

# Low Complexity and High Performance Multiplier based on Multi-Level Approximate Compressor

Vishnumaya C, Sajitha A.S.

*ECE Department, Nehru College of Engineering & Research Centre, Pampady, Thrissur, Kerala*

**Abstract** - Approximate computing has been considered to improve the accuracy-performance trade-off in error-tolerant applications. The proposed work focuses on novel approximate compressors to exploit them for the design of efficient approximate multipliers. The latency of the critical path is reduced to less than three XOR logic gates, which is a considerable boosting for the speed performance. Further, the proposed approximate multiplier is applied to FIR Filter for its performance analysis. The approximate compressors are a key element in the design of power-efficient approximate multipliers. The gate-level delay has been reduced considerably and the carry rippling problem in the compressors is reduced by proposed 5-2 compressor. Further, higher order 6:3 compressor based on symmetric stacking can also be added to the design to further reduce the critical path delay.

**Keywords:** 5-2 Compressor, delay, XOR gate, Carry rippling problem, multiplier

## I. INTRODUCTION

Because of their higher efficiency, parallel multipliers are now extremely used in many high-speed systems such as digital signal processors (DSPs), central processing units (CPUs), and multimedia applications. In most of the CPUs, the multiplier lies in the critical path for signal propagation. To decrease the delay and complexity of such systems, practical design considerations have been pursued over recent years. In a parallel multiplier, the multiplication process is divided into three steps. At first, the partial products (PPs) are generated. Then these products are summed, and the process continues until two rows remain, and at the final stage, the two remaining rows will be added by means of, for example, a carry propagation adder. At the second stage and after generation of PPs, a partial product reduction tree (PPRT) is often employed for efficient summation of the products. Considering the full adder (FA) as the main building block in various multiplication configurations, this block constitutes the basis for much different architecture.

However, the main drawback of FA-based configurations, which limits their usage in today's parallel multiplier design, is the propagation latency for the cascaded cells. Moreover, the most important concern in the design process of a parallel multiplier is the circuit size and power consumption, which is directly related to the number of employed gates used in the various parts of the architecture. An efficient solution to overcome such

drawbacks is to utilize a compressor network instead of FA trees, especially in PPRT. Furthermore, comprehensive analysis depicts that the most significant part of the total delay and power dissipation will belong to this stage in a parallel multiplier. Therefore, the performance enhancement of this stage can significantly improve the speed and lower the power dissipations of the whole system.

Design methodology for ultrahigh-speed 5-2 and 7-2 compressors has already been established. With the help of this procedure, the gate-level delay has been reduced considerably when compared with the previous designs, while the total transistor and gate count remain in a reasonable range. By starting the discussion for the carry rippling problem in  $n - 2$  compressors, the method has been developed for 5-2 compressor and is expanded for 7-2 architecture, which shows 32% and 30% improvement in speed performance for these structures, respectively. Existing method optimizes the partial product summation through row compression techniques in the Wallace tree or Dadda tree.

Higher order 6:3 compressor based on symmetric stacking can also be added to the design to further reduce the critical path delay. It uses 3-bit stacking circuits, which group all of the "1" bits together. The bit stacks are then converted to binary counts, producing 6:3 counter circuits with no XOR gates on the critical path. The overall partial product is divided into MSP (Most Significant Part) and LSP (Least Significant Part). All the compressors are kept exactly in MSP part and approximated in the LSP part. The approximation is done by eliminating the XOR gates to achieve power and area minimization. FIR filter is designed using the proposed multiplier to analyze its performance power and area.

## II. DESIGN CONSIDERATIONS

### A. Literature Review

The art of summing up the numbers with minimum carry propagation delay is one of the common speed improvement techniques utilized in state-of-the-art digital circuits. The basic idea is to reduce three numbers to two numbers with the help of an FA, which is the general definition of a 3-2 counter block. The 4-2 compressor is the simplest form for the realization of an  $n-2$  compressor.

In conventional form, it is constructed by means of cascaded attachment of two FAs. In a similar fashion, a 5-2 compressor is simply obtained by cascading three FAs. Because three FA blocks have been arranged after each other, at least five XOR logic gate-level delay is expected for such structure. By means of the optimizations reported in, the gate-level latency was reduced to four XOR logic gates. There are also other 5-2 compressor circuits which have their own advantages and drawbacks. The common disadvantage between all of the previous works is the rippling of the horizontal outputs (Cout1 or Cout2) for at least three consecutive stages. This problem, which has barely been investigated before, drastically degrades the performance of the compressor cell. Hence, the real gate level latency will be equal to five XOR logic gates for the circuits reported in and . Also, the configuration, which is based on conventional architecture, contains no speed enhancement when compared with . Moreover, the corresponding latencies for compressors Cell. Hence, the real gate-level latency will be more than six XOR logic gates. The rippling problem (for at least three stages) also exists for higher order compressor block. The emphasis over the other works reported in the literature proves the hypothesis. Herein, we are about to demonstrate that if the conventional truth table of each compressor is investigated carefully, then the rippling effect can be reduced considerably. Also, other parameters like power and active area can be maintained at a moderate level.

**B. Carry Rippling Problem**

One of the essential key factors on the performance enhancement of an  $n - 2$  compressor is the reduction of carry rippling issues between adjacent compressor cells (for 5-2 or higher order structures). To clarify such an issue, which directly increases the total delay of the compressor block, we can consider the horizontal cascading of three 5-2 compressor structures, as illustrated in Fig. 1. In Fig. 1(a), there is a logical dependence between Cout2 and Cin1, which is highlighted with the dashed lines for the middle stage compressor.

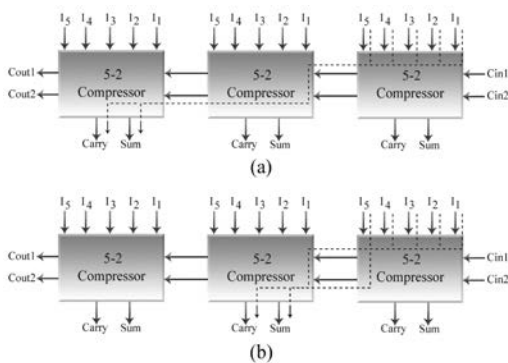


Fig. 1. Horizontal cascading of 5-2 compressors for clarification of carry rippling problem. (a) Three-stage carry rippling. (b) Two-stage carry rippling.

Therefore, the input of the first-level compressor will be rippled to the third-stage compressor structure due to this dependence. This obstacle will ultimately degrade the speed performance of the designed architecture, which can clearly be seen in all of the previously reported works. On the other hand, if the design considerations will be built on the independency of Cout1 and Cout2 outputs from input Cin1 and Cin2 bits, then the structure of Fig. 1(b) will be obtained. As it is evident, the carry rippling has been reduced for two cascaded stages.

**III. 5-2 COMPRESSOR**

As illustrated in Fig, the 5-2 compressor architecture consists of seven inputs and four outputs. The parameters denoted as I1, I2, I3, I4, and I5 constitute the primary inputs, while Cin1 and Cin2 are the secondary inputs that obtain their values from the adjoining compressor of one binary bit order lower in significance. The weightings of all seven inputs are identical. The Sum output possesses the same weight as the inputs, while the other three outputs (Carry, Cout1, and Cout2) contain one binary bit higher order than the other bits. The outputs Cout1 and Cout2 are injected into the adjacent compressor block of higher significance. A set of cascaded six logic gates constitute the critical path of this structure. Therefore, the gate-level delay will optimistically be six XOR logic gates. This is the essential key factor that degrades the performance of such an arrangement.

The existing circuit for 5-2 compressor has been demonstrated in Fig 3. For the better realization of the latency in different paths from inputs to the outputs, two neighbour 5-2 compressor cells must be connected together as the carry rippling finishes in the second compressor block. It must be taken into account that all of the MUX gates in Fig except the one that produces Cout1, and the gate which is fed by Gnd, are channel-ready gates. As a result, they will exhibit a latency equal to 0.25.

Moreover,  $B^-$  is produced by means of an inverter at the output node of the nonfull swing XOR gate that generates B. However, it does not affect the delay of the whole system. The reason is that when B is generated, it is being fed to the gates of TG transistors in the MUX . For high logic values at the input stage of TG, the output capacitor will be charged to  $V_{dd} - V_{th}$  by means of the NMOS transistor. At the same time,  $B^-$  is transferred to the output of the inverter gate. With the help of  $B^-$ , which enables the PMOS transistor, the output capacitor will fully be charged to  $V_{dd}$  value. For low logic states in the input stage of TG, the NMOS path itself discharges the output capacitor to the zero value. The two output XOR-XNOR gates generate both outputs simultaneously, which are designed and wholly discussed in . Because of the similar paths, their different outputs have little influence on the delay and glitch effect of TG waveforms. Also,

static CMOS has been employed to design the architectures of NAND and NOR gates.

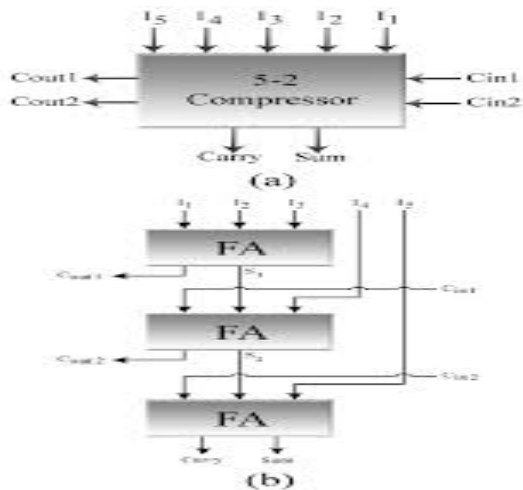


Fig.2. 5-2 compressor. (a) General architecture. (b) Conventional implementation.

Moreover, B<sup>-</sup> is produced by means of an inverter at the output node of the nonfull swing XOR gate that generates B. However, it does not affect the delay of the whole system. The reason is that when B is generated, it is being fed to the gates of TG transistors in the MUX. For high logic values at the input stage of TG, the output capacitor will be charged to V<sub>dd</sub>-V<sub>th</sub> by means of the NMOS transistor. At the same time, B<sup>-</sup> is transferred to the output of the inverter gate. With the help of B<sup>-</sup>, which enables the PMOS transistor, the output capacitor will fully be charged to V<sub>dd</sub> value. For low logic states in the input stage of TG, the NMOS path itself discharges the output capacitor to the zero value. The two output XOR-XNOR gates generate both outputs simultaneously, which are designed and wholly discussed in . Because of the similar paths, their different outputs have little influence on the delay and glitch effect of TG waveforms. Also, static CMOS has been employed to design the architectures of NAND and NOR gates.

If the outputs of these gates are inverted, then the AND/OR gates will be obtained, and since these gates have not been located in the critical path, therefore, they would not affect the latency of the whole system. To calculate the delay of the critical path for the proposed 5-2 compressor in the logic gate level, we refer to the calculations provided in the previous section. As stated above, the critical path belongs to the route starting from I<sub>1</sub> and I<sub>2</sub> and finishing in Carry output. Considering the defined values for the delay, we have  $\tau_{5-2} = \tau_{XORd} + \tau_{MUX} + \tau_{XORf} + \tau_{MUX} + \tau_{MUX}$  (23) where  $\tau_{MUX}$  illustrates the latency of channel-ready MUX. Substitution of the obtained values in (23) results in  $\tau_{5-2} = + 0.25 + 0.75 + 0.5 + 0.25 = 2.75$ . As a result, the latency of the critical path is reduced to less than three XOR logic gates,

which is a considerable boosting for the speed performance.

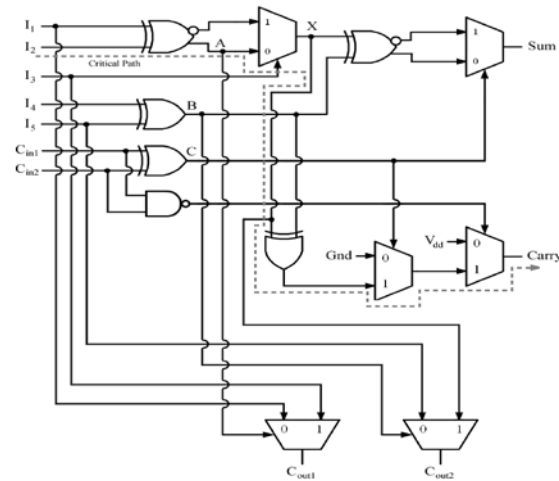


Fig 3.5-2 Compressor using XOR gates.

If the outputs of these gates are inverted, then the AND/OR gates will be obtained, and since these gates have not been located in the critical path, therefore, they would not affect the latency of the whole system. To calculate the delay of the critical path for the proposed 5-2 compressor in the logic gate level, we refer to the calculations provided in the previous section. As stated above, the critical path belongs to the route starting from I<sub>1</sub> and I<sub>2</sub> and finishing in Carry output. Considering the defined values for the delay, we have  $\tau_{5-2} = \tau_{XORd} + \tau_{MUX} + \tau_{XORf} + \tau_{MUX} + \tau_{MUX}$  (23) where  $\tau_{MUX}$  illustrates the latency of channel-ready MUX. Substitution of the obtained values in (23) results in  $\tau_{5-2} = + 0.25 + 0.75 + 0.5 + 0.25 = 2.75$ .

As a result, the latency of the critical path is reduced to less than three XOR logic gates, which is a considerable boosting for the speed performance. For a better comparison, the 5-2 compressor architecture reported in will be analyzed by means of the defined values for the gate-level latency. The critical path of such a structure starts from I<sub>1</sub>, I<sub>2</sub>, and I<sub>3</sub>. The signal will pass through the CGEN1 block, which demonstrates a latency equal to 2. After that and in the second compressor block, a channel-ready MUX gate will transmit the signal to the third adjacent compressor as Cin2 input. Finally, a double-output XOR gate, along with one full swing XOR architecture, will produce the Carry output. Hence, the delay of the critical path will be equal to  $\tau_{5-2} = 2 + 0.25 + 0.75 = 4$ . The comparison of the values obtained in and demonstrates that a significant speed enhancement equal to 32% has been achieved by reducing the carry rippling problem by one stage.

#### IV. 7-2COMPRESSOR

As illustrated in Fig, the 7-2 compressor architecture consists of nine inputs and four outputs. By extending the

concept of 5-2 compressor, the inputs denoted as I1–I7 will organize the primary inputs while Cin1 and Cin2 are the secondary inputs which obtain their values from the adjoining compressor of one binary bit order lower in significance. As shown in Fig, the conventional realization of a 7-2 compressor can be obtained by the serial connection of five FAs. Following the previous discussions, a gate-level delay of seven XOR logic gates is expected for the critical path. The best-reported design with some optimizations could only achieve six XOR logic gates delay, if the lone input (which is applied to the second level FA) is eliminated and its corresponding FA block is replaced with a half adder (HA), then a 6-2 compressor can also be obtained. But the simple fact associated with 6-2 and higher order compressors is that they have more than eight inputs. As a result, the outputs will be at least 4 bits. Therefore, the lower limit for the latency from inputs to the outputs will be four XOR logic gates, which pertains to the Sum output.

To calculate the delay of the critical path, it is evident that the critical path belongs to the route starting from I3 and I4 and finishing in Carry output of adjacent compressor, we have

$$\tau_{7-2} = \tau_{XORd} + \tau_{MUX} + \tau_{MUX} + \tau_{XORd} + \tau_{MUX} + \tau_{AND} + \tau_{MUX}$$

where  $\tau_{XORd} + \tau_{MUX} + \tau_{MUX}$  demonstrates the latency for the generation of Carry 4 – 2 output because its generation is the same as Carry output from the 4-2 compressor.

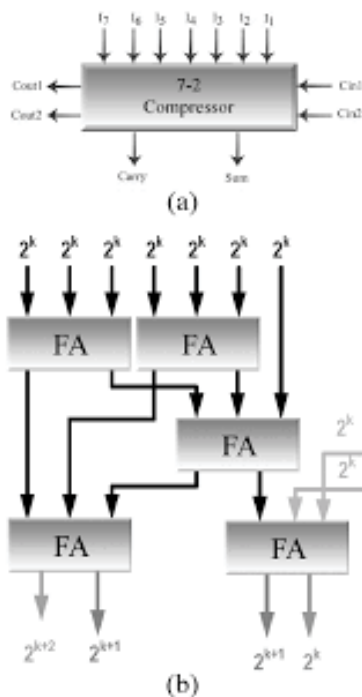


Fig. 4. 7-2 Compressor (a) Schematic (b) Conventional form

By substituting the obtained values for the computation of gate-level delay and also, by assuming the latency of AND/OR gates equal to that of defined for double-output XOR gate, we will obtain  $\tau_{7-2} = 0.25 + 0.25 + 0.25 + 0.25 = 4$ . Therefore, a significant speed enhancement equal to 30% will be achieved by careful design along with the reduction of carry rippling problem by one stage. If the outputs of these gates are inverted, then the AND/OR gates will be obtained, and since these gates have not been located in the critical path, therefore, they would not affect the latency of the whole system.

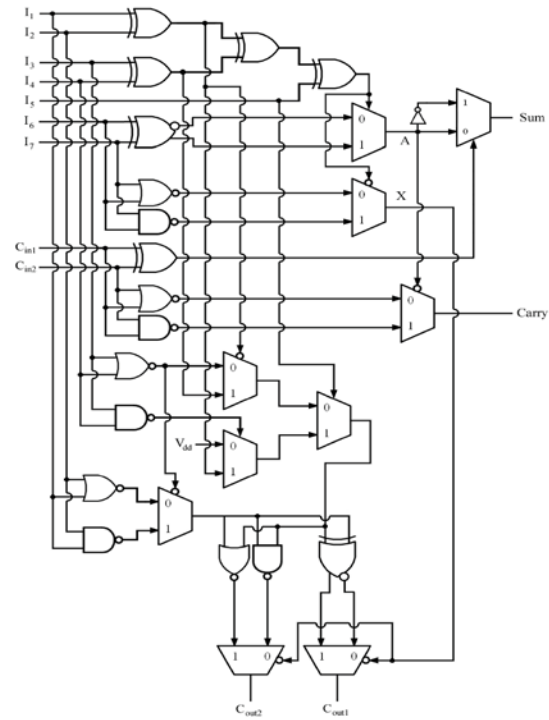


Fig. 5.7-2 Compressor using XOR gate

### V. 6-3 COMPRESSOR

The proposed method uses multi-level exact and approximate compressors for the design of efficient approximate multipliers. The approximate compressors are a key element in the design of power-efficient approximate multipliers. The gate-level delay has been reduced considerably and the carry rippling problem in the compressors is reduced by proposed 5-2 compressor. Further, higher order 6:3 compressor based on symmetric stacking can also be added to the design to further reduce the critical path. It uses 3-bit stacking circuits, which group all of the “1” bits together. The bit stacks are then converted to binary counts, producing 6:3 counter circuits with no XOR gates on the critical path. The overall partial product is divided into MSP and LSP. All the compressors are kept exactly in MSP part and approximated in the LSP part. The approximation is done by eliminating the XOR gates to achieve power and area minimization. FIR filter is designed using the proposed multiplier to analyze its performance power and area.

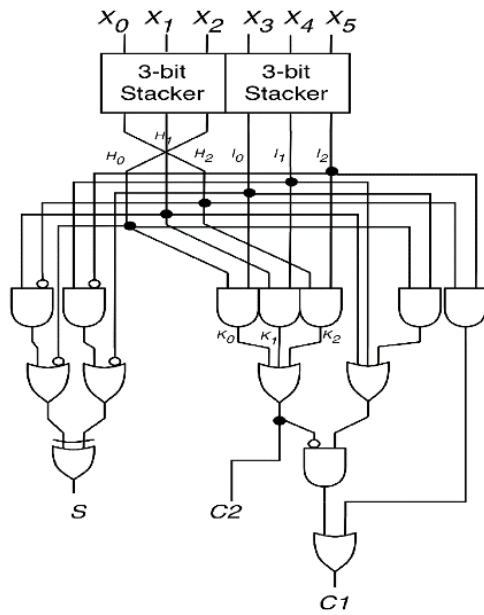


Fig.6.Proposed 6-3 Compressor

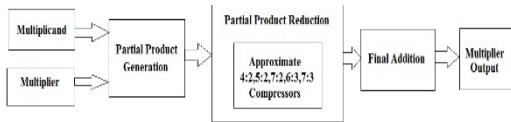


Fig .7.Block diagram

VI. 4-2 COMPRESSOR

$$sum = x_1 + x_3 + x_4,$$

$$carry = x_1 \cdot x_2 + x_3 \cdot x_4.$$

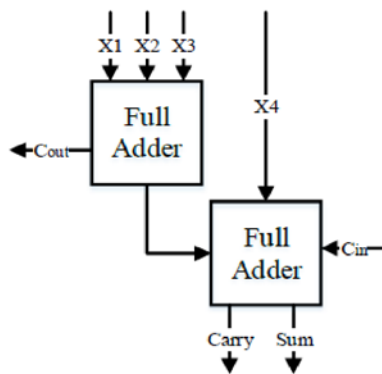


Fig.8. 4-2 Compressor

VII. SIMULATION RESULT

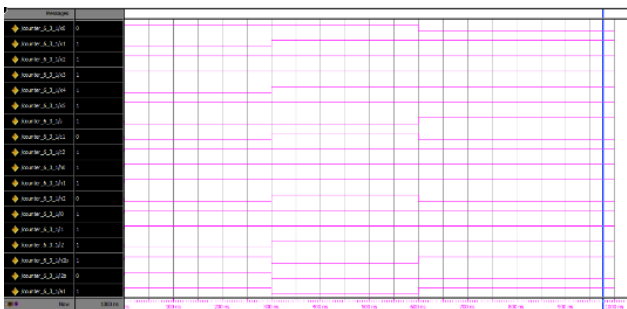


Fig. 9.1 Proposed simulation

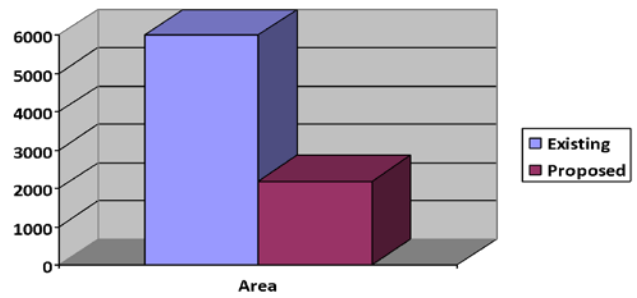


Fig 9.2Performance in terms of area

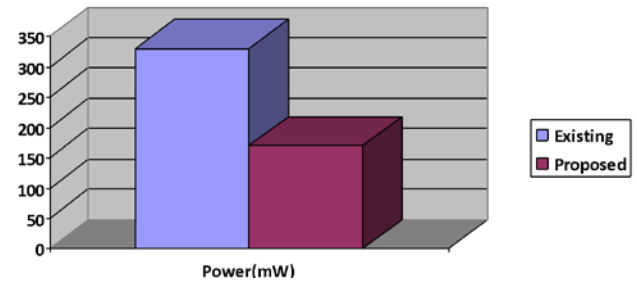


Fig. 9.3Performance in terms of power

VIII. APPLICATIONS

Due to its advantages like Carry delay reduction, 30-40% reduction in Area, power, high performance & delay and low Latency by eliminating the XOR gate in LSP part, this approach can be applied into various DSP applications, Image compressing, smoothing, shaping, Signal processing, Multimedia processing, Machine learning, Data mining and recognition.

IX. CONCLUSION

A new design methodology for speed performance enhancement of  $n - 2$  compressor structures has been developed in this article. By starting from 5-2 compressor with 32% speed improvement, this idea was expanded for 7-2 compressor, which shows a 30% speed enhancement over previously reported works. The proposed system adds 6:3 compressor. Approximate computing technique can reduce major drawbacks in the existing system. This can be employed in DSP application. In addition, by considering the reduced carry ripple problem between the cells of a row and total transistor and gate count, it is clear that the PDP proposed architectures are lower than the previous designs. Also, the provided concept is applicable to other  $n - 2$  compressors. Therefore, high-performance structures can be implemented, which are very suitable for high-speed multipliers. As a consequence, a typical  $16 \times 16$  bit multiplier has been implemented where the simulation results for propagation latency indicate the superiority of proposed compressor blocks.

X. REFERENCES

[1] I. S. Abu-Khater, A. Bellaouar, and M. I. Elmasry, "Circuit techniques for CMOS low-power high-performance multipliers," IEEE J. Solid State Circuits, vol. 31, no. 10, pp. 1535-1546, Oct. 1996.

- [2] O. Kwon, K. Nowka, and E. E. Swartzlander, "A 16-bit  $\times$  16-bit MAC design using fast 5:2 compressors," in Proc. IEEE Int. Conf. Appl.- Specific Syst., Archit., Processors, Jul. 2000, pp. 235–243.
- [3] W.-C. Yeh and C.-W. Jen, "High-speed booth encoded parallel multiplier design," IEEE Trans. Comput., vol. 49, no. 7, pp. 692–701, Jul. 2000.
- [4] A. M. Shams, T. K. Darwish, and M. A. Bayoumi, "Performance analysis of low-power 1-bit CMOS full adder cells," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 10, no. 1, pp. 20–29, Feb. 2002.
- [5] P. J. Song and G. De Micheli, "Circuit and architecture trade-offs for high-speed multiplication," IEEE J. Solid-State Circuits, vol. 26, no. 9, pp. 1184–1198, Sep. 1991.
- [6] A. Fathi, S. Azizian, K. Hadidi, and A. Khoei, "A novel and very fast 4-2 compressor for high speed arithmetic operations," IEICE Trans. Electron., vols. E95, no. 4, pp. 710–712, 2012.
- [7] V. G. Oklobdzija, D. Villeger, and S. S. Liu, "A method for speed optimized partial product reduction and generation of fast parallel multipliers using an algorithmic approach," IEEE Trans. Comput., vol. 45, no. 3, pp. 294–306, Mar. 1996.
- [8] C.-H. Chang, J. Gu, and M. Zhang, "Ultra low-voltage low-power CMOS 4-2 and 5-2 compressors for fast arithmetic circuits," IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 51, no. 10, pp. 1985–1997, Oct. 2004.
- [9] A. Fathi, S. Azizian, K. Hadidi, A. Khoei, and A. Chegeni, "CMOS implementation of a fast 4-2 compressor for parallel accumulations," in Proc. IEEE Int. Symp. Circuits Syst., May 2012, pp. 1476–1479.
- [10] K. Prasad and K. K. Parhi, "Low-power 4-2 and 5-2 compressors," in Proc. Conf. Rec. 25th Asilomar Conf. Signals, Syst. Comput., Nov. 2001, pp. 129–133.
- [11] P. Saha, P. Samanta, and D. Kumar, "4:2 and 5:2 decimal compressors," in Proc. 7th Int. Conf. Intell. Syst., Modelling Simulation (ISMS), Jan. 2016, pp. 424–429.
- [12] D. Balobas and N. Konofaos, "Low-power high-performance CMOS 5-2 compressor with 58 transistors," Electron. Lett., vol. 54, no. 5, pp. 278–280, Mar. 2018.
- [13] A. Najafi, S. Timarchi, and A. Najafi, "High-speed energy-efficient 5:2 compressor," in Proc. 37th Int. Conv. Inf. Commun. Technol., Electron. Microelectron., Opatija, Croatia, May 2014, pp. 80–84.
- [14] M. Tohidi, M. Mousazadeh, S. Akbari, K. Hadidi, and A. Khoei, "CMOS implementation of a new high speed, glitch-free 5-2 compressor for fast arithmetic operations," in Proc. 20th Int. Conf. Mixed Design Integr. Circuits Syst. (MIXDES), Jun. 2013, pp. 204–208.
- [15] G. Caruso and D. Di Sclafani, "Analysis of compressor architectures in MOS current-mode logic," in Proc. 17th IEEE Int. Conf. Electron., Circuits Syst., Dec. 2010, pp. 13–16.
- [16] W. Ma and S. Li, "A new high compression compressor for large multiplier," in Proc. 9th Int. Conf. Solid-State Integrated-Circuit Technol., Oct. 2008, pp. 1877–1880.
- [17] M. Rouholamini, O. Kavehie, A.-P. Mirbaha, S. J. Jasbi, and K. Navi, "A new design for 7:2 compressors," in Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl. (AICCSA), May 2007, pp. 474–478.
- [18] S. Mehrabi, K. Navi, and O. Hashemipour, "Performance comparison of high-speed high-order (n:2) and (n:3) CNFET-based compressors," Int. J. Model. Optim., vol. 3, no. 5, pp. 432–435, Oct. 2013.
- [19] C. Pan, Z. Wang, and C. Sechen, "High speed and power efficient compression of partial products and vectors," J. Algorithms Optim., Oct., vol. 1, no. 1, pp. 39–54, 2013.
- [20] G. Yang, S.-O. Jung, K.-H. Baek, S. Hwan Kim, S. Kim, and S.-M. Kang, "A 32-bit carry lookahead adder using dual-path all-N logic," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 13, no. 8, pp. 992–996, Aug. 2005.
- [21] P. Saha, P. Samanta, and D. Kumar, "4:2 and 5:2 decimal compressors," in Proc. 7th Int. Conf. Intell. Syst., Modelling Simulation (ISMS), Jan. 2016, pp. 424–429.