectures and Features*, Wiley-IEEE Press, 1997.
2. L. Wanhammar, *DSP Integrated Circuits*, Academic Press, 1999.
3. U. Meyer-Baese, *Digital Signal Processing with Field Programmable Gate Arrays*, Springer, 2001.
4. M. Mehendale, S. D. Sherlekar, and G. Venkatesh, "Low-power realization of FIR filters on programmable DSPs," *IEEE Trans. VLSI Systems*, vol. 6, no. 4, pp. 546–553, Dec. 1998.
5. O. Gustafsson and L. Wanhammar, "Design of linear-phase FIR filters with minimum Hamming distance," in *Proc. IEEE Nordic Signal Processing Symp.*, Hurtigruten, Norway, Oct. 4–7, 2002.
6. K. Masselos, P. Merakos, S. Theoharis, T. Stouraitis, and C. E. Goutis, "Power efficient data path synthesis of sum-of-products computations," *IEEE Trans. VLSI Systems*, vol. 11, no. 3, pp. 446–450, June 2003.
7. T. Arslan and A. T. Erdogan, "Data block processing for low power implementation of direct form FIR filters on single multiplier CMOS DSPs," in *Proc. IEEE Int. Symp. Circuits Syst.*, Monterey, CA, 1998, vol. 5, pp. 441–444.
8. K. K. Parhi, "Approaches to low-power implementations of DSP systems," *IEEE Trans. Circuits Syst.–I*, vol. 48, no. 10, pp. 1214–1224, Oct. 2001.
9. P. E. Landman and J. M. Rabaey, "Architectural power analysis: The dual bit type method," *IEEE Trans. VLSI Systems*, vol. 3, no. 2, pp. 173–187, June 1995.
10. S. Ramprasad, N. R. Shanbhag, and I. N. Hajj, "Analytical estimation of transition activity from word-level signal statistics," in *Proc. Design Automat. Conf.*, June 1997, pp. 582-587.
11. M. Lundberg, K. Muhammad, K. Roy, and S. K. Wilson, "A novel approach to high-level swithing activity modeling with applications to low-power DSP system synthesis," *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3157–3167, Dec. 2001.
12. Gnu Linear Programming Kit 4.16, http://www.gnu.org/software/glpk/
13. S. Hong, S.-S. Chin, S. Kim, and W. Hwang, "Multiplier architecture power consumption characterization for low-power DSP applications," in *Proc. IEEE Int. Conf. Elec. Circuits Syst.*, Dubrovnik, Croatia, Sept. 15–18, 2002.
14. S. T. Oskuii, P. G. Kjeldsberg, and O. Gustafsson, "Transition-activity aware design of reduction-stages for parallel multipliers," in *Proc. Great Lakes Symp. on VLSI*, Stresa-Lago Maggiore, Italy, March 11–13, 2007.
15. P. Caputa, H. Fredriksson, M. Hansson, S. Andersson, A. Alvandpour, and C. Svensson, "An extended transition energy cost model for buses in deep submicron technologies," in *Proc. Int. Workshop on Power and Timing Modeling, Optimization and Simulation*, Santorini, Greece, Sept. 15–17, 2004, pp. 849–858,.
16. M. Fischetti, J. J. Salazar Gonzâlez, and P. Toth, "The generalized traveling salesman problem," in *The Traveling Salesman Problem and Its Variations*, Kluwer Academic Publishers, pp. 609–662, 2002.
17. D. M. Kodek, "Performance limit of finite wordlength FIR digital filters," *IEEE Trans. Signal Processing*, vol. 53, no. 7, pp. 2462–2469, July 2005.