Ultra Low-Voltage Low-Power CMOS 4-2 and 5-2 Compressors for Fast Arithmetic Circuits

Chip-Hong Chang, Senior Member, IEEE, Jiangmin Gu, Student Member, IEEE, and Mingyan Zhang, Student Member, IEEE

Abstract—This paper presents several architectures and designs of low-power 4-2 and 5-2 compressors capable of operating at ultra low supply voltages. These compressor architectures are anatomized into their constituent modules and different static logic styles based on the same deep submicrometer CMOS process model are used to realize them. Different configurations of each architecture, which include a number of novel 4-2 and 5-2 compressor designs, are prototyped and simulated to evaluate their performance in speed, power dissipation and power-delay product. The newly developed circuits are based on various configurations of the novel 5-2 compressor architecture with the new carry generator circuit, or existing architectures configured with the proposed circuit for the exclusive OR (XOR) and exclusive NOR (XNOR) [XOR–XNOR] module. The proposed new circuit for the XOR–XNOR module eliminates the weak logic on the internal nodes of pass transistors with a pair of feedback PMOS–NMOS transistors. Driving capability has been considered in the design as well as in the simulation setup so that these 4-2 and 5-2 compressor cells can operate reliably in any tree structured parallel multiplier at very low supply voltages. Two new simulation environments are created to ensure that the performances reflect the realistic circuit operation in the system to which these cells are integrated. Simulation results show that the 4-2 compressor with the proposed XOR–XNOR module and the new fast 5-2 compressor architecture are able to function at supply voltage as low as 0.6 V, and outperform many other architectures including the classical CMOS logic compressors and variants of compressors constructed with various combinations of recently reported superior low-power logic cells.

Index Terms—4-2 compressors, 5-2 compressors, arithmetic circuits, digital multipliers.

I. INTRODUCTION

The semiconductor industry has witnessed an explosive growth of integration of sophisticated multimedia-based applications into mobile electronics gadgetry since the last decade. As the CMOS process technology shrinks, the unity gain cutoff frequency, $f_T$, of the transistors become comparable with that of the GaAs bipolar technology that it is now practical to design sub-1-V radio frequency integrated circuits (RFICs) based solely on the matured low cost, low-power CMOS process [17]. Front-end wireless communication circuitries, traditionally based on analog circuit techniques are also being transferred into the digital domain to offer alluring power savings and high density integration by direct conversion architecture [4]. It is high time we explore the well-engineered deep sub-micron CMOS technologies to address the challenging criteria of these emerging low-power and high-speed communication digital signal processing chips. Fast arithmetic computation cells including adders and multipliers are the most frequently and widely used circuits in very-large-scale integration (VLSI) systems. Microprocessors and digital signal processors rely on the efficient implementation of generic arithmetic logic units and floating point units to execute dedicated algorithms such as convolution and filtering [6],[10],[15],[16]. In most of these applications, multipliers have been the critical and obligatory component dictating the overall circuit performance when constrained by power consumption and computation speed. At the circuit design level, considerable potential for optimizing the power-delay product of the multiplier exists by voltage scaling and through the use of contemporary and new CMOS logic styles for the implementation of its embraced combinational circuits [13]–[15], [19].

A fast array or tree multiplier is typically composed of three subcircuits: a Booth encoder for the generation of a reduced number of partial products; a carry save structured accumulator for a further reduction of the partial products’ matrix to only the addition of two operands; and a fast carry propagation adder (CPA) [9] for the computation of the final binary result from its stored carry representation. Among these subcircuits, the second stage of partial product accumulation, often referred to as the carry save adder (CSA) tree [5], [6], [8], [10], [12], [18], occupies a high fraction of silicon area, contributes most to the overall delay, and consumes significant power. Therefore, speeding up the CSA circuit and lowering its power dissipation are crucial to sustain the performance of the multiplier to stay competitive. Early designs of CSA tree used the Dadda’s column compression technique [18] with the 3-2 counters, or equivalently the full adders to reduce the partial product matrix. To lower the latency of the partial product accumulation stage, 4-2 and 5-2 compressors have been widely employed nowadays for high speed multipliers. Owing to its regular interconnection, the 4-2 compressor is ideal for the construction of regularly structured Wallace tree with low complexity [12], [18]. Several 4-2 compressor circuits have been proposed for low-power applications [3], [5], [7], [8], [10], [12]. Some of these are able to operate at low supply voltages but require excessive number of transistors due to their complementary CMOS structures, others use smaller number of transistors but fail to function at ultra low voltages, or lack the driving capability to drive the next level of
subcircuits. The insertion of additional buffers for every output port to provide the output drive increases the switching activities and hence the power dissipation. Higher input compressors have also been studied by researchers [2], [6], [8], [10] and fast 5-2 compressors have been increasingly employed in large word-size multipliers and high precision multiply-accumulators [5], [6], [12]. Most of the research on high-input compressors focuses on the optimization of circuit structure for high speed applications at standard supply voltages. With trends in VLSI toward deep-submicrometer technology, circuits operating reliably at sub-1 V will soon become a reality. This is because the materials used to form the transistors cannot withstand an electric field of unlimited strength, and as transistors get smaller, the field strength increases if the supply voltage is held constant. Since supply voltage has a quadratic contribution to the power dissipation, lowering the supply voltage is also a lucid means of reducing power consumption. However, the major problem with reducing the supply voltage is that the speed of the circuits is also degraded. Therefore, there is a strong impetus to renew the full custom arithmetic cells to achieve high power efficiency for VLSI circuits operating at ultra low supply voltages.

In this paper, we explore new design methodologies for low-power 4-2 and 5-2 compressor circuits that possess sufficient drivability at ultra low voltages based on the advanced CMOS process technology. By investigating the performances of several fast 4-2 and 5-2 compressor architectures and their underlying building modules, a new composite exclusive OR (XOR) and exclusive NOR (XNOR) [XOR-XNOR] cell is proposed. The 4-2 compressors constructed around the proposed XOR-XNOR cell exhibit superior power efficiency comparing to other configurations of the same architecture. A new fast 5-2 compressor architecture is proposed, together with a new circuit for its carry generator module. This new architecture performs well with almost any configuration of logic styles and its overall performance is the best among the known 5-2 compressor architectures under a realistic simulation environment that truly reflects its actual operability in a tree-structured multiplier.

II. THE 4-2 COMPRESSOR

A. 4-2 Compressor Architectures

A 4-2 compressor has five inputs and three outputs, as shown in Fig. 1. The four inputs $x_3$, $x_2$, $x_3$, and $x_4$, and the output $c_{in}$ have the same weight. The output $c_{out}$ is weighted one binary bit order higher. The 4-2 compressor receives an input $c_{in}$ from the preceding module of one binary bit order lower in significance, and produces an output $c_{out}$ to the next compressor module of higher significance. Different structures of 4-2 compressors exist and they all have to abide by the fundamental equation given as follows:

$$x_1 + x_2 + x_3 + x_4 + c_{in} = \text{sum} + 2 \cdot (c_{out} + c_{carry}).$$  (1)

Besides, to accelerate the carry save summation of the partial products, it is imperative that the output, $c_{carry}$ be independent of the input $c_{in}$.

The conventional implementation of a 4-2 compressor is composed of two serially connected full adders, as shown in Fig. 2. At gate level, high input compressors are anatomized into XOR gates and carry generators normally implemented by multiplexers (MUX). Therefore, different designs can be classified based on the critical path delay in terms of the number of primitive gates. Let $\Delta_{XOR}$ denote the delay of an XOR gate and $\Delta_{CGEN}$ denote the delay of a carry generator. A compressor is said to have a delay of $(m\Delta_{XOR} + n\Delta_{CGEN})$ if its critical path consists of $m$ XOR gates and $n$ carry generators. Since the difference between the delays of widely used XOR gate and carry generator is trivial in an optimized design, the delay of the compressor is more commonly specified as $(m + n)\Delta$. Therefore, the straightforward implementation of a 4-2 compressor of Fig. 2 has a long critical path delay of $4\Delta$ [5]. Additionally, due to the uneven delay profiles of the outputs arriving from different input paths, the CSA tree for the partial product accumulation constructed from such cells tends to generate a lot of glitches.

A 4-2 compressor flattened and optimized at gate level to reduce the critical path delay is shown in Fig. 3 [12]. It uses more than 80 transistors when implemented in conventional or complementary CMOS logic style. As this circuit is capable of functioning below 1 V, it is used as a benchmark for evaluating the performance of other low voltage and low-power 4-2 compressor circuits.
A recent design of 4-2 compressor [3], [7], [10], [12] is derived from the modified equations for the functions of Fig. 2. The three outputs of the design are described as follows:

\[
\begin{align*}
C_{\text{out}} &= (x_1 \oplus x_2) \cdot x_3 + x_1 \cdot x_2 = (x_1 \oplus x_2) \cdot x_3 \\
S &= x_1 \oplus x_2 \oplus x_3 \\
\text{sum} &= s \oplus x_4 + c_{\text{in}} = x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus c_{\text{in}} \\
\text{carry} &= (s \oplus x_4) \cdot c_{\text{in}} + s \cdot x_4 \\
&= (x_1 \oplus x_2 \oplus x_3 \oplus x_4) \cdot c_{\text{in}} \\
&\quad + (x_1 \oplus x_2 \oplus x_3 \oplus x_4) \cdot x_4. 
\end{align*}
\]

The two carry signals carry and \(c_{\text{out}}\) are generated from both the XOR and XNOR functions of the input signals. The sum output is generated by several two-input XOR circuits, some internal signals of which can be used to generate the two carries. Fig. 4 shows the logic decomposition of this 4-2 compressor architecture. It is mainly composed of six modules, four of which are XOR circuits and the other two are 2-1 MUX. Three special XOR-XNOR modules marked with “XOR*” generate both the XOR and XNOR signals simultaneously to other modules driven by them. This design has a critical path delay of \(3\Delta\), which is \(1\Delta\) delay shorter than the conventional implementation. Besides, its outputs feature balanced signal arrival time from each data inputs \((x_1 \text{ to } x_4)\), thanks to the special “XOR*” modules.

B. Circuit Implementations of Building Block Modules

There are several designs [1], [11]–[15] of the XOR* modules, as shown in Fig. 5, for implementing the 4-2 compressor of Fig. 4. As both the XOR and XNOR functions are required in the carry generation circuits, circuits capable of co-generating these two signals, as depicted in Fig. 5, are beneficial to the implementation of the special XOR* modules of Fig. 4. Although there is no need for the XOR* module with inputs of \(x_3\) and \(x_4\) to provide the XNOR output, the same structure as other XOR* modules is still preferred in order to avoid skewed delay paths. In Fig. 5, the numerical values next to the transistors are the channel widths in \(\mu\text{m}\). Minimum feature size is assumed for the channel lengths of the transistors based on the latest Chartered Semiconductor CSM 0.18-\(\mu\text{m}\) CMOS process technology.

The design of Fig. 5(a) has the least number of transistors and consumes very low power [13]–[15]. However, it generates a weak logic “1” at the XNOR node when the primary inputs are both “1”s, which prevents it from functioning reliably at low supply voltage. The design of Fig. 5(b) is able to operate at low voltage, but it is not power efficient [13]–[15]. Both designs [Fig. 5(a) and (b)] use inverter to generate the complementary XOR and XNOR signals, therefore, their outputs skew heavily in time.

The design of Fig. 5(c) consists of two cross-backed XOR and XNOR cells [13]–[15]. It is able to generate the complementary XOR and XNOR outputs simultaneously. However, it performs non full-swing operations for some input patterns causing their corresponding outputs to be degraded by 1 \(V_{th}\). For example, the XNOR output transmits a weak logic “1” when both inputs are “1”s, whereas the XOR output transmits a weak logic “0” when both inputs are “0”s. When the power supply voltage is lower than 1 V, the weak logic transmission will slow down the charging and discharging speed of the driven circuits, or worse, unable to turn on or off the driven transistors as desired. Therefore, it is also not a suitable candidate for ultra low-voltage operation.

The combined XOR–XNOR cell of Fig. 5(d) was proposed in [11], [12]. It is a low-power circuit with the least number of transistors that can output XOR and XNOR concurrently. It eliminates the transmission of weak logic for certain input patterns by virtue of the feedback PMOS–NMOS transistors in the midst of the circuit. Nevertheless, it is still not suitable for low-voltage applications for the following reason. When the inputs change from any other pattern to “00” or “11,” the feedback transistors that were turned off originally will be turned on by both a weak logic driver and a high impedance driver. This transition takes
a long time at very low voltage slightly above $2 \cdot V_{thP}$. Meanwhile, the short-circuit current of the following stage increases tremendously, which lead to a rise in the power dissipation.

We proposed a new circuit for the XOR module, as shown in Fig. 5(e), which is able to generate the XOR and XNOR outputs simultaneously, too. A pair of feedback PMOS-NMOS transistors is added to the XOR–XNOR circuit of Fig. 5(c). It eliminates the weak logic problem encountered by the former circuit when the input pattern is \[00\] or \[11\]. It is able to operate at even lower voltage than Fig. 5(d) because the pairs of serially connected PMOS transistors and serially connected NMOS transistors ensure full-swing logic during the input transitions to \[00\] and \[11\]. Hence, the proposed circuit is rather robust against delay driven voltage scaling. Fig. 6 shows the layout of the proposed new circuit module.

The last XOR module to generate the \[s_{\text{sum}}\] output of Fig. 4 does not need to provide the XNOR output, but it should provide sufficient output current to drive the next stage of 4-2 compressors. Fig. 7 shows two possible XOR gate designs for this simple XOR module [13]–[15]. Fig. 7(a) is a low-power design of XOR function. But the limited drivability prevents it from being used as an output module in a 4-2 compressor. Fig. 7(b) is an XNOR circuit followed by an inverter, which has stronger driving capability than Fig. 7(a). Therefore, Fig. 7(b) is used in our proposed 4-2 compressor to generate the \[s_{\text{sum}}\] signal. As mentioned earlier, the numbers beside the transistors are the optimized channel widths. As some transistors have different optimized sizes for different compressors, the numbers in brackets are the sizes of the transistors optimized for 5-2 compressors, which will be discussed next.

The carry generator modules produce the signals \[c_{\text{carry}}\] and \[c_{\text{carry}}\] which are usually generated by MUX. Several designs are shown in Fig. 8. Fig. 8(a) is widely used in low-power full adder cells [13]–[15]. However, its driving capability is somewhat limited, which causes signal decay when many stages are to be cascaded. So it is not suitable for use in the 4-2 compressors of the CSA tree. Fig. 8(b) improves the driving capability by adding an output buffer at the expense of increasing its power dissipation [13]–[15]. The output buffer formed by the cascaded inverters is designed such that the first inverter is half the size of the output inverter in order to cut down the power dissipation.

The circuit of Fig. 8(c) is a MUX implemented in standard complementary CMOS logic style [3], [19]. Being a complementary CMOS circuit, it is robust against both voltage scaling and transistor sizing. Despite having one inverter lesser than the design of Fig. 8(b), this circuit still delivers sufficient drive to its succeeding circuits through the output inverter. The total number of transistors of this circuit is ten. Although it is two more than that of Fig. 8(b), the silicon areas occupied by both circuits are almost the same. This is because Fig. 8(b) requires more space to segregate the different diffusion areas, which increases the routing complexity of the interconnecting lines. The regularity of the layout of the circuit of Fig. 8(c) is evident from the diagram on its right. The MUX circuits of Fig. 8(d) and Fig. 8(e) are implemented in complementary pass transistor logic (CPL) and dual pass transistor logic (DPL) styles, respectively [10], [19]. Because they are dual-rail circuits, complementary pairs of primary inputs and outputs need to be generated. Although they can generate full-swing outputs, due to the pass transistor structure, they will not provide adequate drivability if many such circuits are cascaded, particularly at
The 5-2 compressor is another widely used building block for high precision and high speed multipliers. The block diagram of a 5-2 compressor is shown in Fig. 11, which has seven inputs and four outputs. Five of the inputs are the primary inputs \(x_1, x_2, x_3, x_4\) and \(x_5\), and the other two inputs, \(c_{\text{in1}}\) and \(c_{\text{in2}}\) receive their values from the neighboring compressor of one binary bit order lower in significance. All the seven inputs have the same weight. The 5-2 compressor generates an output \(s_{\text{sum}}\) of the same weight as the inputs, and three outputs \(c_{\text{out1}}, c_{\text{out2}},\) and \(c_{\text{in2}}\) weighted one binary bit order higher. The outputs, \(c_{\text{out1}}\) and \(c_{\text{out2}}\) are fed to the neighboring compressor of higher significance. All the 5-2 compressors of different designs abide by (6)

\[
x_1 + x_2 + x_3 + x_4 + x_5 + c_{\text{in1}} + c_{\text{in2}} = s_{\text{sum}} + 2 \cdot (\text{carry} + c_{\text{out1}} + c_{\text{out2}}). \tag{6}
\]

Besides, to speed up the carry save summation of the partial products, the output \(c_{\text{out1}}\) must be independent of the input \(c_{\text{in1}}\) and \(c_{\text{in2}}\), and the output \(c_{\text{out2}}\) must be independent of the input \(c_{\text{in2}}\).

A simple implementation of the 5-2 compressor is to cascade three full adders in a hierarchical structure, as shown in Fig. 12, which has a critical path delay of \(6\Delta\).

Fig. 13 shows a modified architecture of the 5-2 compressor [10], which has a critical path delay of \(5\Delta\). The XOR* modules in the figure generate both the XOR and XNOR signals simultaneously, as described in Section II. The style and structure of the circuit share some common attributes as the recently published structural design of full adders [10], [11], [13]–[15] and 4-2 compressors [3], [10], [12].

In spite of the structural differences between the implementations of Fig. 12 and Fig. 13, the formulas to generate the output signals are essentially derived from the same basic architecture of Fig. 12. Each full adder can be logically expressed as

\[
s_{\text{FA}} = a \oplus b \oplus c \tag{7}
\]

\[
c_{\text{FA}} = (a \oplus b) \cdot c + (a \oplus b) \cdot a \tag{8}
\]

where \(a, b,\) and \(c\) are the primary inputs, and \(s_{\text{FA}}\) and \(c_{\text{FA}}\) are the primary outputs of the full adder.

It follows that the outputs and the internal nodes of Fig. 12 can be expressed by the following set of equations:

\[
s_1 = x_1 \oplus x_2 \oplus x_3 \tag{9}
\]

\[
c_{\text{out1}} = (x_1 \oplus x_2) \cdot x_3 + (x_1 \oplus x_2) \cdot x_1 \tag{10}
\]

\[
s_2 = s_1 \oplus x_4 \oplus c_{\text{in1}} \tag{11}
\]

\[
= x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus c_{\text{in1}}
\]
Although this architecture produces different delays, are still controlled by the corresponding stage module. The carry generator is faster than that controlled by a complementary CMOS logic style, which is derived later than the primary inputs, implementing CGEN2 with more complex gate as CGEN1 is not necessary as it will not help to reduce the critical path delay of the tree-structured multiplier any way.

The carry generators CGEN2 in Fig. 15(a) can also be implemented by MUX, as shown in Fig. 15(b). CGEN2 of Fig. 15(a) is better implemented in complementary CMOS logic style, which offers lower capacitance and the driving XOR* modules while the MUX-based carry generator modules of Fig. 15(b) is more suitable for realization with other logic styles.

The output functions of our proposed architecture in Fig. 15 are described by (19)–(22)

\begin{align*}
\text{sum} &= x_1 + x_2 + x_3 + x_4 + x_5 + G_{n1} + G_{n2} \\
\text{c}_{\text{out}1} &= (x_1 + x_2 + x_3 + x_4) \\
\text{c}_{\text{out}2} &= [(x_1 + x_2 + x_3 + x_4) \cdot x_5] + (x_1 + x_2 + x_3 + x_4 + G_{n1}) \cdot G_{n2} \\
\text{carry} &= [(x_1 + x_2 + x_3 + x_4 + G_{n1}) \cdot x_5] + (x_1 + x_2 + x_3 + x_4 + x_5 + G_{n1}) \cdot G_{n2}.
\end{align*}

We proposed a new and faster architecture with the same theoretical critical path delay of 4Δ. The structure of our 5-2 compressor is shown in Fig. 15(a). It uses one less module than the 5-2 compressor of Fig. 14. The mechanisms for generating the carry output signals, such as carry, \(c_{\text{out}1}, c_{\text{out}2}\), are fundamentally different from those of the existing architectures. The carry generator CGEN1 is used to produce the signal \(c_{\text{out}1}\). In anticipation that this carry generator is more likely to be situated in the longest delay path in the tree structured multiplier, it is implemented as a complex gate fed only with the primary inputs. Although several more transistors are used, the speed of this carry generator is faster than that controlled by a XOR* module. The other two carry generators CGEN2 which are used to produce the carry and \(c_{\text{out}2}\), are still controlled by the corresponding XOR* modules. Unlike CGEN1, one of the inputs to CGEN2 is either \(G_{n1}\) or \(G_{n2}\), which comes from an output of the compressor of the preceding stage. Since these signals are often arrived later than the primary inputs, implementing CGEN2 with more complex gate as CGEN1 is not necessary as it will not help to reduce the critical path delay of the tree-structured multiplier any way.

A faster 5-2 compressor proposed in [6] is shown in Fig. 14. This architecture uses a different method to generate the outputs, \(c_{\text{out}1}\) and \(c_{\text{out}2}\). Although this architecture produces different output bit patterns in \(c_{\text{out}1}\) and \(c_{\text{out}2}\) for the same input data, it still abides by (6). It is claimed to have a delay of 4Δ.

Equations (15)–(18) show the output functions:

\begin{align*}
\text{sum} &= x_1 + x_2 + x_3 + x_4 + x_5 + G_{n1} + G_{n2} \\
\text{c}_{\text{out}1} &= (x_1 + x_2) \cdot x_3 + x_4 \\
\text{c}_{\text{out}2} &= [(x_1 + x_2 + x_3 + x_4) \cdot (x_1 \cdot x_2 + x_3 \cdot x_4)] + (x_1 + x_2 + x_3 + x_4 + G_{n1}) \cdot G_{n2} \\
\text{carry} &= [(x_1 + x_2 + x_3 + x_4 + G_{n1}) \cdot x_5] + (x_1 + x_2 + x_3 + x_4 + x_5 + G_{n1}) \cdot G_{n2}.
\end{align*}
power and high-speed 5-2 compressor cells for instantiation at architectural level. For example, a dual-rail 5-2 compressor [10] is proposed for the architecture of Fig. 4, where the XOR modules are implemented as dual-rail MUX to improve the performance.

B. Circuit Implementations of the Proposed 5-2 Compressor Architectures

The design of Fig. 5(e) is recommended for the implementation of XOR* modules to allow low-voltage operation. The XOR* modules do not need strong driving capability for the internal modules driven by them. Therefore, the design of Fig. 7(a) without the output buffer can be used for the XOR* module. The sum output stage is implemented with the circuit of Fig. 7(b) to assure the drivability.

A new complementary CMOS logic style circuit, as shown in Fig. 16, is suggested for the implementation of the CGEN2 module. It resembles the circuit of Fig. 8(c), except that an additional input \( c_{in} \) is introduced and the original input \( a \) is rearranged. Comparing with the design of Fig. 8(c), this design lowers the switched capacitance of the preceding XOR* module due to the reduction of its fanouts from 2 to 1 for both the XOR and XNOR outputs. It is also easier to layout without the cross lines of \( \text{sel} \) and \( \text{sel} \).

The \( C_{out1} \) output of the CGEN1 module receives the primary inputs of \( x_2 \) to \( x_3 \) directly. It can be implemented with the carry generator circuit of the full adder in complementary CMOS logic style, which is shown in Fig. 17(a) [9], [19]. This circuit uses two more transistors than the circuit of Fig. 16, but it bypasses the XOR and XNOR signals from the preceding XOR* module, which has prevented the delay of the \( C_{out1} \) generation from being degraded by about \( 1\Delta \). Fig. 17(b) shows the circuit for the CGEN1 module implemented in CPL logic style [19].

IV. SIMULATION RESULTS

A. Simulation Environment

The simulations are performed by Nassda HSIM 2.0 tool with the option “HSIMSPEED” set to “0.” This option gives the slowest simulation time with the highest accuracy giving results compatible to HSPICE simulation. All the circuits and layouts are targeted for CSM 0.18-\( \mu \)m CMOS technology. Therefore, the circuits are designed and optimized based on this process model.

The simulation environments for the 4-2 compressor and 5-2 compressor circuits are shown in Fig. 19. Each input is driven by buffered signals and each output is loaded with buffers, which provide a realistic simulation environment reflecting the compressor operation in actual applications. The simulation environments for the 4-2 compressor and 5-2 compressor consist of two cascaded 4-2 compressors and three cascaded 5-2 compressors, respectively. These compressors are running in parallel to simulate an actual compressor stage in the CSA tree. More than one compressors are used in the simulation because the critical paths of some data patterns may cross adjacent compressors in the same stage of the CSA tree. The dashed lines in Fig. 19...
indicate the scenario of such potential critical paths. The leftmost compressor of both simulation environments is inspected because it is most probable to have the longest delay. The delay is measured from the earliest input signal reaching 50% of the supply voltage to the latest output signal reaching 50% of the supply voltage for each input cycle. The worst case delay is the largest delay among all input data.

For each circuit, 1024 data are randomly generated by MATLAB to feed into the circuits as input stimuli. The circuits under test are simulated at various supply voltages, ranging from 0.6 to 3.3 V. Each supply voltage corresponds with one of two simulation frequencies (the rate at which data patterns are fed): supply voltages greater than 1.0 V operate at 100 MHz, while supply voltages less than or equal to 1.0 V operate at 10 MHz. It should be noted that the simulation frequency is not the maximum operating frequency of the compressors. In fact, the compressors simulated are capable of operating correctly at much higher frequency than the simulation frequency. The average power consumption of the leftmost compressor is measured for every supply voltage, with the power consumed by the additional buffers excluded from the average power consumption calculation.

B. Simulation Results of 4-2 Compressors

Nine different 4-2 compressor designs are simulated. The first circuit is a full complementary CMOS logic style implementation of Fig. 3, which is used as the benchmark for low-voltage operation. The second to the sixth compressors use the same architecture of Fig. 4, with the XOR modules implemented by the circuit of Fig. 7(b) and the MUX modules implemented by the circuit of Fig. 8(b). They differ mainly in the circuit implementation of the modules. The modules of these five compressors are respectively implemented by the circuits of Fig. 5(a) to Fig. 5(e). The seventh compressor is a hybrid design employing the circuit of Fig. 5(e) for its XOR modules, the circuit of Fig. 7(b) for its XOR modules, and the circuit of Fig. 8(c) as its MUX modules. The eighth and ninth compressors are implemented by CPL and DPL logic styles, using the circuits of Figs. 8(d) and 9, and Fig. 8(e) and 9, respectively as their building blocks. The configurations of the nine compressors are listed in Table I. For brevity, only the figure numbers of the circuits, e.g. 5(a) instead of Fig. 5(a), are shown in the table. The last column in Table I shows the lowest operable voltage for each 4-2 compressor obtained from the simulation, below which the circuit fails to function correctly. The full simulation results for the performance of all the compressors at different supply voltages are tabulated in Tables II–IV. The power efficiency, or the power-delay product (PDP) measured in fJ is defined as the product of the worst-case delay and the average power consumption. This metric provides an indication of the

<table>
<thead>
<tr>
<th>S/N</th>
<th>Design</th>
<th>XOR*</th>
<th>XOR</th>
<th>MUX</th>
<th>Min. V_{\text{op}} (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4-2 CMOS</td>
<td>3</td>
<td>b</td>
<td>b</td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td>4-2 a</td>
<td>5(a)</td>
<td>7(b)</td>
<td>b</td>
<td>0.9</td>
</tr>
<tr>
<td>3</td>
<td>4-2 b</td>
<td>5(b)</td>
<td>7(b)</td>
<td>b</td>
<td>0.6</td>
</tr>
<tr>
<td>4</td>
<td>4-2 e</td>
<td>5(c)</td>
<td>7(b)</td>
<td>b</td>
<td>2.1</td>
</tr>
<tr>
<td>5</td>
<td>4-2 d</td>
<td>5(d)</td>
<td>7(b)</td>
<td>b</td>
<td>0.6</td>
</tr>
<tr>
<td>6</td>
<td>4-2 e</td>
<td>5(c)</td>
<td>7(b)</td>
<td>b</td>
<td>0.6</td>
</tr>
<tr>
<td>7</td>
<td>4-2 hybrid</td>
<td>5(d)</td>
<td>7(b)</td>
<td>b</td>
<td>0.6</td>
</tr>
<tr>
<td>8</td>
<td>4-2 CPL</td>
<td>8(d) + 9</td>
<td>8(d) + 9</td>
<td>8(d)</td>
<td>0.6</td>
</tr>
<tr>
<td>9</td>
<td>4-2 DPL</td>
<td>8(e) + 9</td>
<td>8(e) + 9</td>
<td>8(e)</td>
<td>0.6</td>
</tr>
</tbody>
</table>

### Table II
Comparison of Delay (Nanoseconds) of 4-2 Compressors

<table>
<thead>
<tr>
<th>Design</th>
<th>Delay (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-2 CMOS</td>
<td>6.29, 3.18</td>
</tr>
<tr>
<td>4-2 a</td>
<td>12.7, 5.99</td>
</tr>
<tr>
<td>4-2 b</td>
<td>8.03, 4.05</td>
</tr>
<tr>
<td>4-2 c</td>
<td>1.01, 0.75</td>
</tr>
<tr>
<td>4-2 d</td>
<td>34.7, 7.04</td>
</tr>
<tr>
<td>4-2 e</td>
<td>6.83, 3.43</td>
</tr>
<tr>
<td>4-2 hybrid</td>
<td>7.11, 3.72</td>
</tr>
<tr>
<td>4-2 CPL</td>
<td>6.28, 3.04</td>
</tr>
<tr>
<td>4-2 DPL</td>
<td>6.74, 3.32</td>
</tr>
</tbody>
</table>

### Table III
Comparison of Power (Microwatts) of 4-2 Compressors

<table>
<thead>
<tr>
<th>Design</th>
<th>Power (\mu W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-2 CMOS</td>
<td>0.17, 0.23</td>
</tr>
<tr>
<td>4-2 a</td>
<td>0.33, 0.41</td>
</tr>
<tr>
<td>4-2 b</td>
<td>0.11, 0.16</td>
</tr>
<tr>
<td>4-2 c</td>
<td>1.41</td>
</tr>
<tr>
<td>4-2 d</td>
<td>0.26, 0.32</td>
</tr>
<tr>
<td>4-2 e</td>
<td>0.11, 0.15</td>
</tr>
<tr>
<td>4-2 hybrid</td>
<td>0.11, 0.15</td>
</tr>
<tr>
<td>4-2 CPL</td>
<td>0.13, 0.18</td>
</tr>
<tr>
<td>4-2 DPL</td>
<td>0.12, 0.16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Design</th>
<th>PDP (fJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-2 CMOS</td>
<td>1.05, 0.74</td>
</tr>
<tr>
<td>4-2 a</td>
<td>4.19, 2.45</td>
</tr>
<tr>
<td>4-2 b</td>
<td>0.92, 0.64</td>
</tr>
<tr>
<td>4-2 c</td>
<td>14.2</td>
</tr>
<tr>
<td>4-2 d</td>
<td>9.06, 2.22</td>
</tr>
<tr>
<td>4-2 e</td>
<td>0.77, 0.52</td>
</tr>
<tr>
<td>4-2 hybrid</td>
<td>0.81, 0.58</td>
</tr>
<tr>
<td>4-2 CPL</td>
<td>0.79, 0.54</td>
</tr>
<tr>
<td>4-2 DPL</td>
<td>0.80, 0.53</td>
</tr>
</tbody>
</table>

The power efficiency, or the power-delay product (PDP) measured in fJ is defined as the product of the worst-case delay and the average power consumption. This metric provides an indication of the
energy expended and the life span of the battery when the circuit is operating at its maximum speed. Two best performances at each supply voltage are highlighted in italic and bold print for ease of comparison.

To investigate the performance variations of the same 4-2 architecture due to different implementations of the module, the worst case delay, power dissipation and power-delay product of Designs 2–6 are charted in Fig. 20. It is evident that the worst case delays at low supply voltages of Designs 2 and 5 are much longer than those of the other designs. Designs 2 and 4 consume more power. Consequently, Designs 3 and 6 perform significantly better than the other designs in terms of the power efficiency. In fact, these are the only two designs of Fig. 20 that are able to operate down to 0.6 V. The circuits used to implement the XOR* modules for Designs 3 and 6 are respectively, Fig. 5(b) and our proposed Fig. 5(e). The worst case delay, power dissipation and power-delay product of all designs capable of functioning down to the lowest supply voltage of 0.6 V, including our proposed hybrid 4-2 compressor are charted in Fig. 21. Designs 1, 8 and 9 featuring respectively the CMOS, CPL and DPL logic styles have comparatively shorter worst-case delay. However, they and Design 3 dissipate notably higher power than Designs 6 and 7, which use our proposed XOR* cell of Fig. 5(e). For example, Design 8 (CPL) consumes 12.7% more power at 0.6 V than Design 6 (4-2_e). The average power consumption exacerbates with increasing supply voltages to a surplus of 14.9% at 1.0 V, 19.0% at 1.8 V and 30.5% at 3.3 V. The power efficiency (PDP) of Designs 6 and 7 are comparable to those circuits implemented in CPL and DPL styles. Both CPL and DPL logic styles require the generation of dual-rail signals for each primary input and output, incurring almost twice as many interconnecting lines as the other designs. Taking into account the substantial capacitive load due to the wiring overhead, Design 8 and 9 will not be competitive for the implementation of 4-2 compressors in large parallel multipliers.

C. Simulation Results of 5-2 Compressors

Thirteen different 5-2 compressor designs are simulated. These selected designs are all operable from 0.6 V to 3.3 V and their circuit configurations are tabulated in Table V. Each design is named according to its base architecture and a postfix indicating the types of circuits employed for the constituent modules. For example, the name 5del_ebb of Design 2 implies that it is a 5-2 compressor constructed by the circuits of Fig. 5(e), Fig. 7(b) and Fig. 8(b) for its XOR, XOR, and MUX modules, respectively. The first five designs are 5-2 compressors of 5Δ based on the architecture of Fig. 13. The four designs that followed are the 5-2 compressors of 4Δ based on the architecture of Fig. 14 proposed by Kwon et al. [6]. Designs 6–8 use the optimized complementary CMOS logic style to generate the OR–AND and AND–OR functions while Design 9 (kwon_cpl) generates these two functions with CPL logic style. The last four designs are the 5-2 compressors of using the proposed architecture of Fig. 15. Designs 10, 11, and 13 use the circuits of Fig. 15(b) and Fig. 17(a) for the CGEN1 modules, and the circuit of Fig. 8(b) for the MUX modules. Design 12 is based on the architecture of Fig. 15(a), with a hybrid composition of optimally designed circuits of different

<table>
<thead>
<tr>
<th>S/N</th>
<th>Design</th>
<th>Architecture</th>
<th>XOR</th>
<th>XOR, XOR</th>
<th>MUX, CGEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5del_bbb</td>
<td>13</td>
<td>5(b)</td>
<td>7(b)</td>
<td>8(b)</td>
</tr>
<tr>
<td>2</td>
<td>5del_ebb</td>
<td>13</td>
<td>5(e)</td>
<td>7(b)</td>
<td>8(b)</td>
</tr>
<tr>
<td>3</td>
<td>5del_hybrid</td>
<td>13</td>
<td>5(e)</td>
<td>7(b)</td>
<td>8(c)</td>
</tr>
<tr>
<td>4</td>
<td>5del_cpl</td>
<td>13</td>
<td>8(d)</td>
<td>8(d)</td>
<td>8(d)</td>
</tr>
<tr>
<td>5</td>
<td>5del_dpl</td>
<td>13</td>
<td>8(e)</td>
<td>8(e)</td>
<td>8(e)</td>
</tr>
<tr>
<td>6</td>
<td>kwon_bbb</td>
<td>14</td>
<td>5(b)</td>
<td>7(b)</td>
<td>8(b)</td>
</tr>
<tr>
<td>7</td>
<td>kwon_ebb</td>
<td>14</td>
<td>5(e)</td>
<td>7(b)</td>
<td>8(b)</td>
</tr>
<tr>
<td>8</td>
<td>kwon_ebc</td>
<td>14</td>
<td>5(e)</td>
<td>7(b)</td>
<td>8(c)</td>
</tr>
<tr>
<td>9</td>
<td>kwon_cpl</td>
<td>14</td>
<td>8(d)</td>
<td>8(d)</td>
<td>8(d)</td>
</tr>
<tr>
<td>10</td>
<td>4del_bbb</td>
<td>15(b)</td>
<td>5(b)</td>
<td>7(b), 7(a)</td>
<td>17(a), 8(b)</td>
</tr>
<tr>
<td>11</td>
<td>4del_ebb</td>
<td>15(b)</td>
<td>5(e)</td>
<td>7(b), 7(a)</td>
<td>17(a), 8(b)</td>
</tr>
<tr>
<td>12</td>
<td>4del_hybrid</td>
<td>15(a)</td>
<td>5(e)</td>
<td>7(b), 7(a)</td>
<td>17(a), 16</td>
</tr>
<tr>
<td>13</td>
<td>4del_cpl</td>
<td>15(b)</td>
<td>8(d)</td>
<td>8(d)</td>
<td>17(b), 8(d)</td>
</tr>
</tbody>
</table>
logic styles. It uses the proposed XOR* cell of Fig. 5(e), the pass-transistor style XOR gate of Fig. 7(b), the circuit of Fig. 7(a), the complementary CMOS styled CGEN1 circuit of Fig. 17(a), and the CGEN2 circuit of Fig. 16. Design 13 (CPL) is based on the same architecture as Designs 10 and 11, but uses the circuits implemented in CPL logic style for the XOR, AND, and MUX modules. Its CGEN1 module is implemented with the CPL circuit of Fig. 17(b).

The simulation results of the delay, power and power efficiency of all the compressors are tabulated in Tables VI–VIII. Two best performances at each supply voltage are printed in bold and italic. The performances of all 5-2 compressors based on the architecture are charted in Fig. 22 for comparison. The CPL (5del_cpl) and DPL (5del_dpl) designs have the best worst case delay and power efficiency. Among the non dual-rail designs, 5del_bbb consumes more power and 5del_hybrid has the longest delay. Design 5del_ebb provides the best trade-off between delay, power and power efficiency among the three non dual-rail compressors. Considering all aspects of the performance, this architecture is best implemented with either CPL or DPL circuits.

Fig. 23 compares the performances of various 4Δ 5-2 compressors built upon the architecture proposed by Kwon et al. [6]. The design implemented with CPL circuits continues to perform well in delay than the non dual-rail designs as in the previous comparison but its overall performance is not necessarily better than the others this time. The problem lies in its high power dissipation. Designs kwon_ebb and kwon_ebc consume about 2% to 18% lesser power than the other two designs, but they are slower too. In term of power efficiency, Design kwon_bbb
outperforms the other two non-CPL designs, which have very similar performance.

Fig. 24 shows the performances of several circuits built around our proposed $4\Delta$ 5-2 compressor architecture. Both $4\text{del}_{\text{ebb}}$ and $4\text{del}_{\text{hybrid}}$ outperform $4\text{del}_{\text{bbb}}$ and $4\text{del}_{\text{cpl}}$ remarkably. Their speeds are comparable to the circuit implemented with CPL logic style. They consume 32% to 35% lesser power than $4\text{del}_{\text{cpl}}$ and 10% to 20% lesser power than $4\text{del}_{\text{bbb}}$. At voltage higher than 1.0 V, $4\text{del}_{\text{ebb}}$ has a slight edge over $4\text{del}_{\text{hybrid}}$, whereas at voltage lower than 1.0 V, $4\text{del}_{\text{hybrid}}$ performs slightly better.

Performance differences due to architectural difference are also studied by comparing different compressor architectures using identical configuration for the same anatomized modules. The performances of three 5-2 compressor architectures with configurations of bbb, ebb, hybrid, and CPL are respectively charted in Figs. 25–28. Design $5\text{del}_{\text{dpl}}$ is also added to Fig. 28 for comparison. Our proposed architecture has the best performance among all architectures implemented with non dual-rail logic styles. For modules implemented with CPL or DPL logic style, the architecture of Fig. 13 is better.

Fig. 25 compares the architectures with bbb configuration. The results show that the average power of our proposed $4\Delta$ architecture is 19% to 24% lesser than Kwon’s architecture of Fig. 14, and 20% to 25% lesser than the $5\Delta$ architecture of Fig. 13. Although it is slower, its power-delay product is still far lower than the other architectures.

Fig. 26 shows that, with the configuration ebb, both the average power and worst-case delay of our proposed $4\Delta$ architecture consume 25% to 28% lesser power than Kwon’s architecture, and 20% to 29% lesser power than the $5\Delta$ architecture.

Fig. 27 shows that, with the hybrid configuration, our proposed $4\Delta$ architecture consumes 25% to 28% lesser power than Kwon’s architecture, and 20% to 29% lesser power than the $5\Delta$
architecture. Therefore, and the carry generator delay is also proposed. Different architectures, an architecture. It is also 6% to 23% faster than Kwon’s architecture and 12% to 23% faster than the 5Δ architecture. Therefore, the hybrid configuration is best suited for the proposed architecture.

Fig. 28 shows the comparison of architectures with CPL and DPL configurations. The designs of the 5Δ configurations, 5del_cpl and 5del_dpl, have their performances improved sensibly over the other architectures. Therefore, the CPL and DPL logic styles are more suitable to be implemented on this architecture.

V. CONCLUSION

The architectures of 4-2 and 5-2 compressors are analyzed and different CMOS logic style circuit implementations of their constituent modules are explored. A new low-power circuit with good drivability is proposed for the complex XOR* logic module which is used to co-generate the XOR-XNOR outputs. A novel 5-2 compressor architecture of 4Δ delay is also proposed. A new design of the carry generator cell has been spawned as a result of this unique architecture. In order to realistically assess and compare the figures of merits of different configurations of 4-2 and 5-2 compressors at various supply voltages, new simulation environments are established to ensure the measured performances are still sustainable when these cells are integrated in a CSA tree. The simulation results show that the 4-2 and 5-2 compressors constructed with the novel XOR* cell is able to function down to 0.6 V, and features high speed and low-power characteristics. Our proposed 5-2 compressor architecture outperforms all the other architectures over the range of voltages simulated, particularly when it is configured with the proposed circuits for the XOR* and the carry generator modules. Better performances against other architectures are also attained almost irrespective of the logic styles used for the circuit implementation of their constituent modules.

In summary, a library of excellent power efficiency 4-2 and 5-2 compressor cells based on CMOS process technology has been developed for implementing high speed and low-power multipliers operable at ultra low supply voltages.

REFERENCES


**Chip-Hong Chang** (S’92–M’98–SM’03) received the B.Eng. (Hons) degree in electrical engineering from the National University of Singapore, Singapore, in 1989, and the M.Eng. and Ph.D. degrees in electrical and electronic engineering from Nanyang Technological University, Nanyang, Singapore in 1993 and 1998, respectively.

He worked as a Component Engineer, General Motors, Singapore, and as a Technical Consultant, Flextech Electronics, Singapore, in 1989, and 1998, respectively. He joined the Electronics Design Centre, Nanyang Polytechnic, Singapore as a Lecturer in 1993. Since 1999, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he is currently an Assistant Professor. Dr. Chang has served a number of administrative roles during his academic career. He holds concurrent appointments at the university as the Deputy Director of the Centre for High Performance Embedded Systems, and the Program Director of the VLSI Design and Embedded Systems Research Group, Centre for Integrated Circuits and Systems. His current research interests include low-power arithmetic circuits, design automation and synthesis, and algorithms and architectures for digital image processing. He has published more than 80 refereed international journal and conference papers, and book chapters.

**Jiangmin Gu** (S’01) received the B.Sc. degree in the physics and the M.Eng. degree in electronic engineering and information science from the University of Science and Technology, Hefei, China in 1997, and 2000, respectively. He is currently working toward the Ph.D degree in electrical and electronic engineering at Nanyang Technological University, Singapore.

His research interests are low-power very-large-scale integration design methodologies and optimization of CMOS arithmetic circuits.

**Mingyan Zhang** (S’02) received the B.Eng. degree with first class honors in 2002, from Nanyang Technological University, Singapore, where she is currently working toward the M.Eng. degree.

Her research interests include low-power very-large-scale integration digital circuit design and digital image processing.