Hardware Error Correction using Local Syndromes

Mohamed Hafidhi, Emmanuel Boutillon

Hardware Error Correction using Local Syndromes

Mohamed Mourad Hafidhi and Emmanuel Boutillon
Lab-STICC, UMR 6582, Université de Bretagne Sud 56100 Lorient, France
email: mohamed.hafidhi@univ-ubs.fr, emmanuel.boutillon@univ-ubs.fr

***Abstract*—Increasing the integration density offers the possibility for designers to built very complex system on a single chip. However, approaching the limits of integration, circuit reliability has emerged as a critical concern. The loss of reliability increases with process/voltage and temperature (PVT) variations. Faults can appear in circuits which can affect the system behaviour and lead to a system failure. Therefore it is increasingly important to build more fault tolerant resilient system. This paper [1] proposes a new fault tolerant scheme, the Duplication with Syndrome based Correction (DSC) scheme. Two criteria were considered to evaluate the proposed scheme: the reliability (probability that no error appears in the output of the architecture) and the hardware efficiency of the architecture. Results show that the DSC scheme reduces the complexity by 32%, compared to the classical Triple Modular Redundancy (TMR) scheme, while maintaining a level of reliability closed to the TMR. The paper shows also an example of signal processing applications where the DSC has been used to protect the correlation function and filters inside the tracking loops of the Global Positioning System (GPS) receiver.**

I. INTRODUCTION

Due to the increasing demand for enhancing performance and functionality at reduced area and cost, transistors have been scaled down over the past four decades. This growth has helped to increase the number of transistors per unit area and to optimise the performance and power consumption of circuits. Today chips employ billions of transistors, include multiple processor cores on a single silicon die, run at clock speeds measured in gigahertz, and deliver more than 4 million times the performance of the first ship [1]. Moreover, since power consumption is proportional to the square of the supply voltage V_{dd}, voltage scaling has been started in the late 80s in order to reduce consumption of circuits. During the last decades, V_{dd} was scaled from 5V to 3.3V then to 2.5V and it is predicted to be reduced to 0.64V in 2028 [2].

The increase of integration density with technology scaling has offered the possibility for designers to built very complex system on a single chip, reducing, so, the cost of circuit in term of area and power consumption, and improving their speed and performances. However, approaching the limits of integration, circuit reliability has emerged as a critical concern [3]. A fault or a set of faults may affect the system behaver if they are not masked and they can cause a system failure. Therefore, it is extremely important to protect systems from fault's impact to achieve acceptable reliability and maintain low complexity, cost and power.

[1]This work has received a French government support granted to the COMIN Labs excellence laboratory and managed by the National Research Agency in the "Investing for the Future" program under reference ANR-10-LABX-07-01. It has also received support from the Brittany Region.

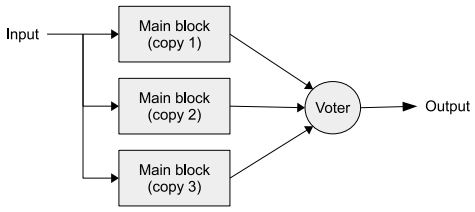


Fig. 1: The TMR concept

Fault-tolerance is the set of measures and techniques that aim to enable continuity of correct service delivered by a system even in presence of errors due to PVT variations coupled with increased advancement of CMOS technology. In the literature, a considerable amount of architectures has been proposed to mask or mitigate errors in different context, but what they have in common is that they all use redundancies to detect different types of errors. Redundancy takes two forms: spacial and temporal. The spatial redundancy refers as a replication of blocks, functions or data in a system. In the temporal redundancy, the same operation is repeated multiple times and by comparing the result in different instant, the presence of errors in the system is detected.

It is John von Neuman that pioneered the idea of using redundancy, in the 1950's, to improve the reliability of systems [4]. Then, the well-known Triple Modular Redundancy (TMR) appear as a similar approach with less complexity. In a TMR system, the original module is replicated three times, and error correction is achieved by a majority vote operation [5]. Fig. 1 illustrates the TMR concept. One of the reason of its popularity is its high ability to protect circuits from all type of errors. However, it is a very costly technique, an overhead of more than 200% compared to the original module. For extremely critical applications, such as space, avionics and healthy applications, where the system cost is less important than its reliability, TMR can be used to protect circuits. Otherwise, the need for resilient technique consuming less power and approaching the performance of the TMR at the same time, is increasing. In the context of signal processing applications, to reduce the area overhead, the authors of [6] propose to add only one additional module to the original module. The additional module gives a reduced precision estimation for the output of the original function and consumes less power than the original. The final output is chosen between the output of the original module and the output of the reduced replica. The scheme is referred as the Algorithmic Noise Tolerance (ANT) and is shown in Fig.2.

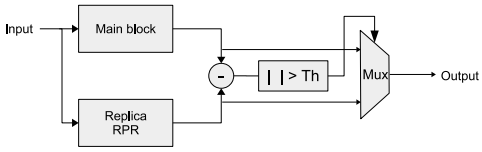


Fig. 2: The ANT scheme

Error Correction Codes have also been proposed to protect memories, [7], and then applied for interconnect networks [8]. All these methods are based on the spacial redundancy. Architectures based on the temporal redundancy usually use the double sampling technique to detect the presence of errors and a re-computation of operations for the recovery procedure [9]. Temporal redundancy comes with a throughput penalty and can be suitable only for application that tolerate spending extra time to recompute operations.

This paper presents a new resilient scheme, named the Duplication with Syndrome based Correction (DSC) scheme; The proposed scheme is based on the duplication of a module and correction using a syndrome based approach. The remainder of this paper is organized as follows. Sec. II introduces the DSC scheme in details and gives informations about its ability of correcting errors. Sec. III presents an example of application where the scheme can be used and the evaluation methodology. Sec. IV provides results of the comparison between DSC and the classical TMR schemes.

II. Principle of the DSC method

Before describing the DSC method, let us define minimum hypothesis on targeted signal processing application required to apply the DSC method.

A. Definitions

Definition 1: Let consider (G,+) a set with an operation (+) that combines any two elements x_1 and x_2 to form another element denoted x_1+x_2. (G,+) is a group if the four following properties are satisfied:

- **Closure**: For all x_1, x_2 in G, the result of the operation, x_1 + x_2, is also in G,
- **Associativity**: For all x_1, x_2 and x_3 in G, (x_1 + x_2) + x_3 = x_1 + (x_2 + x_3),
- **Identity element**: There exists an element e in G such that, for every element x_1 in G, e + x_1 = x_1 + e = x_1. Such an element is unique
- **Inverse element**: For each x_1 in G, there exists an element x_2 in G, commonly denoted -x_1, such that x_1 + x_2 = x_2 + x_1 = e, where e is the identity element.

Definition 2: Let consider two groups (A,+) and (B,+). We define a group homomorphism Ψ: (A,+) $\longrightarrow$ (B,+) where for all x_1 and x_2 in A, Ψ (x_1 + x_2)= Ψ (x_1) + Ψ (x_2).

There is many examples of group homomorphisms in the context of signal processing application, for example multiplication with a given value or polynomial, filtering, (convolution), matrix product, ... Those operations can be performed in any set (real, integer modulo P, Galois Field,...). ...

B. Three Duplication with Syndrome based Correction scheme (3-DSC)

In the literature, all the fault tolerant architectures propose to add extra redundancy to be able to detect occurrence of faults. However, today, many operations and functions exists in replica inside the same design. Let us assume that there exist in a design three identical group homomorphism, F, processing in parallel on three independent inputs, denote x_1, x_2, x_3 to generate $y_1 = F(x_1)$, $y_2 = F(x_2)$ and $y_3 = F(x_3)$ (see Fig.3a). Researchers previously propose to triplicate each operation independently and add a vote majority to mask errors, as shown in Fig 3b. With that, The design that contained three group homomorphisms at the beginning, will contain nine group homomorphisms and three voters. We propose in this paper a resilient scheme that will contain only six group homomorphisms F and a syndrome based corrector instead of three voter to mask errors. The proposed scheme is composed of the three original group homomorphisms, and three other redundant group homomorphisms F computing $z_4 = F(x_1 - x_2)$, $z_5 = F(x_2 - x_3)$ and $z_6 = F(x_3 - x_1)$, as shown in Fig. 4. In this figure, $y_k = F(x_k)$ for k = 1, 2 and 3, are replaced by $z_k = F(x_k)$, for k = 1,2 and 3 since the final y_k values are estimated after the correction mechanism. According to the group homomorphism structure of the fonction, in case of error free computation, the three following equations are all verified

$$\begin{cases} z_4 = F(x_1 - x_2) = F(x_1) - F(x_2) = z_1 - z_2 \\ z_5 = F(x_2 - x_3) = F(x_2) - F(x_3) = z_2 - z_3 \\ z_6 = F(x_3 - x_1) = F(x_3) - F(x_1) = z_3 - z_1 \end{cases} \quad (1)$$

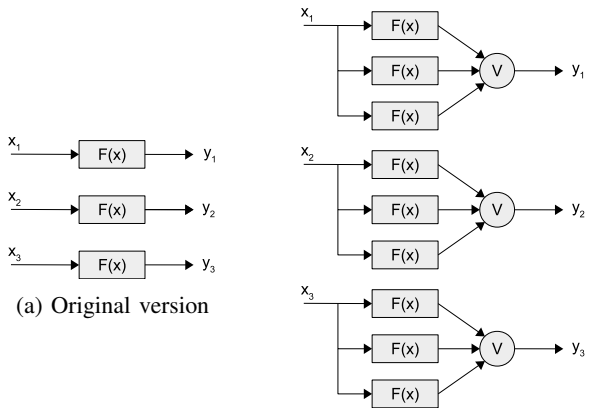
Note that, in case of computation with rounding and/or saturation noise, the strict equality can be relaxed to a distance compatible with the noise computation. Those three inequality can be represented by a triplet of Boolean values, or syndromes, $(s_1 s_2 s_3)$, with s_1 (respectively s_2 and s_3) equals to 1 if the first equation (respectively second and third equation) is not fulfilled.

In a more robust version, it is also possible to further consider Z_r, the parity check sum, defined as

$$Z_r \triangleq z_4 + z_5 + z_6 = (z_1 - z_2) + (z_2 - z_3) + (z_3 - z_1) = 0 \quad (2)$$

In case of no error, the syndromes are all equal to zero and the values of z_1, z_2 and z_3 are simply copied in the outputs y_1, y_2 and y_3.

If a single fault occurs in one of the six modules then the value of the syndromes allows to detect the error and correct it. Let us consider, for example, that an error occurs in the computation of $F(x_1)$ and that the first operator outputs $\tilde{z_1}$, with $\tilde{z_1} \neq z_1$. According to (1), replacing z_1 by the erroneous value $\tilde{z_1}$ in will generate the violation of the equality of the first and last equation of (1), i.e., leading the 3 syndromes $(s_1 s_2 s_3)$ be equal to $(s_1 s_2 s_3) = (101)$. In this case, it is still possible to recover the exact value y_1 by processing $y_1 = z_4 + z_2 = F(x_1 - x_2) + F(x_2) = F(x_1)$. This example can be generalised for any single error, as shown in Table I.



(a) Original version

(b) Protection using TMR

Fig. 3: Protection using the TMR concept

Faulty variable	Syndrome	y_1	y_2	y_3
$\tilde{z}_1$	(101)	$z_4 + z_2$	z_2	z_3
$\tilde{z}_2$	(110)	z_1	$z_5 + z_3$	z_3
$\tilde{z}_3$	(011)	z_1	z_2	$z_6 + z_1$
$\tilde{z}_4$	(100)	z_1	z_2	z_3
$\tilde{z}_5$	(010)	z_1	z_2	z_3
$\tilde{z}_6$	(001)	z_1	z_2	z_3

TABLE I: The syndrome corresponding to each faulty module

Finally, if more than one error occurs, in most of the case, the 3 equations won't be fulfilled. In that case, an error is detected an other correction/mitigation mechanisms at higher level may be activated but those mechanisms are out of the scope of the paper[2]. Nevertheless, it should be note that two errors event of same amplitude may not be detected. For example, if the function F is applied on integer and both z_1 and z_2 are affected by an additive noise of same amplitude n, i.e. $\tilde{z}_1 = z_1 + n$ and $\tilde{z}_2 = z_2 + n$, then equation $z_4 = \tilde{z}_1 - \tilde{z}_2$ will remain correct. The syndrome is thus (011), which implies the erroneous correction of z_3. Note that, in case of this double event of same amplitude, TMR system will also output a wrong value.

To conclude this section, we should mention that the ANT technique presented in [6] can be adapted to the proposed scheme. Instead of doing computation of z_4, z_5 and z_6 in full precision, it is possible to do them in a reduced precision to save area and power dissipation. The drawback is that, in case of error occurring in z_1 for example, the reconstructed value $y_1 = z_4 - z_2$ will have a degraded precision due to the reduced precision of z_4.

C. Generalisation: N-DSC scheme

In this section, a resilient scheme for a design with N group homomorphisms is proposed as an extension of the 3-DSC ($N > 3$). The extension is straightforward: to the N functions

[2]In [10], in case of error, the previous last correct value is given to a feedback filter in order to limit the propagation of error for example

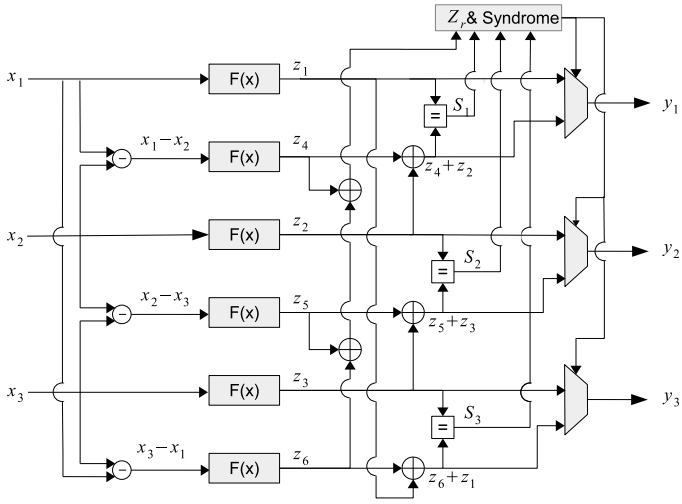


Fig. 4: Protection using 3-DSC

$z_q = F(x_q)$, $q = 1 \dots N$, N additional function are added to generate,

$$\begin{cases} z_{N+q} = F(x_q - x_q), \ \forall \ q = 1 \dots N-1 \\ z_{2N=F(x_N)-F(x_1)} \end{cases} \tag{3}$$

Similarly to the 3-DSC case, in case of no error, the N following equation are fulfilled,

$$\begin{cases} z_{N+q} = z_q - z_{q+1}, \ \forall \ q = 1 \dots N-1 \\ z_{2N} = z_N - z_1 \end{cases} \tag{4}$$

N syndrome s_q, $q = 1 \dots N$ can be defined, with $s_q = 0$ is the q^{th} equation is fulfilled, 1 otherwise.

From the local syndrome s_q and s_{q+1}, it is possible to evaluate if the computation of z_q is correct. In fact, if due to an error, $\tilde{z}_q$ is output instead of z_q, both syndromes s_q and s_{q+1} are equals to 1. In that case, the correct value of y_q can be estimated as $y_q = z_{N+q} + z_{q-1}$. The hardware required to perform those operation is shown in Fig. 5.

To summarise, any single error in the N-DSC module is detected and corrected. Two errors that appear in the $N-DSC$ module can be corrected if and only if their corresponding local syndrome doesn't interact. However, two errors that appear in two adjacent redundant modules z_{N+q} and z_{N+q-1} lead to a wrong estimation of y_q. In fact, if z_{q-1}, z_q and z_{q+1} are correct while $\tilde{z}_{N+q-1}$ and $\tilde{z}_{N+q}$ are faulty the corresponding syndrome $s_{q-1} = 1$ and $s_q = 1$, and thus, according to the correction mechanism, y_q will be estimated as $y_q = \tilde{z}_{N+q} + z_{q-1}$. This type of error can result to non correctable error, which may be problematic if the result of function F is used to feed a feedback loop. In this case we can propose to split the original design in blocks of 3/4/5 functions F and use the 3-DSC, 4-DSC and 5-DSC to protect each blocks.

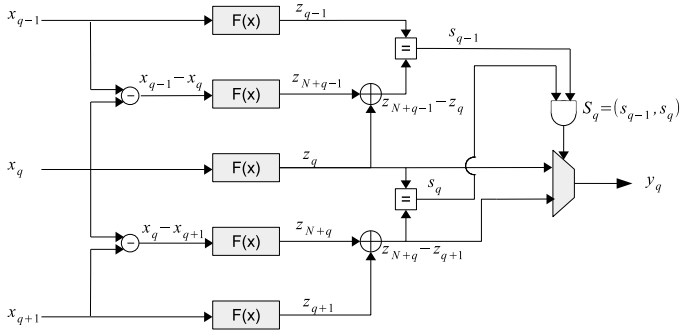


Fig. 5: Correction using Local syndromes for N-DSC scheme

III. Application of DSC in the GPS context

A. Introduction to GPS

The GPS is a well known technology that allows determining both the physical position and the absolute time of a receiver. The position in time and in space is determined thanks to a precise distance measurement with at least four GPS satellites. Each GPS satellite transmits a navigation message at 50-bits/s using the CDMA (Code Division Multiple Access) technology. The analytical expression of the transmitted signal of a satellite a is:

$$e_a(t) = c_a(t)\ d_a(t)\ e^{2\pi j f_{L1} t}$$

where:

- $d_a(t)$: navigation message of the a^{th} satellite,
- $c_a(t)$: a^{th} Coarse/Acquisition (C/A) satellite code with a Binary Phase Shift Keying (BPSK) modulation (i.e. $c_a \in \{-1, 1\}$),
- f_{L1}: the carrier frequency in the L1 GPS Band (Open Sevice).

GPS receiver has to demodulate the navigation message of different satellites in view to make the distance measurement. This involves two essential and sequential process: the acquisition process and tracking process. The acquisition process is the process by which the receiver identifies which satellites are in view. It is a a three-dimensional search to determine the GPS satellite identifier (which is the index of its associated C/A code), the code phase (represented by τ), and the carrier frequency offset due to Doppler effect (represented by f_d).

Since satellites are in continuous motion, the distance between any satellite and the receiver is dynamic. Besides to that, the carrier frequency of the received signal is also constantly changing in time due to Doppler shifts. Therefore, once acquired, GPS signals have to be tracked over time. To note here, any GPS receiver design contain at least four channel tracking module; each module tracks a unique GPS satellite.

To track a satellites' signals, each tracking module is composed of a correlation module and two tracking loops (the carrier tracking loop and the code tracking loop). The carrier tracking loop performs the task of aligning the local generated carrier with the incoming signal while the code tracking loop ensures the time alignment of the local generated codes. Each loop is made of discriminators, filters and generators. The correlation function is computed every 10 ms period to compare local signals with incoming signals. A maximum correlation output is achieved when the two signals are aligned. A simplified representation of the channel tracking module is given in Fig. 6.

B. Robustness of the Correlation Function

Faults when computing the correlation function can produce errors at its output. Because of feedback loops these errors will propagate over time and will corrupt generated carrier and codes. Loss in the signal tracking process can be reported, forcing the receiver to restart the initial signal acquisition procedure. So it is increasingly important to deal with the impact of the faults when they appear in the correlation process. For each tracking channel, the incoming signal is first multiplied by the generated carrier, and, then by three generated codes, c_E, c_P and c_L. As mentioned earlier, any GPS receiver design contain at least four channel tracking module, we duplicate the first multiplication of the four canal. Error correction is achieved to determine the correct X_{P1}, X_{P2}, X_{P3} and X_{P1} as shown in Fig.7. Then, the three multiplication in each tracking are duplicated and protected independently as illustrated in Fig.8.

C. Evaluation

To compare the proposed method to the TMR, we propose to use the approach defined in [11]. This approach characterises an architecture in unreliable hardware by two dimensional criteria: the reliability (probability that no error appear in the output of the architecture $P_{No-error}$ and the hardware efficiency of an architecture (defined as the normalised number of operation per unit area and time unit). To compute the efficiency, the nature of the computation inside the area unit is not specified, it can be a simple multiplier or more complex operation like an FFT transformation or a iterative system .... If we consider an operation that takes n area units and m area clocks to be executed, the efficiency is expressed as,

$$\gamma \triangleq \frac{1}{n \times m}\ \ operation/(areaunit \times timeunit) \quad (5)$$

In the following part, we will focus on the evaluation of the 3-DSC scheme as shown in Fig.7.

Triple Modular redundancy (TMR): Let P_M the error probability in a single module during one clock cycle, and n_v the area cost of the voter. The resulting error probability of the TMR is expressed as,

$$P_{TMR} = 1 - [\ (1-P_R)^3 + 3\ P_R\ (1-P_R)^2\ ] \quad (6)$$

The efficiency of the TMR is determined by,

$$\gamma_{TMR} = \frac{1}{(3.n + n_v).m} \quad (7)$$

The normalised hardware efficiency is defined as,

$$\Gamma_{TMR} = \frac{\gamma_{TMR}}{\gamma_{original\ module}} = \frac{n.m}{(3.n + n_v).m} \quad (8)$$

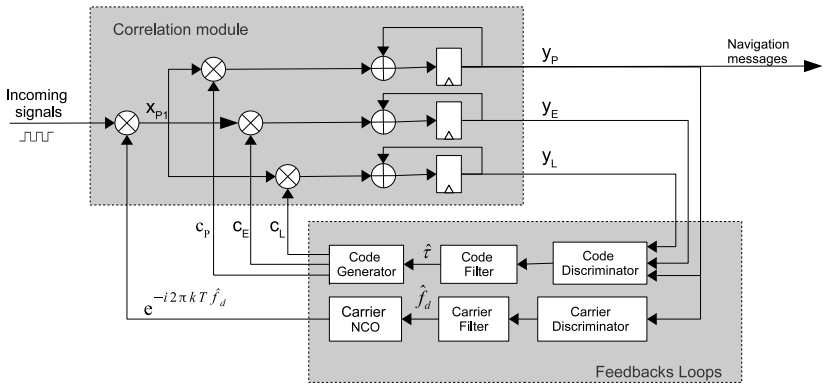


Fig. 6: Generic digital receiver channel block diagram.

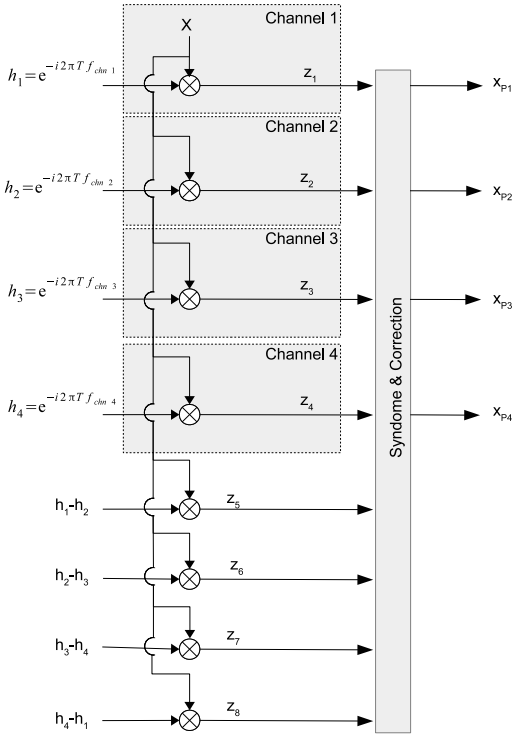


Fig. 7: Protection of the carrier multiplication

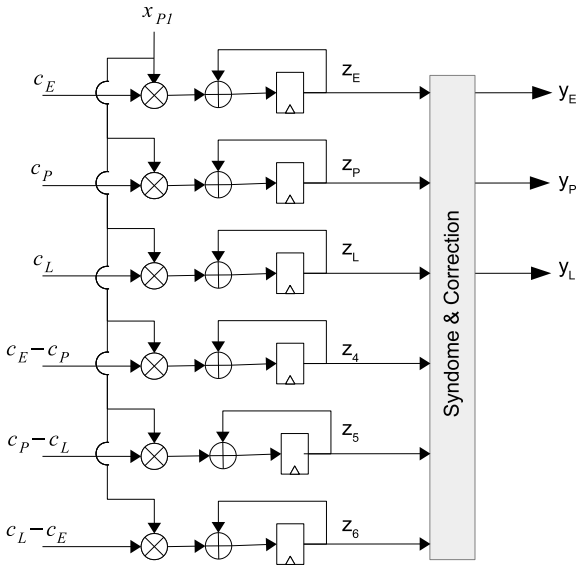


Fig. 8: Protection of the three codes multiplications

Duplication with Syndrome based Correction (DSC): Let consider the output y_1. y_1 is correct on two cases:

- when P_{cs}=0 and z_2 and z_3 are non faulty (i.e S=101)
- when z_1 is non faulty

The error probability in one of the SC scheme outputs is,

$$P_{DSC} = 1 - [\ (1 - P_R) + P_R\,(1 - P_R)^5] \qquad (9)$$

The efficiency of the DSC-scheme is,

$$\gamma_{DSC} = \frac{3}{(6.n + n_c).(m + m_c)} \qquad (10)$$

where n_c and m_c represents the number of area unit and time unit respectively. The normalised hardware efficiency is,

$$\Gamma_{DSC} = \frac{3.n.m}{(6.n + n_c).(m + m_c)} \qquad (11)$$

IV. Performance comparison

The TMR and DSC schemes were evaluated in term of hardware efficiency and reliability. Table II details the logic synthesis results in term of number of cell obtained using Synopsys Design compiler in the 45 nm technology. Varying the probability that an error occurs at the output of a single module, the probability that no error appear in the output of each scheme is compared in Fig. 9. From this figure we can see that the DSC scheme provides robustness closed to the TMR methods. Now, given a probability of an error at the output of a single module fixed to 10^{-3}, results from the analyse of the normalised hardware efficiency of each method are summarised in Fig. 10 as of function of the error probability.

Component	F	Voter	Syndrome and correction
Number of cells	629	64	189

TABLE II: Logic synthesis results in term of number of cell obtained by synthesizing the designs with Synopsys Design compiler in the 45 nm technology

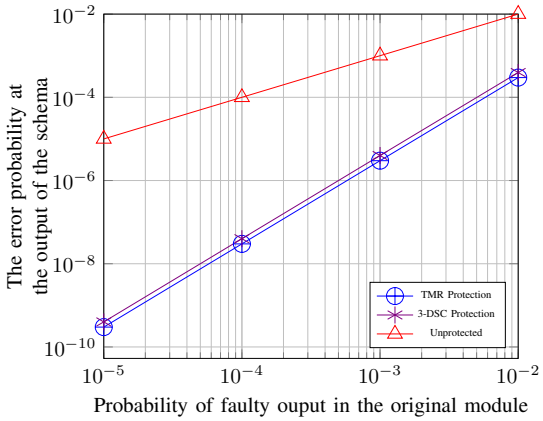


Fig. 9: Performance as a function of the upset probability p.

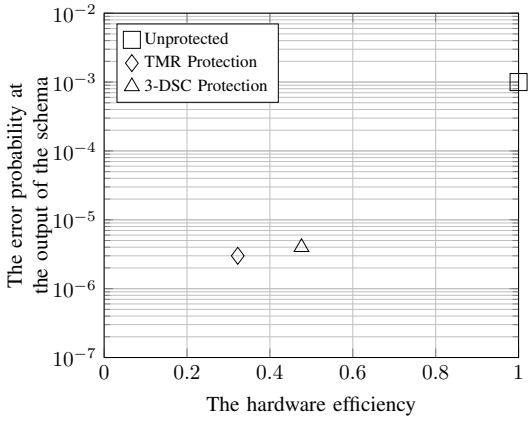


Fig. 10: Complexity and performance of the proposed methods with $p = 10^{-3}$.

V. Conclusion

Our results show an 32% improvement in the complexity with the proposed 3-DSC scheme compared to the classical TMR scheme. Moreover, we guarantee that the reliability offered by the 3-DSC is maintained closed to the TMR. As an example, the 3-DSC was used in the context of a GPS tracking application to protect some internal component such as the correlation function and filters. It has been shown in this paper that the 3-DSC can be extended to design a more general resilient scheme named N-DSC where N redundant modules are added to an original design that contains N identical modules operating on N different data.

References

[1] A. Danowitz, K. Kelley, J. Mao, J. P. Stevenson, and M. Horowitz, "Cpu db: Recording microprocessor history," *Queue*, vol. 10, no. 4, pp. 10:10–10:27, Apr. 2012.

[2] International Technology Roadmap for Semiconductors, "Summary 2013 ORTC Technology Trend Targets," https://www.dropbox.com/sh/z80p2ne051g770t/AADvOmLAEcVwCpxtexqSvE7ra/2013ORTC_SummaryTable.pdf?dl=0.

[3] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "The impact of technology scaling on lifetime reliability," in *Dependable Systems and Networks, 2004 International Conference on*, 2004, pp. 177–186.

[4] D. Bhaduri, S. Shukla, P. Graham, and M. Gokhale, "Comparing reliability-redundancy tradeoffs for two von neumann multiplexing architectures," *IEEE Transactions on Nanotechnology*, vol. 6, no. 3, pp. 265–279, 2007.

[5] C. Winstead, Y. Luo, E. Monzon, and A. Tejeda, "An error correction method for binary and multiple-valued logic," in *Multiple-Valued Logic (ISMVL), 2011 41st IEEE International Symposium on*, 2011, pp. 105–110.

[6] B. Shim, S. R. Sridhara, and N. R. Shanbhag, "Reliable low-power digital signal processing via reduced precision redundancy," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 12, no. 5, pp. 497–510, 2004.

[7] V. Gherman, S. Evain, M. Cartron, N. Seymour, and Y. Bonhomme, "System-level hardware-based protection of memories against soft-errors," in *Design, Automation Test in Europe Conference Exhibition, 2009. DATE '09.*, 2009, pp. 1222–1225.

[8] A. Ejlali, B. M. Al-Hashimi, P. Rosinger, and S. G. Miremadi, "Joint consideration of fault-tolerance, energy-efficiency and performance in on-chip networks," in *Design, Automation Test in Europe Conference Exhibition, 2007. DATE '07*, 2007, pp. 1–6.

[9] M. Nicolaidis, "Time redundancy based soft-error tolerance to rescue nanometer technologies," *17th IEEE Proceedings. VLSI Test Symposium, 1999*, 1999.

[10] M. M. Hafidhi, E. Boutillon, and C. Winstead, "Reducing the impact of internal upsets inside the correlation process in gps receivers," in *2015 Conference on Design and Architectures for Signal and Image Processing (DASIP)*, 2015, pp. 1–5.

[11] Y. Tang, E. Boutillon, C. Jgo, and M. Jzquel, "Hardware efficiency versus error probability in unreliable computation," in *2011 IEEE Workshop on Signal Processing Systems (SiPS)*, 2011, pp. 168–173.