Circuit Design Steps for Nano-Crossbar Arrays: Area-Delay-Power Optimization with Fault Tolerance

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Transactions on Nanotechnology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>TNANO-00011-2019</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>NANOARCH2018ATHENS</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>02-Jan-2019</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Morgul, Muhammed Ceylan; Istanbul Teknik Universitesi Elektrik-Elektronik Fakultesi, Electronic and Communication Engineering Frontini, Luca; Universita degli Studi di Milano Tunali, Onur; Istanbul Teknik Universitesi Elektrik-Elektronik Fakultesi, Electronic and Communication Engineering Vatajelu, Ioana; Institut Polytechnique de Grenoble, EE and CS Ciriani, Valentina; Universita degli Studi di Milano anghel, lorena; INPG, EE and CS Moritz, Csaba Andras; U Mass, Amherst, ECE Stan, Mircea; University of Virginia, ECE Alexandrescu, Dan; IROC Technologies Altun, Mustafa</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Crossbar Arrays, Logic Synthesis, Defect Tolerance, Fault tolerance, Performance Optimization, Memristor Arrays</td>
</tr>
</tbody>
</table>
January 02, 2019

Dear Professor Fabrizio Lombardi,


A preliminary version of this paper, titled “Integrated Synthesis Methodology for Crossbar Arrays” is presented at the 14th ACM/IEEE International Symposium on Nanoscale Architectures. Nearly 50% of material in this manuscript is new:

- At least 70% of the material in the logic synthesis are new;
  - we have proposed completely new technique for two-level logic synthesis for memristive nanoarrays
  - we have investigated multi-level logic synthesis for four-terminal switch based nanoarrays as well as other types (diode, memristor and FET), and we have compared them, first in this study
- At least 50% of the material in the subsection “Defect Tolerance for Four-terminal” is new, we have improved our technique and described in more detail;
- At least 95% of the material in the performance optimization is new, we have made in-depth analysis of complexity of delay and power performances of nanoarray types;
- We have added new section for a case study to increase comprehensibility of complete design methodology;

The subject area of this paper is “Developing a complete integrated synthesis methodology for circuits and architectures”.

Best Regards,

[Signature]

Mustafa Altun, PhD
Assistant Professor
Director of the Emerging Circuits and Computation Group
Istanbul Technical University

Email: altunmus@itu.edu.tr
Web: www.ecc.itu.edu.tr
Circuit Design Steps for Nano-Crossbar Arrays: 
Area-Delay-Power Optimization with Fault Tolerance 

M. Ceylan Morgul, Luca Frontini, Onur Tunali, E. Ioana Vatajelu, Valentina Ciriani, Lorena Anghel, Csaba Andras Moritz, Mircea R. Stan, Dan Alexandrescu, and Mustafa Altun

Abstract—Nano-crossbar arrays have emerged to achieve high performance computing beyond the limits of current CMOS. They offer area and power efficiency in courtesy of their easy-to-fabricate and dense physical structures consisting of regularly placed crosspoints as computing elements. Depending on the used technology, a crosspoint behaves as a diode, a memristor, a field effect transistor, or a four-terminal switching device. In this study, we comparatively elaborate on these technologies in terms of their capabilities for computing in terms of area, delay, and power consumption. Also, we consider fault tolerance capabilities of the arrays. Due to the stochastic nature of nano-fabrication, nanoarrays have much higher fault rates compared conventional technologies such as CMOS. As a result, this study introduces a synthesis methodology that considers basic technology preference for switching crosspoints and defect or fault rates of the given nanoarray as well as their effects on performance metrics including power, delay, and area.

Index Terms—Crossbar Arrays, Logic Synthesis, Defect Tolerance, Fault Tolerance, Performance Optimization, Memristor Arrays

I. INTRODUCTION

Nano-crossbars have emerged to be an alternative technology to CMOS [30]. They are fabricated with relatively cheap bottom-up nano-fabrication techniques rather than using pure lithography based conventional production. Due to the novel manufacturing techniques, produced fabrics are regular and in dense forms that result in area and power efficient structures [9] [2].

M. Ceylan Morgul and Mustafa Altun are with the Department of Electronics and Communication Engineering, and Onur Tunali is with the Department of Nanoscience and Nanoeengineering of Istanbul Technical University, Istanbul, Turkey e-mail: {morgul, onur.tunali, altunmus}@itu.edu.tr

Luca Frontini and Valentina Ciriani are with the Dipartimento di Informatica, Università degli Studi di Milano, Milan, Italy e-mail: {luca.frontini, valentina.ciriani}@unimi.it

E. Ioana Vatajelu and Lorena Anghel are with TIMA laboratory, Grenoble-Alpes University, Grenoble, France e-mail: {ioana.vatajelu, lorena.anghel}@imag.fr

Csaba Andras Moritz is with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, Massachusetts, USA e-mail: andras@ecs.umass.edu

Mircea R. Stan are with the Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, Virginia, USA e-mail: mircea@virginia.edu

Dan Alexandrescu is with IROC Technologies, Grenoble, France e-mail: dan.alexandrescu@irotech.com

This work is part of a project that has received funding from the European Union’s H2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement #691178, as well as supported by the TUBITAK-Career project #113E760.

Currently, computing is achieved with crosspoints behaving like switches, either as two-terminal or four-terminal. This is illustrated in Figure 1. Depending on the used technology, a two-terminal switch behaves either as a diode [13], a resistive/memristive switch [21], or a field effect transistor (FET) [22]. Diode and resistive switches correspond to the crosspoint structure in Figure 1(a); here, the switch is controlled by the voltage difference between the terminals. Figure 1(b) shows a FET based switch; here, the red line represents the controlling input. This is a unique opportunity that allows us to integrate well developed conventional circuit design techniques into nano-crossbar arrays. Finally, a novel four-terminal switch is demonstrated in Figure 1(c) and technology development is presented and analyzed in a recent paper [19] with detailed TCAD simulations. In mentioned paper, switching lattice can be fabricated using CMOS-like technology. It has four terminal either connected each other or not. Preferred state is actualized with a controlling input, which is not shown in the Figure 1(c) and has a separate physical formation from the crossbar, that is thoroughly explained for different technologies in [3] [19].

To illustrate different computing approaches, we show examples for the implementation of $f_{XOR_2} = x_1 \overline{y_2} + \overline{x_1} y_2$ in Figure 2. Logic synthesis models for diode and memristor based crossbars are quiet similar to Programmable Logic Array (PLA) as can be seen in Figure 2(a) and 2(b). Memristor based crossbars have one major difference that establishing the output goes through several states/loops (for further information,
Fig. 2. Implementation of $f_{XOR}$ with different nano-crossbar types: crosspoint as a) diode, b) memristor, c) FET, and d) four-terminal switch.

check [29]).

For FET based crossbars, each logic function product and dual function product is realized by a separate column, as seen in Figure 2(c). Each input is assigned to a row for controlling all the FETs on corresponding row.

Finally, a four-terminal based crossbar; here every crosspoint performs switching on all four directions. Control lines of crosspoints are not shown in Figure 2(d), yet detailed explanation of control lines can be found in [3] [19].

Regarding emerging technologies and nano-fabrication, fault rates are much higher for nano-crossbars, as expected, compared to those of conventional CMOS circuits [10]. Therefore, during logic synthesis, consideration of faults and defects is mandatory. This applies for the integration of both diode, FET based or novel four-terminal based logic synthesis methodologies. For this reason, researchers focus on challenges including defect and variance tolerances [25] [17]. Defect and variance tolerant approaches are closely related to logic realization and performance optimization, respectively.

Taking mentioned issues into account, we have developed a complete integration methodology for logic synthesis, defect tolerance and performance optimization. This methodology is designed as a step-by-step guide to combine modular research approaches into an entire production pipeline.

This work combines detailed survey for all steps of the methodology with new approaches and techniques for the missing parts, so a complete synthesis methodology is achieved. Surveyed and newly studied materials are listed below.

Surveyed items:

- Single and multi-output logic synthesis for diode, memristor, and FET based crossbars, and single-output logic synthesis of four-terminal lattice.

Newly studied items:

- Defect tolerance technique for diode, memristor, and FET based crossbars.
- Performance (delay-power) analysis of memristor arrays.

Fig. 3. Integrated synthesis methodology scheme for nano-crossbar arrays.

- Defect tolerance technique for four-terminal lattice.
- A greedy optimization algorithm for two-level single-output memristor crossbar (logic synthesis).
- Defect tolerance technique for four-terminal lattice.
- Performance (delay-power) analysis of diode, FET and four-terminal arrays.

Details of the Integration Methodology

As briefly explained above, nano-fabrication produces switching nano-crossbar arrays with structures or individual components having varied properties. Mentioned factors introduce random characteristics of which need to be carefully considered by synthesis process. For instance, a competent synthesis methodology must consider basic technology preference for switching elements and defect or fault rate of the
A diagram summary demonstrating every step of the method with annotation showing the certain research tasks is given in Figure 3.

First step of the synthesis process of a nano-crossbar involves the decision of switching technology which will be explained elaborately in Section II. Main purpose is to determine which of the diode/memristor, FET, or four-terminal based components are to be used. This step is one of the most important procedures determining the size of the nano-crossbar. Production with diode/memristor based technologies as well as with FET are explained and then logic synthesis design with four-terminal based switches is given.

Second step of the synthesis process of a nano-crossbar covers the permanent faults (defects forming in the course of fabrication) and the tolerance aspects, which will be described in Section III. Main purpose is to obtain a valid realization of a given logic function using two distinct approaches titled as defect-aware and defect-unaware. First method employs faults existing in nano-crossbar during the realization of logic function hence the name aware. Second method avoids the faults by attempting to find a fault-free region of nano-crossbar at the beginning so realization of given logic function is straightforward at the end.

Third step of the synthesis process of a nano-crossbar covers the performance analysis and optimization, which will be explained in Section IV. Main purpose is to analyse delay and power consumption of arrays by showing their dependencies on the properties of the given function as well as specifics of the used technology.

Final step of the synthesis process of a nano-crossbar involves the analysis of transient faults. Main purpose is to determine the effect of transient faults to the operational capacity of nano-crossbar and calculate fault tolerance performance.

Section V is devoted to further elucidate the proposed synthesis process with a case study.

II. LOGIC SYNTHESIS

In the beginning of the logic synthesis process, the crossbar technology, i.e., diode, memristor, FET, or four-terminal, needs to be determined based on the following criteria:

- Crossbar size (number of rows and columns)
- Number of outputs (single or multiple function realization)
- Fabrication complexity
- Power and delay specifications
- Application specifications

A decision can be made on the importance and priority of the listed items, depending on the application. For example, if an application includes also memory, then the memristor technology can be chosen for the realization of logic functions, since the memristor can also be used as a memory unit. Thus, these components can be fabricated using the same technique.

On the other hand, the realization of a logic function using a diode or a memristor based crossbar requires less number of crosspoints than those in the FET based crossbar. However, the FET based designs consume less power than the diode/memristor-based designs. Moreover, the four-terminal based crossbar includes less number of crosspoints than other crossbar designs [15].

As follows in the first part, we examine the logic synthesis techniques developed for diode, memristor, and FET based nanoarrays in the literature; we formulate their crossbar sizes needed to implement given functions. We also present results for four-terminal switch based arrays to synthesize multi-output functions.

In the second part, we present a new two-level synthesis technique for memristor based nano arrays and compare it with other techniques in the literature.

A. Area Comparisons for Different Crossbar Technologies

We present the logic synthesis step of the integration methodology, considering only the number of crosspoints in the crossbar arrays, which is actually the size of the crossbar array. The size of an crossbar array including diode, memristor, FET, and four-terminal is given as follows, where $n$ is the number of logic functions (the number of outputs); $f_i$ denotes the $i^{th}$ logic function, and $f_i^D$ stands for its dual with $1 \leq i \leq n$.

- **Diode:**\((\# \text{ of products of all } f_i) + n\) × \((\# \text{ of literals in } f) + n\)
- **Memristor:** \((\# \text{ of products of all } f_i) + n\) × \((\# \text{ of literals in } f) + 2n\) [worst-case]
- **FET:** \((\# \text{ of literals in } f) \times \((\# \text{ of products of all } f_i) + (\# \text{ of products of all } f_i^D)\)\)
- **Four-terminal:** \((\# \text{ of products of all } f_i) + n − 1\) [worst-case]

As mentioned in Section I, the logic synthesis on diode and memristive based crossbars is similar to the PLA like synthesis. Thus, the techniques, such as product sharing and phase changing used in the PLA design, are also applicable in these designs. Since, the array sizes can be further reduced using the product sharing, these array size formulations can be considered as an upper bound for the logic synthesis techniques.

For the single and multiple output function realization, the synthesis methodology for FET crossbar does not allow us to produce multi-level logic synthesis, only two-level approach can be used [23]. However, multi-level logic synthesis approach is applicable for the diode and memristive crossbars [27]. Therefore, the optimization on the array size still demands further research for the diode and memristor based designs.

The logic synthesis using four-terminal crossbars, generally known as switching lattices, is a new method. As shown in [3], Altun presented a useful logic synthesis technique for the switching lattices. However, this method cannot find the optimal solution in terms of the lattice size. Therefore, new specific logic synthesis methodologies are needed to be presented. As shown in [12] and [15], optimal synthesis
methodologies are provided. In addition, there are decomposition based techniques such as XOR based [14], [7], p-circuit [5] and dimension reducibility [6] decompositions as well.

However, all of these studies were only focused on the realization of a single logic function using switching lattices. On the other hand, in [1], three main steps are presented to realize the multiple functions using switching lattices. These steps are given as follows: 1) find the realization of each logic function using a switching lattice; 2) merge these lattices into a single lattice; 3) check if these lattices can be realized using a smaller number of rows and columns such that the final lattice includes a small number of four-terminal switches. This article is the first to present the sizes of diode, memristor, FET, and four-terminal based crossbar arrays on the multiple output functions. The results are given in Table I. Note that the results given in bold under the four-terminal column indicates that they are found using the approximate algorithm of [1], following the three steps described above. On the other hand, the other results are found using a divide and conquer method, based on the divide and synthesize (DS) method of [1], following the first two steps described above.

As can be observed from Table I, the four-terminal based crossbar arrays include significantly less number of crosspoints when compared to the diode, memristor, and FET based crossbar arrays.

### B. Proposed Algorithms for two-level multi output synthesis for Memristive Arrays

In memristive crossbar arrays, function outputs’ and their negations are produced [29]. However production order changes the array size. We have called phase_combination-0 and phase_combination-1 realization, if realization is happened on an output itself and its negation, respectively (note that in phase_combination-1, first its negation is produced [28]). Based on this property array size of memristive crossbars can be optimized.

In our previous study [28], we have proposed a greedy algorithm, which only considers output, individually and doesn’t consider output interrelation (no product sharing at analysis). We can call this algorithm as “initial Greedy Algorithm (GA-initial)”. On the other hand our new greedy algorithm considers product sharing with the change of phase.

Proposed Greedy Algorithm (PGA): investigates outputs collectively by changing phase of outputs, one at a time, starting from phase_combination-0 and phase_combination-1. Compares them; if any of the changed phase combination has less # of total product; then keeps that phase and continues searching with changing other outputs phase one at a time. Algorithm steps are as follows:

1) Step: Compare product numbers of phase_combination-0 and phase_combination-1. Set Reference_Number as minimum product number and Save the phase combination as Reference_Phase,

2) Step: Create two sets of candidate phase combination; set0: which has only one phase-1 and set1: which has only one phase-0.

3) Step: Find the phase combinations which yields minimum number of products for the function in set0 and set1.

4) Step: Decide; the phase combination found in step-2 is in set0 or set1. If it is in set0 further changes will be to 1, and vice versa.

5) Step: Decide; if the product number of the phase combination is less than Reference_Number. If Yes, make that number of product as Reference_Number and Save the phase combination as Reference_Phase; If NO, Jump to step-7

6) Step: According to decision in step-3, Create a set includes phase combinations by changing phases of outputs other than the one which has been already changed in Reference_Phase compared to phase_combination-0 or phase_combination-1.

7) Step: While Reference_Number is higher or equal to found in step-5, repeat step-4

8) Step: Finalize phase combinations as Reference_Phase. For example, if we are given a function which has three outputs. First we check phases 000 and 111 (phase_combination-0 and phase_combination-1). Then calculate product numbers of phase in phase sets; set0: 001,010,100 (changing zeros

### TABLE I

<table>
<thead>
<tr>
<th>Benchmarks</th>
<th>Diode</th>
<th>Memristor</th>
<th>FET</th>
<th>Four-Terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td>rd3</td>
<td>442</td>
<td>560</td>
<td>819</td>
<td>127</td>
</tr>
<tr>
<td>squar5</td>
<td>594</td>
<td>884</td>
<td>900</td>
<td>108</td>
</tr>
<tr>
<td>bw</td>
<td>1900</td>
<td>3300</td>
<td>1824</td>
<td>441</td>
</tr>
<tr>
<td>inc</td>
<td>897</td>
<td>1280</td>
<td>2731</td>
<td>235</td>
</tr>
<tr>
<td>rd73</td>
<td>2210</td>
<td>2620</td>
<td>4638</td>
<td>606</td>
</tr>
<tr>
<td>mshipx</td>
<td>437</td>
<td>600</td>
<td>1334</td>
<td>126</td>
</tr>
<tr>
<td>sqq10</td>
<td>840</td>
<td>1032</td>
<td>1540</td>
<td>165</td>
</tr>
<tr>
<td>ex3p</td>
<td>19823</td>
<td>19956</td>
<td>32664</td>
<td>2664</td>
</tr>
<tr>
<td>rd84</td>
<td>5180</td>
<td>6240</td>
<td>10960</td>
<td>2320</td>
</tr>
<tr>
<td>clip</td>
<td>2875</td>
<td>3528</td>
<td>6256</td>
<td>685</td>
</tr>
<tr>
<td>apex4</td>
<td>16835</td>
<td>25480</td>
<td>72335</td>
<td>7308</td>
</tr>
<tr>
<td>sao2</td>
<td>1488</td>
<td>1764</td>
<td>3672</td>
<td>476</td>
</tr>
<tr>
<td>ex1010</td>
<td>8820</td>
<td>11800</td>
<td>57960</td>
<td>5958</td>
</tr>
</tbody>
</table>

** Time exceeds 600 seconds

### TABLE II

<table>
<thead>
<tr>
<th>Benchmarks</th>
<th>Basic [29]</th>
<th>GA-initial [28]</th>
<th>PGA</th>
<th>Optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>rd3</td>
<td>560</td>
<td>416</td>
<td>416</td>
<td>416</td>
</tr>
<tr>
<td>squar5</td>
<td>884</td>
<td>858</td>
<td>832</td>
<td>780</td>
</tr>
<tr>
<td>bw</td>
<td>1280</td>
<td>1280</td>
<td>1280</td>
<td>1280</td>
</tr>
<tr>
<td>inc</td>
<td>2620</td>
<td>2620</td>
<td>1940</td>
<td>1940</td>
</tr>
<tr>
<td>mshipx</td>
<td>600</td>
<td>700</td>
<td>600</td>
<td>600</td>
</tr>
<tr>
<td>sqq10</td>
<td>1032</td>
<td>648</td>
<td>548</td>
<td>648</td>
</tr>
<tr>
<td>ex3p</td>
<td>19564</td>
<td>19312</td>
<td>19312</td>
<td>**</td>
</tr>
<tr>
<td>rd84</td>
<td>6240</td>
<td>6072</td>
<td>4584</td>
<td>4584</td>
</tr>
<tr>
<td>clip</td>
<td>3528</td>
<td>3388</td>
<td>3332</td>
<td>3332</td>
</tr>
<tr>
<td>sao2</td>
<td>1764</td>
<td>1176</td>
<td>1176</td>
<td>1176</td>
</tr>
<tr>
<td>ex1010</td>
<td>11800</td>
<td>11800</td>
<td>11800</td>
<td>11800</td>
</tr>
<tr>
<td>ala4</td>
<td>25696</td>
<td>16544</td>
<td>16544</td>
<td>16192</td>
</tr>
<tr>
<td>b12</td>
<td>2544</td>
<td>1872</td>
<td>1776</td>
<td>1776</td>
</tr>
<tr>
<td>table5</td>
<td>11136</td>
<td>11136</td>
<td>11136</td>
<td>**</td>
</tr>
<tr>
<td>vg2</td>
<td>7854</td>
<td>7854</td>
<td>7854</td>
<td>7854</td>
</tr>
</tbody>
</table>

Transactions on Nanotechnology
IEEE TRANSACTIONS ON NANOTECHNOLOGY (TNANO) SPECIAL SECTION/ISSUE ON IEEE / ACM INTERNATIONAL SYMPOSIUM ON NANOARCH 2018
to one), and set 1: 111,110,101,011 (Changing ones to zero) (Note that, we change one at a times). Lets say we compared product numbers and found that phase 110 yields the least number product. Then we continue from the phase 110, with changing ones to zero. Means, we check 100,010 and compare the results with result of phase 110. If there is one which yields less product number chose it, or chose the phase 110 for the final phase combination of the function.

To evaluate the algorithms and compare the results, we use espresso and MATLAB on a 3.20 GHz Intel Core i7 CPU (only single core used) with 4GB memory. Results are shown in Table II. Algorithm performances differs function to function. For six of the functions PGA results less area size than GA-initial. For the rest they give same result. GA-initial initials (clip and sao2) of the functions GA-initial results less area size than PGA. For the rest they give same result. GA-initials major advantages is that it has almost no time cost. It can be used for any case.

III. DEFECT/FAULT TOLERANCE

In this section, we will investigate defects or faults with categorizing them as permanent (naming defects) and transient (naming faults). As mentioned in Section I, crossbars tend to be fabricated with defects. Also, particular transient faults can occur in the field. Defect tolerance basically means finding defect-free region or crosspoint which can still be employed during logic synthesizing. On the other hand, faults can only be tolerated by redundancy, since they occur transiently. Yet sensitivity analysis can be made for both types. Defect model can be found in Figure 4 demonstrating stuck-at-0 (open) and stuck-at-1 (close). Their features can be summarized as:

- **Permanent Faults** occur mostly in fabrication and are tolerated in post-fabrication by redundancy and reconfigurability (mapping).
- **Transient Faults** occur in field and are tolerated in field by only redundancy

A. Defect Tolerance for Diode, Memristor and FET

Defect tolerance is achieved by realizing a target logic functions on a defective crossbar using row and column permutations. This problem is considered as NP-complete [20]. For the worst-case, \(N!M!\) permutations are required to find a successful mapping for \(N \times M\) crossbar. Algorithms in the literature use defect-unaware or defect-aware approach.

Defect-unaware algorithms aim to find the largest possible \(k \times k\) defect-free sub-crossbar from a defective \(N \times N\) crossbar where \(k \leq N\). The algorithms are inefficient for high fault rates - obtained \(k\) values are much smaller than \(N\). In this regard, defect-aware algorithms perform much more satisfactorily [25]. Detailed analysis of both approaches can be found in [26].

Defect-aware considers the defect characteristics (stuck-at-0 or stuck-at-1), then decide which switch to employ during the mapping. In our previous work [25], we have proposed an efficient heuristic algorithms which aims to match defected crossbar and the function solution crossbar. For this, it defines crossbars as matrix. Therefore it can perform sorting, matching and backtracking steps efficiently. It makes repetition for a limit of permutation. This controls heuristic feature of the algorithm.

B. Defect Tolerance for Four-terminal

Four-terminal defect tolerance demands a different approach than the methods we have covered so far. For this reason, we present a novel method, which is firstly introduced in this paper. The proposed method utilizes a priori sensitivity analysis of crossbar (lattice) to specify critical switches, and strengthens them with proposed mitigation factors. The same naming conventions are applicable, regarding defects which are categorized as stuck-at-0 (SA0) and stuck-at-1 (SA1). Furthermore, we describe a new defect that can be considered for crossbars (lattices), that consists in changing a switching literal in a cell \(c_{i,j}\) of the lattice with a literal that is in a adjacent cell (i.e., \(c_{i-1,j}, c_{i+1,j}, c_{i,j-1}, \text{or } c_{i,j+1}\)). We denote this fault model as Adjacent Cellular Fault Model (ACFM). In addition, we follow the same terminology adopted in [3] and [12] by addressing crossbar as lattice and switch as cell to be consistent and emphasize the distinction of four-terminal approach. Finally, it should be noted that as opposed the previous sections, we provide a more detailed explanation due to original technical contribution presented in this section.

Defect Injection Methodology

This section introduces the algorithms for the defect injection in the SAFM and in the ACFM.

The two procedures are repeated for each cell of the lattice. Once the "defective" lattice is obtained, both the algorithms generate all the possible \(2^n\) inputs (where \(n\) is the number of variables). For each input, the simulation algorithms compare the given output with the correct one. The two defect injection algorithms (for the SAFM and the ACFM) differ in the calculation of number of defective outputs in the two fault models, as explained in the following subsections.

Fault Model 1: Stuck at Faults (SAFs)

The sensitivity of the decomposition algorithm on a given crossbar is analyzed face to SA0 and SA1 being the model widely adopted today for memristive type of crossbar. As there is no consensus today on the fault distribution, we have, first of all, chosen a uniform distribution for each type of SA0 and SA1 [8]. The fault density considered may be up to 10% of the crossbar, all faults being independent, as reported in [8]. The fault injection in the above lattices is performed substituting

![Nano-Crossbar Array](image-url)
a single cell with an always stuck-at 1 (SA1) or stack-at-0 (SA0) cell.

Let $E^0_{i,j}$ (resp., $E^1_{i,j}$), with $1 \leq i \leq r$, $1 \leq j \leq s$, be the number of defective outputs with a SA0 (resp., SA1) in the cell $(i,j)$ of the given lattice. Note that $0 \leq \{E^0_{i,j}, E^1_{i,j}\} \leq 2^n$. Moreover, when $E^0_{i,j}$ (resp., $E^1_{i,j}$) is equal to 0 we have that, for any possible input, the lattice output is never changed by the SAF in the cell $c_{i,j}$. In this case, we call the cell $c_{i,j}$ robust w.r.t. SA0 (resp., SA1). Let $R^0$ (resp., $R^1$) be the total number of defective cells w.r.t. SA0 (resp., SA1) in the simulation. Finally, let $E^0 = \sum_{i=1}^{r} \sum_{j=1}^{s} E^0_{i,j}$ (resp., $E^1 = \sum_{i=1}^{r} \sum_{j=1}^{s} E^1_{i,j}$) be the total number of defective outputs with SA0 (resp., SA1) in the simulation. For an example of function $f$ in the simulation. For an example of function $f$ in the simulation. For an example of function $f$ in the simulation.

**Fault Model 2: Adjacent Cellular Faults (ACFs)**

In the classical CFM [11] for CMOS circuits it is assumed that a fault modifies the behavior of exactly one node $v$ in a given circuit $C$ and that the modified behavior is still combinational. This fault for a switching lattice $L$ can be described as follows: a cellular fault in $L$ is a tuple $(c_{i,j}, l_e, l_f)$, where $c_{i,j}$ is the cell of the lattice $L$ (i.e., fault location), $l_e$ is the correct controlling literal in $c_{i,j}$, and $l_f$ (≠ $l_e$) is the faulty controlling literal. We denote adjacent cellular fault a cellular fault where the faulty literal $l_f$ is in an adjacent cell. More precisely:

**Definition 1**: Let $l_{h,k}$ be the literal in the cell $c_{h,k}$ of a lattice $L$. We have that:

1. A Left Adjacent Cellular Fault (L-ACF) is the cellular fault $(c_{i,j}, l_{i,j}, l_{i,j-1})$.
2. A Right Adjacent Cellular Fault (R-ACF) is the cellular fault $(c_{i,j}, l_{i,j}, l_{i,j+1})$.
3. A Bottom Adjacent Cellular Fault (B-ACF) is the cellular fault $(c_{i,j}, l_{i,j}, l_{i-1,j})$.
4. A Top Adjacent Cellular Fault (T-ACF) is the cellular fault $(c_{i,j}, l_{i,j}, l_{i+1,j})$.

Let $E^0_{i,j}$ (resp., $E^0_{i,j}$, $E^1_{i,j}$, and $E^2_{i,j}$), with $1 \leq i \leq r$, $1 \leq j \leq s$, be the number of defective outputs with a L-ACF (resp., R-ACF, B-ACF, and T-ACF) in the cell $c_{i,j}$ of the given lattice. Let $R^0$ (with $a \in \{L, R, B, T\}$) be the total number of robust cells w.r.t. $a$-ACF in the lattice. Finally, let $E^0 = \sum_{i=1}^{r} \sum_{j=1}^{s} E^0_{i,j}$ be the total number of defective outputs with $a$-ACF in the simulation.

**Metrics used for Sensitivity Analysis**

In order to evaluate the sensitivity of a lattice to SAF and ACNF defects, we propose two metrics. The first one measures the average number of defective outputs considering sensitive cells to defects only. The second one measures the average number of defective outputs in the entire lattice. Note that the total number of cells is the area of the lattice (i.e., $r \cdot s$), the number of non-robust cells for SA0 (resp., SA1) is $r \cdot s - R^0$ (resp., $r \cdot s - R^1$), and $2^n$ is the total number of inputs. Moreover, the number of non-robust cells for $a$-ACF (with $a \in \{L, R, B, T\}$) is $r \cdot s - R^0$.

The sensitivity of defective cells is the total number of inputs that give an uncorrected output ($E^i$, with $i \in \{0, 1, L, R, B, T\}$) divided by the total number of inputs ($2^n$), for each non-robust cell. The metric can be expressed as: $S_{C}^{i} = E^i / (2^n \cdot (r \cdot s - R^i))$ for each fault $i \in \{0, 1, L, R, B, T\}$.

The sensitivity of lattice is the total number of inputs that give an uncorrected output divided by the total number of inputs for each cell: In particular, $S_{L}^{i} = E^i / (2^n \cdot r \cdot s)$, with $i \in \{0, 1, L, R, B, T\}$.

**Benchmarks and Simulations**

The defect simulations have been run on a machine with two AMD Opteron 4274HE for a total of 16 CPUs at 2.5 GHz and 128 GByte of main memory, running Linux CentOS 7. The benchmarks functions are expressed in PLA form and are taken from a subset of LGSynth93 [31]. A total of about 580 functions were considered, and each output of a function is implemented as a separate Boolean function.

The software used for simulations is written in C++. We used ESPRESSO to implement the method described in [3], and a collection of Python scripts for computing minimum-area lattices by transformation to a series of SAT problems, to simulate the results reported in [12]. Each SAT execution is stopped after ten minutes.

In Table III, we report a sample of benchmark functions and their sensitivity values, according to the metrics presented before. In particular, Table III refers to lattice synthesized as described in [3] and [12]. The benchmarks that are present in Table III with dual method were stopped after ten minutes of SAT execution, but that was not the case for the rest.

More precisely, in both methods, the first column reports the name and the number of the considered output of each function. The following columns report dimension ($r \times s$) required for the synthesis of a given function according to each decomposition method, and the number of input variables $n$. Columns from 4 to 11 refers to Stuck At fault model, columns from 12 to 27 to Adjacent Cellular fault model showing the total number of errors $E$, the Sensitivity of defective cells $S_C$, the Sensitivity of lattice $S_L$, and the percentage of robust cells $\%R / r \times s$.

Table IV describes the overall results for the benchmarks we have considered. It also shows the average values for the considered metrics. We can note that the percentage of cells that are considered robust according to our metrics is higher in the first approach [3]. This is due to the more constrained structure of the lattices produced by the first synthesis method. Indeed, the method proposed in [3] computes a lattice for $f$
Mitigation by Defect Avoidance

From the above results, it can be seen that the two analyzed mapping algorithms show different sensitivities of the output of a given function. As a matter of fact, the more restrictive an algorithm is in terms of area (closer to optimal solution), the higher the defect sensitivity of the output to cell defect. It is mandatory to include the mapping algorithm defect-avoidance heuristics, but hardware-level defect tolerant scheme may also be necessary, especially in the case of high defect densities. Redundancy schemes that can be used are inspired from memory testing and repairing structures published in the early 2000 [16]. They can be used at a column level, or block level, as the basic computation unit of memristor array is not the memristive cell as in classical CMOS based memories, but an entire column. Therefore, several redundant columns can be added to the initial design to be able to detect and replace the memristive RAM affected columns but also to overcome the potential SAD that affect the redundant parts. Mapping of logic functions on crossbar arrays are thus divided into two main phases: mapping phase to write the parameters of functions in the memristive array and a read operation to check the results from the crossbar. The objective here is to identify at the writing time, if common literals and other multiple-choice literals of the function have to be mapped on highly critical cells.

Mitigation for Stuck at Faults:

In order to mitigate the sensitivity of a lattice to SAD, we propose the following possible strategy applied to the synthesis method proposed in [3] which has been proven as less sensitive to SAD impact on the output functions: (1) For a given mapped function, if a potential SA0, SA1 defect affects a robust cell identified by the defect injection campaign, the lattice still computes the correct output, thus we do not need any mitigation with defect tolerant design. (2) However, if an injected defect occurs in a multiple-choice cell, if a different literal can be chosen to make the cell robust, we change the literal with the new one. (3) Otherwise, if the injected SA0 defect is proven as being critical for the output value, the column that contains that defective cell has to be replaced by spare columns. In case of an SA1 the row that contains the defective cell has to be replaced by a spare row. Note that, in this case, the output still provides a correct function \( f \) from top to bottom, but the function from left to right could be changed and become a function which will not be dual of \( f \) anymore.

As an example, consider the lattice synthesized in Figure 6(a) with \[ f = x_4x_5x_6 + x_4x_5x_6x_7 + x_4x_5x_6x_7 + x_4x_5x_6x_7 \] by using synthesis method presented in [3]. The example shows one case of mitigation of 3 independent SAD affecting the lattice implementing the function, yielding an approximate 10% defects. In Figure 6, SA1 cells are marked in blue and SA0 cells a remarked in red.

To avoid output errors due to these SAD we have used the following strategy:

1. Identify robust cells for a given function mapping. Example: the defect in first row, fifth column is non-influential on the value of the output (robust cell), the sensitivity map, obtained through defect injection campaign, shows that this cell is not sensitive to SA1 for the mapped function.

2. Identify the swapping of literals during synthesis process on a column of an high sensitive cell. Example: the defect
Mitigation for Adjacent Cellular Faults:

We can note that adjacent cells containing the same controlling literal are robust to adjacent cellular faults. For this reason, in this fault model, it is of huge importance to maximize the number of adjacent cells with the same controlling literal. For this purpose, the columns and rows, obtained by the synthesis method proposed in [3], can be permuted. The properties of the lattice synthesized with [3] guarantee that row and column permutations do not effect the Boolean function computed by the switching lattice.

C. Transient Fault Tolerance

Regarding transient faults, there are two approaches: redundancy based and manipulation of logic function to obtain a more fault tolerant design. Nevertheless, it should be noted that since nano-crossbars are in the early stage of development, there is a lack of in-field data regarding transient faults. For this reason, second approach towards transient fault tolerance is methodology independent and rather focuses on the intrinsic features of the given logic functions. As mentioned, transient faults can be tolerated with redundancy. Inserting redundant components can be constructed with adding extra rows and/or columns as shown in [18] [4].

Furthermore, provided certain conditions, particular logic functions are inherently tolerant to transient faults limited to certain switches of crossbar as show in [25]. Mentioned inherent tolerance capability varies depending on the design. For example, consider a design having low logic inclusion ratio (IR), meaning less number of crosspoints of a crossbar are used, that means this design is more tolerant to stuck-at-0 faults. Logic synthesis (design) should be made with regard to this, since there are multiple design solution.

Similarly, if the target function can be realized with high IR, then a technology, the one which tends to have stuck-at-1 faults, should be preferred. For example function \( f = x_1x_2 + x_2x_3 + x_3x_4 \) can also be written as \( f' = x_1x_2\bar{x}_3 + x_2x_3 + \bar{x}_3x_4 = f \), therefore IR can be increased.

Fault sensitivity analysis can be made using Monte Carlo analysis, yet it is costly. Since we know the dynamics of the fault tolerance we can calculate sensitivity (fault tolerance performance) directly with algebraic equations. Equation parameters consists of crossbar dimensions, inclusion ratio, fault occurrence possibility and number of tolerable crosspoint. For further information please refer to [25].

IV. PERFORMANCE OPTIMIZATION

Area performance of crossbar is previously investigated in Section II. Here, we will analyze the delay and power depending on number of products and/or literals. In our previous study [28], we have concluded this analysis for memristive crossbars.

Here we have used an approach of multiplying resistive and capacitive loads for delay. This helps us to calculate maximum frequency with \( 1/delay \). In the delay analysis for diode, we see that “number of columns” and “load resistor” dominate the capacitive load and resistive load respectively. For memristive crossbars, it is proportional to constant \( 7 \). Considering FET, we need to calculate resistive loads and capacitive loads from the longest path for the worst-case scenario. Lastly, four-terminal \( delay \) is directly proportional to constant \( 7 \). Considering FET, longest path on the lattice can be calculated as explained in [3]:

\[
L_{\text{long path}} = \begin{cases} 
R \quad & R \leq 2 \cup C \leq 1 \\
3\frac{R-2}{2}C + \frac{2+(-1)^R+(-1)^C}{2} \quad & R > 2 \cap C > 1 
\end{cases}
\]

where \( R \): number of rows, \( C \): number of columns, \( L_{\text{long path}} \): length of the longest path on the lattice.

Delay Analysis:

- Diode: \( \propto (\text{load resistor}) \times (\# \text{ of literals in } f) \)
- Memristor: \( \propto \text{constant} : 7 \) [28]
### Performance Optimization

#### Section IV

- **FET**: \( \propto (\text{degree of the largest product in } f \text{ and } f^D) \times ((\# \text{ of products in } f) + (\# \text{ of products in } f^D) + (\# \text{ of literals in } f) \)
- **Four-terminal**: \( \propto L_{\text{long path}} \)

Power consumption is total energy used in unit time. Therefore, we estimate consumed total energy in a period, afterwards we divided it with delay (delay is proportional with one period maximum frequency). Because we have estimated circuit will use in maximum frequency.

For power of diode crossbar is dominated by static power. Thus, it is inversely proportional only with load resistor. Load resistor can be considered as 10 times of diode’s inner resistance. Memristor’s total energy consumption is proportional to total memristor count, because it is assumed that all memristors are changes their states one time only [28]. Total energy consumption of FET crossbars is directly proportional to the capacitance of output node and nodes related to the output node. For four-terminal, we assume all crosspoints consume energy at worst-case scenario. Finally, we divided energy consumption by the delay to find the complexity of power consumption.

#### Power Analysis

- **Diode**: \( \propto 1 / (\text{load_resistor}) \)
- **Memristor**: \( \propto (\text{total memristor count in crossbar})/7 \) [28] [24]
- **FET**: \( \propto 1 / (\text{degree of the largest product in } f \text{ and } f^D) \) (@ maximum frequency)
- **Four-terminal**: \( \propto ((\# \text{ of products in rows}) \times (\# \text{ of products in columns})) / L_{\text{long path}} \)

### V. A Case Study for Benchmark Function Squar5

In this section, we demonstrate the whole process with an example. We have chosen the benchmark function Squar5 which has 5 inputs, 25 products and 8 outputs in pla format. Diagram of the process is given in Figure V.

First, we need to evaluate different technologies such as diode, FET, memristor or four-terminal. Using logic synthesis, we can generate function descriptions in crossbar (lattice) form and obtain area sizes utilizing equations in Section II. Results are as follows:

<table>
<thead>
<tr>
<th>Technology</th>
<th>Choice</th>
<th>Area Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diode</td>
<td>FET</td>
<td>594</td>
</tr>
<tr>
<td>Memristor</td>
<td>Memristor</td>
<td>832</td>
</tr>
<tr>
<td>Four-terminal</td>
<td>4-Terminal</td>
<td>160</td>
</tr>
</tbody>
</table>

#### Defect/Fault Tolerance

- **Diode**: \( \# \text{ of products of all } f_{oi} + n ) \times ( (\# \text{ of literals in } f) + n ) = (25 + 8) \times (10 + 8) = 594. \)
- **Memristor**: \( ((\# \text{ of products of all } f_{oi}) + n ) \times ( (\# \text{ of literals in } f) + 2n ) = (25 + 8) \times (10 + 16) = 858. \) However, our proposed greedy algorithm PGA decreases the size to 832.
- **FET**: \( (\# \text{ of literals in } f + n ) \times ( (\# \text{ of products of all } f_{oi}) + (\# \text{ of products of all } f_{oi}^D) ) = (10 + 8) \times (25 + 25) = 900. \) Coincidently, its dual also has the same number of products.
- **Four-terminal**: \( (\# \text{ of products in } f \times (n + 1) ) \times ( (\# \text{ of products of all } f_{oi}) + (\# \text{ of products of all } f_{oi}^D) ) = 5 \times (25 + 8) = 160. \)

First term is chosen according to the product which has the maximum number of literals.

Secondly, we use function descriptions and defect map of a crossbar. Applying the defect tolerant logic mapping methods in Section III, it is possible to measure defect/fault performance. Key point is that diode, FET and memristor have a rich literature of methods for reaching 100% defect tolerance for defect rates up to 10% [26] [28]; nevertheless four-terminal is at its infancy in terms of defect tolerance. In this paper, defect performance of only single output functions are studied and we are planning to extend the work into multi output functions as well in the future.

Lastly, we conduct a performance optimization specific to technology dependent delay and power parameters of crossbars. Since related equations are presented in immediate section, only result are given:

- **Diode**: Delay is \( R \times 10 \) (Load_resistor is shown with \( R \)). Power is \( 1/R \). If assume \( R \) is 10 times than a diode’s inner resistance, then Delay is 100 and Power is 1/10.
- **Memristor**: Delay is constant and 7. Power is 133/7 \( \approx 19 \). The number 133 denotes the number of memristors used in the crossbar.
- **FET**: Delay is 300 and power is 1/5. Including the next four-terminal, FET has the largest delay.
- **Four-terminal**: The longest path is 73, so is the delay. Power is 160/73 \( \approx 2 \).

Note that, values only lights the complexity of delay and power, they are not real elapsed time and power consumption. Here we assume, they all fabricated with same technology. For
instance, in order to calculate the delay of diode crossbars, we need to know value of $R$ and the source voltage. To provide an overall evaluation, four-terminal is the most advantageous choice in terms of area size. However, defect tolerance is poor especially regarding the complexity and computation power needed to conduct experiments for four-terminal. Delay of memristive crossbars is predictable, it doesn’t depend on function. Therefore, it could be considered as the most advantageous choice in terms of delay. For power consumption, diode seems to have least consumption, yet it is static power also depends source voltage. On the other hand, FET crossbars has only dynamic power consumption with complementary architecture.

VI. CONCLUSION AND DISCUSSION

In this study, we present a synthesis methodology for crossbar arrays having crosspoints working as FET, diode/resistive/memristive, or four-terminal switch based devices. We cover “logic synthesis”, “defect/fault tolerance”, and “area-power-delay performance optimization” steps. Presented synthesis methodology provides optimization algorithms for each step of the process as well as their relations and trade-offs.

As a future work, a fully automated software tool will be developed. It takes the target function, the used technology, and the performance specifications (area, delay and/or power consumption) as input and return the optimized crossbar-arrays structure as an output. This software tool to be developed will benefit from the optimization algorithms introduced in this study. It will also make the technology mapping using technology data.

Acknowledgement: We thank Levent Aksoy for his helps to synthesize multi-output functions with four-terminal switch based arrays.

REFERENCES

Integrated Synthesis Methodology for Crossbar Arrays*

1st M. Ceylan Morgul
3rd Onur Tunali
10th Mustafa Altun
Istanbul Technical University
Istanbul, Turkey
{morgul, onur.tunali, altunmus}@itu.edu.tr

2nd Luca Frontini
5th Valentina Ciriani
Università degli Studi di Milano
Milan, Italy
{luca.frontini, valentina.ciriani}@unimi.it

4th E. Ioana Vatajelu
6th Lorena Anghel
TIMA laboratory
Grenoble-Alpes University
Grenoble, France
{ioana.vatajelu, lorena.anghel}@imag.fr

7th Csaba Andras Moritz
University of Massachusetts, Amherst
Massachusetts, USA
andras@ecs.umass.edu

8th Mircea R. Stan
University of Virginia
Charlottesville, Virginia, USA
mircea@virginia.edu

9th Dan Alexandrescu
IROC Technologies
Grenoble, France
dan.alexandrescu@iroctech.com

ABSTRACT
Nano-crossbar arrays have emerged as area and power efficient structures with an aim of achieving high performance computing beyond the limits of current CMOS. Due to the stochastic nature of nano-fabrication, nano arrays show different properties both in structural and physical device levels compared to conventional technologies. Mentioned factors introduce random characteristics that need to be carefully considered by synthesis process. For instance, a competent synthesis methodology must consider basic technology preference for switching elements, defect or fault rates of the given nano switching array and the variation values as well as their effects on performance metrics including power, delay, and area. Presented synthesis methodology in this study comprehensively covers the all specified factors and provides optimization algorithms for each step of the process.

CCS CONCEPTS
• Hardware → Emerging architectures; Emerging tools and methodologies;
• Hardware → Crosspoint architectures; Crossbar Arrays

KEYWORDS
Crossbar Arrays, Logic Synthesis, Defect Tolerance, Fault Tolerance, Performance Optimization, Memristor Arrays

1 INTRODUCTION
Nano-crossbars are emerged to be an alternative technology besides CMOS [25]. They are fabricated with relatively cheap bottom-up nano-fabrication techniques rather than using purely lithography based conventional production. Due to the novel manufacturing techniques, fabrics yield to be in regular and dense form [7]. Because of their structure and technology, they are area and power efficient [1].

Currently, computing is achieved with crosspoints behaving like switches, either as two-terminal or four-terminal. This is illustrated in Figure 1. Depending on the used technology, a two-terminal switch behaves either as a diode [11], a resistive/memristive switch [18], or a field effect transistor (FET) [19]. Diode and resistive switches correspond to the crosspoint structure in Figure 1(a); here, the switch is controlled by the voltage difference between the terminals. Figure 1(b) shows a FET based switch; here, the red line represents the controlling input. This is a unique opportunity that allows us to integrate well developed conventional circuit design techniques into nano-crossbar arrays. Finally, a novel four-terminal switch is given in Figure 1(c). The controlling input, not shown in the figure, has a separate physical formation from the crossbar that is thoroughly explained for different technologies in [2].

To illustrate their computing approaches, we show examples for the implementation of $f_{XOR} = x_1 \overline{x_2} + \overline{x_1}x_2$ in Figure 2. Logic

---

*This work is part of a project that has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement #691178, as well as supported by the TUBITAK-Career project #113E760.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

NANOARCH ’18, July 17–19, 2018, Athens, Greece
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-5815-6/18/07...
https://doi.org/10.1145/3232195.3232211
synthesis models for diode and memristor based crossbars are very much PLA-like as can be seen in Figure 2(a) and 2(b). Memristor based crossbars have one major difference that logic computation is made through several states/loops (for further information, check [24]). For FET based crossbars, each product of the function or function’s dual is realized by a separate column, as seen in Figure 2(c). Each inputs is assigned to a row to control the FETs on the corresponding row. Another type is a four-terminal based crossbar; here every crosspoint performs switching on all four directions. Crosspoints’ control lines are not shown in Figure 2(d), yet detailed explanation of control lines can be found in [2].

Regarding emerging technologies and nano-fabrication, fault rates are much higher for nano-crossbars, as expected, compared to those of conventional CMOS circuits [8]. Therefore, during logic synthesis, consideration of faults and defects is mandatory. This applies for the integration of both diode, FET based or novel 4-terminal based logic synthesis methodologies. For this reason, researchers focus on challenges including defect and variance tolerances [21] [10]. Defect and variance tolerant approaches are closely related to logic realization and performance optimization, respectively.

Taking mentioned issues into account, we have developed a complete integration methodology for logic synthesis, defect tolerance and performance optimization with variance tolerance. This methodology is designed as a step-by-step guide to combine modular research approaches into an entire production pipeline. The overview of proposed integrated methodology is explained in Section 2. Following sections, we have demonstrated and reviewed the current methods with the exception of Section 4.2. Since four-terminal design parts from the rest in terms of defect tolerance, we present a preliminary and novel approach for defect tolerance of four-terminal crossbars which is anticipated as an initial step in further research.

2 PROPOSED INTEGRATION METHODOLOGY

As briefly explained in the previous section, nano-fabrication delivers switching nano-crossbar arrays with structures or individual components having varied properties. Mentioned factors introduce random characteristics of which need to be carefully considered by synthesis process. For instance, a competent synthesis methodology must consider basic technology preference for switching elements, defect or fault rate of the given nano-crossbar and the variation values. Presented synthesis methodology in this study comprehensively covers the all specified factors and provides optimization algorithms for each step of the process. A schematic summary demonstrating every step of the method with annotation showing the certain research tasks is given in Figure 3.
First step of the synthesis process of a nano-crossbar involves the decision of switching technology which will be explained elaborately in Section 3. Main purpose is to determine which of the diode/memristor, FET, or four-terminal based components are to be used. This step is one of the most important procedures determining the size of the nano-crossbar. Production with diode/memristor based technologies as well as with FET are explained and then logic synthesis design with four-terminal based switches is given.

Second step of the synthesis process of a nano-crossbar covers the permanent faults (defects forming in the course of fabrication) and the tolerance aspects, which will be described in Section 4. Main purpose is to obtain a valid realization of a given logic function using two distinct approaches titled as defect-aware and defect-unaware. First method employs faults existing in nano-crossbar during the realization of logic function hence the name aware. Second method avoids the faults by attempting to find a fault-free region of nano-crossbar at the beginning so realization of given logic function is straightforward at the end.

Third step of the synthesis process of a nano-crossbar covers the variation minimization, which will be explained in Section 5. Main purpose is to minimize the overall delay by considering the individual variation values of components used in logic synthesis. Heuristic algorithms in the literature produce results very close to theoretical lower bound.

Delay minimization stage of the fourth step can be modified to include power optimization and fault tolerance as well. An objective function added to the variation tolerant algorithms guides the process to minimize both delay and power values of a nano-crossbar. Nevertheless, since this is a multi-level optimization, performance of the algorithm diminishes to an extent. In order to execute a parametric optimization, trade-offs must be decided considering technology dependent features and other conditions for the objective function. Furthermore, fault tolerance mechanism can be appended to the delay minimization process. It is possible to avoid faulty elements by assigning infinite variation values to them. Since the heuristic is based on minimizing the variation values of the nano-crossbar, fault avoidance is inherently established.

Final step of the synthesis process of a nano-crossbar involves the analysis of transient faults. Main purpose is to determine the effect of transient faults to the operational capacity of nano-crossbar and calculate fault tolerance performance.

### 3 LOGIC SYNTHESIS

At the beginning of logic synthesis process, crossbar technology must be determined based on certain fundamental criteria:

- Crossbar size limits (area) and Function size
- Fabrication Complexity
- Function output number (Multi or single output realization)
- Power and Delay Specifications
- Application specification (memory based etc.)

Decision should be made on the importance of the listed items; this could change depending on the application. For example, if an application has also memory unit, then it will be smart to choose memristor technology for logic unit. Since memristor could be used on memory unit as well, then they can interact more smoothly and the same fabrication technique could be used for both.

On the other hand, realization of a function with diode or memristor based crossbar requires less area than FET based option. However FET has better power performance over other technologies. In addition to all, four-terminal based crossbar performs better results on most of the function in terms of area [13].

In this section, we survey logic synthesis step of the integration methodology by considering only area size of the crossbar arrays. Array size formulations, which are given below, could be a guideline to this.

Array size formulations for single output function $f$ (where $f$’s dual is $f^D$):

- for **Diode**: $(\text{number of products in } f + 1) \times ((\text{number of literals in } f) + 1)$
- for **Memristor**: $(\text{number of products in } f + 2) \times ((\text{number of literals in } f) + 2)$
- for **FET**: $(\text{number of literals in } f) \times ((\text{number of products in } f) + (\text{number of products in } f^D))$
- for **4-terminal**: $(\text{number of products in } f^D) \times (\text{number of products in } f)$

For the single and multi output function realization, synthesis methodology for FET crossbar does not allow us to produce multi-level logic synthesis, only two-level approach can be used [20]. However, multi-level logic synthesis approach is applicable for diode and memristive crossbars [23]. Therefore, area optimization still demands further research for FET systems. Furthermore, as mentioned in Section 1, logic synthesis on diode and memristive crossbars is similar to PLA like synthesis. So the same approaches (which for the PLA) are applicable such as product sharing, phase changing etc.

Logic synthesis on four-terminal crossbars (lattices) is relatively a new method and technology. As shown in [2], Altun presented a useful logic synthesis technique for four-terminal crossbars (lattices). Yet mentioned method does not warrant the fact that produced lattice has optimal solution in terms of area. Therefore, new specific logic synthesis methodologies are needed to be presented. As shown in [9] and [13], optimal synthesis methodologies are provided. In addition, there are decomposition based techniques such as XOR based [12] [6], p-circuit [4] and dimension reducibility [5] decompositions as well.

### 4 DEFECT/FAULT TOLERANCE

In this section, we will investigate defects or faults with categorizing them as permanent (naming defects) and transient (naming faults). As mentioned in Section 1, crossbars tend to be fabricated with defects. Also, particular transient faults can occur in the field. Defect tolerance basically means finding defect-free region or cross-point which can still be employed during logic synthesizing. On the other hand, faults can only be tolerated by redundancy, since they happens transiently. Yet sensitivity analysis can be made for both types. Defect model can be found in Figure 4 demonstrating stuck-at-0 (open) and stuck-at-1 (close). Their features can be summarized as:

- **Permanent Faults** occur mostly in fabrication and are tolerated in post-fabrication by redundancy and reconfigurability (mapping).
4.1 Defect Tolerance for Diode, Memristor and FET

Defect tolerance is achieved by realizing a target logic functions on a defective crossbar using row and column permutations. This problem is considered as NP-complete [17]. For the worst-case, N!M! permutations are required to find a successful mapping for N × M crossbar. Algorithms in the literature use defect-unaware or defect-aware approach.

Defect-unaware algorithms aim to find the largest possible k × k defect-free sub-crossbar from a defective N × N crossbar where k ≤ N [27]. The algorithms are inefficient for high fault rates -obtained k values are much smaller than N [27]. In this regard, defect-aware algorithms perform much more satisfactorily [21]. We have performed detailed analysis of algorithms in [22].

Defect-aware considers the defect characteristics (stuck-at-0 or stuck-at-1), then decide which switch to employ during the mapping. In our previous work [21], we have proposed an efficient heuristic algorithms which aims to match defected crossbar and the function solution crossbar. For this, it defines crossbars as matrix. Therefore it can perform sorting, matching and backtracking steps efficiently. It makes repetition for a limit of permutation. This controls heuristic feature of the algorithm.

4.2 Defect Tolerance for Four-terminal

Four-terminal defect tolerance demands a different approach than the methods we have covered so far. For this reason, we present a novel method, which is firstly introduced in this paper. The Method utilizes a prior sensitivity analysis of crossbar to specify critical switches, and strengthens them with proposed mitigation factors. The same naming conventions are applicable, regarding defects which are categorized as stuck-at-0 (SA0) and stuck-at-1 (SA1). In addition, we follow the same terminology adopted in [2] and [9] by addressing crossbar as lattice and switch as cell to be consistent and emphasize the distinction of four-terminal approach. Finally, it should be noted that as opposed the previous sections, we provide a more detailed explanation due to original technical contribution presented in this section.

4.2.1 Defect Injection Methodology. We perform a defect injection with uniform distribution to lattice reaching defect densities up to 10%. Every cell (a four-terminal switch) is presumed to have only SA0 or SA1. Once the "defective" lattice is obtained, the algorithm generates all the possible 2^n inputs (where n is the number of variables). For each input, the simulation algorithm compares the given output with the correct one. Let E_0^n (resp., E_1^n), with 1 ≤ i ≤ r, 1 ≤ j ≤ s, be the number of defective outputs with a SA0 (resp., SA1) in the cell (i, j) of the given lattice. Note that 0 ≤ |E_0^n|, E_1^n| ≤ 2^n. Moreover, when E_0^n (resp., E_1^n) is equal to 0 we have that, for any possible input, the lattice output is never changed by the SAF in the cell (i, j). In this case, we call the cell (i, j) robust w.r.t SA0 (resp., SA1). Let R^0 (resp., R^1) be the total number of robust cells w.r.t. SA0 (resp., SA1) in the lattice. Finally, let E_0 = Σ_{i=1}^{r} Σ_{j=1}^{s} E_0^{ij} (resp., E_1 = Σ_{i=1}^{r} Σ_{j=1}^{s} E_1^{ij}) be the total number of defective output with SA0 (resp., SA1) in the simulation. For an example of function f = x_4 x_5 x_7 + x_4 x_6 x_7 + x_4 x_5 x_6 x_7 + x_4 x_6 x_7 + x_4 x_5 x_7, realized in Figure 5(a) (with the method in [2]), in the Figure 5(b) (resp., 5(c) shows the map containing E_0^{ij} (resp., E_1^{ij}) in each cell.

4.2.2 Metrics used for Sensitivity Analysis. In order to evaluate the sensitivity of a lattice to SA0 and SA1 defects, we propose two metrics. The first one measures the average number of defective outputs considering sensitive cells to SA0 or SA1 only. The second one measures the average number of defective outputs in the entire lattice. Note that the total number of cells is the area of the lattice (i.e., r x s). The number of non-robust cells for SA0 (resp., SA1) is r x s - R^0 (resp., r x s - R^1), and 2^n is the total number of inputs. (1) Sensitivity of defective cells is the total number of outputs that give an uncorrected output (E_0 and E_1) divided by the total number of inputs (2^n), for each non-robust cell (r x s - R^0 or r x s - R^1). In the case of SA0 the metric can be expressed as: S_0 = E_0/(2^n(r x s - R^0)). The same reasoning can be done for SA1 defect sensitivity. (2) Sensitivity of lattice is the total number of inputs that give an uncorrected output divided by the total number of inputs for each cell, in the case of SA0 is: S_L = E_0/(2^n(r x s)). The SA1 case is analogous.

4.2.3 Benchmarks and Simulations. The defect simulations have been run on a machine with two AMD Opteron 4274HE for a total of 16 CPUs at 2.5 GHz and 128 GByte of main memory, running Linux CentOS 7. The benchmarks functions are expressed in PLA form and are taken from a subset of LGSynth93 [26]. A total of about 580 functions were considered, and each output of a function is implemented as a separate Boolean function. The software used for simulations is written in C++. We used ESPRESSO to implement the method described in [2], and a collection of Python scripts for computing minimum-area lattices by transformation to a series of SAT problems, to simulate the results reported in [9]. Each SAT execution is stopped after ten minutes.
Table 1: A sample of benchmark functions synthesized with [2] and [9] approaches and their sensitivity values

<table>
<thead>
<tr>
<th>Name</th>
<th>r × s</th>
<th>n</th>
<th>E₀</th>
<th>S₀₀</th>
<th>S₀₁</th>
<th>L₀</th>
<th>E₁</th>
<th>S₁₀</th>
<th>S₁₁</th>
<th>L₁</th>
<th>%R₀</th>
<th>%R₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>add6(1)</td>
<td>6×6</td>
<td>4</td>
<td>19</td>
<td>0.06</td>
<td>0.03</td>
<td>47%</td>
<td>9</td>
<td>0.06</td>
<td>0.02</td>
<td>75%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>alu2(2)</td>
<td>11×10</td>
<td>8</td>
<td>462</td>
<td>0.03</td>
<td>0.02</td>
<td>35%</td>
<td>121</td>
<td>0.02</td>
<td>0.01</td>
<td>80%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b1(1)</td>
<td>3×6</td>
<td>7</td>
<td>28</td>
<td>0.02</td>
<td>0.01</td>
<td>44%</td>
<td>73</td>
<td>0.03</td>
<td>0.03</td>
<td>6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dc2(0)</td>
<td>4×6</td>
<td>7</td>
<td>117</td>
<td>0.05</td>
<td>0.04</td>
<td>17%</td>
<td>162</td>
<td>0.08</td>
<td>0.05</td>
<td>33%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>exam(5)</td>
<td>6×11</td>
<td>9</td>
<td>1868</td>
<td>0.07</td>
<td>0.06</td>
<td>17%</td>
<td>131</td>
<td>0.02</td>
<td>0.01</td>
<td>74%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e(2)</td>
<td>12×12</td>
<td>5</td>
<td>70</td>
<td>0.03</td>
<td>0.02</td>
<td>51%</td>
<td>14</td>
<td>0.03</td>
<td>0</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Synthesis with Quantified Boolean Logic [9]

<table>
<thead>
<tr>
<th>Name</th>
<th>r × s</th>
<th>n</th>
<th>S₀₀</th>
<th>S₀₁</th>
<th>L₀</th>
<th>E₁</th>
<th>S₁₀</th>
<th>S₁₁</th>
<th>L₁</th>
<th>%R₀</th>
<th>%R₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>add6_G_1</td>
<td>5×3</td>
<td>4</td>
<td>31</td>
<td>0.15</td>
<td>0.13</td>
<td>13%</td>
<td>32</td>
<td>0.14</td>
<td>0.13</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>alu2_G_2</td>
<td>7×3</td>
<td>8</td>
<td>464</td>
<td>0.1</td>
<td>0.09</td>
<td>14%</td>
<td>384</td>
<td>0.08</td>
<td>0.07</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>b11_G_1</td>
<td>3×5</td>
<td>7</td>
<td>45</td>
<td>0.03</td>
<td>0.02</td>
<td>2%</td>
<td>128</td>
<td>0.07</td>
<td>0.07</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>dc2_G_0</td>
<td>4×4</td>
<td>7</td>
<td>104</td>
<td>0.06</td>
<td>0.05</td>
<td>13%</td>
<td>132</td>
<td>0.07</td>
<td>0.06</td>
<td>13%</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Overall results of the simulations

<table>
<thead>
<tr>
<th>Synthesis Method</th>
<th>Average</th>
<th>Average</th>
<th>n</th>
<th>S₀₀</th>
<th>S₀₁</th>
<th>L₀</th>
<th>E₁</th>
<th>S₁₀</th>
<th>S₁₁</th>
<th>L₁</th>
<th>%R₀</th>
<th>%R₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>[2]</td>
<td>30</td>
<td>6</td>
<td>0.05</td>
<td>0.05</td>
<td>20%</td>
<td>0.06</td>
<td>0.05</td>
<td>29%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[9]</td>
<td>15</td>
<td>7</td>
<td>0.07</td>
<td>0.06</td>
<td>9%</td>
<td>0.07</td>
<td>0.07</td>
<td>8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Table 1, we report a sample of benchmark functions and their sensitivity values, according to the metrics presented before. In particular, Table 1 refers to lattice synthesized as described in [2] and [9]. The benchmarks that are present in Table 1 with dual method were stopped after ten minutes of SAT execution, but that was not the case for the rest.

More precisely, in both methods, the first column reports the name and the number of the considered output of each function. The following columns report dimension (r × s) required for the synthesis of a given function according to each decomposition method, and the number of input variables n. Columns from 4 to 7 refers to SA0 defect metrics (resp., columns from 8 to 11 to SA1 metrics) showing the total number of errors E₀, the Sensitivity of defective cells S₀, the Sensitivity of lattice S₀₁ and the percentage of robust cells %R₀/r × s.

Table 2 describes the overall results for the benchmarks we have considered. It also shows the average values for the considered metrics. We can note that the percentage of cells that are considered robust according to our metrics is higher in the first approach [2]. This is due to the more constrained structure of the lattices produced by the first synthesis method. Indeed, the method proposed in [2] computes a lattice for f and its dual that is in general less compact than the lattice given by [9] (see, the column Average area in Table 2). Moreover, we can note that the sensitivity of the lattice to stuck-at-defects (SAD) is quite low for both methods. In fact, the experiments show that, in general, non-robust cells -in presence of a SAD- compute a defective output for a very limited number of inputs.

4.2.4 Mitigation by Defect Avoidance. From the above results, it can be seen that the two analyzed mapping algorithm shows different sensitivities of the output of a given function. As a matter of fact, the more restrictive an algorithm is in terms of area (results closer to optimal solution), the higher the defect sensitivity of the output to cell defect of SA0 or SA1. It is mandatory to include in the mapping algorithm defect-avoidance heuristics.

In order to mitigate the sensitivity of a lattice to SAD, we propose the following possible strategy applied to the synthesis method proposed in [2] which has been proven as less sensitive to SAD impact on the output functions: (1) For a given mapped function, if a potential SA0, SA1 defect affects a robust cell identified by the defect injection campaign, the lattice still computes the correct output, thus we do not need any mitigation with defect tolerant design. (2) However, if an injected defect occurs in a multiple-choice cell, if a different literal can be chosen to make the cell robust, we change the literal with the new one. (3) Otherwise, if the injected SA0 defect is proven as being critical for the output value, the column that contains that defective cell has to be replaced by spare columns. In case of an SA1 the row that contains the defective cell has to be replaced by a spare row. Note that, in this case, the output still provides a correct function f from top to bottom, but the function from left to right could be changed and become a function which will not be dual of f anymore.

As an example, consider the lattice synthesized in Figure 6(a) with \( f = x_4 x_5 x_7 + x_4 x_6 x_7 + x_5 x_6 x_7 + x_4 x_6 x_7 + x_4 x_5 x_7 \) by using synthesis method presented in [2]. The example shows one case of mitigation of 3 independent SAD affecting the crossbar implementing the function, yielding an approximative 10% defects. In Figure 6, SA1 cells are marked in blue and SA0 cells a remarked in red.

4.3 Transient Fault Tolerance

Regarding transient faults, there are two approaches: redundancy based and manipulation of logic function to obtain a more fault tolerant design. Nevertheless, it should be noted that since nanocrossbars are in the early stage of development, there is a lack of in-field data regarding transient faults. For this reason, second approach towards transient fault tolerance is methodology independent and rather focuses on the intrinsic features of the given logic functions.

Figure 6: a) defect-free lattice; b) lattice with defects: SA0 in red and SA1 in blue; and c) lattice with the defect fixed.

\[ f = x_4 x_5 x_7 + x_4 x_6 x_7 + x_5 x_6 x_7 + x_4 x_6 x_7 + x_4 x_5 x_7 \]
As mentioned, transient faults can be tolerated with redundancy. Inserting redundant components can be constructed with adding extra rows and/or columns as shown [16] [3].

Furthermore, provided certain conditions, particular logic functions are inherently tolerant to transient faults limited to certain switches of crossbar as show in [21].

Mentioned inherent tolerance capability varies depending on the design. For example, consider a design having low logic inclusion ratio (IR), meaning less number of crosspoints of a crossbar are used, that means this design is more tolerant to stuck-at-0 faults. Logic synthesis (design) should be made with regard to this, since there are multiple design solution.

Similarly, if the target function can be realized with high IR, then a technology, the one which tends to have stuck-at-1 faults, should be preferred. For example function \( f = x_1x_2 + x_2x_3 + x_3x_4 \) can also be written as \( f' = x_1x_2 \overline{x_3} + x_2x_3 + \overline{x_3}x_4 = f \), therefore IR can be increased.

Fault sensitivity analysis can be made using Monte Carlo analysis, yet it is costly. Since we know the dynamics of the fault tolerance we can calculate sensitivity (fault tolerance performance) directly with algebraic equations. Equation parameters consists of crossbar variation tolerant delay values, different optimization algorithms for switching components of the arrays. Future direction. Another future direction is performing sensitivity analysis for switching nanoarrays. In Very Large Scale Integration (VLSI-Soc), 2016 IEEE International Conference on, IEEE, 1–6.

5 VARIANCE TOLERANCE AND PERFORMANCE OPTIMIZATION

Another aspect of the crossbar fabrication is that every crosspoint does not have the same property in terms of dimension, doping, etc. [14]. It is called variance of crosspoints. This affects threshold voltages and ON and OFF resistances as well as capacitance values. It means that delay and power performances are changing. Decision of which crosspoint switches are going to be used during logic design plays a crucial role in performance optimization. To achieve variation tolerant delay values, different optimization algorithms have been proposed [28] [15]. These algorithms aim to optimize the worst-case delay values in logic mapping. They use Gaussian distribution to model variances.

On the other hand, the literature lacks variation tolerant power optimization algorithms. Considering that fault tolerance mechanism can be appended to the delay minimization process, variation-power-delay optimizations are needed. This can be considered as a future direction. Another future direction is performing sensitivity analysis for switching components of the arrays.

6 CONCLUSION AND DISCUSSION

In this study, we present a synthesis methodology for crossbar arrays having crosspoints working as FET, diode/resistive/memristive, or four-terminal switch based devices. We cover "logic synthesis", "defect/fault tolerance", and "variation-area-power-delay performance optimization" steps. Presented synthesis methodology provides optimization algorithms for each step of the process as well as their relations and trade-offs.

REFERENCES