Techniques and prospects for fault-tolerance in post-CMOS ULSI

Yangyang Tang†‡, Gopalakrishnan Sundararajan†, Chris Winstead†, Emmanuel Boutillon∗,
Christophe Jégo† and Michel Jézéquel‡

∗Université de Bretagne Sud, UMR CNRS 6258 Lab-STICC, Lorient, France
†Dept. of Electrical and Computer Engineering, Utah State University, Logan, Utah 84322
‡Institut Polytechnique Bordeaux, UMR CNRS 5218 Lab-IMS, Bordeaux, France
§Institut TELECOM/TELECOM Bretagne, UMR CNRS 6258 Lab-STICC, Brest, France

Abstract—This paper presents a survey of fault-masking techniques suitable for tolerating short-duration transient upsets in minimum-scale switching devices. Two types of fault masking are considered. The first type, coded dual-modular redundancy (cDMR), represents a family of parity-checking methods suitable for correcting a low rate of transient upsets. The second type, Restorative Feedback (RFB), is a triple-modular solution suitable for compensating a higher rate of transient upsets. We show that cDMR can be used efficiently for crossbar-style logic, but is not efficient in general for all logic styles. By contrast, RFB offers a fixed redundancy, and can be applied in general to any logic circuit. Finally, we propose novel circuits for ternary Muller C implementation based on carbon nanotube FET devices.

I. INTRODUCTION

When devices are scaled to minimum physical dimensions, very small energy is available for storing and communicating logic levels between gates. Additionally, when devices are integrated at maximum density, the resulting high thermal density may seriously increase transient noise. Logic circuits must therefore operate with reduced signal-to-noise ratios that result in a high rate of momentary logic faults and memory state upsets. Transient signal upsets can be caused by thermal noise, electromagnetic interference (e.g. from switching events in neighbouring devices on the same chip), timing failures in synchronous pipelines, power supply noise, high-energy particle collisions and other sources [1]–[4]. Various techniques are available to mask the occurrence of momentary faults for binary and multiple-valued logic system [5]–[11].

In this work we present a survey of methods for tolerating transient upsets in nano-CMOS and post-CMOS ULSI systems. We first present a coded dual-modular redundancy (cDMR) technique, which has low redundancy in some cases. The cDMR method is permissible if certain constraints are imposed on the logic synthesis. The required constraints sometimes result in more complex logic synthesis, which may limit the generality of cDMR solutions.

Triple-Modular Redundancy (TMR) [12] provides a more general solution for protecting black-box logic modules without requiring any constraints on the logic synthesis method. TMR-style solutions have a fixed redundancy of two, whereas the cost of cDMR varies for each synthesized circuit. Hence it may be preferable to use TMR because it provides more certain characteristics during architecture planning.

The Restorative Feedback method (RFB) [13] is a TMR-style solution which is specialized for correcting transient faults in digital logic. The RFB method was previously shown to be applicable to multiple-valued logic circuits. In this paper, we present novel ternary implementations of the Restorative Feedback (RFB) method, which is comprised of Muller C-elements. We present two ternary C-element implementations using carbon nanotube FET (CNTFET) devices, which may be used to realize a multiple-valued RFB solution. A review of alternative implementations and tradeoffs is also presented.

II. CODING DUAL-MODULAR REDUNDANCY TECHNIQUES

One of the authors (Winstead) has proposed an LDPC-coded Fault Compensation Technique (LFCT) [14] which achieves reliable operation in the presence of transient and permanent defects. The LFCT method uses Gaudet and Rapley’s theory [15] of stochastic decoding to correct errors that appear at the output of some logic computation.

The LFCT architecture is depicted in Fig. 1. In the proposed LFCT system, a logic function \( f_1 \) is subject to transient internal upsets, causing one or more errors to appear on the function’s outputs, \( s \). To correct these error, a second function \( f_2 \) is introduced which maps the input data \( u \) to a vector of parity-check bits \( r \), so that the output \([s \ r]\) forms a codeword according to a traditional block error correction design. In this system, we assume that the input vector \( u \) is error-free. If the number of errors in \([s \ r]\) is sufficiently small, then they are corrected by the ECC module, resulting in a correct output vector \([\hat{s} \ \hat{r}]\). Finally the corrected data \( \hat{s} \) can be propagated to other logic modules for additional computations.

![Architecture of the LFCT method](image.png)

Figure 1: Architecture of the LFCT method. The original logic function \( f_1 \) is composed with a block ECC code to create the parity-mapped function \( f_2 \). Outputs are then corrected by the ECC circuit.
If the logic function is constrained to use a crossbar synthesis architecture, as commonly proposed for post-CMOS electronics [16], [17], then the fault statistics are suitable for using the LFCT technique. To explain this constraint, we may contrast the two circuits shown in Fig. 2. Both circuits implement a two-bit adder function. Fig. 2(a) shows a traditional ripple-carry implementation, and Fig. 2(b) shows a crossbar implementation. In the ripple-carry implementation, a single gate error may propagate to the several of the output signals. An example of error propagation is indicated by the \( \star \) symbol in Fig. 2(a). In the LFCT system, error-propagation may induce many simultaneous faults in the \([s\ r]\) codeword, which are not likely to be correctable.

In the crossbar implementation, as shown in Fig. 2(b), logic is implemented by fabrics of AND-logic and OR-logic. The “dots” in Fig. 2(b) indicate the placement of junctions which physically implement the logic operations. In this style of implementation, every operation is associated with only a single output. If a momentary fault occurs at some junction, it will propagate only to a single output. This style of logic guarantees that single-error events are correctable. The major disadvantage of crossbar logic is that the operation counts are not optimal. The crossbar adder in Fig. 2(b), for instance, has 57 separate operations. Crossbar logic does not generally obtain minimized gate complexity, but it offers improved reliability by eliminating error propagations. The added gate complexity could be considered as a form of redundancy. Furthermore, if crossbar-style logic is used with the LFCT method, then the logic function can be duplicated only once to provide error correction.

The authors (Tang, Winstead, Boutillon, Jégo and Jézéquel) have developed a specialized LFCT method, called the LDPC Stochastic Decoding (LSD) architecture [18], which can be considered as a circuit-level implementation of the LFCT concept. To evaluate the LSD solution, an example system was simulated using a variety of error-correcting codes. All codes are rate 1/2, meaning that there is one redundant parity bit for each systematic bit. The Bit Error Rate (BER) results for the LDPC codes applying LSD algorithm are shown in Fig. 3. In our simulations, the output bits \( s \) and \( r \) from \( F(x) \) and \( E \cdot F(x) \), respectively, are assumed to have a uniform independent error probability of \( \alpha \). Internal errors are also modeled in the correction circuit, with a uniform error probability equal to \( \epsilon \). Stuck-at faults were also inserted in uniformly random positions in \([s\ r]\) with a fault rate of \( \beta \), here, \( \beta = 0.001 \).

As long as \( \alpha < 0.05 \), the output \( s \) from \( F(x) \) can be recovered with a significantly reduced error probability. As \( \alpha \) is reduced below \( \beta \), the performance becomes dominated by the gate-level fault probability \( \epsilon \). The results show that, in the case of \( \epsilon = 10^{-5} \), the LSD architecture introduces gains about two orders of magnitude by comparison with uncoded data (output \( s \) from \( F(x) \)). Consequently, the proposed decoding method is able to suppress the resulting error probability to a level equal to the decoder’s internal fault rate.

Figure 2: Implementations of a two-bit binary adder function, representing a traditional ripple-carry design (a), and a crossbar design suitable for some nanoelectronic device families (b). The \( \star \) symbol indicates the occurrence of an error which propagates to multiple signals.

Figure 3: Simulation results for rate-1/2 LDPC codes based on LSD architecture with five iterations. The hard-fault rate \( \beta \) is 0.001.
III. TMR Method: Restorative Feedback

TMR methods require duplicating a logic function twice, and require no constraints on how the logic functions are synthesized (i.e. TMR methods can correct cases of error-propagation). Although cDMR requires only a single duplication, it requires suboptimal synthesis that may significantly increase the operation count. Therefore TMR methods are more attractive to be used with traditional non-crossbar logic styles. A variant of TMR, referred to as Restorative FeedBack (RFB), was previously described by one of the authors (Winstead) [13]. The RFB method is based on Muller C-element gates, as shown in Fig. 4. The RFB circuit is derived from the theory of stochastic iterative decoding, and achieves better performance than traditional majority-based TMR for correcting momentary upsets. The RFB method is also applicable to ternary logic systems as well as conventional binary logic circuits. A ternary RFB circuit was previously described using CMOS semi-floating gate circuits.

In this section, we consider a post-CMOS implementation based on carbon nanotube FET devices (CNTFETs). We first review the basics of ternary logic, and then investigate the CNTFET transistor, its operation and multi $V_{th}$ design in the context of CNTFET devices. Muller C Element is presented afterwards. At last, two proposed circuits for implementing ternary C-element are illustrated, as well the simulation results.

A. CNTFET Transistor

Fig. 5 shows the cross-section of a CNTFET transistor. CNTFET transistors are constructed by replacing the silicon channel of traditional MOSFETs with semiconducting carbon nanotubes (CNT). The CNTs deployed in the channel region can be either metallic or semiconducting depending on the angle of atom arrangement along the tube. Similar to traditional MOSFETs, CNTFETs have four terminals: drain, gate, source and back-gate [19].

B. Overview of Ternary Logic

In this subsection, we review the basics of ternary logic. Ternary logic functions are defined as those having importance if a third value is introduced into the binary logic value. Ternary Logic values consists of three states (1, 2, 3) representing false, undefined and true respectively [20]. Table I shows the logic values and the corresponding voltage levels used in this paper to represent ternary logic values.

The inverter is a basic logic gate used in digital design. In ternary logic, there are three types of inverters: positive ternary inverter (PTI), standard ternary inverter (STI) and negative ternary inverter (NTI). The truth table for all the inverters mentioned above is shown in table II.

The current through the CNTFET transistor is controlled by adjusting the device parameters such as gate length, number of nanotubes, chirality vector and the pitch distance [21]. The operation of CNTFETs is very similar to traditional MOSFETs except for differences in the device orientation. In contrast to MOSFETs, the source and drain terminals of CNTFETs transistors are not interchangeable. This is because doping variations are introduced in the source and drain regions during lithography rendering it impossible to interchange them [22].

For CNTFET circuits, multiple $V_{th}$ transistors can be achieved by tuning the chirality of each transistor. Even though CNTFET are relatively new, plenty of existing literature has documented the feasibility of adjusting the $V_{th}$ of transistors by growing different nanotube diameters [23].

The $V_{th}$ of the intrinsic CNT channel is given by:

$$V_{th} \approx \frac{E_g}{2e} = \frac{\sqrt{3}}{3} \frac{aV_{x}}{eD_{CNT}}$$

where $a = 2.49 \times 10^{-10}$ m, $V_x = 3.033$ eV, and

$$D_{CNT} = \frac{\sqrt{3}a_0}{\pi} \sqrt{n^2 + m^2 + mn}$$

where $a_0 = 0.142$ nm is the inter-atomic distance between each carbon atom and its neighbor and $(m, n)$ is called the chirality vector that describes the structure of a carbon nanotube [24].

C. Muller C Element

In this subsection, we describe the binary and ternary Muller C elements and its operation. The C element has been employed in asynchronous circuits design [25]. Fig. 6 shows the circuit and the schematic of the Muller C element. The

---

**Figure 4:** The RFB circuit based on Muller C-element gates. The $S$ gate is a storage element.

**Figure 5:** Side view of a CNTFET transistor.

**Figure 6:** Schematic of the Muller C element.
Table I: Ternary Logic Symbols

<table>
<thead>
<tr>
<th>voltage-Level</th>
<th>Logic Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.5V_{DD}</td>
<td>1</td>
</tr>
<tr>
<td>V_{DD}</td>
<td>2</td>
</tr>
</tbody>
</table>

Table II: Ternary Inverter Truth Table

<table>
<thead>
<tr>
<th>Input X</th>
<th>STI</th>
<th>PTI</th>
<th>NTI</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table III: Muller C-element truth table. The $S$ value indicates that the C-element holds its state (i.e. does not change value) for the corresponding pattern of inputs.

<table>
<thead>
<tr>
<th>Input a</th>
<th>Input b</th>
<th>Output c</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$S$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$S$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>$S$</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>$S$</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>$S$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$S$</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Muller C-element is a fundamental circuit element widely used for control synchronization in asynchronous designs. In general, a C-element is a state holding circuit which is transparent when all its inputs are equal, and holds the previous output otherwise. The C element consists of two inputs $a$, $b$ and one output $c$. Table III shows the truth table of the binary Muller C element. The logic equation describing the behaviour of the Muller C element is described as follows:

$$c = \overline{c} \ (a + b) + a \cdot b$$

The ternary Muller C element is similar to its binary counterpart except for the fact that the inputs and the output could take three logic values. Hence there will be an increase in the number input combinations. Table IV shows the truth table of ternary Muller C element. From the table, we find that a new value is latched at output $c$ when the two inputs $a$ and $b$ are equal such that $c = a = b$. Otherwise, the ternary Muller C retains its old value.

### D. Static Complementary CNTFET Logic Implementation

Static complementary logic design is a widely used design style in MOS technology because of its robustness and low power consumption. In CNTFETs also static complementary design consisting of N-CNTFETs and P-CNTFETs can be employed resulting in significant advantages. In this design, we use a combination P-CNTFETs and diode connected CNTFETs. Fig. 7 shows the circuit diagram of the Static Complementary CNTFET Muller C element. The initial portion of the Muller C consists of five P-CNTFETs and five N-CNTFETs. The inverter latch consisting of cross coupled inverters are also implemented using static complementary CNTFET design from literature [26].

The operation of the circuit is as follows. When the input voltages are below 300 mV P-CNTFET transistors $T_1$, $T_2$, $T_4$ and $T_5$ turn on thereby pulling the node cout to a logic high value. This logic value is passed on to cross coupled inverters thereby pulling the output $c$ to a low value. When the input voltages are raised above 300 mV transistors $T_4$, $T_5$, $T_9$ and $T_{10}$ are off and the combination of transistors $T_1$, $T_2$, $T_7$ and $T_8$ along with the diode connected transistors $T_3$ and $T_6$ produce a voltage drop of 0.45 V or logic 1 at node cout. This intermediate logic value is fed to the inverter latch thereby producing a voltage drop of 0.45 V or logic 1 at the output $c$. As the input voltages are raised above 600 mV, transistors $T_4$ and $T_5$ are off and transistors $T_9$ and $T_{10}$ are on which pull the node cout to a logic low value. This low logic value is fed to the cross coupled inverters thereby driving the node output to a high value. The Transistors $T_1$, $T_2$, $T_7$ and $T_8$ have diameter of 1.487 nm. Transistors $T_4$, $T_5$, $T_9$ and $T_{10}$ have diameter of 0.783 nm. Diode connected transistors $T_3$ and $T_6$ have diameter of 1.018 nm.

E. CNTFET Muller C Element with different back gate biased voltages

Fig. 8 shows the circuit diagram of CNTFET based Muller C element based on different back gate biasing voltages. In this implementation, two transistors are reduced compared to
Two supply voltages are used in this implementation. The first supply voltage is 0.45V and the second supply voltage is 0.9V. When the input voltage is below 0.3V, T1, T2, T3 and T4 are on and the remaining transistors are off. A high output voltage resulting from first part is then fed to the inverter latch thereby resulting in low output voltage at the output c. The output voltage is 0V. When the input voltage is between 0.3V and 0.6V T5 and T6 are on and T1, T2 are operating in sub-threshold region. As discussed in [27] lower threshold CNTFET devices operating in sub-threshold region with forward biasing offers high drive current which is equal in measure compared to the drive current provided by higher threshold CNTFET device operating in super-threshold region. Therefore, the output is the first supply voltage VDD1 which is 0.45V. When the input voltages are above 0.6V then transistors T1, T2, T3 and T4 are off and low voltage value appears at the cout which is fed to the inverter latch resulting in high output at c.

The designs were built and evaluated using HSPICE circuit simulator. The technology node used in this paper to evaluate the two designs was 32 nm. Stanford CNTFET device models were used to design we perform simulations in HSPICE using CNTFET parameter models by Stanford Nanoelectronics Group in the 32nm technology node [28]. All the designs were simulated at supply voltage of 0.9 V and at room temperature of T = 27 C.


In this subsection we present our simulation results. The static and the resistive load CNTFET designs were built and simulated using HSPICE. Transient simulations were run on the two designs to verify the functionality of the designs. Figure 9 and Figure 10 show the transient simulation results for Complementary CNTFET design. From the transient simulations we observe that when the first input a is held at a steady state and input b is swept from logic low to a high value the output changes state when the two inputs are at the same logic level. When the inputs a and b are at a different logic level the output holds the previous latched logic value. The inputs are swept from logic low to logic high and then changed to logic low in constant steps and the output is observed for a full low to high and a high to low transition. From the analysis, we can conclude that the ternary C element works as intended.

Table V shows the average delay, average power and the Power Delay Product (PDP) for the two proposed CNTFET ternary Muller C designs. From the table, we could see that though back gate design dissipates less power compared to the complementary CNTFET design it has a larger propagation delay which results in a higher power delay product.

IV. CONCLUSION

While electronic circuit is downsized into the nanometer scale, efforts are necessary to jointly optimize robustness and efficiency, while accounting for the probabilistic nature of nanometer devices. Among all available fault-tolerant methods to deal with transient fault occurrence, two latest works, the DMR-based LDPC-coded Fault Compensation Technique (LFCT) and TMR-based Restorative Feedback (RFB) technique, are first reviewed. Finally, two novel implementations of ternary Muller C-element using carbon nanotube FETS (CNTFETs) are presented for the purpose of error-resilience in multi-valued logic circuit.

ACKNOWLEDGEMENT

This work was supported by the US National Science Foundation under awards ECCS-0954747 and CCF-0916105.

REFERENCES


