UNIVERSIDAD POLITÉCNICA DE MADRID
ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE TELECOMUNICACIÓN

RESISTIVE RAM: SIMULATION AND MODELING
FOR RELIABLE DESIGN

Ph.D. Thesis

Fernando García Redondo
B.S. Electronic Engineering
M.Sc. in Electronic Systems Engineering
2017

2017 – FERNANDO GARCÍA REDONDO

To be printed in B5 paper.

This work is licensed under a [Creative Commons](#) “[Attribution-NonCommercial-ShareAlike 3.0 Unported](#)” license.

Some parts of this work, such as some of the developed software (Section 3.4) may be licensed as MIT.

DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA
ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE TELECOMUNICACIÓN
UNIVERSIDAD POLITÉCNICA DE MADRID

RESISTIVE RAM: SIMULATION AND MODELING
FOR RELIABLE DESIGN

Ph.D. Thesis

Author:

Fernando García Redondo

B.S. Electronic Engineering

M.Sc. in Electronic Systems Engineering

Advisor:

María Luisa López Vallejo

Full Professor, Dpto. de Ingeniería Electrónica

Universidad Politécnica de Madrid

2017

PH.D. THESIS: Resistive RAM: Simulation and Modeling for Reliable Design

AUTHOR: Fernando García Redondo

ADVISOR: María Luisa López Vallejo

The Tribunal appointed to judge the above thesis,
composed by the following doctors:

PRESIDENT:

VOCALS:

SECRETARY:

agree granting the qualification:

Madrid, June 5th, 2017
Tribunal Secretary

ABSTRACT

Resistive switching technologies are regarded as one of the most promising alternatives to future storing and computing systems. Characteristics like nonvolatility, very-high-storage density, low power consumption, as well as the compatibility with traditional CMOS circuits, raise the attention of device manufacturers, circuit designers and system architects. In the recent years, extensive and thorough studies regarding RRAM materials, manufacturing, cell and array structures, as well as circuit modules and system architectures have been conducted. However, as many emerging technologies, the design of these novel circuits not only inherits the CMOS-caused reliability issues, but adds new challenges that must be addressed and mitigated.

In this setting, this thesis focuses on the modeling, simulation and characterization of memristor devices and circuits used for the design of reliable applications. Approaching the topic from three different perspectives —device, circuit and system levels, we develop solutions to important issues that otherwise would compromise the performance and correct behavior of the circuits.

As an introduction, the thesis begins reviewing the key concepts and characteristics of resistive switching technologies: formulation and appearance, behavior and characteristics of physical devices, applications and modeling, and, finally, the main reliability issues that need to be addressed.

The first part of the thesis approaches the modeling, simulation and characterization from the device abstraction level. We have developed a characterization framework —and related simulation methodology— for the definition, analysis and improvement of present and future memristor compact models. Closing the gap between devices manufacturing and circuit designers, this framework allows to study memristor compact models, including characteristics such as variability, temperature dependence or multi-level capabilities, for the later generation of the improved SPICE subcircuit. Combined with the related methodology, the framework contributions can be summarized as: the characterization of the best-known memristor models, generation of SPICE subcircuits, systematic modifications to improve SPICE code leading to significant simulations speed-ups, analysis of parameter variations, temperature and input dependence, support for multi-level devices, and, finally, mechanisms that improve simulation convergence, computational load and the evolution of system variables. In second place we have developed an RRAM SPICE compact model. This model relies on the idea of independently describing conductive processes and switching mechanisms, and it is

able to resemble both unipolar and bipolar physical devices. In this case, the state switching is described using electrical magnitudes —voltage, current, flux, charge or energy. The model considers complex device characteristics such as multi-level capabilities, time-dependent variability, device degradation or data retention. Additionally, the model is accurately fitted to two different physical devices, one exhibiting unipolar behavior and one displaying bipolar characteristics.

The second part of the thesis, Chapter 5, is devoted to the characterization and analysis of RRAM-CMOS hybrid circuits. We first discuss the problems that appear in circuits using RS and short-node CMOS technologies, as well as the lack of tools and methodologies to address these issues. Then, as a solution to such problem, we present a design and characterization strategy together with an ad-hoc simulation framework. Using the presented multi-metric evaluation paradigm, we are able to measure the performance and reliability of circuits that simultaneously suffer from the effects of several uncertainty sources. Therefore, automating the always difficult evaluation of the effects caused by PVT variations, aging, degradation, data retention or radiation. Different cases of study, from digital modules to RRAM-CMOS hybrid circuit, highlight the solution capabilities.

Moving up to the system level, the third part of the thesis deals with the design of reliable RRAM-CMOS circuit applications. Using the resources presented in previous parts, we provide solution to two different issues affecting RRAM memories. First, we introduce a novel architecture that is able to mitigate the dramatic effects caused by temperature in RRAM memories. Regardless of the circuit temperature, this writing architecture dynamically adapts the writing voltage amplitude to the environmental conditions, correctly achieving the desired operation. Moreover, given that different writing schemes imply distinct cell lifetime, power consumption, or other characteristics, the architecture adapts the writing scheme to specific application objectives. By dynamically loading different writing profiles, the driver selects among a myriad of writing strategies improving the system performance, power consumption, data retention or HRS/LRS ratio durability. Finally, memories exhibiting non-volatility characteristics may impose a security breach in highly secure applications. The last contribution comes as a solution to overcome this drawback, and provides RRAM memories with an additional security mechanism. Data-encryption may not be sufficient, so we have developed an auto-erasable memory architecture, to secure RRAM memories from physical and firmware attacks. This way, and complementing encryption approaches, this architecture provides an additional security layer, preventing the access to sensitive data. Additionally, an optional emergency power supply based on on-chip capacitors, allows the designers to include an additional power source to be used in case a physical disconnection from the main power supply line occurs.

RESUMEN

Las tecnologías emergentes de resistencia variable o *resistencia-commutativa*, *Resistive-Switching (RS) technologies*, están consideradas como una de las alternativas más prometedoras para el futuro de aplicaciones electrónicas de almacenamiento y sistemas de computación. Características tales como su no-volatilidad, alta densidad de integración, bajo consumo, así como su compatibilidad con circuitos CMOS tradicionales, despierta el interés de fabricantes de dispositivos, diseñadores de circuitos y arquitectos de sistemas. En estos últimos años, variados y exhaustivos estudios se han centrado en el análisis de materiales para la fabricación de dispositivos RRAM, procesos de fabricación o estructuras de célula. Asimismo, se han dedicado vastos esfuerzos al diseño de circuitos y arquitecturas que hagan uso de estas tecnologías. Sin embargo, y como en otras tecnologías emergentes, el diseño de estos novedosos circuitos adolece no sólo de los problemas de fiabilidad experimentados por los dispositivos CMOS, sino que también se añaden nuevas fuentes de problemas, y por tanto de nuevos retos, que deben ser analizados y solventados.

En este contexto, este trabajo de tesis se centra en el modelado, simulación y caracterización del memristor, así como de circuitos fabricados usando estas tecnologías y que luego son empleados para el diseño de aplicaciones fiables. Analizando la problemática desde tres perspectivas diferentes —nivel de dispositivo, circuito y sistema—, presentamos solución a cuestiones que, de otra manera, comprometerían las prestaciones del circuito y su correcto funcionamiento.

A modo de introducción, la tesis comienza revisando los aspectos más importantes de las tecnologías RS: conceptos básicos, formulación y morfología, comportamiento y características de dispositivos físicos, principales tipos de modelos, aplicaciones más importantes y, finalmente, los problemas de fiabilidad que requieren ser tratados.

La primera parte de la tesis aborda el modelado, simulación y caracterización desde la perspectiva del dispositivo. Hemos desarrollado un entorno software de caracterización —así como una metodología de simulación asociada— para la definición, análisis y mejora de modelos compactos de memristor, tanto presentes como futuros. Eliminando la barrera entre fabricantes de dispositivos y diseñadores de circuitos, este entorno permite el estudio de modelos compactos de memristor, teniendo en cuenta características como variabilidad, dependencia con temperatura o capacidades multi-nivel, teniendo como objetivo la generación de un subcircuito *SPICE* mejorado. Las principales contribuciones relacionadas con el software y su

metodología asociada son: la caracterización de los modelos de memristor más conocidos; la generación y mejora sistemática de subcircuitos *SPICE* dando lugar a elementos circuitales más rápidos y con menos problemas de convergencia; el análisis de variaciones de parámetros; el estudio de su dependencia con la temperatura y estímulos; y por último, el soporte de dispositivos multi-nivel. En segundo lugar hemos desarrollado un modelo compacto *SPICE* de dispositivos RRAM. Basado en la idea de la descripción independiente de los procesos de conducción y mecanismos relativos al cambio de estado, este modelo es capaz de emular el comportamiento de todo tipo de dispositivos RS físicos. Haciendo uso de magnitudes eléctricas —voltaje, corriente, flujo, carga o energía— es capaz de describir desde la dinámica del dispositivo a complejos efectos tales como comportamiento multi-nivel, dependencia con el tiempo de la variabilidad, la degradación de las características del dispositivo o el tiempo de retención de dato. Este modelo se adecúa al proceder y propiedades de dispositivos físicos tanto unipolares como bipolares.

La segunda parte de la tesis está consagrada al análisis y caracterización de circuitos híbridos RRAM-CMOS. En primer lugar discutimos la problemática asociada al uso de dispositivos RS y tecnologías CMOS en tamaños nanométricos, así como a la falta de herramientas y metodologías para hacer frente a estas dificultades. Posteriormente, y como solución a estos problemas, hemos ideado una estrategia para el diseño y la caracterización de este tipo de circuitos, que se encuentra acompañada por un entorno CAD *ad-hoc* de simulación. Aplicando un paradigma de evaluación de múltiples métricas, conseguimos caracterizar las prestaciones y la fiabilidad de circuitos que sufren simultáneamente de variadas y heterogéneas fuentes de incertidumbre. Con esta estrategia automatizamos la siempre difícil evaluación de los efectos causados por variaciones PVT, envejecimiento, degradación, retención de datos o radiación. Para recalcar las capacidades de la solución propuesta, se describen diferentes casos de estudio, que incluyen módulos digitales, analógicos y circuitos híbridos RRAM-CMOS.

Escalando a nivel de sistema, la tercera parte de la tesis se centra en el diseño de aplicaciones fiables que hagan uso de circuitos híbridos RRAM-CMOS. Usando el conocimiento y recursos presentados en las partes anteriores, damos solución a dos graves problemas que afectan a las memorias RRAM. Primero, presentamos una novedosa arquitectura que es capaz de mitigar los efectos causados por variaciones de temperatura en este tipo de memorias. Sin importar la temperatura del circuito, la arquitectura de escritura adapta dinámicamente la amplitud del voltaje de escritura a las condiciones del entorno, consiguiendo realizar la operación con éxito. Además, y en relación al hecho de que diferentes esquemas de escritura conllevan efectos y características distintos —velocidad, consumo, etc.—, la arquitectura adapta el esquema de escritura en función de los objetivos de la aplicación. Así,

cambiando dinámicamente el perfil de escritura, el *driver* selecciona la estrategia de escritura más apropiada, mejorando las prestaciones del sistema ya sea en velocidad, consumo, vida útil de las células, retención de dato o ratio HRS/LRS. Por otra parte, las memorias no-volátiles conllevan una brecha en la seguridad de aplicaciones altamente seguras, donde la encriptación de datos puede no ser suficiente. La última contribución de la tesis aborda este gran inconveniente, y para ello dotamos a las memorias RRAM con una capa de seguridad adicional. La solución consiste en una arquitectura que, en caso de ataques físicos o de *firmware*, borra el contenido de las memorias RRAM de manera inmediata. De esta manera, y como complemento a soluciones basadas en encriptación, la arquitectura otorga un nivel adicional de seguridad impidiendo el acceso a datos sensibles. Adicionalmente, presentamos una fuente de energía opcional, que, basada en condensadores integrados en el chip, proporciona la energía de borrado en casos donde el ataque físico desconecte el chip de la fuente de energía principal.

Contents

Abstract	i
-----------------	----------

Resumen	iii
----------------	------------

Agradecimientos	xxiii
------------------------	--------------

Introduction	1
---------------------	----------

1 Introduction	1
-----------------------	----------

1.1 Motivation	3
--------------------------	---

1.2 Objectives	5
--------------------------	---

1.3 Nomenclature	6
----------------------------	---

1.4 Contributions	7
-----------------------------	---

1.4.1 Journal Papers	7
--------------------------------	---

1.4.2 Licensed Software	8
-----------------------------------	---

1.4.3 International Conference Papers	8
---	---

1.5 Document Structure	10
----------------------------------	----

2 Introduction to Resistive Switching Technologies	13
---	-----------

2.1 Chua's Memristor	15
--------------------------------	----

2.2 Resistive Switching Technologies	17
--	----

2.2.1 Device Structure and Behavioral Operation	17
---	----

2.2.2	Categories of RS Technologies	18
2.2.3	Switching Mechanisms Summary	19
2.2.4	Conduction Processes Summary	20
2.2.5	Multi-level (ML) and Multi-bit (MB) capabilities	22
2.2.6	RRAM-CMOS Hybrid Chips	23
2.3	Memristor Models	24
2.3.1	Dynamic Simulation Models	24
2.3.2	Kinetic Monte Carlo Simulation Models	24
2.3.3	Continuum Models	24
2.3.4	Percolation Models	25
2.3.5	Compact Models	25
2.4	Resistive Switching Applications	26
2.4.1	Conventional Applications	26
2.4.2	Novel Computing Paradigms	28
2.5	Reliability Issues in RRAM and CMOS Hybrid Circuits	31
2.5.1	Introduction to Uncertainty Sources in CMOS Technologies	31
2.5.2	Introduction to Uncertainty Sources in RRAM Technologies	32
2.6	Summary	35

I Modeling and Simulation of RRAM Devices 37

3 Modeling and Characterization of RS Technologies 39

3.1	Introduction	39
3.2	Review of Best Known Compact Models	42
3.2.1	Linear ion drift	42
3.2.2	Non-linear ion drift models with different window functions	43
3.2.3	Simmons' tunneling barrier model	43
3.2.4	Yakopcic's model	43
3.2.5	VTEAM model	44
3.2.6	Simplified Simmons' barrier model	44
3.2.7	Eshraghian's MIM & threshold model	44
3.2.8	1-Dimensional CF, Arizona University - Stanford University	44
3.2.9	2-Dimensional CF, Stanford University - Beijing University	45
3.2.10	1-Dimensional Filament Growth Model, Politecnico de Milano University	45
3.2.11	Review's Summary	45
3.3	Systematic Improvement of Model Netlists	50
3.4	Memristor Application Framework MAF	53
3.4.1	Model Library	54

3.4.2	Device Characterization Module	55
3.4.3	Level Characterization Module	58
3.4.4	Subcircuit Editor Module	59
3.5	MAF Design Methodology	59
3.5.1	Model Developers	59
3.5.2	Circuit Designers	63
3.6	Speedup Results	68
3.7	Conclusions	69
3.8	Related Contributions	71
4	SPICE Modeling of Resistive Switching Technologies	73
4.1	Introduction	73
4.2	Developed Model	77
4.2.1	Conduction Module Modeling	77
4.2.2	State Module Modeling	80
4.2.3	Extensibility	85
4.2.4	Verilog-A Implementation	86
4.3	Non-Ideal Effects	86
4.3.1	Related Work	86
4.3.2	Proposed Solution	88
4.4	Fitting and Simulation of Physical Devices	90
4.4.1	Physical Devices	90
4.4.2	Energy Threshold	90
4.4.3	Conductive Bridge RAM (CBRAM) Bipolar Device	92
4.4.4	Metal Oxide Based RAM (Ox-RAM) Unipolar Device	92
4.4.5	Non-Ideal Effects	95
4.5	Conclusions	99
4.6	Related Contributions	100

II Characterization and Analysis of RRAM-CMOS Hybrid Circuits **101**

5	Reliable Simulation of RRAM-CMOS Hybrid Circuits	103
5.1	Introduction	103
5.2	Overview of the Presented Solution	105
5.3	Strategy for Reliable Analyses of Circuits	107
5.4	Circuit-Level Simulation Framework	112
5.4.1	Functional Structure	113
5.4.2	Circuit Modifiers: Model Enhancement Schemes and Radiation Schemes Implementation	117

5.4.3	Simulation Modes	121
5.4.4	Variability, Corners and Different Technologies Handling	122
5.4.5	Metrics Implementation	125
5.5	Validation Against Experimental Data	127
5.5.1	Static Analysis Validation	128
5.5.2	Dynamic Analysis Validation	132
5.5.3	Variability Analysis Validation	136
5.6	Digital Module Characterization	137
5.6.1	Sequential Circuits Granularity Considerations	139
5.7	Analog Module Characterization	142
5.7.1	Bandgap Voltage Reference	143
5.7.2	Power Supply Monitor	150
5.8	RRAM Read Circuit Characterization	155
5.8.1	Original Technology Models	156
5.8.2	Target Circuit to Be Characterized	157
5.8.3	Design and Characterization Process	157
5.9	Conclusions	170
5.10	Related Contributions	172

III Reliable Design of RRAM Based Systems and Applications **175**

6	Reliable, Reconfigurable RRAM Writing Architecture	177
6.1	Introduction	177
6.2	Resistive Switching Dependence on Temperature	181
6.3	Dynamic Profile Thermal-Adaptive Write Architecture	184
6.3.1	Optimal Voltage Amplitude Selection	186
6.3.2	Profile-Temperature Handler Core	186
6.3.3	Calibration of Physical Devices	189
6.3.4	Optimum Voltage Amplitude Selection Algorithm	190
6.3.5	Dynamic Profiles	191
6.3.6	Voltage References System	192
6.3.7	Optional Modules	193
6.4	Full System Implementation	194
6.4.1	Profile and Temperature Handler Core	194
6.4.2	Implementation of the Voltage References System	195
6.5	Temperature Aware, Reliable Oriented Profiles Results	197
6.5.1	Required Voltages and Power Consumption Results	199
6.5.2	Reliability and Speed-Up Results	201

6.6	Conclusions	205
6.7	Related Contributions	205
7	RRAM Memory Secured Against Physical Hacks	207
7.1	Introduction	207
7.2	Information Securing Architecture	209
7.2.1	Erasing Circuitry	211
7.2.2	Erasing Procedure	211
7.2.3	Erasing Control	214
7.2.4	Optional Emergency Erase Power Supply	215
7.3	Temperature Influence	217
7.4	Fast Resistive Switching Modeling	218
7.5	Implementation and Results	219
7.5.1	Architecture Standard Erasing Procedure	220
7.5.2	Architecture's Optional Erasing Procedure	225
7.6	Conclusions	231
	Conclusions and Future Work	233
8	Concluding Remarks	233
8.1	Future Work	236
	Appendices	239
	Appendix A MAF Implementation Details	239
A.1	Electric Device Model	239
A.1.1	Magnitudes Modeling	240
A.1.2	Parameters	240
A.1.3	Model Simulation Control	240
A.2	Simulation Scenarios and PVT Variations	241
A.3	Thread Management	241
A.4	Non-Convergence Problems Handling	242
A.5	Data Population Control	246
	Appendix B Multi-level Device LTSPICE Code	247
	Appendix C Radiation Modeling and Tools	251
C.1	Overview of Radiation Models and Tools	251
C.1.1	CMOS Technologies: Physical and Transistor Level	252
C.1.2	Gate and System Level	253

C.1.3	Devices Not Supported in Standard Work Flow: RRAM and Analog Modules	253
Appendix D	Simulation Framework Implementation Details	255
D.1	Simulation Load and Data Management	255
D.1.1	Critical Parameter Value Binary Search	255
D.1.2	Sweep Parameter Design Space Division	257
D.1.3	Masked and Omitted Nodes or Instances	258
D.1.4	Granularity	258
D.1.5	Computational Acceleration Techniques	259
D.1.6	Efficient Computation Example	261
Bibliography		263

List of Figures

1.1	Thesis contents	11
2.1	Chapter 2 scope	14
2.2	The memristor and involved electrical magnitudes.	15
2.3	Ox-RAM structure	17
2.4	Memristor switching	19
2.5	Different Modeling Approaches	25
2.6	Computing paradigms	29
3.1	Chapter 3 scope	40
3.2	Chapter 3 overview	40
3.3	Traditional versus presented memristor subcircuits	50
3.4	MAF functional scheme	54
3.5	MAF Device Characterization Mode	55
3.6	MAF Device Characterization	56
3.7	MAF fitting functionality	57
3.8	MAF Level Characterization Mode	58
3.9	MAF importing data measures	60
3.10	Gross space exploration using MAF	61
3.11	MAF fitting	62
3.12	MAF final fitting	62
3.13	MAF methodology	63
3.14	3×3 crossbar circuit	64
3.15	State variable evolution	65
3.16	State variable evolution with variations	66

3.17	MAF Multi-level histograms	67
3.18	State variable evolution in 3×3 crossbars	68
3.19	MAF Speed-up results	69
4.1	Chapter 4 scope	75
4.2	Compact modeling approach	77
4.3	2 bit multi-level device	78
4.4	N -state conduction module	79
4.5	Device State Handler Module	80
4.6	Hysteresis loop examples.	81
4.7	Soft-switching mechanism example	84
4.8	Soft-switching mechanism example	84
4.9	CF's local temperature simulation	85
4.10	Energy expended in both <i>SET</i> and <i>RESET</i> processes	91
4.11	Energy expended in both <i>SET</i> and <i>RESET</i> processes	91
4.12	Bipolar CBRAM subcircuit and $i(v)$ relation	93
4.13	Unipolar Ox-RAM subcircuit and $i(v)$ relation	94
4.14	Example of Random Telegraph Noise (RTN) dependency on cycling	95
4.15	Example of cycle dependent variability	96
4.16	Variability effects dependent on cycling and electrical stress	97
4.17	Non-ideal effects dependent on cycling and electrical stress	98
5.1	Chapter 5 scope	106
5.2	Standard workflow design	107
5.3	Presented analyses strategy	108
5.4	Models Redefinition	109
5.5	Evaluation of user defined metrics	110
5.6	Efficient definition of the design space	111
5.7	Developed circuit characterization framework	113
5.8	Example Circuit	114
5.9	Example Circuit, stage 2	114
5.10	Example Circuit, stage 3	115
5.11	Example Circuit, stage 4	115
5.12	Example Circuit, stage 5	116
5.13	Substitution scheme examples	117
5.14	Example of hierarchy tree structure	118
5.15	Example of dependency tree structure	119
5.16	Example of modifiers algorithm	120
5.17	<i>Basic Analysis Mode</i> workflow	121
5.18	Comparison between traditional and proposed solutions.	127
5.19	Q_{crit} dependence on τ_2	129

5.20	Q_{crit} dependence on transistors size	130
5.21	SER dependence on cell dimensions	131
5.22	Normalized read ($RSNM$) and write ($WSNM$)	132
5.23	Dynamic Q_{crit} in a nominal SRAM	133
5.24	SRAM Variability and radiation dynamic analysis	134
5.25	SRAM Variability and radiation dynamic analysis	135
5.26	DICE flip-flop schematic	137
5.27	Radiation impacts modeling in registers	138
5.28	Non-propagated SEU example	141
5.29	Registered SEU example	142
5.30	Simplified bandgap schematic	143
5.31	Bandgap output voltage error for each corner	145
5.32	Q_{crit} for different corners in wide temperature ranges	146
5.33	Bandgap variability and temperature results	146
5.34	Bandgap error under radiation	148
5.35	Mean upset ratio dependence on temperature and Q_{coll}	149
5.36	Power supply monitor schematic	150
5.37	Power supply monitor nominal behavior	151
5.38	Power monitor MC Analysis, wide temperature range	151
5.39	Power monitor Q_{crit} dependence on corners	152
5.40	Power monitor global upset rate	153
5.41	Power monitor metrics en example nodes	154
5.42	Schematic of the studied RRAM-CMOS hybrid circuit	155
5.43	RRAM read circuit transients	156
5.44	Scheme illustrating how using a SPICE subcircuit we redefined the RRAM original model.	158
5.45	RRAM model enhancement: Endurance	159
5.46	RRAM model enhancement: TID	159
5.47	RRAM read circuit: EMs and PM metrics	161
5.48	RRAM read circuit: PM_1 dependences	162
5.49	RRAM read circuit: Metrics evaluation	163
5.50	RRAM circuit: Environmental and stress evaluation	164
5.51	RRAM circuit: reliability characterization	165
5.52	RRAM circuit: reliability characterization, partial results	167
5.53	RRAM circuit: reliability characterization, partial results	168
5.54	RRAM circuit: reliability characterization with $SEEs$	169
6.1	Chapter 6 scope	178
6.2	Scheme of the used compact model [HWL⁺14 , HCL⁺14 , LJH⁺15]	181
6.3	Required time to perform operations using a fixed operation voltage.	183

6.4	Developed writing architecture	184
6.5	Integration of the writing architecture in 2D and 3D morphologies	185
6.6	Main core architecture	187
6.7	Sensing-voltage amplitude selection process flow	188
6.8	Temperature computation modes	188
6.9	Proposed calibration methodology	189
6.10	Voltage amplitude selection algorithm	190
6.11	Required voltage amplitudes	191
6.12	Dynamic profile loading schedule	192
6.13	<i>Voltage references system</i>	193
6.14	Implemented 40 nm bandgap voltage reference.	196
6.15	<i>Voltage reference system</i> response under variability	196
6.16	Required voltages and consumed energies for every profile	200
6.17	Comparison between traditional and proposed SET operation	201
6.18	Operation error dependence on temperature	202
6.19	Achieved speed-ups	203
7.1	Chapter 7 scope	208
7.2	Proposed information securing architecture	210
7.3	Circuitry driving the erasing pulses to the memory crossbar	211
7.4	Equivalent circuit for a crossbar RRAM array during erase process	212
7.5	Circuitry driving the erasing pulses to the memory crossbar	213
7.6	Erasing process	214
7.7	Example of functioning under physical attack. The CMOS capacitor discharges through the memory cells, erasing them.	216
7.8	1T – 1R Fast switching behavior of the used compact model	218
7.9	Low Resistance State (LRS) to High Resistance State (HRS), standard erasing procedure, 16 × 16 RRAM block	221
7.10	LRS to HRS, standard erasing procedure, 32 × 32 RRAM block	222
7.11	LRS to HRS, standard erasing procedure, 64 × 64 RRAM block	222
7.12	500 Monte Carlo simulations at 220 K circuit temperature.	224
7.13	LRS to HRS, emergency erasing procedure, 16 × 16 RRAM block	226
7.14	LRS to HRS, emergency erasing procedure, 32 × 32 RRAM block	227
7.15	LRS to HRS, emergency erasing procedure, 64 × 64 RRAM block	227
7.16	Different discharging ratios for different memory sizes.	228
7.17	Different discharging ratios for different memory sizes.	228
7.18	Variability effects at different temperatures	230
A.1	MAF Thread management	242
A.2	MAF Thread diagram	244
A.3	MAF Data population control	246

D.1 Binary search example 257

D.2 n bit adder 259

D.3 Simulation summary results for 1 bit and 14 bit adders 260

List of Tables

2.1	Basic conduction processes	21
2.2	Variables and parameters for the expressions shown in Table 2.1.	22
3.1	Memristor compact models	46
3.2	Review of memristor models presented in the literature	48
3.3	MAF Speed-up results summary	70
4.1	Models comparison.	76
5.1	Q_{crit} in standard register	139
5.2	Q_{crit} in DICE register	140
5.3	Q_{crit} in registers	141
5.4	DICE Chain Summary.	142
6.1	Conductive Filament (CF) cylinder sizes and resistance, by profile	197
6.2	Different profiles and maximum temperature deviation they can assume and mitigate.	198
6.3	Calibrated voltage amplitudes required for each profile.	202
7.1	Percentage of erasing accomplishment in 500 MC, using standard erasing procedure, 64×64 RRAM block	223
7.2	Percentage of erasing accomplishment in 500 MC, using emer- gency erasing procedure, 64×64 RRAM block	229
D.1	14 bit adder inventory	261
D.2	Critical charge results comparison between adders	262

Acronyms

1R	1-Memristor
1T1R	1-Transistor-1-Memristor
1S1R	1-Selector-1-Memristor
BE	Bottom Electrode
CAM	Content Addressable Memory
CBRAM	Conductive Bridge RAM
CIM	Computation In Memory
CMOS	Complementary Metal-Oxide-Semiconductor
CF	Conductive Filament
CRS	Complementary Resistive Switches
ECM	Electro Chemical Memory
EM	Error Metric
FPGA	Field Programmable Gate Array
HRS	High Resistance State
IoT	Internet of Things
LRS	Low Resistance State

MB	Multi-Bit
MC	Monte Carlo
MIM	Metal/Insulator/Metal
MOM	Metal/Oxide/Metal
ML	Multi-Level
NVM	Nonvolatile Memory
RRAM	Resistive RAM
Ox-RAM	Metal Oxide Based RAM
PDK	Process Design Kit
PM	Performance Metric
PVT	Process Voltage and Temperature
PUF	Physical Unclonable Function
PWL	Piecewise Linear
RNG	Random Number Generator
RS	Resistive Switching
RTN	Random Telegraph Noise
SET	Single Event Transient
SEU	Single Event Upset
SER	Single Event Rate
SPICE	Simulation Program with Integrated Circuit Emphasis
SRAM	Static RAM
TCAM	Ternary CAM
TE	Top Electrode
TID	Total Ionizing Dose
VCM	Valence Change Memory
vNA	von Neumann Architecture

Agradecimientos

Acabando el documento, me encuentro volviendo la mirada atrás y repasando todas las personas e instituciones que, sin duda, lo han hecho posible. En primer lugar, me gustaría dar las gracias a las instituciones y proyectos que han financiado el trabajo de investigación recogido en este documento:

- El Ministerio de Economía, Industria y Competitividad del Gobierno de España, por medio de los proyectos TOLERA 2 (TEC2012-31292), RED VARIABLES (TEC2014-53909), RAD-HARQ (IPT-2012-0422-370000), TOLERA (TEC2012-31292) y RED NANOVAR (TEC2011-15599-E).
- El Laboratorio de Systems Integrados de la Universidad Politécnica de Madrid, en donde he realizado la mayor parte del trabajo de investigación de la tesis, y he podido iniciarme en la docencia.
- El Consejo Social de La Universidad Politécnica de Madrid, promoviendo las estancias y movilidad de los investigadores.
- Aquellas personas que desinteresadamente han contribuido de una u otra manera en los proyectos de software libre usados durante la tesis. Sin estas herramientas no habría podido obtener los resultados aquí presentados.

Querría agradecer a Marisa López Vallejo todo el esfuerzo, tiempo y consejo que me ha dedicado durante toda esta etapa. Sería injusto decir sólo que ha sido tutora y directora de tesis, cuando en realidad he podido disfrutar de una maravillosa profesora, jefa y compañera de trabajo. Ya hace mucho tiempo, pero otros dos grandes profesores también espolearon mis inquietudes: María y Rafa.

Este camino ha estado plagado de muy buenos momentos que he compartido con los miembros del LSI, desde el director, Carlos López Barrio, a compañeros

con los que día a día he trabajado. No han sido pocos los cafés, comidas o vinos, ni las veces en las que me han echado una mano. Por ello, y por todo lo que he aprendido de ellos, me gustaría dar las gracias a Pablo Ituero, Pablo Royer, Carlos, Pablo Barrio, Hernán, Roberto, Victor Iglesias y Miguel Ángel.

Me gustaría también dar las gracias a Mireia Bargallo, Francesca Campabadal (IMB-CNM), Montserrat Nafria, Javier Martín Martínez y Albert Crespo (UAB) por el acceso a muestras de dispositivos unipolares y equipos de medida. Similarly, I change to English to acknowledge Kees De Groot, Robert P. Gowers, Katrina Morgan and Junqing Fan, from Southampton University, for supplying the bipolar devices. De igual manera, me gustaría agradecer a Ricardo Rianza (UPM) y a Sebastià A. Bota su sabiduría y sugerencias.

The following sentence is written in English, though it would have the same meaning either in Italian or in Greek: I would like to thank the ‘Mediterranean’ people back at Southampton. It was a short time but I am really glad that I had the chance to know you. Simmone, Ilaria, María, Isha... I own you a lot of espressos.

Desde mucho antes de llegar a la Escuela, he tenido la suerte de estar rodeado de personas con las que he crecido y disfrutado. Sandra, y luego Giaco, son buen ejemplo. Junto con todos los que han pasado por el Delta, forman mi segunda familia. Fido, Nico y David; mis ingenieras electrónicas, Sara y Ceci; los cafés y las charlas con Dani. Desde más atrás aún me acompañan, aun cada uno por su lado, Bea, Narx, Sperman y Juancho. Por supuesto, siempre al pie del cañón Edu, Mery, Mora, Emilio, Ángel, Vicky, los peques... y sobre todo Txus. Con todos ellos puedo decir que he tenido el placer de compartir mis mejores experiencias y recuerdos.

Javi es, aparte de Rebe, con quien más tiempo he compartido en estos últimos cinco años. Con pocas personas he podido discutir más, y de pocas personas he podido aprender tanto. Gracias, Javi.

También quiero dar las gracias a mis tíos y abuelos, incondicionales en todas mis andanzas.

A mis padres, pues siempre han sido el ejemplo a seguir, y son por tanto tan culpables como yo de que haya acabado escribiendo esta tesis.

A mi hermana, a la que la palabra luchadora se queda corta.

A Rebe. Hay mil y una razones por las que darte las gracias, pero la más importante de todas es que siempre estás ahí.

Fernando García Redondo
Madrid, 31 de Marzo, 2017

PARA REBECA Y ELENA.

Good news, everyone.

Professor, Futurama.

1

Introduction

The world of electronics is living a time of headlong changes. We are witnessing a revolution, similar to the popular spreading of the personal computer back in the 1970s, or the widespread commercialization of the first consumer laptops during the 1990s. Historically, full workstations were meant to bring the application change to the consumer. On the contrary, nowadays there is a trend toward less centralized but more numerous and interconnected systems. On the one hand, the most powerful servers are starting to take advantage of a large number of highly heterogeneous processors, communicated through high speed interfaces. On the other hand, minimal but omnipresent applications interconnect and provide the most simple household appliance with intelligent capabilities. Together with mobile electronic applications, the Internet of Things (IoT) is drastically modifying the processing, communication and storing needs of today's systems [Aus15]. Unavoidably, and led by this change of paradigm, electronic systems are evolving and, in order to meet these present and future requirements, novel ways of design and emerging technologies are to be used [JKK⁺16].

Since Complementary Metal-Oxide-Semiconductor (CMOS) technologies first appeared in the 1960s, their use and importance have done nothing but increase. Gordon Moore's Law accurately predicted CMOS evolution, and it was not until the first decade of the 2000s that, due to both scaling processes and power dissipation in high transistor densities, this law started to decline [MAA⁺07]. After this moment, the predicted increase in performance and the power budget could no longer be maintained by rising the clock frequency. Designers had to look for innovative solutions: first frequency and voltage scaling, then core replication techniques, and nowadays heterogeneous design strategies. However, starting from the late 2010s, when this thesis was conceived, miniaturization problems such as variability, high levels of power densities or uncertainties in the manufacturing processes imposed new restrictions on CMOS design that require innovative and original approaches [Oat14]. In any case, there is a hard physical limit that will mark the end of CMOS scaling capabilities and, although the manufacturers are working hard to extend it as much as possible, it is not in a far away future that the smallest CMOS transistor is fabricated [FMM⁺12].

There is yet another factor that must be considered: electronic systems require to be more and more reliable while, in the contrary, current nanotechnologies display inherent quantum effects that undermine their trustworthiness.

Under these circumstances where the CMOS is near its limit and there is high need of reliable systems, scientists are looking for alternative technologies that will complement, accompany and outlive CMOS. However, although different technologies have been proposed and seem very promising, they all face two strong problems: First, they need to compete against the CMOS fabrication process which has had the industry support over the last decades. Second, emerging technologies are still immature and suffer from serious reliability issues. This hinders the possibility of going beyond the prototype circuit. Among these emerging technologies, we can remark Spin-transfer Torque RAMs (STT-RAMs), Phase Change RAMs (PC-RAMs) or Resistive RAMs (RRAMs), which are promising candidate solutions not only for data storage but standard data paths and neuromorphic computing systems [EBC⁺15].

This thesis is devoted to one of the most promising of these emerging technologies: RRAMs, *memristor* (*memory-resistors*) or *Resistive Switching* (RS), which are different denominations for a set of devices whose resistance varies depending on the device state and stimuli. The memristor was postulated in 1971 by Leon Chua [Chu71], and it was in 2008 when HP labs reported the first device behaving as a memristor [SSSW08]. Since then numerous researchers dedicate big efforts to the manufacturing, characterization and application of these devices.

The memristor possesses unique characteristics that make it an interesting approach for improving current application and conceive novel ones. It is not

just the idea of speeding-up current electronics, but being able to keep the energy consumed by binary information processing and storage, that definitely makes RRAMs one of the candidates destined to play a leading role in the next generation of low-power, high-speed and high-density electronics [EBC⁺15, JKK⁺16]. More specifically, multibit storage, reconfigurable hardware interconnection, or innovative processing architectures, such as hardware implemented neural networks or material implication designs are well adapted to the behavior [EBC⁺15, JKK⁺16] of memristors.

1.1 Motivation

This section gathers the main ideas that inspired the work presented in this thesis.

In the electronic design field there are several hierarchy levels involved in the final circuit fabrication. In particular, from bottom to top, we may distinguish between the technology-manufacturing level, the device level, the circuit level and the architecture level. Independent groups of people work at each of these levels. The traditional work-flow of CMOS design has well established interfaces between these levels [Wes10].

However, emerging technologies are pressing to be included in this work flow, and this implies changes at all the hierarchy levels and their interfaces. For instance, nowadays there is a clear gap between the physical level and device modeling level. A fundamental step in the integration and standardization of an emerging technology is the mathematical description using differential equations of the device behavior that leads to a compact model. These compact models are the mechanism to instantiate physical devices with Simulation Program with Integrated Circuit Emphasis (SPICE) circuit simulators. On the one hand there are research groups dedicated to the development of new compact models [BBB09, YTSP13a]. On the other hand there are research groups devoted to manufacturing new devices that require to be described by a compact model to be included in the circuit design work-flow [PSB⁺09, AP11].

Concerning the model developers, the behavior of memristor devices depends on a wide set of variables, and it also exhibits extremely non-linear dynamics. During the modeling stages, these two factors may lead to reliability and convergence issues. Therefore, *the first motivation of this thesis was the lack of tools and methodologies that support model construction.*

The wide range of materials and processes involved during manufacturing stages leads to a variety of devices displaying different characteristics. This affects the second group of researchers, the device growers. Usually, they require to manually extract the device key characteristics and fit the desired model. Given the huge number of variables involved, this task may be tedious and inaccurate.

Automating the search of a model and a set of parameters that best fit the lab measurements is the second challenge that motivated this thesis.

In addition, current compact models for memristors lack the ability to model complex effects such as variability, life endurance or data retention dependence on the electrical stress and time. Furthermore, the SPICE compact models present in the literature have room for improvement regarding computational load and convergence problems. *Thus, our third motivation is the need for accurate and efficient mechanisms to simulate memristors.*

At circuit level, the traditional characterization of CMOS circuits only considered the effects of single or dual uncertainty sources. Consequently, reliability analyses individually considering such effects may be obviating problematic situations, therefore neglecting possible errors. Moreover, the inclusion of specific effects such as radiation, requires a big amount of manual work that hinders the characterization of big circuits. *Our fourth motivation is the lack of tools automating the analysis of the combined effects caused by multiple uncertainty sources.*

Expanding this issue, traditional work-flows, but more specifically circuit design tools, are only conceived for CMOS technologies, leaving out of consideration emerging technologies. This idea is the spur behind the *fifth motivation, which aims at the necessity of including emerging technologies within standard circuit design flows.*

Leaving circuit level and focusing at system level, the electronic applications based on resistive switching technologies do not usually take into account non-ideal effects such as temperature dependence or degradation. The study of such effects first on circuit modules and then on whole systems is a motive itself. *Thus, the sixth motivation relies on the idea of exploring how memristor non-ideal behaviors may compromise the final application.*

Finally, we cannot avoid being curious about the completely new and attractive systems built around memristors. Hence *finding new applications would be the final motive inspiring this thesis.*

As we have seen, there is an urgent need to incorporate emerging technologies in the standard electronics work flow. New device characterization, circuit design and system verification tools are required to overcome the dramatic effects caused by the reliability problems related to novel technologies. These tools, together with the suitable methodologies, will speed up the already changing electronics design flow. We may then summarize the ideas underlying the motivations of this thesis in the following sentence:

There is a clear demand of models, tools and design strategies to support the inclusion of emerging technologies, such as RRAM devices, in present and future electronic applications.

1.2 Objectives

In view of this change of trend in electronic design, dozens of research groups involving both universities and companies have assigned numerous resources to the development of products based on emerging technologies. Even with all this drawn attention, resistive switching is still an immature technology, and therefore it is a long way to go, and there is a lot of work to be done, before nowadays research helps bringing the next generation of reliable RS applications.

In this compelling context and to provide a complete solution to the design of reliable RRAM based applications, this thesis addresses the matter from three different abstraction levels: device, circuit and system level. We leave out the most physical task of manufacturing RS devices. Instead, we will address the individual device modeling, the RRAM-CMOS circuit simulation, and the RRAM based system design, thus, covering the whole design process. Specifically, the main goals of this thesis are the following:

1. The development of support tools for the characterization of memristor models, filling the gap between circuit simulators and device manufacturing.
2. The development of a SPICE compact model of RRAM devices, fitting *Ox*-RAM and CBRAM cells and exhibiting both bipolar and unipolar behaviors, while considering RRAM intrinsic problems such as temperature dependence, variability and endurance.
3. The development of a design and analyses strategy of circuits containing RRAM devices. This strategy must contemplate these emerging devices together with more traditional CMOS technologies. Additionally, and aiming at reliability, this strategy must be able to study and characterize the combined effects caused by concurrent error sources. Therefore, the developed strategy must provide solution for the shortcomings of traditional work-flows.
4. The development of a simulation and analysis framework for RRAM-CMOS hybrid circuits, supporting the presented design and analysis strategy, and automating the required simulations and analyses.
5. The design and development of a reconfigurable architecture for the reliable write operation of RRAM arrays working in extreme environments, suffering from *Process Voltage and Temperature (PVT)* variations, and meeting variable application requirements.
6. The design of a secure memory architecture, hardened against both software and hardware attacks, as well as to direct tomography inspections.

1.3 Nomenclature

There are certain terms in the thesis that can result somehow ambiguous for the reader. The relative recent manufacturing of resistive switching devices has led to a discussion on the terminology to use when different entities or systems are referred to.

In order to clarify them, we next include a set of definitions, referring their most common use in the scientific literature, as they are used in this thesis:

- **MEMRISTOR:** *memory-resistor*, fourth basic circuit element. This two terminal component, postulated in 1971 by Leon Chua [Chu71], has a non-linear charge-flux characteristic relation. This translates into the variation of its resistance level depending on its previous state and inputs.
- **MEMRISTIVE SYSTEM OR DEVICE:** A broad generalization of *memristors* [CM76], whose dynamics are characterized by Lissajous figures.
- **RESISTIVE SWITCHING:** *RS*, intra-device and reversible resistance change phenomena produced, among others, by the current/voltage stimulus applied through/across two or more of its terminals [LLN15]. Phenomenon that characterizes a *memristor* [VS16].
- **RESISTIVE RAM:** *RRAM*, *ReRAM*, direct application of a *memristor* as a memory element [LLN15, Iel16].

Throughout the thesis the terms *memristor*, *memristive device* and *resistive switching device* are used interchangeably. Even not accurately enough, given that it was the first application of the memristor, the RRAM term is also employed to denote a *memristor* [Iel16] —which may act, among others, as a storage cell. In the present thesis depending on the context it will be clear which meaning we denote.

Three different circuit simulators have been employed for the simulation of the experiments described in Chapters 3, 4 and 5: Cadence Spectre, and the SPICE simulators GnU NGSpice and Linear Technologies LTSpice. These software suites simulate circuit netlists that are composed of sentences describing the analyzed circuit. In this thesis we interact with the inputs, outputs and control mechanisms of these simulators, and thus the following terms may ease the reading of the document:

- **SPICE MODEL:** SPICE sentence(s) describing a circuit element, either a single component or a more complex subcircuit, with a set of default parameters, that allows to be instantiated one or more times within the circuit.
- **SPICE SUBCIRCUIT:** SPICE sentence(s) describing a delimited set of circuit elements and related parameters that grouped form a circuitual module with

a single identity as a result. Similarly to component models, subcircuits can be instantiated multiple times within the circuit.

- **SPICE INSTANCE:** SPICE sentence(s) representing a circuit element located within a netlist node context. All its parameters have been either implicitly or explicitly particularized.
- **SIMULATION SCENARIO:** Individual netlist to be simulated, composed by an organized group of circuit elements —instances— together with a particularized set of environmental, component-related and simulation parameters.

All the above terms, defined for SPICE compatible circuit simulators are applicable to the circuit level simulator Cadence Spectre.

1.4 Contributions

The objectives shown in Section 1.2 have directed the work carried out during the development of this thesis. As a result, the following thesis contributions have already been published, licensed and / or presented at international conferences:

1.4.1 Journal Papers

1. **On the Design and Analysis of Reliable RRAM-CMOS Hybrid Circuits.**

Fernando García-Redondo; Marisa López-Vallejo.

IEEE Transactions on Nanotechnology. Accepted and recommended for publication (April 2017).

2. **Reconfigurable Writing Architecture for Reliable RRAM Operation in Wide Temperature Ranges.**

Fernando García-Redondo; Pablo Royer; Marisa López-Vallejo; Hernán Aparicio; Pablo Ituero; Carlos A. López-Barrio.

IEEE Transactions on Very Large Scale Integration (VLSI) Systems. Volume: 25, Issue: 4, Pages 1224 - 1235. April 2017. DOI: 10.1109/TVLSI.2016.2634083.

3. **SPICE Compact Modeling of Bipolar/Unipolar Memristor Switching Governed by Electrical Thresholds.**

Fernando García-Redondo, Robert P. Gowers, A. Crespo-Yepes, Marisa López-Vallejo, and Liudi Jiang.

IEEE Transactions on Circuits and Systems I: Regular Papers, Volume: 63, Issue: 8, Aug. 2016. DOI: 10.1109/TCSI.2016.2564703. Impact Factor 2015: 2.40

4. **Building Memristor Applications: From Device Model to Circuit Design.**

Fernando García, Marisa López-Vallejo, Pablo Ituero.

IEEE Transactions on Nanotechnology, Volume: 13 , Issue: 6, Pages 1154 - 116. November 2014. DOI: 10.1109/TNANO.2014.2345093. Impact Factor 2014: 1.8.

1.4.2 Licensed Software

1. **Memristor Application Framework**

Public Repository: <https://github.com/fgr1986/maf>

Framework for the characterization and use of memristor compact models.

2. **Circuit Reliability Framework**

Public Repository:

https://github.com/fgr1986/circuit_reliability_framework

Reliability oriented framework for the characterization of CMOS/RRAM hybrid circuits.

3. **VLSI Memristor/RRAM SPICE Compact Model**

Public Repository:

https://github.com/fgr1986/vlsi_memristor_compact_model

SPICE and Spectre model of unipolar and bipolar memristors, considering endurance and variability.

1.4.3 International Conference Papers

1. **CAS-T Lecture: Reconfigurable Writing Architecture for Reliable RRAM Operation in Wide Temperature Ranges.**

Fernando García-Redondo; Pablo Royer; Marisa López-Vallejo; Hernán Aparicio; Pablo Ituero; Carlos A. López-Barrio.

IEEE international Symposium on Circuits and Systems, (ISCAS 2017), Baltimore, MD, Usa, May 28-31 2017.

2. **CAS-T Lecture: SPICE Compact Modeling of Bipolar/Unipolar Memristor Switching Governed by Electrical Thresholds.** *Fernando García-Redondo, Robert P. Gowers, A. Crespo-Yepes, Marisa López-Vallejo, and Liudi Jiang.*

IEEE international Symposium on Circuits and Systems, (ISCAS 2017), Baltimore, MD, Usa, May 28-31 2017.

3. Characterization of Analog Modules: Reliability Analyses of Radiation, Temperature and Variations Effects.

Fernando García-Redondo, Hernán Aparicio-Cerqueira, Marisa López-Vallejo, Pablo Ituero and Carlos A. López Barrio.

Conference on Design of Circuits and Integrated Systems 2016 (DCIS 2016) Granada(Spain) 23-25 November 2016.

4. Reliable Design Methodology: The Combined Effect of Radiation, Variability and Temperature.

Fernando García-Redondo, Marisa Lopez-Vallejo, Hernán Hernán Aparicio Cerqueira, Pablo Ituero.

13th IEEE International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD 2016), Lisbon (Portugal). 27 - 30 June 2016.

5. Taxonomy of Power Supply Monitors and Integration Challenges.

Pablo Ituero, Marisa López-Vallejo, Hernán Aparicio, Fernando Garcia-Redondo.
Mixed-Signal Testing Workshop (IMSTW), 2016 IEEE 21st International.

6. A Thermal Adaptive Scheme for Reliable Write Operation on RRAM Based Architectures.

Fernando García-Redondo, Marisa López-Vallejo, Pablo Ituero.

Conference: 2015 33rd IEEE International Conference on Computer Design (ICCD), October 2015, New York.

7. Evolution of radiation-induced soft errors in FinFET SRAMs under process variations beyond 22nm.

Pablo Royer, Fernando García-Redondo, Marisa López-Vallejo.

Conference: Nanoscale Architectures (NANOARCH), 2015 IEEE/ACM International Symposium on. Boston, MA, USA (July 2015).

8. A Tool for the Automatic Analysis of Single Events Effects on Electronic Circuits.

Fernando García-Redondo, Marisa López-Vallejo, Pablo Royer, Javier Agustín.
5th European Workshop on CMOS Variability, PATMOS VARI 2014, Mallorca (Spain).

9. Four-Injector Variability Modeling of FinFET Predictive Technology Models.

Pablo Royer, Marisa López-Vallejo, Fernando García-Redondo, Carlos A. López Barrio.

5th European Workshop on CMOS Variability, PATMOS VARI 2014, Mallorca (Spain).

10. Model Validation and Simulation Framework for Novel Nanometer Devices.

Fernando García-Redondo, Marisa López-Vallejo, Pablo Ituero.

Conference on Design of Circuits and Integrated Systems 2012 (DCIS 2012) Avignon (France).

11. A CAD Framework for the Characterization and Use of Memristor Models.

Fernando García, Marisa López-Vallejo, and Pablo Ituero.

International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design 2012 (SMACD2012) Seville, Spain. 19 - 21 September 2012.

12. Temperature Sensor Placement Including Routing Overhead and Sampling Inaccuracies. *Pablo Ituero, Fernando García, and Marisa López-Vallejo.*

International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design 2012 (SMACD2012) Seville, Spain. 19 - 21 September 2012.

1.5 Document Structure

Synthesizing the thesis contents, Figure 1.1 represents the thesis scope, organized in chapters. Chapter 1 has been omitted given its introductory character. The diagram follows a bottom-up structure, where each layer represents a different abstraction level. The topics addressed by this thesis leave out of the scope the RS device manufacturing, and start just after it.

Chapter 2 introduces the memristor device concept, presents its mathematical nature and describes the characteristics of different physical manufactured devices. The best well-known memristor-based applications are also exposed. The chapter concludes with the main challenges to be addressed regarding RS technologies.

Chapter 3 describes the developed memristor simulator. This simulator is able to characterize RS devices depending on its behavioral definition and characteristics, establishing the influence of the model parameters, the effects caused by

Figure 1.1: Diagram displaying the reliability concepts addressed in each chapter of this thesis, organized by abstraction level.

variability and, studying its dynamics, evaluating the model reliability. Therefore, the developed simulator is conceived as a useful tool to compare different memristor models, establish their parameter boundaries and find the model that best suits the experimental device measures.

Chapter 4 presents a state of the art, fully SPICE-compatible memristor model, fitting both unipolar and bipolar devices, while being able to emulate physical aspects such as temperature dependence, variability or device degradation. The validation against two different physical memristors is also presented.

Chapter 5 tackles with the reliability issues located at the circuit abstraction level, and presents a strategy for the definition of analyses focused on the design of reliable solutions that use RRAM-CMOS hybrid circuits. It starts by presenting the challenges that need to be addressed by nowadays circuit designers, reviewing the most common uncertainty sources attached to both RS and small CMOS node technologies. After that, the three stage solution, based on the efficient definition of models, metrics and design space is presented, followed by the simulation and analysis framework that has been developed to automate each analysis required by the proposed strategy. We later present an experiment to validate the framework results against Static RAM (SRAM) experimental data. This validation stage is followed by two cases of study, involving the characterization of both digital and analog modules. Finally, a thorough case of study involving the reliability and performance characterization of an RRAM read driver illustrates the utility of this strategy.

The results of applying the developed tools, analyses strategies and methodologies are gathered in Chapters 6 and 7. First, we present a reconfigurable driver architecture that allows different writing approaches to be dynamically applied. By using this architecture, and depending on the working scenario and application needs we can employ reliability oriented profiles targeting power reduction the increment of durability or the improvement of reliability. Moreover, we fight the

RS operation dependence on temperature, and guarantee the correct behavior of the driven cells in a wide range of temperatures. Secondly, and also following the presented design guidelines, we have designed a secure RRAM crossbar architecture, with the ability of self-erasing in case of hack or attack.

Finally, Chapter 8 summarizes the main contributions of the thesis and establishes some guidelines for future continuation of the ideas presented here.

Just Shut Up And Take My Money.

Fry, Futurama.

2

Introduction to Resistive Switching Technologies

Postulated by Leon Chua in 1971 [Chu71], the memristor is the fourth basic electrical component —along with resistors, capacitances and inductances— displaying a predictable relationship between its resistance and the electrical charge traversing it. Depending on the direction of the current flowing through the device, its resistance increases or decreases. When unbiased, the device acts as a nonvolatile memory, storing the actual resistance value until once again a voltage stimulates it. In a binary-value memristor, two resistance values are stored. The first one, R_{ON} is related to the *LRS*, while R_{OFF} refers the *HRS*, being R_{ON} orders of magnitude lower than R_{OFF} . Although in 1968 resistive switching phenomena had been reported in TiO_2 thin films [Arg68], it was not until the year 2008 when Stanley Williams’ research group from Hewlett-Packard Labs officially presented a device also fabricated using layered TiO_2 , as the first actual implementation of a memristor [SSSW08].

Figure 2.1: Chapter 2 scope and perspective within the Thesis.

After such discovery several memristor based applications appeared. Memristors opened up a whole new world of possibilities. Their scalability abilities, low power consumption, nonvolatile storage, or their compatibility with standard CMOS process attracted the attention of the research community.

Therefore, since HP Labs' discovery, many academia and industry research groups have been working on the development of memristive structures using both semiconductors and, most often, metal oxides as isolator materials. The exponential growth of memristor research has required of several and solid reviews [MKW12, EBC⁺15, LLN15, CLC⁺15, YC16] to gather and organize the scientific contributions and challenges of RRAM technologies.

Even the most controversial aspects, regarding the complex atomic level physics under the resistive switching mechanisms have been thoroughly studied. Those studies clarify the basic processes that provide memristors with nonvolatility characteristics. In this area there are remarkable works [LLN15, Iel16, KB16] condensing the laborious task of in detail describing and modeling the switching processes that take place at atomic level.

Thus, as Figure 2.1 displays, to help understanding why the memristor is shaking up nowadays electronics this Chapter provides an introduction to the resistive switching technologies context. First Chua's concept of memristor is introduced, followed by the enumeration of the most common physical technologies experimenting resistive switching. After that, the best well known memristor models will be introduced. This review of models is followed by the description of the most common memristive based applications. The last point of the Chapter encompasses the main challenges to be addressed in order to bring this technology to final consumer products.

(b) Electrical magnitudes and electronic components relating them [Chu71].

Figure 2.2: The memristor and involved electrical magnitudes.

2.1 Chua's Memristor

As conceived by Leon Chua [Chu71], the *memristor*, represented by the symbols shown in Figure 2.2a, is an electrical component ruled by a non-linear relation between its charge and its flux. Figure 2.2b, represents the main electrical magnitudes, and how circuit components relate them. In this symmetric scheme, the memristor complements the set of classical passive devices: resistor, capacitor and coil; devices described by their key electrical properties the resistance R , the capacitance C and the inductance L , respectively.

The first of the electrical properties of the memristor is the *memristance*. The memristance can be understood as the capacity for which the device can remember the last resistance (ratio of voltage and current) it had before it was disconnected (from any source of voltage and / or current). The second electrical property, the *memconductance*, stands to the inverse of the memristance. Similarly, it may be understood as the property for which the device can remember the last conductance it had before it was disconnected from the source.

As we introduced in Figure 2.2b, the memristor gets defined by its non-linear charge-flux relation,

$$g_m(\varphi_{mr}, q_{mr}) = 0, \quad (2.1)$$

being v_{mr} and i_{mr} the voltage across and the current through the memristor, and $\varphi'_{mr}(t) = v_{mr}(t)$, $q'_{mr}(t) = i_{mr}(t)$ the flux and charge standard electromagnetic relations.

In case the memristor is charge / flux controlled, this relationship takes the form

$$\varphi_{mr} = \phi(q_{mr}) \quad (2.2)$$

or

$$q_{mr} = \sigma(\varphi_{mr}) \quad (2.3)$$

respectively.

In a charge controlled memristor [Chu71], the device is governed by

$$q'_{mr} = i_{mr} \quad (2.4a)$$

$$v_{mr} = R_{mr}(q_{mr})i_{mr}, \quad (2.4b)$$

while if we refer to a memristive system ([CM76, DPCV09]), the equations are

$$x' = f(x, i_{mr}) \quad (2.4c)$$

$$v_{mr} = R_{mr}(x, i_{mr})i_{mr}. \quad (2.4d)$$

In these equations, x is the vector representing the internal state variables.

The equations referred to flux controlled memristors or memristive systems are

$$\varphi'_{mr} = v_{mr} \quad (2.5a)$$

$$i_{mr} = G_{mr}(\varphi_{mr})v_{mr}, \quad (2.5b)$$

for the memristors, and

$$y' = g(y, v_{mr}) \quad (2.5c)$$

$$i_{mr} = G_{mr}(y, v_{mr})v_{mr} \quad (2.5d)$$

for the mem-systems. In that case, y corresponds to the dependent state vectors which define the memconductance.

Following the nomenclature of [RT10, GRR12], memristance and memconductance get defined by

$$R_m(q_{mr}) = \frac{d\phi(q_{mr})}{dq_{mr}} \quad (2.6)$$

for charge-controlled memristors, and

$$G_m(\varphi_{mr}) = \frac{d\sigma(\varphi_{mr})}{d\varphi_{mr}} \quad (2.7)$$

for flux-controlled memristors.

Figure 2.3: Schematic of Ox-RAM and CBRAM in their three main states.

If the last applications are invertible, then

$$G_m(\varphi_{mr})^{-1} = R_m(\sigma(\varphi_{mr})). \quad (2.8)$$

These equation systems match most devices' behavior [GRLVI14], and thereby both physical and phenomenological models specify both state and conductive equations defining how the model runs [RT10, GRR12]. Further analysis of the different equations that define the behavior of the models presented in literature can be found in Section 2.3.

2.2 Resistive Switching Technologies

2.2.1 Device Structure and Behavioral Operation

The manufacturing of the first TiO_2 based bipolar memristor by Hewlett-Packard Labs triggered a race for the identification of layered materials that, trapped among suitable electrodes, exhibit memristive properties and behavior. RRAM devices consist of an insulating layer -either metal oxide or a solid electrolyte semiconductor- placed between a *Top Electrode (TE)* and a *Bottom Electrode (BE)*. Both TE and BE electrodes are traditionally built using different metallic depositions, either single layered or as stacks.

Figure 2.3 shows the schematic structure and three main operation states of a physical memristor device. Prior its *electroforming*, the device is initially restrained to the *Pristine State*. The device's electroforming process consists in the generation of one or more *CFs* formed by dielectric breakdown. It is to be noted that the

number and morphology of such CFs depend on the device's materials, sizing and input stimuli. Given that the electrical and reliability characteristics of the device depend on such CFs, the understanding of the underlying physics is a subject of great interest [LLN15, Iel16, KB16]. For simplicity, we will refer to both single instances and to the plausible tree of CFs as an individual CF.

With the CF correctly developed, electrically connecting TE with BE, the device manifests an increased conductance by placing in parallel to the insulating layer a usually ohmic-like low resistance path. This device's high conductance state is denominated LRS. The HRS is achieved after the reset operation takes place. Here the CF is dissolved, leading to an extremely low electric conductance between both electrodes. Usually, R_{ON} and R_{OFF} terms refer to the resistance values achieved at LRS and HRS respectively.

The key idea toward which all memristive based applications are built relies on the eligibility to control this CF growing and dissolving. This gives the resistive switching technologies the ability to recover from LRS to HRS throughout multiple cycles. The maximum number of cycles the device is able to cycle through without significantly reducing its recover abilities is called device lifetime, defining its endurance.

2.2.2 Categories of RS Technologies

Traditionally, memristor physical devices can be classified into two different sub-categories [Iel16]: *Ox-RAM* and *CBRAM*. Figure 2.3 describes the key differences between the two of them. Basically, *Ox-RAM*'s filament consists of oxygen vacancies organized within the metal oxide layer while *CBRAM*'s CF is formed by metal atoms such as fast-diffusive *Ag*, *Pt* or *Cu* ions [YC16]. The terms *Valence Change Memory (VCM)* and *Electro Chemical Memory (ECM)* establish a parallel terminology for *Ox-RAM* and *CBRAM* respectively [MKW12].

Leaving out the differences of the underlying physics, these two types of devices share in common multiple characteristics: unipolar and bipolar switching, compatibility with mainstream silicon CMOS technology process, or even radiation hardening capabilities [MHP⁺14, MFH⁺15]. Regarding the most notable differences, and under a gross generalization assumption, in *Ox-RAM* cells the R_{ON}/R_{OFF} ratio may be smaller than in *CBRAM* cells. On the other hand, regarding cycling endurance, *Ox-RAM* cells usually give wider lifetime spans than *CBRAM*.

A myriad of materials has demonstrated RS behaviors, although the most common ones involve TiO_x [PSB⁺09, AP11], HfO_x [ABGI14, Iel16] or TaO_x [ZGJ⁺15, KCLL14] for *Ox-RAM* cells and a wide variety including carbides as *aSiC* [ZJHD14], chalcogenides such as *GeSe* [YW11] or oxides as SiO_x [MMB⁺15]

Figure 2.4: Memristor switching modes depending on its polarity. The compliance current I_c limits the maximum current flowing through the device after the SET process.

or Al_2O_x [BDF⁺14] for CBRAM. A wider and most thorough review of the materials used, together with their characteristics can be found in [EBC⁺15, LLN15].

Finally, a third less common memristor subcategory, corresponds to nonfilamentary devices [EBC⁺15, LLN15, Iel16]. In this subcategory, uniform resistivity changes occur across the entire device active area. Besides the non-filamentary conductions, two key differences are to be noted compared against filamentary devices: first, no electroforming stage is required; and second, the conductance of these devices -mostly manganites, cuprates and titanates- highly depends on the active area.

2.2.3 Switching Mechanisms Summary

Two different operation methods produce resistive switching, differing in the polarity of both set and reset operations: unipolar and bipolar. Figure 2.4 describes how both set and reset processes take place under unipolar positive voltage; while for bipolar devices set and reset processes must alternate their polarity.

Unipolar switching processes have been explained by the purely thermal acceleration of redox transitions, producing the connection and disconnection of the device CF with the corresponding electrode. Due to the high temperatures achieved during set transitions, the chemical reduction of the metal oxides produces the CF growth. During the reset, it is the oxidation of the metallic filament, highly dependent on the temperature achieved inside the CF, that produces its dissolution. Additionally, the electric field applied across the device electrodes plays a key role as a threshold triggering the set transition. This is due to the local formation of an electronic filament supporting local Joule heating for redox reactions [Iel16].

In contrast to unipolar switching, bipolar switching has been explained in terms of ionic migration, assisted by the temperature and the electric field [Iel16]. During set transitions, ionized defects form the CF, shaping it depending on the device's thermal maps. On the contrary, and also following temperature driven patterns, reset processes involve the migration of the ion defects towards the negatively biased electrode, destroying the CF and increasing the gap with the electrode.

Additionally, it is to be noted that, as Figure 2.4 describes, in both unipolar and bipolar switching set processes, in order to control the CF size and avoid destructive breakdowns, currents must be limited by a *compliance current* I_c .

Further information regarding switching mechanisms can be found in the review works [LLN15, EBC⁺15, Iel16, KB16], where the physical mechanisms involved in both transitions are described in depth.

2.2.4 Conduction Processes Summary

The discrete (either binary or multi-level) or continuous resistance values a memristor can experiment are defined by the different conduction processes. These processes depend on environmental conditions like temperature, material characteristics like barrier heights or activation energies, and the applied input stimuli.

The electrical characteristics of the device are then subjugated to the conduction processes happening simultaneously inside both CF and insulator. In most cases, the voltage-current relationship is described by basic conduction processes. However, many different processes may occur at the same time, being some of them dominant over the others depending on the working conditions —such as circuit temperature or electric field— and on the memristor state [YLH⁺15, GRGCY⁺16].

Table 2.1 describes the main conduction processes reported to occur in different RRAM devices [SN07, VGJM⁺14, EBC⁺15, GRGCY⁺16, KB16, LLN15], including both Ox-RAM and CBRAM cells. This summary addresses the current density (J) dependence on both temperature, electric field and applied voltage for each conduction process. The variables, parameters and constants describing the equations involved in the conduction processes can be found in Table 2.2.

Finally, we would like to highlight that to read the resistance value stored in an RRAM device, translating it to the corresponding voltage value, an input voltage or current is to be applied. Therefore, depending on the involved conductive processes, this read value will vary with the applied stimuli subjugated by the relationships described in Table 2.1.

Table 2.1: Basic conduction processes in memristor electrode-insulators contacts and their relation with voltage, electric field and temperature.

Process	Expression
Generalized Trap Assisted	$J \propto a_1 \left(\frac{V}{d}\right)^{a_2}$
Generalized Tunneling (I)	$J \propto \exp\left(\frac{-g}{g_0}\right) \sinh\left(\frac{V}{V_0}\right)$
Generalized Tunneling Temperature Dependent (II)	$J \propto \exp\left(\frac{-g}{g_1}\right) \sinh\left(\frac{V}{kTV_1}\right)$
Tunneling	$J \propto \mathcal{C}^2 \exp\left[-\frac{4\sqrt{2m^*}(q\phi_B)^{3/2}}{3q\hbar\mathcal{C}}\right]$
Schottky Thermionic Emission	$J = A^{**}T^2 \exp\left[-\frac{q\left(\phi_B - \sqrt{\frac{q\mathcal{C}}{4\pi\epsilon}}\right)}{kT}\right]$
Poole-Frenkel Emission	$J \propto \mathcal{C} \left[-\frac{q\left(\phi_B - \sqrt{\frac{q\mathcal{C}}{\pi\epsilon}}\right)}{kT}\right]$
Ohmic	$J \propto \mathcal{C} \exp\left[-\frac{\Delta E_{ac}}{kT}\right]$
Ionic Conduction	$J \propto \frac{\mathcal{C}}{T} \exp\left[-\frac{\Delta E_{ai}}{kT}\right]$
Space-charge-limited	$J \propto \frac{9\epsilon\mu V^2}{8d^3}$
QPC	$I = \frac{2e}{h} \left\{ eV + \frac{1}{\alpha} \ln \left[\frac{1 + \exp\left(\alpha(\phi_B - \beta eV)\right)}{1 + \exp\left(\alpha(\phi_B + (1 - \beta)eV)\right)} \right] \right\}$

Table 2.2: Variables and parameters for the expressions shown in Table 2.1.

Parameter/Magnitude	Description
V	Voltage applied between the electrodes.
T	Temperature.
d	Insulator thickness.
$\mathcal{E}$	Electric field in insulator.
g	Gap distance.
a_i, g_j, V_k	Matching parameters.
ϕ	Material barrier height.
ϵ	Insulator permittivity.
A^{**}	Effective Richardson constant.
e	Electron charge.
h	Plank constant.
$\hbar$	Reduced Planck constant (Dirac constant).
ΔE_{ac}	Activation energy of electrons.
ΔE_{ai}	Activation energy of ions.
α	Parameter related to barrier thickness.
β	Parameter related to the fraction of dropped voltage.

2.2.5 Multi-level (ML) and Multi-bit (MB) capabilities

Multi-Level (ML) capability in resistive switching devices refers to the ability of a RS device to be programed in a multiple set of values between its HRS and LRS [KCS⁺16, LLW⁺16, SBB⁺16, KLK⁺15, KNC⁺15, PDS⁺15]. This ability can either refer a discrete set of states S_i or a continuous function —therefore achieving every possible state between HRS and LRS states. When functioning as a two terminal device, ML resistance states can be seen as *Multi-Bit (MB)* capabilities. As an illustrative example, a memory cell being able to be programed between a set of 8 resistance levels, is said to behave as a 3-bit memory.

Specially important regarding non-conventional computing approaches, as neural networks, is the ability to be programed within a continuous range of states. Consequently, the device's state can be controlled producing a non-linear but continuous function at the output, similar to the response neurons produce.

We would like to remark that in order to achieve specific ML states, the programing strategies differ from those that only allow either HRS or LRS states to be programed. Usually, such ML programing strategies require progressive pulses to be applied to the cells.

2.2.6 RRAM-CMOS Hybrid Chips

The progressive migration of the current applications toward memristive designs, together with the implementation of novel approaches requires the integration of RRAM subcircuits or modules within CMOS architectures, resulting in RRAM-CMOS hybrid circuits.

The most common memristive structures are fully compatible with standard CMOS processes, allowing to grow memristors on top of CMOS transistor contacts [JKK⁺16, CLC⁺15, Iel16, YC16]. The most common RRAM-CMOS device integration comes as the addition of an access transistor to a memristor, developing *1-Transistor-1-Memristor (1T1R)* RRAM cells, proving how easily RRAM-CMOS hybrid structures may interact.

Following with this idea, and as many other applications —summarized in Section 2.4— the second major memristive application requires to adapt current neuromorphic circuit architectures to include RRAM based crossbars [JKK⁺16]. In this sense, every memristive system benefits from RRAM technology scaling and easy CMOS integration. Cells size of $6F^2$ —with F referring to the process size— can be achieved using *1T1R* structures, while this area can be reduced to $4F^2$ using a crossbar architecture with selector devices [Iel16].

Crossbar circuits [CLC⁺15, Iel16, YC16] with different cell structures and/or device materials have been reported. High speed and large capacity (up to 16 Gb by Micron/Sony or 32 Gb by ScanDisk/Toshiba) RRAM prototype chips have been achieved [YC16], reassuring the resistive switching technologies capabilities regarding future large storage-processing crossbar applications.

More importantly, a new research thread, focusing on the high density integration of devices, has achieved two different 3D architectures [EBC⁺15, Iel16, YC16]: stacked 3D horizontal cross-point arrays and 3D vertical cross-point arrays. The first one is based on stacking conventional 2D horizontal cross-point arrays, while the second vertical structure is composed of RRAM devices sandwiched between pillar and multilayer plane electrodes.

However, even if the future of electronics seems promising and shining, most applications are still in initial development stages. As Section 2.5 involves, the added challenge of solving not only RRAM but CMOS short node technology problems, complicates the long path to go before reliable applications reach the final consumer.

2.3 Memristor Models

Since its first implementation, a wide percentage of research resources have been dedicated to the accurate modeling of memristors.

The complex mechanisms involved in memristor behavior have diversified the modeling lines, each of which aims at different aspects. Therefore, in the literature we find switching dynamics models, electrical models and compact models. Additionally, the models gathered within these categories can fit the characteristics of *Ox*-RAM, CBRAM, or be common to both types of device.

However, the following classification responds to the abstraction level the model aims to characterize and emulate. This allows us to get a better perspective on the issue, and to review them depending on the modeling goal, starting with atomistic approaches and ending with circuital schemes.

2.3.1 Dynamic Simulation Models

On the atomic scale, simulation methods such as density functional theory (DFT) and molecular dynamics (MD) are used [AVS⁺16, PLPP16, KB16]. This approach is, in terms of computational load, extremely expensive. Consequently, instead of solving the whole switching process, these simulations are more commonly limited to problems such as LRS and HRS electronic structure variations and/or the calculus of migration barriers [Was16].

2.3.2 Kinetic Monte Carlo Simulation Models

Kinetic Monte Carlo (KMC) simulation methods are the bridge between atomistic and continuum methods [KB16]. Usually, due to the statistical nature of the KMC method, it can be used to study the variability and failure statistics [Was16].

2.3.3 Continuum Models

Continuum models are based on partial differential equations, that once solved describe the device's behavior. Memristor devices are characterized by Poisson equations together with drift-diffusion equations [Was16, KB16]. Regarding thermal characterization, which is a key issue since most switching and conduction processes depend on temperature, these models consider local Joule heating. To compute the partial differential equations, the simulators use finite element method (FEM), finite difference method (FDM), or finite volume method (FVM).

Continuum models usually are the step taken before the development of compact models [KB16]. Thus, the continuity equations that allow to analyze mechanisms not directly observable in experimental data also serve as physics basis during the development of accurate analytical and compact models.

Figure 2.5: Modeling approaches: compact versus DFT, KMC or continuum models.

2.3.4 Percolation Models

Percolation approaches describe the behavior of connected clusters in random graphs. Both LRS and HRS states can be considered as the appearance and disappearance of a percolating cluster in a resistor network. Thus, as gathered by [LLN15, Iel16] review articles, several percolation-like trap-and-thermal assisted models appeared, modeling the resistive switching process through conductive trees connecting the electrodes of the device.

2.3.5 Compact Models

Each of the above approaches is more accurate —even some target limited— than compact models. However, the computational load associated with the simulation of memristive devices at circuit levels leave them out of consideration during circuit design stages.

Described by Figure 2.5, a compact model is a system of parameterized equations describing the behavior of the modeled device. Therefore, the set of equations defining the behavior of a memristor model must meet the system (2.1), introduced in Section 2.1.

Circuit simulators considering analog components, from SPICE (and its variants HSPICE, PSPICE etc.) to Verilog-a, make use of compact models to compute how the devices interact whenever electrical magnitudes are involved. However, the variety of distinct system applications integrating memristors leads to some different requirements for each design. For example, a small variation of the size of a given circuit can result in a substantial increase in the simulation time. Therefore, an effective memristor model will be the one which satisfies several requirements, achieving a good trade-off between accuracy and computation efficiency.

Chapter 3 describes in detail the main characteristics of the best well known models presented in literature.

2.4 Resistive Switching Applications

The characteristics of RS devices have pushed application architects and circuit designers towards the improvement and conception of conventional and novel applications. Thus, beyond the conventional applications a myriad of previously non-considered uses of Nonvolatile Memory (NVM) are arising. This section introduces the most common applications making use of RS technologies.

2.4.1 Conventional Applications

Memory Applications

The most straight-forward application of memristor devices is information storing. Giving its low-operation power, high-density and speed, memristors are definitely called to play a key role in the next generation of memories. To the date different companies such as Micron, Sony, ScanDisk or Toshiba have reported completely functional memory solutions up to 32 Gb [YC16]. Both traditional 2D but also the novel 3D integration approaches [CLY16] have great potential to be used not only in massive storage blocks but as fast micro-controller caches and other specialized memories [CLC⁺15].

Further designs such as memristor based *Content Addressable Memory* (CAM) cells have been proposed, achieving 256b-word length Ternary CAM (TCAM) with search time as low as 1ns fabricated using 2.5T1R cell structure [LHL⁺16].

Another attractive characteristic of RS technologies relies on the ability of most resistive memory cells to develop not only two but multiple resistance levels. This phenomenon is referred to as *ML* or *MB* capabilities. Thus, a cell can exhibit more than two resistance states between LRS and HRS, which can be used to store multiple bits within a single memory [LLN15, KKK⁺15].

There is a huge amount of works in the literature describing the progress made in this area. More information can be found in the reviews [MKW12, EBC⁺15, LLN15, CLC⁺15, YC16, VS16].

Basic RRAM Cell Structures Four basic RRAM cell structures are widely used in large RRAM arrays. Each cell structure has inherent advantages and disadvantages, which will be discussed in Section 2.5.

- *1-Memristor (1R)* Cells, composed of a single memristor, are the basic structure forming large cross-point arrays. As introduced before, with an area as low as $4F^2$, using this approach high integration densities are achieved.
- *1-Transistor-1-Memristor (1T1R)* Cells, composed of a cell selection transistor in series with a single memristor. This scheme allows the isolation of the

selected cell from the other unselected cells present in the array. The *World Line* (WL) controls the transistor's gate. Therefore, together with the *Bit Line* (BL), the tuning of the WL voltage controls the write current that is delivered to the cell. This scheme can be developed consuming $6F^2$ area.

- *1-Selector-1-Memristor (1S1R)* Cells, composed of a two terminal selector device -simple diodes or more complex layered components- in series with a single memristor [CKC⁺16, AJDG16]. This scheme improves 1R cell characteristics.
- *Complementary Resistive Switches (CRS)* Cells consist of two opposed memristors in series, pairing the BE of the first of them with the BE of the second [YMO⁺15]. With this cell structure, different architecture level problems are minimized.

FPGA Architectures

As well as storage applications, *Field Programmable Gate Array* (FPGA) may benefit from RS technologies. Here, memristors do not only replace the traditional SRAM *Look Up Tables* (LUTs), but they can also replace the transmission gates and be used to propagate data-path signals.

The huge area interconnection overhead, together with its volatility, is one of the main concerns in standard SRAM based FPGAs, resulting in high power consumption and slow operation speed. RRAM based FPGAs, usually denoted by *rFPGA*, *mrFPGA* or simply *RRAM-FPGA*, make use of the high ratios between RRAM resistance levels and zero standby consumption to give birth to more efficient and faster FPGA applications [TKGD16, HZHL16, JKK⁺16]

Common Digital and Analog Designs

Reviews [PV10, MKW12, EBC⁺15, LLN15, CLC⁺15, YC16, VS16] describe the taxonomy of conventional applications that have been ported to memristive circuits. A wide variety of standard applications have been reported, mostly digital but also analog circuits. Memristive digital applications include [PV10, MKW12]:

- Threshold comparators.
- Schmitt triggers.
- Difference amplifiers.
- Multistate pipeline registers.
- Flip Flops.

Within the analog processing domain, the most common memristive applications include [MKW12]:

- Chaos Circuits.
- Filters.
- Biomedical -integrator- circuits.

Security Primitives

RRAM stochastic characteristics, carefully detailed in Section 2.5, compromise the behavior of the circuit. However, designers have taken advantage of this variable behavior, proposing RRAM alternatives to random pattern generation subcircuits working as security primitives.

The variability appearing in switching processes, as well as the fluctuations experienced during read processes, may then define the randomness designers are looking for. As a result, recent works like [BAC⁺16, GCLY16, DSB16] use RRAM cells to design such as *Physical Unclonable Function (PUF)* and *Random Number Generator (RNG)* primitives.

2.4.2 Novel Computing Paradigms

Late years novel computation solutions have appeared to address the problem of processing the extremely large data streams coming from the present generation of big data and IoT information sources. This volume of data, unconceivable time ago, required out of the box solutions that have given birth to different computing paradigms. Memristor based applications come again as a natural solution to implement those computing architectures.

Figure 2.6 represents a simplified classification of nowadays paradigms of computing. The traditional *von Neumann Architecture (vNA)* is composed of general purpose processors —*central processing unit (CPU)*— that together with *application specific processors (ASP)* interact with the memory blocks [Aus15]. Halfway between vNA and *non-von Neumann Architectures (non-vNA)* *stateful implication logic architectures* based on memristors can be implemented either as vNA or as non-vNA architectures, and it is the first step to integrate computation capabilities directly in memory.

On the other hand, the design of pure non-vNA architectures using memristors is currently under the spotlight [JKK⁺16]. *Stateful Logic Computing* modules, but more generally *Computation In Memory (CIM)* are architectures that allow specific functions (either boolean logic or more complex such as large matrix

Figure 2.6: Computing paradigms scheme.

multiplication) to be computed within the memory. Finally, the so called *approximate computing* (AP), *cognitive computing*, *neuromemristive computing* or most commonly, *neuromorphic computing*, refer to the brain-like novel approaches.

Memristor Based Logic Computing

Memristor-based logic circuit design concepts [VS16] radically change with the traditional computing paradigm. We could define *Memristor Based Logic* or *Stateful Logic* as the architectures that simultaneously use memristors as both storage element and basic in-place processing unit [ZYW⁺13]. Consequently, in-memory boolean operations can be accomplished and store this operation value at the same place. Once again, low power consumption, speed and nonvolatility characteristics make the memristor the natural component to be used in this computing scheme. Depending on its practical implementation, *memristor based logic* can be categorized as a von Newman Architecture, or be included in the definition of *Computation-In-Memory* [JKK⁺16]. To the present day, a wide range of stateful logic approaches has been presented [JKK⁺16]. We can subdivide these architectures between *Boolean Logic*, *Implication Logic* or *Threshold/Majority*. The works [VS16, JKK⁺16] carefully review the best known approaches/frameworks presented in the literature: *IMPLY*, *MRL*, *MAGIC*, *Threshold logic*, *CMOS-like* and *Parallel input processing*. Each of the above differs in the basic unit computing the operation (memristor, memristor+resistor, memristor+diodes), the computation procedure, and the auxiliary circuitry involved.

Computation-In-Memory

From a more general perspective, CIM or *Memory-Intensive Architecture* solutions define the architectures in which the data storage and the processing are indistinguishable processes that take place within memory structures [JKK⁺16]. Following this definition we can encompass the previously described *stateful logic* modules as CIM architectures. However, we must remark that CIM solutions are not restricted to compute boolean logic functions, On the contrary complex (analog/digital) functions such as matrix arithmetic operations, including additions or multiplications can be implemented using CIM architectures [SMC⁺15, HXN⁺15, HYN⁺16, PLA⁺16, NHL⁺17, JKW17]. These architectures are called to overcome the limitations of standard CMOS von Neumann Architectures, promising high throughput and energy efficiency. Once the challenges ahead are solved, CIM architectures are then expected to lead the next computing architectures revolution.

Neuromorphic Computing

Neuromorphic systems were firstly conceived at the end of the 1980s as silicon-based analog circuits that function as the building blocks of a *spiking neural network* (SNN) [JKK⁺16]. Since then, different schemes involving FPGAs and ad-hoc *neural processor unit* (NPU) ASICs have been used in this non von Neumann computing architecture. Compared against traditional CPU approaches, a neuromorphic system is highly asynchronous, which implies the asynchronous operation of each neuronal circuit only when it is triggered by external inputs. Among other advantages such as noise tolerance, neuromorphic architectures are able to handle NP problems otherwise, using traditional approaches, impossible to solve. Therefore classification problems, pattern recognition or approximate computing —filtering, analog or multistate multiplication— are the most common applications to be accomplished by neuromorphic processors. Big efforts such as the *True North* from IBM or the NPU processor developed by Qualcomm highlight the importance of these novel computation approaches.

Neuromorphic processors are built using a huge amount of basic units acting like neurons. Consequently, the interactions between adjacent neuron-units emulate the biological synapses. In this sense, to implement learning protocols in neuromorphic systems, artificial synaptics procedures must be firstly accomplished before the system is set up. Memristors, in contrast to the other passive components or transistors, have been successfully employed to simulate this behavior. Their analog response, threshold-like behavior and signal integrating capabilities, make memristor technologies the natural solution to emulate neural networks [JKK⁺16, VS16].

2.5 Reliability Issues in RRAM-CMOS Hybrid Circuits

Years of experience in CMOS circuit design have served to identify PVT variations, radiation and aging as the most concerning issues in nowadays CMOS circuits [KBK⁺11, FHS15, SBK15]. The study of their related problems is completely established in standard design methodologies, and commercial CAD tools provide support for the characterization of the effects caused by PVT, and in some cases, aging.

However, RRAM is an immature technology compared to CMOS. In the same way variability greatly affects CMOS performance and reliability, RRAM suffers, among others, from intrinsic problems such as limited lifetime [CSL⁺16], variability [PLPP16], fixed-time data retention [CTKP16], or alteration of stored data during write and read operations of half selected cells [CLY16].

To help understand the technology challenges researchers are facing, this section reviews the most common uncertainty sources that may compromise RRAM-CMOS hybrid circuits.

2.5.1 Introduction to Uncertainty Sources in CMOS Technologies

Nowadays the ever pressing circuit miniaturization is inseparably attached to reliability issues. PVT caused errors are the best known side effect of technology scaling, with a reduction on the confidence of the circuit. As stated before, the study of PVT variations is completely established in standard CMOS circuit design methodologies, and commercial CAD tools. Only a wider support for emerging devices is desirable, accelerating the inclusion of these technologies in the standard workflows.

Radiation is another source of uncertainty. This is not only for space applications, because, as technologies scale down, ground radiation levels may affect the circuit behavior. The design methodologies related to radiation-hardened circuits rely on radiation tests to completely assess their radiation hardening levels. Usually, these processes require several iterations, which are inevitably expensive both in terms of cost and time. Furthermore, in smaller technologies, variability and temperature dependence can compromise both the standard behavior of the circuit as well as its radiation hardening capabilities. The simultaneous analyses of both radiation effects and variability are nothing but gaining attention [KBK⁺11, FHS15].

Finally, electrical and thermal stress during the transistor life alters the device behavior through time. Transistor aging greatly affects the chances of failure, and, for this reason, CAD software solutions start to support the simulation of these effects. For instance, compact models of transistors suffering from aging

are extended to include radiation long time effects —called Total Ionizing Dose (TID) [SBK15].

2.5.2 Introduction to Uncertainty Sources in RRAM Technologies

RRAM intrinsic uncertainties vary in form and severity due to the wide range of materials used in RRAM fabrication. However, we can identify the following sources as recurrent in most devices. This section gathers some of these concerning issues, all related to the RRAM device intrinsic nature.

Variability

As in many other technologies, variability in memristors is one of the most concerning issues to be solved in memristive applications. Depending on the device characteristics, three distinct kinds of variability are expected to occur in resistive switching devices [EBC⁺15]: *Device-to-Device Variability*, *Intra-Device Variability* and *Read Current Fluctuations*, also called *Random Telegraph Noise (RTN)*. Following this issue, several works have studied the variability patterns of different RRAMs [GVR⁺16, GWZ⁺15, EBC⁺15].

Crossbar Related Errors

Crossbar memory blocks come with two major architecture disadvantages. First, by applying the selection voltages to the chosen row or column lines during write operations, the state of half selected cells —located in these specific row or column— may be altered [YBJH16]. Second, depending on a) the state of the crossbar elements, and b) the crossbar read/write driver scheme, the operation current may be driven through multiple spurious paths instead of directed through the desired one. This *sneak paths* leak write and read currents worsening the circuit capabilities [EBC⁺15, LLN15, KB16, YC16, JKK⁺16].

Memory blocks based on $1R$ cells come with the advantage of not requiring any additional internal device to select the desired cell. However, given the *voltage-current* device behavior, applying the selection voltages to the chosen row and column lines would also apply a voltage difference in the cells located in these specific row and column, modifying their stored values. Consequently, voltages higher than the required aggravate spurious writings and may trigger the selector threshold.

The widely used $V_w/2$ and $V_w/3$ writing schemes [HMM13, VSSH16, CLY16, YBJH16] suffer from this spurious selection and expose cells to possible non-desired state alterations [LJH⁺15]. At the same time, this partial selection produces

a leakage phenomena that should be addressed [CY15]. More in depth information about how this stress affects data retention can be found in [FYL⁺10, RFBP15].

To overcome these problems, circuit level and system level solutions have been proposed. At circuit level, most of them relate the addition of different elements to the $1R$ cell scheme. A first solution introduces a transistor within the cell, building up $1T1R$ cells. $1T1R$ based schemes, as shown in [YTSP13a, BHS⁺14], introduce an additional path to access the transistor gate at each RRAM cell. In exchange of the additional logic, the design avoids accessing non-target cells, improving reliability. The inconvenient area increase coming with the additional CMOS transistor, together with the supplementary *selection* line has led to the investigation of alternative crossbar morphologies.

Most recent efforts focus on the implementation of $1S1R$ cells, with the novelty of introducing an extremely non-linear device inside the cell, acting as a selector [SKN⁺14, CKC⁺16]. Equivalent to two in series diodes, these innovative cells are intended to provide a customizable voltage threshold starting from which the cell would be selected. A similar idea is followed when *Complementary Resistive Switches* cells are used [YMO⁺15].

Data Alteration During Read Operations

If large enough, the voltage applied during read operations may trigger the switching mechanism, altering the geometry of the device CF, with the consequent resistance change. The characteristics of the threshold voltage [HWL⁺14] that triggers this switching depends on the device material, geometry and manufacturing process, and therefore the involved circuitry must ensure that the stored data is not altered.

Temperature Dependence

The local temperatures reached within the isolating material and CF play a fundamental role in the resistive switching mechanisms and conductive processes [WWS⁺11, AKG⁺14, YGCR15, SNK15]. The physical processes described in Section 2.2 are tightened to the thermal evolution of the device. Works like [VGC⁺16, PWG⁺17] verify with experimental data how the oxygen and metal ion migration mechanisms are affected and catalyzed by the instant temperature [Iel16]. Thus, in order to ensure their reliability, memristive based applications are to be hardened against thermal changes.

Lifetime Reduction

RRAM endurance is defined as the number of consecutive *SET-RESET* operations that a single device supports with no serious degradation of HRS/LRS. This feature is a major concern more especially in storage applications, and as a result it has been thoroughly studied [HCW⁺13, LLN15, BAW⁺15, Iel16, KB16, PAR16]. The results indicate that operation voltage amplitudes, and therefore switching speed, have a huge impact on the device degradation. Analytical models describing how the device endurance is dependent on the characteristics of the writing operation have been reported [HCW⁺13]. Tahui Wang's research team has presented, prior a thorough study [CSL⁺16], a different cycle-stress degradation model that describes the relation between HRS stress current and device degradation. Thus, we can relate this non-desired effect to, among others, the use of larger voltages. The use of higher amplitudes would then reduce the RRAM lifetime. In a similar way, [Cho16] relates endurance capabilities and the operation duty cycle, and therefore the applied energy. Similar conclusions can be found in recent works [BAW⁺15, YC16], relating excessive operation voltage amplitudes with endurance reduction.

Regarding *1T1R* cells, the study accomplished in [MCYC16] indicates how different configurations of word-line (WL), bit-line (BL) and source-line (SL) voltages greatly affect the devices lifetime.

Summarizing, stress-caused lifetime reduction has been raised as one of the most concerning issues to be addressed before RRAM technologies get solidly established. A myriad of studies have focused on accurately describing, modeling and palliating RRAM degradation. To overcome this problem, at circuit level alternative schemes like [KYS⁺16] improve oxide devices endurance. At system architecture level, additional measures have been proposed. The work in [PAR16] evaluates the impact of RRAM read/write endurance on system failure, and even proposes reconfiguration approaches in which damaged resistive cells would be substituted by healthy ones. Even though this novel approach extends the system lifetime, it significantly multiplies the consumed area.

Latency

Different write schemes result in different operation characteristics. As seen in [MCYC16], at design time a hard trade-off between HRS/LRS ratio, operation latency, device endurance, power consumption and data retention must be solved. A non desired increase of an operation latency could lead to a bit misread/miswrite, and therefore the variables that define the operation delay must be studied under any possible working condition.

Radiation

In the same way as in CMOS technologies, the most common studies of RRAM circuits working under radiation either involve *Single Event Transients (SETs)*, or the effects of *TID* [EBC⁺15]. Even though usually RRAM technologies present a better behavior in this kind of hostile environments than CMOS, memristors also suffer from SET [CKP⁺14, BHS⁺14, LMBY15] and TID [LWHL15, MHP⁺14, MFH⁺15] effects.

Current Overshoot

Current overshoot is a non-desired effect due to parasitic capacitance during set transition that affects resistance switching phenomena in RRAM arrays [AMW⁺16]. To overcome this problem, two direct measures can be taken. First, minimize the parasitic capacitances, and second, optimize the writing pulse parameters, in order to achieve specific switching speeds and avoid high power consumptions.

2.6 Summary

This section has introduced the reader to resistive switching technologies, describing the memristor theoretical model and presented the physical devices together with their main characteristics and equivalent models. After that, the main RRAM based applications have been reviewed. Finally, we presented the most common uncertainty sources affecting RRAM-CMOS hybrid circuits.

We have seen how RRAM devices, though one of the most promising technologies for next generation electronics, are affected not only by the traditional PVT problems but also by intrinsic limitations that may compromise the whole circuit reliability. As we introduced in the motivation - Section 1.1, the goal of this thesis is to provide support tools and methodologies for the design of the next generation of reliable memristive applications.

Part I

Modeling and Simulation of RRAM Devices

You Win Again Gravity.

Zapp Brannigan, Futurama.

3

Modeling and Characterization of Resistive Switching Technologies

3.1 Introduction

Both industry and academia communities are endeavoring to find novel applications for resistive switching devices and achieve precise models that describe their behavior. First, these days the number of patents and research articles reporting the direct application of memristors does nothing but increase, and relates both to improved versions of already developed systems —such as memories or security primitives— and to novel applications and approaches —such as new computing paradigms. And second, the modeling, simulation and prediction of the behavior of memristors are fundamental pieces of the development and validation of new applications. In 2008, after the first device was implemented, HP Labs introduced the first model based upon the few available fitting points from their memristor [SSSW08]. In the following years, other works put forward improved models that best tracked the dynamic characteristics of that device. In 2011 HP Labs presented

Figure 3.1: Chapter 3 scope and perspective within the Thesis.

Figure 3.2: Scheme showing how model developers and circuit designers would benefit from support tools when dealing with emerging devices.

the first model that departing from a physical formulation —the tunnel effect— significantly improved accuracy [AP11]; at the same time, new devices began to appear and their models also employed physics-based formulations. In the last few years several patents have presented new memristive structures and the modeling community is working towards generalized models that are able to predict any kind of device behavior.

In spite of the big effort carried out by model developers, circuit and system designers still have serious difficulties to integrate and validate memristors. Models are immature and controlling convergence, accuracy, speed and computational load of simulations is specially hard and tricky. In that context, Figure 3.1 defines the present Chapter’s scope: to address challenges and problems researchers face working at this abstraction level.

As it is normal at the early stages of a technology, especially when dealing with nano-structures, there is still a gap between technologists, model developers and circuit designers. In an attempt to bridge this gap, Biolek et al. [BBB09] presented in 2013 several models of memory devices described in SPICE, the traditional standard for circuit designers; they provided with interesting features such as the possibility to control the time step, soft transition functions, simulation parameters

control and behavioral integrators. Still, the approach is not generalizable, it just includes three models and the designer needs to go through several trial and error phases to select the correct time step and control integral limits. Figure 3.2 displays the described context, where both model developers and circuit designers would benefit from a support framework, providing non-converging control methods, parallel computation, and considering PVT variations.

The contributions of this chapter aspire to help circuit/system designers in the integration of the memristor, while at the same time aiding model developers in the validation of their proposals. We introduce the Memristor Application Framework, MAF hereafter, that encompasses three modules corresponding to the stages in the standard work methodology: 1) device model characterization and validation for a given application scenario; 2) memristor level characterization, computation and selection of stimuli and operation time; 3) SPICE/Spectre subcircuit generation. The main features of MAF that contribute to the state of the art are the following:

- It includes the best-known memristor models and future upcoming models can be easy to introduce.
- It provides both SPICE and Spectre —the *de facto* industrial standard from Cadence— subcircuits.
- It employs systematic modifications to improve SPICE code —i.e. aggregation/disaggregation of sources, aggregation of variables, etc.— that are translated into significant simulation speed-ups.
- MAF allows the analysis of parameter variations, an important aspect in memristor validation. The memristor, as well as other nano-devices, is highly affected by variability due to its extremely non-linear behavior.
- The framework provides the designer with total control over convergence, computational load and the evolution of system variables.

The binaries of Memristor Application Framework are available for downloading at the research group resources page [[GRLVR15](#)]. The most up to date source code, as well as the code version history can be found at the developer's repositories [[GR16](#)].

In this chapter we first review the best known compact models presented in the literature. After that, in Section 3.3 we collect the proposed systematic improvements related to the subcircuits netlist. Section 3.4 describes the framework, focusing on the main modules and functionalities. Next, Section 3.5 describes the related design methodology and presents an example of use. In Section 3.6 we detail the performance improvements achieved by the subcircuit modifications.

Finally, we would like to point that Appendix A gathers MAF implementation details, including:

- Electric device modeling.
- Scenarios modeling and PVT.
- Thread Management.
- Non-convergence problems handling.
- Data population control.

3.2 Review of Best Known Compact Models

We now proceed to review the most important memristor compact models presented in the literature to date, ordered by the publication date of their first version. Most of them have been incorporated within the MAF's model library, as will be described in Section 3.4.

It is to note that some of them may refer a partial aspect of the whole set of dynamics of the modeled memristor, e.g. switching modeling. The following memristor models enumeration gathers then models that have been proposed to be simulated within a wide variety of circuit simulators as well as in numerical calculus simulators.

Although some models base the resistance change on the alteration of the ion concentration in the device regions, most physical models consider the resistance variations due to the growing/destruction of an ion/oxygen vacancy *Conductive Filament (CF)*. The modeling of the memristor behavior then relies on the geometrical evolution of the filament, which is formed/destroyed under specific circumstances. The behavior of the remaining models depends on the ion/oxygen vacancies density of varying 2-dimensional or 3-dimensional regions. Next, we briefly describe their highlights.

3.2.1 Linear ion drift

Linear ion drift model, firstly presented in [SSSW08], represents a generic charge controlled memristor. This model, whose state variable time derivative linearly depends on the current, can be expressed through a polynomial equation system. Therefore, this was the first analytical description of *memristance* dynamics, relating it with the resistance switching phenomena naturally occurring in some layered nanoscale systems. Additionally, this work introduces the non-linear drift model, which is described more deeply in later works.

3.2.2 Non-linear ion drift models with different window functions

Extending the linear ion drift model, and moved by the simplicity of using a single state variable dependent on the charge of the device, several research groups have developed compact models reporting non-linear drift of the dopant area [SSSW08, JW09, BBB09, PP11, ZHL16]. Though simple, these models present a good first approximation to the behavior of memristors, and therefore several memristor-based applications are developed using them [EM11, HMM13]. Non-linearity characteristics are given by shaping the state variable using different window functions. Some of these functions have limitations, like the *hard switching condition* when modeling the boundary conditions [SSSW08, BBB09], or lack of modeling the threshold effect in the memristor behavior.

3.2.3 Simmons' tunneling barrier model

One of the first and most complete memristor models [PSB⁺09, AP11], emulates Simmons' tunneling barrier phenomena reported in a Metal/Insulator/Metal (MIM) junction combined with an in series resistance. The memristor layered structure behaves following the Simmons tunneling theory [Sim63], matching the experimental results [AP11]. In this case, the state variable leading to the resistance switching refers the tunnel barrier width. For simplicity, the model developers obviate the —experimentally observed— diode-like asymmetry for which under positive and negative bias schemes the device conducts differently. Despite the accuracy and the solid physical bases, this model often experiences convergence problems due to the complexity associated to the computation of transient simulations.

3.2.4 Yakopcic's model

Yakopcic's model [YTSP13a] is characterized by a single state variable whose evolution depends on the applied voltage. Controlled by two voltage thresholds, the state variable experiments an exponential drift dependence on the applied voltage, and is also windowed to control both speed and overflows when approaching to its boundaries. Additionally, the state variable model introduces thresholds. The conduction is described using a hyperbolic sinusoid shape, therefore following and simplifying the previous MIM memristors. If the parameter set is correctly set-up, the model can fit the characteristic $i(v)$ curves of several physical devices [YPL⁺08, JCE⁺10]. Moreover, in [YTSP13a] the authors report one of the first modeling approaches regarding parameter variations. As a final note, this model has been used in development of memristor-based neuromorphic systems.

3.2.5 VTEAM model

The Voltage ThrEshold Adaptive Memristor model [KRFK15] also derives from the Simmons tunneling barrier model. Like Yakopcic's model, the TEAM model proposes a simplification of complex MIM structures like the ones presented in [PSB⁺09, AP11], and introduces voltage thresholds to control the state variable evolution. However, compared to the Yakopcic's exponential dependence of the state variable dynamics, the evolution of the state variable in the VTEAM model is based on a polynomial curve. Compared to Simmons' tunneling barrier model, the conduction mechanism is simplified, which, consequently, reduces the computational load.

3.2.6 Simplified Simmons' barrier model

This model, presented in [KFKW13], is a modification of the Simmons' tunneling barrier [PSB⁺09, AP11]. Developed by VTEAM model's authors, this model was built in order to compare the computational load and accuracy of VTEAM with Simmons' tunneling barrier model. The $i - v$ relation is simplified to a linear relation. Also, the current dependence on the state variable is reduced from the complex system presented in [PSB⁺09] to a single exponential. Therefore, performance is improved and most convergence problems are solved, while, on the other hand, the accuracy is reduced.

3.2.7 Eshraghian' MIM & threshold model

Eshragian's model [EKC⁺12] presents a complete and deep study of the underlying mechanisms of the memristor dynamics. Based on this work, Eshraghian et al. present an accurate, highly customizable and robust (from the convergence and overflow point of view) memristor model. The model follows a conduction emulation approach based on tunneling. The state variable evolution obeys programming thresholds coupled to a complex function dependent on the applied voltage. The SPICE source code going with the analytical model aims to be directly used in memristor-based large systems.

3.2.8 1-Dimensional CF, Arizona University - Stanford University

Researchers from Arizona University -initially working at Stanford University- developed a physics-based RRAM Verilog-A model [CY15, YU15, JWY⁺16]. Validated against physical devices, this Verilog-A model describes the resistive switching effect using a simplified 1-dimensional conductive filament. Considering this approach, there is no definition on the lateral size of the filament. In this model, the auxiliary state variable is the gap distance, which is defined as the distance

between the top electrode and the CF. This gap determines the RRAM resistance through electron tunneling conduction mechanism. The model dynamics consider thermal dependence and variability emulation. It is to be noticed that this model is under constant development, and, what is more important, its source code is kindly open to the research community under Creative Commons Attribution-NonCommercial License. This characteristic has lead the model to be one of the most used by researchers and circuit designers.

3.2.9 2-Dimensional CF, Stanford University - Beijing University

Stanford University together with Beijing University researchers developed a Mott–Gurney ion-transport based RRAM Verilog-A model [HWL⁺14, HCL⁺14, LJH⁺15]. Today, and together the 1 – D CF Arizona University model, this 2 – D model is one of the most widely accepted models for CBRAM devices.

Compared against 1 – D CF approaches, this model considers the CF as an evolving cylinder, considering both its width and length. Similarly to the 1 – D Arizona - Stanford model, it considers thermal and variability effects, as well as parasitic effects, including both parasitic resistance of switching layer and electrodes, and parasitic MIM capacitance. Regarding its license, although all rights are reserved and the user needs to be granted access to the sources by the authors, this model is available to the research community, with the consequent popularity increase.

3.2.10 1-Dimensional Filament Growth Model, Politecnico de Milano University

Complementing the last two reviewed models, Dr. Ielmini’s team proposed a generalized analytical model [Iel11, ABGI14, Iel16] for the voltage-controlled switching in RRAMs. In this case, the evolving dimension refers the CF width growing and dissolution. The switching kinetics are completely dictated by the voltage, which controls the energy barriers involved and the local Joule heating. In this model, both temperature and I_c current play a key role in both set and reset transitions.

This model was later extended in [ABGI14], to consider different CF evolutions (requiring 2 – D to describe the processes) as well as to include variability within the simulation capabilities.

3.2.11 Review’s Summary

Finally, as a summary of the reviews’ contents, Tables 3.1 and 3.2 detail the switching and conductive characteristics of each reviewed model.

Table 3.1: Review of memristor compact models presented in the literature: State variables dependence. The naming of variables and parameters can be found at each specific model publication.

Model	State Variables Dynamics
Linear ion drift [SSSW08]	$\frac{dw}{dt} = \mu_v \frac{R_{on}}{D} i(t)$
Non-linear ion drift [SSSW08, JW09, BBB09, PP11, ZHL16]	$\frac{dw}{dt} = \mu_v \frac{R_{on}}{D} i(t) f(w), \text{ where}$ $f(w) = 1 - (\frac{2w}{D} - 1)^{2p}, \text{ [JW09],}$ $f(w) = 1 - (\frac{w}{D} - stp(-i))^{2p}, \text{ [BBB09],}$ $f(w) = j(1 - ((w - 0.5)^2 + 0.75)p), \text{ [PP11]}$
Yakopcic [YTSP13a]	$\frac{dx}{dt} = g(v(t))f(x(t)), \text{ where}$ $g(v) = \begin{cases} A_p(e^{v(t)} - e^{V_p}) & \text{if } v(t) > V_p \\ -A_n(e^{-v(t)} - e^{V_p}) & \text{if } v(t) < V_n \\ 0 & \text{otherwise} \end{cases}$ $f(x) = \begin{cases} e^{-\alpha_p(x-x_p)}(\frac{x_p-x}{1-x_p} + 1) & \text{if } x > x_p \\ e^{\alpha_n(x+x_n-1)}(\frac{x}{1-x_n}) & \text{if } x < 1 - x_n \\ 1 & \text{otherwise} \end{cases}$

Model	State Variables Dynamics
Simmons tunneling barrier [PSB⁺09]	$\frac{dw}{dt} = \begin{cases} f_{off} \sinh(\frac{ i }{i_{off}}) e^{-e \frac{w - a_{off}}{w_c} - \frac{ i }{b} - \frac{w}{w_c}} & \text{if } i > 0 \\ -f_{on} \sinh(\frac{ i }{i_{on}}) e^{-e \frac{a_{on} - w}{w_c} - \frac{ i }{b} - \frac{w}{w_c}} & \text{otherwise} \end{cases}$
VTEAM [KRFK15]	$\frac{dx}{dt} = \begin{cases} k_{off} (\frac{i(t)}{i_{off}} - 1)^{\alpha_{off}} e^{-e \frac{x - a_{off}}{w_c}} & \text{if } i > i_{off} \\ k_{on} (\frac{i(t)}{i_{on}} - 1)^{\alpha_{on}} e^{-e \frac{a_{on} - x}{w_c}} & \text{if } i < i_{on} \\ 0 & \text{otherwise} \end{cases}$
Simplified Simmons barrier [KRFK15]	$\frac{dw}{dt} = \begin{cases} f_{off} \sinh(\frac{ i }{i_{off}}) e^{-e \frac{w - a_{off}}{w_c} - \frac{ i }{b} - \frac{w}{w_c}} & \text{if } i > 0 \\ -f_{on} \sinh(\frac{ i }{i_{on}}) e^{-e \frac{a_{on} - w}{w_c} - \frac{ i }{b} - \frac{w}{w_c}} & \text{else} \end{cases}$
Eshraghian MIM with threshold [EKC⁺12]	$f(w) = sf_0 + sf_m(1 - (2w - 1)^{2p})$ $\frac{dw}{dt} = \begin{cases} 0 & \text{if } v > 0, w \leq w_{max} \text{ or } v < 0, w \geq w_{min} \\ f_{on}(1 - \frac{v}{2\phi_0}) e^{f(w)\phi_0(1 - \sqrt{1 - \frac{v}{2\phi_0}})} & \text{else if } v > 0 \\ -f_{off}(1 + \frac{v}{2\phi_0}) e^{f(w)\phi_0(1 - \sqrt{1 + \frac{v}{2\phi_0}})} & \text{else if } v < 0 \end{cases}$
1D CF Arizona - Stanford [CY15, YUY15, JWY⁺16]	$\frac{d\langle g \rangle}{dt} = -v_0 \exp(-\frac{E_A}{kT}) \sinh(\gamma \frac{a_0}{t_{ox}} \frac{qV}{kT})$ $g _{t+\Delta t} = \int (\frac{d\langle g \rangle}{dt} + \delta_g \chi(t)) dt$ $\gamma = \gamma_0 - \beta g^\alpha, \quad T = T_0 + V I R_{th}$

Model	State Variables Dynamics
	$T = T_0 + V I R_{th}$
	RESET: min dx between
	$\frac{dx}{dt} = a f \exp\left(\frac{-(E_i - \gamma Z e V)}{kT}\right)$
2D CF Stanford - Beijing [HWL ⁺ 14, HCL ⁺ 14, LJH ⁺ 15]	$\frac{dx}{dt} = a f \exp\left(\frac{-E_h}{kT}\right) \sinh\left(\frac{\alpha_h Z e V_r}{x k T}\right)$
	SET:
	$\frac{dx}{dt} = -a f \exp\left(\frac{-(E_a - \alpha_a Z e V/x)}{kT}\right)$
	$\frac{dw}{dt} = 2a(1 + a(w)) f \exp\left(\frac{-(E_a - \alpha_a Z e V/x_0)}{kT}\right)$
1D CF Model, Milano [Iel11, ABGI14, Iel16]	$\frac{d\Phi}{dt} = A \exp\left(-\frac{E_{A0} - \alpha q V}{kT_0(1 + V^2/(8T_0 \rho k_t h))}\right)$
	$T = T_0 + \frac{V^2}{8 \rho k_t h}$

Table 3.2: Review of memristor models presented in the literature: Conduction processes. The naming of variables and parameters can be found at each specific model publication.

Model	$i(t) - v(t)$ relationship
Linear ion drift [SSSW08]	$v(t) = \{R_{on} * \frac{w(t)}{D} + R_{off} * (1 - \frac{w(t)}{D})\} i(t)$
Non-linear ion drift [SSSW08, JW09, BBB09, PP11, ZHL16]	$v(t) = \{R_{on} * \frac{w(t)}{D} + R_{off} * (1 - \frac{w(t)}{D})\} i(t)$

Model	$i(t) - v(t)$ relationship
Simmons tunneling barrier [PSB ⁺ 09]	Large system according [Sim63, PSB ⁺ 09, AP11]
Yakopcic [YTSP13a]	$i(t) = \begin{cases} a_1 x(t) \sinh(bv(t)) & \text{if } v(t) > 0 \\ a_2 x(t) \sinh(bv(t)) & \text{else} \end{cases}$
VTEAM [KRFK15]	$\lambda = \log(R_{off}/R_{on})$ $v(t) = R_{on} \exp\left(\frac{\lambda}{w_{off} - w_{on}} (w - w_{on})\right) i(t)$
Simplified Simmons barrier [KRFK15]	$\lambda = \log(R_{off}/R_{on})$ $v(t) = R_{on} \exp\left(\frac{\lambda}{w_{off} - w_{on}} (w - w_{on})\right) i(t)$
Eshraghian MIM with threshold [EKC ⁺ 12]	$i(t) = a_1 w^n \sinh(b_1 v(t)) + a_2 (e^{b_2 v(t)} - 1)$
1D CF Arizona - Stanford [CY15, YU15, JWY ⁺ 16]	$I = I_0 \exp\left(\frac{-g}{g_0}\right) \sinh\left(\frac{V}{V_0}\right)$
2D CF Stanford - Beijing [HWL ⁺ 14, HCL ⁺ 14, LJH ⁺ 15]	$I_{hop} = I_0 \pi r^2 \exp\left(\frac{-x}{x_t}\right) \sinh\left(\frac{V_{gap}}{V_t}\right)$ $I_{CF} = \frac{V_{CF} \sigma_{CF} \pi w^2}{4(x_0 - x)}$
1D CF Model, Milano [Iel11, ABGI14, Iel16]	$I = V / \left(\frac{4\rho t_{ox}}{\pi \Phi^2}\right)$

(a) Traditional subcircuit schemes.

(b) Custom subcircuit schemes.

Figure 3.3: Conventional (3.3a) memristor subcircuit scheme and custom scheme (3.3b).

3.3 Systematic Improvement of Model Netlists

The authors of some of the models described in Section 3.2 did not provide a direct circuitual realization. We have developed netlists describing them not only in SPICE but also in Spectre.

Furthermore, we have introduced systematic improvements in the circuit scheme of all considered models. These advances help the low level circuit simulator to better handle the improved incidence and loop matrices [GRR12]. Consequently, new model schemes lead to important speedups in both SPICE and Spectre simulations, and, as detailed in Section 3.6, to reduce or mitigate convergence problems.

In order to explain the systematic improvements that we propose, let us first review the structure and behavior of a basic memristor subcircuit (a netlist that is instantiated in a higher hierarchy circuit). Figure 3.3 displays the three stages a memristor subcircuit can be divided into, corresponding to the conductance, state variables, and auxiliary variables respectively.

First, the device $v - i$ relationship is traditionally modeled as two terminals connected by a current generator in series with a resistor R_{aux} . The current generator, instance of a SPICE non-dependent source, is governed by a expression with the form

$$i = r(v, \vec{x}, \vec{a}), \quad (3.1)$$

where $\vec{x} = \{x_1, \dots, x_q\}$ and $\vec{a} = \{a_1, \dots, a_p\}$ are the state and auxiliary variables, respectively.

Second, for all models, the state variables evolution is determined by a set of differential equations which depend on voltage (voltage controlled memristor) or current (current controlled memristor), the state and the auxiliary variables:

$$\frac{dx_j}{dt} = f_j(v, i, \vec{x}, \vec{a}). \quad (3.2)$$

As seen in the central part of Figure 3.3a, with the aim of modeling each state variable a current generator takes the value given by equation (3.2). A capacitor placed in series with the current generator performs the state variable integration. The circuit simulator internally performs the integration of the capacitor current. This way, the voltage at node x_j takes the form

$$x_j = \frac{1}{C} \int_{t_0}^t f_j(v, i, \vec{x}, \vec{a}) dt + v(x_j)|_{t_0}. \quad (3.3)$$

The x_j voltage is scaled by changing the capacitance C . Also, by specifying the initial voltage $v(x_j)|_{t_0}$ at node x_j , the state variable initial condition gets fixed.

Finally, each of the variables that are not determined by a differential equation is modeled by a voltage source with the expression

$$a_i = g_i(v, i, \vec{x}, a_0, \dots, a_{i-1}, a_{i+1}, \dots, a_p) \quad (3.4)$$

Figure 3.3a summarizes the previous scheme. Next, we describe the four systematic optimizations introduced in the subcircuits which are represented in Figure 3.3b.

Variable Integration Mechanism

The use of a capacitor together with a current generator to perform a variable integration makes the simulator manage two different components in each simulation step, which duplicates the convergence assessments.

SPICE-like simulators allow the utilization of several mathematical functions. Among these functions the designer can choose integrators. By removing the capacitor-current generator set, and placing instead a single voltage generator whose value refers to the differential equation, the circuit simulator only manages one component. Therefore, the computational load is reduced.

Variable Control

The developed subcircuit schemes must provide the mechanism to ensure that some state variable values are constrained. Some subcircuit models [KFKW13] delegate in a set of diode-voltage generators to impose these bounds. Other subcircuits limit the ranges using switches [RC10]. This scheme involves several circuit components with the corresponding computational overhead.

In contrast, we can eliminate these components by using the simulator boundary functions. Furthermore, this mechanism can be integrated in the same definition as the differential equation modeling, simplifying the subcircuit structure. Consequently, we can reduce the computational load systematically eliminating some components. This way a state variable can be defined, initialized and controlled in a single SPICE sentence:

```
1 EW w 0 value={min(min_value , max(idt(diff_eq , w_init) ,
    max_value))}.
```

Common Values Aggregation

In case that, during the computation of the state or auxiliary variables, a value depending on an auxiliary variable is computed several times, the circuit simulator is performing the same operation multiple times. Let $h(a_l)$ be an operation evaluated several times. We propose to use a new variable b_l , which takes the form of a voltage source V_{b_l} in the subcircuit, instead of computing every time step the same value. This way we have the following system

$$\begin{aligned} v_m &= g_m(v, i, \vec{x}, a_0, \dots, a_{m-1}, a_{m+1}, \dots, a_p, h(a_l)) \\ &\vdots \\ v_n &= g_n(v, i, \vec{x}, a_0, \dots, a_{n-1}, a_{n+1}, \dots, a_p, h(a_l)) \end{aligned}$$

transformed into

$$\begin{aligned} b_l &= h_l(a_l) \\ v_m &= g_m(v, i, \vec{x}, a_0, \dots, a_{m-1}, a_{m+1}, \dots, a_p, b_l) \\ &\vdots \\ v_n &= g_n(v, i, \vec{x}, a_0, \dots, a_{n-1}, a_{n+1}, \dots, a_p, b_l). \end{aligned}$$

which creates the new set of variables $\vec{b} = \{b_0, \dots, b_z\}$.

Current and voltage sources aggregation and disaggregation

In several cases a state variable is modeled as a contribution of distinct voltage or current sources. If these sources do not have convergence problems the circuit simulator needs to handle several components, which implies checking different convergence or boundary conditions. The aggregation into a single source can reduce the computational load. On the other side, there are sources computing complex functions that can lead into convergence problems. Those functions can be split to isolate the convergence problem.

3.4 Memristor Application Framework MAF

Continuing with the idea of aiding the design community, we have developed a memristor application framework (MAF) that encompasses all the stages necessary to include a memristor model in the standard circuit design flow. MAF is developed in Java. Therefore, the application can be run in several OS.

As seen in Figure 3.4, MAF functionalities are mainly covered by three main modules: *Device Characterization*, *Level Characterization* and *Subcircuit Editor*. The first two are different simulators with which the user is able to fully characterize each of the memristor models included in the model library. Each module is illustrated with a simple example in Section 3.5.2.

Before describing MAF main modules we will explain how both simulators—*Device Characterization* and *Level Characterization*—take advantage of the handling capabilities of parameter variations. The framework helps analyzing the fluctuation of parameters which can destabilize the memristor behavior. For example, a slight variation on the threshold of the device may lead the memristor to act as a normal resistor. By contrast, under the presence of spurious voltage, a value stored in a memristor can be altered. MAF simulators provide the way to study the dynamics of a specific memristor in presence of parameter variations. We have included *Corner analysis* and *Monte Carlo (MC)* methods. In both methods the variation of voltage, temperature (depending on the memristor model) and model inherent parameters is given by a probabilistic deviation from the mean value. Therefore, instead of a specific memristor excited by a single voltage, several simulations are concurrently executed, performing a deeper analysis. This way, a wide set of scenarios which differ in the values of the device nominal parameters, the temperature (if present) and the voltage fluctuations are created and simulated.

Figure 3.4: *MAF functional scheme, showing the main modules and the interaction with external software.*

3.4.1 Model Library

As depicted in Figure 3.4, MAF characterization modules rely on the framework model library. The first developed version of MAF included seven of the models described in Section 3.2, while considering their improved subcircuit netlists from Section 3.3. Over the years the number of models presented in literature increases; thus, the successive MAF versions have extended their model library, addressing circuit designers needs [GR16]. It is likely that new memristor devices, and their corresponding models, will appear. The inclusion of these memristors in the models library is a key point in the extensibility of the framework.

To deal with this problem, the MAF model library and engine have been carefully designed as a modular system to allow the scalability of its functionalities. MAF is composed of primitives which have been programmed to act as independent functional units, interacting between them by exchanging data models and scenarios.

Figure 3.5: *MAF* Yakopcic's model [YTSP13a] device characterization: Monte Carlo simulation, $i(v)$ curve.

3.4.2 Device Characterization Module

The *Device Characterization Module* is a transient simulator that follows the device behavior equations, providing a powerful environment to study the response of the models included in the database once stimulated by different patterns of voltage and taking into account parameter variations.

It is well known that some of these models require from small time steps during the simulation in order to obtain valid results [AP11]. Additionally the computation of some variables strongly depends on the timing of the simulation, so to avoid computation overflow problems and non-desired behaviors, our simulator guarantees that the time step is small enough to make the device behave correctly while ensuring the convergence of the simulation. The simulator also includes several techniques [GRLVIL12] to solve and recover from the non-convergence problems derived from the variability of nanotechnologies, assuring the viability of the memristor behavior and the simulation convergence.

Figure 3.5 shows, as an example, the behavior of Yakopcic's model [YTSP13a] suffering PVT variations. Here, the standard hysteresis $i(v)$ curve follows different slopes and thresholds depending on the parameter stochastic variations. To help the model developer and circuit designer characterize the dynamics of the device, useful transients such as the ones shown in Figure 3.6 in detail describe state variables and physical magnitudes evolution.

(a) State variable transient, study under variations.

(b) Resistance levels transient, study under variations.

Figure 3.6: Monte Carlo results given by MAF Device Characterization module.

Figure 3.7: MAF Fitting functionality. The blue pointed line refers the measured curve, while the red line corresponds to the simulated model.

Fitting Module

To find the parameter values of the different compact models that best match the behavior of physical devices is a problem that consumes time and computational resources [KRFK15] that may not lead to the optimum solution. This problem gets magnified when dealing with complex resistive switching structures whose behavior depends on a myriad of parameters, and whose conductance mechanism may meet several processes depending on the working conditions. The model developer then deals with several models and endless variables, trying to find the parameter set that best fits the device measured curves. To that end, MAF includes a useful fitting module. This functionality provides model developers with the required assistance and support fitting algorithms. Figure 3.7 shows an example fitting a bipolar *aSiC* device using Ielmini's CF growth model [Iel11, ABGI14, Iel16] combined with Schottky barrier conductance model.

The *Device Characterization Module* accomplishes a design space exploration: first it varies and simulates the model depending on its parameters. Then, it computes the error comparing the simulation curves against measured data. At

Figure 3.8: *MAF Level characterization Monte Carlo simulation. Histogram corresponding to the multi-level characterization showing the time to reach specific resistance levels.*

this point, the tool can either explore the whole design space given by the user, or, on the contrary, iteratively reduce the error using a Simulated Annealing heuristic algorithm. At the end of the process, the simulator reports the set of parameters and model that best fits the physical device behavior.

3.4.3 Level Characterization Module

The memristance presents a highly non-linear relation with the flux. Therefore, the designer needs high accuracy in the flux value and related parameters. Thus, with the aim of setting a precise resistance value (e.g. to perform a multi-level storage or a weighted operation), the designer has to accurately know the amplitude and duration of the writing signal. With a fixed set of pulse amplitudes, the pulse length required to store a specific value will depend on the memristor speed.

For this reason, one of the main issues when designing memristive applications is, given a supply voltage, determining the time it takes for the device to perform an accurate state change. The *Level Characterization* module allows a comprehensive study of the memristance-flow relationship when writing a desired value in the memristor. Consequently, this module helps in the study of the time evolution of memristance and model state variables. Also, the tool allows the automated calculation of the pulse lengths required to store the desired values.

Additionally, with this module the designer is able to acquire and study the mean values and maximum deviations of the operation timing. This, together with the histogram and the state variable-memristance evolution plots, helps find the

operation characteristics and define the safety margins. To illustrate the module features, Figure 3.8 displays the statistical results regarding the time needed to reach a specific resistance level under a constant voltage. This experiment characterizes the multi-level storage capabilities of Yakopcic's model [YTSP13a] suffering PVT variations.

3.4.4 Subcircuit Editor Module

This module loads the customized parameters brought from the characterized memristor and creates the appropriate subcircuit. In this stage the user is able to modify the parameters to include additional components or to change the subcircuit structure. Furthermore, the designer can automatically generate the code to perform Monte Carlo & Corner analysis in the external circuit simulator. As result, the generated subcircuit is ready to be referenced by the circuit netlist.

3.5 MAF Design Methodology

In this section we put forward a recommended methodology for model or circuit designers that wish to use MAF.

3.5.1 Model Developers

The goal moving model developers is the definition of a compact model that accurately mimics the measured behavior of a physical device. Thus, the process starts with the sketching of the set of equations and variable ranges that, with more or less precision, are expected to approximately meet the device properties. This stage uses the experience provided by previous literature models and the knowledge on the device materials and characteristics.

This system of equations is translated to MAF and immediately simulated for a coarse characterization. MAF's fitting module is responsible for the generation of accurateness and possible non-convergence reports. Iteratively, and always using the graphical and statistical information given by this module, the model developer refines the variables evolution progressively delimiting the parameter ranges. Once the model correctly meets the measured voltage-current curves, as well their transients, if the experimental data is thorough enough, statistical properties such as conductance or switching parameter variations can be studied. At this point the device characterization module is able to simulate the effects that individual or combined parameter variations have on the voltage-current fingerprint.

At the end of the process, using the subcircuit editor module the model developer can create the SPICE equivalent code representing the defined model.

Figure 3.9: Semiconductor analyzer measurements imported in MAF.

Example of Use

The present example of use describes the methodology for the definition and characterization of the compact model meeting a physical bipolar device. In particular, we will be working with a bipolar $Cu/a-SiC/TiN$ CBRAM which we characterized at Southampton University under the supervision of Professor K. De Groot.

The goal of this example is the accurate fitting of the behavior of the device. The experimental setup includes a wafer that allows the devices to be individually accessed using probes by a semiconductor analyzer.

We start by gathering together concepts and knowledge. We know the materials and dimensions of the layered structure. Additionally, similar structures based on amorphous silicon carbides whose conduction processes are dominated by ohmic (in the LRS) and thermo-ionic field emission over a Schottky barrier—Schottky emission— (in the HRS) have been reported. Regarding the switching, we will follow the filament growing model controlled by electric field and thermal activation [Iel11, ABGI14, Iel16].

In a second stage we characterize the physical device using the semiconductor analyzer, and later import the data in MAF. Figure 3.9 describes the representative curve related to a complete *SET-RESET* cycle.

The third stage is devoted to the definition of the differential equations system that, in principle, will meet the device dynamics and characteristics. Basically, the system is composed of the equations related to the CF evolution [Iel11, ABGI14,

Figure 3.10: Gross space exploration using MAF.

Iel16] and the ohmic and Schottky conduction mechanisms that can be found in Table 2.1. Then, these equations are systematically translated to the syntax defining MAF’s electrical device models —Appendix A.

In the forth step the physical parameters such as dimensions, activation fields, barrier potentials, etc. are set. The remaining parameters (i.e. *free parameters*), which include thermal coefficients, effective Richardson constant or the maximum filament dimensions, are initially guessed and set within reasonable ranges.

The fifth step involves iterative approximations in which the model developer accomplishes gross simulations varying the list of values that the *free parameters* may have. Figure 3.10 shows a simple example of how the device characterization module handles these simulations and later reports. At every step, the fitting module gives us the best parameter set extracted from these gross simulations. As a representative example, Figure 3.11 shows the best fitting automatically extracted from the simulations displayed in Figure 3.10, for both switching thresholds and LRS conduction mechanism.

This process can be either manually guided (introducing the list of values for the *free parameters* to be evaluated) or accomplished in an automated way using the SA heuristic capabilities. As seen in Figure 3.12, this SA heuristic is specially useful for the final stage, in which fine-grain simulations extract the parameter set that best suits the measurements.

Figure 3.11: MAF fitting capabilities in intermediate iteration.

Figure 3.12: Reported fitted behavior.

Figure 3.13: MAF methodology to be followed by the circuit designer.

3.5.2 Circuit Designers

Figure 3.13 illustrates the methodology to be followed by circuit designers. The methodology departs from an application in which one or several memristors need to be used. The application has certain restrictions which translate into a set of constraints for the memristors. The designer is assumed to be acquaintance with the models so as to reduce the group of possible candidates. As seen in the scheme, the user can select the optimal model by comparing the behavior of the candidate models using coarse-grained MAF's *Device Characterization* simulations.

Once the memristor model and specifications are defined, in the next stage the designer performs a refinement of the memristor's parameters, stimuli, timing and powering conditions. This precise adjustment is performed employing MAF's *Device Characterization* and *Level Characterization* simulators. In case the analysis does not converge, the memristor is not fast enough, or if some parameter variation disrupts the operation, the parameters should be refined. MAF offers its own graphical results representation as well as formated-data exportation for external software such as GnuPlot and Matlab.

Figure 3.14: Crossbar circuit, showing the target cell and the half-selected cells susceptible of being altered.

When the initial requirements are satisfied, a SPICE / Spectre subcircuit memristor netlist is generated, allowing the designer to customize it, include the subcircuit in the global circuit netlist and proceed to its simulation.

Example of Use

To show the framework capabilities, we present a simple case of use targeting one of the most common memristive applications: a crossbar memory. More concretely, we will perform an 8-level writing operation in a 3×3 RRAM custom netlist, as shown in Figure 3.14. In this crossbar scheme, the writing operation involves two different pulse generators and nine cells, each composed of one memristor. Each memristor cell can be modified to store eight different resistance values, from “0” to “7”, linearly separated. Initially, all cells have a value of “0”. Rows and columns are numbered from 1 to 3. In this example, the desired operation will write a “7” in the second cell of the second row of the RRAM array (position [2, 2]).

Crossbar array structures have structural problems. As we saw in Section 2.5, one of the main problems refers to the spurious voltage which feeds the cells placed next to the selected RRAM. These voltages can modify the state variables in non-desired cells, and consequently, alter the stored values. To minimize those non-desired effects, we will use the half-voltage writing scheme shown in Figure 3.14: Non selected RRAM cells will be affected by lower spurious voltages.

Figure 3.14 displays the circuit scheme, the target cell is represented in dark blue, while the cells that are sensitive to data alterations are drawn in light blue.

The purpose of this example is twofold:

1. Determining all the parameters to perform the full state change—including feeding voltage and timing characteristics.

Figure 3.15: State variable w evolution at different pulse voltages. In all cases, the pulse lengths used are 20 ns with 8 ns slopes and a delay of 5 ns.

2. Analyzing the impact on the non-selected cells when the writing operation is performed.

Following the presented methodology, first we define the design constraints. In our example we will assume that the circuit power supply limits the maximum voltage with which the memristor is fed, besides we set a limit in the duration of the writing operation, thus we have:

1. Voltage supply should not exceed 2 V.
2. The operation should be performed in 15 ns.

The next step is selecting the most appropriate memristor model. We will use Eshraghian's model [EKC⁺12] with the default set of parameters because of its good trade-off between accuracy and computational load, as well as the presence of a threshold in its behavior, which will help minimizing the cell data corruption effect.

We proceed to simulate with the MAF *Device Characterization* module different transient simulations to set the feeding voltage, taking into account that the model must be fast enough to perform the required operation. Figure 3.15 shows simulation results. As can be seen, with 1.75 V we can completely alter the state variable (bounded between “1” and “0”), and therefore, reach with this writing voltage either maximum / minimum memristance values.

(a) General view of the state variable evolution.

(b) Detailed view of the state variable evolution.

Figure 3.16: State variable w in a Monte Carlo analysis, general view (3.16a) and detailed view (3.16b) showing ΔT and Δw margins due to variability.

Performing a Monte Carlo analysis with the *MAF Device Characterization* module we study how the variation of the feeding voltage and other parameters of the model affect the maximum storage levels (Δw) and the writing operation length (ΔT). Figure 3.16 shows an example of the evolution of the state variable w along the time. In the detailed view both ΔT and Δw are presented. Evaluating those values we are able to refine the parameters of the memristor's model and estimate the variability of the device.

Focusing on the pulse length of the writing operation, we carry out another Monte Carlo simulation using the *MAF Level Characterization* module. Figure 3.17 shows the histogram with the pulse lengths occurrences required to perform multi-

Figure 3.17: Monte Carlo results given by MAF Level Characterization module.

level storage over 250 process scenarios. Based on this information, a designer is able to determine the operation security margins.

After all the memristor variables have been settled, as well as the operation voltage and timing are chosen, using the *Subcircuit Editor* module we automatically generate the corresponding SPICE subcircuit and proceed to include it in the global netlist. This subcircuit will be instantiated in each RRAM cell at the global SPICE netlist.

At this stage, we can verify in SPICE that the non-desired effects of a spurious alteration of previously stored data in adjacent cells do not appear. Figure 3.18 displays the writing operation accomplished in both the target cell and an adjacent cell. As shown, the previous value stored in the adjacent cell is almost unaltered—notice the change in the left y axis—thus the impact is negligible.

(a) State variable w in a the target selected cell. (b) State variable w in a half selected cell.

Figure 3.18: SPICE simulation of the RRAM crossbar array. State variable evolution w in the target cell (the desired value is correctly stored) and in an adjacent cell (the state variable is almost unaltered).

3.6 Speedup Results

We have tested two different scenarios to illustrate the impact that the presented subcircuit modifications have in the circuit simulator performance. The first scenario stands as the transient simulation of a sinusoidal voltage feeding a single memristor. The second scenario takes place as a 3×3 Resistive RAM (RRAM) crossbar array. In this crossbar scheme, while the target is to write a single cell, the whole operation affects the nine memristors. In both cases, with the objective of increasing the simulation loads, the transient simulation step length has been limited to 1 ps . Dealing with the computational load of different models, it is impossible to guarantee a specific performance; however, with the computation of a large amount of steps, we extract a general trend that covers most cases and makes visible remarkable differences between models.

To broaden the scope of the analysis, we have performed the test in both LTSpice and Spectre. The results show important speed-ups. In some cases, simulating the *Simmons tunneling barrier* memristor model, the simulation could

Figure 3.19: Speed-up results. Note that model [AP11] did not converge before applying our presented subcircuit modifications.

not converge even using time steps as short as 1 ps. On the contrary, when enhancing the original subcircuit proposed by the authors with the subcircuit improvements presented in Section 3.3, the simulation succeeded. In the cases where both conventional and new model simulations converge, we achieve a 2 times speed-up in a simulation using 1 or 9 memristors. All simulations were run in an i7 950 (3.6 Ghz) processor machine equipped with 12GB RAM. Table 3.3 shows the percentage of time saved when using our presented scheme instead of the conventional one, while Figure 3.19 corresponds to the graphical representation of these speed-ups. As can be seen, the most complex models obtain a greater benefit of the proposed subcircuit modifications in the reduction of the computational load.

3.7 Conclusions

Since the memristor was first built in 2008 a myriad of devices and models have appeared in the literature. However, the integration of these memristive devices in circuits is still not properly solved because their models are not mature and it is a hard task for the designer to control convergence, accuracy, speed and computational load of simulations.

In this Chapter we have presented a novel solution that bridges the gap between technology scientists, memristor model developers and circuit designers that want to incorporate memristors in their standard work flow. We have presented the MAF (Memristor Application Framework) that is conceived to assist the memristor community throughout all design stages: model design and characterization,

Table 3.3: *Speed-up results summary. Saved time ratio between our presented scheme and original models.*

Model	1 RRAM	1 RRAM	3 × 3 RRAM
	LTSpice	Spectre	LTSpice
Non-linear	20.84 %	3.15 %	20 %
Yakopcic	35.36 %	0.01 %	21.44 %
Modified HP-Simmons	12.28 %	22.83 %	27.15 %
HP-Simmons	NaN	NaN	45.87 %
TEAM	30.01 %	18.74 %	29.33 %
Eshraghian	11.16 %	3.57 %	4.5 %

subcircuit generation and simulation. It includes two simulators, i.e. *Device Characterization* and *Level Characterization*, that cover all the steps required to get a certain memristor model ready for application development, from model selection and validation to the fine adjustment of each parameter and variable even in the presence of process or conditions variations. MAF automatically generates a SPICE / Spectre subcircuit of the analyzed memristor to ease its integration in application circuits.

The framework includes a library with the best-known memristor models and it is simple to introduce improved upcoming models. More importantly, it helps model developers during the arduous task of tuning and testing their models to fit physical devices. Additionally, systematic modifications have been carried out to improve the SPICE code of each model providing better convergence (solving scenarios where conventional subcircuits generate overflows) and significant simulations speedups.

As illustrated by examples of use, the framework provides the designer with total simulation control, overcoming usual problems in the integration of memristors.

Regarding the following chapters, the MAF framework has been a key tool for the development of the SPICE model described in Chapter 4. In the same way, this support tool and methodology helped obtaining the thermal and stimuli dependence shown in the experiments gathered in Chapter 6. Through those experiments, MAF's versatility gets ratified. Moreover, since its first version was opened to the research and teaching community, many teachers and scientist have reported MAF use; thus, proving its usefulness as a stand-alone tool as well as giving the guide lines to future RRAM support software.

3.8 Related Contributions

- **Building Memristor Applications: From Device Model to Circuit Design.** *Fernando García, Marisa López-Vallejo, Pablo Ituero.* Nanotechnology, IEEE Transactions on, Volume:13 , Issue: 6, Pages 1154 - 116. November 2014. DOI: 10.1109/TNANO.2014.2345093. Impact Factor 2014: 1.8.
- **Memristor Application Framework**
Public Repository:
<https://github.com/fgr1986/maf>
Framework for the characterization and use of memristor compact models.
- **Model Validation and Simulation Framework for Novel Nanometer Devices.**
Fernando García-Redondo, Marisa López-Vallejo, Pablo Ituero.
Conference on Design of Circuits and Integrated Systems 2012 (DCIS 2012) Avignon (France).
- **A CAD Framework for the Characterization and Use of Memristor Models.**
Fernando García, Marisa López-Vallejo, and Pablo Ituero.
International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design 2012 (SMACD2012) Seville, Spain. 19 - 21 September 2012.

My story is a lot like yours, only more interesting 'cause it involves robots.

Bender, Futurama.

4

SPICE Modeling of Resistive Switching Technologies

4.1 Introduction

The creation of subcircuit compact components ready to be used by SPICE-like circuit simulators is a key issue for the design of architectures based on resistive switching technologies. In this sense, the trade-off between accuracy and simulation length becomes a critical factor when a large number of cells are computed within crossbars or substantial neural networks. We have seen how works like [BGS⁺15] accurately studied and reviewed the electrical behavior of memristor devices, gathering information of their internal dynamic processes for the design of a compact model. Section 3.2 summarized the characteristics of the most popular models.

Along this Chapter we will consider two different model approaches depending on how the subcircuit insights emulate the device behavior: physical models and phenomenological models. Physical compact models rely on the description of

current-voltage relationships together with their dependence on a given set of internal variables (conductive filament geometry, dopant volume elongation, ion migration probability, tunneling distances, etc.) matching the characteristics of a specific physical device. Depending on the grade of complexity, physical approaches include a broad variety of models, from the simpler and more flexible ones that are more inaccurate [BBB09, XZW⁺14, ZHL16], to the more accurate but complex ones able to model the physical behavior with high precision [ATB⁺15, LJH⁺15, KLH⁺15, Mir15, CY15].

The compendiums [EKC⁺12, GRLVI14, BGS⁺15] describe in more detail examples of both accurate and inaccurate models. Consequently, physical models that match different devices generally present accuracy problems, while complex approaches, with a suitable fitting of the physical component behavior, can only fit a narrow set of devices. Additionally, as the associated computational load increases with the model complexity, the trade-off between speed and accuracy restricts the use of some of the models. Moreover, in some cases the high non-linearity of the model behavior requires extremely short simulation time steps, cluttering the transient computation, and occasionally leading to convergence problems related to hard-switching conditions [GRLVI14].

On the other hand, phenomenological compact models focus on fitting the electrical magnitude relationships, and describing their evolution using mathematical variables which are not necessarily related to the physical variables. Phenomenological approaches obtain a broad flexibility regarding the range of described devices [GRLVI14]. More recent solutions update and improve earlier versions such as [YTSP13a, ZSK14, KRFK15], standing as a powerful solution for general modeling, and an effective application to neuromorphic computing [YMT⁺14].

In this Chapter we present a novel SPICE compact model based on the independent modeling of the device conduction mechanisms and device state. For that task we make use of MAF framework, presented in Chapter 3. First, taking advantage of the model characterization capabilities of the framework, we refine the presented model behavior. Second, using MAF's fitting module, we match the model response to two different physical devices. Thus, and as Figure 4.1 describes, we will work at the device abstraction level to provide the SPICE solution ready to be simulated within standard circuit simulators.

The presented model fully relies on physical magnitudes and their inter-relations, defining the tuple electrical and resistive switching behavior. While voltage and current thresholds are supported, the state modeling bases the device switching on electrical charge, flux or energy thresholds. It is applicable to both bipolar and unipolar devices, widening the range of devices that can be described, and therefore the presented approach rivals in flexibility with the

Figure 4.1: Chapter 4 scope and perspective within the Thesis.

most suitable phenomenological approaches [YMT⁺14, YTSP13a]. In turn the conduction module accurately fits the conduction mechanisms of the physical component.

Fully compatible with SPICE simulators, the developed model does not require implementation in Verilog-A or CMI to accurately match physical components. Although components written in Verilog-A or CMI allow more complex schemes to be emulated, they have a major drawback: widely used SPICE simulators do not allow pre-compiled code execution, restricting its simulation on several platforms.

The key features of the developed model are condensed in Table 4.1. This Table summarizes the key characteristics of our model comparing it against the most commonly used compact models presented in literature. The characteristics and detailed description of each of the considered models can be found in Section 3.2.

Once introduced the key capabilities of the presented model, we describe the development insights. Additionally, we dedicate Section 4.3 to the variability handling methods included within the modeling applicabilities. The simulation results fitting two different physical devices will be shown in Section 4.4. Finally, Appendix B includes the LTSpice source code describing a multi-level bipolar memristor.

Table 4.1: *Model Comparison Against Best Well Known Models in Literature.*

Model	Physical or Phenomenological	Requires Verilog-A	Bipolar or Unipolar
Proposed	Physical	No	Both
Arizona, Stanford & Beijing Univ.	Physical	Yes	Bipolar
Yakopcic's	Phenomenological	No	Both
Biolek's & Dopant Drift	Physical	No	Bipolar
VTEAM's	Phenomenological	Yes	Bipolar
Model	State Magnitude	Multi-level Aware	Pristine State Aware
Proposed	Consumed energy	Explicit	Yes
Arizona, Stanford & Beijing Univ.	CF width or gap	CF implicit	No
Yakopcic's	Auxiliar	Adaptable	No
Biolek's & Dopant Drift	Dopant area length	No	No
VTEAM's	Dopant area length	Adaptable	No
Model	Variability Aware	Degradation Aware	Temperature Aware
Proposed	Static & dynamic	Static & dynamic	Yes
Arizona, Stanford & Beijing Univ.	Static	Static	Yes
Yakopcic's	Static	No	No
Biolek's & Dopant Drift	No	No	No
VTEAM's	Static	No	No

Figure 4.2: Compact Modeling Design: bi-port component independently handling conduction module (left) and state module (right).

4.2 Developed Model

The model is composed of two different subcomponents which allow matching to the device behavior (Figure 4.2). The *Conduction Module*, addressing the memristor component interface through the ports *Plus* and *Minus*, models the dynamic resistance; in turn, the *State Module* handles the device state.

The component behavior is defined by the transient signal set, composed of the input voltage $v(t)$, the current flowing through the device $i(t)$, and the device state vector $s(t)$. This vector contains the information related to cycle count, switching thresholds, energy and resistance state. In this chapter we will simplify the notation of the electrical and state equations, and their dependence on time will become implicit: consequently, as an example, $v(t)$ will be denoted as v .

The equations describing the signal relation of the different submodules can be written using the generalized system

$$\begin{aligned} i &= f(v, s) \\ \dot{s} &= g(v, i, s). \end{aligned} \quad (4.1)$$

This equation system matches the behavior of most devices [GRLV14], and thereby both physical and phenomenological models specify f and g functions to define how the model runs.

4.2.1 Conduction Module Modeling

The Conduction Module is responsible for the computation of the current that flows through the device. This depends on the voltage and memristor state. Denoting the first component of the state vector $s^{(1)}$ as the discrete value referring the device level, within our developed model the current function

$$i = f(v, s) \quad (4.2)$$

Figure 4.3: Example of a simulated 2 bit multi-level device, with three different Schottky (l_0 , l_1 and l_2) and one ohmic conduction schemes (l_3).

takes the form, for different N conduction levels,

$$i = \begin{cases} f_1(v, s) & \text{if } s^{(1)} = s_1 \\ f_2(v, s) & \text{if } s^{(1)} = s_2 \\ \dots & \\ f_N(v, s) & \text{if } s^{(1)} = s_N. \end{cases} \quad (4.3)$$

This approach, completely covering the device state design space, allows multi-level device modeling such as [CLL⁺15]. Figure 4.3 shows an example of a 2 bit multi-level memristor simulated using the presented scheme. The bipolar device experiences three partial *SET* events followed by their corresponding partial *RESET*s. Compared against different approaches with no explicit multi-level definition (Table 4.1), the presented model eases tuning the conduction properties of each level, including degradation and variability. As an example, the Schottky characteristics exhibited in Figure 4.3 were captured with no effort.

For a device with two-resistance levels and no pristine state description, equation (4.3) gets simplified to the basic system

$$i = \begin{cases} f_{on}(v, s) & \text{if } s^{(1)} = s_{on} \\ f_{off}(v, s) & \text{if } s^{(1)} = s_{off} \end{cases} \quad (4.4)$$

where f_{on} and f_{off} are the low and high resistance state currents, respectively.

The state currents f_j are modeled using different methods depending on the specific device. As an example, the low resistance state dependence on voltage is usually simplified to a single resistance with temperature dependence in both oxide-based [LJH⁺15, KLH⁺15, YLH⁺15] and semiconductor based devices [ZJHD14].

Figure 4.4: *N*-state conduction module scheme using behavioral sources.

However, the model accepts different schemes to be used, such as an in-series diode-resistance approach, improving the simulation accuracy [BGS⁺14].

Even though some works model the conduction mechanisms of the higher resistance states using diode-resistance schemes [Mir15], in most cases the voltage-current relation is described by basic conduction processes like the ones described in Table 2.1 (Chapter 2). Modeling such processes using behavioral current/voltage sources allows us to accurately mimic the device behavior. Some devices require complex schemes where not only individual conduction processes take place but concurrent processes occur [YLH⁺15]. Our model fully covers this requirement allowing multi-contribution schemes, where distinct current sources contribute to define the global one.

In case the device conduction process is not defined by an analytical function, or further measurement data is required, we could model the varying resistance of the device as a behavioral resistor. The behavioral resistor value is determined by an auxiliary function defined using the circuit simulator “*user defined functions*” syntax [GRLVR15]. Taking advantage of the SPICE simulator capabilities, the behavioral resistor may refer a Piecewise Linear (PWL) function defining the voltage and state dependent resistance based on previous voltage-current measurements that build up the pairs $v_i - r_i$.

Summarizing, our model accepts individual / cumulative conduction schemes with which it models the current flow at multi-level states for different devices/technologies. The general subcircuit of a memristor device with N different resistive levels modeled using N independent current behavioral sources is shown in Figure 4.4. This approach allows us to include additional helpful effects such as the compliance current or control sentences [GRLVI14]. An example of a constraint current source modeling an N state device using the approach described in Figure 4.4 would be represented by the following equation:

$$i(v, s) = \max(-cI, \min(cI, f_j(v, s))). \quad (4.5)$$

Figure 4.5: Presented Device State Handler Module.

In this equation, the current through the device is limited by the compliance current parameter cI .

Similarly, the scheme shown in Figure 4.4 allows the modeling of parasitic-overshoot effects [LJH⁺15, YLH⁺15, SYK⁺15]. To that end, parasitic impedances Z_1 and Z_2 are attached in series and parallel to the global current source.

4.2.2 State Module Modeling

Several works propose models whose behavior depends on voltage or current thresholds that trigger different mechanisms unleashing a state change in the device [GRLVI14, YTSP13a, KRFK15].

In this work we propose an extension of this concept, broadening the threshold mechanisms that define the switching. Figure 4.5 describes the components the *State Handler Module* is composed of. Basically, the device state will depend on the previous state and the device electrical inputs. This scheme is compatible with the simple approach of voltage/current boundary threshold, leaving the magnitudes energy, charge and flux out of the threshold management and using instant voltages and currents to trigger the switching mechanisms.

Moreover, the model is enriched, allowing energy, charge or flux to be used on the simulated device as the triggering magnitude that releases the switching events. Following the idea shown in [JJL⁺12, JCM12, YTSP13b, GO15, MMMD⁺15], this model computes the energy, charge or flux levels at which the device experiences a state change. Using individual submodules, we calculate the energy, charge or

Figure 4.6: Characteristic hysteresis loops of a charge controlled memristor fed with different sine voltages. The compact model was built following the scheme shown in Figure 4.5, using a charge handling module to control the device state.

flux applied to the device in order to define the switching, the cycle and event count, building up the whole state vector s . Following this approach Figure 4.6 shows the characteristic pinched hysteresis loop for a charge controlled device similar to [BBB09].

Compared to Verilog-A implementations [LJH⁺15, KLH⁺15], where different variables can be used, in SPICE-like subcircuit design we need to use signals to store the related information. Therefore, the energy, charge, flux or cycle/event count computations require individual signals, usually in the form of node voltages, to define their value.

As we introduced in Chapter 3, instead of using capacitors to integrate magnitudes [XZW⁺14], the energy computation can be accelerated [GRLVI14] using the time integral function provided by the SPICE-like simulator *idt*. On the other hand a user defined function, *switching*, performs the switching conditions verification. Taking energy as the magnitude managing the switching, *switching* would be described by equation (4.6).

$$s_{next}^{(1)} = \text{switching}(E, P, s_{current}^{(1)}), \quad (4.6)$$

There we name P as the matrix whose elements p_{jk} define the energy thresholds to switch from state j to state k . This state-controlling scheme was used for the multi-level simulation shown in Figure 4.3. Appendix B includes the SPICE code associated with the *switching* function.

Pristine State Modeling

Using the presented conduction and switching schemes allows us to model the pristine (conductive filament preforming) state. The ability of including this additional state enriches the model capabilities.

By defining the conduction mechanism and forming energy (e_p), the pristine state can be easily included in the netlist:

$$s_{next}^{(1)} = \begin{cases} s_{pristine}^{(1)} & \text{if } E < e_p \\ switching(E, P, s_{current}^{(1)}) & \text{otherwise.} \end{cases} \quad (4.7)$$

Cycle and Event Count Computation

As explained in Section 4.3, the ability to compute how many switching events a device has suffered is essential to accurately model its degradation. Using the magnitude that defines the switching —energy, charge or flux— we propose a novel scheme that allows the device cycling count to be stored. For the sake of simplicity, equations 4.8 and 4.9 will suppose devices with two states and an initial HRS. Let E be the computed energy, parameters p_{set} and p_{reset} the required energy levels to perform the *SET* and *RESET* respectively, and the parameter p_{cycle} the required energy per switching cycle, defined as $p_{cycle} = p_{set} + p_{reset}$. For unipolar devices we can compute the state vector s as

$$\begin{aligned} E &= \int_0^t v \times i \, dt \\ cycles &= \left\lfloor \frac{E}{p_{cycle}} \right\rfloor \\ extra_set &= \begin{cases} 1 & \text{if } (E - cycles \times p_{cycle}) > p_{set} \\ 0 & \text{otherwise} \end{cases} \\ events &= 2 \times cycles + extra_set \end{aligned} \quad (4.8)$$

The same methodology can be applied to bipolar devices. However, positive contributions E_+ are required to be treated independently from negative contributions E_- :

$$\begin{aligned} cycles &= \left\lfloor \frac{E_-}{p_{reset}} \right\rfloor \\ events &= \left\lfloor \frac{E_-}{p_{reset}} \right\rfloor + \left\lfloor \frac{E_+}{p_{set}} \right\rfloor \end{aligned} \quad (4.9)$$

Soft Switching Events

We can avoid hard-switching events by introducing an additional variable b that softens the transition between states. As we saw in Section 3.2, this variable would act similarly to the auxiliary elongation variables that both *linear ion drift*, *non-linear ion drift*, and *Yakopcic's* models use. Depending on the physical device to be emulated, the variable b can be defined in a multitude of ways.

For a device experiencing a transition from state $s_j^{(1)}$ to state $s_{j+1}^{(1)}$, where i_j and i_{j+1} are their corresponding current levels, the current switching can be softened if the global current is redefined as

$$i = \dots + w_j(b) \times i_j + w_j(1 - b) \times i_{j+1} + \dots \quad (4.10)$$

where $w_j(b)$ are the softening functions that control the different state's current contributions. The softening variable b can be described based on the corresponding energy thresholds

$$\begin{aligned} a1 &= \left\{ \frac{E_j}{e_{a1}} \right\} \\ a2 &= \left\{ \frac{E_j}{e_{a2}} \right\} \\ b &= \begin{cases} \min(1, \max(0, a1)) & \text{if } s^{(1)} = s_j \\ 1 - \min(1, \max(0, a2)) & \text{if } s^{(1)} = s_{j+1} \end{cases} \end{aligned} \quad (4.11)$$

where E_j is the total/partial energy (depending on the device type), and e_{a1} and e_{a2} the energy thresholds required to perform the transitions $s_j^{(1)} \rightarrow s_{j+1}^{(1)}$ and $s_{j+1}^{(1)} \rightarrow s_j^{(1)}$ respectively.

The softening functions $w_j(b)$ make use of the state variable b to control the partial contributions of i_j and i_{j+1} . The model designer is able to customize these functions to better adapt the switching mechanism to the physical measures.

Figure 4.7 describes one of the possible softening schemes for bi-state devices, together with the produced effects. In these example functions the partial contributions depend on two thresholds, b_th_set and b_th_reset , and the softening parameter p , which manages the switching slope in the following expressions:

$$\begin{aligned} f_{set}(b) &= \begin{cases} 1 & \text{if } b \geq b_{th_set} \\ \left(\frac{b}{b_{th_set}} \right)^p & \text{if } b < b_{th_set} \end{cases} \\ f_{reset}(b) &= \begin{cases} 1 & \text{if } b \leq b_{th_reset} \\ \left(1 - \frac{b - b_{th_reset}}{1 - b_{th_reset}} \right)^p & \text{if } b > b_{th_reset} \end{cases} \end{aligned} \quad (4.12)$$

The effects caused by these soft-switching functions are represented by the detail view located in Figure 4.8.

Figure 4.7: Example of soft-switching mechanism: Different p values produce the shown effects on the auxiliary variable b during set and reset processes.

Figure 4.8: Example of soft-switching mechanism: the softening functions make use of the auxiliary variable b to soften the state switching. Different thresholds and p values produce the shown effects during reset process.

Figure 4.9: Local temperature within the CF during the SET process and computed voltage-current relation, varying circuit temperature T_0 .

4.2.3 Extensibility

The modular design of the presented scheme allows its extension in order to consider additional phenomena.

As an example, temperature dependence has proven a key factor in the device behavior [LJH⁺15, YLH⁺15, YGCR15]. Each of these extensions can be considered as independent behavioral sources for their later reference in dependent signals. The cited temperature influence can be directly incorporated within our model in both conduction mechanisms and energy thresholds computation by simply altering the nominal expressions of the behavioral source. Figure 4.9 presents the results after the incorporation of local thermal behavior due to Joule-heating [LJH⁺15, KLH⁺15].

This modularity has been key in order to emulate important behaviors such as variability or device degradation mechanisms (Section 4.3). Future extensions that allow the emulation of non-considered effects such as explicit current-overshot modeling —Section 2.5— would also take advantage of this modular approach.

4.2.4 Verilog-A Implementation

Once the conduction mechanisms and the threshold parameters are defined, the present model can be easily ported to the Verilog-A language, with the consequent speed up of circuit simulation. Each behavioral source is directly translated into an independent variable, following the approach shown in [LJH⁺15].

Additionally, Verilog-A is more flexible regarding the use of user defined functions or variable data types (such as arrays), which eases the implementation of the methods adopted to reproduce variability effects shown in the next section.

However, even though some circuit simulators such as Cadence Spectre allow Verilog-A code execution, widely used SPICE simulators like LTSpice do not allow Verilog-A co-simulation, accepting SPICE-like code only. Consequently, the research community may benefit from the SPICE developed compact model.

4.3 Modeling Non-Ideal Effects: Variability, Life endurance and State Retention

Resistive switching technologies suffer from intrinsic sources of uncertainties that may compromise the circuit reliability. Variability, device degradation and data retention are among the most concerning issues to be addressed. It is then essential to develop accurate compact models that take into account the effects caused by these uncertainty sources. The hard task of translating the dynamic relations between device characteristics/state, cycling and working conditions is complicated but necessary. With that aim, and complementing the presented SPICE model, in this section we describe a novel emulation scheme that allows the emulation of physical complex effects caused by the degradation of devices, data retention and variability.

4.3.1 Related Work

Variability Modeling

Variability was one of the first uncertainty sources whose effects on RRAM technologies were reported. Thus, several research teams have focused on the development of compact models accounting for RRAM variability [GVR⁺16, GWZ⁺15, PPS⁺15, EBC⁺15, LJH⁺15]. Before introducing the most complete works, we describe the three distinct kinds of variability that are expected to occur in resistive switching devices [EBC⁺15]:

- *Inter-Device Variability.* Device to device variations in size, thickness, ion-concentrations, etc.

- *Intra-Device Variability*. Small cycle-to-cycle fluctuations caused by the stochastic nature of generation and recombination of oxygen vacancies and ion migration.
- *Read Current Fluctuations*, also called *Random Telegraph Noise (RTN)*, is a variation in the measured current under constant bias during reading operations. It can be caused either by the electron capture and emission processes that inherently exist in oxides with high defect concentrations, or by atomic changes in the conducting filament.

Inter-Device Variability can be handled via the Monte Carlo mechanisms included in the circuit simulator. Therefore, we will focus on how the other approaches have been modeled in the literature.

Verilog-A Modeling Compact models developed in Verilog-A language require from either a specific Verilog-A simulator or a Spice-like simulator that first compiles the model and then uses the binary to be simulated within the circuit. Even though some SPICE simulators are able to handle this kind of models, widely used simulators as LTSpice lack from the ability of including Verilog-A code. The most complex but accurate intra-device variability approaches are the models developed at Arizona University, [JWY⁺16] and at Stanford & Beijing Universities [LJH⁺15]. These two models thoroughly describe the statistical variations of the conductive filament evolution. Using Monte Carlo simulations, these models allow obtaining the statistical values of both *High Resistance State (HRS)* and *Low Resistance State (LRS)*, achieving accurate distributions. An improved version of the Stanford & Beijing's variability model has been recently reported in [LHG⁺16]. Here, not only intra-device but device-to-device variability is considered. However, these models have three major drawbacks:

- First, they are developed in Verilog-A language.
- Second, they do not consider RTN.
- Third, variability modeling is static, and therefore not dependent on the stress time or device cycle-to-cycle.

SPICE Modeling Opposite to Verilog-A models, some research groups have focused on purely SPICE approaches [CYMMR⁺16], and to study intra-device statistical variations only use native SPICE components. However, such studies

- Only account for conductivity modeling, leaving out the device's switching.
- Only consider intra-device variability.
- Again, variability is only statically modeled.

Degradation and Data Retention Modeling

The degradation of RRAM cells shortens their lifetime. The figure of merit that determines the cell endurance is the maximum number of supported switching cycles. As seen in [MCYC16], aside the physical characteristics of the cell, RRAM endurance is closely related to the programming/reading operation schemes. Therefore the characteristics of the current feeding the cells have great influence in the cell lifetime.

The same conclusions are extracted from [HCW⁺13], the first analytical model for endurance degradation in metal oxide RRAMs. This model, developed by researchers at the Beijing University is an early approach that later would evolve into [LJH⁺15]. This first approximation included an endurance degradation formulation based on the progressive *over SET* and *over RESET* effects that worsen the device's conductive filament. Similarly to the variability modeling, this solution, though effective, does not provide variability dependence on time. By considering both effects as independent, variability is not altered no matter the level of degradation. Moreover, it is to be noticed that in this model the combination of the endurance degradation modeling with the Verilog-A analytical model—that accomplishes the transient operation—is carried out outside the circuit simulator. This demands a feedback loop between an external handler and the circuit simulator to allow *over SET* and *over RESET* effects to be calculated.

Complementing Beijing's model, in [BAW⁺15] Ielmini's team presented an analytical model describing the relationship between the device's endurance and the voltage characteristics of the cycling operation. This model also states that the endurance of a device highly depends on the operation characteristics.

4.3.2 Proposed Solution

In accordance with the rest of the model, we present a solution that can be fully deployed in SPICE-like simulators, not requiring additional code to be compiled—Verilog-A, CMI—and most importantly, without requiring from external tools to feedback the circuit simulator, at every step of the loop, with the externally computed degradation [HCW⁺13].

Variability, Degradation and Data Retention Modeling

Three distinct kinds of variability may occur in RRAMs. Each of these types of variability [EBC⁺15, GWZ⁺15, BAW⁺15], as well as the degradation [HCW⁺13, MCYC16] of the device endurance depend on

- Device's physical characteristics.
- Device's state (pristine, s_{on} , s_{off} , s_i).

- Electrical stress: Programming/read operation characteristics and number of cycles or time.

In order to provide trustful and accurate results, the proposed modeling approach takes into account all of the above dependences. The voltage or energy thresholds defining the switching, the current flowing through the device or any other signal requiring variability injection can be modeled as sources whose value is composed of nominal and variable contributions. Being P the matrix p_{jk} containing the thresholds that define the switching from the state $j \rightarrow k$, the variable contribution is described through the functions $r_j()$, $w()$:

$$\begin{aligned} i &= \begin{cases} r_1(v, s) f_1(v, s) & \text{if } s^{(1)} = s_1 \\ \dots & \\ r_N(v, s) f_N(v, s) & \text{if } s^{(1)} = s_N \end{cases} \\ \dot{P} &= w(v, i, s, P). \end{aligned} \quad (4.13)$$

By using the above equation we guarantee that the dependence on the device's characteristics and state, cycling and electrical stress is provided.

In the SPICE subcircuit, r_i , w can be modeled as probability density functions [GVR⁺16, GWZ⁺15] coming from analytical models. In addition, we can extract from physical measurements the mean and standard deviation values. To generate the probability density functions we make use of random number generator functions built into the circuit simulators. If the circuit simulator does not support this feature, we can extract the random pattern by adding an additional zero amplitude voltage source affected by uniform/shaped noise.

On the other hand, in cases where the random data of noise and parameter variability do not follow a suitable function, we can use PWL or behavioral user defined functions to directly use the scattered data. Following this approach, these sources can directly inject random patterns extracted from physical measurements. Another alternative to the PWL defined sources are the specific noisy source components that allow values defined via input files.

Compared against the models previously presented in the literature, the proposed modeling scheme admits defining how the device degrades or how variability changes throughout time/cycling or stress. The equations (4.13) also provide the ability to define how the device gets degraded depending on its workload, and consequently, the determinability of its stored data. Summarizing, we provide a more accurate variability/degradation modeling:

1. Static variability/degradation emulation, extracted from physical measurements [GWZ⁺15, LJH⁺15, JWY⁺16] and fitting statistical properties non-varying through time.

2. Dynamic variability/degradation emulation, modeled using time/cycle and electrical stress dependent user defined functions. The statistical properties can evolve through time. Consequently, the variability and degradation of the device represent more accurately the real behavior of an RRAM working through a specific period of time under given stress conditions. Therefore, statistical characteristics such as standard deviation (σ) and mean (μ) values become dynamic ($\sigma(t, s, v, i)$, $\mu(t, s, v, i)$).

Moreover, as reflected in equations (4.13), the explicit declaration of the different conduction mechanisms eases the tuning of variability and degradation functions at each resistive level. It is to be highlighted that to the date, no RRAM's SPICE model presented in literature matches the capabilities of the presented approach.

4.4 Fitting and Simulation of Physical Devices

4.4.1 Physical Devices

To show the model features, two physical devices have been studied and fitted: a bipolar *Cu/a-SiC/TiN* and a unipolar *Ni/HfO₂/Si* device. Moreover, these devices not only illustrate unipolar/bipolar switching behaviors, but also Ox-RAM and CBRAM memristor structures, highlighting the model capabilities.

The unipolar *Ni/HfO₂/Si* Ox-RAM device was manufactured by M. Bargallo and F. Campabadal at the IMB-CNM (Barcelona). The measurements were accomplished during a research stay at UAB University Labs (Barcelona), following the measure procedure and methodology suggested by M. Nafria and J. Martín Martínez research teams (A. Crespo-Yepes and M. Maestro).

Most of the bipolar *Cu/a-SiC/TiN* CBRAM measurement data were provided by K. De Groot research team (Robert P. Gowers, K. Morgan, and J. Fan) from Southampton University. The remaining measures were taken during a research stay following Robert P. Gowers advise, also at Southampton University Labs.

4.4.2 Energy Threshold

Figure 4.10 shows the measured energy levels required to perform consecutive *SET-RESET* processes for the considered unipolar device. Similarly, Figure 4.11 regards the bipolar device. The simulations shown in the following experiments have been performed using Cadence Spectre and Linear Technologies LTSpice circuit simulators, replicating the measurement experiments applied to the physical devices. The full source code of those models, together with their SPICE and

Figure 4.10: SET and RESET scatter and histogram of measured energy levels required to perform switching in the unipolar device. The values have been normalized using the maximum value of both processes.

Figure 4.11: SET and RESET scatter and histogram of measured energy levels required to perform switching in the bipolar device. The values have been normalized using the maximum value of both processes.

Spectre implementations can be found in [GRLVR15]. The parameter set fitting was automatically accomplished using the MAF simulator [GRLVI14].

4.4.3 CBRAM Bipolar Device

The first fitted device shows a reverse Schottky emission as its basic HRS conduction process, even though it displays a leakage phenomena. The model is able to fit a minimum asymmetric behavior regarding its polarity. The LRS is modeled using two different resistors depending on the voltage polarity. Figure 4.12a presents the simulated subcircuit, while Figure 4.12b displays the achieved fitting compared against measured data. The simulated voltage follows the voltage source that fed the device during its characterization.

4.4.4 Ox-RAM Unipolar Device

The conductivity modeling of the considered unipolar device follows the guidelines from [Mir15]: two pairs of in series resistance-diodes match both LRS and HRS (Figure 4.13a). The state handling approach described by System (4.8) perfectly fits the switching behavior of unipolar devices as shown in Figure 4.13b. In this case, and emulating the physical measurement procedure, the simulation input voltage reproduces the sawtooth stimulus feeding the memristor.

(a) Subcircuit of bipolar CBRAM using four different conduction submodules: off and on states for both positive and negative voltages.

(b) Bipolar CBRAM, voltage-current relationship together with the voltage-current transients. On the upper graph the different slopes show how the asymmetric behavior, depending on the input voltage polarity, was perfectly mimicked.

Figure 4.12: Bipolar CBRAM subcircuit and $i(v)$ relation

(a) Subcircuit of Unipolar Ox-RAM using four different conduction submodules: off and on states for both positive and negative voltages.

(b) Unipolar Ox-RAM, voltage-current relationship together with the voltage-current transients.

Figure 4.13: Unipolar Ox-RAM subcircuit and $i(v)$ relation

Figure 4.14: Simulated $v - i$ curve showing the RTN dependence on device's state and input voltage. Scatter displays the physical measurements.

4.4.5 Non-Ideal Effects

To highlight the capabilities of the presented approach, we gather here different experiments progressively emulating different kinds of variability and endurance, therefore describing the complex relationships between degradation, variability, device physical state and time/cycle, suffered by real world devices. The experiments refer to the bipolar $\text{Cu}/a\text{-SiC}/\text{TiN}$ device, as we had access to the early pristine state prior the CF electro-forming, therefore highlighting the model feature of adapting the variability scheme to the device physical state. Taking advantages of the schemes presented in Section 4.3, the three representative kinds of variability are considered: RTN, intra-device variability and inter-device variability. These non-ideal effects are complemented by a final experiment considering the degradation of the device CF. It is to be noticed that in all the experiments the emulated variability and lifetime degradation depends on the electrical stress, time/cycle and the memristor physical state, thus meeting the real devices evolution.

First, the emulated dependence of RTN on both the device state (pristine, HRS or LRS) and the input voltage is shown in Figure 4.14. RTN was modeled using noise injectors defined by probability density functions. We would like to highlight how the noise levels perfectly match the patterns typically extracted from physical measurements when semiconductors analyzers are used.

Second, we complemented this RTN experiment by simulating intra-device variability, taking into account its dependency on the device's cycling. As seen in Figure 4.15, this experiment shows how variability worsens through time. The evolution of the conductance and threshold parameters was accomplished using PWL functions directly related to physical measurements.

Figure 4.15: Simulated $v - i$ curve of a device exhibiting both RTN and intra-device variability. The variability of the conduction mechanisms and switching thresholds varies with the device's cycling and feeding voltage.

Third, we used the MC circuit simulator capabilities to vary the device-to-device parameters. Thus, and as can be seen in Figure 4.16, we emulated *Inter-device* variability. The curves show two adjacent devices with distinct parameters, that suffer from different levels of intra-device variability.

Finally, Figure 4.17 shows how degradation can be modeled as a function dependent on the RRAM programming characteristics. This final example combines every kind of variability effects with the emulated degradation, using noise injectors, behavioral sources, user defined functions and PWL functions. For these simulations only SPICE-level simulators were used, without requiring the use of

Figure 4.16: Simulated $v - i$ curve of a device exhibiting both RTN, intra-device variability (cycle-to-cycle), and inter-device variability (devices A and B).

external tools or Verilog-A code. The full code regarding these experiments can be found in [GR16].

Figure 4.17: Degradation dependence on the characteristics of the programming pulses for a device suffering RTN, intra-device and inter-device variability. The variability and degradation ratio vary with the cycling and feeding voltage.

4.5 Conclusions

We have provided a customizable, physical memristor SPICE compact model, that accurately fits both unipolar and bipolar devices without requiring a Verilog-A compatible simulator. The conduction modeling allows multi-level description, each level being able to be characterized using diverse contributions. The component state modeling is based on the processes that trigger resistive switching (device electrical thresholds, energy, charge or flux), while omitting complex geometry/internal process computations.

This design simplifies the arduous work of translating physical device characteristics to the circuit compact model, and therefore helps device manufacturers to simulate state of the art devices. Not only is the model able to handle variability but also it describes how variability/parameter degradation evolves with time, making durability simulations straightforward. Its modular scheme allows the consideration of additional phenomena like temperature effects, improving memristor SPICE simulations and making the analysis more reliable. The condensed contributions of the presented model are as follows:

- Accurate modeling of dynamic resistance cleanly mimicking physical device response.
- Effective switching behavior for Ox-RAM and CBRAM bipolar/unipolar devices.
- Availability of cycle and switching event count information for all time, becoming a resource for additional characteristics such as variability.
- Conventional variability awareness: the model provides mechanism to emulate variability, using static probabilistic density functions —whose main characteristics remain time and cycle invariant.
- Provision of variability dynamics and resistive state retention handling, defining how the device degrades through time/cycle stress. Both parameter values and statistic properties vary through time. Thus, our model allows the user to establish how the variability functions and mechanisms vary during a transient depending on the device cycle/time.
- Explicit support for multi-level storage. Unlike CF based models [LJH⁺15, KLH⁺15, CY15], the explicit definition of multiple states allows more detailed level descriptions. This also allows the modeling of *Pristine State* -the initial *High Resistive State (HRS)* prior any electroforming-.

- The full source code for those bipolar and unipolar devices, together with additional materials can be found at the research group resources page [GRLVR15] as well as at the developer's repositories [GR16].

Finally, the chapter concludes with the validation of the model against two different physical devices: Ox-RAM unipolar and CBRAM bipolar devices, therefore highlighting the model versatility.

4.6 Related Contributions

- **SPICE Compact Modeling of Bipolar/Unipolar Memristor Switching Governed by Electrical Thresholds.** *Fernando García-Redondo, Robert P. Gowers, A. Crespo-Yepes, Marisa López-Vallejo, and Liudi Jiang.* Circuits and Systems I: Regular Papers, IEEE Transactions, Volume: 63, Issue: 8, Aug. 2016. DOI: 10.1109/TCSI.2016.2564703. Impact Factor 2015: 2.40

- **VLSI Memristor/RRAM SPICE Compact Model**

Public Repository:

https://github.com/fgr1986/vlsi_memristor_compact_model

SPICE and Spectre model of unipolar and bipolar memristors, considering endurance and variability.

- **CAS-T Lecture: SPICE Compact Modeling of Bipolar/Unipolar Memristor Switching Governed by Electrical Thresholds.** *Fernando García-Redondo, Robert P. Gowers, A. Crespo-Yepes, Marisa López-Vallejo, and Liudi Jiang.*

IEEE international Symposium on Circuits and Systems, (ISCAS 2017), Baltimore, MD, Usa, May 28-31 2017.

Part II

Characterization and Analysis of RRAM-CMOS Hybrid Circuits

At The Risk Of Sounding Negative, No.

Leela, Futurama.

5

Reliable Simulation of RRAM-CMOS Hybrid Circuits

5.1 Introduction

The first part of this thesis has covered the modeling and characterization of resistive switching devices. With the aim of designing, simulating and characterizing circuits involving CMOS and RRAM technologies, in this part we move up from device to circuit level. In this sense, we will analyze how different uncertainty sources —introduced in Chapter 2— alter the trustworthiness of the circuit, focusing on reliability and performance. In order to achieve functional RRAM-CMOS hybrid systems, the problems affecting both RRAM and CMOS technologies cannot be independently managed, and require from a unified process gathering design, manufacturing, testing and verification.

First, we have the uncertainty sources affecting CMOS circuits design. The designer has access to the technology *Process Design Kit (PDK)*, which includes thorough documentation about tested CMOS models characterized against PVT

variations. Consequently, standard design methodologies usually consider variations. However, when additional error sources such as radiation or aging appear, traditional design flows either do not contemplate their characterization or independently analyze the effects caused by each of those issues. Let us focus on short-node CMOS technologies, in which the effects of aging and radiation are particularly important. Transistor aging compromises the circuit performance over the medium and long term, and therefore its characterization and mitigation is of great importance. Starting with BSIM3 Level 49 or 53, SPICE models are allowed to include, though not in a completely accurate way, the modeling of *Hot Carriers Injection (HCI)* and *Negative Bias Temperature Instability (NBTI)* [PBG⁺]. Therefore, some PDKs include aging models approximating the evolution of transistor parameters through time. However, these aging models tend to obviate the stress caused to the transistor by the applied voltage and working temperature. Regarding radiation, the most complex studies involve *Single Event Transients (SETs)* together with some parameter variability [KBK⁺11, FHS15], or the effects of *Total Ionizing Dose* associated with aging [SBK15]. In order to obtain more accurate characterizations, the most complete research studies combine them by pairs —radiation against variability, aging versus power supply, etc. We dedicate Appendix C to provide the context on how radiation and the pair radiation-aging have been traditionally treated in circuit design. There we describe how for CMOS technologies, although aided by design methodologies and tools, radiation simulation is a problem that is insufficiently supported: it requires the manual intervention of the user, and therefore the automation of the thousands of analyses is a pressing need.

Thus, the first motivation behind this Chapter is that, although the modeling and simulation of PVT variations in CMOS technologies is fully established, the accurate characterization of aging and radiation effects, and more importantly the combined effects of these uncertainty sources is still a matter of study. Moreover, at the day of writing this thesis, the simple analysis of radiation effects leaving out other uncertainty sources requires the designer to manually emulate the radiation occurrence. Moreover, to compute the performance and ensure the correct behavior of a given circuit, this manual radiation emulation requires to be accomplished in each susceptible node, varying the impact/accumulated charge as well as the working and environmental conditions. Besides these considerations, CMOS technologies beyond 65 nm demand variability analysis paired to radiation hardness level characterization studies [KBK⁺11]. Furthermore, the circuit designer lacks specific tools to incorporate emerging devices and analog components into the standard design flow, and more especially, in their integration within critical applications. In this sense, system level analyses would be benefited from more sophisticated error input models: if each system module is accurately

characterized, a *Single Event Rate (SER)* study would be enough to characterize complex circuits. To this effect, an in depth statistical analysis of the conditions where functional errors are generated by radiation occurrences is required.

On the other hand, traditional circuit level analyses did not cover the RRAM problems and limitations seen in Section 2.5. For example, and with reference to radiation and aging most studies characterizing RRAM circuits focus individually on each uncertainty source. Therefore both RRAM degradation [HCW⁺13, LLN15, Iel16, KB16, PAR16] and radiation are then individually approached [HCW⁺13, LLN15, Iel16, KB16, PAR16]. Concerning radiation, most studies individually consider SET [CKP⁺14, BHS⁺14, LMBY15] and TID [LWHL15, MFH⁺15] effects. But in the RRAM's case not only degradation and radiation are the concerns to be addressed. The immaturity of the technology has traditionally left out of the scope radiation, focusing on manufacturing improved devices to solve other problems such as data retention or low power operation. At circuit design level, and probably because of their short life since its discovery, traditional circuit level analyses did not cover the full of RRAM problems and limitations. Here, early predictive SPICE models are still immature, and far from being part of solid RRAM PDKs. Then, the second motive inciting the work presented in this Chapter is that circuit designs based on these predictive models lack the methodologies and support tools for the full design of complete systems.

Furthermore, when integrated with CMOS technologies in order to build RRAM-CMOS hybrid circuits, the effects of uncertainty sources in both technologies may catalyze unexpected functional errors compromising the whole system behavior. Therefore, the reliability and performance of RRAM-CMOS hybrid circuits must be evaluated from a global perspective, where circuit designers consider, characterize and mitigate the harmful effects due to the combination of uncertainty sources. Summarizing, the reliability of an RRAM-CMOS hybrid circuit can only be guaranteed by assessing the whole picture at early design stages.

5.2 Overview of the Presented Solution

To support the above requirements, in this chapter we propose a two-fold solution: an analyses strategy and a support framework for designers that require accurate analyses, studying the reliability of RRAM-CMOS hybrid circuits. Based on a multi-metric evaluation paradigm, our solution aims to study the combined effects that both RRAM-CMOS intrinsic limitations and external environmental factors, such as temperature or radiation, produce on the systems.

To that end, we redefine the analyses to be carried out in order to characterize both performance and reliability, and provide a simulation framework that accomplishes the study of RRAM-CMOS circuits working under multiple uncertainty

Figure 5.1: Chapter 5 scope and perspective within the Thesis.

sources. The proposed reliability analyses compute the circuit characteristics facing non-ideal conditions (radiation environments, variability, aging) and their dependence on different device, circuit, operation or environment parameters.

Consequently, and as Figure 5.1 displays, making use of the emerging device models and support tools introduced in the first part of this thesis, we now present a solution targeting circuit level design stages. Compared against traditional approaches, where a circuit functioning error is defined either by bit upsets [KBK⁺11, FHS15] (CMOS) or large changes in cell resistive levels [CSL⁺16, CTKP16] (RRAM), our proposal employs additional, user defined metrics. CF alterations, analog signals deviation, leakage currents, signal delays or frequency variations, may then define the design trustworthiness and performance of RRAM-CMOS hybrid applications, widening the target circuits to be analyzed.

To aid the circuit designer we provide a circuit-level simulation framework. By automating the characterization processes we save design time, and avoid possible errors created by the designer in stages requiring manual interaction. The tool, interfacing with the low level circuit simulator, employs user defined metrics to measure the circuit response, while extracting the *critical* values of key parameters that cause misbehavior or, on the contrary, optimize the circuit behavior. Bit flip in memories, offsets or deviations in DACs or voltage references, or frequency shifts inside oscillators are examples of applicable metrics. The critical values of, for example, collected charge that cause this bit flip, worst particle impact instants on sequential circuits, or the maximum temperature at which the offset in a voltage reference does not exceed a threshold, are automatically computed by the analysis module.

Till now, we have reviewed the context and motivation, describing how traditional approaches no longer fulfill RRAM-CMOS hybrid circuit designers. The remaining of the chapter first presents the proposed analyses strategy and framework supporting the methodology. Second, the framework is validated against experimental data. Finally, and to illustrate the proposal, several experiments

Figure 5.2: Scheme representing a standard workflow accomplished during the stages of a traditional methodology. Grey boxes represent stages requiring user actions. Optional stages are plotted with discontinued borders.

characterize different circuits. First, we analyze a digital circuit. Next, an analog module is characterized. And third, we present a full RRAM-CMOS circuit characterization.

5.3 Strategy for Reliable Analyses of Circuits

The computational complexity associated to the automated exploration of the design space totality, involving the evaluation of every single possible scenario, dealing with each parameter particularization, environmental configuration and working condition leaves the approach completely out of consideration. The simple task of manually generating the independent netlists associated to each scenario involves a huge amount of resources. Given this unmanageable situation, it becomes clear that the design space definition needs to be carefully declared.

On the other hand, the simplified analyses accomplished within traditional methodologies by well established tools are not able to correctly describe an RRAM-CMOS based circuit in presence of multiple uncertainty sources. In this case, a more complex and refined approach is required. Figure 5.2 represents the workflow that usually accompanies these traditional methodologies. Our proposal addresses this dilemma relying on three key ideas:

Figure 5.3: Scheme of the presented analyses strategy. Blue boxes represent the stages at which models, metrics or the design space are defined. Analysis stages are represented by white boxes.

- The use of more suitable device models.
- The simultaneous evaluation of multiple user-defined performance-reliability metrics.
- The efficient definition of the design space.

With these three concepts, and making use of the designer knowledge, we build up an analyses strategy that accomplishes accurate studies otherwise computationally untraceable. To support the proposed strategy, and automating the whole characterization process, we present an open-source simulation and analysis framework.

Figure 5.3 gathers the proposed three-stage strategy to follow during a reliability-oriented RRAM-CMOS circuit characterization.

Models Redefinition

Even the most complete device models may lack from the ability to emulate a specific and required capability. Regarding reliability, a device model may not contemplate the device behavior under the influence of a given uncertainty source. To overcome this limited scope, and seeking the evaluation of the combined effects

Figure 5.4: Models redefinition scheme.

of the different non-desired effects, we propose a *Models Redefinition* stage. This stage, represented by Figure 5.44, supplements the circuit analyses—defined in a later stage—with additional model capabilities not considered before.

In this sense we can make use of the most up to date research works to revamp the standard models into enriched ones, now incorporating the new emulated skill. To that end, redefinition schemes usually employ either voltage/current controlled voltage / current sources (VCVS, C CVS, VCCS, CCCS) [SBK15] or independent behavioral injectors sources [RLV14, GRGCY⁺16].

For instance, the large trajectory in CMOS technologies has promoted to incorporate aging effects in standard PDK models. However, the use of the most accurate models [MMM⁺14], or the consideration of additional facets such as TID effects [SBK15] would widen the scope of later analyses. As will be shown in Section 5.8, this methodology can be used to either substitute or enhance RRAM models, redefining their behavior introducing, for example, endurance or TID effects [WWBH⁺14, MFH⁺15].

User Defined Metrics

The second stage of the proposed strategy is based on the simultaneous evaluation of multiple performance and reliability metrics: Individual and independent *error definitions*—such as a memory bit flip—can no longer be applied, as this approach is too simple to gather all required aspects. Therefore, in our proposal the designer defines the evaluation metrics according to the circuit characteristics.

Once defined, and as described in Figure 5.5, these metrics will be used by the analyses algorithms to assess the robustness of the circuit working under non ideal conditions, as well as to measure its performance.

Figure 5.5: Evaluation of user defined metrics.

By interacting with the SPICE-level circuit simulator, we are able to analyze, among others:

- Voltage/current signal amplitudes.
- Signal delays.
- External compiled modules, such as internal real type magnitudes within a Verilog-A model.
- Periods and frequencies.
- Measures or computations involving the previous variables (mean, max or min values, mean errors etc.).

Using them we are able to define multiple *Error Metrics (EMs)*, each one referring to a situation that would lead to a system functional error (Stage 2 in Figure 5.3). During the circuit analyses, those *error metrics* are simultaneously evaluated to guarantee the system correct functioning. Regarding RRAM technologies, most RRAM SPICE compatible subcircuits are defined either using standard SPICE components [GRLVI14, ATB⁺15, GRGCY⁺16] or Verilog-A instances [CY15, YU15]. We propose a close interaction between the reliability/performance analyses and the low level circuit simulator, gaining access to both current and voltage signals as well as to Verilog-A magnitudes. Consequently, we can benefit from those custom *error metrics*, using, for instance, CF geometric changes, dopant areas drifts or changes in resistance state to define a functional error.

At the same time, and taking advantage of these analyzed magnitudes, we can define *Performance Metrics (PMs)*, that properly measure the circuit efficiency against different aspects: RRAM switching time, CF degradation ratio under stress, CMOS logic delays etc.

Figure 5.6: Efficient definition of the design space. The initial vast design space is progressively reduced using the results obtained after a series of iterative simulations.

Efficient Design Space Exploration

The third pillar sustaining the proposed strategy involves the definition of the design space to be explored, making the circuit characterization computationally achievable while guaranteeing the accuracy of the study. To that end, once *EM* and *PM* metrics have been defined, and the models are properly selected, an initial gross exploration of the design space takes place.

The process is described by Figure 5.6. There, rough but fast simulations combining uncertainty sources allow the designer to get an estimation on the design space areas worth of being analyzed in depth: the best set of design parameters, the worst case studies, etc. The purpose of this exploration is to constrain the design space to those scenarios of relevance for our design and environmental conditions, lighting the total number of transient simulations required to guarantee the correct circuit functioning. Thus, making use of the designer knowledge obtained from the previous gross analyses, the design space can be efficiently defined. As seen at the bottom of Figure 5.3, the circuit can be correctly characterized and evaluated with respect to *EM* and *PM* metrics, accomplishing global, fine grain multi-parameter analyses considering multiple and concurrent uncertainty sources.

5.4 Circuit-Level Simulation Framework

To support the proposed analyses strategy, a reliability and performance oriented framework has been developed. This circuit-level framework has been designed with two goals in mind:

1. Automation of the problem analysis, the generation and managing of the required simulations, and the processing and analysis of the computed reliability/performance results. If such automation is not present, to cover the full design space the user should manually create the pertinent circuit modifications —which may involve occurrences at each circuit node, then substitute the models by the enhanced ones and iteratively simulate each of the varying working conditions. Therefore, this automation makes affordable the required analyses, saving time and avoiding errors that could be manually created.
2. Improvement of performance. Efficiency is a major requirement given the huge number of simulations to be accomplished.

To put things in perspective, let's consider the example of a circuit hardened against radiation, that requires to be characterized. Conventional methods for parametric analysis of radiation usually relies on the SPICE circuit simulator for simple parameter sweeps, studying a low number of parameters in a sequential manner. For example, in recent works where binary search algorithms are applied to compute the critical charge value, the study is limited to only two parameters [RBW⁺15]. In this context, the authors reported that the computational load relative to the parameter sweep resolution was high enough to justify the use of a high end workstation. Even leaving out variability, simulation length is a restrictive factor.

But we must consider variability. When variability is taken into account, the Monte Carlo analyses required by process variations simulation combined with each dimension sweep computation lead to a huge amount of individual transients to be considered. To deal with this, the framework includes speedup features that allow a fast and intelligent simulation. Appendix D gathers the framework implementation details, describing, among others:

- Implemented algorithms.
- Parallelization and acceleration techniques.
- Design space division.
- Simulation Granularity.

The appendix also includes a 14-bit adder case of study, illustrating the framework efficiency when characterizing the circuit.

Figure 5.7: Functional graph of the reliability framework. System inputs include the netlist and configuration files. The core of the framework, delimited by the dotted line, is composed of several modules that handle the analysis and characterization of the given circuit.

5.4.1 Functional Structure

Relying on the SPICE level Cadence Spectre circuit simulator, the tool automates the simulation of different tests and performs post simulation analyses to extract the relevant information, related to the evaluation of the user defined *EM* and *PM* metrics. To that end, the framework manages every simulation and post-analysis, and after they are accomplished, it generates the reliability and performance report. In this report the framework identifies the most sensitive working conditions where the circuit is vulnerable to errors, and characterizes the performance regarding different working or design parameters.

Figure 5.7 shows the software functional structure. It has been developed using a modular scheme that helps the extension of the framework functionalities. Additionally, this modularity eases the integration of novel component models, the inclusion of new uncertainty sources, or the definition of different types of metrics. The description of the main functional structures is accompanied by a simple but clarifying example, that helps the reader to visualize the purpose of each functional stage. Figure 5.8 describes this exemplifying circuit, composed

Figure 5.8: Example circuit to be characterized.

Figure 5.9: Redefinition of RRAM instances in the example circuit.

of an RRAM crossbar, the accessing and reading circuitry, and an RRAM based register.

As seen in the functional graph in Figure 5.7, the system inputs are the circuit netlist, technology information and experiment configuration files (metrics definition, radiation environment, experiment variables etc.).

Following the functional diagram, the main box corresponds to the different modules the *Reliability and Performance Analyses Tool* is composed of. First, we find the *Circuit Parser and Circuit Analysis* modules, responsible for the identification of circuit components, establishing a hierarchical relationship between models, instances and their interconnection forming *subcircuits*.

Depending, first, on the models redefinition —specified in Stage 1 of the methodology— the radiation (if any) impact characteristics and impacted node, and second, on the design space exploration method, different *scenarios* are created. Each scenario considers the original circuit including the model modifications, concrete radiation impact emulation, and a particularized set of design, environmental and working parameters. *Circuit modifier* module is a key components in the process, as it is responsible for the alteration or redefinition of those models that require the emulation of additional characteristics, such as RRAM endurance,

Figure 5.10: Redefinition of CMOS instances in the example circuit.

Figure 5.11: Simulation of radiation impact in the example circuit.

aging, TID, or SET [BHS⁺14, RLV14, SBK15, GRGCY⁺16]. In this step, the framework introduces the variability —via CMOS PDKs, parameters influencing the enhanced RRAM-CMOS models, etc. Therefore, and using the pertinent *circuit modifiers*, the *scenario handling algorithm* systematically creates the different circuit scenarios. Figures 5.9 and 5.10 illustrate this process. In first place (Figure 5.9) RRAM cells and RRAM based components are substituted by instances of models properly redefined. These redefinitions take into account the desired uncertainty sources. Similarly, in second place, CMOS models are redefined (Figure 5.10).

In cases where the modified component is replicated within the circuit, the circuit modifier, making use of the hierarchical structures, propagates the model redefinition to dependent instances on upper levels of the circuit. Each model redefinition or radiation occurrence, together with the different parameter sets to be evaluated, are translated into a new modified netlist ready to be simulated and analyzed. At the end of this process, the framework will automatically have redefined the models of every required component instance. Continuing with the example, Figure 5.11 represents a particle impacting in the node i located in the upper column decoder. This radiation occurrence will be translated to an

Figure 5.12: Pool of scenarios to be simulated.

individual scenario that gathers every accomplished model redefinition together with the radiation emulating mechanism. The same process would be repeated until the radiation of every node has been accomplished, creating the pool of scenarios shown in Figure 5.12.

Following this approach, and once every uncertainty source has been considered, at the end of the procedure the tool has created a myriad of different netlists. These individual netlists, each one representing a different scenario with different error source characteristics, design parameters, or environmental conditions, are simulated. The *simulations handler*, interacting with the low-level circuit simulator, coordinates the scenarios concurrent computation. Once each scenario simulation has ended, the *reliability & performance results analyzer*, by applying the *EM* and *PM* metrics defined in Stage 2 of the methodology, extracts reliability and performance information from the transients and proceeds to process the results and generate statistical reports and plots.

Thanks to this framework, the designer automates the arduous, time consuming and error prone labor of generating, simulating and processing a large set of individual netlists. This is a major issue considering the multitude of emulated uncertainty sources, as well as the handling of individual combinations of parameter values particularizations. Furthermore, since the framework has been built around the multi-metric evaluation paradigm, efficient parallel algorithms directly provide the user with all reliability and performance reports, with no need of additional post-processing [GRLVAI16].

As a summary, and connecting the framework —available for download at [GRLVR15]— with the presented analyses strategy, the developed tools stand as an open solution for the pressing needs of RRAM-CMOS hybrid circuit designers. The combination of the proposed analyses together with the developed framework extends the circuit knowledge, improving traditional characterizations. We hereafter describe the main framework details, including circuit modifiers, simulation modes, metrics implementation and variability, corners and technologies handling.

Figure 5.13: Substitution scheme examples: a) Bias dependent radiation subcircuit scheme [KSA⁺09] and b) Switch controlled radiation scheme [MEV⁺08].

5.4.2 Circuit Modifiers: Model Enhancement Schemes and Radiation Schemes Implementation

Circuit modifiers are mechanisms implemented in the framework that allow the enhancement of the different models as well as the emulation of different uncertainty sources such as radiation occurrences. Consequently, they produce a modification on the original circuit giving birth to a new *modified circuit*. The presented methodology and analyses strategy (Section 5.3) uses injector sources and/or model substitutions with the aim of providing additional behaviors or characteristics to the evaluated components. The use of Verilog-AMS and C++—via *Compiled Model Interface* (CMI), *Verilog Procedural Interface* (VPI), *Programming Language Interface* (PLI) and *Netlist Compiled Functions* (NCF) instances and models—in the definition of the modifiers behavior guarantees the compatibility with previous approaches whereas it opens a wide range of possibilities such as TID/SET co-simulation or the study of complex aging effects on radiation hardening levels [KSA⁺09, BHS⁺14, SBK15].

The modular design of the framework eases future extensions in order to deal with future memristor models, new variability or degradation schemes, as well as more complex radiation and aging schemes, such as layout dependent injections or multiple transient injections. Moreover, we can take advantage of the netlist level enhancements presented in Section 3.3, reducing the computational load and improving the global simulation convergence.

Model Enhancement and Radiation Types

Our framework considers two powerful modification schemes which, following the reconfigurability provided, model almost any kind of required behavior—variability, degradation, TID, aging—and radiation SET phenomena. Depending

Figure 5.14: Example of hierarchy in a tree-like circuit structure.

on the technology and uncertainty source or characteristic to be emulated, the model enhancements and/or radiation occurrences can be modeled following two approaches:

- *Injection Scheme.* This first approach is mainly used to emulate single radiation occurrences or modifications altering sole components. The tool places a current/voltage source at each node susceptible of being affected. The current source injects any waveform the user defines in order to accurately emulate the desired phenomena. This approach has been used in the experimental Sections 5.5 5.6, 5.7 and 5.8 to emulate the different SET radiation impacts.
- *Substitution Scheme.* The tool replaces any instance susceptible of being modified with the instance of an already enhanced subcircuit. This scheme is extremely useful to emulate novel phenomena such as TID or degradation effects in RRAM devices, as seen in Sections 5.5 and 5.8.

In case the substitution scheme is used to emulate effects strongly linked to the particular characteristics of the simulated devices —such as the stressing voltage for aging, or complex radiation sources, or induced currents highly dependent on the technology— the accuracy of results for these emulating schemes usually rely on a previous calibration. Figure 5.13 shows two examples of this kind of sources: Bias dependent radiation subcircuit scheme[KSA⁺09] and switch controlled radiation scheme [MEY⁺08]. Many other complex models can be included using this approach such as [ME07], or the TID-aging emulation approach [SBK15].

Figure 5.15: Example of dependency in a tree-like circuit structure.

Model Enhancement and Radiation Algorithm

An efficient algorithm has been implemented in order to suitably find the circuitual components susceptible of being enhanced or radiated, and proceed to their modification. Instead of brute force search and replace, our approach makes use of the circuit hierarchy and component-dependency structures to evaluate, notify and propagate the different circuit modifications.

This algorithm is responsible for accurately analyzing the parsed circuit netlist, generating an assemble set of circuit instances structured as relational trees — Figure 5.14. Subsequently the inheritance analysis takes place, organizing the different circuit elements, along with the analysis and control statements and building relational tree structures. Finally, this module performs an instance dependency analysis which is required by the modifier algorithm. The dependency analysis allows, as shown in Figure 5.15, to directly access the primitives, models, subcircuits, etc. bypassing non-desired statements and providing a useful search engine.

Once a circuit modification takes place, the algorithm recursively searches among the circuit elements independently of the number of levels the circuit has. Making use of the relational data structures, it finds the altered nodes or modified instances, and propagates the modification occurrence to dependent instances on upper levels of the circuit.

Figure 5.16 illustrates the way the modification algorithm works. The target circuit contains two buffer subcircuit instances placed together: `i_buff_1` and `i_buff_2`. In turn, each buffer instance is composed of a chain of two inverters. In order to characterize the circuit hardening capabilities against radiation, we will emulate particle impacts in every circuit node, producing SETs that require to be analyzed. Therefore the selected radiation scheme is node injection. For

Figure 5.16: Example of how the modifier algorithm behaves.

simplicity, we will consider v_{in} , vdd and gnd as nodes not-susceptible of being radiated. The modification process over this circuit is as follows:

- The tool finds the first component or node susceptible of being modified: the output of the first inverter in the buffer `i_buff_1`. The modifiers create an altered buffer subcircuit in which the first inverter is radiated `i_alt_1_1`. Right after this the modifiers substitute the original instance, `i_buff_1`, with the radiated one, `i_alt_1_1` and export the radiated netlist.
- The modifiers propagate the subcircuit buffer alteration through the circuit: the buffer `i_buff_2` gets substituted by `i_alt_1_2` and a new radiated netlist is generated.
- No more buffer instances are found, and therefore the modifiers radiate the next susceptible node, the second inverter of the first buffer `i_buff_1`. A new altered buffer subcircuit is created, `alt_2`, and the instance `i_buff_1` is replaced by a radiated one `i_alt_2_1` generating another radiated netlist.
- The modifiers propagate the last subcircuit buffer alteration: the second instance buffer `i_buff_2` is substituted by a radiated instance `i_alt_2_2`, generating the last netlist.

At this point no more nodes are susceptible of being radiated, and four altered netlists emulating a SET occurrence at different places are ready to be simulated. This simple example shows how the framework makes use of hierarchical and inheritance structures to efficiently create the required scenarios. Therefore using this approach it avoids brute force search and replace routines after every iteration.

Figure 5.17: Basic Analysis Mode workflow. Blue labels refer to independent threads, while the blue box refers to the specific considered scenario. The extracted reliability and performance results are transferred to the simulation control for later report generation.

5.4.3 Simulation Modes

There are three analysis modes provided by the framework, *Basic Analysis*, *Multi-Dimensional Analysis* and *Variability Aware Multi-Dimensional Analysis*.

The *Basic Analysis* is based on the transient analysis accomplished by the low-level circuit simulator. This enhanced simulation mode finds in seconds the results of the circuit evaluation against the user defined metrics, being able to extract the critical values of the key parameters that mainly influence the device behavior. This mode, represented by the workflow displayed in Figure 5.17, allows the user to customize the accuracy and granularity of the simulation. The computed results include whether the user metrics are satisfied, the maximum and punctual errors, etc; identifying the most sensible circuit nodes and components. The two remaining simulation modes rely on this *Basic Analysis*.

To provide the reliability dependence on circuit or environmental parameters, the *Multi-Dimensional Analysis* accomplishes the space exploration considering the involved parameters as different dimensions to be analyzed. The analyzed dimensions can refer to circuit or design parameters such as transistor sizes, environmental parameters as temperature or the radiation energy or impact time, or control parameters like signal convergence boundaries. The characterization results are then treated and the relation with the circuit behavior extracted, providing a powerful analysis for circuits where variability is not a major issue. In this mode, several *Basic Analysis*, differing in the evaluated parameter set, are concurrently simulated. The automated analysis of the behavior regarding the studied parameters aids the user to find the best/worst parameter set cases for the circuit. Moreover, this simulation scheme is able to handle static design parame-

ters —such as RRAM geometry— together with dynamic parameters like RRAM degradation or the radiation strike time. Thus, this simulation mode correlates these parameters with the circuit behavior, allowing to find the time windows where the circuit module is more vulnerable, and how the generated pulse is propagated interfering with the adjacent modules. The statistical results obtained from this analysis assist the study of error masking between circuit submodules.

Each of the parameter set-ups analyzed by the *Multi-Dimensional Analysis* is susceptible of suffering from variability. For example, the common temperature and voltage sweep analysis that usually is computed for the radiation characterization of voltage references, is not enough given that analog circuits relying on differential pairs are extremely sensible to technology variability. The most complex simulation mode, *Variability Aware Multi-Dimensional Analysis* adds the capability of studying how PVT or radiation/environmental variations affect the circuit. As introduced before, variability stands as a great liability when designing reliable systems. For each dimension, parallel Monte Carlo simulations are carried out. Those variation-aware scenarios are compared to the nominal ones, providing data on how those statistical variations affect the circuit behavior. This simulation scheme computes the statistical relationships between affected nodes, devices and parameters and it also analyzes how parameter variations impact on the reliability under radiation conditions. Involving Monte Carlo techniques, the computational load associated with *Variability Aware Multi-Dimensional Analysis* is the highest and most time consuming of the three available simulation modes.

Different characterization goals require different simulation modes. To illustrate the simulation modes capabilities Sections 5.6, 5.7 and 5.8, will show three distinct cases of study.

5.4.4 Variability, Corners and Different Technologies Handling

Usually, when designing a circuit using a specific commercial technology a set of design rules are given by the technology PDK. These design rules specify, among other restrictions, how the different models are defined, their parameter mean values, statistical variations and fabrication/simulation boundaries (corners). On the contrary, when using immature compact models emulating emerging technologies behavior, the situation changes drastically. On the contrary, predictive models do not have the same testing degree, and therefore they may be trusted only under certain environmental and working conditions. Here, the mechanisms emulating variability, aging, or temperature dependence are of vital importance, and, as for the compact model presented in Chapter 4, modularity and accessibility is a matter that under no circumstances should be neglected.

The developed framework has been designed to be able to work with both standard PDKs and predictive models, bringing then the gap between solid CMOS and new RRAM modeling approaches. The simulation of standard *Multi-Dimensional Analysis*, and *Variability Aware Multi-Dimensional Analysis* relies on both approaches.

Corners and Different Technologies Handling

During the development of this part of the thesis, we had to deal with the design of digital and analog modules hardened against PVT variations and radiation. During their characterization stages, and as most industry circuit designers are required to do, we had to compare distinct designs varying the technology corners and even contrasting different technology PDKs themselves.

Especially for analog design flows, and probably for future RRAM design flows, it is particularly important to be able to measure circuit performance and to detect functional faults when exploring the whole design space, taking into account each technology corner. Section 5.7 will show an example of the framework capabilities regarding the automated analysis of the radiation properties using several technology corners as well as technology defined parameter variability.

Our circuit-level framework can be configured to use different technology corner files, automating these common tasks. This way, aside from the separate simulation of each corner, the framework provides the automated comparison of the analysis results between the diverse corners. Therefore, the framework is not restricted to the use of individual technologies. By adapting the system inputs to match the technology specific models and PDKs, the user defined metrics can be applied to each of them with comparison purposes. This is of extreme help for the circuit designer: instead of a study where hundreds of simulations require to be manually simulated, handled and post-processed [RBW⁺15], our framework automates every technology comparison [RGRLV15].

Variability

To emulate variability low-level circuit simulators extract from the technology PDKs the statistical functions that define how the different model parameters vary. On the contrary, predictive models either include mechanisms to vary parameters within themselves or to assign values created externally by the low-level circuit simulator. Our framework addresses both approaches, allowing RRAM-CMOS hybrid circuits to be analyzed under variability.

The technology PDKs and the predictive models information can be included as a whole within the experiment, but also the customization of each technology parameter variation can be specified. Additionally, user defined parameters (and

their variations) regarding design, simulation or radiation can be independently handled. Therefore the user is able to customize how parameters vary at different levels:

- PDKs. Either as a whole set of technology design rules, or with the independent selection of specific parameters.
- Technology intrinsic uncertainty sources, including parameter variations. These parameter values can be defined by analytical functions but also extracted from experimental measures as well as from specific device level simulators —Section 3.4.
- Environmental and radiation parameters, including variations: analytical functions, values read from tables. Using the provided methods —probability density functions, PWL functions with measured data— the user can specify how environmental and radiation variables such as temperature, particle impact rate, collected charge etc. are modeled within the system.
- Time or cycle dependent user defined design parameter variations: analytical functions, values read from tables. This approach allows to use variability and degradation schemes such as the one presented during the RRAM read circuitry characterization gathered in Section 4.3.
- Time or cycle dependent user defined simulation control parameters, including mathematical solvers and convergence control parameters. This case is especially interesting in experiments considering long transients of circuits involving tens of thousands of devices, reducing the simulation time and disk space.

The two last methods are especially interesting, since they provide variability and degradation schemes more in accordance with reality, improving the overall analyses accuracy. The use of one or another variability emulation scheme depends on the circuit under design, as well as on the characterization to be accomplished. For example, the validation case of study presented in Section 5.5 will show an example of how the framework is able to independently handle parameter variations specified in the technology design rules. There, the statistical variations of the temperature, P transistor and N transistor widths are isolated from the technology files, and the upset ratio dependence on these variations can be determined. Similarly, Sections 5.5, 5.6, 5.7 and 5.8 will emulate different variability types, highlighting the framework capabilities.

5.4.5 Metrics Implementation

As introduced in Section 5.3, the second key point in which the analyses strategy is based on is the simultaneous use of user defined metrics. The developed framework grants this multi-metric evaluation paradigm, and extending the standard analyses flexibility. Depending on its definition, using standard XML files, each error metric *EM* may trigger (a) functional error(s) in the circuit under study. In the same way, the analyses consider the evaluation of performance metrics *PMs*, thus measuring how well the circuit behaves regarding a specific aspect.

Traditionally, we have seen how circuit misbehaviors were considered as simple bit-flips. However, to give support to RRAM based circuits as well as analog modules, the computation of functional errors must change, in order to consider the new error definitions. To that end, our solution manages more sophisticated error definitions, widening the analyses range and extending the types of circuits to be analyzed.

The algorithms that check whether both user defined *EM* and *PM* metrics are met require low level interaction with the base circuit simulator. Depending on the metric type, the process involves transient read and evaluation and / or access to low-level circuit evaluations. The designer may take advantage of calculus expressions —peak values, delays etc. In this case, for the last version of the framework, we make use of Cadence Ocean expressions engine, performing evaluations over signals with no additional computational load penalty. In these cases the framework would retrieve their data, and next it would proceed to evaluate the corresponding metric.

In cases where the metric evaluation involves the study of transient signals —no matter its nature, being electrical magnitudes or real magnitudes from Verilog-A or any other pre-compiled module— the metric computation requires from the following steps:

1. First, the transient convergence is checked.
2. Second, the transient signal is prepared for later stages, going from the first time step to the last one. As the simulation time steps probably differ from the golden scenario to be compared to, an interpolation process takes place, standardizing the target signal to the golden one.
3. Third, while analyzing the signal evolution through time, the metric evaluation is computed. The algorithm checks:
 - Whether the signal has reached absolute thresholds —absolute values.
 - Whether the signal exceeds a threshold difference with the golden magnitude at the same time —relative error.

As we stated before, a key characteristic of the developed framework is its flexibility. With regard to *EM* metrics, these two checks may refer to two different procedures:

- The performed check is accomplished punctually, so the metric evaluation is independent of the simulated time —punctual errors. Therefore, when using this scheme should any error be detected the metric raises an alert, immediately reporting the simulations handler about the error.
- The performed check is dynamically evaluated, considering the specific time step within a time window. Thus, the metric evaluation result depends on a *customizable time window* during which the circuit may recover from a possible functional error. This evaluation mode is especially useful for the evaluation of transient signals with high noise levels and glitches — maintained errors. Compared to the previous scheme, when using this maintained errors detection the metric only reports the error if the circuit does not recover within the specified time window. On the contrary, if inside this window the circuit recovers from the error, the metric logs the event but does not inform about it.

To summarize the framework functioning, Figure 5.18 synthesizes the flow describing the simulation process, post-analysis and report of results, compared against the traditional analysis flow. Synthesized, the four main differences are:

- Single versus multiple scenario analyses —comparison between aged modules, comparison between individually radiated nodes, etc.
- Single goal versus multi-metric approach.
- Individual parameter sweeps versus multiple parameter sweeps, results coupled by pairs, creating hierarchically organized reports.
- Timing aware reports: errors/performance by time window, etc.

As an ending note we would like to remark that given the variety of problems when evaluating not only RRAM-CMOS hybrid circuits but every emerging technology, the framework stands as a helpful solution to fully characterize the circuit. Its flexibility allows the study of a vast range of metrics, using both Ocean evaluations and signal transients through the customizable metric parameters: absolute thresholds, relative margins and time windows. Therefore, the use of the presented metrics makes the framework more versatile and able to adapt to future circuits and technologies.

Figure 5.18: Comparison between traditional single-goal oriented analysis flow and the proposed multimetric, multiscenario analysis flow.

5.5 Validation Against Experimental Data

The main purpose of this section is to validate the results obtained by our framework comparing them to results of experimental works. Experimental results [TBAS14], with silicon data from SRAM cells radiated with alpha particles, are reproduced using a similar CMOS technology sharing the same 65nm node size. Static analyses varying pulse characteristics and transistors dimensions are carried out in the next sections, the validation is completed with the resulting soft error rate calculation, matching the figures from the referenced work.

The SRAM cell is then studied under dynamic operation, achieving the same qualitative conclusions than similar works, carried out with a different technology. Finally the capabilities of our framework are shown performing a combined analysis of radiation and process variations of a cell during dynamic operation, comparisons with other works show that we are able to reproduce equivalent results with our tool.

All the experiments within this section, as well as the simulations included in the following sections have been simulated using four virtual i7 cores equipped with 16GB RAM.

5.5.1 Static Analysis Validation

A complete static analysis of 65nm SRAM cells differing in the transistor dimensions was presented in [TBAS14]. There, the different cells were first simulated and then fabricated, radiated with alpha particles and finally measured to study how SRAM transistor sizes affect their radiation hardening capabilities. We validate our framework by contrasting our results against the experimental data presented in that work, kindly transferred by Sebastià A. Bota and Jaume Segura from UIB University.

Induced Pulse Characteristics Analysis

Alpha-particle impacts produce current pulses of around 50ps length in this technology [TBAS14]. We setup the framework to model those pulses using the well known double exponential current sources [DWB⁺07, FCMG13]. Figure 5.19 shows simulated results, displaying the critical charge (Q_{crit}) linear dependency on the collection time constant of the junction (τ_2 in the double exponential definition). This linear dependency is in accordance with the experimental results shown in [TBAS14].

Transistor Size Analysis

The second part of the validation consists in a two-dimensional analysis of the critical charge dependency on the nominal dimensions of the two kinds of transistors (nMOS and pMOS) that build the SRAM cell. For this purpose we kept $\tau_1 = 1$ ps and $\tau_2 = 10$ ps as the time constants of the double exponential, which results in a pulse length of around 50ps. Figure 5.20 shows the linear dependency of Q_{crit} with widths of the nMOS and pMOS transistors respectively, while the width of the other transistor type is given different constant values of 15 μm 23 μm and 30 μm . In addition, in each figure both positive and negative induced charges are considered.

These dimensional analyses reproduce the experimental results extracted from the irradiations carried out in [TBAS14]. The simulated Q_{crit} and its dependence on the transistors size are in accordance with the experimental measures. Both parameters affecting the induced radiation as the time constant or circuit dependent parameters such as the transistor widths have been analyzed, the results obtained

Figure 5.19: Critical charge dependence on τ_2 parameter for both positive and negative charge injections. The injected pulse width can be approximated to $5\tau_2$.

by our tool match the ones reported in the cited work, which validates this kind of analysis.

SER

As a last verification step for this type of analysis, the critical charge values obtained by our framework are used to calculate the soft error rate of the cells.

We compare our results against [TBAS14], using the same equation, radiation flux (κ), and electron and holes charge collection efficiency (η_e and η_h):

$$SER = \kappa \cdot (A_{diff,n} \cdot e^{\frac{-Q_{crit,n}}{\eta_e}} + A_{diff,p} \cdot e^{\frac{-Q_{crit,p}}{\eta_p}}) \quad (5.1)$$

$$\kappa = 8.03 \cdot 10^7 \alpha - particles \cdot m^{-2} \cdot s^{-1} \quad (5.2)$$

$$\eta_e = 2.02 fC \quad (5.3)$$

$$\eta_h = 0.79 fC \quad (5.4)$$

where $A_{diff,n}$ and $A_{diff,p}$ denote the n and p sensitive diffusion areas of the internal nodes of our cells, that may present small differences with the cells of the

(a) Q_{crit} dependence on N-transistors size, w_n . (b) Q_{crit} dependence on P-transistors size, w_p .**Figure 5.20:** SRAM radiation analysis: Simulated Q_{crit} and its dependence on transistor size and injected charge polarity.

Figure 5.21: SER dependence on SRAM cell dimensions: simulations vs experimental results [TBAS14]. The cell types S, M and L stand for Small (0.15 μm), Medium (0.23 μm) and Large (0.3 μm) transistor widths, where the two letters refer to N-transistor and P-transistor respectively.

reference work. $Q_{crit,n}$ and $Q_{crit,p}$ are the negative and positive critical charges obtained as the result of simulations.

Figure 5.21 presents the comparison between the experimental results measured in [TBAS14] and the simulated results obtained using our framework. As can be seen the simulated SER properly fits the experimental data, only small differences persist, that can be explained by the small differences between the simulated cells and the ones from the reference work.

Small differences can also be observed between the simulation data and the radiated results in the reference work. As our results are in the range of these small differences, we can consider that this experiment validates our framework as far as SER calculation is considered, matching both simulation and experimental data from [TBAS14]. In this experimental study, the circuit was built using a 65 nm technology similar to the PDK used in our simulations. It can be noticed how the simulated results go in accordance with the experimental ones. Even small differences can be noticed, the SER values correctly match the different cell types, and, moreover, these short differences can be described as a positive bias almost constant between cell types.

Figure 5.22: Normalized read (RSNM) and write (WSNM) static noise margins for different combinations of N transistor (w_n) and P transistor (w_p) widths.

SRAM Cell Selection

The framework can also be configured to extract user-defined metrics not necessarily related to radiation, and perform the same dimensional analyses of their dependence on circuit parameters.

This is the case of the read, write, and hold static noise margins (SNM) [SLL87] that tell us the robustness of SRAM cells during the different operations.

We have setup the software to extract these parameters for 6T SRAM cells, for different width combinations of the transistors.

Assuming that pass-gate and pull-down transistors share the same diffusion and thus have the same width, Figure 5.22 shows that the optimal trade-off for nominal read and write static noise margins is obtained for similar values of w_p and w_n widths. We keep widths values of $w_n = w_p = 0.2 \mu m$ for the cell to be analyzed under dynamic operation, these widths are in consonance with those used in [TBAS14].

5.5.2 Dynamic Analysis Validation

In addition to circuit parameters, such as the width of the transistors, the framework allows to analyze a circuit during dynamic operation. We setup the tool to find out the dependence of the critical charge with the particle impact time. The goal is then to compare the results obtained with our framework against those presented in [GZD⁺11]. For this, we simulate successive write and read cycles that cover all the situations, analyzing which are the most vulnerable instants during SRAM operation.

Figure 5.23: Q_{crit} characterization of the SRAM nominal scenario during different operations. The set of cell operations shown are: H for hold, W for write and R for read. Zoomed areas highlight the disturbance introduced by the activated pass-gates.

Figure 5.23 shows the critical charge (Q_{crit}) obtained at different operations for the nominal scenario. It can be seen that during write operations, the critical charge is heavily increased, above the maximum value considered, as the contents of the cell are forced externally. The zoomed part of the figure shows how during a read operation the cell radiation hardening is first worsened due to disturbance introduced by the activated pass-gates, this happens even for impacts taking place before the actual read has begun. As the read operation carries on, the radiation hardening capabilities of the cells increase again.

Despite obtaining different absolute numbers, since a different technology is used, the results obtained during this simulation, with the different critical charge values during write and hold, and its transient evolution during read operation, are totally in accordance with those shown in [GZD⁺11]. Although the reference work only presents simulation data, the good matching of the transient evolution of the critical charge can be regarded as a prove of validity, especially considering that we are dealing with internal signals would make it difficult, if not impossible, to validate the results with such a high level of detail.

Figure 5.24: Worst case upset ratio (radiation error for the pair collected charge/impact time) taking into account variability, for negative charged impacts. The 2D projection over Qcoll/tImpact plane compares the nominal scenario where no variability is considered against the variations-aware simulations.

Figure 5.25: Worst case upset ratio (radiation error for the pair collected charge/impact time) taking into account variability, for positive charged impacts. The 2D projection over Qcoll/tImpact plane compares the nominal scenario where no variability is considered against the variations-aware simulations.

5.5.3 Variability Analysis Validation

Although SRAM cells are very simple circuits that do not require from hierarchical approaches to simulate their radiation hardening, the high number of simulations required to analyze them dynamically takes advantage of the automation brought by our framework. As seen in the previous section, the evolution of the critical charge with the different read and write cycles was automatically generated and could be reproduced for different cases.

When variability is considered, the number of simulations ramps up. Therefore, the automation, together with speed-up improvements capabilities detailed in Appendix D, becomes particularly necessary. 80,000 Monte Carlo simulations were carried out taking into account variations in transistor widths, operating temperature and supply voltage. The short length of the induced pulses compared to the 17 ns transient length forced the simulator to limit its maximum simulation time step, leading to an enormous computational overload. The implemented parallel computation allowed the whole set of simulations and their later analysis to be carried out in 2 hours 53 minutes.

The results of those Monte Carlo simulations considering radiation effects under variability conditions are condensed in Figures 5.24 and 5.25, provided by the framework. They display the upset ratio dependence on the collected charge value, Q_{coll} , while varying the particle impact time. There, it can be noticed how different impact times lead to different computed functional error rates. The figures represent the worst case reliability results for both positive and negative induced charges. As shown, the variability in transistor width and temperature parameters produces undesired effects on the critical charge calculus. For comparison, the nominal critical charge dependence on the impact time — Figure 5.23— is reproduced in blue in both 2D graphs. This way, the border showing charges leading to a failure has become a progressive transition along a range of charges that switch from a 0% to a 100% failure probability due to mismatch. As a result the time window boundaries at which the cell is vulnerable are widened, and the induced charges that could lead to a bit upset are reduced in about 1 fC.

This example illustrates how parameter variations cannot be ignored when evaluating radiation hardening capabilities. Again, these results show that our framework is able to reproduce those results presented in [GZD⁺11], combining PVT with transient simulations, in a fully automated way.

Figure 5.26: *DICE flip-flop schematic presented in [SBB⁺09].*

5.6 Digital Module Characterization

This section contains a guided example which shows the proposed tool capabilities when analyzing digital modules composed of a hierarchy of components, in which some of the framework characteristics such as granularity control can be applied. The purpose of the experiment is then the characterization and comparison between two circuits accomplishing the same function—a flip-flop—differing in their radiation hardening characteristics. Relying on the framework, we will determine the radiation hardness of both of them. The first circuit is a twelve-transistor standard flip-flop [Wes10], built on commercial 40 nm technology. The second circuit is a hardened flip-flop with a *dual interlocked cell (DICE)* structure [SBB⁺09]. The hardening principle is described in Figure 5.26, where the latch is improved by using the dual interlocked feedback circuitry underneath. Clock edges are always subject to error when dealing with timed modules. The DICE flip-flop should be more robust in non-clock transition instants, and we make use of our tool to measure that improvement.

The first step is the experiment configuration:

- Technology: Commercial 40 nm.
- Definition of the module timing: 10 ns clock period.
- Particle impact modeling: We will use the double exponential injection scheme [DWB⁺07, FCMG13] in which the particle impact generates a cur-

Figure 5.27: Radiation impacts modeling in the analyzed registers.

rent at the impacted node with the form

$$i_{inj} = I_0 * (e^{-(t-t_i)/\tau_2} - e^{-(t-t_i)/\tau_1}) \quad (5.5)$$

$$I_0 = Q_{coll}/(\tau_2 - \tau_1). \quad (5.6)$$

Parameters τ_1 and τ_2 depend on the technology. During our experiment we will consider the values $\tau_1 = 1E - 11$ s and $\tau_2 = 2E - 10$ s. Q_{coll} refers to the collected charge at the impacted node. The last parameter, t_i , is the particle impact instant.

- Maximum and minimum charge range: In our experiment, this parameter will vary between $1E - 14$ C and $1E - 11$ C.
- Definition of the particle impact instants t_i : between 2.5 ns and 17.5 ns avoiding the situations when both clock and input have transitions, as shown in Figure 5.27.
- Observable signals: The chosen observable signal is the output of each flip-flop.
- Definition of our circuit *radiation error*. We will consider a radiation error as the occurrence of a non-desired bit flip due a particle impact. This bit change should at least last 1 ns to be considered a Single Event Upset (SEU).
- Definition of injectable circuit components and nodes. Each component is susceptible of being radiated. Each node but the output, our observable signal, is susceptible of being impacted by a particle.

With the above configuration we proceed to simulate the experiment. After the radiation analysis, the results extracted from the report confirm the radiation robustness of the DICE scheme facing the standard flip-flop shown in [SBB⁺09].

Table 5.1: Standard Register. Critical charge at each node varying the particle time of impact. ★: Not affected by the maximum charge ($1E - 11$ C). ABMC: Affected by the minimum charge ($1E - 14$ C)

t_i (ns)	Node 0	Node 1	Node 2	Node 3	Node 4	Node 5
2.5	3.05E-12	★	★	★	3.90E-13	9.69E-12
5	ABMC	★	★	★	2.04E-14	1.51E-13
7.5	ABMC	★	★	★	1.67E-14	★
12.5	★	ABMC	★	3.8E-14	1.38E-12	1.43E-14
15	★	ABMC	★	4.54E-14	8.44E-12	★
17.5	★	★	★	★	9.06E-12	★

t_i (ns)	Node 6	Node 7	Node 8	Node 9	Node 10	Node 11
2.5	★	1.91E-14	★	1.55E-14	1.55E-14	★
5	★	★	★	ABMC	1.55E-14	★
7.5	★	★	★	ABMC	3.56E-14	★
12.5	1.30E-12	6.12E-14	★	★	★	ABMC
15	1.55E-14	★	★	★	★	ABMC
17.5	3.56E-14	★	★	★	★	★

Tables 5.1 and 5.2 summarize the results ordered by the internal nodes of the circuit modules. From their thorough analysis we extract the data summary shown in Table 5.3 which supports the DICE radiation endurance regarding the standard register: DICE structures, even having more nodes perceptible of being radiated, present a better radiation hardness behavior.

Moreover, we also deduce the worst impact time t_i at which our module is more sensitive. As expected, clock rising edges are the most sensitive instants (5 ns, 15 ns).

5.6.1 Sequential Circuits Granularity Considerations

Let us consider a new scenario with a chain of five DICE registers which maintains the same timing as the example before. In case we want to extrapolate the radiation characteristics of the single DICE analysis, we should not consider the worst cases of the isolated register alone but as an ensemble of timed modules. Otherwise, two different consequences can appear:

1. If the SEU at the output of the first register does not meet the rising edge of the clock signal, it is not captured by the second register, shown in Figure 5.28.

Table 5.2: DICE Register. Critical charge at each node varying the particle time of impact. ★: Not affected by the maximum charge ($1E - 11C$). ABMC: Affected by the minimum charge ($1E - 14C$)

t_i (ns)	Node 0	Node 1	Node 2	Node 3	Node 4	Node 5
2.5	★	★	★	★	★	★
5	★	★	★	★	★	★
7.5	★	★	★	★	★	★
12.5	1.03E-13	★	★	★	★	1.91E-14
15	2.27E-12	★	★	★	★	2.04E-14
17.5	2.27E-12	★	★	★	★	★

t_i (ns)	Node 6	Node 7	Node 8	Node 9	Node 10	Node 11
2.5	★	★	★	★	★	★
5	★	★	★	★	★	★
7.5	★	★	★	★	★	★
12.5	★	★	★	1.3E-12	★	★
15	★	★	★	1.22E-12	★	★
17.5	★	★	★	1.3E-12	★	★

t_i (ns)	Node 12	Node 13	Node 14	Node 15	Node 16	Node 17
2.5	★	4.98E-13	6.15E-13	★	★	★
5	★	4.54E-14	2.4E-14	★	★	★
7.5	★	1.53E-12	★	★	★	★
12.5	★	★	★	★	★	★
15	★	★	★	ABMC	★	ABMC
17.5	★	★	★	★	★	1.03E-13

2. If the SET inside the first register does not produce a SEU (it is not long enough), but it matches the rising edge of the clock signal, a SEU might occur in the second register, as shown in Figure 5.29.

Even if the probability of impact in a specific node at the worst time could be really small, regarding sequential circuits we should consider the critical amount of collected charge as a time dependent function $Q_{crit}(t)$. For a simple chain of two registers, being $Q_{reg}(t)$ each register critical charge, and d_{reg} the delay between its input and output, we would consider Q_{pair} , the critical charge of the module, as

$$Q_{pair}(t) = f_{min}(Q_{reg}(t), Q_{reg}(t - d_{reg})). \quad (5.7)$$

Table 5.3: Q_{crit} at each node, number of sensitive nodes and ratio between sensitive nodes and nodes susceptible of being radiated for each register scheme. ★: Not affected by the maximum charge ($1E - 11C$). ABMC: Affected by the minimum charge ($1E - 14C$). # SN: Number of sensitive nodes.

t_i (ns)	Standard			DICE		
	# SN	Ratio	Min Q_c	# SN	Ratio	Min Q_c
2.5	6	0.5	1.55E-14	2	0.11	4.98E-13
5	5	0.42	ABMC	2	0.11	2.4E-14
7.5	4	0.33	ABMC	1	0.05	1.53E-12
12.5	7	0.85	ABMC	3	0.17	1.91E-14
15	5	0.42	ABMC	5	0.28	ABMC
17.5	2	0.17	3.56E-14	3	0.17	1.03E-13

Figure 5.28: The SEU generated in the first register will not be propagated to the second one. The SEU length is not long enough to be captured in the next clock edge (20ns).

However, the complexity of equation (5.7) grows depending on the number of register stages, complicating the calculation of the amount of charge that would destabilize our circuit behavior. Once again our radiation tool proves useful systematically solving the problem: The design space exploration of the five stage chain, involving fifty impacted nodes analyzed at six different instants was swiftly computed. It took over five minutes of a virtualized bi-core Intel i7 with 8 GB RAM

Figure 5.29: The SEU generated in the first register before 20ns has been propagated till the output of our registers chain. Should any later module read the chain output, the error would leave the studied module being stored in the system.

Table 5.4: DICE Chain Summary. ★: Not affected by the maximum charge ($1E - 11$ C). ABMC: Affected by the minimum charge ($1E - 14$ C).

t_i (ns)	# Sensitive nodes	Min Q_c
2.5	0	★
5	11	$2.16E-14$
7.5	7	$2.03E-14$
12.5	2	$1.91E-14$
15	10	ABMC
17.5	5	$4.78E-14$

to generate the radiation results shown in Table 5.4. As was predictable, some SEUs were corrected by posterior stages (2.5 ns impact time). In other cases some SET propagations ended up becoming a SEU which previously was not present.

5.7 Analog Module Characterization

The next case of study focuses on the characterization of an analog circuit working under sources of errors prior to its fabrication. This analog circuit is in turn composed of other analog submodules: a power sensor monitoring the stability

Figure 5.30: Simplified schematic of the analyzed bandgap voltage reference circuit.

of a bandgap voltage reference. This example allows us to illustrate how the reliability characteristics of a circuit submodule influence its periphery behavior.

Following that approach we will first analyze a bandgap voltage reference (Section 5.7.1) that is later used to build a power monitor (Section 5.7.2). Both circuits were designed by Hernán Aparicio [AILV15].

5.7.1 Bandgap Voltage Reference

This circuit is a widely employed voltage reference due to its low dependence on temperature and process parameters [Raz02]. The principle behind it consists on balancing the negative temperature coefficient -CTAT (Complementary To Absolute Temperature)- of a PN junction current with the positive temperature coefficient -PTAT (Proportional To Absolute Temperature)- of the thermal voltage, V_T . A reference voltage is generated by adding two voltages that have temperature coefficients opposite to each other with suitable multiplication constants. The resulting voltage is ideally independent on the temperature [Raz02].

Our module is represented by the schematic shown in Figure 5.30, and has been designed and sent to fabrication using a 40 nm commercial technology. At the time of writing the present document, the chips have been received from fabrication, so all the presented results correspond to circuit simulations. As can be seen, its full custom design allows a compact area of $51.4 \mu\text{m} \times 30.9 \mu\text{m}$. This circuit module produces a constant voltage non-dependent on temperature or voltage supply variations. This particular circuit, by tuning the internal resistors

R_{eq1} , R_{eq2} , R_{eq3} and the capacitor C_{eq1} , is designed to minimize the voltage variations produced by temperature changes in this specific technology.

This module, though simple, is enough to highlight the capabilities of the proposed analyses strategy-framework when applied to analog circuits. In the presence of radiation, the differential pair composed by Q1-Q2 gets destabilized, altering the output voltage reference VBG. Additionally, different technology corners produce distinct responses, a side effect that requires to be addressed. Consequently, the main purpose is then to study how both neutron-induced radiation and variability unbalance the voltage given by the bandgap at different temperatures.

Bandgap Reliability Analyses

To be used as the *golden* behavior, we simulated under nominal conditions the non-altered circuit module using the *typical-typical* technology corner, whose transients define the *golden* circuit functioning.

As introduced before, our analyses strategy includes three stages: the models redefinition stage, the metrics definition stage, and the efficient definition of the design space. In this example, as we are only using solid tested commercial models no model redefinition is required. Thus, in Stage 2, we defined the metrics to be used during the reliability analyses. The bandgap has been designed to supply a stable voltage to a library of components designed in the same 40 nm technology. Consequently, two different metrics are specified based on how the bandgap output VBG may cause misbehavior in other library components:

- A voltage deviation of more than 3 mV, not recovering in a time window of 0.5 ns is considered as a failure.
- A punctual voltage deviation of more than 80 mV is also considered a functional failure.

In the following experiments, both metrics are independently applied no matter the considered uncertainty sources. For example, if a particle strikes producing an overvoltage of more than 80 mV over the *golden* response, even if the output gets stabilized after a short time, metric 2 raises a failure. However, if a specific technology corner working at a certain temperature produces an output deviation in the range of 3 – 80 mV lasting less than 0.5 ns, no threshold error metric is violated, and no functioning failure is considered.

These metrics are fed to the reliability framework together with radiation configuration files. The analyses specified by Stage 3 in our methodology are automatically carried out. Although only the circuit variation mode results are

Figure 5.31: Bandgap output voltage deviation for each technology corner at different temperatures. Technology corners: typical-typical (tt), fast-fast (ff), slow-slow (ss), fast-slow (fs) and slow-fast.

presented, each individual report regarding all injected-nodes, studied parameters and individually analyzed magnitudes are available to the user.

Figure 5.31 shows how the framework analyzes the circuit dependence on different technology corners and temperature. As seen in the graph, within the considered temperature range of 1°C to 100°C , the maximum output voltage variation for each individual corner is below 2 mV .

The rest of Stage 3 analyses characterize the hardening levels against heavy ion radiation. In this case of study we consider neutron impacts as the radiation source to be emulated. Radiation experiments consider the critical charge Q_{crit} as the charge quantity (Q_{coll}) that collected by the circuit produces a functional error. Traditionally this functional error is given by a SEU, when a bit flip occurs. However, our improved simulator allows sophisticated metrics to declare whether an error has taken place. Therefore, and using the previously defined metrics, a maintained voltage difference at the bandgap output of 3 mV or more, or a punctual voltage difference of more than 80 mV would trigger a failure.

The collected charge characteristics for this technology size under neutron induced radiation are obtained from [NUT⁺12]. Figure 5.32 displays the automated Q_{crit} study considering the different corners. As expected, and given the temperature aware design, the critical charge does not depend on this parameter but on the analyzed corner.

Next, we analyzed how mismatch, together with process variations, destabilize the circuit behavior. Worst case variability conditions, extracted from the technology files were applied. Figure 5.33 shows how variability increases in one

Figure 5.32: Computed critical charge for positive and negative charge injections. The different corners response is simulated in a wide temperature range.

Figure 5.33: Variability results of an automated Monte Carlo analysis varying all considered technology/design parameters. The statistical variations are extracted from the technology PDK. The upsets ratio is presented together with the error statistics: mean, maximum, minimum and quartiles.

(mean) and even two orders of magnitude (worst cases) the output error. As can be seen, the results records are automatically generated including quartile and upsets statistics studied for each parameter. It should be noticed that the error threshold for this kind of analysis is automatically determined using the third quartile error value for each parameter at the individually analyzed magnitudes (80 mV in Figure 5.33 for the pair *temperature-output voltage*). Moreover, this

functionality allows considering several magnitudes for sources of errors. The simulation and individual/statistical analysis of 50,000 simulations was fulfilled in 13 s.

Usually, only temperature / supply-voltage dependence [KG14], together with the technology corners are considered in the analyses when characterizing bandgap modules. However, when combined with radiation, further analyses are required. Using the error threshold obtained from the previous analysis, the framework computes the radiation hardening characteristics under variability conditions and their dependence on the studied design/technology parameters. Figure 5.34 shows the results obtained and their comparison against the golden scenario.

The variability-radiation characterization is completed with the upsets ratio computation, shown in Figure 5.35. The framework capabilities are highlighted comparing these results with those shown in Figure 5.32, the conventional radiation analysis. In this late analysis more than 1,800,000 individual simulations were computed. The total time of the simulation, individual/statistical analysis was approximately 12 h.

As expected, given the bandgap intrinsic topology [Raz02], we can see in Figure 5.34 how performance is almost equivalent at low and high temperatures. The use of middle-size transistors gives additional protection against radiation. We can see how the mean value of the maximum achieved error at the output raises from near 60 mV to 80 mV when the collected charge Q_{coll} varies from 1 fC to 50 fC. However, when considering the upsets ratio, this study reveals how for Q_{coll} values higher than 10 fC the upsets ratio considerably increases; and even that the output voltage error does not noticeably rise, the system trustworthiness may be compromised.

(a) Negative injection.

(b) Positive injection.

Figure 5.34: Mean output voltage error dependence on temperature and collected charge. The Monte Carlo analyses involved 25,000 simulations per injection—positive and negative—and used the statistical variations given by the technology PDK.

(a) Negative injection.

(b) Positive injection.

Figure 5.35: Mean upset ratio dependence on temperature and collected charge.

Figure 5.36: Schematic of the power supply monitor together with the bandgap voltage reference.

5.7.2 Power Supply Monitor

After characterizing the bandgap module, we proceeded to analyze the reliability characteristics of the whole circuit, with the output of the bandgap voltage reference monitored by a power supply monitor. The selected circuit, a power supply monitor has been recently presented as a reference-free, low power solution, working in a wide temperature range [AILV15].

Figure 5.36 describes the circuit, which, leaving out the bandgap, is composed of 32 MOS devices together with auxiliary resistors and capacitors. The feedback loop included in the monitor, a scheme extremely sensitive to voltage variations, makes this circuit ideal to display how variations, temperature and radiation may cause its malfunctioning. Basically, this circuit audits upper and lower variations in the VDD power supply line, sensing possible deviations and activating the alarm signals overshoot and undershoot, respectively (Figure 5.37). In [AILV15] a conventional design methodology was followed, accomplishing temperature sweeps together with corner analysis. In this case of study we extend the previous work investigating how the combined effects of variability, temperature and radiation may cause malfunctions.

Figure 5.37: Power supply monitor nominal behavior.

Figure 5.38: Monte Carlo analysis together with temperature dependence using user defined metrics.

First, we study the golden simulation regarding the nominal behavior of the monitor, whose functioning is represented by the transients of Figure 5.37. As can be seen, the slight upper and lower variations in the power line trigger the sensing signals overshoot and undershoot.

Following the second pillar of our methodology, in the second stage we define the metrics to evaluate the performance of the circuit. In this example, we will define a circuit failure as the scenario cases where the sensing signals overshoot and/or undershoot suffer from: (1) A bit flip compared against the nominal scenario, or (2), a delay of more than 1 ns with respect to the nominal simulation. If either (1) or (2) failures affect overshoot and/or undershoot, we consider a behavioral upset.

Using these metrics, the analyses belonging to Stage 3 take place. For simplicity, in this case of study we will consider only two configurable parameters: *temperature* and *particle impact instant* or *tImpact*. Figure 5.38 shows the reliability results provided by the framework regarding how process variations affect the circuit in a wide temperature range. The graph shows the computed upsets (up

Figure 5.39: Power supply monitor: computation of Q_{crit} dependence on corners, considering both positive and negative charged impacts, and showing mean and median computations.

to 1% on the higher temperatures) and the affected signals. In this case only the overshoot signal was affected enough to lead to a functional error.

Regarding the study of the combined effects of radiation, corners and parameter sweeps, the collected charge characteristics for this technology size under neutron induced radiation are obtained from [NUT⁺12]. Figure 5.39 presents the statistical values of critical charge (Q_{crit}) for each corner/temperature pair. Q_{crit} can be described as the minimum charge (Q_{coll}) that collected by the circuit produces a functional error. Three patterns were extracted depending on the ion strike instant t_{Impact} , related to an over-voltage A, non-deviation B, and under-voltage C situation (Figure 5.37).

Using these previous results, we can reduce the computational load of the global analysis by restricting the explored design space –Stage 3 in the method-

Figure 5.40: Power supply monitor circuit global upset rate under real conditions.

ology. As can be extracted from the mean values in Figure 5.39, for each corner, most nodes are not affected by the maximum Q_{crit} value. For the most sensitive nodes, the temperature dependence is obtained using the median values function shape. Therefore we will reduce the temperature resolution at the middle of the temperature range where the circuit is more reliable. Regarding t_{Impact} reliability dependence we will use three different ion strike instants (A, B and C).

At the end of this analysis, the developed tool automatically generates reports regarding reliability dependence on radiation, variability and temperature/ion strike instant. The combined effects results can be globally studied on the circuit or, by contrary, access them by node/submodule. The same applies to user defined metrics, where both global upsets and individual metric reports are generated. This allows designers to identify and correct the possible problems. Figure 5.40 shows the circuit global upset, using the mean of the three ion strike instants (A, B and C) and considering the impacts at every node. We can extract that for Q_{crit} values near $10 fC$ the power supply monitor is going to start to fail with a 30% probability. This ratio grows at lower temperatures. For high temperature we find a failure rate of 1% approximately. As an example of node report, the graphs in Figure 5.41 display representative patterns in the error metric at two distinct nodes.

Figure 5.41: Power supply monitor circuit metric error for specific nodes.

It can be noticed how the critical charge 10 fC , value that triggered the bandgap error at the output —Figure 5.35— is key and establishes a threshold in the feedback circuitry of the whole power supply monitor circuit.

To conclude, after we have defined the experiment configuration files, and compared against other reliability studies [KBK⁺11, QBRB13, FHS15, AILV15], the proposed tool automatically analyzes the combined effects of multiple error sources, thoroughly describing how they affect the different circuit modules. Together with those reliability analysis improvements, the framework has automated the circuit hierarchy analysis, as well as the unreliability source injection or parameter investigation, systematizing what otherwise would require from the manual generation, simulation, and post analysis of each independent scenario.

Figure 5.42: Schematic of the studied RRAM-CMOS hybrid circuit. The resistive read scheme is based on a sense amplifier, a load resistor $-R_{LOAD}$ - and auxiliary decoupling capacitors $-C_{IN}$, C_{OUT} -.

The simulation and post processing were computed in an *i7* machine fitted with 6 virtualized cores. The time required for completing the whole analyses was 56 *h*, including the analyses that did not consider variability (approximately 2 *h*), and the most time consuming analysis, considering every uncertainty source, where 1,458,000 transients were automatically generated, altered, analyzed, and the reliability patterns extracted in 54 *h*. The efficient definition of the design space using the first approximation analyses was of extreme importance, allowing to reallocate simulation resources from the most reliable impact time and parameter sets to the most unreliable, increasing the study robustness.

5.8 RRAM Read Circuit Characterization

The previous cases of study have shown how the developed solution fully supports and extends the traditional analysis of digital circuits, and how analog circuits may benefit from the multi-metric analyses paradigm, especially when considering simultaneous uncertainty sources.

The last case of study has been designed to illustrate the full capabilities of the presented strategy and framework. To that end, the present section describes the characterization of an RRAM read circuit. The RRAM-CMOS hybrid circuit, whose schematic is described in Figure 5.42, translates the selected RRAM cell's state to a binary voltage value.

Figure 5.43: Transient plot representing standard HRS and LRS read operations for the studied RRAM-CMOS hybrid circuit. The upper graph represents the voltage reference and the read voltage when reading HRS and LRS resistances.

The basic functioning of the developed read driver is displayed in the signals transient shown in Figure 5.43. LRS and HRS resistance values vary the voltage dropped in the node V_{IN} . When compared against the voltage at the node V_{REF} , the driver's output varies to a "1" or to a "0" respectively.

This circuit has been analyzed taking into account the following uncertainty sources: CMOS process variability, temperature and voltage variations, RRAM variability, TID, degradation of endurance capabilities, and finally, neutron particle impacts.

5.8.1 Original Technology Models

RRAM Verilog-A Model

In this work we employ the physics-based RRAM Verilog-A model developed by Arizona State University researchers [CY15] [YU15]. Validated against physical devices, this Verilog-A model describes the resistive switching effect using a simplified 1-dimensional conductive filament. The CF dissolution and growth gets reflected on the variation of gap distance between the CF's end and the opposite electrode. Being I_0 and k_i the model conduction parameters [CY15], the RRAM equivalent resistance will depend on both gap length $-gap-$ and read

voltage $-v_{READ}$ — following

$$R_{READ} = I_0 \cdot k_0 \cdot \exp\left(-\frac{gap}{k_1}\right) \cdot \sinh\left(\frac{v_{READ}}{k_2}\right). \quad (5.8)$$

Additionally, the RRAM dynamics accurately model effects such as variability, thermal dependence or data retention failures under constant stress [CY15].

CMOS Models

The CMOS parts of the circuit shown in Figure 5.42 has been designed using a 40 nm commercial CMOS technology. These technology models provide the emulation of variability effects, thermal dependence and aging.

5.8.2 Target Circuit to Be Characterized

Using the above technologies, we have designed a simple driver to read the resistive state of a selected RRAM cell in a crossbar structure. The crossbar involves 16×16 RRAM devices, each of them coupled with a set of one capacitor and three resistances to provide parasitic effects. The column and row selectors circuitry are composed of four inverters and forty eight nand gates, while the low voltage sense amplifier involves nine transistors. The read operation is accomplished as follows: The read current flowing through the RRAM cell generates a voltage difference $-V_{IN}$ across a load resistor $-R_{LOAD}$. This voltage is compared against the reference level V_{REF} using an ad-hoc sense amplifier, determining if the RRAM cell is closer to HRS or to LRS. Being R_{OFF} and R_{ON} the resistance values of an RRAM in either HRS or LRS, the driver's output V_{OUT} would provide either HRS_VAL or LRS_VAL transients (Figure 5.43) when V_{IN} takes the values V_{IN1} or V_{IN2} respectively.

5.8.3 Design and Characterization Process

Models Redefinition

Applying the design strategy presented in Section 5.3, we start redefining the RRAM model in order to provide endurance degradation and TID emulation, thus building up an enhanced model.

The models redefinition process is illustrated by Figure 5.44. As we will emphasize at each point, the models are based on physical mechanisms and are suitable for the analyzed devices. As an illustrative example, Listing 5.1 describes the SPICE wrapper definition of the required degradation scheme. All remaining SPICE and Spectre subcircuits are available at [GRLVR15].

Figure 5.44: Scheme illustrating how using a SPICE subcircuit we redefined the RRAM original model.

Listing 5.1: Model redefinition to include durability

```

1 *****
2 ** http://vlsi.die.upm.es/memristor_spice_model
3 ** Use with Arizona University Verilog-A Model:
4 ** [CY15] Pai-yu Chen and Shimeng Yu.
5 *****
6 .SUBCKT emulated_durability_rram Plus Minus PARAMS:
7 **** Verilog-A parameters of original model
8 + gap_ini=1.367e-9 ** (HRS). For LRS gap_ini=0.5e-9
9 ** gap boundaries no degradation considered
10 + gap_max=1.7e-9 gap_min=0.1e-9
11 **** improved model
12 **** endurance parameters
13 + cycle_c=1
14 ** cf degradation ratios (metters per cycle)
15 + cf_hrs_deg=(3e-14)/1e3 cf_lrs_deg=(1e-14)/1e3
16 ** Cell fail cycle
17 + fail_cycle=f_fail_cycle( gap_max, gap_min, cf_hrs_deg,
18     cf_lrs_deg )
19 + gap_max_cycle = f__max_gap( cycle_c, fail_cycle,
20     gap_max, gap_min, cf_hrs_deg, cf_lrs_deg )
21 + gap_min_cycle = f__min_gap( cycle_c, fail_cycle,
22     gap_max, gap_min, cf_hrs_deg, cf_lrs_deg )
23 + gap_ini_cycle=max( min( gap_ini, gap_max_cycle),
24     gap_min_cycle )
25 **** end of improved model parameters
26 **** rram instantiation
27 rram0 Plus Minus gap_min=gap_min_cycle gap_max=
28     gap_max_cycle gap_ini=gap_ini_cycle
29 .ENDS emulated_durability_rram

```

Figure 5.45: RRAM model enhancement with endurance emulation as a function of cycling. Red and blue signals represent HRS and LRS resistance levels degradation.

Figure 5.46: RRAM model enhancement with TID emulation as a function of the total dose received. Red and blue signals represent HRS and LRS resistance levels degradation.

RRAM Endurance Degradation Emulation. Three distinct degradation patterns may affect RRAM cells [PAR16]. In this work we will use the most common one, which involves the progressive degradation of CF restoration capabilities, altering both HRS and LRS resistances [BAW⁺15, PAR16]. To incorporate endurance to the used RRAM Verilog-A model [CY15], we have developed a sub-circuit that translates the main endurance parameter -cycle_number- into the standard Verilog-A model parameters -the maximum and minimum gap elongation, gap_max and gap_min respectively-. Figure 5.45 represents this approach.

TID Effects. TID effects on RRAMs get translated into a slight change of mean and deviation values of both HRS and LRS resistances and/or switching voltages [MFH⁺15, LWHL15]. We have followed the same approach used to emulate the

device endurance degradation, building up a subcircuit and altering HRS and LRS resistance values. In this case, the redefined model varies the conductance Verilog-A parameter I_0 depending on the irradiation dose [LWHL15], producing the variation of the conduction properties as shown in Figure 5.46.

RRAM Variability. Even though all three variability types can be emulated using the injectors approach presented in [RLV14, GRGCY⁺16], *Device-to-Device* variability has been modeled following the original scheme shown in [CY15]. In this case HRS and LRS resistances mostly depend on an internal parameter, whose probability distribution is automatically handled by the proposed framework. To complement this approach, and using the guides shown in [GRGCY⁺16], *Intra-Device* variability and *RTN* current injectors emulate the remaining variability types.

Single Event Radiation Emulation. Similarly to the previous model enrichments, the radiation emulation is accomplished by redefining circuit nodes: a single event occurrence is traditionally emulated by a current injector representing the collected charge in an impact. This arduous task is automatically performed by the framework: first, each node susceptible of being radiated is identified, and then independently radiated and analyzed. For the present experiment we have considered neutron impacts as the radiation source affecting our circuit. The collected charge effects for a 40 nm technology working under neutron induced radiation are modeled using the well known *double exponential* current source approach [NUT⁺12].

Performance-Reliability Metrics Definition

Figure 5.47 represents the different metrics used to evaluate the circuit reliability and performance. The considered *EMs* refer to the data stored in the RRAM cell. The first metric, *EM_0*, manifests a failure in the case the stored value is altered, while the second one, *EM_1*, raises a malfunction when this value is misread. If during the on-going analysis either one or both failures are declared, the scenario will be tagged as unreliable.

The first performance metric, *PM_0*, has been defined to evaluate the delay between read enable signal *READ_EN* and the output *V_OUT* edges. The second *PM*, *PM_1*, refers the sense amplifier *input voltage margin*, given by the voltages *V_IN1* and *V_IN2*. By using higher read voltages, the sense amplifier produces narrower *V_OUT* pulses. Finally, *PM_2* measures output pulse widths. Furthermore, for most analyses these *PM* metrics can also be used to define a *EM*.

Figure 5.47: EMs and PM metrics used in the evaluation of the case of study. Notice how EM_0 errors do not consider recoverable glitches, while EM_1 metric evaluates the CF gap even out of the read window time span.

More in detail, during the following experiments, our specific metrics are defined as:

- EM_0 : Being gap_0 the gap value at the start of the read operation, an error is raised if the gap is altered by $0.05nm$. This threshold may be considered small during a single operation, but consecutive reads would exponentially alter this distance.
- EM_1 : Being D_0 the binary value related to the resistive state stored in the RRAM cell, an error is declared with the condition this value does not match the signal V_{OUT} . To avoid glitches, this misread should be not punctual, and maintained at least half the system clock period -1ns in the experiments- during the *read_window* (Figure 5.47).
- $PM_0_EM_2$: Delay between $READ_EN$ and V_{OUT} signal edges, being lower delays more desirable. Scenarios and configurations producing a delay surpassing the 10% of the system clock period are considered as non-reliable.

Figure 5.48: PM_1 approximation dependence on v_{READ} and R_{LOAD} .

- $PM_1_{EM_3}$: $V_{IN1} - V_{IN2}$ voltage difference, being larger voltages more desirable. Voltage differences lower than 20 mV are considered as non-reliable.
- $PM_2_{EM_4}$: Output pulse widths below a 50% of the system clock period are considered as non-reliable.

Design Space Definition

With the support of our framework and following the proposed strategy, we can define the most suitable design space to be explored.

Component Values Selection. In order to determine the optimum R_{LOAD} value, we calculated the read margin dependence on both v_{READ} and R_{LOAD} parameters. Figure 5.48 shows how larger v_{READ} values lead to higher PM_1 , and enlightens the optimal R_{LOAD} value.

To select the best v_{READ} value, we simultaneously evaluated all considered metrics. In Figure 5.49, yellow shadowed areas (below 0.15V) display the EM errors, in this case triggered by a misread (system failure during the evaluation of EM_1). Consequently, voltages higher than 0.15V are needed. In the same way, if only PM_1 were contemplated, higher read voltages would be better.

However, and highlighting the importance of the multi-metric evaluation paradigm, the system performance and reliability can be improved. As seen in Figure 5.49, to reduce punctual glitches in the output (EM_1), as well as the exponential non desired growth of the CF gap (EM_0), lower voltages are recommended. The same applies to the performance metric PM_0 , which -until 0.8V- recommends the same. With the above information we can set the system v_{READ} voltage at 0.2V, value used in later simulations.

Figure 5.49: Metrics evaluation for v_READ . All punctual failures (recoverable) or failures not surpassing the considered thresholds are plotted in blue. Not recoverable failures or failures surpassing considered thresholds are plotted in red. Yellow shadowed areas represent EM errors.

Endurance, Circuit Temperature and TID. To complete the design space definition we addressed the circuit correct functioning when varying the environmental and stress conditions. All possible parameter combinations were automatically simulated and analyzed, and finally, statistical reports were automatically generated. Every scenario was statistically processed, and to ease their interpretation, the framework automatically grouped them into three different maps: A, B and C (Figure 5.50). Maps labeled with A gather the temperature versus endurance EM weighted-mean results, considering different TID doses. Similarly, B and C maps represent temperature versus TID and TID versus endurance respectively.

Figure 5.50: Environmental and stress conditions evaluation. Blue areas represent non upsets while color progression towards red indicates EM errors.

It can be extracted that the progressive degradation of the CF, with the consequent increase/decrease of LRS/HRS resistance levels, is established as the most harmful error source. We can identify a threshold starting from which resistance values cannot be properly read. Hence, it raises the cycle number as the key parameter that triggers both EM_0 and EM_1 errors, and consequently, we can identify and enclose the environmental and stress conditions compromising the circuit functioning. With these reports we could study the conditions leading to unreliable behaviors while reducing the computational load associated with large parameter ranges.

Figure 5.51: Reliability characterization, no individual neutron impact considered. The reports are summarized by paired parameters.

PVT, Endurance and TID Characterization Results

With the design space properly defined, we accomplished the most computationally hungry characterization, simultaneously studying TID, endurance and temperature. Making use of the framework, every analysis was automatically accomplished, computing partial and summary reports, hierarchically organized by paired parameters, and allowing to retrieve individual transient and statistic reports. Next we present a representative set of those reports, following a top-to-bottom approach. First, Figure 5.51 describes the upsets ratio summaries for both HRS and LRS reads after *EMs* evaluation, grouped again by the previous A, B and C parameter pairs. These general summaries are computed as the weighted mean of the corresponding partial results, S_i . Here, every contribution S_i represents the partial reports regarding individual scenarios, sharing specific parameter evaluations.

These general reports, which extract the essential effects of the uncertainty sources and dependence on parameters, are the most useful and prime source of information a circuit designer may have. With those simple diagrams, the framework user is able to address the circuit reliability and performance dependence on each possible parameter pair. Summarizing, this general view gives the designer the big picture on how the system behaves, which may be a too much simplified view. To that end, in detail reports are also provided.

To clarify the weighted mean leading to these summaries, Figure 5.52 shows how each S_i contribution regards a specific particularization of the remaining parameters under evaluation. In this case the mean upset summary, computed for the parameter pair *Temperature* versus *Cycle-count* is decomposed into several scenarios S_0 to S_{10} , each of them related to a specific TID dose.

Finally, we illustrate how the upsets maps are computed. The circuit designer may also study individual metric maps, extracted from the myriad of —also accessible— transients. Figure 5.53 shows how representative EM_0 and EM_1 metrics, both evaluating S_{10} partial analysis, are displayed.

Figure 5.52: Reliability characterization, no individual neutron impact considered. HRS (A) summary map —from Figure 5.51— is composed computing individual scenario contributions — S_i . To illustrate the metrics evaluation, the partial map S_{10} displays its EM_0 and EM_1 results.

Figure 5.53: Reliability characterization, no individual neutron impact considered. HRS (A) summary map has been decomposed into S_i contributions. To illustrate the metrics evaluation, partial map S_{10} displays its EM_0 and EM_1 results.

Figure 5.54: SEEs Critical charge compromising the studied circuit reliability in presence of additional uncertainty sources: TID, endurance and temperature. In these graphs, and contrary to Figure 5.51, higher values —displayed in blue— represent more hardened scenarios, being able to support higher radiation doses.

Full Characterization Results

Complementing these results, we emulated the existence of *Single Event Effects* (SEEs). Figure 5.54 shows the A parameter pair report summarizing the critical charge Q_{crit} dependence on the considered uncertainty sources. Q_{crit} can be described as the minimum charge (Q_{coll}) that collected by the circuit produces a functional error. Consequently, higher Q_{crit} values represent more hardened -and therefore more reliable- scenarios.

In contrast to Figure 5.51 maps, and being j the circuit nodes radiated, here each S_i report is composed by $S_{i,j}$ partial contributions. Similarly to the processes described by Figures 5.52 and 5.53, the circuit designer is able to access not only the general but also the partial reports and every one of the individual analyses. However, in these analyses we find each report reproduced by the number of radiated nodes. Therefore providing information on how the circuit behaves depending on the impacted node for each set of considered parameters, in a wide range of impact instants, while including the effects of aging, endurance, PVT variations and TID.

To sum up, we automatically obtained conclusive reports describing the overall behavior of the system while being able to access every specific partial result. The circuit designer gains instant access to the root of the possible problems that may compromise the system correct functioning. To properly characterize the circuit, during these two last analyses we estimated that 6,600,000 short transients were automatically simulated and processed, achieving a throughput of 30.55 *trans/sec*, and requiring less than 70 *h*.

5.9 Conclusions

The characterization of present and future circuits involving short-node CMOS technologies and, more especially, emerging devices, requires changes in both methodologies and supporting tools. On the one hand short-node CMOS technologies suffer the effects caused by uncertainty sources. Traditional methodologies and characterization tools do not consider their combined effects, undermining the accuracy of the reliability studies. Closely related to this issue, novel and more accurate approaches taking into account the combined effects of paired uncertainty sources require from tedious manual steps that limit their application.

On the other hand, emerging technologies suffer from intrinsic uncertainties, compromising the circuit reliability and performance. Compared against the better known CMOS problems, these effects impact the circuit differently. The methodologies for the characterization of traditional circuits are not able to handle the needs attached to these emerging devices. More importantly, the inclusion of these novel technologies within standard circuit design flow is hindered by the lack of solid compact models that consider multiple uncertainty sources.

To address these issues, in this Chapter we have presented a two fold solution, based on a circuit analyses strategy supported by a characterization framework. This technology independent solution focus on the study of RRAM-CMOS hybrid circuits suffering from the non-desired effects produced by simultaneous uncertainty sources. The main contributions of this work are the following:

- We have presented a new analyses strategy that simultaneously employs multiple user-defined metrics to measure circuit performance and reliability. Additionally, this strategy relies on the redefinition and use of enhanced models allowing different uncertainty sources to be emulated. Finally, the third pillar this strategy is based on the efficient definition of the design space, allowing the circuit to be correctly and thoroughly characterized.
- We have presented a simulation framework [GRLVR15] that, interacting with SPICE simulators, automates the performance and reliability of RRAM-CMOS circuits. Compared to standard design flows, where only two-paired unreliability sources are considered, the proposed framework improves the analyses depth. This approach allows the study of possible interactions between different technology or environmental error sources that may not be detected by traditional approaches, involving individual or paired error source analyses.
- We have presented a solution that is technology-independent, where both analyses strategy and support tool allow the simulation of other emerging technologies and analog modules. Furthermore, a wide range of variability or endurance schemes, radiation impact sources, as well as the simulation of other uncertainty sources can be incorporated.

We validated our reliability simulator against experimental measures. Moreover, we illustrated the proposed strategy through three different cases of study. First, we analyzed the hardening levels against radiation of a digital module, extending traditional methodologies. Next, a thorough case of study, involving a power supply sensor circuit suffering from PVT variations together with radiation impacts highlighted the presented solution capabilities when characterizing analog circuits. Finally, illustrating the full power of the developed solution, we presented a case of study of an RRAM read driver. The characterization considers different design parameters as well as RRAM inherent constraints or environmental error sources: RRAM lifetime, temperature, CMOS and RRAM variability and radiation—both TID and single particle impacts.

5.10 Related Contributions

- **On the Design and Analysis of Reliable RRAM-CMOS Hybrid Circuits.**

Fernando García-Redondo; Marisa López-Vallejo.

IEEE Transactions on Nanotechnology. Accepted and recommended for publication (April 2017).

- **Circuit Reliability Framework**

Public Repository:

https://github.com/fgr1986/circuit_reliability_framework

Reliability oriented framework for the characterization of CMOS/RRAM hybrid circuits.

- **Characterization of Analog Modules: Reliability Analyses of Radiation, Temperature and Variations Effects.**

Fernando García-Redondo, Hernán Aparicio-Cerqueira, Marisa López-Vallejo, Pablo Ituero and Carlos A. López Barrio.

Conference on Design of Circuits and Integrated Systems 2016 (DCIS 2016) Granada(Spain) 23-25 November 2016.

- **Reliable Design Methodology: The Combined Effect of Radiation, Variability and Temperature.**

Fernando García-Redondo, Marisa Lopez-Vallejo, Hernán Hernán Aparicio Cerqueira, Pablo Ituero.

13th IEEE International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD 2016), Lisbon (Portugal). 27 - 30 June 2016.

- **Taxonomy of Power Supply Monitors and Integration Challenges.**

Pablo Ituero, Marisa López-Vallejo, Hernán Aparicio, Fernando Garcia-Redondo.

Mixed-Signal Testing Workshop (IMSTW), 2016 IEEE 21st International.

- **Evolution of radiation-induced soft errors in FinFET SRAMs under process variations beyond 22nm.**

Pablo Royer, Fernando García-Redondo, Marisa López-Vallejo.

Conference: Nanoscale Architectures (NANOARCH), 2015 IEEE/ACM International Symposium on. Boston, MA, USA (July 2015).

- **A Tool for the Automatic Analysis of Single Events Effects on Electronic Circuits.**

Fernando García-Redondo, Marisa López-Vallejo, Pablo Royer, Javier Agustín.

5th European Workshop on CMOS Variability, PATMOS VARI 2014, Mallorca (Spain).

- **Four-Injector Variability Modeling of FinFET Predictive Technology Models.**

Pablo Royer, Marisa López-Vallejo, Fernando García-Redondo, Carlos A. López Barrio.

5th European Workshop on CMOS Variability, PATMOS VARI 2014, Mallorca (Spain).

Part III

Reliable Design of RRAM Based Systems and Applications

*Now, My Usual Fee Is 500 Bucks. But Seeing As How It's You,
I'm Gonna' Need It In Advance.*

Bender, Futurama.

6

Reconfigurable Writing Architecture for Reliable RRAM Operation in Wide Temperature Ranges

6.1 Introduction

Chapter 2 presented the interesting properties that make memristors a solid alternative to present applications and a promising candidate for future ones. However, Section 2.5 also pointed out how multiple error sources may cause dramatic effects on RRAM-CMOS hybrid systems, endangering the application functioning. To solve this issue, throughout Chapter 3, Chapter 4, and Chapter 5, we have presented methodologies, models and support tools that handle these uncertainties, first starting at the device abstraction level, then addressing the problem at circuit abstraction level.

Figure 6.1: Chapter 6 scope and perspective within the Thesis.

Chapter 6 and Chapter 7 make use of this knowledge and tools, employing them at system and application abstraction levels. Figure 6.1 describes how the contents of the present Chapter can be seen within the thesis context.

In this Chapter we will focus on solving, at architecture level, the problems caused by the highly-nonlinear thermal dependence of memristive cells. Throughout Sections 1.1 and 2.4 we showed how in nowadays —and predictably in future— designs memory blocks grow both in number and in size, and how RRAM stands as a natural solution for the next generation of nonvolatile memories —as well as distinct crossbar based applications. Thus, it is of critical importance to provide reliable solutions to these applications when ported to RRAM based circuits.

Therefore, RRAM must deal with several integration obstacles before been established as the next memory standard. In this sense, performance and reliability are two key factors that can determine the success of RRAM devices over other novel technologies. As we saw in Section 2.5, process variations, crossbar design and non-linear selector devices inclusion are the most concerning issues that nowadays limit direct application of resistive-switching technologies. To overcome the related design challenges, reliable [GLmG⁺15, KYS⁺16], fast [CWC⁺15], and low-power [HMM13, GLmG⁺15, CWC⁺15] writing schemes have been proposed. Although most efforts focus on device to device and cycle to cycle cell variations, temperature is also a critical issue to be addressed in order to ensure the correct behavior of RRAM circuits.

The wide temperature range under which resistive switching devices operate constricts the memories trustworthiness. The conductive filament, defining the resistance state, grows and dissolves within the memristor with processes highly affected by its temperature. This local temperature plays a fundamental role in the resistive switching mechanisms and conductive processes [WWS⁺11, AKG⁺14, YGCR15, SNK15]. Thus, together with the electrical inputs, the circuit working temperature describes the dynamics of the conductive filament heating processes [Iel16]. Consequently, characteristics such as switching speed [GRLV15] or data-

retention ability [CI15] highly depend on the working temperature. To the end of our knowledge, until [GRLVI15] no driver had dealt with temperature caused problems.

But not only does the writing scheme and working conditions affect reliability. In the same manner, the device endurance strongly depends on the writing voltage characteristics [HCW⁺13, BAW⁺15]. The use of a scheme able to guarantee the memory block operations does not ensure an optimal endurance for the devices and may cut back the useful lifetime. As presented in [MCYC15b, MCYC15a, MCYC16], depending on the application, reliable, power saving or faster profiles may be used. Therefore, usually a trade-off between these characteristics is fixed at design time [KMG⁺14].

Looking ahead, it is very likely that RRAM based applications would be employed in very different environments. In this sense, from nonvolatile memories to reconfigurable devices, parallel datapaths, neural networks and approximated-computing processors, memristive devices will be required to work under varying conditions. If the working conditions vary -temperature range, computation load, etc.- the target application may benefit from a different and dynamically loaded profile. For example, aerospace circuits working at different altitudes experience environmental temperature changes. In such environments, RRAM data retention characteristics could improve by dynamically adapting the used writing amplitudes. The same principle could be applied to control the memory block endurance before employing drastic system level solutions relying on replicated hardware [PAR16]. Taking advantage of a more suitable writing amplitude, related to a different HRS/LRS ratio [MCYC16], the memory lifetime would improve.

It would be highly desirable that, once the reliability is guaranteed, the application user could take advantage of a set of writing configurations [MCYC16], maximizing the performance towards a specific application target.

To address these issues, in this Chapter we present a writing driver adaptable to the memory working conditions that, accordingly to specific application objectives, can dynamically load different writing profiles. This reconfigurable capabilities are built on top of the required hardware to ensure achieving correct operations, no matter the working temperature, in return for a small area overhead.

The circuit temperature catalyzes the switching mechanisms that define the device state. For low temperatures, higher voltages and/or larger pulses (involving more power consumption) are required. Otherwise, using standard writing schemes would result in writing failures. Therefore, for critical applications working at temperatures as low as 200 K, we must ensure the correctness of the operation. Similarly to *Dynamic Voltage* and *Frequency Scaling* —DVS, DFS— techniques, the dynamic driver adapts the RRAM memory writing scheme behavior to flexible targets, opening a wide range of programming strategies

[MCYC15b, MCYC15a, MCYC16] to be dynamically applied. Thus, the driver selects among a myriad of writing strategies improving the system performance, power consumption, reliability, data retention abilities and HRS/LRS ratio durability, always considering the circuit working temperature as a critical parameter for the memory reliability.

The developed architecture has been modularly designed, ensuring its compatibility with nowadays resistive-switching architectures. Thus, the dynamic driver works for both *1-Transistor-1-memristor (1T1R)*, *1R* and *1-selector-1-memristor (1S1R)* configurations, built in standard *2D* or stacked *3D* architectures [YC16, CLY16].

In this chapter we first analyze the impact that the circuit working temperature has on resistive memories reliability. We present the relationships between voltage amplitudes, pulse lengths and temperature, always guarantying the correct memory behavior. Using the model jointly developed by Stanford, Arizona and Beijing university [LJH⁺15], we extract the writing requirements under which, at a specific temperature, the memory is able to be correctly written. This model takes into account the parasitic capacitance/resistance that appears together with the resistive switching structure.

In a second stage we present the proposed driver, describing the hardware modules that allow using dynamic profiles and detailing how it guaranties the circuit reliability while working on a wide range of temperatures.

In order to enlighten the driver capabilities, different writing profiles are analyzed:

- Profile sets with dissimilar HRS/LRS ratio, providing different device endurance [MCYC15b, MCYC15a, MCYC16].
- Profile sets offering extra protection against temperature misreads that may be caused by unexpected changes in the die thermal flow, or device to device thermal properties variations.

We bring into comparison their reliability capabilities and power consumption, and show the speed-ups this driver achieves when compared against other schemes.

To develop these ideas, we structure the chapter as follows. First we describe the memristor dependency on the temperature, and how the selected model handles this thermal dependency. Second, the proposed driver is presented together with the novel dynamic profiles loading system. After that, we define the analysis to determine the temperature resolution to be used. Based on the selected temperature resolution, we describe the full system implementation and related area cost. Section 6.5 presents the achieved speed-ups and power consumptions for different profiles. Finally, some concluding remarks are discussed in Section 6.5.

Figure 6.2: HRS and LRS state representation together the used compact model.

6.2 Resistive Switching Dependence on Temperature

The present chapter uses the well known Verilog-A compact model of metal-oxide RRAM developed by Stanford University and Beijing University [LJH⁺15]. This model has been validated through several oxide devices and covers most physical variables/mechanisms involved in the resistive switching process. More importantly, this model emulates the physical device dependence on temperature [HWL⁺14, HCL⁺14, LJH⁺15]. As seen in Figure 6.2, the device is composed of a conductive filament, located between the top and bottom electrodes, and surrounded by the MIM area. The device state will be between LRS and HRS states, depending on whether the CF is formed or dissolved. The transitions from HRS to LRS and LRS to HRS are called *SET* and *RESET* respectively. The used compact model describes the device using a conductive cylinder of variable height and width to emulate the CF dynamics.

The CF evolution is defined by the generation/recombination of oxygen vacancies and oxygen ions at the edge of the CF. For the *SET* operation, this model considers that the CF first grows in height until it reaches the opposite electrode, and then it widens the cylinder base. Translated to the section scheme in Figure 6.2, first the gap x is reduced and then the width w would reach its maximum. On

the other hand, the *RESET* process is modeled using x gap alone: the application of a negative voltage will result on the CF dissolution, and therefore on the increment of x until its maximum.

Both the applied voltage and CF temperature play key roles in CF dynamics [WWS⁺11, AKG⁺14, YGCR15, SNK15], catalyzing the involved processes. Thorough studies revealing the significant role of the circuit working temperature -and therefore the RRAM electrodes initial temperature- can be found in [VGC⁺16, PWG⁺17]. Being a_h, b_i, c_j, d_k constants, the speed of the CF evolution during the *SET* is determined by the following equations:

$$\frac{dx}{dt} = a_1 \exp\left(\frac{-a_2 - a_3 V}{T_{CF}}\right) \quad (6.1)$$

$$\frac{dw}{dt} = \frac{b_1}{w} \exp\left(\frac{-b_2 - b_3 V}{T_{CF}}\right) \quad (6.2)$$

while the *RESET* process speed is defined by

$$\frac{dx1}{dt} = c_1 \exp\left(\frac{-c_2 - c_3 V}{T_{CF}}\right) \quad (6.3)$$

$$\frac{dx2}{dt} = d_1 \exp\left(\frac{-d_2}{T_{CF}}\right) \sinh\left(\frac{d_3 V}{T_{CF}}\right) \quad (6.4)$$

$$\frac{dx}{dt} = \min\left(\frac{dx1}{dt}, \frac{dx2}{dt}\right) \quad (6.5)$$

On the other hand, the current flowing through the device heats the CF. Naming R_{th} as the CF equivalent thermal resistance, the CF thermal behavior is modeled by the Joule-heating equation:

$$T_{CF} = T_0 + IV R_{th} \quad (6.6)$$

where T_0 refers the circuit working temperature.

The complex non-linear relation between the device state and the circuit operating temperature and applied voltage is displayed in Figure 6.3. As can be noted, the boundary conditions imposed by T_0 greatly influence the device dynamics. Both *SET* and *RESET* operations have been simulated at each pair of voltage amplitude and circuit temperature. The graphs represent the writing time required to completely perform the corresponding operation. Voltage amplitude-temperature pairs at which the operation is not correctly achieved have been computed but omitted to improve the graph readability. In agreement with [VGC⁺16, PWG⁺17], it can be concluded that, given a fixed operation length, to succeed in writing at lower temperatures demands higher amplitudes. Therefore, and most especially in aerospace applications where working at extremely low temperatures is a requirement, the traditional two level amplitude scheme, one for *SET*, one for *RESET*, can no longer be applied.

Figure 6.3: Required time to perform standard SET and RESET operations using a fixed operation voltage, being T_0 the circuit working temperature. Operations where the CF did not achieve the desired geometry within the maximum considered (1.2 us), and therefore requiring wider pulses, are plotted in red.

Figure 6.4: Developed writing architecture. The figure shows an S -sensors network connected to the temperature-handler core, together with the different voltage references supplier. The rows/columns demux-mux circuitry provides access to an $M \times N$ RRAM memory. For 1R and alike schemes the column demux and switch-mux accessing the transistor gates in 1T1R is not needed.

6.3 Dynamic Profile Thermal-Adaptive Write Architecture

We have seen how using RRAM based systems has two major issues regarding writing amplitudes and working temperature, that need to be addressed. Firstly, higher voltages are needed to ensure the correct writing of memory blocks working at low temperatures. Secondly, the use of amplitudes larger than required would lead into functional errors, endurance and leakage problems.

The developed writing architecture is based on a writing driver that provides the most suitable writing amplitude depending on the working scenario and the application strategy. This dynamic writing scheme, similarly to *DVS* and *DFS* techniques, adapts the resistive switching blocks behavior to the circuit working temperature and application requirements. Consequently, our solution increases the system adaptability and releases the circuit designer from the hard trade-offs early imposed by a rigid, single amplitude driver.

Standard 2D RRAM Blocks Distribution

Standard 3D RRAM Blocks Distribution

Figure 6.5: Presented architecture interconnection with memory blocks in standard 2D and stacked 3D morphologies.

The functioning of the presented driver relies on a *sensing-acting* approach. The simplified process is as follows: the operation starts with the memory block temperature sensing, followed by the selection of the most suitable writing voltage amplitude, and ends performing the writing operation.

The modular architecture of the writing driver is represented in Figure 6.4. It relies on four main modules:

- S temperature-sensors surrounding the memory blocks.
- A light communication network [ILVMO12].
- The *profile-temperature-handler core* or PTHC.
- The *voltage references system* or VRS.

The architecture has been designed to adapt to and be included within standard RRAM array schemes. Figure 6.5 describes the architecture interconnections on both 2D and 3D stacked RRAM arrays. There, the interconnection and distribution of surrounding sensors, submodules, and light-network guarantees design compatibility with both *1T1R* and *1R/1S1R* memory blocks: CMOS temperature sensors are placed next to the standard read and write access circuitry, close to the resistive memories. Low power CMOS temperature sensors present in the literature [ILVLB13], or add-hoc sensor solutions can be interconnected using a light network [ILVMO12].

This network scheme serves for calibration, sensing and transmission of any thermal information to the *profile-temperature-handler core*. Additionally, and

thanks to its time-multiplexing scheme, this network could also drive supplementary information, such as devices aging [ILVMO12], providing valuable intelligence to be used for the dynamic profile selection by the *profile-temperature-handler core*.

Using the proposed driver, all temperature-caused errors are fully mitigated, achieving the correct writing of the desired value. During this operation, a specific voltage amplitude is dynamically selected, ensuring that the writing process uses the most suitable voltage to accomplish the process regarding the loaded profile target.

6.3.1 Optimal Voltage Amplitude Selection

Writing operation starts by loading a temperature-voltage amplitude profile in the *profile and temperature handler core*, targeting a previously defined configuration (endurance, data retention, power consumption, etc.).

When a write operation starts, first, the *profile and temperature handler core* -or *PTHC*- decides whether a new temperature sensing should be performed or it relies on a previous measure. If a new temperature shall be acquired, a set of S different temperature measures are taken by the sensors surrounding the memory. These measurements are driven to the main core where the target RRAM cell temperature is computed. The required voltage amplitude is selected based on this temperature among a discrete set of values, and once driven to the selected cell, the writing operation takes place. The algorithm defining which values are to be included within this voltage amplitudes set is described in Section 6.3.4.

6.3.2 Profile-Temperature Handler Core

The *profile and temperature handler core*, represented by Figure 6.6, is in charge of the remaining operations.

The writing evaluation process of the voltage amplitude that best suits the environmental conditions is described by Figure 6.7. First, the *sampling control module* manages whether or not a new sensing measurement should take place. Based on the expected heat flow H_{max} , a sampling frequency f_s defines the error estimation [IGRLV12]

$$\epsilon_{sampling} = \frac{H_{max}}{f_s}. \quad (6.7)$$

Usually, given the large inertia the temperature flow has, a low frequency is enough to correctly sense the memory temperature changes, barely affecting energy consumption and system performance. The default hardware implementation of *Sampling control module*, which controls f_s , relies on a simple configurable end-count counter.

Figure 6.6: Main core architecture interfacing complementary modules: different voltage references system and optional variation aware module [JJMK10, GLmG⁺ 15].

On the other hand, regarding spatial-caused errors, silicon imposes a heat flow behavior [IGRLV12] that leads to a spatial error contribution described by

$$\epsilon_{spatial} = T_0(1 - \exp(\frac{-2r}{k})) \quad (6.8)$$

Here, T_0 denotes the hot-spot temperature, r refers to the distance from which the temperature is sensed and k is related to the thickness of the circuit package. The main core, using the measures information together with the distance between the cell and the different sensors is able to compute the cell temperature (*temperature computation module*). The use of Manhattan distance simplifies the computation of the temperature as the row/column information can be directly applied, but other distances can be employed.

As seen in Figure 6.8, using the S different temperature measurements, several schemes can be computed at the temperature-handler core: from using the sensed temperature of the nearest sensor

$$T_{RRAM} = T_i | d_i = \min(d_j), j \in [0, S - 1] \quad (6.9)$$

to weight the sum of all sensed temperatures. This multi-contribution computation scheme is described by

$$T_{RRAM} = \sum_i T_i w_i \frac{d_i}{D_{total}}, i \in [0, S - 1] \quad (6.10)$$

Figure 6.7: Sensing-voltage amplitude selection process flow.

Figure 6.8: Temperature computation modes.

where, for a sensor S_i , T_i and w_i stand for its sensed temperature and configurable weight respectively, d_i is the distance to the target cell and D_{total} is the sum of distances from the cell to each sensor. An optimal distribution based on area, power consumption of both sensors and network, temperature evaluation scheme and achieved accuracy can be calculated using the algorithm described in [IGRLV12].

If a new temperature measurement is commanded, the *PTHC* module first de-multiplexes the time-multiplexed thermal information broad-casted by the network, and reconstructs the temperature sensed by each one of the S sensors.

Using these temperature reads and the cell location, the *temperature computation module* computes the cell temperature. This temperature value is stored in the *computed temperature buffer* to be used until —commanded by *sampling control module*— its value is overridden. This temperature value decodes the most suitable writing voltage amplitude index stored in the look-up table —*LUT*. Finally, this voltage index selects the desired amplitude, wired from the *different voltage references system* to the row/column drivers, achieving the satisfactory writing of the desired value in the target cell.

Figure 6.9: Proposed calibration methodology for physical devices.

6.3.3 Calibration of Physical Devices

The temperature-voltage amplitude relationships can be extracted by means of two different procedures. If the device model is accurate enough and the parameter variations during the fabrication stages do not compromise the memory behavior, a temperature characterization can be accomplished using circuit simulators as shown in Section 6.2 for $1R$, $1S1R$, and $1T1R$ models.

Alternatively, a calibration process of the physical memory block can take place. Compared to the maps shown in Figure 6.3 and given the impossibility of accessing the CF sizes, in this approach the memristor resistance should be measured instead. Figure 6.9 explains the methodology for the characterization of physical devices. Using a stable thermal chamber, the circuit can be heated to the required temperature. For each circuit temperature considered, a full write-read-reset cycle should be performed. At the end of the measures, a map showing the achieved memristance dependencies on the circuit temperature, like the ones shown in Figure 6.3, is obtained.

The calibration process ends with the selection of the optimum voltage amplitude at each temperature. Due to variability in the memory blocks a calibration should be accomplished for each one. Without needing an additional adjustment stage, during this same process the set of sensors surrounding the memory can also be calibrated, reducing the sensing error and achieving more reliable measures.

Figure 6.10: Illustration of the voltage amplitude selection algorithm, applied over maps —like the one shown in Figure 6.3.

6.3.4 Optimum Voltage Amplitude Selection Algorithm

Instead of picking the minimum voltage that at each temperature fulfills the write operation, we propose to include a safety margin regarding temperature, gaining robustness against thermal misreadings. The selection algorithm can be applied both at early and later design stages such as post-fabrication calibration.

Lets first define g as the analyzed magnitude (CF width during *SET* or gap height during *RESET*), and a the value of g at which the write operation is correctly performed. Chosen $T = \{t_j\}, j \in [0, p-1]$ and $V = \{v_i\}, i \in [0, q-1]$ as the equidistant set of temperatures/voltages at which we have measured/simulated the device, the considered temperature and voltage resolutions are defined as $\Delta t = t_j - t_{j-1}$ and $\Delta v = v_i - v_{i-1}$ respectively. We simulate/measure the memory at each pair (t_j, v_i) defining the binary matrix $D = \{d_{ij}\}, i \in [0, p-1], j \in [0, q-1]$, where

$$d_{ij} = \begin{cases} 1 & \text{if } g(v_i, t_j) > a \\ 0 & \text{elsewhere} \end{cases} \quad (6.11)$$

At the temperature t_j the required voltage c_j to safely write the desired value is defined as

$$c_j = \min(v_i) \mid d_{ij-n} \& \dots \& d_{ij-1} \& d_{ij} \& d_{ij+1} \& \dots \& d_{ij+n} \quad (6.12)$$

Then, the voltage set $C = \{c_j\}$ ensures a correctly achieved writing operation at any temperature t within the range $[t_0, t_{p-1}]$.

For a fixed thermal resolution, the parameter n defines the robustness of the writing scheme against temperature variations due to a fast heat flow in the memory: For each considered temperature t_j , the algorithm finds the minimum required voltage that is able to fulfill a proper writing operation at t_j as well as at the neighbor temperatures $t_{j-n}, \dots, t_{j+n}$. Therefore the higher the value of

Figure 6.11: Voltage amplitudes required for 1R and 1T1R cells schemes.

n , the wider the temperature misreading range the system is able to mitigate. Figure 6.10 represents a graphical description of the algorithm for $n = 2$. Figure 6.11 displays the voltage amplitudes obtained when applying this algorithm to the thermal characterization results of Figure 6.3. A temperature resolution of 2°C and a robustness value of $n = 1$ were used.

6.3.5 Dynamic Profiles

In [GRLVI15] a new methodology for the calibration of physical devices is presented together with an algorithm for the selection of optimum voltage amplitudes regarding thermal-variations protection and power consumption.

As introduced in Section 6.1, those voltages can be specified not only to ensure a correct operation no matter the working temperature, but to optimize different memory characteristics: device endurance, data retention, thermal-variations extra protection, power consumption or a solid HRS/LRS ratio. Different writing profiles suit different applications. Depending on the application orientation, a distinct profile would be desirable to match a particular target. Our proposal allows to dynamically change those oriented profiles, optimizing the memory behavior given a desired application target and/or environmental characteristics.

Figure 6.12: Dynamic profile loading schedule, profile-writing-module. During the read-only mode the new profile, containing the information relative to every considered temperature, is sequentially stored.

The *profile writing module* subcircuit, shown in Figure 6.6, provides this functionality by safely replacing the values stored in the temperature-voltage LUT for those related with the new profile to be loaded.

Profile Writing Process

The proposed scheme adds the hardware needed for the reconfiguration of the LUT locating the temperature-voltage decoding information.

The *profile writing module* handles the process as detailed in Figure 6.12. To avoid possible writing problems, after the chip selector signal —`pwm_CS` in Figure 6.6— enables the profile-writing module, the resistive memory block is momentarily set to read-only mode. When the reconfigure signal, `pwm_r`, is activated, the new temperature-voltage decoding codes are in-series stored within the LUT. At the end of the process, and after `pwm_r` signal is deactivated, the resistive memory block is restored to write/read mode, and the updated profile can be used.

6.3.6 Voltage References System

The *different voltage references system* provides the required U discrete stable voltage amplitudes. To guarantee the stability of these signals, a bandgap voltage reference module has been used.

Figure 6.13 describes the implemented system. Instead of using U bandgap modules to provide U different outputs, we employed a single bandgap that stabilizes the voltage V_{BG} , followed by a CMOS voltage divider that generates the remaining amplitudes. This approach reduces the consumed area to near

Figure 6.13: Voltage references system, composed of a single bandgap voltage reference, followed by a U -stage CMOS voltage divider to provide the U different voltage outputs.

$1/U$ of the full version. Additionally and due to the n-transistor proximity, by using the CMOS voltage divider we reduce process variation effects. At design stage, by tuning R_{eqX} and C_{eq1} equivalent resistors and capacitors the bandgap robustness against PVT can be optimized [GRALV⁺16].

6.3.7 Optional Modules

Variability in RRAMs is one of the most concerning issues to be solved. Both device to device and cycle to cycle fluctuations can compromise the system reliability. To fight against variability caused problems, systems like [JJMK10, GLmG⁺15, CWC⁺15] provide variability protection during write operations.

The modular design of the proposed driver allows embedding this kind of solutions. Figure 6.6 shows how the writing voltage leaving the *temperature handler core* would be controlled by the *variation aware module*. Using the reference i_{ref} and feedback i_{sensed} currents, the *variation aware module* controls the operation state and handles the writing voltage accordingly.

6.4 Full System Implementation

To study the hardware sizing and area cost related to the proposed driver, different implementations have been built using commercial 40 nm CMOS technology. Along this section we will describe in detail the *profile writing module* as well as the study of the hardware insights and sizing.

6.4.1 Profile and Temperature Handler Core

LUT implementation

The area overhead introduced by the *profile-temperature-handler core* is mostly dominated by both LUT's and combinational hardware contributions. This LUT size for *SET* and *RESET* will depend on the resolution as well as on the considered temperature range.

$$LUT_{size} = p = \#T = \frac{temp_{min} - temp_{max}}{resolution} \quad (6.13)$$

where p defines the temperatures set size T .

On the other hand, the word length of the stored values will depend on the cardinal of set U , where $U \subseteq C$ is the subset of different voltages in equation (6.12).

$$word_{length} = \lceil \log_2(\#U) \rceil \quad (6.14)$$

The cardinal $\#U$ also determines the voltage references module size. For example, in the case of the hardware implementation of the required voltages shown in Figure 6.11, ten voltage levels are sufficient to correctly write a value on both $1R$ and $1T1R$ at any temperature in the industrial thermal operation range. Finally, the mux, demux and decoder sizes are specified by both $\#U$ and $\#T$. The sizes of these modules can be constrained limiting the voltage amplitudes resolution in exchange of an increment in the power consumption. This can be easily seen in Figure 6.3: if a desired voltage amplitude is suppressed, higher voltages can be found that meet the constraints imposed by the algorithm previously presented. Summarizing, several design variables can be freely chosen while taking into account the trade-off between area, temperature resolution, misread robustness and power consumption.

Different systems with dissimilar temperature resolutions have been designed and simulated. We extracted from these results that a temperature resolution of $2^\circ C$ is enough to mitigate all thermal-variation caused writing errors. As the considered temperature range goes from 200 K to 380 K, using 7 bits to decode the selected amplitude, a 128 word LUT, with 4 bits per word was implemented. The use of a compact SRAM implementation only requires $14 \times 63 \mu m^2$. Those

memories whose RRAM cells dependence on temperature requires from higher resolution, could easily increase the LUT size: for 1°C and 0.5°C resolutions, the LUT area would increase by $\times 2$ and $\times 4$ respectively.

Remaining Circuitry

The rest of the circuitry has been first written in *HDL* for its later synthesis using standard cells provided by the 40 nm technology.

The functioning of the interconnection network [ILVMO12] eases the scalability of the incorporated temperature sensors. As the information is time-multiplexed in the transmission, and the temperature computations are performed and buffered over fixed registers, the area cost of increasing the number of sensors only enlarges the word-length of some used registers. Thus, being 7 the bits used to encode the temperature, the cost of using $S = 2^R$ sensors would be translated into a $7 + R + 1$ registers word-length.

Furthermore, as the *temperature computation module* requires a division during the temperature calculus, defining S as a power of two allows us to avoid a costly $1/S$ divider and we can use instead a simple register shifter, saving both power and area. The synthesized *profile and temperature handler core* without its LUT consumes a total area of $36.63 \times 30.18 \mu\text{m}^2$.

6.4.2 Implementation of the Voltage References System

Figure 6.14 shows the full custom compact design -only $51.4\mu\text{m} \times 30.9\mu\text{m}$ - of the used bandgap voltage reference [GRALV⁺16]. This bandgap has already been characterized regarding temperature, voltage, radiation and parameter variations in Section 5.7.1. Thus, we can ensure the correctness of this module behavior in the considered temperature range. As introduced in the Section 6.3, due to the voltage divider regular design together with the n-transistor proximity, the voltage references system is hardened against process variations. Figure 6.15 represents the accomplished Monte Carlo simulations, considering PVT. For the experiments shown in next section the following voltage levels supplied the *voltage reference system*: $V_{\text{REG}} = 3.3\text{V}$, $V_{\text{SS_BG}} = 2.1\text{V}$, $V_{\text{SS_DIV}} = 1\text{V}$. Using the above supply signals, all the 17 voltage levels required to cover both $1R$ and $1T1R$ cell configurations, from 2.6V to 1V , were correctly generated.

Figure 6.14: Implemented 40 nm bandgap voltage reference.

Figure 6.15: Voltage reference system response under variability. Four outputs of the voltage reference system, showing high reliability against PVT. 1000 Monte Carlo runs.

Table 6.1: CF cylinder sizes relation with device's resistance, using a read voltage of 0.1 V. Relative sizes compared against widest (w_{MAX}) cylinder and larger gap (x_{MAX}) [LJH⁺ 15], being 3.65 M Ω and 5 K Ω the HRS and LRS model boundaries.

Profile	CF Relative Size	HRS	LRS	Ratio
P_{1X}	70%	390 K Ω	10.26 K Ω	38
P_{2X}	80%	822 K Ω	7.86 K Ω	104.6
P_{3X}	90%	1.72 M Ω	6.21 K Ω	276.9

6.5 Temperature Aware, Reliable Oriented Profiles Simulation and Results

As introduced before, several works have thoroughly studied writing profiles focused on durability or power consumption. In this work we will analyze the power consumption and performance impact of nine independent profiles, resulting from the combination of three distinct pairs of $R_{ON} - R_{OFF}$ ratios and three different protection levels against temperature variation. Supposing each $R_{ON} - R_{OFF}$ ratio has a characteristic durability [MCYC15b, MCYC15a, MCYC16], the user gains the ability of dynamically selecting the desired hardening level against temperature variations while choosing a more suitable configuration regarding the power consumption/endurance trade-off [MCYC15b, MCYC15a, MCYC16]. For each considered profile P_{ij} , i refers the HRS-LRS resistance pair and j the level of protection against thermal variations.

Section 6.2 described how the compact model emulates the device CF evolution using a conductive filament characterized by both magnitudes x and w . Table 6.1 shows the different HRS-LRS resistance levels used in this experiment and their relationship with individual combinations of the cylinder sizes.

We use the metrics s_{Pxy} and r_{Pxy} to measure the quality of the simulated operation. Using the computed sizes $-x_{final}, w_{final}-$ of x and w CF magnitudes, and naming x_{Pxy}, w_{Pxy} as the CF sizes at which for a specific profile we consider the operation as correctly accomplished, the metrics can be defined as follows:

$$s_{Pxy} = \frac{\max(0, w_{final} - w_{Pxy})}{w_{Pxy}} \times 100(\%) \quad (6.15)$$

$$r_{Pxy} = \frac{\max(0, x_{final} - x_{Pxy})}{x_{Pxy}} \times 100(\%) \quad (6.16)$$

In those operations where s_{Pxy} or r_{Pxy} values are greater than zero the writing has not been correctly accomplished, and therefore a *writing error* arises. s_{Pxy} or r_{Pxy} values near 100% mean that the writing operation has not been able to store the desired value.

Table 6.2: *Different profiles and maximum temperature deviation they can assume and mitigate.*

Profile	MAX temperature deviation Δ_j
P_{X1}	$0^\circ C$
P_{X2}	$5^\circ C$
P_{X3}	$10^\circ C$

As introduced in Section 6.4, the sensors and temperature evaluation scheme were sufficient to provide reliable temperature estimates and to mitigate all errors withing the $2^\circ C$ resolution. However, embedded critical systems may require to decrease its working frequency -and therefore the temperature sampling frequency- which would lead to an increment of the computed temperature error. Additionally, due to aging effects the temperature sensor could increase its error. To address those scenarios, we consider profiles with additional protection against sensor malfunctioning errors or sudden and unexpected local temperature changes, which may cause errors not able to be mitigated. To that end we define three distinct P_{Xj} profiles with a different protection level against temperature variations. Therefore, as shown in Table 6.2, using P_{Xj} the driver will mitigate all possible malfunctions caused by errors in the temperature estimation Δ_j .

Resistive switching speed is one of the concerns to be dealt with during design stages. In the following experiments we fix the writing operation to have a maximum length of $1 \mu s$. To obtain the optimal voltage amplitudes that guarantee the correct operation in each considered temperature, we followed a specific RRAM cell characterization methodology. First, we simulated the updated RRAM model ($1R$ and $1T1R$ cell configurations) in Cadence Spectre simulator, using a $40 nm$ commercial technology for CMOS components, and sweeping both circuit temperature and writing amplitude parameters. Given the capabilities extracted from the commercial CMOS *process design kit* models, for $1T1R$ cells we considered a temperature range starting at $220 K$. With those results we built 2-dimensional arrays containing the resulting values of x and w CF magnitudes.

We filter, using the previously defined metrics s_{Pxy} and r_{Pxy} , the circuit temperature-voltage amplitude pairs at which the writing operation is not correctly performed. Finally, we obtained the optimum temperature-voltage values matrices M_{ij} related to each profile P_{ij} after applying the voltage selector algorithm (6.12).

6.5.1 Required Voltages and Power Consumption Results

Tables 6.1 and 6.2 gather the characteristics of each of the nine characterized profiles. Figure 6.16 summarizes the voltages and consumed energies per operation, for both $1R$ and $1T1R$ cells depending on the configurable profile used by the driver. Higher HRS/LRS ratio profiles, related with data retention oriented strategies, use larger amplitudes and therefore consume more power than low HRS/LRS -and usually higher device endurance- profiles [MCYC16].

As expected, lower temperatures require higher voltages and energies. It can be noticed how in $1T1R$ schemes, the in-series transistor modulates the current through the memristor, while diminishes the effective applied voltage. This effect raises two immediate consequences: First, for the same applied voltage $1T1R$ cells are slower than $1R$. Second, slightly higher operation voltages are required.

For SET processes, the use of discrete voltage amplitudes and pulse lengths creates an increment in power consumption, producing the saw-tooth like curves, that can be compared against the most energy efficient curve. All required voltages can be coded using 4 *bits* as the profiles consuming the larger number of different voltages, P_{3X} during SET process in $1T1R$ cells, only required 11 levels. Therefore, the LUT area estimation described in Section 6.4 is enough to store all considered profiles. In the same way, the temperature-voltage amplitudes relationship extracted from the initial study of Section 6.2 is presented for every simulated profile. For each profile, an almost decreasing linear staircase approximation can be used to model this voltage-temperature dependency. However, it must be remarked that, as the simulated temperature gets closer to the CMOS *process design kit* limit, this approximation can no longer be used. As shown in the $1T1R$ cell SET process, lower temperatures exhibit an exponential shape.

Figure 6.16: Required voltages and consumed energies for every considered profile at each working temperature, for both 1R and 1T1R cells.

Figure 6.17: SET operation comparison between the proposed approach (in blue) and the traditional one (in red) at $T_0 = 200K$. Traditional schemes would require from larger or consecutive writing pulses.

6.5.2 Reliability and Speed-Up Results

Until [GRLV15], drivers presented in the literature varied the *SET/RESET* processes duration to correctly achieve the desired operation in presence of PVT variations using two distinct schemes. A first one enlarges/decreases the writing pulse until the operation is correctly accomplished [JJMK10, GLmG⁺15]. A second scheme consecutively performs write-read operations until the device state is satisfactorily switched [AGHS12], approach that can be complemented with incremental amplitudes [KYS⁺16]. Both protection mechanisms dramatically reduce the system performance in presence of high temperature variations. To correctly perform the switching at temperatures lower than the nominal (temperature at which the driver is calibrated, usually 300 K), both first and second variable pulse-length approaches would require from larger operation lengths. The transient displayed in Figure 6.17 compares a *SET* operation using the proposed approach and the traditional one.

We have computed the inaccuracy of using a scheme considering fixed voltage amplitudes, no matter the working temperature conditions, together with a maximum writing pulse of t_{op} . In the same manner, we have simulated the time required to perform *SET/RESET* processes for each P_{iX} dynamic profiles, comparing the operation speed when using both the proposed scheme and [JJMK10]. The first step in the procedure to simulate the writing approach [JJMK10] corresponds to the calibration of *SET* and *RESET* writing amplitudes. Using the fixed 1 μs pulse length we calculated the minimum amplitudes that completely accomplish the desired operation, for each considered profile. Those amplitudes, captured in Table 6.3, were used to compute the operation inaccuracy under fixed voltage

Table 6.3: Calibrated voltage amplitudes required for each profile.

Cell Type	Operation	Required Amplitude		
		P_{1X}	P_{2X}	P_{3X}
1R	SET	1.25	1.25	1.3
	RESET	1.85	2.1	2.3
1T1R	SET	1.3	1.3	1.3
	RESET	1.9	2.1	2.35

Figure 6.18: Operation error dependence on temperature. While the scheme using the proposed driver is able to correctly behave under low temperatures, the classic fixed-voltage amplitude driver produces operation errors.

pulse length and amplitudes, as well as the pulse length required to perform the device switching at each considered temperature, for the different profiles.

Figure 6.19: Achieved speed-ups comparing the proposed scheme against variable-pulse-length schemes presented in literature under low-temperature conditions.

Fixed-Amplitude Scheme Error Results

To characterize how inaccurate a fixed voltage amplitude scheme can behave working under a fixed writing time, we simulated both 1R and 1T1R memories using $1 \mu\text{s}$ as the limiting operation length. Figure 6.18 shows the errors obtained after applying s_{Pxy} and r_{Pxy} metrics to the simulation results. It can be extracted that low temperatures produce a dramatic effect undermining the fixed-amplitude driver efficacy.

Speed-Up Results

We compare the minimum time required to correctly achieve a writing operation using a fixed-voltage amplitude scheme with variable pulse length against our fixed $1 \mu\text{s}$, dynamic amplitude driver. To this purpose we computed the pulse length that satisfies both s_{Pxy} and r_{Pxy} metrics, ensuring 0% error at each considered temperature.

In our proposed architecture no matter the working temperature our driver guarantees the system $1 \mu\text{s}$ operation, while works like [JMK10, AGHS12, KYS⁺16]

would enlarge the operation length, or even repeat it until the cell achieves the desired state.

Given the cell dependence on the circuit working temperatures, for traditional drivers we find two different temperature sub-ranges, divided by the *room temperature* at which the driver writing voltages have been set.

For temperatures over room temperature, in schemes like [JMK10], even though the device switching may occur earlier choking the cell voltage feeding, the writing operation timing would be guarded. Consequently, no timing difference would exist compared against our driver.

However, significant speed-ups -ratio between required time using the cited literature approaches and our proposed one- appear at lower temperatures. As shown in Figure 6.19, speed-ups near to $700\times$ -1R- and $70\times$ -1T1R- have been accomplished depending on the used profile, showing one the great advantages of using a variable amplitude scheme.

6.6 Conclusions

Reliability, power consumption and device durability are the most pressing issues to be addressed to reinforce RRAM as the leading technology of future storage and novel computing paradigms. In particular the close relationship between writing process characteristics and RRAM features impose hard trade-offs during system design process. On the other hand, the variations of circuit temperature strongly affect RRAM behavior, catalyzing the switching processes and, therefore, the memory block characteristics.

In this Chapter we have made use of the methodologies and tools that support the characterization of RRAM devices and RRAM based circuits presented in the previous Chapters. Working at system abstraction level, we have presented an innovative adaptive writing architecture for RRAM memories, able to correct any writing error that may appear during the process due to temperature variations. We also provide the driver with useful characteristics, introducing reconfiguration capabilities, therefore allowing different writing approaches to be dynamically applied. By using the presented solution and depending on the working scenario, different profiles oriented towards a specific application goal —such as power saving, enhanced endurance or higher reliability— can be employed. Consequently, more efficient and optimized configurations can be utilized in critical applications such as aerospace circuits, while being able to ensure the correct behavior in a wide range of temperatures.

Nine different profiles, varying the cell HRS/LRS ratio and the level of protection against temperature sensing errors have been analyzed, characterizing their power consumption and required voltage amplitudes. Additionally, we have compared the proposed approach against other drivers able to handle RRAM thermally-caused miss-behaviors, obtaining notable writing operation speed-ups near to 700× in certain working temperatures.

6.7 Related Contributions

- **Reconfigurable Writing Architecture for Reliable RRAM Operation in Wide Temperature Ranges.**

Fernando García-Redondo; Pablo Royer; Marisa López-Vallejo; Hernán Aparicio; Pablo Ituero; Carlos A. López-Barrio.

IEEE Transactions on Very Large Scale Integration (VLSI) Systems. Volume: 25, Issue: 4, Pages 1224 - 1235. April 2017. DOI: 10.1109/TVLSI.2016.2634083.

- **CAS-T Lecture: Reconfigurable Writing Architecture for Reliable RRAM Operation in Wide Temperature Ranges.**

Fernando García-Redondo; Pablo Royer; Marisa López-Vallejo; Hernán Aparicio; Pablo Ituero; Carlos A. López-Barrio.

IEEE international Symposium on Circuits and Systems , (ISCAS 2017), Baltimore, MD, Usa, May 28-31 2017.

- **A Thermal Adaptive Scheme for Reliable Write Operation on RRAM Based Architectures.**

Fernando García-Redondo, Marisa López-Vallejo, Pablo Ituero.

Conference: 2015 33rd IEEE International Conference on Computer Design (ICCD), October 2015, New York.

- **Temperature Sensor Placement Including Routing Overhead and Sampling Inaccuracies.** *Pablo Ituero, Fernando García, and Marisa López-Vallejo.*

International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design 2012 (SMACD2012) Seville, Spain. 19 - 21 September 2012.

*Everyone, I Have A Very Dramatic Announcement. So Anyone
With A Weak Heart Should Leave Now. Goodbye.*

Professor, Futurama.

7

Auto-Erasable RRAM Memory Architecture Secured Against Physical and Firmware Attacks

7.1 Introduction

Chapter 6 presented an architecture to ensure that every write operation was correctly accomplished, no matter the circuit working temperature. Additionally, it also provided the ability to reconfigure the writing scheme to better meet the application requirements. Summarizing, we improved and made reliable an already existing, traditionally CMOS application, when ported to RRAM technologies. By contrast, as represented by Figure 7.1, this chapter takes advantage of RRAM technologies in order to deploy a new system level application.

Every year new attacks and secure blocks-countermeasure systems appear. In particular, and leaving software out of consideration, to defend memories from unauthorized physical access and hardware interfaces working under malicious

Figure 7.1: Chapter 7 scope and perspective within the Thesis.

firmware has proven extremely challenging [Wei14]. Nonvolatile memories, unlike standard CMOS SRAMs, retain the stored information whenever a system power-off occurs. This characteristic, very attractive to standard storing applications, brings additional challenges, in order to avoid sensitive information from being unauthorizedly accessed. In that sense, big efforts are being made to encrypt and protect valuable stored data.

However, even encrypted memories may be compromised if either secure keys are directly accessed or seized via *side channel attacks* (SCAs). Then, it is of critical importance that the next generation of large storage memories is able to protect the contained critical information from being hacked [KKSK15]. The transactions accomplished within the memory leave power traces that could be used to extract keys from the cryptosystems. To that end obfuscation schemes are used to counter this kind of attacks [KKR15].

Moreover, RRAM based cache memories may contain secure keys granting access to other memories [KKR15], and it has been proven that, under an unexpected poweroff, encryption based solutions give time windows wide enough to access sensitive data [CS11]. To prevent this kind of low-level attacks commercial solutions implement secure schemes for CMOS based nonvolatile memories. Should any firmware-level attack be detected, these protocols either delete within few clock cycles special memory blocks, or obscure their read using fuse layers [Mic]. As an example, commercial mobile devices designed to accomplished monetary transactions [Ver17], such as credit card payments, include specific hardware providing erasing routines, to secure the most sensible memory blocks containing system, client and server keys.

However, when accessed physically, flash memories are still susceptible of being de-caped for the acquisition of images reveling the digital states of each cell present in the memory. Regarding RRAM technologies, in-situ TEM images may reveal the CF that denotes the resistive state of the RRAM cell [HKJ⁺15, LWP⁺15].

Thus, by de-caping the memory, attackers have access to the stored sensitive information.

As seen, previous works focused on the definition of encryption mechanisms [KSK15] or obfuscation schemes [KKR15, Mic] to hinder the access to stored data. However, it is to be noticed that in many mobile commercial solutions such obfuscation schemes, either based on erasing, encrypting, or both, require external power supply batteries to accomplish the securing operation, in order to correctly finish the process should any attack deprive of the main power source.

In this Chapter we go even further, presenting a novel architecture that secures the information of sensitive memory blocks. If either a firmware-level or physical-level attack is detected, instead of erasing one by one the memory cells, our proposal erases the whole memory almost instantly by simultaneously exposing the array to an erasing pulse. Even more important, if under any circumstances the circuit is deprived of the power supply, an optional emergency power supply embedded within the system allows the erasing operation to still take place. Therefore, moving the external battery to a power source embedded within the chip, the physical attack to the supply lines gets even more complicated. Thus, we provide highly sensitive applications with the hardware required to ensure that their most critical data remain safe from being unauthorizedly accessed.

The Chapter is structured as follows: First we describe the architecture structure and characteristics, focusing on the erasing procedure and detailing the functioning of the different modules. After that, we introduce the effects caused by temperature and its influence on the securing system. The characteristics of the memristor model emulating the fast switching devices to be included in the system are outlined in Section 7.4. Later, we will present the analyses accomplished to design and characterize the architecture behavior depending on the system size and environmental characteristics. Finally, some conclusions will be drawn.

7.2 Information Securing Architecture

LRS and HRS states depend on whether the CF of the cell is fully deployed or partially/fully destroyed. Based on this premise, two methods give access to the stored information. First, we could read the cell resistance sensing the flowing current after applying a read voltage. Second, accessing the CF state by TEM imaging would provide us the vision of a deployed or fully/partially destroyed CF. To forbid the access to the contained information, we can grow or dissolve every CF, therefore erasing the memory, so each cell is indistinguishable from others.

The proposed architecture comprehends the modules and circuitry required to erase, within a single, short in time operation, the sensitive information stored in an RRAM memory block. We base our proposal on the ability of different

Figure 7.2: Proposed information securing architecture. The switch S_{ER} , closes if the signal sys_OK changes from V_{DD} to 0, erasing the RRAM memory. This event occurs whether a monitor—Temperature monitor, Authentication unit or Voltages monitor—raises a problem. If the pin V_{ERASE} is disconnected—opening S_V_ERASE and lowering v_erase_ok — sys_OK would also change to 0, and therefore the charge stored in the CMOS capacitor C_{EPS} would also erase the memory.

RRAM devices to switch from LRS to HRS and vice versa in short periods of time [YGG⁺12, Pai13]. We present the additional modules, environmental sensors and driver circuitry required to switch every cell present in the memory to either LRS or HRS in a single operation. Therefore, thanks to this procedure the attackers would be unable to retrieve the data previously stored in the nonvolatile memory.

The architecture is composed of three main blocks: a *Sensitive RRAM Memory* to be secured, together with an *Erasing control* and the so called *Emergency erase power supply* modules. The *Erasing control* is in turn composed of the *Voltages monitor*, the *Temperature monitor* and the *Authentication unit*. *Erasing control* and *Emergency erase power supply* subsystems are respectively managing and performing the erasing procedure. As shown in Figure 7.2, interfacing these modules, but out of the scope of this chapter, we find the *Encryption Unit*. There are different research lines providing solutions for RRAM memory encryption [KSK15]. These works, defining the memory encryption and decryption processes, complement the presented architecture. From now on we will focus on the automated erasing procedure.

Figure 7.3: Circuitry driving the erasing pulses to the memory crossbar.

7.2.1 Erasing Circuitry

During the erasing procedure, the erasing voltage V_{ERASE} is driven to the memory through specific circuitry. Figure 7.3 describes the design of the proposed erasing driver. Designed to not interfere with the standard functioning row and column selectors, the driver short-circuits every row line to V_{ERASE} , and every column line to gnd, thus starting the erasing procedure. To achieve this, a set of transistors controlled by the sys_OK signal interconnect the different lines.

7.2.2 Erasing Procedure

Traditionally, to erase an RRAM crossbar—drive every cell to a known state, either LRS or HRS—the cells are sequentially accessed. On the contrary, our approach simultaneously erases the whole memory, gaining critical time otherwise used by the attackers to access the information remaining from being erased.

The architecture is based on the high switching speeds and threshold based behavior found in RRAM devices [YGG⁺12]. If every CF of the cells is grown or destroyed in a short time span, any later attempt to access the encrypted-decrypted information either by firmware or manual cell stimulation would be worthless. Even de-cap & TEM image processes would not work: in cases where residues of partially destroyed CF remain within the structures, such residues would be equivalent in size among cells, not being able to distinguish whether the cell previously stored a 0 or a 1.

Figure 7.4: Equivalent circuit for a crossbar RRAM array during erase process. For each cell RT_{i_j} , RB_{i_j} and Cap_{i_j} refer the line/cell parasitic resistances and capacitance. The current i_{erase} is composed of partial currents, each weighted by the different equivalent impedances Zeq_{i_j} .

To accomplish this goal, the *Erasing control* module first detects the physical or firmware attack attempt over the sensitive memory, and then the architecture erasing modules proceed to generate the voltage pulses required to erase the memory. To do so, two different erasing schemes could, theoretically, be used:

1. Switch every cell from their previous state to LRS —HRS to LRS— triggering the different *SET* processes, when required.
2. Switch every cell from their previous state to HRS —LRS to HRS— triggering the different *RESET* processes, when required.

To accomplish the first scheme, every cell must be fed with positive voltage differences between their terminals. To fulfill the second scheme, negative voltages are to be driven.

To erase the memory, our architecture short-circuits every row line to the erasing voltage and every column line to ground, and then it would use either (1) or (2) voltage schemes. Figure 7.4 describes the equivalent memory circuit during this erasing stage. Here, every cell is fed by the same erasing voltage, and the current flowing through each branch i_j depends on the resistance value Zeq_{i_j} .

Comparison Between HRS to LRS and LRS to HRS Schemes

Usually, a single *SET* process is faster and consumes less energy than a *RESET* process: First, a *SET* process involves lower power consumptions, as the instant power inversely depends on HRS until the resistance is switched achieving LRS. On the contrary, the instant power during *RESET* processes inversely depends

Figure 7.5: Current divisor effects occurring in an RRAM memory when connected to the proposed erasing architecture.

on LRS, therefore involving higher currents. Second, in most devices the *SET* threshold voltage is lower than the *RESET* one.

However, as our architecture short-circuits every cell, the total power consumption is the sum of individual cell contributions. That means that, as we are going to erase a full memory filled with both cells in LRS and HRS, we cannot suppose any initial state.

Following the first scheme, erasing every cell towards LRS, a positive voltage difference is applied to every cell, no matter its state, during the whole process. Consequently, those cells at HRS would consume low energy during their switch. But on the contrary cells already at LRS, not needing to be switched, will be consuming high amounts of power. Moreover, let's consider the memory shown in Figure 7.5, containing a random pattern of HRS and LRS cells. When the erasing operation starts, LRS cells drive most of the erasing current. Inversely, HRS cells are given negligible currents, and therefore, no switching will occur. Being most of the cells in a LRS state, the memory equivalent resistance would be extremely low, and most of the erasing current would be driven directly towards ground. Thus the erasing operation performing *SETs*, leading every cell towards their LRS, is discarded.

On the other hand, when using the second erasing scheme the negative voltage v_{erase} produces *RESET* transitions in every RRAM cell, changing those in LRS to HRS. As the cell array is equivalent to parallel resistances, the current (and consequently the energy) supplied to each cell depends on its resistance value.

Figure 7.6: Erasing process. Cells in LRS are progressively switched toward HRS.

Thus, those cells with lower resistances would drive higher currents. On the contrary, those cells already at HRS, not needing to be switched, will contribute to the overall power consumption with short power quantities, as they produce lower currents than LRS cells. This translates into a compensated switching process, leading every cell to converge towards the same resistance, completely destroying any CF and therefore erasing any previous information.

Figure 7.6 represents the process. In the example, $Z_{eq_a_b} < Z_{eq_a_b+1}$, so acting as a current divider $i_{a_b} > i_{a_b+1}$. Progressively, this higher current contribution leads the LRS to switch faster towards its HRS. Consequently, at the end of the process, every cell resistance converges towards the same value.

Summarizing, the switching speed and power consumed will depend on the used scheme and the characteristics and the state of the different cells.

7.2.3 Erasing Control

The *Erasing Control* module manages the erasing procedure by using the control signal `sys_OK`.

During the circuit normally functioning, `sys_OK` signal is at high level. This value opens the switch `S_ER`, therefore isolating the RRAM memory from the erasing power sources. If any undesired error is detected, the erasing voltage is driven to the RRAM block by the same switch. Consequently, the erasing process only takes place if any unauthorized operation is attempted, or an unpredicted and non-standard behavior appears, a condition represented by a low level in signal `sys_OK`.

Three different submodules—the *Voltages monitor*, the *Temperature monitor* and the *Authentication unit*—control the authorized access and monitor the working conditions of the memory. Therefore, the control signal `sys_OK` is the product of different status signals referring the correct functioning of the system:

- Signal `vDD_OK` is the output of two ad-hoc power supply monitors [ALLV15]. Their functioning, described in Section 5.7.2, is in charge of sensing and reporting unexpected changes in voltage lines. The *Voltages monitor* module handles both instances, and produces the signal `vDD_OK` to capture the correct status of both `V_ERASE` or `V_DD` signals. This module drops `vDD_OK` to low level if either `V_ERASE` or `V_DD` lines experiment a disconnection or wide amplitude variations which may be related to physical attacks. To avoid situations where a false alarm may cause an erasing, this monitor neglects glitches, focusing on pronounced and sustained voltage decrements.
- Most complex physical attacks require the circuit temperature to be reduced to minimum levels in order to prevent — or slow down— secure systems to encrypt and erase sensitive information. Signal `temp_OK`, generated by the *Temperature monitor*, goes to low level if the sensed circuit temperature either reaches a critical value T_{crit} or decreases with an unexpected ratio T'_{crit} . Combining temperature thresholds together with the analysis of temperature evolution avoids raising false alarms.
- `Auth Unit`, controlling the signal `auth_OK`, it is responsible for correctly authenticating whether the memory is securely accessed or not. To that end it verifies an external key against a unique private key, ensuring that the firmware-operation is correctly signed. Otherwise, if an unauthorized access is attempted, `auth_OK` goes to low level triggering the memory erasing procedure.

7.2.4 Optional Emergency Erase Power Supply

Unexpected power-offs or metal de-cap actions would disconnect the memory from the voltage supply lines. Usually, critical applications include additional batteries, external to the processors, that provide the power required to perform securing routines should any unpredicted behavior happen. Point of sale (POS) devices and Mobile POS systems are a clear example of this kind of applications. These devices create secure connections to accomplish secure transactions. Usually, POS systems include such auxiliary batteries to supply emergency power until the information inside the device's memory is correctly secured. However, if auxiliary batteries are also attacked, the system could be compromised [KKSK15].

We propose an additional countermeasure based on CMOS capacitors or *Metal/Oxide/Metal* (MOM) super-capacitors, that can also be achieved through RRAM growing procedures [VGC⁺16]. It should be noticed that large capacitances require large areas. However, if instead of using CMOS layers the capacitor is built employing MOM capacitors the area gets reduced. Compared against the

Figure 7.7: Example of functioning under physical attack. The CMOS capacitor discharges through the memory cells, erasing them.

external battery approach, being within the chip package as a part of the circuit, our proposal hinders the attack, as the emergency power supply would erase the memory before it can be disconnected.

These capacitors, built next or on top of RRAM crossbar layers, can provide the energy needed to perform an emergency erasing operation. Despite the area cost, these capacitors have sufficient charge to completely erase the memory in a short time period. Therefore, during normal functioning, these capacitors are charged. However, when a physical attack happens these capacitors discharge, erasing the sensitive memory block. Figure 7.7 represents how the Cap_EPS CMOS capacitor charges, and later supplies the charge required to *RESET* the cells during an attack, without requiring external power. Here we can see how if the main line that feeds V_{ERASE} is cut, the erasing voltage can be provided using the energy stored within the capacitor.

Behaving like a stand-alone power source, the charge stored in the capacitor is limited. Therefore, the implementation of the erasing architecture must operate consuming low energy. As described in Section 7.2.2, probably the memory to be erased would contain numerous cells in LRS state, situation that may lead to an incomplete erasing. But considering the capacitor as the erasing power source, this problem would be magnified. Consider the feeding of a high number of parallel memristors in LRS while the rest of the memory switches from HRS to LRS: in this scenario the erasing operation would consume a great amount of energy, which would drain the energy stored within the capacitor. Moreover, the capacitor discharging time constant τ depends on the memory equivalent resistance. Thus, switching every cell towards LRS would discharge the capacitor faster than

switching towards HRS without achieving the erasing operation. Consequently, such erasing scheme is again discarded.

To sum up, the *Emergency Erase Power Supply* subsystem stands as an optional module providing the system with an additional security layer, always recalling the area price we pay implementing the CMOS/MOM capacitors. If this optional subsystem is to be used, even in situations where the erasing circuitry is disconnected from the main power source, the architecture erases the memory resetting every cell to their HRS state.

7.3 Temperature Influence

A drastic circuit cooling during an attack may degrade the system securing mechanisms. Therefore, the performance of the proposed architecture must be studied regarding temperature variations.

Chapters 2, 5 and 6 described how the CF forming and dissolving processes that cause the abrupt resistance change in RRAMs are mainly controlled by the applied electric field and the local thermal behavior [Iel11, Iel16]. The high temperatures reached within the device catalyze both oxide vacancies and metal ions migrations that control the CF geometry evolution. Formed and dissolved CFs relate to LRS and to HRS states respectively. Summarizing, these highly non-linear dynamics produce abrupt resistive switchings depending on the RRAM technology characteristics, voltage inputs and working environment. In particular low circuit temperatures slow down RRAM switching. Then, it was necessary that the used model accurately emulates the effects caused by temperature [CY15], also considering the influence the circuit working temperature has on the CF evolution. For further reading, thorough studies on the conduction properties and voltage threshold evolution dependence on temperature for HfO_2 devices can be found in [VGC⁺16, PWG⁺17].

On the other hand, and as seen in Chapter 5, temperature is a critical factor that determines the performance of CMOS devices [Wes10]. At higher temperatures, the performance of CMOS transistors is worse than at lower temperatures. If we simplify the transistor as a simple switch, both sub-threshold –during *OFF*– and leakage currents increase with temperature. On the other hand, during *ON* stages the carrier mobility reduces with temperature. Therefore, the *ON* currents, both in linear and saturation regions, decrement when temperature rises.

Summarizing, temperature may produce dramatic effects on the performance of the proposed architecture. Making use of the design strategy and framework presented in Chapter 5, in the following section we will analyze the implementation details and present the architecture simulation results.

Figure 7.8: Reproduced from [Pai13]. 1T – 1R Fast switching behavior of the used compact model. V_{WL} , V_{BL} and V_{SL} refers the voltages at word, bit and selection lines.

7.4 Fast Resistive Switching Modeling

For the experiments accomplished in this chapter we employ the physics-based RRAM Verilog-A model developed by Arizona State University researchers [CY15, YU15, JWY⁺16], previously used in Chapter 5 RRAM read circuit characterization. Two reasons backup this decision. First, as we have previously said, the model correctly emulates thermal dynamics. Second, this compact model accurately fits the behavior of HfO₂ devices exhibiting extremely fast switching speeds [YGG⁺12]. Figure 7.2, reproduced from [Pai13], illustrates this fast switching response for a 1T – 1R cell. We will make use of the fast switching processes of these devices to accomplish the emergency memory erasing. In order to recall its features, we now summarize its main characteristics. Validated against physical devices, this model describes the resistive switching effect using a simplified 1– dimensional conductive filament. The CF dissolution and growth gets reflected on the variation of gap distance between the CF’s end and the opposite electrode.

Consequently, we will base our following experiments on an RRAM model fitting the behavior of physical HfO₂ devices exhibiting fast state transitions.

7.5 Implementation and Results

Lower voltages require larger pulses to trigger *SET* and *RESET* processes. However, even though larger voltages are preferred, a hard constraint is that the erasing voltage should be correctly handled by the used CMOS technologies. Additionally, lower temperatures lead to slower switchings [Iel11, Iel16, CY15], while CMOS devices perform worse at higher temperatures. In the light of such temperature dependences, the architecture implementation will be characterized in the standard military temperature range from 220 K to 400 K.

On the other hand, regarding operating voltages, our proposed solution has been designed using a 40 nm commercial technology. As introduced in Chapter 6, the nominal voltage of the transistor models we use is 2.5 V. However, the technology PDK states that voltages up to 2.875 V (2.5 V + 15%) can be used inside the chip, always taking into account that raising the working voltage engraves aging effects and may reduce the transistor life. Erasing voltages up to 2.8 V will be considered. However, such voltages will be set aside just for the *Erasing Driver*. It is important to remember that such circuitry only would be used if an attack occur. Therefore, under standard functioning no degradation will affect the driver. On the other hand, the remaining modules and circuitry —*Erasing Control*— will be fed by standard 2.5 V and 1.8 V, guaranteeing their endurance throughout time.

Then, there is a trade-off between speed —critical in this kind of applications— the size of transistors present in the erasing driver, and the used voltage. The circuit temperature —influencing the CF heating and therefore the switching speed— array size and device characteristics determine the voltage to be used. Moreover, if the optional capacitor is to be included, we need to consider its size and area as two decisive parameters, cluttering the already complicated design parameters election. To that end, we will apply the considerations and simulation framework presented in Chapter 5, automating and supporting the whole process.

The device model we have used accurately fits fast-switching physical devices [CY15]. As previously seen in Section 7.4, the model normally operates between the 50 k Ω of the LRS (relative to a CF gap of 0.5 nm) and a 1 M Ω resistance in the HRS (relative to a CF gap of 1.37 nm). Therefore, as the standard set of parameters matches the required physical device behavior, in the experiment, we will consider those two values as the normal operating boundaries. In the same way, we will also adhere the model maximum and minimum CF elongations, relative to gap distance of $gap_{min} = 0.1$ nm and $gap_{max} = 1.7$ nm.

Additionally, we added the following parasitic components

$$RT_i_j = 10\Omega, RB_i_j = 10\Omega, Cap_i_j = 1fF$$

as the top and bottom line parasitic resistances and parallel capacitance, in order to properly describe large arrays behavior and RRAM cells current overshoot effects. In this Section we first study the influence of erasing voltage, array size and working temperature on the architecture performance and reliability. After that, we analyze the architecture functioning, finding the voltages required to erase any cell within 5 ns, then frustrating any attack almost immediately.

We proceeded to explore the wide design space that considers erasing voltage, circuit working temperature, sizes of memory and driver transistors. The used memristor model already considers the unreliability sources —temperature dependence— we would like to deal with in the experiments. Therefore, and following the presented strategy for reliable analyses, we defined the metric to be used to determine the degree of CF destruction. Given that the memristor model considers a 100% CF destruction when the gap reaches gap_{max} , we define this metric using the model's maximum $-gap_{max} = 1.7 \text{ nm}$, for HRS— and minimum $-gap_{min} = 0.1 \text{ nm}$:

$$operation_percent_reset(gap) = 100 \frac{gap - gap_{min}}{gap_{max} - gap_{min}}. \quad (7.1)$$

In every case the erasing operation time is strictly restricted to 5 ns, and therefore only scenarios achieving the 100% of the erasing operation in such short time are desired.

After that, we accomplished gross explorations of the design space in order to acquire knowledge about the circuit response depending on its parameter set. In particular, we first selected the CMOS capacitor value that was able to correctly erase the considered memories at low temperatures.

7.5.1 Architecture Standard Erasing Procedure

The architecture standard erasing procedure does not implement the optional capacitor as power supplier. On the contrary, it relies on the *Voltage Monitor* module to detect any voltage variation and trust the erasing operation to be accomplished before the power is fully drained.

Using this architecture, and unlike traditional erasing procedures in which the cells are independently and sequentially accessed, our erasing procedure does not leave floating cells. Therefore no matter how large the array is, sneak paths [KB16, YC16, EBC⁺15, JKK⁺16, LLN15] can be neglected.

However, as every cell is connected in parallel, the number of cells determines the equivalent resistance of the crossbar. Following a resistance divisor scheme, extremely low crossbar equivalent resistance values would lead the erasing driver transistors to drop the effective voltage applied to the cells. Therefore, the size of the crossbar does influence the driver size. Moreover, the drawback effects caused

LRS -> HRS scheme, 16x16 RRAM Block

Figure 7.9: Results describing the achieved erasing percentage during the erasing operation, varying the erasing voltage as well as the width of the driver transistors, for a 16×16 RRAM block.

by the parasitic components inherent to crossbar lines, cell electrodes, and cell spurious capacitances, do depend on the memory size.

To that end, we will analyze the worst-case scenario for the LRS to HRS erasing scheme, supplying the erasing voltage from the V_{ERASE} line, which is considered as a stable source. In such scheme, the worst-case relates a scenario where every cell is required to be altered towards the opposite state. The exploration of the design space leads to the following results. Figures 7.9, 7.10 and 7.11 describe the quality of the erasing operation when every cell within the memory block is driven towards their HRS state, therefore fully destroying their CF. It can be seen how both the erasing voltage and the width of the transistors are key design factors to be taken into account depending on the size of the memory that is to be erased. As expected, larger memories require wider transistors to correctly drive the erasing currents.

Figure 7.10: Results describing the achieved erasing percentage during the erasing operation, varying the erasing voltage as well as the width of the driver transistors, for a 32×32 RRAM block.

Figure 7.11: Results describing the achieved erasing percentage during the erasing operation, varying the erasing voltage as well as the width of the driver transistors, for a 64×64 RRAM block.

Table 7.1: *Percentage of erasing using the standard erasing scheme in a 64×64 RRAM block using $35 \mu\text{m}$ width transistors, under variability for wide temperature range. 500 Monte Carlo simulations were accomplished per temperature.*

Circuit Temperature	Operation Accomplishment %
220 K	100%
256 K	100%
292 K	100%
328 K	100%
364 K	100%
400 K	100%

Final Implementation and Temperature and Variability Dependence

We will now characterize the standard erasing scheme under temperature variations and variability, and use it over a 64×64 RRAM block, simulating the worst-case scenario to ensure the degree of reliability of the proposed approach.

Thanks to the information provided by Figure 7.11, we choose to implement the erasing driver using $35 \mu\text{m}$ width transistors together with 2.8 V erasing voltage. From the figure, we could extract that under nominal temperature and widths down to $32 \mu\text{m}$ at 2.8 V would be sufficient. However, temperature and variability deteriorate the circuit performance so we will increase the width to gain reliability. With these two set-ups we will now analyze the erasing driver dependency on temperature as well as on variability—both affecting CMOS and RRAM. Similarly to the variability modeling scheme employed in Section 5.8, to evaluate the system performance we emulated variability using the PDK information for CMOS and redefined the RRAM model via model wrappers. More specifically, we considered normal density probability functions for the initial gap, local field enhancement factor and the gap velocity parameters, all of them centered in their nominal values and varying following a $\sigma_i = 12.5\%\mu_i$.

We accomplished 500 Monte Carlo simulations for each temperature, obtaining the results shown in Table 7.1. It can be seen how for the implemented architecture, and using peak voltages of 2.8 V , we mitigate the effects caused by variability and temperature, hence correctly achieving the desired operation in 100% of the cases.

Complementing the table, Figure 7.12 shows the superposition of the different transients for the 500 simulations relative to 220 K circuit temperature. We can see how every erasing procedure was completely accomplished in less than 5 ns . It is to remark that higher temperatures catalyze the effects caused by variability. The variation in the effective voltage fed to each memristor cell is more pronounced, and these voltage differences lead to different switching speeds. On the contrary,

Figure 7.12: CF gap evolution of the slowest cell within the 64×64 memory during 500 Monte Carlo simulations at different circuit temperatures. Erasing operation starts at $t = 0$ ns.

the transistors behave better at lower temperatures. At $T_0 = 220$ K, almost every cell is fed the same effective voltage, and the different switching speeds are due to the RRAM cell variability.

7.5.2 Architecture's Optional Erasing Procedure

The present section extends the previous analyses, and characterizes the architecture performance and reliability when the optional emergency capacitor supplies the energy required to erase the memory. To that end, we will also consider the worst case scenario of the LRS to HRS erasing scheme.

If the emergency erasing power supply is to be used, the array size greatly influences the erasing reliability. For the sake of simplicity, let us suppose that the transistors present in the driver are wide enough so that their equivalent resistances can be neglected compared against the resistive memory block. Therefore, we can easily approximate the impedance seen by the capacitor. As the initial state of the worst-case scenario cells is LRS, the approximated initial resistance of the memory block is

$$Z_{init}^{-1} = \sum_{i=1}^M \sum_{j=1}^N \frac{1}{Z_{eq_i_j}}, \quad (7.2)$$

and supposing every cell in the same resistance —related to LRS— state

$$Z_{init} = \frac{Z_{eq_0_0}}{M \times N}. \quad (7.3)$$

The emergency capacitor will discharge following a τ time constant that depends on both the capacitance and the impedance shown by the capacitor. This impedance varies as the cells CFs dissolve, which makes τ vary with time, becoming $\tau(t)$. Therefore, the emergency erasing voltage supplied by the CMOS capacitor will initially discharge with the ratio

$$\tau(t)|_{t_0} = \tau_{init} = Z_{init} C. \quad (7.4)$$

Consequently, the larger the array, the faster the erasing voltage drops, maybe not supplying enough current to completely erase the whole memory. This discharge constant varies through time as the cells switch towards their HRS. In this way, if every cell is completely erased, the final discharge time constant is equivalent to

$$\tau(t)|_{t_{end}} = \tau_{end} = Z_{end} C, \quad (7.5)$$

where Z_{end} refers the memory equivalent impedance when every cell has its CF destroyed.

Figures 7.13, 7.14 and 7.15 describe the reliability results for a 16×16 , 32×32 and 64×64 memory blocks, erased using the emergency capacitor as power source, pre-charged using 2.8 V. It is to be noticed that the capacitance required by the erasing operation strongly depends on the memory size, a dependency whose effects dramatically influence the design parameters on the larger memories — Figure 7.15. Focusing on the 64×64 RRAM erasing results, we can also highlight

Figure 7.13: Results describing the achieved erasing percentage during the erasing operation, varying the emergency capacitance as well as the width of the driver transistors, for a 16×16 RRAM block.

an optimal design space area, relative to the plateau located in the transistor width range of $80 - 120 \mu\text{m}$. Larger values would involve higher gate capacitances, worsening the overall performance.

In order to illustrate the capacitor discharging dependence on the crossbar state, Figures 7.16 and 7.17 describe the transient evolution of the gap in the CF, and the related capacitor discharging —voltage fed by the emergency capacitor— for different capacitances and memory sizes.

Figure 7.14: Results describing the achieved erasing percentage during the erasing operation, varying the emergency capacitance as well as the width of the driver transistors, for a 32×32 RRAM block.

Figure 7.15: Results describing the achieved erasing percentage during the erasing operation, varying the emergency capacitance as well as the width of the driver transistors, for a 64×64 RRAM block.

Figure 7.16: Different discharging ratios for different capacitances. Erasing operation starts at $t = 2$ ns.

Figure 7.17: Different discharging ratios for different memory sizes. Erasing operation starts at $t = 2$ ns.

Table 7.2: *Percentage of erasing using the emergency erasing scheme in a 64×64 RRAM block using $100 \mu\text{m}$ width transistors, under variability for wide temperature range. 500 Monte Carlo simulations were accomplished per temperature.*

Circuit Temperature	Operation Accomplishment %
220K	100%
256K	100%
292K	100%
328K	100%
364K	100%
400K	100%

Final Implementation and Temperature and Variability Dependence

Following the methodology described for the architecture with standard erasing scheme, we will use a 64×64 RRAM block to characterize the performance of the architecture, simulating the worst-case scenario to ensure the degree of reliability of the proposed approach. To define the size of the final circuit to be implemented, we extract the transistors width and capacitor sizes from Figure 7.15. In particular, we select these parameters from the blue plateau related to the design space area where the erasing percentage reaches its maximum. Therefore, we will use $100 \mu\text{m}$ transistors coupled with a 300 pF erasing capacitor. The capacitor has been designed using a metal-oxide-metal primitive within the commercial CMOS technology, whose total dimension reaches $520.54 \mu\text{m} \times 517.56 \mu\text{m}$. Even though this approach involves an important area consumption, it does not require the even bigger external battery, which is also more easily attacked.

We accomplished 500 MC simulations for each temperature, obtaining the results shown in Table 7.2. Using the above implementation — $100 \mu\text{m}$ transistors width coupled with a 300 pF erasing capacitor— and charging the capacitor with 2.8 V , the effects caused by variability and temperature are mitigated, therefore achieving the erasing operation in every case. Even the slower cells achieve their CF destruction. Complementing this graph, Figure 7.18 shows the superposition of the different transients for the 500 simulations relative to the simulated circuit temperatures. Again, we can notice how the effects caused by variability increase with temperature: the cloud of CF gap transients widens with respect to the nominal erasing procedure. However, as the used transistors are wider than those related to the standard erasing scheme, the effects caused by variability are less pronounced. If we compare the results shown in Figure 7.18 against those presented in Figure 7.12, we can notice how the wider transistors of the emergency approach drop less voltage, therefore feeding higher effective voltages

Figure 7.18: Variability effects on the evolution of the CF gap for the emergency power supply erasing scheme at different temperatures. The image displays the CF gap evolution of the slowest cell within the 64×64 memory during 500 MC simulations at different circuit temperatures. Erasing operation starts at $t = 0$ ns.

to the resistive cells and achieving faster operations. If we would like to increment the speed of the standard scheme, transistors wider than $35 \mu m$ would be needed.

7.6 Conclusions

In this Chapter we have presented an auto-erasable architecture for RRAM memories that obscures the stored information if physical or firmware attacks happen. The idea complements previously reported encryption approaches, and gives an additional security layer to prevent the access to the most sensitive stored information. An optional emergency power supply scheme based on on-chip capacitors, gives the option to include, at the cost of consuming a great portion of chip area, an additional power source securing the memory if any physical de-cap is attempted. This optional power supply scheme would be comparable to the external batteries used under emergency situations where standard power lines have been disconnected. We must highlight that our power supply, placed within the chip, is much less exposed than an external battery.

The particular architecture implementation uses a commercial 40 *nm* CMOS technology and an RRAM model calibrated against HfO_2 physical devices. We analyzed the reliability and performance characteristics when design and environmental parameters are varied. Additionally, we studied the sizing of the architecture against the worst-case scenario, providing a solution in a wide temperature range. Finally, we secured a full 64×64 memory suffering from variability, preventing accesses to the stored data no matter the working temperature using either standard erasing scheme or the optional one employing the emergency capacitor as power source.

Hooray! A Happy Ending For The Rich People.

Zoidberg, Futurama.

8

Concluding Remarks

RRAM stands under the spotlight, exactly right now when researchers are defining the road to be followed, when the design of future electronic systems changes its direction towards a new paradigm. The great potential of memristive technologies relies on their high density, low power consumption, multi-level capabilities, non-volatility, and easy integration within standard CMOS circuits, guaranteeing a viable manufacturing. Thus, memristors are strong candidates for the next generation of memories and processors for novel computer paradigms.

However, even though the road ahead is promising, it is also riddled with challenges. As many emerging technologies, RRAM manufacturing processes, models and applications are still immature. Error sources such as variability or radiation, that have been extensively and thoroughly studied for CMOS circuits, also produce important effects in RRAM devices, endangering the applications trustworthiness. Moreover, when considering memristors, new error sources arise: CF degradation, current overshoot, data alteration during read operations or sneak paths are some of the most concerning reliability issues to be considered. Therefore, the new generation of RRAM-CMOS hybrid circuits inherits the well known CMOS

but also new RRAM related reliability issues, that must be addressed and mitigated. Thus, to overcome these difficulties, the design of reliable applications requires from support tools and strategies. From the device modeling perspective these mechanisms are required to build accurate compact models for physical devices. In order to correctly characterize sub-systems at circuit level, both methodologies and simulation suites are needed if the error sources affecting RRAM are to be considered. Only making use of this set of tools and strategies will the designer achieve reliable applications. Therefore the methodologies and design decisions applied at system level must always take into consideration the analyses carried out at both circuit and device levels.

In this context, this thesis has addressed the modeling and simulation of RRAM components for the design of reliable circuits. We have analyzed the subject from different abstraction levels —device, circuit and system levels— bringing solutions to important issues that otherwise would compromise the circuits correct behavior. In this thesis we have started by presenting resistive switching technologies. We introduced the foundations of the memristor, such as the initial Chua's concept, and immediately moved to describe the behavior and characteristics of manufactured physical RRAM devices: structure, categories, as well as their switching and conduction mechanisms. We also presented the best known memristor models, which were followed by the review of the most common RRAM based applications. We ended this introduction by linking the description of the reliability issues affecting RRAM and CMOS circuits with this thesis motivation.

In the first part of the thesis we dealt with the problematic of working with an emerging technology that is still immature. Thus, we started studying the memristor compact modeling and simulation matter from the device level. In first place we presented a characterization framework and design & simulation methodology for the definition, analysis and improvement of present and future compact models. The Memristor Application Framework (MAF) has been presented as a support tool for both model and circuit designers, closing the gap between the devices experimental measures, the behavioral description via a differential equations system, to give birth to the final SPICE-like compact model. Making use of MAF tools and methodology, we developed an RRAM SPICE compact model. This model, relying on the idea of independently describing conductive processes and switching mechanisms, is able to resemble both unipolar and bipolar physical devices. Built using a modular design, this model allows different electrical magnitudes —voltages, currents, flux, charge or energy— to describe both the switching and the device aging. Additionally, this modularity allows the model to be extended in order to emulate future RS technologies characteristics. As an example, thermal dynamics were included. One of the most important contributions regarding the model is that without requiring Verilog-A,

it accurately emulates time-dependent variability, multi-level capabilities, device degradation or data retention. To validate the model we presented accurate fittings of two different physical devices, one with unipolar and one with bipolar nature.

Using the concepts and knowledge of the first part, we devoted the second part of the thesis to the characterization and analysis of RRAM-CMOS hybrid circuits. Thus, we integrated the knowledge regarding device abstraction level into circuit level. We first discussed the problems to be addressed when designing circuits using RS and short-node CMOS technologies. As a solution to the lack of methodologies and support tools aiming at the reliable design of circuits with these hybrid components, we first provided the circuit designer with a strategy to define reliable analyses of RRAM based circuits. To support it, we also developed an ad-hoc simulation framework based on a multi-metric evaluation paradigm. Using this multi-metric approach, the framework is able to measure the performance and reliability of circuits that simultaneously suffer from the effects of several error sources. Therefore, the designer gets aid in the always difficult task of evaluating the effects caused by PVT variations, aging, degradation, data retention or radiation. We validated the solution capabilities against experimental measures of physical memories, and to illustrate the framework and methodology capabilities, we later described three different cases of study: a digital circuit, an analog subsystem, and finally an RRAM-CMOS circuit.

The third part of the thesis deals with the design of reliable circuit applications at system level. Using the resources previously described, we have presented solutions to two different issues affecting RRAM memories. First, we have designed and evaluated a novel architecture that is able to mitigate the dramatic effects caused by temperature in RRAM memories. No matter the circuit temperature, this writing architecture dynamically adapts the writing voltage amplitude to the environmental conditions, correctly achieving the desired operation. But what is more important, as different writing schemes imply distinct cell lifetime, power consumption, or other characteristics, our architecture adapts the driver to specific application objectives. By dynamically loading different writing profiles, the driver selects among a myriad of writing strategies improving the system performance, power consumption, reliability, data retention abilities and HRS/LRS ratio durability. Second, non-volatility represents a challenge when designing ultra secure applications. Nonvolatile memories containing secure keys can become a weak point in the liability chain of a secure application, making it vulnerable to attacks. As data-encryption may not be sufficient, we developed an auto-erasable memory architecture, to secure RRAM memories from physical and firmware attacks. Complementing encryption approaches, this architecture provides an additional security layer, preventing the access to sensitive data. If any kind of firmware or physical attack occurs, the erasing procedure automatically erases

the whole memory in a single stage, avoiding the expense of sequentially erasing individual cells. Moreover, an optional emergency power supply, based on on-chip capacitors, allows the designers—with a large area trade-off—to include an additional power scheme to be used should any physical de-cap be attempted.

8.1 Future Work

During the development of this thesis, the number of publications, workshops, and research groups working to bring the next generation of RRAM based applications has nothing but increased. It is then reasonable to foretell a bright future to resistive switching technologies. Device manufacturing, device modeling, circuit or system level design, each area of research is experiencing a moment of changes, and it is therefore a good time to contribute to this small electronics revolution. To conclude this document, we would like to discuss some lines on the work that can follow the ideas presented here.

Short-term future work involves the direct improvement of the developed tools. The MAF framework presented in Chapter 3 has been made public, and, to this day, several researchers and students from different universities make use of it. It would be then a great achievement to get to new developers, so they can contribute to the project expanding it. The inclusion of new models, the improvement of the simulation core, and a tighter interaction with the circuit simulator would benefit from a larger number of software developers. Finally, and given the huge amount of researchers finding resistive switching behavior in new and improved layered structures, the MAF fitting module could benefit from different heuristic algorithms to find the best solution to the parameter matching problem.

In relation with this last feature, memristor manufacturers would take advantage of an automated stage that would provide feedback between the SPICE simulator using the memristor model presented in Chapter 4, and MAF. When finding the set of parameters that best fits measured data, and given the different computation methods present in both simulators, a co-simulation process can further detect convergence issues, as well as provide more accurate results. Additionally, the access to experimental results would open a wide range of possibilities. The knowledge given by measures data regarding operation at extreme temperatures, different size devices, variability and degradation, etc., would help to create a predictive PDK for different RRAM technologies.

Moving up to circuit simulation level, it is clear that there is a long way before methodologies and tools like the one proposed in Chapter 5 gets standardized in commercial solutions and design flows. Therefore there is room for improvement in this field. In particular, collaborating with teams developing commercial so-

lutions would enrich the framework functionalities, and therefore better aiding circuit designers. Additionally, and similarly to MAF maintenance and expansion, the simulation framework would benefit from the contribution of additional developers. Here, the inclusion of a mechanism distributing the computational load between different machines in a high performance cluster is a clear line to be explored.

Thus the ultimate objective would be to continue with the development of tools and strategies that, integrated in the standard circuit design flows, would lead to the formulation of reliable RRAM based applications.

On the other hand, and making use of the presented contributions long-term future work will be composed of new research lines, focused on the development of new and more reliable applications. Gaining perspective on the present and future of electronic designs, a myriad of new RRAM applications, or improved adaptations of traditional CMOS systems, will appear. In this sense we plan to continue working on the design of reliable systems as the ones presented in Chapters 6 and 7. Memristor devices are candidates to be used as key elements within IoT systems. Therefore, the first research line to be opened is the migration of IoT applications towards RRAM based architectures.

Second, a particular area that would benefit from the presented contributions would be non-conventional computing approaches. These days we see how artificial intelligence and machine learning are gaining more and more attention. Neural networks built with memristors, as well as other brain-like computing architectures are exciting research lines whose maths are sustained on the accurate modeling and simulation of emerging devices.

Finally, and also considering the direct application of the developed tools and methodologies, the design of novel reconfigurable hardware architectures also stands in the horizon as a research line particularly interesting.

If, For Any Reason You're Not Satisfied, I Hate You.

Sales Clerk, Futurama.

MAF Implementation Details

As Figure 3.3b introduced, MAF was designed following a modular structure built around the device models definition. Starting from that, the particular set of model, environment, stimuli, and simulation parameters are gathered forming an entity *scenario* ready to be simulated and processed by a thread. In this Appendix we describe the technical details regarding the modeling of these electric devices as well as the handling of the concurrent simulation of the different scenarios.

A.1 Electric Device Model

The framework uses a basic structure to characterize the desired model. Taking into account that a single device may be modeled by several implementations, each attending distinctive behaviors under certain conditions, different structures can be included within the framework model library.

The model electric structure is composed of the specific device behavior, a collection of parameters and their ranges, a set of magnitudes and state variables, and a simulation control. Using the above, the considered model provides the

implemented simulators a specific method to compute the response step to the given stimuli, usually in form of differential equations that represent the device behavior.

A.1.1 Magnitudes Modeling

The electrical magnitudes involved in the characterization of the device can be categorized into two groups:

- Primary magnitudes. In this group the designer can include the most relevant magnitudes that are present in all devices. This group includes voltages at specific nodes, currents or impedances.
- Secondary magnitudes. Several electrical magnitudes may be needed in the characterization of some devices, but that is not the case in some others. An example of these secondary magnitudes can be charge, internal voltages or flux.

The second group of variables involves state variables, in which conductance and switching mechanisms usually rely, determining the model behavior. Each magnitude or state variable implementation is defined by the current magnitude value, the recent magnitude history, a reference to the preferred integration method—in order to provide the magnitude integral over the time, and the global history of its time evolution—in case a specific simulation stage needs to be recovered.

A.1.2 Parameters

The device parameters characterize a specific device model instance depending on its specific properties. That way, each device model instance differs from others in the value of its parameters. The definition and implementation of a parameter object within the framework—instead of the inclusion of numeric values in the device behavior—helps both in the extensibility of the models and in the instantiation of different scenarios to be simulated. Following this approach, the accuracy of the simulation results is guaranteed.

A.1.3 Model Simulation Control

The simulation control included in each model determines the simulation computational load and the possible non-convergence problems. Thus, each model must have a method to control the timing requirements in each computational step. That method involves a mechanism for controlling the success of the simulation step and the control of the incremental and decremental ratios of the timing resolution needed.

A.2 Simulation Scenarios and PVT Variations

As introduced before, the *scenario* object is, therefore, the basic instance to be simulated.

In presence of PVT variations, the fluctuations of parameter values could destabilize the device behavior. During the creation of the PVT variation scenarios the device parameters values could be constant, or by contrast, experiment fluctuations. To study the dynamics of a specific component in presence of parameter variations, together with the standard *Corner Analysis* simulation process, we have included the *Monte Carlo Analysis* method, where parameter variation is given by a probabilistic deviation from the mean value. Therefore, instead of a single scenario, several simulations are executed. This way, a wide set of scenarios which differ in the values of the device parameters, the temperature (if the device model depends on it), and the voltage fluctuations is determined. A broader analysis is performed, revealing the parameter limits between which the model behaves as expected, or, on the other hand, in which zones of the design space the models lead to non-convergence problems or non-desired operations, playing a key role in circuit validation.

In order to increase the framework performance, each simulation is related to a thread which runs in parallel making good use of the simulation platform capabilities. The framework optimizes the usage of platform resources, allowing to reduce the total simulation time and processing independently each of the considered scenarios.

A simulation control manages every scenario simulation, recording the stability of each of them. This way this control supervises the concurrency of the different simulations, and once each thread is finished, it independently processes the results and calculates the global statistical data. Next section thoroughly describes this approach. The execution of transient simulations is controlled by a simulation manager that handles non-convergence issues, and carries out data population control. This simulation management is described next.

A.3 Thread Management

We have implemented a thread management system that follows a three level hierarchy. Each level is in charge of the creation, management and destruction of the instances in the next level. The hierarchy solution is described by the scheme shown in Figure A.1. Each set of model and environmental parameters ready to be characterized by either the *Device Characterization* or the *Level Characterization* simulators define a *scenario*. The developed framework has the ability of automat-

Figure A.1: *MAF* Thread Manager.

ically handle the huge amount of scenarios that concurrently runs, managing the available resources and coordinating the threads' interactions.

Going from top to bottom in the system hierarchy, the first level includes the main execution process (commonly related to a Graphical User Interface, GUI) who manages the generation of a *Thread Manager (TM)* object.

In the second level, the TM is responsible for the management (creation, control and destruction) of simulations pertaining to different scenarios. Each of these simulations corresponds to a third level thread fork. Likewise, TM is responsible for managing the concurrency and access to resources of each simulation.

As mentioned before, the scenarios that must be considered in Monte Carlo Analysis simulations are assigned to the third thread level where they run in parallel. Each of them differs from the others in input voltage, temperature and device parameters. Once each simulation is finished, the TM processes the results, the possible stability problems and notifies the GUI or main process, in order to display or export the final reports.

A.4 Non-Convergence Problems Handling

There are different methods to control the system convergence depending on the type of simulation that the framework supports. Our framework is focused on solving transient analysis. For this reason we have considered three techniques to deal with non-convergence problems:

1. Complex scenarios are related to long simulation durations. Some possible non-convergence problems could not be detected until advanced states in

the transient simulation. The first non-convergence technique consists on the simulation of a simplified scenario, where, using non-accurate timing, possible non-convergence situations can be quickly discerned. This way, the designer is able to save time focusing on the involved problematic magnitudes.

2. The second technique is associated to a Thread Manager (TM) which is responsible for supervising the correct computation of the different simulations, keeping record of the list of detected problems and the computation statistics.
3. Finally, with the aim of dealing with the location, handling and correction of punctual non-convergence problems a complete simulation work-flow has been considered.

Next we will detail the operation scheme of a single simulation thread. The presented scheme improves the performance of the simulation, reduces and narrows possible convergence issues and provides higher resolution when required.

Figure A.2: MAF Single simulation thread diagram.

Figure A.2 illustrates the scheme of a single simulation thread. In the first stage, the scenario is set up by its TM parent, including the involved input signals, the initial timing conditions and the specific device parameters in this scenario.

The computation of the steps automatically begins after this set-up stage, and it will not end until its timing module indicates that the simulation is completed or an unrecoverable non-convergence problem appears. Therefore, until the timing module indicates the simulation end, every step is computed, checked, and evaluated:

- When step i has been correctly computed and the simulation converges, several checks are performed. In case the simulation is completed, the thread notifies the TM its success and the results are processed and exported. Otherwise, if the simulation has not reached its end, the software tests if a time step increment can be accomplished. In that case, the pertinent timing modifications are performed, the magnitudes and state variables are updated, the scenario backup is renewed and we proceed to compute the next step, $i + 1$. On the other hand, if the time step cannot be incremented, the timing module does not get modified. The set of magnitudes and state variables are updated, and we proceed to compute the next step, $i + 1$.
- By contrast, if step i has not been correctly computed, the software checks whether a time step reduction can be done. In that case, the timing is updated, the magnitudes and variables are not updated and the simulation continues computing step i with the new timing. If the time step cannot be reduced, the timing module checks the viability of recovering from a backup scenario. In cases where the scenario has already been recovered, the simulation gets flagged as unsuccessful, the thread proceeds to notify and the results are exported.

Following those steps it is guaranteed that the simulated data is saved both in convergent and non-convergent processes. Thus, the designer can characterize the behavior of the device and study possible problems in the simulation, tracking the problem until the moment of the failure.

In the same way, the developer is able to access the time evolution of the magnitudes involved. This is useful for both the circuit designer and model developer when a new device model is added to the library because, as seen in Section A.1, the device model object is responsible for deciding whether the step has been correctly performed. Additionally, the device model also decides if a time step modification would be desirable in order to improve accuracy or performance.

Figure A.3: *Data Population Control additional step.*

A.5 Data Population Control

MAF simulators work using configurable buffers that store the computation of magnitude values as well as previous steps history. Should the workstation RAM size be larger, the tunable size of these memory structures adapts, providing advantages over more modest workstations. These advantages mainly refer to the reduction of the simulation time, to the ability to recover from more errors—see Figure A.2— as well as the ability to graphically represent more scenarios.

After the analysis ends, the data of every single simulation can be saved in hard disk. Additionally, when a huge amount of data is processed, if the results require to be graphically represented, streaming the simulation data to disk becomes necessary. As a drawback, this step introduces latency compared against the pure RAM buffers approach.

Thus, when large amounts of data are processed and stored, a data population control system becomes essential. Figure A.3 displays the inclusion of the data population control in the thread diagram of Figure A.2. That way we can store larger amounts of data, regardless of the number and size of the simulations. Finally, this approach also eases the exportation of data to external mathematical software.

I Don't Want To Live On This Planet Anymore.

Professor, Futurama.

B

Multi-level Device LTSPICE Code

To illustrate the proposed model structure, we present the LTSPICE source code related to the memristor characterized in Figure 4.3. Additional resources such as the full source code of charge controlled memristor, unipolar and bipolar devices, temperature dependency and variability aware experiments can be downloaded from http://vlsi.die.upm.es/memristor_spice_model. Compatible simulators: Linear Technologies LTSPICE and Cadence Spectre.

```

*****
2  ** http://vlsi.die.upm.es/memristor_spice_model
  ** 2b-multilevel v1.0, 21/01/2016
4  ** switching defined by energy thresholds
  ** Three Schottky and one ohmic levels
6  *****
  .SUBCKT memristor Plus Minus PARAMS:
8  +pi=3.1415926 Kb=1.38e-23 q=1.6e-19 eps0=8.85e-12
  +area=6.4e-9 d=4e-8 scl1=5 scl2=7 scl3=9 AA1=1e6 AA2=2
    e6 AA3 =5e7 Ub=0.9
10 +epsr=8
  +Ron=300 T0=300 ktemp=1 cI=1e-4 ronp=6.1k ronnn=8.5k epsi
    =epsr*eps0
12 ** energy threshold multilevel parameters
  +p_th_1_2=0.5e-13 p_th_2_3=p_th_1_2+1e-8 p_th_3_4=
    p_th_2_3+0.5e-5
14 +p_th_4_3=-1e-7 p_th_3_2=p_th_4_3-1e-8 p_th_2_1=p_th_3_2
    -1e-9
  ** internal voltages
16 +v_s1=0e-7 v_s2=1e-7 v_s3=2e-7 v_s4=3e-7

18 *****
  ** Energy computation
20 *****
  * total energy
22 Etpospower tpp 0 value={idt( IF( V(Plus, Minus)>0 & v(pp)
    )<p_th_3_4, I(Gcond)*V(Plus,Minus), 0) ,0 )}
  Etnegpower tnp 0 value={idt( IF( V(Plus, Minus)<0 & v(np)
    )>p_th_2_1, abs(I(Gcond))*V(Plus,Minus), 0),0 )}
24 * relative energy
  Epospower pp 0 value={idt( IF( V(Plus, Minus)>0, I(Gcond)
    )*V(Plus,Minus), 0) ,0, v(trig_p)>0)}
26 Enegpower np 0 value={idt( IF( V(Plus, Minus)<0, abs(I(
    Gcond))*V(Plus,Minus), 0),0, v(trig_n)>0)}
  E_trigger_p trig_p 0 value={IF( v(s)==v_s1 & v(pp)>=
    p_th_3_4, 1e-7, 0)}
28 E_trigger_n trig_n 0 value={IF( v(s)==v_s4 & v(np)<=
    p_th_2_1, 1e-7, 0)}

```

```

30 *****
** state computation
32 *****
Estate s 0 value={ switching(V(Plus, Minus), V(pp), V(np
    ), p_th_1_2, p_th_2_3, p_th_3_4, p_th_4_3, p_th_3_2,
    p_th_2_1, v_s1, v_s2, v_s3, v_s4) }
34 Rstate s 0 r=100k

36 *****
** Conduction processes
38 *****
Eon on 0 value={V(Plus, Minus)}
40 Ron on 0 r=ron
Goff1 0 off1 value={ sgn(V(Plus, Minus))*area*AA1*ktemp
    *(T0**2)*exp(-q*Ub/(Kb * T0))*exp( sqrt(abs(V(Plus,
    Minus)))*(sc11 + q / (Kb * T0) * ( sqrt(q / (d * 4 *
    pi * epsi)))) ) }
42 Goff2 0 off2 value={ sgn(V(Plus, Minus))*area*AA2*ktemp
    *(T0**2)*exp(-q*Ub/(Kb * T0))*exp( sqrt(abs(V(Plus,
    Minus)))*(sc12 + q / (Kb * T0) * ( sqrt(q / (d * 4 *
    pi * epsi)))) ) }
Goff3 0 off3 value={ sgn(V(Plus, Minus))*area*AA3*ktemp
    *(T0**2)*exp(-q*Ub/(Kb * T0))*exp( sqrt(abs(V(Plus,
    Minus)))*(sc13 + q / (Kb * T0) * ( sqrt(q / (d * 4 *
    pi * epsi)))) ) }
44 Raux1 off1 0 r=1k
Raux2 off2 0 r=1k
46 Raux3 off3 0 r=1k
Gcond Plus Minus value={ max( -cI, min(cI, IF(V(s)==v_s1
    , i(Goff1), IF(V(s)==v_s2, i(Goff2), IF(V(s)==v_s3, i
    (Goff3), i(Ron) ) ) ) ) ) }

```

```

*****
50 ** event counters
*****
52 E_p_events p_events 0 value={floor( V(tpp)/p_th_3_4 )}
E_n_events n_events 0 value={floor( V(tnp)/p_th_2_1 )}
54
*****
56 ** switching function
*****
58 .func switching(v,pp,pn,th_1_2,th_2_3,th_3_4,th_4_3,
    th_3_2,th_2_1,v_s1,v_s2,v_s3, v_s4) {
+     IF(v>=0,
60 +         IF(pp>=th_3_4, v_s4,
+         IF(pp>=th_2_3, v_s3,
62 +         IF(pp>=th_1_2, v_s2,
+         v_s1) ) ),
64 +         IF(pn>=th_4_3, v_s4,
+         IF(pn>=th_3_2, v_s3,
66 +         IF(pn>=th_2_1, v_s2,
+         v_s1)) ) ) }
68 .ENDS memristor

```

What do you say wanna go around again?

Fry, Futurama.

Radiation Modeling and Tools

As stated in Section 2.5, the study of PVT variations in CMOS technologies is completely established, thus meaning that methodologies and tools already provide the support required to design and fabricate circuits considering PVT caused problems. On the other hand, in that same Section RRAM intrinsic problems have been reviewed and described. Radiation is then a specific problem that require the attention and time of circuit designers. Consequently, we dedicate the next section to provide the context on how radiation have been traditionally treated in circuit design.

C.1 Overview of Radiation Models and Tools

The objective is to give the big picture regarding how circuit designers model this source of errors, and how their effects are usually emulated, focusing on the weaknesses that these traditional approaches bring to RRAM-CMOS hybrid circuits.

Many efforts have been devoted to both experimental characterization and simulation of circuits under radiation. Although until recent years these efforts have been focused on CMOS technologies, more and more resources are dedicated to understanding the behavior of RRAM devices working in radiation environments.

In depth reviews on modeling the radiation effects on semiconductors at transistor level can be found in [ME07, FCMG13]. Next we briefly review the main models and tools presented in the literature at different abstraction levels.

C.1.1 CMOS Technologies: Physical and Transistor Level

TCAD [Syn] is extensively used as the reference when validating the effect of radiation on semiconductors due to its accuracy and the broad range of applications supported. However, the enormous computational load associated to a simple device simulation restricts its application to the device level. Additionally, TCAD modeling is strictly associated with each transistor design in a specific technology.

Monte Carlo (MC) Based 3D Physical Simulators allow to characterize the strikes of radiation particles in specific electronic devices. Several works take advantage of this kind of analysis, from tool-kit/libraries like Geant-4 to complete solutions like MRED, MUSCA-SEP or MULLASIS [RWA⁺13]. Even though this is a computation hungry technique, it has been successfully used to analyze error rates in several circuit modules [WSW⁺08]. Even with the smallest technologies, these simulators have proven to be valid to study SER relations with different kind of impacting particles [HAR15].

SPICE-level radiation source models came up as a solution to the need of including radiation effects in circuit simulators. Radiation can be emulated by the injection of a shaped current, either internally or externally simulated, in the affected device. This is the case of the well-known double exponential [DWB⁺07]. The transistor affected by radiation can be also substituted by a modified model. In [MEY⁺08] the charge collection mechanism in fully depleted CMOS/SOI devices is accurately addressed by a *Switch Based Model*. The *Equivalent Collection Model* [ME07] uses variational calculus to solve the equation which determines the induced current shape. Another subcircuit model [KSA⁺09] for a 90nm commercial technology dynamically controls the bias-dependent induced current.

However, smaller technologies impose a new way to study the matter. Robust designs should consider both radiation and process variations early in the design cycle. As an example, in [KBK⁺11] 65nm and 90nm SRAM performance under process variations is compared.

Experimental data, together with physical simulators such as TCAD, have shown how planar technologies beyond $65nm$ exhibit charge collection effects that lead to a variation in the induced current compared to the classical double exponential [FCMG13]. Even though double exponential is suitable enough for most analysis, some studies may require more accurate injection schemes.

In order to increase the accuracy of the emulated transient, alternative current injection methods have been presented. Mixed-mode simulators [CC05] generate a dependent injected current waveform by interconnecting the SPICE-like simulator with the physical tool. SPICE analysis and 3D physical simulation are performed simultaneously, and this feedback allows more accurate results. On the other hand, the use of complex numerical solvers slows down the SPICE simulator, and consequently constrains the size of the circuit under study. As alternative, an efficient estimation method is presented in [BSH⁺12], reducing the computation time compared to 3D physical simulators in design flows where SETs are not the primary concern.

The induced current characterized in different emerging non-planar devices such as Si FinFET, III-V FinFET, and III-V Hetero-junction tunnel FET (HTFET), can be modeled using a double exponential current injection [LCDN14].

C.1.2 Gate and System Level

Once the SET has been modeled, the pulse propagation through the circuit must be evaluated [NBS⁺08]. Novel methods [GBBS12] deal abstractly with the problem by considering analytic models at the data-path level, accelerating the simulation of the propagation through the individual modules.

Works like [QBRB13, BHAS15] put SER instead SET/SEU under the spotlight, being able to predict the system tolerance under given radiation conditions. In [FHS15] a cross-layer methodology for the SER estimation under PVT variations is presented, considering digital glitches generation, propagation and latching.

The advantages of those simulations are scalability, capacity to study error mask-effects between components, and their ability to be hardware accelerated using FPGAs [LOEGV⁺07]. However, as this kind of tools statistically approach the radiation characterization, and sometimes even using an excessive simplification of the radiation effects, it makes them not suitable to be applied to predictive technologies or emerging devices studies.

C.1.3 Devices Not Supported in Standard Work Flow: RRAM and Analog Modules

As we have seen, memristors, as well as other emerging devices such as spin-transfer torque RAMs or phase change RAMs are progressively being incorporated

into standard circuits. Several studies measure their characteristics regarding SET [CKP⁺14, BHS⁺14, LMBY15] and TID [LWHL15, WWBH⁺14, MHP⁺14, MFH⁺15]. However, they lack from support tools when designing full systems, and are not integrated in the standard commercial software tools.

Finally, radiation effects are usually studied from the digital point of view, focusing on how the pulse is generated and propagated until it is masked or it generates a bit upset in a memory element. However, analog circuits are as susceptible as digital, and radiation can become an inadmissible error source in the implemented application. Even though studies have been presented analyzing radiation effects in analog circuits [SMHB06, RDR⁺14], the industry needs specific tools to automatically characterize the radiation hardening characteristics of the studied module.

I do. [Futurama's Ending]

Leela, Futurama.

Circuit-Level Simulation Framework Implementation Details

D.1 Simulation Load and Data Management

As introduced in Section 5.4, the framework includes speedup features that allow a fast and intelligent simulation.

D.1.1 Critical Parameter Value Binary Search

A modified version of the well known [\[RBW⁺15\]](#) binary search algorithm has been implemented to quickly estimate the values of the key parameters that may affect the circuit behavior.

The algorithm's pseudo code described in Listing D.1:

Listing D.1: Binary search algorithm pseudo-code

```
// initial value computation and scenario set up
2 key_param_value = get_initial_value(key_param, false);
  set_up_scenario(key_param, key_param_value);
4 end_of_iteration = false;
  // iterative process
6 while(!end_of_iteration){
  // compute
8 compute_transient_analysis();
  // metrics application and results extraction
10 is_circuit_reliable = apply_metrics();
  // compute parameter next value
12 old_param_value = key_param_value;
  key_param_value = get_param_value(key_param,
    is_circuit_reliable);
14 // compute exit condition
  end_of_iteration = is_param_accurate_enough(
    key_param_value, old_param_value, accuracy);
16 }
```

Within this process, each iteration benefits from `get_param_value`, a heuristic method that determines the midpoints. Being N a natural odd number greater than 2, a_n and b_n , with $a_n < b_n$, the boundaries of the considered parameter values at the n th iteration, the heuristic midpoint scheme is selected by the user among the following:

- Linear scheme: The next parameter value is selected as $next_value = (a_n + b_n)/2$.
- Logarithm scheme #1: The next value is given by $next_value = \exp(\log(b_n) - \log(a_n))/(N - 1)$.
- Logarithm scheme #2: The next value is given by $next_value = \exp(\log(b_n) - \log(a_n))/(N - 1) * N$.

Therefore, for a specific accuracy a lower computation time is required. Figure D.1 shows an example of the lineal scheme applied to this algorithm.

Additionally, the stop condition computation relies on a customizable method, `is_param_accurate_enough`, which is responsible for the refinement of the critical value calculus. A higher accuracy would increase the number of iterations, and therefore would result in slower analyses.

Figure D.1: Injected charge evolution for critical charge calculus using the implemented binary search algorithm with a linear value selection scheme.

D.1.2 Sweep Parameter Design Space Division

The division of the design space of each dimension is independently handled, allowing non-linear schemes to increase the accuracy at some regions. Following similar division schemes as the one shown in the previous section, the user is able to select the method among:

- Linear scheme. Convenient for a most parameters, linearly explores the parameter design space.
- Logarithm scheme. Convenient for exponential dependencies between parameters, such as critical charge-feeding voltage in SRAMs [RBW⁺15], or the filament width dependence on the barrier lowering coefficient in some oxide RRAMs [BHS⁺14, MFH⁺15].
- Square root based division of the parameter values range. Convenient for quadratic relationships such as amplitude/length of writing pulses applied to *amorphous* SiC based RRAMs [MFH⁺15].

- Manually inserted parameters from a list given by the user. Convenient for dynamic analysis, in which the user is able to determine the most sensible timing (clock edges, memory operations, etc.).

D.1.3 Masked and Omitted Nodes or Instances

Each radiation occurrence or model redefinition is translated into a new modified netlist ready to be simulated and analyzed. However, some emulated radiation occurrences do not influence the signals involved in the user defined error metrics. This is the case of injections at sources or ground nodes or nodes/circuitual instances whose paths are not connected to the sensible signals.

The early element analysis of the input netlist shown in Figure 5.7 builds up a relational graph of the elements and nodes that are accessed by the framework for the nodes/instances masking algorithm. This masking process also accepts a user given list of instances not to be radiated. This way the user can isolate the instances of lower influence in the circuit global functioning saving their simulation time; or on the contrary, simulate only selected modules to individually study their influence in the reliability analysis. The time saved by this masking methods depends then on the circuit elements interconnections as well as on the signals defining the error metrics.

D.1.4 Granularity

A heuristic analysis has been developed to avoid the simulation of non-desired nodes, subcircuit elements, instances or even referenced circuits. The statistical data obtained from the impact time dependence of the different submodules aids the tuning of the granularity configuration.

The circuit element analysis notifies the designer in case several instances of the same model sharing identical parameters are found. At this point, the designer can choose whether to perform the original analysis or, on the contrary move to a modular approach. In this modular method, the replicated instance is characterized, and the resulting properties are extended to the remaining instances.

This technique saves the time spent in the analysis of the replicated modules, but as a drawback, some propagation errors can be unintentionally skipped. As shown in Section D.1.6, for this 14 bit adder example, the inspection of the individual node results shows pulses propagation trough the carrying lines not considered in the 1 bit module. For that reason, after the analysis of individual modules, the framework presents both stand alone radiation occurrences and node/parameter/global statistical results.

Figure D.2: n bit adder used to measure the framework computational performance.

The use of this methodology is specially convenient in modules with high density of replicated modules, such as pipelined structures or memories, and becomes essential for the characterization of larger circuits.

D.1.5 Computational Acceleration Techniques

Parallelization techniques have been widely used by high end simulators focused on computationally expensive analysis. The automation of the nodes/circuit elements injection, the multi-parameter analysis capabilities, and the handling of different variability schemes requires from parallelization and simulations control. To solve the issue a *CPU/RAM load sensing based scheduler* is in control of the independent thread creation and intercommunication. The scheduler intermittently asks the operating system for the available resources, knowing at each time which of them are accessible to be used by additional threads.

Additionally, the framework admits the compilation of some *AHDL*, *Verilog-A* and *CMI/PLI* components such as electrical sources, technology resistances and capacitances, etc., that reduces the time used during computation processes. The same applies to *Verilog-A* and *CMI/PLI* components that requires from a prior compilation. The developed tool is able to communicate with the underlying circuit simulator building up a data base of compiled components that, correctly configured, can be shared between simulations, saving the compilation time of each component.

(a) Critical charge dependence on temperature for both 1 bit and 14 bit adders. The scatter set of individual radiated nodes is presented in blue. Yellow and red lines present the mean and the worst values of critical charge.

(b) Computation time and speedup comparison between 1) Granularity and Parallel Approach; 2) Parallel Approach; 3) Traditional approach.

Figure D.3: Simulation summary results for 1 bit and 14 bit adders.

Table D.1: 14 bit adder inventory. The *ahdl injector element* refers to the current source used to emulate the radiation impacts.

Circuit element family	Quantity
probe	1
bsim4 transistor	820
ahdl injector	1
capacitor	15
vsorce	30

D.1.6 Efficient Computation Example

We have selected a middle size example, a 14-bit registered adder to show the improvements the framework brings regarding simulations performance. Figure D.2 displays the modular circuit scheme used. The 14 bit registered adder complexity gets detailed by the circuit element count of Table D.1. The timing results provided refer to the simulation of this circuit module in a 4 core 24GB RAM virtualized *i7* computer. In each simulation, the user defined error metric is particularized as any occurrence of an incorrect bit at any register output.

When applied to the 14 bit adder, the framework detected different instances replicated: from logic gates to the pair full adder followed by a register. We selected to individually analyze a single full adder as the key module, and extrapolate its results to the 14 full adder instances.

Figure D.3a shows the critical charge dependence on temperature for both 1 bit and 14 bit adders. The scatter data is presented together with the mean and the worst cases scenarios. The speedups obtained after applying the presented methods are compared against traditional analysis in Figure D.3. While the traditional approach took more than 33 hours, the proposed simulation scheme only required 25 minutes, achieving speedups up to 80 $\times$.

The error derived from the extrapolation of 1 bit adder characteristics to the 14 bit adder are collected in Table D.2. There, critical charge statistical functions of mean, mode, median and the worst cases are compared.

To sum up, it can be seen how granularity allows us to decide whether to gain huge speed-ups we can get rid of the simulation of replicated modules, always taking into account the involved loss of accuracy.

Table D.2: Critical charge results comparison between 14b adder simulation and the approximation using 1b adder simplification.

Critical Charge Metric	Percent Error
Negative Injections Q_{crit} Mean	11.6 %
Negative Injections Q_{crit} Worst Case	3.91 %
Negative Injections Q_{crit} Mode	0 %
Negative Injections Q_{crit} Median	0 %
Positive Injections Q_{crit} Mean	5.42 %
Positive Injections Q_{crit} Worst Case	0 %
Positive Injections Q_{crit} Mode	0 %
Positive Injections Q_{crit} Median	0 %

Bibliography

- [ABGI14] Stefano Ambrogio, Simone Balatti, David C. Gilmer, and Daniele Ielmini. Analytical Modeling of Oxide-Based Bipolar Resistive Memories and Complementary Resistive Switches. *IEEE Trans. Electron Devices*, 61(7):2378–2386, jul 2014.
- [AGHS12] Fabien Alibart, Ligang Gao, Brian D Hoskins, and Dmitri B Strukov. High precision tuning of state for memristive devices by adaptable variation-tolerant algorithm. *Nanotechnology*, 23(7):075201, feb 2012.
- [AILV15] Hernan Aparicio, Pablo Ituero, and Marisa Lopez-Vallejo. 2.64 pJ reference-free power supply monitor with a wide temperature range. In *2015 Int. Work. C. Var.*, pages 9–12. IEEE, sep 2015.
- [AJDG16] Ahmedullah Aziz, Nicholas Jao, Suman Datta, and Sumeet Kumar Gupta. Analysis of Functional Oxide based Selectors for Cross-Point Memories. *IEEE Trans. Circuits Syst. I Regul. Pap.*, pages 1–14, 2016.
- [AKG⁺14] Chiyui Ahn, Seyoung Kim, Tayfun Gokmen, Oliver Dial, Mark Ritter, and H. S Philip Wong. Temperature-dependent studies of the electrical properties and the conduction mechanism of HfOx-based RRAM. *Proc. Tech. Progr. - 2014 Int. Symp. VLSI Technol. Syst. Appl. VLSI-TSA 2014*, pages 1–2, 2014.
- [AMW⁺16] Stefano Ambrogio, Valerio Milo, Zhongqiang Wang, Simone Balatti, and Daniele Ielmini. Analytical Modeling of Current Overshoot in Oxide-Based Resistive Switching Memory (RRAM). *IEEE Electron Device Lett.*, 37(10):1268–1271, oct 2016.
- [AP11] Hisham Abdalla and Matthew D. Pickett. SPICE modeling of memristors. In *2011 IEEE Int. Symp. Circuits Syst.*, pages 1832–1835. IEEE, may 2011.
- [Arg68] F. Argall. Switching phenomena in titanium oxide thin films. *Solid. State. Electron.*, 11(5):535–541, may 1968.

- [ATB⁺15] Alon Ascoli, Ronald Tetzlaff, Zdenek Biolek, Zdenek Kolka, Viera Biolkova, and Dalibor Biolek. The Art of Finding Accurate Memristor Model Solutions. *IEEE J. Emerg. Sel. Top. Circuits Syst.*, 5(2):133–142, jun 2015.
- [Aus15] Todd Austin. Keynote talk I: Ending the Tyranny of Amdahl’s Law. In *2015 33rd IEEE Int. Conf. Comput. Des.*, pages vii–ix. IEEE, oct 2015.
- [AVS⁺16] M Azzaz, E Vianello, B Sklenard, P Blaise, A Roule, C Sabbione, S Bernasconi, C Charpin, C Cagli, E Jalaguier, S Jeannot, S Denorme, P Candelier, M Yu, L Nistor, C Fenouillet-Beranger, and L Perniola. Endurance/Retention Trade Off in HfOx and TaOx Based RRAM. In *2016 IEEE 8th Int. Mem. Work.*, number c, pages 1–4. IEEE, may 2016.
- [BAC⁺16] Simone Balatti, Stefano Ambrogio, Roberto Carboni, Valerio Milo, Zhongqiang Wang, Alessandro Calderoni, Nirmal Ramaswamy, and Daniele Ielmini. Physical unbiased generation of random numbers with coupled resistive switching devices. *IEEE Trans. Electron Devices*, 63(5):2029–2035, may 2016.
- [BAW⁺15] Simone Balatti, Stefano Ambrogio, Zhongqiang Wang, Scott Sills, Alessandro Calderoni, Nirmal Ramaswamy, and Daniele Ielmini. Voltage-Controlled Cycling Endurance of HfOx-Based Resistive-Switching Memory. *IEEE Trans. Electron Devices*, 62(10):3365–3372, oct 2015.
- [BBB09] Zdeněk Biolek, Dalibor Biolek, and Viera Biolková. SPICE Model of Memristor with Nonlinear Dopant Drift. *Radioengineering*, 18(2):210–214, 2009.
- [BDF⁺14] A. Belmonte, R. Degraeve, A. Fantini, W. Kim, M. Houssa, M. Jurczak, and L. Goux. Origin of the current discretization in deep reset states of an Al₂O₃/Cu-based conductive-bridging memory, and impact on state level and variability. *Appl. Phys. Lett.*, 104(23):233508, jun 2014.
- [BGS⁺14] Juli Blasco, Nestor Ghenzi, Jordi Suae, Pablo Levy, and Enrique Miranda. Modeling of the hysteretic I-V characteristics of TiO₂-based resistive switches using the generalized diode equation. *IEEE Electron Device Lett.*, 35(3):390–392, 2014.
- [BGS⁺15] J. Blasco, N. Ghenzi, J. Suñé, P. Levy, and E. Miranda. Equivalent circuit modeling of the bistable conduction characteristics in electroformed thin dielectric films. *Microelectron. Reliab.*, 55(1):1–14, jan 2015.
- [BHAS15] Ghaith Bany Hamad, Syed Rafay Hasan, Otmane Ait Mohamed, and Yvon Savaria. Characterizing, modeling, and analyzing soft error propagation in asynchronous and synchronous digital circuits. *Microelectron. Reliab.*, 55(1):238–250, 2015.
- [BHS⁺14] William G Bennett, Nicholas C Hooten, Ronald D Schrimpf, Robert A Reed, Michael L Alles, En Xia Zhang, Stephanie L. Weeden-Wright, Dimitri Linten, Malgorzata Jurczak, and Andrea Fantini. Dynamic Modeling of Radiation-Induced State Changes in HfO₂/Hf 1T1R RRAM. *IEEE Trans. Nucl. Sci.*, 61(6):3497–3503, 2014.

- [BSH⁺12] W. G. Bennett, R. D. Schrimpf, N. C. Hooten, R. a. Reed, J. S. Kauppila, R. a. Weller, K. M. Warren, and M. H. Mendenhall. Efficient Method for Estimating the Characteristics of Radiation-Induced Current Transients. *IEEE Trans. Nucl. Sci.*, 59(6):2704–2709, dec 2012.
- [CC05] CFD Research Corporation and (CFDRC). TNanoTCAD, 2005.
- [Cho16] Hyun-Sik Choi. Main degradation mechanism in AsTeGeSiN threshold switching devices. *Microelectron. Reliab.*, 56:61–65, jan 2016.
- [Chu71] Leon O. Chua. Memristor, The Missing Circuit Element. *IEEE Trans. Circuit Theory*, 18, 1971.
- [CI15] Nicola Ciocchini and Daniele Ielmini. Pulse-Induced Crystallization in Phase-Change Memories Under Set and Disturb Conditions. *IEEE Trans. Electron Devices*, 62(3):847–854, mar 2015.
- [CKC⁺16] Simone Cortese, Ali Khiat, Daniela Carta, Mark E. Light, and Themistoklis Prodromakis. An amorphous titanium dioxide metal insulator metal selector device for resistive random access memory crossbar arrays with tunable voltage margin. *Appl. Phys. Lett.*, 108(3):033505, jan 2016.
- [CKP⁺14] Dakai Chen, Hak Kim, Anthony Phan, Edward Wilcox, Kenneth LaBel, Stephen Buchner, Ani Khachatrian, and Nicolas Roche. Single-Event Effect Performance of a Commercial Embedded ReRAM. *IEEE Trans. Nucl. Sci.*, 61(6):1–7, 2014.
- [CLC⁺15] Meng Fan Chang, Albert Lee, Pin Cheng Chen, Chrong Jung Lin, Ya Chin King, Shyh Shyuan Sheu, and Tzu Kun Ku. Challenges and circuit techniques for energy-efficient on-chip nonvolatile memory using memristive devices. *IEEE J. Emerg. Sel. Top. Circuits Syst.*, 5(2):183–193, 2015.
- [CLL⁺15] Wenbo Chen, Wenchao Lu, Branden Long, Yibo Li, David Gilmer, Gennadi Bersuker, Swarup Bhunia, and Rashmi Jha. Switching characteristics of W/Zr/HfO₂/TiN ReRAM devices for multi-level cell non-volatile memory applications. *Semicond. Sci. Technol.*, 30(7):075002, 2015.
- [CLY16] Pai-yu Chen, Zhiwei Li, and Shimeng Yu. Design Tradeoffs of Vertical RRAM-Based 3-D Cross-Point Array. *IEEE Trans. Very Large Scale Integr. Syst.*, pages 1–8, 2016.
- [CM76] Leon O. Chua and Sung Mo Kang. Memristive Devices and Systems. *Proc. IEEE*, 64(2), 1976.
- [CS11] Siddhartha Chhabra and Yan Solihin. i-NVMM: A Secure Non-Volatile Main Memory System with Incremental Encryption. In *Proceeding 38th Annu. Int. Symp. Comput. Archit. - ISCA '11*, page 177, New York, New York, USA, 2011. ACM Press.
- [CSL⁺16] Yueh-ting Chung, Po-cheng Su, Wen-jie Lin, Min-cheng Chen, and Tahui Wang. SET/RESET Cycling-Induced Trap Creation and SET-Disturb Failure Time Degradation in a Resistive-Switching Memory. *IEEE Trans. Electron Devices*, 63(6):2367–2373, jun 2016.

- [CTKP16] Simone Cortese, Maria Trapatseli, Ali Khiat, and Themistoklis Prodromakis. On the origin of resistive switching volatility in Ni/TiO₂/Ni stacks. *J. Appl. Phys.*, 120(6):065104, aug 2016.
- [CWC⁺15] Meng-Fan Chang, Jui-jen Wu, Tun-fei Chien, Yen-chen Liu, Ting-chin Yang, Wen-chao Shen, Ya-chin King, Chrong Jung Lin, Ku-feng Lin, Yu-der Chih, and Jonathan Chang. Low VDDmin Swing-Sample-and-Couple Sense Amplifier and Energy-Efficient Self-Boost-Write-Termination Scheme for Embedded ReRAM Macros Against Resistance and Switch-Time Variations. *IEEE J. Solid-State Circuits*, 50(11):2786–2795, nov 2015.
- [CY15] Pai-yu Chen and Shimeng Yu. Compact Modeling of RRAM Devices and Its Applications in 1T1R and 1S1R Array Design. *IEEE Trans. Electron Devices*, 62(12):4022–4028, dec 2015.
- [CYMMR⁺16] A. Crespo-Yepes, J. Martin-Martinez, I Rama, M. Maestro, R Rodriguez, M Nafria, X. Aymerich, M B Gonzalez, and F. Campabadal. Intra-device statistical parameters in variability-aware modelling of resistive switching devices. In *2016 Jt. Int. EUROSIO Work. Int. Conf. Ultim. Integr. Silicon*, number Figure 2, pages 84–87. IEEE, jan 2016.
- [DPCV09] Massimiliano Di Ventra, Yuriy V. Pershin, Leon O. Chua, and Massimiliano Di Ventra. Circuit Elements With Memory: Memristors, Memcapacitors, and Meminductors. *Proc. IEEE*, 97(10):1717–1724, jan 2009.
- [DSB16] Jayita Das, Kevin Scott, and Sanjukta Bhanja. MRAM PUF: Using Geometric and Resistive Variations in MRAM Cells. *ACM J. Emerg. Technol. Comput. Syst.*, 13(1):1–15, may 2016.
- [DWB⁺07] S. DasGupta, a. F. Witulski, B. L. Bhuvu, M. L. Alles, R. a. Reed, O. a. Amusan, J. R. Ahlbin, R. D. Schrimpf, and L. W. Massengill. Effect of well and substrate potential modulation on single event pulse shape in deep submicron CMOS. *IEEE Trans. Nucl. Sci.*, 54(6):2407–2412, 2007.
- [EBC⁺15] Arthur H. Edwards, Hugh J Barnaby, Kristy A Campbell, Michael N Kozicki, Wei Liu, and Matthew J Marinella. Reconfigurable Memristive Device Technologies. *Proc. IEEE*, 103(7):1004–1033, 2015.
- [EKC⁺12] Kamran Eshraghian, Omid Kavehei, Kyoung-Rok Cho, James M. Chappell, Azhar Iqbal, Said F. Al-Sarawi, and Derek Abbott. Memristive Device Fundamentals and Modeling: Applications to Circuits and Systems Simulation. *Proc. IEEE*, 100(6):1991 – 2007, 2012.
- [EM11] IE E Ebong and Pinaki Mazumder. Self-controlled writing and erasing in a memristor crossbar memory. *Nanotechnology, IEEE Trans. ...*, 10(6):1454–1463, 2011.
- [FCMG13] Veronique Ferlet-Cavrois, Lloyd W. Massengill, and Pascale Gouker. Single Event Transients in Digital CMOS—A Review. *IEEE Trans. Nucl. Sci.*, 60(3):1767–1790, jun 2013.

- [FHS15] Bahar Farahani, Seyedamin Habibi, and Saeed Safari. A cross-layer SER analysis in the presence of PVT variations. *Microelectron. Reliab.*, 55(7):1013–1027, 2015.
- [FMM⁺12] Martin Fuechsle, Jill a Miwa, Suddhasatta Mahapatra, Hoon Ryu, Sunhee Lee, Oliver Warschkow, Lloyd C L Hollenberg, Gerhard Klimeck, and Michelle Y Simmons. A single-atom transistor. *Nat. Nanotechnol.*, 7(4):242–6, 2012.
- [FYL⁺10] Z Fang, H Y Yu, W J Liu, N Singh, and G Q Lo. Resistive RAM Based on Hfo x and its Temperature Instability Study. *World Acad. Sci. Eng. Technol.*, 48(12):905–907, 2010.
- [GBBS12] Xavier Gili, Salvador Barcelo, Sebasti à A. Bota, and Jaume Segura. Analytical Modeling of Single Event Transients Propagation in Combinational Logic Gates. *IEEE Trans. Nucl. Sci.*, 59(4):971–979, aug 2012.
- [GCLY16] Ligang Gao, Pai-yu Chen, Rui Liu, and Shimeng Yu. Physical Unclonable Function Exploiting Sneak Paths in Resistive Cross-point Array. *IEEE Trans. Electron Devices*, 63(8):3109–3115, aug 2016.
- [GLmG⁺15] Amirali Ghofrani, Miguel-angel Lastras-montaña, Siddharth Gaba, Melika Payvand, Wei Lu, Luke Theogarajan, and Kwang-ting Cheng. A Low-Power Variation-Aware Adaptive Write Scheme for Access-Transistor-Free Memristive Memory. *ACM J. Emerg. Technol. Comput. Syst.*, 12(1):1–18, aug 2015.
- [GO15] Gargi Ghosh and Marius K Orlowski. Write and Erase Threshold Voltage Interdependence in Resistive Switching Memory Cells. *IEEE Trans. Electron Devices*, 62(9):2850–2856, sep 2015.
- [GR16] Fernando Garcia-Redondo. <https://github.com/fgr1986> [Fernando García-Redondo Repositories], 2016.
- [GRALV⁺16] Fernando Garcia-Redondo, Hernan Aparicio, Marisa Lopez-Vallejo, Pablo Ituero, and Carlos Lopez-Barrio. Characterization of analog modules: Reliability analyses of radiation, temperature and variations effects. In *2016 Conf. Des. Circuits Integr. Syst.*, pages 1–5. IEEE, nov 2016.
- [GRGCY⁺16] Fernando Garcia-Redondo, Robert P. Gowers, Albert Crespo-Yepes, Marisa Lopez-Vallejo, and Liudi Jiang. SPICE Compact Modeling of Bipolar/Unipolar Memristor Switching Governed by Electrical Thresholds. *IEEE Trans. Circuits Syst. I Regul. Pap.*, 63(8):1255–1264, aug 2016.
- [GRLVAI16] Fernando Garcia-Redondo, Marisa Lopez-Vallejo, Hernan Aparicio, and Pablo Ituero. Reliable design methodology: The combined effect of radiation, variability and temperature. In *2016 13th Int. Conf. Synth. Model. Anal. Simul. Methods Appl. to Circuit Des.*, pages 1–4. IEEE, jun 2016.
- [GRLVI14] Fernando Garcia-Redondo, Marisa Lopez-Vallejo, and Pablo Ituero. Building Memristor Applications: From Device Model to Circuit Design. *IEEE Trans. Nanotechnol.*, 13(6):1154–1162, nov 2014.

- [GRLVI15] Fernando Garcia-Redondo, Marisa Lopez-Vallejo, and Pablo Ituero. A thermal adaptive scheme for reliable write operation on RRAM based architectures. In *2015 33rd IEEE Int. Conf. Comput. Des.*, pages 367–374. IEEE, oct 2015.
- [GRLVIL12] Fernando Garcia-Redondo, Marisa Lopez-Vallejo, Pablo Ituero, and Carlos López Barrio. A CAD framework for the characterization and use of memristor models. In *2012 Int. Conf. Synth. Model. Anal. Simul. Methods Appl. to Circuit Des.*, pages 25–28. IEEE, sep 2012.
- [GRLVR15] Fernando García-Redondo, Marisa López-Vallejo, and Pablo Royer. <http://vlsi.die.upm.es/resources> [LSI Online CAD Resources], 2015.
- [GRR12] Fernando García-Redondo and Ricardo Riazza. The tractability index of memristive circuits: Branch-oriented and tree-based models. *Math. Methods Appl. Sci.*, 35(14):1659–1669, 2012.
- [GVR⁺16] Daniele Garbin, Elisa Vianello, Quentin Raffhay, Mourad Azzaz, Philippe Candelier, Barbara DeSalvo, Gerard Ghibaudo, and Luca Perniola. Resistive memory variability: A simplified trap-assisted tunneling model. *Solid. State. Electron.*, 115:126–132, jan 2016.
- [GWZ⁺15] Alessandro Grossi, Damian Walczyk, Cristian Zambelli, Enrique Miranda, Piero Olivo, Valeriy Stikanov, Alessandro Feriani, Jordi Sune, Gunter Schoof, Rolf Kraemer, Bernd Tillack, Alexander Fox, Thomas Schroeder, Christian Wenger, and Christian Walczyk. Impact of Intercell and Intracell Variability on Forming and Switching Parameters in RRAM Arrays. *IEEE Trans. Electron Devices*, 62(8):2502–2509, aug 2015.
- [GZD⁺11] Alessio Griffoni, Paul Zuber, Petr Dobrovolny, Philippe J. Roussel, Dimitri Linten, Michael L. Alles, Ronald D. Schrimpf, Robert a. Reed, Lloyd W. Massengill, Daisuke Kobayashi, Eddy Simoen, and Guido Groeseneken. Impact of process variability on the radiation-induced soft error of nanometer-scale SRAMs in hold and read conditions. *Proc. Eur. Conf. Radiat. its Eff. Components Syst. RADECS*, pages 195–201, 2011.
- [HAR15] G Hubert, L Artola, and D Regis. Impact of scaling on the soft error sensitivity of bulk, FDSOI and FinFET technologies due to atmospheric radiation. *Integr. VLSI J.*, 50:39–47, jun 2015.
- [HCL⁺14] Peng Huang, Bing Chen, Haitong Li, Zhe Chen, Bin Gao, Xiaoyan Liu, and Jinfeng Kang. Parameters extraction on HfO₂ based RRAM. In *2014 44th Eur. Solid State Device Res. Conf.*, volume 4, pages 250–253, 2014.
- [HCW⁺13] P. Huang, B. Chen, Y. J. Wang, F. F. Zhang, L. Shen, R. Liu, L. Zeng, G. Du, X. Zhang, B. Gao, J. F. Kang, X. Y. Liu, X. P. Wang, B. B. Weng, Y. Z. Tang, G.-Q. Lo, and D.-L. Kwong. Analytic model of endurance degradation and its practical applications for operation scheme optimization in metal oxide based RRAM. In *2013 IEEE Int. Electron Devices Meet.*, pages 22.5.1–22.5.4. IEEE, dec 2013.

- [HKJ⁺15] William A. Hubbard, Alexander Kerelsky, Grant Jasmin, Edward R White, Jared Lodico, Matthew Mecklenburg, and Brian C. Regan. Nanofilament Formation and Regeneration During Cu/Al₂O₃ Resistive Memory Switching. *Nano Lett.*, 15(6):3983–3987, jun 2015.
- [HMM13] SJ J Seok-Jin Ham, HS S Hyun-Sun Mo, and Kyeong-Sik KS S Min. Low-power VDD/3 write scheme with inversion coding circuit for complementary memristor array. *IEEE Trans. Nanotechnol.*, 12(5):851–857, sep 2013.
- [HWL⁺14] Peng Huang, Yijiao Wang, Haitong Li, Bin Gao, Bing Chen, Feifei Zhang, Lang Zeng, Gang Du, Jinfeng Kang, and Xiaoyan Liu. Analysis of the Voltage-Time Dilemma of Metal Oxide-Based RRAM and Solution Exploration of High Speed and Low Voltage AC Switching. *IEEE Trans. Nanotechnol.*, 13(6):1127–1132, nov 2014.
- [HXN⁺15] Said Hamdioui, Lei Xie, Hoang Anh Du Nguyen, Mottaqiallah Taouil, Koen Bertels, Henk Corporaal, Hailong Jiao, Francky Catthoor, Dirk Wouters, Linn Eike, and Jan van Lunteren. Memristor Based Computation-in-memory Architecture for Data-intensive Applications. *Proc. 2015 Des. Autom. Test Eur. Conf. Exhib.*, pages 1718–1725, 2015.
- [HYN⁺16] Adib Haron, Jintao Yu, Razvan Nane, Mottaqiallah Taouil, Said Hamdioui, and Koen Bertels. Parallel matrix multiplication on memristor-based computation-in-memory architecture. In *2016 Int. Conf. High Perform. Comput. Simul.*, pages 759–766. IEEE, jul 2016.
- [HZHL16] Kejie Huang, Rong Zhao, Wei He, and Yong Lian. High-Density and High-Reliability Nonvolatile Field-Programmable Gate Array With Stacked 1D2R RRAM Array. *IEEE Trans. Very Large Scale Integr. Syst.*, 24(1):139–150, jan 2016.
- [Iel11] Daniele Ielmini. Modeling the Universal Set/Reset Characteristics of Bipolar RRAM by Field- and Temperature-Driven Filament Growth. *IEEE Trans. Electron Devices*, 58(12):4309–4317, dec 2011.
- [Iel16] Daniele Ielmini. Resistive switching memories based on metal oxides: mechanisms, reliability and scaling. *Semicond. Sci. Technol.*, 31(6):063002, jun 2016.
- [IGRLV12] Pablo Ituero, Fernando Garcia-Redondo, and Marisa Lopez-Vallejo. Temperature sensor placement including routing overhead and sampling inaccuracies. In *2012 Int. Conf. Synth. Model. Anal. Simul. Methods Appl. to Circuit Des.*, pages 69–72. IEEE, sep 2012.
- [ILVLB13] Pablo Ituero, Marisa López-Vallejo, and Carlos López-Barrio. A 0.0016 mm² 0.64 nJ leakage-based CMOS temperature sensor. *Sensors (Basel)*, 13(9):12648–12662, 2013.
- [ILVMO12] Pablo Ituero, Marisa Lopez-Vallejo, Miguel Ángel Sánchez Marcos, and Carlos Gómez Osuna. Light-Weight On-Chip Monitoring Network for

- Dynamic Adaptation and Calibration. *IEEE Sens. J.*, 12(6):1736–1745, jun 2012.
- [JCE⁺10] Sung Hyun Jo, Ting Chang, Idongesit Ebong, Bhavitavya B Bhadviya, Pinaki Mazumder, and Wei Lu. Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett.*, 10(4):1297–301, apr 2010.
- [JCM12] Chul Moon Jung, Jun Myung Choi, and Kyeong Sik Min. Two-step write scheme for reducing sneak-path leakage in complementary memristor array. *IEEE Trans. Nanotechnol.*, 11(3):611–618, 2012.
- [JJL⁺12] Chul-moon Jung, Kwan-hee Jo, Eun-sub Lee, Huan Minh Vo, and Kyeong-sik Min. Zero-Sleep-Leakage Flip-Flop Circuit With Conditional-Storing Memristor Retention Latch. *IEEE Trans. Nanotechnol.*, 11(2):360–366, mar 2012.
- [JJMK10] Kwan Hee Jo, Chul Moon Jung, Kyeong Sik Min, and Sung Mo Kang. Self-adaptive write circuit for low-power and variation-tolerant memristors. *IEEE Trans. Nanotechnol.*, 9(6):675–678, 2010.
- [JKK⁺16] Doo Seok Jeong, Kyung Min Kim, Sungho Kim, Byung Joon Choi, and Cheol Seong Hwang. Memristors for Energy-Efficient New Computing Paradigms. *Adv. Electron. Mater.*, 201600090:1600090, aug 2016.
- [JKW17] Yuning Jiang, Jinfeng Kang, and Xinan Wang. RRAM-based parallel computing architecture using k-nearest neighbor classification for pattern recognition. *Sci. Rep.*, 7(March):45233, mar 2017.
- [JW09] Yogesh N. Joglekar and Stephen J. Wolf. The elusive memristor: properties of basic electrical circuits. *Eur. J. Phys.*, 30(4):22, jul 2009.
- [JWY⁺16] Zizhen Jiang, Yi Wu, Shimeng Yu, Lin Yang, Kay Song, Zia Karim, and H.-S. Philip Wong. A Compact Model for Metal–Oxide Resistive Random Access Memory With Experiment Verification. *IEEE Trans. Electron Devices*, 63(5):1884–1892, may 2016.
- [KB16] Michael N Kozicki and Hugh J Barnaby. Conductive bridging random access memory—materials, devices and applications. *Semicond. Sci. Technol.*, 31(11):113001, nov 2016.
- [KBK⁺11] A. V. Kauppila, Bharat L. Bhuva, J. S. Kauppila, Lloyd W. Massengill, and W. T. Holman. Impact of process variations on SRAM single event upsets. *IEEE Trans. Nucl. Sci.*, 58(3):834–839, 2011.
- [KCLL14] Sungho Kim, ShinHyun Choi, Jihang Lee, and Wei D Lu. Tuning Resistive Switching Characteristics of Tantalum Oxide Memristors through Si Doping. *ACS Nano*, 8(10):10262–10269, oct 2014.
- [KCS⁺16] Wonjoo Kim, Anupam Chattopadhyay, Anne Siemon, Eike Linn, Rainer Waser, and Vikas Rana. Multistate Memristive Tantalum Oxide Devices for Ternary Arithmetic. *Sci. Rep.*, 6:36652, nov 2016.

- [KFKW13] Shahar Kvatinsky, Eby G Friedman, Avinoam Kolodny, and Uri C. Weiser. TEAM: ThrEshold Adaptive Memristor Model. *IEEE Trans. Circuits Syst. I Regul. Pap.*, 60(1):211–221, jan 2013.
- [KG14] Rafal Kleczek and Pawel Grybos. Low voltage area efficient current-mode CMOS bandgap reference in deep submicron technology. In *2014 Proc. 21st Int. Conf. Mix. Des. Integr. Circuits Syst.*, number 1, pages 247–251, 2014.
- [KKR15] Ganesh Khedkar, Dhireesha Kudithipudi, and Garrett S. Rose. Power Profile Obfuscation Using Nanoscale Memristive Devices to Counter DPA Attacks. *IEEE Trans. Nanotechnol.*, 14(1):26–35, jan 2015.
- [KKSK15] Sachhindh Kannan, Naghmeh Karimi, Ozgur Sinanoglu, and Ramesh Karri. Security Vulnerabilities of Emerging Nonvolatile Main Memories and Countermeasures. *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, 34(1):2–15, jan 2015.
- [KLH⁺15] J F Kang, H T Li, P Huang, Z Chen, B Gao, X Y Liu, Z Z Jiang, and H.-S. P. Wong. Modeling and design optimization of ReRAM. In *20th Asia South Pacific Des. Autom. Conf.*, volume 1, pages 576–581, 2015.
- [CLK⁺15] Kyung Min Kim, Seung Ryul Lee, Sungho Kim, Man Chang, and Cheol Seong Hwang. Self-Limited Switching in Ta₂O₅/TaO_x Memristors Exhibiting Uniform Multilevel Changes in Resistance. *Adv. Funct. Mater.*, 25(10):1527–1534, 2015.
- [KMG⁺14] Ibrahim Kazi, Pascal Meinerzhagen, Pierre Emmanuel Gaillardon, Davide Sacchetto, Yusuf Leblebici, Andreas Burg, and Giovanni De Micheli. Energy/reliability trade-offs in low-voltage ReRAM-based non-volatile flip-flop design. *IEEE Trans. Circuits Syst. I Regul. Pap.*, 61(11):3155–3164, 2014.
- [KNC⁺15] Athanasios Kiouseloglou, Gabriele Navarro, Alessandro Cabrini, Guido Torelli, and Luca Perniola. Current pulse generator for multilevel cell programming of innovative PCM. In *2015 Int. Conf. IC Des. Technol.*, pages 1–4. IEEE, jun 2015.
- [KRFK15] Shahar Kvatinsky, Misbah Ramadan, Eby G Friedman, and Avinoam Kolodny. VTEAM: A General Model for Voltage-Controlled Memristors. *IEEE Trans. Circuits Syst. II Express Briefs*, 62(8):786–790, aug 2015.
- [KSA⁺09] Jeffrey S. Kauppila, Andrew L. Sternberg, Michael L. Alles, A. Matt Francis, Jim Holmes, Oluwale a. Amusan, and Lloyd W. Massengill. A Bias-Dependent Single-Event Compact Model Implemented Into BSIM4 and a 90 nm CMOS Process Design Kit. *IEEE Trans. Nucl. Sci.*, 56(6):3152–3157, dec 2009.
- [KYS⁺16] Kyung Min Kim, J. Joshua Yang, John Paul Strachan, Emmanuelle Merced Grafals, Ning Ge, Noraica Davila Melendez, Zhiyong Li, and R. Stanley Williams. Voltage divider effect for the improvement of variability and endurance of TaOx memristor. *Sci. Rep.*, 6(October 2015):20085, feb 2016.

- [LCDN14] Huichu Liu, Matthew Cotter, Suman Datta, and Vijaykrishnan Narayanan. Soft-error performance evaluation on emerging low power devices. *IEEE Trans. Device Mater. Reliab.*, 14(2):732–741, 2014.
- [LHG⁺16] Haitong Li, Peng Huang, Bin Gao, Xiaoyan Liu, Jinfeng Kang, and H. S. Philip Wong. Device and Circuit Interaction Analysis of Stochastic Behaviors in Cross-Point RRAM Arrays. jun 2016.
- [LHL⁺16] Chien-Chen Lin, Jui-Yu Hung, Wen-Zhang Lin, Chieh-Pu Lo, Yen-ning Chiang, Hsiang-jen Tsai, Geng-hau Yang, Ya-Chin King, Chrong Jung Lin, Tien-Fu Chen, and Meng-Fan Chang. 7.4 A 256b-wordlength ReRAM-based TCAM with 1ns search-time and 14× improvement in wordlength-energyefficiency-density product using 2.5T1R cell. In *2016 IEEE Int. Solid-State Circuits Conf.*, volume 0, pages 136–137. IEEE, jan 2016.
- [LJH⁺15] H Li, Z Jiang, P Huang, Y Wu, B Gao, X Y Liu, and J F Kang. Variation-Aware, Reliability-Emphasized Design and Optimization of RRAM Using SPICE Model. In *Des. Autom. Test Eur. Conf. Exhib. (DATE), 2015*, pages 1425–1430, 2015.
- [LLN15] Jae Sung Lee, Shinbuhm Lee, and Tae Won Noh. Resistive switching phenomena: A review of statistical physics approaches. *Appl. Phys. Rev.*, 2(3), 2015.
- [LLW⁺16] Yu-Hsuan Lin, Ming-Hsiu Lee, Jau-yi Wu, Yu-yu Lin, Feng-Ming Lee, Dai-ying Lee, Kuang-Hao Chiang, Erh-kun Lai, Tseung-yuen Tseng, and Chih-yuan Lu. A Novel Varying-bias Read Scheme for MLC and Wide Temperature Range TMO ReRAM. *IEEE Electron Device Lett.*, 3106(c):1–1, 2016.
- [LMBY15] Rui Liu, Debayan Mahalanabis, Hugh J Barnaby, and Shimeng Yu. Investigation of Single-Bit and Multiple-Bit Upsets in Oxide RRAM-Based 1T1R and Crossbar Memory Arrays. *IEEE Trans. Nucl. Sci.*, 62(5):2294–2301, oct 2015.
- [LOEGV⁺07] C. Lopez-Ongil, L. Entrena, M. Garcia-Valderas, M. Portela, M.a. Aguirre, J. Tombs, V. Baena, and F. Munoz. A Unified Environment for Fault Injection at Any Design Level Based on Emulation. *IEEE Trans. Nucl. Sci.*, 54(4):946–950, aug 2007.
- [LWHL15] Shih-Hung Lin, You-Lin Wu, Yu-Huei Hwang, and Jing-jenn Lin. Study of radiation hardness of HfO₂-based resistive switching memory at nanoscale by conductive atomic force microscopy. *Microelectron. Reliab.*, 55(11):2224–2228, nov 2015.
- [LWP⁺15] Rui Liu, Huaqiang Wu, Yachuan Pang, He Qian, and Shimeng Yu. Experimental Characterization of Physical Unclonable Function Based on 1 kb Resistive Random Access Memory Arrays. *IEEE Electron Device Lett.*, 36(12):1380–1383, dec 2015.

- [MAA⁺07] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, C.-H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neirynck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, C. Thomas, T. Troeger, P. Vandervoorn, S. Williams, and K. Zawadzki. A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100Packaging. In *2007 IEEE Int. Electron Devices Meet.*, pages 247–250. IEEE, 2007.
- [MCYC15a] Manqing Mao, Yu Cao, Shimeng Yu, and Chaitali Chakrabarti. Optimizing latency, energy, and reliability of 1T1R ReRAM through appropriate voltage settings. In *2015 33rd IEEE Int. Conf. Comput. Des.*, volume 85287, pages 359–366. IEEE, oct 2015.
- [MCYC15b] Manqing Mao, Yu Cao, Shimeng Yu, and Chaitali Chakrabarti. Programming strategies to improve energy efficiency and reliability of ReRAM memory systems. In *2015 IEEE Work. Signal Process. Syst.*, pages 1–6. IEEE, oct 2015.
- [MCYC16] Manqing Mao, Yu Cao, Shimeng Yu, and Chaitali Chakrabarti. Optimizing Latency, Energy, and Reliability of 1T1R ReRAM Through Cross-Layer Techniques. *IEEE J. Emerg. Sel. Top. Circuits Syst.*, pages 1–12, 2016.
- [ME07] David G. Mavis and Paul H. Eaton. SEU and SET Modeling and Mitigation in Deep Submicron Technologies. In *2007 IEEE Int. Reliab. Phys. Symp. Proceedings. 45th Annu.*, pages 293–305. IEEE, apr 2007.
- [MEY⁺08] A. Makihara, T. Ebihara, T. Yokose, Y. Tsuchiya, Y. Amano, H. Shindou, R. Imagawa, Y. Takahashi, and S. Kuboyama. New SET Characterization Technique Using SPICE for Fully Depleted CMOS/SOI Digital Circuitry. *IEEE Trans. Nucl. Sci.*, 55(6):2921–2927, dec 2008.
- [MFH⁺15] K. A. Morgan, J. Fan, R. Huang, L. Zhong, R. P. Gowers, L. Jiang, and C. H. de Groot. Switching kinetics of SiC resistive memory for harsh environments. *AIP Adv.*, 5(7):077121, jul 2015.
- [MHP⁺14] Katrina A Morgan, Ruomeng Huang, Kenneth Potter, Chris Shaw, William Redman-White, and C. H. De Groot. Total Dose Hardness of TiN/HfOx/TiN Resistive Random Access Memory. *IEEE Trans. Nucl. Sci.*, 61(6):2991–2996, dec 2014.
- [Mic] Microchip Atmel. <http://www.atmel.com>.
- [Mir15] E. Miranda. Compact Model for the Major and Minor Hysteretic I-V Loops in Nonlinear Memristive Devices. *IEEE Trans. Nanotechnol.*, 14(5):787–789, sep 2015.

- [MKW12] Pinaki Mazumder, Sung Mo Kang, and Rainer Waser. Memristors: Devices, Models, and Applications [Scanning the Issue]. *Proc. IEEE*, 100(6):1911–1919, jun 2012.
- [MMB⁺15] E Miranda, A Mehonic, J Blasco, J. Sune, and A J Kenyon. Multiple Diode-Like Conduction in Resistive Switching SiO_x-Based MIM Devices. *IEEE Trans. Nanotechnol.*, 14(1):15–17, 2015.
- [MMM⁺14] Chenyue Ma, Hans Jurgen Mattausch, Kazuya Matsuzawa, Seiichiro Yamaguchi, Teruhiko Hoshida, Masahiro Imade, Risho Koh, Takahiko Arakawa, and Mitiko Miura-Mattausch. Universal NBTI Compact Model for Circuit Aging Simulation under Any Stress Conditions. *IEEE Trans. Device Mater. Reliab.*, 14(3):818–825, 2014.
- [MMMD⁺15] M. Maestro, J. Martin-Martinez, J. Diaz, A. Crespo-Yepes, M.B. Gonzalez, R. Rodriguez, F. Campabadal, M. Nafria, and X. Aymerich. Analysis of Set and Reset mechanisms in Ni/HfO₂-based RRAM with fast ramped voltages. *Microelectron. Eng.*, 147:176–179, nov 2015.
- [NBS⁺08] Balaji Narasimham, Bharat L. Bhuvu, Ronald D. Schrimpf, Lloyd W. Massengill, Matthew J. Gadlage, W. Timothy Holman, Arthur F. Witulski, William H. Robinson, Jeffrey D. Black, Joseph M. Benedetto, and Paul H. Eaton. Effects of Guard Bands and Well Contacts in Mitigating Long SETs in Advanced CMOS Processes. *IEEE Trans. Nucl. Sci.*, 55(3):1708–1713, jun 2008.
- [NHL⁺17] Leibin Ni, Hantao Huang, Zichuan Liu, Rajiv V Joshi, and Hao Yu. Distributed In-Memory Computing on Binary RRAM Crossbar. *ACM J. Emerg. Technol. Comput. Syst.*, 13(3):1–18, mar 2017.
- [NUT⁺12] Hideyuki Nakamura, Taiki Uemura, Kan Takeuchi, Toshikazu Fukuda, Shigetaka Kumashiro, and Tohru Mogami. Scaling effect and circuit type dependence of neutron induced single event transient. *IEEE Int. Reliab. Phys. Symp. Proc.*, pages 1–7, 2012.
- [Oat14] Anthony S. Oates. Will reliability limit Moore’s law? In *Proc. Tech. Progr. - 2014 Int. Symp. VLSI Technol. Syst. Appl.*, pages 1–1. IEEE, apr 2014.
- [Pai13] Shimeng Yu Pai-Yu Chen. ASU RRAM Verilog-A Compact Model. 1(3):3–6, 2013.
- [PAR16] Peyman Pouyan, Esteve Amat, and Antonio Rubio. Memristive Crossbar Memory Lifetime Evaluation and Reconfiguration Strategies. *IEEE Trans. Emerg. Top. Comput.*, 6750(c):1–1, 2016.
- [PBG⁺] CR. Parthasarathy, A. Bravaix, C. Guérin, M. Denais, and V. Huard. Design-In Reliability for 90-65nm CMOS Nodes Submitted to Hot-Carriers and NBTI Degradation. In *Integr. Circuit Syst. Des. Power Timing Model. Optim. Simul.*, pages 191–200. Springer Berlin Heidelberg, Berlin, Heidelberg.

- [PDS⁺15] A. Prakash, D. Deleruyelle, J. Song, M. Bocquet, and H. Hwang. Resistance controllability and variability improvement in a TaOx-based resistive memory for multilevel storage application. *Appl. Phys. Lett.*, 106(23):233104, 2015.
- [PLA⁺16] P. E. Gaillardon, L. Amaru, A. Siemon, E. Linn, R. Waser, A. Chattopadhyay, and G. De Micheli. The Programmable Logic-in-Memory (PLiM) computer. *2016 Des. Autom. Test Eur. Conf. Exhib.*, pages 427–432, 2016.
- [PLPP16] Francesco Maria Puglisi, Luca Larcher, Andrea Padovani, and Paolo Pavan. Bipolar Resistive RAM Based on HfO₂: Physics, Compact Modeling, and Variability Control. *IEEE J. Emerg. Sel. Top. Circuits Syst.*, 6(2):171–184, jun 2016.
- [PP11] Themistoklis Prodromakis and Boon Pin Peh. A Versatile Memristor Model With Nonlinear Dopant Kinetics. *IEEE Trans. Electron Devices*, 58(9):3099–3105, 2011.
- [PPS⁺15] Amit Prakash, J-S. Park, J Song, S-J. Lim, J-H. Park, J Woo, E Cha, and Hyunsang Hwang. Multi-state resistance switching and variability analysis of HfOx based RRAM for ultra-high density memory applications. In *2015 Int. Symp. Next-Generation Electron.*, volume 27, pages 1–2. IEEE, may 2015.
- [PSB⁺09] Matthew D. Pickett, Dmitri B. Strukov, Julien L. Borghetti, J. Joshua Yang, Gregory S. Snider, Duncan R. Stewart, and R. Stanley Williams. Switching dynamics in titanium dioxide memristive devices. *J. Appl. Phys.*, 106(7):074508, 2009.
- [PV10] Yuriy V. Pershin and Massimiliano Di Ventra. Practical approach to programmable analog circuits with memristors. *IEEE Trans. Circuits Syst. I Regul. Pap.*, 57(1):1857–1864, 2010.
- [PWG⁺17] Eduardo Pérez, Christian Wenger, Alessandro Grossi, Cristian Zambelli, Piero Olivo, and Robin Roelofs. Impact of temperature on conduction mechanisms and switching parameters in HfO₂-based 1T-1R resistive random access memories devices. *J. Vac. Sci. Technol. B, Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.*, 35(1):01A103, jan 2017.
- [QBRB13] Heather M. Quinn, Dolores a. Black, William H. Robinson, and Stephen P. Buchner. Fault Simulation and Emulation Tools to Augment Radiation-Hardness Assurance Testing. *IEEE Trans. Nucl. Sci.*, 60(3):2119–2142, jun 2013.
- [Raz02] B. Razavi. *Design of Analog CMOS Integrated Circuits*. Tata Mcgrow-Hill Edition, 2002.
- [RBW⁺15] F.R. Rosa, R.M. Brum, G. Wirth, F. Kastensmidt, L. Ost, and R. Reis. Impact of dynamic voltage scaling and thermal factors on SRAM reliability. *Microelectron. Reliab.*, 2, jul 2015.
- [RC10] Ádám Rak and György Cserey. Macromodeling of the Memristor in SPICE. *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, 29(4):632–636, apr 2010.

- [RDR⁺14] Fabien Roig, L Dusseau, P Ribeiro, G Auriel, N. J-H. Roche, A Privat, J.-R. Vaille, J Boch, F. Saigne, R Marec, P Calvel, F Bezerra, R Ecoffet, and B Azais. Study and Modeling of the Impact of TID on the ATREE Response in LM124 Operational Amplifier. *IEEE Trans. Nucl. Sci.*, 61(4):1603–1610, aug 2014.
- [RFBP15] Nagarajan Raghavan, Daniel D. Frey, Michel Bosman, and Kin Leong Pey. Statistics of retention failure in the low resistance state for hafnium oxide RRAM using a Kinetic Monte Carlo approach. *Microelectron. Reliab.*, pages 2–6, jul 2015.
- [RGRLV15] Pablo Royer, Fernando Garcia-Redondo, and Marisa Lopez-Vallejo. Evolution of radiation-induced soft errors in FinFET SRAMs under process variations beyond 22nm. In *Proc. 2015 IEEE/ACM Int. Symp. Nanoscale Archit.*, pages 112–117. IEEE, jul 2015.
- [RLV14] Pablo Royer and Marisa Lopez-Vallejo. Using pMOS Pass-Gates to Boost SRAM Performance by Exploiting Strain Effects in Sub-20-nm FinFET Technologies. *IEEE Trans. Nanotechnol.*, 13(6):1226–1233, nov 2014.
- [RT10] Ricardo Riaza and Caren Tischendorf. Semistate models of electrical circuits including memristors. *Int. J. Circuit Theory Appl.*, 38, 2010.
- [RWA⁺13] R. a. Reed, R. a. Weller, A. Akkerman, J. Barak, W. Culpepper, S. Duzellier, C. Foster, M. Gaillardin, G. Hubert, T. Jordan, I. Jun, S. Koontz, F. Lei, P. McNulty, M. H. Mendenhall, M. Murat, P. Nieminen, P. O’Neill, M. Raine, B. Reddell, F. Saigne, G. Santin, L. Sihver, H. H. K. Tang, P. R. Truscott, and F. Wrobel. Anthology of the Development of Radiation Transport Tools as Applied to Single Event Effects. *IEEE Trans. Nucl. Sci.*, 60(3):1876–1911, jun 2013.
- [SBB⁺09] Blaine Stackhouse, Sal Bhimji, Chris Bostak, Dave Bradley, Brian Cherkauer, Jayen Desai, Erin Francom, Mike Gowan, Paul Gronowski, Dan Krueger, Charles Morganti, and Steve Troyer. A 65 nm 2-billion transistor quad-core itanium processor. *IEEE J. Solid-State Circuits*, 44(1):18–31, 2009.
- [SBB⁺16] P K Sarkar, S Bhattacharjee, A Barman, A Kanjilal, and A Roy. Multilevel programming in Cu/NiO y /NiO x /Pt unipolar resistive switching devices. *Nanotechnology*, 27(43):435701, oct 2016.
- [SBK15] I. Sanchez Esqueda, H J Barnaby, and M P King. Compact Modeling of Total Ionizing Dose and Aging Effects in MOS Technologies. *IEEE Trans. Nucl. Sci.*, 62(4):1501–1515, aug 2015.
- [Sim63] John G. Simmons. Generalized Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film. *J. Appl. Phys.*, 34(6):1793, 1963.
- [SKN⁺14] Sung Hyun Jo, Tanmay Kumar, Sundar Narayanan, Wei D Lu, and Hagop Nazarian. 3D-stackable crossbar resistive memory based on Field Assisted

- Superlinear Threshold (FAST) selector. In *2014 IEEE Int. Electron Devices Meet.*, number 408, pages 6.7.1–6.7.4. IEEE, dec 2014.
- [SLL87] E. Seevinck, F.J. List, and J. Lohstroh. Static-noise margin analysis of MOS SRAM cells. *IEEE J. Solid-State Circuits*, 22(5):748–754, oct 1987.
- [SMC⁺15] A Siemon, S Menzel, A Chattopadhyay, R Waser, and E Linn. In-memory adder functionality in 1S1R arrays. In *2015 IEEE Int. Symp. Circuits Syst.*, volume 1, pages 1338–1341. IEEE, may 2015.
- [SMHB06] A. L. Sternberg, L. W. Massengill, M. Hale, and B. Blalock. Single-Event Sensitivity and Hardening of a Pipelined Analog-to-Digital Converter. *IEEE Trans. Nucl. Sci.*, 53(6):3532–3538, dec 2006.
- [SN07] S.M. Sze and Kwok K. Ng. *Physics of semiconductor device*. 2007.
- [SNK15] Hiroki Sasakura, Yusuke Nishi, and Tsunenobu Kimoto. Temperature dependence of conductance in NiO-based resistive switching memory showing two modes in the forming process. *Appl. Phys. Lett.*, 107(23):233510, dec 2015.
- [SSSW08] Dmitri B Strukov, Gregory S Snider, Duncan R. Stewart, and R. Stanley Williams. The missing memristor found. *Nature*, 453(7191):80–83, may 2008.
- [SYK⁺15] Maheshwar Pd. Sah, Changju Yang, Hyongsuk Kim, Bharathwaj Muthuswamy, Jovan Jevtic, and Leon Chua. A Generic Model of Memristors With Parasitic Components. *IEEE Trans. Circuits Syst. I Regul. Pap.*, 62(3):891–898, mar 2015.
- [Syn] Synopsys TCAD. <http://www.synopsys.com>.
- [TBAS14] Gabriel Torrens, Sebastia a. Bota, Bartomeu Alorda, and Jaume Segura. An Experimental Approach to Accurate Alpha-SER Modeling and Optimization Through Design Parameters in 6T SRAM Cells for Deep-Nanometer CMOS. *IEEE Trans. Device Mater. Reliab.*, 14(4):1013–1021, dec 2014.
- [TKGD16] Xifan Tang, Gain Kim, Pierre-Emmanuel Gaillardon, and Giovanni De Micheli. A Study on the Programming Structures for RRAM-Based FPGA Architectures. *IEEE Trans. Circuits Syst. I Regul. Pap.*, 63(4):503–516, apr 2016.
- [Ver17] Verifone. <http://www.verifone.com/>, 2017.
- [VGC⁺16] Cesar Vaca, Mireia B. Gonzalez, Helena Castan, Hector Garcia, Salvador Duenas, Francesca Campabadal, Enrique Miranda, and Luis A. Bailon. Study From Cryogenic to High Temperatures of the High- and Low-Resistance-State Currents of ReRAM Ni–HfO₂–Si Capacitors. *IEEE Trans. Electron Devices*, 63(5):1877–1883, may 2016.
- [VGJM⁺14] M. A. Villena, M. B. González, F. Jiménez-Molinos, F. Campabadal, J. B. Roldán, J. Suñé, E. Romera, and E. Miranda. Simulation of thermal reset transitions in resistive switching memories including quantum effects. *J. Appl. Phys.*, 115, 2014.

- [VS16] Ioannis Vourkas and Georgios Ch. Sirakoulis. Emerging Memristor-Based Logic Circuit Design Approaches: A Review. *IEEE Circuits Syst. Mag.*, 16(3):15–30, 2016.
- [VSSH16] Ioannis Vourkas, Dimitrios Stathis, Georgios Ch Sirakoulis, and Said Hamdioui. Alternative Architectures Toward Reliable Memristive Crossbar Memories. *IEEE Trans. Very Large Scale Integr. Syst.*, 24(1):206–217, jan 2016.
- [Was16] D. Waser, R.; Ielmini. *Resistive Switching*. Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim, Germany, jun 2016.
- [Wei14] Stephen Weis. Protecting Data In-Use from Firmware and Physical Attacks. *Blackhat.Com*, 2014.
- [Wes10] David Weste, Neil H.E. and Harris. *CMOS VLSI Design: A Circuits and Systems Perspective*. 3rd edition, 2010.
- [WSW⁺08] Kevin M. Warren, Andrew L. Sternberg, Robert a. Weller, Mark P. Baze, Lloyd W. Massengill, Robert a. Reed, Marcus H. Mendenhall, and Ronald D. Schrimpf. Integrating Circuit Level Simulation and Monte-Carlo Radiation Transport Code for Single Event Upset Analysis in SEU Hardened Circuitry. *IEEE Trans. Nucl. Sci.*, 55(6):2886–2894, dec 2008.
- [WWBH⁺14] Stephanie L. Weeden-Wright, William G Bennett, Nicholas C Hooten, En Xia Zhang, Michael W. McCurdy, Michael P King, Robert A Weller, Marcus H Mendenhall, Michael L Alles, Dimitri Linten, Malgorzata Jurczak, Robin Degraeve, Andrea Fantini, Robert A Reed, Daniel M Fleetwood, and Ronald D Schrimpf. TID and Displacement Damage Resilience of 1T1R HfO₂/Hf Resistive Memories. *IEEE Trans. Nucl. Sci.*, 61(6):2972–2978, dec 2014.
- [WWS⁺11] Christian Walczyk, Damian Walczyk, Thomas Schroeder, Thomas Bertaud, Ma X Gorzata Sowinska, Mindaugas Lukosius, Mirko Fraschke, Dirk Wolansky, Bernd Tillack, Enrique Miranda, and Christian Wenger. Impact of Temperature on the Resistive Switching Behavior of Embedded HfO₂-Based RRAM Devices. *IEEE Trans. Electron Devices*, 58(9):3124–3131, 2011.
- [XZW⁺14] Kai Da Xu, Yong Hong Zhang, Lin Wang, Meng Qing Yuan, Yong Fan, William T. Joines, and Qing Huo Liu. Two Memristor SPICE Models and Their Applications in Microwave Devices. *IEEE Trans. Nanotechnol.*, 13(3):607–616, may 2014.
- [YBJH16] Kyung Jean Yoon, Woorham Bae, Deog-Kyoon Jeong, and Cheol Seong Hwang. Comprehensive Writing Margin Analysis and its Application to Stacked one Diode-One Memory Device for High-Density Crossbar Resistance Switching Random Access Memory. *Adv. Electron. Mater.*, page 1600326, sep 2016.
- [YC16] Shimeng Yu and Pai-Yu Chen. Emerging Memory Technologies: Recent Trends and Prospects. *IEEE Solid-State Circuits Mag.*, 8(2):43–56, 2016.

- [YGCR15] E. Yalon, A. Gavrilov, S. Cohen, and D. Ritter. Validation and Extension of Local Temperature Evaluation of Conductive Filaments in RRAM Devices. *IEEE Trans. Electron Devices*, 62(11):3671–3677, 2015.
- [YGG⁺12] Yang Yin Chen, Bogdan Govoreanu, Ludovic Goux, Robin Degraeve, Andrea Fantini, Gouri Sankar Kar, Dirk J. Wouters, Guido Groeseneken, Jorge A. Kittl, Malgorzata Jurczak, and Laith Altimime. Balancing SET/RESET Pulse for >10e10 Endurance in HfO₂ 1T1R Bipolar RRAM. *IEEE Trans. Electron Devices*, 59(12):3243–3249, dec 2012.
- [YLH⁺15] P. Yan, Y. Li, Y. J. Hui, S. J. Zhong, Y. X. Zhou, L. Xu, N. Liu, H. Qian, H. J. Sun, and X. S. Miao. Conducting mechanisms of forming-free TiW/Cu₂O/Cu memristive devices. *Appl. Phys. Lett.*, 107(8):083501, aug 2015.
- [YMO⁺15] Yuanfan Yang, Jimson Mathew, Marco Ottavi, Salvatore Pontarelli, and D. K. Pradhan. Novel complementary resistive switch crossbar memory write and read schemes. *IEEE Trans. Nanotechnol.*, 14(2):346–357, 2015.
- [YMT⁺14] C Yakopcic, M. McLean, T.M. Taha, R Hasan, and D Palmer. Memristor-based neuron circuit and method for applying learning algorithm in SPICE. *Electron. Lett.*, 50(7):492–494, mar 2014.
- [YPL⁺08] J. Joshua Yang, Matthew D. Pickett, Xuema Li, Douglas a a Ohlberg, Duncan R. Stewart, and R. Stanley Williams. Memristive switching mechanism for metal/oxide/metal nanodevices. *Nat. Nanotechnol.*, 3(7):429–33, jul 2008.
- [YTSP13a] Chris Yakopcic, Tarek M. Taha, Guru Subramanyam, and Robinson E. Pino. Generalized Memristive Device SPICE Model and its Application in Circuit Design. *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, 32(8):1201–1214, aug 2013.
- [YTSP13b] Chris Yakopcic, Tarek M. Taha, Guru Subramanyam, and Robinson E. Pino. Memristor SPICE model and crossbar simulation based on devices with nanosecond switching time. *Proc. Int. Jt. Conf. Neural Networks*, 2013.
- [YU15] Pai-Yu Chen Shimeng Yu and Arizona State University. ASU RRAM Verilog-A model, 2015.
- [YW11] Shimeng Yu and H. S Philip Wong. Compact modeling of conducting-bridge random-access memory (CBRAM). *IEEE Trans. Electron Devices*, 58(5):1352–1360, 2011.
- [ZGJ⁺15] Lu Zhang, Ning Ge, J. Joshua Yang, Zhiyong Li, R. Stanley Williams, and Yiran Chen. Low voltage two-state-variable memristor model of vacancy-drift resistive switches. *Appl. Phys. A*, 119(1):1–9, apr 2015.
- [ZHL16] Jinxiang Zha, He Huang, and Yujie Liu. A Novel Window Function for Memristor Model With Application in Programming Analog Circuits. *IEEE Trans. Circuits Syst. II Express Briefs*, 63(5):423–427, may 2016.

- [ZJHD14] L. Zhong, L. Jiang, R. Huang, and C. H. De Groot. Nonpolar resistive switching in Cu/SiC/Au non-volatile resistive memory devices. *Appl. Phys. Lett.*, 104(9):1–5, 2014.
- [ZSK14] Le Zheng, Sangho Shin, and Sung-mo Steve Kang. Modular Structure of Compact Model <newline/>for Memristive Devices. *IEEE Trans. Circuits Syst. I Regul. Pap.*, 61(5):1390–1399, may 2014.
- [ZYW⁺13] Xuan Zhu, Xuejun Yang, Chunqing Wu, Nong Xiao, Junjie Wu, and Xun Yi. Performing Stateful Logic on Memristor Memory. *IEEE Trans. Circuits Syst. II Express Briefs*, 60(10):682–686, oct 2013.

*On Jubilee Street there was a girl named Bee
She had a history, but she had no past...*