The design and analysis of novel integrated phase-change photonic memory and computing devices

Submitted by Emanuele Gemo to the University of Exeter
as a thesis for the degree of Doctor of Philosophy
in Physics and Engineering
In February 2021

Supervisors:
Prof. C. David Wright, University of Exeter, UK
Dr Anna Baldycheva, University of Exeter, UK
Prof. H. Bhaskaran, University of Oxford, UK

Examiners:
Prof. M. Aziz, University of Exeter, UK
Prof. L. Pavesi, University of Trento, IT

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ....

Abstract

The current massive growth in data generation and communication challenges traditional computing and storage paradigms. The integrated silicon photonic platform may alleviate the physical limitations resulting from the use of electrical interconnects and the conventional von Neumann computing architecture, due to its intrinsic energy and bandwidth advantages. This work focuses on the development of the phase-change all-photonic memory (PPCM), a device potentially enabling the transition from the electrical to the optical domain by providing the (previously unavailable) non-volatile all-photonic storage functionality. PPCM devices allow for all-optical encoding of the information on the crystal fraction of a waveguide-implemented phase-change material layer, here $Ge_2Sb_2Te_5$, which in turn modulates the transmitted signal amplitude.

This thesis reports novel developments of the numerical methods necessary to emulate the physics of PPCM device operation and performance characteristics, illustrating solutions enabling the realization of a simulation framework modelling the inherently three-dimensional and self-influencing optical, thermal and phase-switching behaviour of PPCM devices. This thesis also depicts an innovative, fast and cost-effective method to characterise the key optical properties of phase-change materials (upon which the performance of PPCM devices depend), exploiting the reflection pattern of a purposely built layer stack, combined with a smart fit algorithm adapting potential solutions drawn from the scientific literature.

The simulation framework developed in the thesis is used to analyse reported PPCM experimental results. Numerous sources of uncertainty are underlined, whose systematic analysis reduced to the peculiar non-linear optical properties of $Ge_2Sb_2Te_5$. Yet, the data fit process validates both the simulation tool and the remaining physical assumptions, as the model captures the key aspects of the PPCM at high optical intensity, and reliably and accurately predicts its behaviour at low intensity, enabling to investigate its underpinning physical mechanisms.

Finally, a novel PPCM memory architecture, exploiting the interaction of a much-reduced $Ge_2Sb_2Te_5$ volume with a plasmonic resonant nanoantenna, is proposed and numerically investigated. The architecture concept is described and the memory functionality is demonstrated, underlining its potential energy and speed improvement on the conventional device by up to two orders of magnitude.

List of Publications

Journal papers

1. Nathan Youngblood, Carlos Ríos, Emanuele Gemo, Johannes Feldmann, Zengguang Cheng, Anna Baldycheva, Wolfram H. P. Pernice, C. David Wright, and Harish Bhaskaran. "Tunable volatility of $Ge_2Sb_2Te_5$ in integrated photonics," *Adv. Funct. Mater.*, vol. 29, no. 11, p. 1807571, Mar. 2019.
2. Emanuele Gemo, Santiago Garcia-Cuevas Carrillo, Carlota Ruíz De Galarreta, Anna Baldycheva, Hasan Hayat, Nathan Youngblood, Harish Bhaskaran, Wolfram H. P. Pernice, and C. David Wright. "Plasmonically-enhanced all-optical integrated phase-change memory," *Opt. Express*, vol. 27, no. 17, p. 24724, Aug. 2019.
3. Santiago G.-C. Carrillo, Emanuele Gemo, Xuan Li, Nathan Youngblood, Andrew Katumba, Peter Bienstman, Wolfram H. P. Pernice, Harish Bhaskaran, and C. David Wright. “Behavioral modeling of integrated phase-change photonic devices for neuromorphic computing applications,” *APL Mater.*, vol. 7, no. 9, p. 091113, Sep. 2019.
4. Xuan Li, Nathan Youngblood, Zengguang Cheng, Santiago G.-C. Carrillo, Emanuele Gemo, Wolfram H. P. Pernice, C. David Wright, and Harish Bhaskaran. “Experimental investigation of silicon and silicon nitride platforms for phase-change photonic in-memory computing,” *Optica*, vol. 7, no. 3, p. 218, Mar. 2020.
5. Carlota Ruíz de Galarreta, Ivan Sinev, Arseny M. Alexeev, Pavel Trofimov, Konstantin Ladutenko, Santiago G.-C. Carrillo, Emanuele Gemo, Anna Baldycheva, Jacopo Bertolotti, and C. David Wright. “Reconfigurable multilevel control of hybrid all-dielectric phase-change metasurfaces,” *Optica*, vol. 7, no. 5, p. 476, May 2020.
6. Emanuele Gemo, Samir V. Kesava, Carlota Ruíz De Galarreta, Liam Trimby, Santiago G.-C. Carrillo, Moritz Riede, Anna Baldycheva, Arseny M. Alexeev, and C. David Wright. “Simple technique for determining the refractive index of phase-change materials using near-infrared reflectometry,” *Opt. Mater. Express*, vol. 10, no. 7, p. 1675, Jul. 2020.

7. Joaquin Faneca, Santiago G.-C. Carrillo, Emanuele Gemo, Carlota Ruíz de Galarreta, Thalía Domínguez Bucio, Frederic Y. Gardes, Harish Bhaskaran, Wolfram H. P. Pernice, C. David Wright, and Anna Baldycheva. “Performance characteristics of phase-change integrated silicon nitride photonic devices in the O and C telecommunications bands,” *Opt. Mater. Express*, vol. 10, no. 8, p. 1778, Aug. 2020.
8. Carlota Ruíz de Galarreta, Santiago G.-C. Carrillo, Yat-Yin Au, Emanuele Gemo, Liam Trimby, Joe Shields, Euan Humphreys, Joaquin Faneca, Lu Cai, Anna Baldycheva, Jacopo Bertolotti, and C. David Wright. “Tunable optical metasurfaces enabled by chalcogenide phase-change materials: from the visible to the THz,” *J. Opt.*, vol. 22, no. 11, p. 114001, Nov. 2020.
9. Emanuele Gemo, Joaquin Faneca, Santiago G.-C. Carrillo, Anna Baldycheva, Wolfram H. P. Pernice, Harish Bhaskaran, and C. David Wright. "(Invited) A plasmonically-enhanced route to faster and more energy-efficient phase-change integrated photonic memory and computing devices,” *J. Appl. Phys.*, vol 129, no. 11, p.110902, Mar. 2021.
10. Santiago G.-C. Carrillo, Alessio Lugnan, Emanuele Gemo, Peter Bienstman, Wolfram H. P. Pernice, Harish Bhaskaran, and C. David Wright. “System-level simulation for integrated phase-change photonics,” *J. Light. Technol*, 2021, *under peer review*.

Conference papers and presentations

1. Arseny M. Alexeev, Carlota Ruíz De Galarreta, Santiago G.-C. Carrillo, Ivan S. Sinev, Anton K. Samusev, Emanuele Gemo, V. Karthik Nagareddy, Yat-Yin Au, and C. David Wright. “Tunable Dielectric Metadevices Enabled by Phase-Change Materials,” in *Proceedings of E\PCOS Conference*, 2017, pp. 39–40.
2. C. David Wright, Santiago G.-C. Carrillo, Carlota Ruíz De Galarreta, Emanuele Gemo, Liam Trimby, Arseny. M. Alexeev, Yat-Yin Au, V. Karthik Nagareddy, Anna Baldycheva, Jacopo Bertolotti, and Martin Lopez Garcia. "Phase-Change Metadevices for the Dynamic and Reconfigurable Control of Light." In *Novel Optical Materials and Applications*, 2018, pp. NoTh2D-4, Optical Society of America.

3. Emanuele Gemo, Nathan Youngblood, Zengguang Cheng, Carlos Ríos, Matthias Stegmaier, Anna Baldycheva, Wolfram H. P. Pernice, Harish Bhaskaran and C. David Wright. “Modelling phase-change integrated photonic devices,” in *Proceedings of E\PCOS Conference*, 2018, pp. 45–46. (Oral presentation)
4. Emanuele Gemo, Santiago G.-C. Carrillo, Nathan Youngblood, Zengguang Cheng, Carlos Ríos, Matthias Stegmaier, Anna Baldycheva, Wolfram H. P. Pernice, Harish Bhaskaran and C. David Wright. “Simulation study of on-chip phase-change all-optical memories,” in *Proceedings of SNAIA Conference*, 2018, pp. 59–60. (Oral presentation)
5. C. David Wright, Harish Bhaskaran, Wolfram H. P. Pernice, Santiago G.-C. Carrillo, Emanuele Gemo, Anna Baldycheva, Zengguang Cheng, Xuan Li, Carlos Ríos, Nathan Youngblood, Johannes Feldmann, Nico Gruhler, and Matthias Stegmaier. “Integrated Phase-change Photonics: A Strategy for Merging Communication and Computing.” In *Proceedings of Optical Fiber Communication Conference*, 2019, pp. M1D-3, Optical Society of America.
6. Emanuele Gemo, Santiago G.-C. Carrillo, Joaquin Faneca, Nathan Youngblood, Anna Baldycheva, Harish Bhaskaran, Wolfram H. P. Pernice, and C. David Wright. “A plasmonic route towards the energy scaling of on-chip integrated all-photonic phase-change memories,” in *Proceedings of E\PCOS Conference*, 2019, pp. 154–155. (Oral presentation)
7. Santiago G.-C. Carrillo, Emanuele Gemo, Nathan Youngblood, Xuan Li, Andrew Katumba, Peter Bientsman, Wolfram H. P. Pernice, Harish Bhaskaran, and C. David Wright. “A behavioural model for integrated phase-change photonics devices,” in *Proceedings of E\PCOS Conference*, 2019, pp. 154–155.
8. Gemo, Emanuele, Santiago G.-C. Carrillo, Joaquin Faneca, Carlota Ruíz de Galarreta, Hasan Hayat, Nathan Youngblood, Anna Baldycheva, Wolfram H. P. Pernice, Harish Bhaskaran, and C. David Wright. "Sub-wavelength plasmonic-enhanced phase-change memory." In *Photonic and Phononic Properties of Engineered Nanostructures X*, 2020, vol. 11289, p. 112891E, International Society for Optics and Photonics. (Oral presentation)

Additional research-based activities and achievements

- Feb. 2017, Oxford – Research visit at Bhaskaran Laboratories, Department of Materials – University of Oxford
- Nov. 2017, Exeter – Infrared sensing and imaging: Challenges for next generation materials and devices (poster presentation: "On-chip All-Photonic Phase-Change Memories")
- Apr. 2018, Exeter – RMS Photonic and Optoelectronic Materials Conference (poster presentation: "Investigation of Optical and Thermal Properties for On-Chip Phase-Change Material Photonic Memory applications")
- Apr. 2018, Grasmere - Lake District – 2nd International Symposium on Doped Amorphous Chalcogenides and Devices (oral presentation: "Simulation of on-chip integrated phase-change all-photonic memories")
- Jul. 2019, Pisa – ePIXfab Silicon Photonics Summer school,4th Edition (poster presentation, "On-Chip Integrated Optical Phase-Change-Memories")
- Jul. 2019, Exeter – IONS 2019 conference (poster presentation: "Integrated phase-change all-photonic memory: a simulation study"; *Best Poster Award, 1st Prize*)
- Aug. 2019, Oxford – 1st Conversations in Photonics: Future of Photonic Computing (poster presentation: "Sub-wavelength plasmonic-enhanced phase-change memory")
- Oct. 2019, Exeter – Exeter-Bath Knowledge Transfer Conference (oral presentation: "Plasmonic nanoantenna enabled - extreme volume reduction of photonic phase-change memories")
- Dec. 2019, Paris – SNAIA (*Smart Nanomaterials: Advances, Innovation and Application*) 2019 Conference, organization tasks

Acknowledgements

I’ve spent four years at the University of Exeter, and a good chunk of the work there carried out is here presented, as a culmination of many trials and a few triumphs. But none of it would have been possible without the help of a long list of people, to whom I owe my most sincere gratitude. I’d like to thank them all, without any specific order, as each contribution is as important as the others. But, being writing a purely sequential form of communication, I can’t do otherwise, as I presume that sticking here a nice drawing would hardly pass as a proper way to express appreciation, especially in this context. So here I am.

David and Anna, my supervisors, who not a single time let a word out about my incompetence, and continuously supported, advised, taught and guided me along these four years. Your dedication made me a better researcher, and a better person. Working with you has been a privilege I’ll always be thankful for. I also thank David for the opportunity to participate in the Fun-COMP research project, and to work with the collaborators thereof (particularly Harish, Wolfram, Abu, Peter and their research groups). It has been an extremely interesting and formative experience, for which I am truly grateful.

Hasan, Santiago and Carlota are the members of David’s research group with whom I worked closely since my beginning. I’d like to put in writing how much I looked up to you. You are humble, profoundly knowledgeable and, foremost, always there whenever I have that one stupid question bothering me. Most of what I know today is what you taught me: this work is yours as well.

I’m also indebted to Dr Hooper, Prof. Aziz, Prof. Hrkac, and Dr Papatheou, for sharing their time and knowledge to help me go through some of the theoretical and technical challenges of this research.

My sincere appreciation goes to the EPSRC CDT in Metamaterials, the CDT Directors (and to Rosie as well!), for actually granting me the possibility to work here.
And to Tiziana and Giovanni, my M.Sc. supervisors, who very much helped me going through the Master’s Degree to get here.

Joaquin, Liam, Zahid, Kunyapat, William the English Scot, the past and new members of David’s group (Arseny, Karthik, Toby B, Toby O, Joe S, Euan, Joe P), and a few more: those are also people to which my thanks are very much due. For the many moments spent together talking about science, exchanging ideas, pretending to philosophise, blabbering nonsense and drinking a beer together – thanks.

To my family, and closest friends (Enrico, Giampietro, Nicola, Mirko, Sara, Stefano, Valentina) for being always there, no matter what distance there might be (and to the internet for making this distance less daunting). To my parents, for being the only real-life heroes I believe in. And to the person without whom I would not have even dared to walk down this path, Daria (you’ll always be the one hand I reach out for). To you all, thank you.

Index

Frontpage and declaration of authorship i

Abstract iii

List of Publications v

  Journal papers v

  Conference papers and presentations vi

  Additional research-based activities and achievements viii

Acknowledgements ix

Index xi

List of Abbreviations xix

List of Figures xxi

List of Tables xxxix

Chapter 1. Introduction and Motivation 1

  1.1. Data storage and processing challenges 1

  1.2. Information storage technology 4

  1.3. Phase-change materials and phase-change memory devices 9

  1.4. N-vN phase-change computational devices 12

  1.5. Thesis objectives 14

  1.6. Thesis outline 15

Chapter 2. Theoretical background 17

  2.1. Physical mechanisms in integrated photonics platforms 17

    2.1.1. Electromagnetic wave propagation and optical waveguide design 17

    2.1.2. Material optical properties 23

    2.1.3. Plasmonic effects 27

  2.2. Thermal diffusion and thermal boundary resistance 29

  2.3. Phase-change materials 32

    2.3.1. Phase transition mechanisms 33

2.3.1.1. Homogeneous nucleation.. 35

2.3.1.2. Heterogeneous nucleation 37

2.3.1.3. Growth mechanism 39

2.3.1.4. Viscosity in phase-change materials 41

2.3.2. A short overview of chalcogenides phase-change alloys............. 43

2.3.3. $Ge_2Sb_2Te_5$ phase-change material 45

2.4. Integrated all-photonic memory 52

2.4.1. Memory configuration 53

2.4.2. Information storage and retrieval 54

2.4.3. Additional functionalities 56

2.4.4. Computation with PC photonic memory devices 57

Chapter 3. Experimental methodologies and computational techniques 61

3.1. Experimental techniques for photonic device test and characterization 61

3.1.1. Integrated photonic chip test and characterization........................ 61

3.1.2. Characterization techniques for optical material properties 63

3.2. Finite Element modelling 65

3.2.1. COMSOL environment 66

3.2.2. Approximations, initial conditions and boundary conditions.......... 67

3.2.3. Meshing techniques.. 69

3.3. Modelling of the phase-switching mechanism 71

3.3.1. JMAK model 72

3.3.2. Classical Nucleation-Growth (CNG) model................................... 74

3.3.3. Master Rate equation model 76

3.3.4. Gillespie Cellular Automata (GCA) model 77

Chapter 4. Advanced modelling development 83

4.1. 3D phase-change modelling 83

4.1.1. 3D cellular automata 84

4.1.1.1. Explicit unidimensional coordinate system 86

4.1.1.2. Inverse Transform Sampling algorithm enhancement 88

4.1.2. 3D Gillespie Cellular Automata 90

4.1.2.1. Degenerate case: 2D cellular automata 90

4.1.2.2. Comparison between 2D and 3D model outputs 93

4.1.2.3. Computational requirement of the GCA 3D model 96

4.1.3. 3D Cellular Automata based on the Classical Nucleation-Growth 98

4.1.3.1. From rates to probability: definition of the CNG cellular automata core equation set 99

4.1.3.2. Dissociation rate and probability 100

4.1.3.3. Algorithmic steps of the CNG cellular automata 101

4.1.3.4. CNG cellular automata: simulation results 102

4.1.3.5. Comparison of 3D CNG and GCA models 105

4.1.3.6. Discussion and conclusion 107

4.2. Integration of FEM software and phase-change modelling 109

4.2.1. COMSOL framework 109

4.2.1.1. Interfaces 110

4.2.1.2. Geometry 110

4.2.1.3. Mesh 111

4.2.1.4. Physics and features 111

4.2.1.5. COMSOL Study setup 113

4.2.2. MATLAB framework 114

4.2.2.1. Time-dependent user-provided variables 115

4.2.2.2. Data exchange between COMSOL and PC models 115

4.2.2.3. Adaptive time-step determination method 116

4.3. Conclusion 119

4.3.1. Note on the applicability and suitability of the 3D-CNG model.... 120

Chapter 5. A novel method to estimate the complex index parameters from reflection measurements 123

5.1. Why is a quick method for the complex index estimation needed? ... 123

5.2. Technique description 124

5.2.1. Theoretical elements and uniqueness of the solution in FP resonance condition 124

5.2.2. Methodology 127

5.2.3. Numerical solver 128

5.3. Experimental confirmation 131

5.3.1. $Ge_2Sb_2Te_5$ 132

5.3.1.1. Amorphous GST 132

5.3.1.2. Crystalline GST 133

5.3.1.3. Comparison of the crystal and amorphous phases fit results 134

5.3.2. GeTe 135

5.3.2.1. Amorphous GeTe 135

5.3.2.2. Crystalline GeTe 136

5.3.3. Discussion 137

5.4. Conclusion 138

Chapter 6. Understanding the Read and Write processes in the conventional phase-change photonic memory architecture 139

6.1. Phase-change photonic memory: state of the research 139

6.2. Simulation of the Read operation 140

6.2.1. PPCM configuration 141

6.2.2. Crystalline unit cell simulation results and analysis 141

6.2.2.1. Influence of the PPCM geometry 143

6.2.2.2. Influence of the optical properties 143

6.2.2.3. Influence of the unit cell thickness 146

6.2.3. Adjusted simulation: results and conclusion 146

6.3. Simulation of the Write operation 148

6.3.1. Read operation and confirmation of the model reliability 149

6.3.2. Temperature-dependent optical properties................................ 150

6.3.3. Material properties for the thermal model 150

6.3.4. Crystalline unit cell: optical heating and cooling simulations....... 151

6.3.4.1. Simulation results: thermal parameters influence................. 152

6.3.4.2. Simulation results: absence of thermal boundary resistance 154

6.3.4.3. Analysis of the optical heating simulation results 155

6.3.5. Full simulation of Write and Erase processes............................ 156

6.3.5.1. Numerical experiment description....................................... 157

6.3.5.2. Results and analysis 157

6.4. Effect of high-energy probing: tunable volatility................................ 160

6.4.1. PPCM configuration.. 160

6.4.2. Unit cell: amorphous mark definition.. 161

6.4.3. Simulation results 161

6.4.4. Simulation data analysis.. 162

6.5. Discussion of simulation results and conclusion 164

6.5.1. Hypothesis: non-linear optical effects 164

6.5.2. Conclusion.. 167

Chapter 7. A novel architecture: plasmonically-enhanced unit cell.... 169

7.1. Introduction: use of optical resonant structures................................ 169

7.1.1. Dielectric resonant structures 169

7.1.2. Hybrid configuration: implementation of metallic nanoantennas. 170

7.2. Investigation of the plasmonic device concept optical behaviour 171

7.2.1. Test device: single bar nanoantenna... 172

7.2.2. Test device: dimer bar nanoantenna 174

7.2.3. Nanoantenna cross-section and nanogap sizes 175
7.2.3.1. Gap distance 176
7.2.3.2. Bars’ thickness 176
7.2.3.3. Bars’ width.. 177
7.2.3.4. Results analysis and discussion.. 178
7.2.4. Effect of a GST inclusion 179
7.3. Nanoantenna design optimization 181
7.3.1. Alternative nanoantenna geometries 182
7.3.1.1. Bowtie nanoantenna 182
7.3.1.2. Dimer-disc nanoantenna 184
7.3.2. Encapsulated configuration 185
7.3.3. Waveguide-embedded configuration 187
7.3.4. Alternative nanoantennas materials.. 190
7.4. Investigation of the memory functionality 192
7.4.1. Device design and Read operation.. 192
7.4.2. Write operation 194
7.4.3. Erase operation 197
7.4.4. Multi-level storage... 200
7.4.5. Performance comparison with the conventional architecture...... 202
7.5. Summary.. 204
Chapter 8. Final Remarks 207
8.1. Conclusion 207
8.2. Outlook... 209
8.2.1. Development of a geometrically-optimized, tau-leaping, GCA model 209
8.2.2. GST temperature- and intensity-dependent optical properties ... 210
8.2.3. Experimental verification of the plasmonically-enhanced unit cell 211

8.2.4. Adding the mixed-mode functionality to the all-photonic plasmonically-enhanced phase-change memory 212

Appendices.. 213

Appendix I. PDE formalism and solving methods 213

I.1. Theoretical elements of the finite element method 213

I.2. Equation system solvers 215

Appendix II. Pointer matrix generation and 3D cellular automata implementation 217

II.1. Generation of the Pointer matrix 217

II.2. 3D cellular automata framework example 219

II.3. Periodic boundary conditions 220

II.4. 3D framework: unit cell relative distance 220

Appendix III. Artificial reference method 221

III.1. Method description.. 221

III.2. Results and discussion.. 223

Appendix IV. Use of multiple nanostructures along the waveguide 225

IV.1. Double nanoantennas.. 225

IV.2. Multiple nanoantennas... 227

IV.3. Conclusion 229

Appendix V. Embedded bow-tie and dimer-discs plasmonic nanoantennas231

Appendix VI. Fabrication tolerances of the dimer-discs device................... 233

Appendix VII. Insertion of an anti-diffusion barrier in the dimer-discs device 235

References... 237

List of Abbreviations

AFM : atomic force microscopy 131

CDF : cumulative distribution function 79

CFL : Courant-Friedrichs-Lewy 70

CNG : classical nucleation-growth 74

DNN : deep neural network 3

DOF : degrees of freedom 112

EF : enhancement factor 172

e.m. : electromagnetic 23

FEM : finite element method 65

FOM : figure-of-merit 197

FP : Fabry-Pérot 124

GCA : Gillespie - cellular automata 77

GST : $Ge_2Sb_2Te_5$ 32

GT : Gíres-Tournois 124

IC : integrated circuits 2

IL : insertion loss 53

ITO : indium tin oxide 53

ITS : inverse transform sampling 79

JMAK : Johnson-Mehl-Avrami-Kolmogorov 72

LSPR : localised surface plasmon resonance 29

ML : multi-level 200

MRE : master rate equation 76

MSE : mean squared error 127

MZI : Mach–Zehnder interferometer 144

NLO : non-linear optical 50

N-vN : non-von Neumann 3

ODE : ordinary differential equation 76

PC : phase-change 9

PCM : phase-change memory 9

PDE : partial differential equation 67

PEC : perfect electric conductor 68

PIC : photonic integrated circuits 2

PPCM : all-photonic phase-change memory 61

RAM : random access memory 7

R.I.U. : refractive index unit 50

RSA : reverse saturable absorption 165

SA : saturable absorption 165

SEM : scanning electron microscopy 53

TBR : thermal boundary resistance 31

TE : transverse electric 21

TM : transverse magnetic 21

TMM : transfer matrix method 127

vN : von Neumann 1

WDDM : wave division de-multiplexing 2

WDM : wave division multiplexing 2

List of Figures

Figure 1.1. Memory hierarchy, based on latency time, showing the separation between typology of storage. Starting from the top (register class), each class offers a generally higher capacity – lower cost-per-bit, at the expense of higher latency and lower bit-rate. Features such as volatility, capacity, bit-rate and commonly found technology are also reported. Numerical data retrieved from Ref. [23].. 6

Figure 1.2. Ovonic EEPROM device, reprinted from [40]. a) Schematics of the memory device as reported on the patent application. The original caption reads: ”Cross-sectional view of a memory element of the present invention having a first and second contact each adjoining the volume of memory material”. b) Provided captions for the indices in (a).. 11

Figure 1.3. Number of scientific publications (per year) inherent to the hardware implementation of neuromorphic computing concepts, showing the exponential growth of such interest since the definition and dissemination of the concept from Carver Mead [12] (Data retrieved from Ref. [47]). 12

Figure 1.4. Concept of the phase-change integrated photonic memory (reprinted with permission from [67]). Original caption: “Information is stored in the phase state of the GST section on top of the nanophotonic waveguide. Both reading and writing of the memory can be performed with ultrashort optical pulses because the guided light interacts with the GST via its evanescent field. In the readout, data are encoded in the amount of optical transmission through (along) the waveguide because the two crystallographic states of the GST exhibit a high contrast in optical absorption”... 14

Figure 2.1. Examples of integrated platform waveguide configurations. To provide the necessary visual contrast, the substrate is coloured in green, and the core is presented in grey. The cladding is not portrayed.. 17

Figure 2.2. TE and TM modes profiles of a SiN-on-silica rib waveguide, optimized for C-Band transmission (width of 1300 nm, height of 330 nm, etch step of 165 nm). TE mode solutions are reported on the left (a,c,e); TM mode solutions are reported on the right (b,d,f). (a,b) 2D representation of the transversal electric field solution (normalized), as calculated with finite-element methods (colour bar reported on the right). (c,d) 1D plot of the solution, as extrapolated at the waveguide mid-hight cut-line (represented in (a) as “x cut-line”). (e,f) 1D plot of the solution, as extrapolated at the waveguide vertical symmetry axis (represented in (a) as “y cut-line”). These also include the exponential fit of the field components outside the waveguides (black dashed lines). 22

Figure 2.3. Representation of the nucleus - electron schematics, upon which the Drude-Lorentz model is based. 24

Figure 2.4. Electric permittivity and fit models. a) SiO_2: experimental data from [81] (green dots), fit model adapted from [82] (continuous lines). b) Ag: experimental data from [83] (green dots); model from [84] (continuous lines).. 26

Figure 2.5. SPP mechanism schematics. a) prospect of the metal-insulator interface, and electric field lines (red: TM input field, blue: SPP field). b) Plot of the field magnitude as a function of the vertical position, showing the relative higher magnitude on the insulator side (arising from the field continuity condition when $\varepsilon 2' > |\varepsilon 1'|$) and the exponential decay trend... 28

Figure 2.6. Visual representation of the nano-scale defects at the interface between solids, here depicted in two distinct colours (left: macroscopic picture, right: nano-scale representation), giving rise to the localized decrease of the thermal conductivity phenomenologically described as thermal boundary resistance (TBR). 31

Figure 2.7. Phase transition process. a) Phase fraction f (blue, continuous line) and temperature T (red, dashed line) as a function of the provided energy Q. Notably, temperature remains constant at the transition temperature Ttr for the duration of the process, where instead f shifts linearly from 0 (*A* phase) to 1 (*B* phase). b) Phase diagram for an XY binary compound (stoichiometry vs temperature plot). An example composition ($\chi = 0.66$) is highlighted, whose transition from the A phase (solid) towards the B phase (liquid) takes place via the intermediate mixture of Y-rich solid phase and X-rich liquid phase, when between the liquidus line (orange) and solidus line (green). 34

Figure 2.8. Free energy difference plots for a nucleation event, as a function of the spherical cluster radius. The black line depicts the contributions as appearing in Eq. 2.38, whereas the red line depicts the free energy variation during the nucleation process (i.e. increase of the spherical cluster radius), and which identifies a critical radius rc and critical energy $\Delta G(rc)$. 37

Figure 2.9. Heterogeneous nucleation, taking place on a grain boundary. The angle θ results from the equilibrium condition found in the release of the surface energy σAA and in the formation of the *AB* surfaces with energy σAB, thus leading to the formation of a distorted spherical shape (lenticular). 38

Figure 2.10. Comparison between the Gibbs energy variation resulting from the homogeneous and heterogeneous nucleation mechanisms (with a favourable surface energy balance, i.e. sub-unit shape factor). It is visible how the energy barrier is rescaled in the heterogeneous mechanism, but the related critical radius value does not vary. 39

Figure 2.11. Angell plot of the viscosity η as a function of the temperature T, for four different values of the fragility parameter m. Reprinted from [95]. 42

Figure 2.12. $Ge_xSb_yTe_z$ stoichiometry diagram, reporting some of the available compositions and their variation of the complex index at the wavelength $\lambda = 405$ nm. Reprinted from [119] 45

Figure 2.13. T-T-T diagram for $Ge_2Sb_2Te_5$ (adapted from [150]). 4 GST phases (amorphous, crystalline fcc, crystalline hexagonal and liquid) are here depicted. A 5 minutes vertical line and a 10 years vertical line are also reported, to highlight the annealing process conventionally carried out in lab conditions and the amorphous GST long-time phase retention. 46

Figure 2.14. Amorphization and recrystallization processes schematic. a) Recrystallization from the amorphous phase, via a long-duration, relatively low-power, pulse. The temperature increase is steadily increasing throughout the pulse duration, allowing to reach and surpass $Tcryst$ but remaining below Tm, thus triggering the crystallization process. b) Melt-quench amorphization process. The relatively high power pulse, of short duration, allows increasing the PC material temperature above melting point. At the pulse withdrawal, the temperature rapidly decays, quenching the molten phase into the amorphous matrix. 47

Figure 2.15. GST n and κ optical properties (a, c: refractive index; b, d: extinction coefficient), for amorphous (a,b) and crystalline (c,d) phases. To provide an intuitive picture of the variability of the GST optical properties depending on the composition, a few other datasets have been reported from the works of Gholipour [153], Park [154] and Shportko [155]. 48

Figure 2.16. GST optical properties as a function of the crystal fraction. a) Refractive index; b) extinction coefficient. The crystal fraction χ value is color-coded for each data plot (colour bar at the centre). The calculation has been performed via Eq. 2.54, from the experimentally determined data reported in Figure 2.15 (black line). 49

Figure 2.17. 2D plot of temperature- and crystal fraction-dependent GST optical properties at 1550 nm. Data is derived from Eq. 2.54 and Eq. 2.55, assuming crystal-like optical properties for the liquid phase. a) Refractive index, n (ranging from 3.97 to 6.11). b) Extinction coefficient, κ (ranging from 0.045 to 1.96). 51

Figure 2.18. Static electrical conductivity of GST. a) Dependency of the conductivity on temperature and phase, from various references (reprinted with permission from [169]). Red lines correspond to the amorphous and liquid phase; green lines to the fcc phase; blue lines to the hexagonal phase. b) Experimental IV plot of amorphous GST in a mushroom cell (reprinted with permission from [170]). Circles represent the data collected during the current increase; squares represent the data collected during the current decrease. 52

Figure 2.19. Phase-change photonic memory standard configuration (reprinted with permission from [67]). a) False-coloured *scanning electron microscopy* (SEM) image of the unit cell, fabricated on top of a rib waveguide. b) Cross-sectional representation of the device, with a superimposed simulated waveguide electric field TE mode profile. 53

Figure 2.20. Multi-level storage demonstration on a PC photonic memory conventional configuration. a) 8-level (3 bit) memory addressing on a 5 µm unit cell, by independently optimized pulsing schemes (reprinted with permission from [67]). b) 34 levels (>5 bit) storage, achieved by double-step programming pulse (shown in inset) on a 4 µm unit cell (reprinted with permission from [68]). 55

Figure 2.21. False-coloured SEM images of a) micro-ring implemented and b) 1D-photonic cavity implemented variants of the PC photonic memory conventional configuration, aiming for added functionality and increased energy performance. (a) is reprinted with permission from [67]. (b) is reprinted with permission from [174]. 56

Figure 2.22. Additional variants of the PC photonic memory device conventional configuration a) Wu et Al. split unit cell configuration (top: chip-scale microscopy image; bottom: SEM image of the cell sub-units). Reprinted with permission from [175]. b) Mixed-mode device optical and SEM images (reprinted with permission from [176]), which allows both electrical and optical read-write operativity....... 57

Figure 3.1. Optical bench setup schematics. The probe laser operates in CW mode, and is guided to the device via polarization controller and an optical circulator. The probe output is then collected and delivered to the detector, passing through a second optical circulator and a band-pass filter. The probe is here also used for real-time monitoring, to compensate for misalignment. The pump laser is guided towards a polarization controller first, then modulated via EOM, and amplified through EDFA. The pump is then filtered via band-pass filter, and through the optical circulators is sent towards and collected from the photonic chip. The output is then collected from the oscilloscope. The oscilloscope also drives the electrical pulse generator, which provides the necessary signal shape to the EOM. 62

Fig. 3.2. Gillespie Cellular Automata, simulation results of the annealing performed at 573 K (300 °C). (a) Crystal fraction χ versus time t (s) during the annealing process. Details of the progress are shown as insets, during the growth phase (top-left) exhibiting a smooth behaviour, and at completed recrystallization (bottom-right), showing the stochastic attachment and detachment of monomers at $\chi \cong 1$. (b-d) Images showing the GST crystallization evolution, starting with a pure amorphous material (grey), and ending with the full recrystallization (coloured crystallites). Figures were obtained at (b) $t = 5 \times 10 - 7$ s, $\sim 3.1 \times 103$ steps, $\chi = 0.05$; (c) $t = 2 \times 10 - 6$ s, $\sim 2.8 \times 104$ steps, $\chi = 0.5$; (d) $t = 10 - 4$ s, $\sim 6.1 \times 105$ steps, $\chi = 0.9997$. 80

Figure 4.1. Schematics of the original and proposed GCA neighbourhood indexing strategies. a) GCA term matrix and coordinate system of the Moore neighbourhood, colour-coded so to distinguish first neighbours (yellow) and second neighbours (blue) of the highlighted monomer (green). b) Corresponding representation via *unidimensional array*, conjugated with the correlated *Pointer matrix*. The proposed column-wise decomposition is inferable observing the monomer letter indices. The coordinate system here considers a single-valued notation, explicitly stored in the pointer matrix. The example shows the process for the same neighbouring monomer identification as in (a) (“right” monomer). Here, the identification retrieves the coordinate from the Pointer matrix at the position dictated by the (m) coordinate (green square). 87

Figure 4.2. Surficial monomers representation. a) Original implementation, negating the surfaces influences by expanding the matrix by one monomer unit on each direction, and restricting the calculation onto the inner area. b) Proposed solution, which assigns a predefined dummy neighbour where needed via Pointer matrix. The red features in a) and b) depict the out-of-boundary neighbour selection process. 87

Figure 4.3. Schematics of the CDF update method, and ITS activity $a0$ calculation. a) Previous implementation, which recalculated the CDF array after variation of any of its elements. b) Proposed method, which splits the CDF array and performs the cumulative sum operations only where any variation is observed (red lines), added to one further cumulative sum to yield the aimed activity value. 89

Figure 4.4. Performance comparison (log-log plot) of the inverse transform sampling methods. The blue line corresponds to the former single-shot implementation; the red line reports the data for the proposed split-array method. Insets reports the same datasets plotted on cartesian axis. It is clearly visible that, for larger sized arrays, the proposed methodology can reduce the calculation time by up to 3 orders of magnitude, whereas for smaller arrays the single-shot method is still the most suitable choice. 90

Figure 4.5. 3D GCA, simulation results of the 2D degenerate case (a-c), and comparison to the previous model results (d-f) (reprinted from [220]). The process here illustrated is the annealing of an amorphous region at a constant temperature of 404 K. Each panel represents the phase spatial distribution at different crystal fraction (a,d: 0.05, b,e: 0.2, c: 0.97 (0.999 for f)); crystal grains are visible as coloured areas over the uniform background. A good agreement of timings, grain size and grain distribution is reached. 91

Figure 4.6. 3D GCA simulation results of the 2D degenerate case annealing process at 680 K (a-c), and comparison to the previous model results (d-f) (reprinted from [220]). As in Figure 4.5, each panel row identifies a unique crystal fraction threshold (a,d: 0.05, b,e: 0.2, c: 0.97 (0.99 for f)). Once more, timings, grain size and grain distribution of the current model and Ashwin's model find a good correlation, except for the last panel (due to the different monitored crystal fraction). 92

Figure 4.7. Determination of the average (median) grain size from the simulation output at the annealing temperatures of 220 °C (a), 250 °C (b), and 300 °C (c). (d) Histogram reporting the distribution of the grain sizes from the simulated experiment, and comparison of the median data with the experimental one reported in Xu et al. [148]. 93

Figure 4.8. Comparison of the annealing simulation result at fixed temperature (500 K) with the 2D and 3D GCA models. The 2D simulation (a-c) is carried out on a 64 × 64 monomers surface, and the 3D simulation (d-f) emulates a 64 × 64 × 10 monomers volume. The analysis of the grain size distribution is reported in the histogram plots in (g-i), with the blue dataset reporting the 2D case analysis, and the red dataset for the 3D case. Data is reported for the crystal fraction values of $\chi = 0.2$ (a,d,g), $\chi = 0.5$ (b,e,h), and $\chi = 0.97$ (c,f,i). 94

Figure 4.9. Comparison of the simulation results for a 2D and 3D GCA models, applying a temperature gradient between the left (500 K) and right (700 K) boundaries. The experiment and figure follow otherwise the same steps described in Figure 4.8. 95

Figure 4.10. GCA calculation time, comparison between 2D and 3D cases as a function of temperature. All simulations account for the same number of monomers: 64 × 64 monomers for the 2D case; 8 × 8 × 64 for the 3D case. For the GCA model, datapoints above 800 K are not present; the depicted line trend is based on an estimated time of 10^3 seconds for 800 K. 96

Figure 4.11. Number of events (nucleation, growth and dissociation) per simulation time, during a simulated annealing process of an initially amorphous domain at different temperatures. Each of the plots represents a separate annealing process at constant temperature (temperature is reported in plot). The crystal fraction plot is also plotted (red, dashed) with its relative axis on the right of each plot. Legend reported in (a). 97

Figure 4.12. CNG model, events probability (per second – unit cell), as calculated with Eq. 4.1(a-c), Eq. 4.2(a-c) and the simulation parameters illustrated in Table 3.1 and Table 4.1 (the dissociation rate is here averaged over the possible neighbourhood configurations.). 103

Figure 4.13. CNG model, annealing result at a constant temperature of 407 K, for the 2D (a) and 3D (b) cases. c) Comparison of the grain size distributions, for the 2D (blue) and 3D (red) cases. Note the log scale on the vertical axis. d) Comparison of the crystal fraction variation as a function of time, same colour code as for (c). 104

Figure 4.14. CNG model, annealing result with a linear increasing temperature from the left boundary (500 K) to the right boundary (700 K), for the 2D (a) and 3D (b) cases. The temperature (linear) gradient is visually represented as a black arrow in (a). c) Comparison of the grain size distributions, for the 2D (blue) and

3D (red) cases. d) Comparison of the crystal fraction variation as a function of time (same colour code as in (c)). 104

Figure 4.15. GNG and GCA simulation comparison, on a (256 × 256 × 64 monomers) thin layer, with a linearly increasing temperature from the left xz boundary (500K) to the right xz boundary (700 K). Temperature is represented in (a) as a black arrow. a-b) CNG and GCA simulation outputs respectively, at $\chi = 0.97$. c) Grain size distribution, for (blue) CNG and (red) GCA outputs. d) Crystallization time as a function of the crystal fraction, for (blue) CNG and (red) GCA models. 105

Figure 4.16. GNG and GCA simulation comparison, on a highly asymmetric geometry, combined with a z-dependent linear temperature gradient (shown in (a) as a black arrow). a-b) CNG and GCA simulation outputs respectively, at $\chi = 0.97$. c) Grain size distribution, for (blue) CNG and (red) GCA outputs. d) Crystallization time as a function of the crystal fraction, for (blue) CNG and (red) GCA models. 106

Figure 4.17. GNG and GCA simulation comparison, on a thin layer with a de-centred 0D heat source. Temperature boundaries and gradients are represented in (a) as black arrows. a-b) CNG and GCA simulation outputs at $\chi = 0.3$. c) Grain size distribution, for (blue) CNG and (red) GCA outputs. d) Crystallization time as a function of the crystal fraction, for (blue) CNG and (red) GCA models. 107

Figure 4.18. Typical PPCM device COMSOL model. To provide a visual aid, the core of the waveguide is highlighted in blue. a) Geometry. b) Meshing results 111

Figure 4.19. Simulation framework schematics. a) Simplified logical flowchart. b) Data flow diagram (user data is not reported for sake of clarity). Legend on the bottom right. 115

Figure 4.20. Time-step generation, example and schematics. The red line corresponds to the ideal variable function, unknown to the algorithm but employed to generate the most suitable time-step. Each iteration, the algorithm extrapolates the variable value at a set of test locations, with the view to not allow the function to increase over a user-defined delta (height of the rectangles). Time-step increase with respect to the previous iteration is limited as $\tau i = 1 + h\tau i - 1$. Time-step values below or above user-defined boundaries are also adjusted accordingly. In case that the calculated value differs from the previous one (above the predefined delta), the time-step is further adjusted and the value is recalculated or interpolated. This particular example also shows the limitation of the algorithm, as the minimum time step is not well suitable to follow the particularly steep variable dynamics. 118

Figure 5.1. Schematics of the experimental device setup and method principles. (a) Gires-Tournois etalon schematics, and e.m. field reflection and refraction patterns. (b) Reflectance spectra for 4 sets of cavity material optical properties, which highlight the optical properties unique impact on the reflection modulation. 126

Figure 5.2. Schematics of the methodology principles, based on the simulated *vs* experimental squared error calculation for any possible cavity optical property value (inferable from the x and y axis). The colour-coded area depicts the aforementioned error, with the lower values found in red, and the highest ones plotted in blue. The fit procedure is restricted to the optical property values obtainable via weighted average of the literature data (here reported as black crosses). The search space area is here highlighted by its white boundaries. This frame reports the process for a single wavelength: the fit procedure is nevertheless carried out on each of the investigated wavelengths at the same time. 128

Figure 5.3. Reflectance fit – n/κ determination method, algorithmic flowchart 130

Figure 5.4. Amorphous GST, reflectance fit results and n/κ data extraction. In each panel, literature data (here from Refs. [66], [153], [155], [175], [244], [245]) is reported in un-saturated colours, grey to green as the fit method weight increases. a) Experimental (black dotted line) and calculated data. The comparison line (blue dashed line) is calculated based on the ellipsometrical estimation of the n/κ parameters. b-c) Refractive index and extinction coefficient respectively, with the ellipsometric data (blue dashed line) superimposed for confirmation of the fit quality. 133

Figure 5.5. Fit results for the crystalline GST sample, reported in analogy with the previous case. (a) Reflection spectra. (b) Refractive index. (c) Extinction coefficient. 134

Figure 5.6. Fit results for the amorphous GeTe cavity sample, reported in analogy with the previous cases. (a) Reflection spectra. (b) Refractive index. (c) Extinction coefficient. Literature data is taken from Refs. [155], [244], [246]–[248]. 135

Figure 5.7. Fit results for the crystalline GeTe cavity sample, reported in analogy with the previous cases. (a) Reflection spectra. (b) Refractive index. (c) Extinction coefficient. Literature data is taken from Refs. [155], [244], [246]–[248]......... 136

Figure 6.1. PPCM device schematics and functioning principles. A section of the waveguide is portrayed (indigo), comprising the GST unit cell on its top surface (emerald). A Write-Erase pulse sequence (orange) is delivered trough the waveguide, and a set of read pulses (green) highlights the consequent change of the device transmissivity, here enabling a binary memory functionality. 139

Figure 6.2. Simulation results for the Read operation, obtained upon variation of a single parameter amongst the ones initially assumed. a) Transmission upon variation of the waveguide-related parameters (thickness, width, etching step, and refractive index). b) Transmission calculated varying one of the geometrical parameters of the unit cell: thickness, width, length. The inset shows the extended range calculation for the unit cell thickness. c) Results upon variation of both real and imaginary components of the unit cell refractive index (visible on the x and y axis respectively). The experimental transmission value ($\mathrm{T} = 0.798$) is highlighted as a black line. d) Results upon variation of both real and imaginary components of the cap layer refractive index (visible on the x and y axis respectively). In (c,d), the initially adopted refractive indices are marked with a black cross............. 143

Figure 6.3. Transmission as a function of the unit cell refractive index, data from Figure 6.2(c) with additional markers showing the available literature data for the unit cell material (GST) at 1550 nm. Literature data are taken from Refs. [153], [155], [163], [175], [244], [245], [258]; markers labels report the manuscripts first authors. 145

Figure 6.4. Read simulation, electric field norm distribution across the two orthogonal cut planes oriented along the waveguide propagation axis (colour bar on the right). The propagation vector is oriented from left to right. (a,c) zx waveguide top-surface cut plane; (b,d) zy symmetry cut plane. (a,b) 1 μm unit cell, thickness of 6 nm, transmission $\mathrm{T} = 0.796$. (c,d) 5 μm unit cell, thickness of 8 nm, transmission $\mathrm{T} = 0.189$. 147

Figure 6.5. Experimental (normalized) transmission (left axis, continuous black line) and output pump power (right axis, dotted red line) measured during Write / Erase on a 2 μm unit cell. Image adapted from Ref. [71] 148

Figure 6.6. Simulation of the Read operation for a 2 μm unit cell. a) Parametric sweep simulated transmission as a function of the ratio between unit cell and cap layer thickness. A matching value to the experimental data is found with a unit cell / cap layer thickness ratio R = 0.79, corresponding to a unit cell thickness of 7.79 nm. b) e.m. field representation (on the waveguide top surface, zx orientation) corresponding to the unit cell thickness of 7.79 nm. 149

Figure 6.7. Experimentally determined (normalized) transmission data, Write operation transient and cool-down 1/e decay time estimation. (a) Data from Ref. [71]. (b) Data from Ref. [67].. 152

Figure 6.8. Simulation results for the high-energy pump stimulation (14.1 mW, 10 ns) on the 2 μm unit cell (legends on top panels, column-wise). Each column focuses on the variation of a single parameter, specified on top. a-c): Device transmission profiles. d-f): Peak unit cell temperature (thick line), average unit cell temperature (thin lines). In (a-c) and (d-f), the red dotted lines represent the same dataset, to provide a visual clue for sake of comparison. The grey arrows guide the eye from the results obtained with the lowest parameter value to the results obtained with the highest ones, also quantifying the impact of the variation of those. 153

Figure 6.9. Simulation results, assuming absence of thermal boundary resistance on the GST interfaces, using a 14.1 mW, 10 ns pump pulse on the 2 μm unit cell (legends on top panels, column-wise). Each column focuses on the variation of a single parameter, specified on top. A,b): Device transmission profiles. c,d): Peak unit cell temperature (thick line), average unit cell temperature (thin lines). In (a,b) and (c,d), the red dotted lines represent the same dataset, to provide a visual clue for sake of comparison.. 155

Figure 6.10. Comparison of the transmission (a) and temperature (b) transient after delivery of a 10 ns / 14.1 mW optical pulse to a crystalline, 2 μm – long unit cell.. 156

Figure 6.11. Simulation results for the Write / Erase process on a 2 μm unit cell, by use of a 14.1 mW / 25 ns Write pulse and a 14.1 mW / 25 ns, 5.64 / 100 ns Erase pulse, as reported in [71]. a) Transmission, simulated (red dotted line) and experimental (black line, reported from Figure 6.4). b) PPCM reflection, scattering and absorption coefficients. c) Temperature profiles calculated on the unit cell: peak temperature, average temperature, and peak / average temperature of the non-crystal phase (amorphous or molten GST). d) Crystal fraction. Insets report the phase distribution (orange, crystalline; white, amorphous) as viewed from the unit cell top surface. 158

Figure 6.12. Tunable volatility experiment: determination of the unit cell amorphous mark. a) Simulated temperature distribution after delivery of the combined pump and probe signals on the 1 μm unit cell. The melting temperature isothermal line is highlighted in red. b) Definition of the model domains, based on the determined isothermal line which yields the pre-determined transmission $level\ 1$.. 161

Figure 6.13. Experimental and simulated normalized optical contrast as a function of time, for a 1 μm unit cell under stimulation of a continuous high-energy probe. 162

Figure 6.14. Thermal behaviour of the PPCM during high-power probe induced recrystallization. a) Simulated unit cell simulated peak temperature as a function of time, for the three adopted probe powers. b) Analysis of the dependency of the crystallization time $t\chi$ as a function of the inverse of the temperature (bottom axis, coloured markers) and probe power (top axis, black markers). The linear fit is calculated from the inverse-temperature plot. 163

Figure 6.15. Comparison of the normalized contrast data (from Refs. [68], [71]) during Write and Erase sequences, both delivering an initial 14.1 mW pulse. The transient indicates very similar dynamics, regardless of the different initial conditions (c.f. Figure 6.5). 166

Figure 7.1. Numerical analysis of the single bar nanoantenna structure (width of 40 nm, thickness of 30 nm, variable length). a) device optical behaviour, as a function of the bar length: transmission T (left axis, black continuous line); reflection R, scattering S, Ag absorption A (right axis). Inset shows a schematic of the analysed configuration. b,c) electric field on the orthogonal cut plane passing through the bars mid-point (the width of the plot corresponds to the waveguide width). The z component of the field is colour-mapped (blue representing negative values; red for positive values); arrows represent the x and y components (log scale).. 173

Figure 7.2. Numerical analysis of the dimer bar nanoantenna structure (width of 40 nm, thickness of 30 nm, gap distance 10 nm, variable length). a) device optical behaviour, as a function of the bar length: transmission T (left axis, black continuous line), compared to the previous result (grey line); reflection R, scattering S, Ag absorption A (right axis). Inset illustrates the device schematics. b,c) electric field on the orthogonal cut plane passing through the bars mid-point (the width of the plot corresponds to the waveguide width). The z component of the field is colour-mapped (blue representing negative values; red for positive values); arrows represent the x and y components (log scale). d,e) norm of the electric field (blue to red, normalized colour scale) corresponding to the solutions plotted in (b) and (c) respectively. 175

Figure 7.3. Effect of the gap distance variation upon device transmission (a) and enhancement factor (b). Gap distance g values are indicated in (b). 176

Figure 7.4. Effect of the bars thickness variation upon device transmission (a) and enhancement factor (b). Bars thickness t values are indicated in (b)...... 177

Figure 7.5. Effect of the bars width variation upon device transmission (a) and enhancement factor (b). Bars width w values are indicated in (b). 177

Figure 7.6. Summary of the variation of minimum transmission T (a) and (inverse of) maximum EF (b), calculated at dipolar resonant configuration, as a function of the investigated parameter (normalized scale): gap distance g, bars thickness t, and bars width w. 178

Figure 7.7. Schematics of the GST-implemented dimer bar nanoantenna..... 180

Figure 7.8. Simulation results for the amorphous and crystalline GST-implemented dimer bar nanoantenna, as a function of the length of the bars. (a) Transmission T. (b) Enhancement factor EF. (c) Transmission difference $\Delta\mathrm{T}$ calculated from the data reported in (a). (d) Device Minimum insertion loss IL; the two marks highlight the values calculated at the two peaks locations found in (c). 180

Figure 7.9. Simulated optical properties of a dimer-bars nanoantenna, of width 30 nm, and thickness 30 nm, as a function of the length of the bar. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $\Delta\mathrm{T} = \mathrm{T}cry - \mathrm{T}am$. (d) Theoretical minimum insertion loss, $\mathrm{IL} = 10 \cdot \mathrm{log10maxT}GST$. The markers underline the values calculated at the $\Delta\mathrm{T}$ maxima found in (c)... 182

Figure 7.10. Schematics of the bow-tie nanoantenna device, fabricated on the waveguide top surface. 183

Figure 7.11. Simulated optical properties of a bow-tie nanoantenna, of 30 nm, as a function of the prism base (equilateral triangle) height. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $\Delta\mathrm{T} = \mathrm{T}cry - \mathrm{T}am$. (d) Theoretical minimum insertion loss, $\mathrm{IL} = 10 \cdot \mathrm{log10maxT}GST$. The markers underline the values calculated at the $\Delta\mathrm{T}$ maxima found in (c)... 183

Figure 7.12. Dimer-discs nanoantenna schematics. 184

Figure 7.13. Simulated optical properties of a dimer-discs nanoantenna, of 30 nm, as a function of the diameter of the discs. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $\Delta\mathrm{T}$. (d) Minimum insertion loss IL. The markers underline the values calculated at the $\Delta\mathrm{T}$ maxima found in (c). 185

Figure 7.14. Schematics of the encapsulated dimer-bar unit cell. 186

Figure 7.15. Optical characteristics of the encapsulated plasmonically-enhanced unit cell (dimer-bars configuration), as a function of the bars' length. The color legend is reported in (a) only. (a) Transmission for crystalline (continuous line) and amorphous (dashed line) GST inclusion. (b) EF values, calculated at the GST inclusion volume. (c) Transmission difference $\Delta\mathrm{T}$, with the two peaks of each individual line highlighted by a cross and a circle. (d) Minimum IL, with the IL values calculated at the $\Delta\mathrm{T}$ peaks highlighted as in (c). 186

Figure 7.16. Schematics of the embedded dimer-bars device design. 188

Figure 7.17. Simulated optical properties of a dimer-bars nanoantenna, of width 30 nm, and thickness 30 nm, as a function of the length of the bars. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $\Delta\mathrm{T}$; a comparison with the on-top configuration is provided (d) Theoretical minimum insertion loss, $\mathrm{IL} = 10 \cdot \mathrm{log10maxT}GST$. The markers underline the values calculated at the $\Delta\mathrm{T}$ maxima found in (c). Data of the on-top configuration is also provided in (c) and (d), for sake of comparison. 188

Figure 7.18. Optical behaviour of the dimer-bars embedded configuration, as a function of bars length and vertical displacement. The vertical axis is centred at the etching step location (waveguide mid-height); the nanoantenna develops vertically starting from the position indicated on the vertical axis. (a) $\Delta\mathrm{T}$, calculated between crystalline and amorphous configurations. (b) Insertion loss [dB]. 189

Figure 7.19. Embedded dimer-bars device optical behaviour as a function of the nanoantenna material. Two markers (vertical and diagonal crosses) are used, reporting the two values corresponding to the first and second $\Delta\mathrm{T}$ peaks respectively. (a) Peak $\Delta\mathrm{T}$ values. (b) IL calculated at the peak $\Delta\mathrm{T}$ configuration. 191

Figure 7.20. Schematics of the nanoantenna design (a-c), observed from the three orthogonal cross sections cutting through the nanogap centre, and

corresponding simulated norm of the electric field (d-f) for the crystalline GST and amorphous GST inclusion (colour bar on the right, $\log_{10}$ values in V/m units).192
Figure 7.21. Spectral response of the proposed device configuration, in the 1350 nm – 1700 nm range. a) Device transmission for amorphous (red, dashed line) and crystalline (blue, continuous line) GST phases. b) derived $\Delta\mathrm{T}$ (green, left axis) and IL (orange, right axis). c) EF values, same colour scheme as in (a). 194
Figure 7.22. Switching dynamics during Write operation for the optimized plasmonically-enhanced unit cell described in Section 7.4.1. The Write pump is a rectangular pulse of 1 mW / 2 ns, reported in both plots as a grayed area to provide a visual guide (vertical scale in (a), right axis). (a) Temperature on the unit cell and Ag resonators. (b) Left: crystal fraction χ. Right: device transmission variation $\Delta\mathrm{T}$. 195
Figure 7.23. Write operation loss (a) and temperature (b) distribution plots; data taken at the end of the Write pulse (2 ns). The optical loss density 2D plot (a) is calculated on the cut plane passing through the unit cell central point. 196
Figure 7.24. Spatial GST phase distribution across the unit cell (the plots report the crystalline phase only). 196
Figure 7.25. Switching dynamics during the Erase operation, starting from the crystal fraction value of $\chi = 0.18$ obtained after the Write operation. The Erase pulse profile is reported in both plots as a greyed area to provide a visual guide (vertical scale in (a), right axis). (a) Peak, average and minimum temperature on the unit cell and resonators. (b) Left: crystal fraction χ. Right: device transmission variation $\Delta\mathrm{T}$. 199
Figure 7.26. Crystalline GST spatial phase distribution throughout the unit cell, for the proposed 4-levels memory storage, as calculated from the Programming pulse sequences illustrated in Figure 7.27. 201
Figure 7.27. Multi-level storage demonstration, crystal fraction data as a function of time. The pulse delivered power as a function of time is also reported as a greyed area, to guide the eye along the operations process. a) Reset operation, shifting any arbitrary initial state to the value of $\chi = 0.18$. b) Programming operation, returning the previously determined state to any of the four predetermined crystal fraction values, within a tolerance of ± 0.02. Inset demonstrates the quasi-linear correlation between crystal fraction and device $\Delta\mathrm{T}$. 202

Figure 8.1. Electrically connected, embedded nanoantenna device concept and results. a) Schematic representation. b) Optical results of the un-connected device (phase-dependent transmission data, left axis; optical contrast $\Delta\mathrm{T}$, right axis), as a function of the bars length; the waveguide wall boundary is reported to guide the eye. c) Optical results of the connected device, for a bar length of 650 nm (the axis choice is the same as in (b)). The data is reported as a function of the ITO-bar overlap size. 212

Figure A.1. Representation via unidimensional array and Pointer matrix of a 3D volume. The abbreviated notations L (left), R (right), U (up), D (down), T (top), B (bottom) are here adopted for sake of conciseness. The red features, illustrating the selection of the *bottom-up-left* neighbour of the highlighted green monomer (at the centre of the volume), demonstrates how the procedure is carried out exactly as in the 2D cases illustrated in Figure 4.1(b) and Figure 4.2(b). 219

Figure A.2. Artificial reference calculation process, for the crystalline GeTe sample case. (a) Colour-coded data representation on the $\varepsilon1/\varepsilon2$ plane ($\varepsilon1$ ranging from 10 to 50; $\varepsilon2$ ranging from 0 to 40). The first column reports the data calculated at the resonant wavelength; columns 2 to 6 show additional frames as a function of the wavelength (values reported above). Top row: |ΔR| (blue-to-red pointing to higher-to-lower |ΔR| values), with crosses representing the selected Refs. [155], [244], [246]–[248]. Middle row: guess Gaussian distribution G (normalized probability data, blue-to-red pointing to lower-to-higher probability), with the circle representing the Refs. Average and the arrow corresponding to the shift vector v. Bottom row: combined probability C = ΔR×G (blue-to-red pointing to lower-to-higher probability), with the cross representing the C weighted average position. (b) Combined space n and κ values (crosses) calculated from the coordinates of the C weighted average values, as shown in (a)-bottom row. Continuous green lines are the 3rd order polynomial fits of the calculated data, and correspond to the resulting artificial reference. 222

Figure A.3. Algorithmic flowchart for the artificial reference determination. 223

Figure A.4. Alternative fit of the crystalline GeTe data, after construction of the artificial reference based on the literature data reported in Figure 5.7. The black dotted line reports the measured reflectance spectrum. The blue dashed line reports the data related to the ellipsometric estimation of n/κ. The red line reports the fit results for each panel. The orange line highlights the artificial reference amongst the reference pool datasets. a) Reflectance. b) Refractive index. c) Extinction coefficient. 224

Figure A.5. Alternative fit of the crystalline GST data, after construction of the artificial reference based on the literature data reported in Figure 5.5, and reported with the format adopted in the previous (crystal GeTe) case. a) Reflectance. b) Refractive index. c) Extinction coefficient. 224

Figure A.6. Simulation results for two plasmonically-enhanced (dimer-bars) PPCM devices, fabricated on the waveguide top surface, as a function of the center-to-center distance (see inset in (b) for a graphical representation). The degenerate case data (corresponding to a single unit cell) is reported via markers at the $x = 0$ coordinate. a) Transmission data, for the amorphous (red, dashed) and crystalline (blue, continuous) GST. b) Left (green): theoretical $\Delta\mathrm{T}$. Right (orange): theoretical minimum IL. c) EF on the first and second sub-unit cells for the amorphous (red) and crystalline (blue) GST cases. d) GST absorption on the two sub-unit cells, for the amorphous (red) and crystalline (blue) cases. (e,f) Electric field distribution plot (on the sub-unit cells cut plane, top view), calculated at the distance of 280 nm. 226

Figure A.7. Optical behaviour of interacting plasmonically-enhanced PPCM devices, with a centre-to-centre distance of 280 nm, as a function of the number of sub-units. In (a) and (b), the extrapolated behaviour of non-interacting sub-units is also reported (thin lines), as a source of comparison. a) Transmission data, for crystal (blue) and amorphous (red) GST. b) Left (green): $\Delta\mathrm{T}$. Right (orange): IL. c) Average EF for amorphous (red) and crystalline (blue) GST sub-unit cells. d) Average GST absorption for crystalline (blue, left axis) and amorphous (red, right axis) phase. 227

Figure A.8. Simulated optical properties of an embedded bow-tie nanoantenna, of thickness 30 nm, as a function of the prism base height. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $\Delta\mathrm{T}$ (d) Minimum insertion loss. The markers underline the values calculated at the $\Delta\mathrm{T}$ maxima found in (c). Data of the on-top configuration is also provided in (c) and (d). 231

Figure A.9. Simulated optical properties of an embedded dimer-discs nanoantenna, of thickness 30 nm, as a function of the discs' diameter. (a) Transmission data with and without GS inclusion. (b) EF, with and without GS inclusion. (c) Theoretical maximum $\Delta\mathrm{T}$ (d) Minimum insertion loss. The markers underline the values calculated at the $\Delta\mathrm{T}$ maxima found in (c). Data of the on-top configuration (reported from Figure 7.13) is also provided in (c) and (d). 232

Figure A.10. Fabrication tolerance analysis of the proposed dimer-disc nanoantenna, with respect to displacement from the waveguide centre, and from tilting angle θ from the intended orthogonal orientation. a) Simulated $\mathrm{n}\Delta\mathrm{T}$, as a function of displacement and tilt angle. b) Fit of the data as a function of the tilt angle, calculated at displacement $s = 0$. c) Fit of the data as a function of the displacement, calculated at $\theta = 0$. 233

List of Tables

Table 1.1. Shortlist of some promising emerging memory technologies exhibiting non-volatile data retention (Data retrieved from Refs. [23]–[28]) 8

Table 3.1. Parameters adopted for the GCA phase transition simulation demonstrated in Fig. 3.2. 81

Table 4.1. Parameters adopted for the CNG phase transition simulation. 103

Table 6.1. Simulated Read operations of the devices reported in Ríos et al. [65]. Matching unit cell (effective) thickness values are highlighted in red. 148

Table 7.1. Optical properties of the nanoantenna investigated materials. 190

Table 7.2 Summary of plasmonic phase-change optical cell optical behaviour at 1550 nm 193

Table 7.3. Comparison of proposed and experimental PPCM and GST-enabled switch devices performance. The worst and best performance is highlighted in green and red respectively, for each column. *W* refers to the Write process; *E* refers to the Erase process. 203

Table A.1. Summary of the plasmonically-enhanced PPCM behaviour, after implementation of a 3 nm Si anti-diffusion barrier 235

Chapter 1. Introduction and Motivation

1.1. Data storage and processing challenges

Information storage and computational power demands have seen a drastic increase in recent decades. The penetration of the mobile market, the advantages brought by IoT[1] technology [1], the spread of the Cloud Computing paradigms [2] and the success of user-generated content platforms [3], internet based-entertainment providers [4] and social media [5], are the main factors that create the current massive volume of data distributed through the internet, which in the current year is estimated to sum to 40 ZB [6].

Such high demand for data streaming and data processing translates into various technological challenges. One major limitation is found in the difficulty of the current datacentres to accommodate and distribute such a large amount of data. New datacentres need to be continuously created or innovated, in order to upkeep the continuously increasing demands. But the expansion of current facilities is not free of charge: tightly bound to storage and processing power, also energy consumption is an intrinsic problem which seeks solutions, as it is directly linked with the environmental impact (i.e. CO_2 emission). For instance, it is estimated that, as of 2020, 12% of the global emission is due to datacentres only [7], and at the current rate of growth of data demand, a figure of 33% of the world energy use for network access is predicted to be reached in a decade [8].

Turning the attention from a global perspective to the single device's capabilities, it is also possible to outline similar technological challenges that require further research to meet today's demand in storage, computing power, and energy efficiency. The majority of the computing devices, from the simplest calculator to large-scale facilities, are characterized by their adherence to the *von Neuman architecture* (vN), named by its author, John von Neumann, which described it first in 1945 [9]. The basis of the vN architecture lays on the separation of the computing framework into a few main components: a *processing unit,* which is, in turn, comprising the *arithmetic* (or *logic*) *unit*, and the *control unit*; a *memory unit*, onto which the processing unit fetches data and instructions through the

---

[1] Internet of Things.

memory channel; a *mass storage unit*, reserved to hold the long-term data; and the *input* and *output* devices, being the user interfaces.

Whilst this revealed to be a very successful architecture, it carries along its intrinsic drawback, evidenced by John Backus in 1977 [10], the so-called *von Neumann bottleneck*. Backus highlighted how the logical unit idles whilst waiting for the overhead time imposed by the data shuttling from and to the processing unit through the memory channel. Many mitigation approaches have been implemented (e.g. implementation of cache-level memory, or pre-fetching algorithms), but regardless of the efforts, the limited bandwidth of the metallic interconnects natively persists.

An option to side-step the von Neumann bottleneck lays onto the use of optical signalling. Optical fibres set in motion the development of new data sharing paradigms, a merit of the signal encoding onto different wavelengths travelling down a single optical channel (a process referred to as *wave division multiplexing/ demultiplexing* (WDM/WDDM)). The direct consequence of this approach is the bandwidth increase, but additional and significant benefits can be found in the non-thermal loss and the increased signal speed. All these are also suitable characteristics for *integrated circuits* (IC) devices (which, integrating optical components, would be referred to as *photonic integrated circuits* (PIC) devices), as they overcome the inherent limitations of the electrical counterparts (e.g. heat generation, limited bandwidth, fabrication complexity, and costs).

Even though the field of photonics finds a fundamental limitation of the size of the light-manipulating device, bound to the physical law of diffraction limit, the advantages brought by the fast and potentially lossless signal carrier (instead of electrons onto a conventional conductor or semiconductor) make the photonic approach a sought after solution in several applications, where scalability can be treated as a secondary element. A significant example is found on the first experimental validation of on-chip photonic data transfer, illustrated in Sun et al. [11], a fundamental milestone on the path leading to on-chip optical processing. In this work, the authors described the implementation of an optical channel for the data communication between processing and memory units, alleviating the electrical connection bandwidth limitation.

A second viable option to overtake the issue bound to the von Neumann bottleneck can be found after detangling from such architecture (i.e. not physically separate memory and computing units), a feature that von Neumann himself deemed unfeasible [9] due to the technological and theoretical challenges. The non-von Neumann architecture (n-vN) concept has recently found renovated interest, particularly in the hardware implementation of memory devices capable to perform computing primitives, such as *machine learning* paradigms, or, more in general, *in-memory* computing.

In-memory computing collects all techniques which allow one to perform computational tasks directly onto the memory device. The key idea is to exploit the physical attributes of the memory devices, take advantage of the high degree of parallelism allowed by the memory circuitry and array-organization, and perform certain operations without the need to continuously transfer data to the processing unit. The applicability of such techniques is not universal though, as blurring the boundary between processor and memory carries advantages where a high degree of parallelism is required as with very large datasets, or other memory-centric tasks; but it brings also challenges when a more traditional processing approach is required.

In-memory computing paradigms contain also the branch which aims to analyse, replicate, and exploit the processing characteristics found in biological units: the multi-disciplinary *neuromorphic* computing. Neuromorphic computing is a term coined in the late 80s and popularized by Carver Mead [12], referring to analog or digital integrated systems that embody features found in biological network processing. The hardware implementation of a neuromorphic device replicates one of the many biologically inspired computing architectures, such as the *Deep Neural Network* (DNN) model. Any DNN consists of a collection of connected units, called *artificial neurons*, each one characterized by its set of *inputs*, *outputs*, *weights,* and *threshold functions*. Each neuron delivers a signal when a certain condition is met (based on the input weights and the neuron threshold function), and the cascading processing of the provided inputs from the neuron collection results in an emerging behaviour which can perform a wide range of processes, usually devised to extract hard-to-find higher-level features from raw data. DNNs approaches natively offer a high degree of parallelism, asynchronous processing, and more importantly, a potentially much-improved energy efficiency, albeit only after a (usually) energy-hungry training process.

Whilst several software implementations of DNNs are currently employed in a wide range of human and natural science fields, industrial and commercial applications [13], the hardware implementation is particularly sought after, as it has the potential to become a fast and energy effective method for the manipulation of large datasets such as for sensing, data classification, pattern recognition, signal processing, time series prediction, forecast, or parallel arithmetic operations [14]–[16].

Regardless of the vN or n-vN implementation, or electronics vs photonics approach, a boost of the memory element (i.e. improving its performance metrics) is also key to the next leap towards improved computing paradigms, due to the growing divergence between processing unit and memory unit speed, referred to as *memory wall* [17]. Transistor technology is nowadays the basis for most of the technological standards for processing units and memory units' components. Transistors density followed the famous Gordon Moore's prediction [18] for roughly 40 years, and so did the performance of both memory and processing units. But today, the physical limit towards transistor's gate size reduction, as well as the fabrication profitability, are approaching a hard ceiling and such prediction fails to be satisfied [19], [20]. A different perspective on the transistor scaling focuses on energy consumption, the Dennard scaling law [21], which correlates the increase of the power density with the transistor downsizing. Due to charge leakage increasing with the approach to nano-scale sized transistors, scaling in conventional devices also sees the breakdown of such prediction, leading to increased energy consumption and heat dissipation.

All of the above-discussed issues lead not only to a requirement for advancement towards faster, greener, and overall better processing units but also for considerable memory functionality improvement - due to both architectural and technological reasons.

1.2. Information storage technology

Since the first draft depicting the vN architecture, a rigid classification of the memory concept has persisted. More precisely, each of the vN units requires a certain amount of memory, specifically optimized for a characteristic scope. The output unit requires to hold the information for a long time duration, for the user to be capable to "asynchronously" retrieve it. The logic and control units also need a certain memory capability, to perform the needed operations, and the external memory unit is devised to store the data portion containing the information to be processed and instructions on how to perform such process. The ideal memory device, capable to supply the needs of each of those units, would satisfy the following demands:

- Low energy and high speed write operations
- Low energy and high speed read operations
- Retention of state without bias power / non-destructive readout operation
- Holding of state for long time durations
- High cycling endurance
- Low footprint/high memory density (enhanced scalability)
- Ease of fabrication
- And, importantly, low cost of fabrication.

A single memory device architecture cannot supply the best performance for each scope, due to the technology-bound compromise between data retention, speed, and energy efficiency. A significant example of this can be found on the schematics of a Flash memory device, whose base memory unit comprises a *floating-gate transistor* (FGMOS). The FGMOS is a modified version of the standard MOSFET device, where a metallic layer is inserted (i.e. floating gate) in the middle of the insulator layer between transistor's body and control gate. Memory effect is achieved by trapping electrons in the floating gate, by application of a sufficiently high external voltage to increase tunnelling probability across the insulator potential barrier [22]. Even standing by this simplified description, and reminding how the tunnelling probability is directly proportional to the applied electric field but also an inverse function of the potential barrier width [22], the difficulty to achieve both low write energy and long data retention is evident. These two performances must meet at a compromise point, with the insulator layer being sufficiently thin to allow reasonable-voltage writing operations, and yet being thick enough to retain a percentage of the trapped carriers for the engineered duration. The time necessary for the write operation would eventually be optimized to achieve a predetermined minimum charge value, setting the balance between memory speed, energy requirement and retention time.

Classification of the available memory technology is, therefore, a natural consequence of performance compromises, and various parameters may be adopted to build a memory taxonomy (e.g., data retention time, access time, technology, active material, etc.). Amongst all, the information rate is one of the issues previously cited as having the largest impact on future technology development; for such, a useful parameter to adopt for memory classification purposes is the latency time (i.e. time lag between input request and input reception). In fact, the most common memory hierarchy is based upon the lag time. A panoramic of such classification is reported in Figure 1.1.

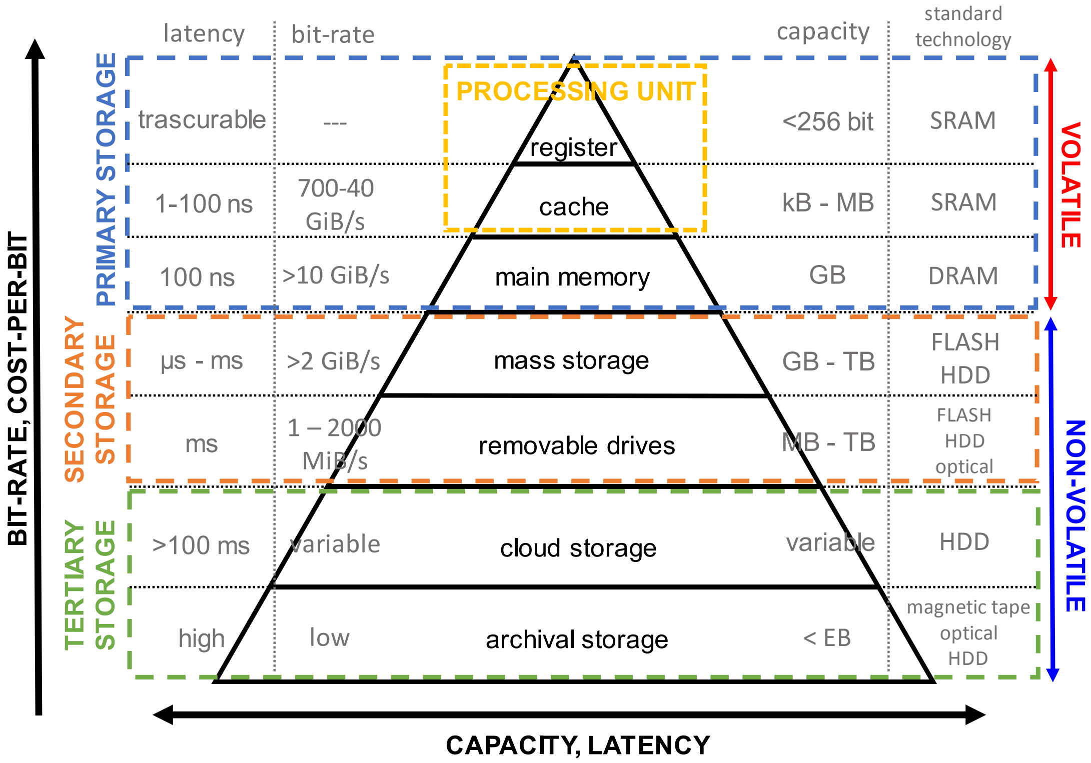

latency
bit-rate
capacity
standard technology
PRIMARY STORAGE
SECONDARY STORAGE
TERTIARY STORAGE
PROCESSING UNIT
trascurable
---
register
<256 bit
SRAM
1-100 ns
700-40 GiB/s
cache
kB - MB
SRAM
100 ns
>10 GiB/s
main memory
GB
DRAM
µs - ms
>2 GiB/s
mass storage
GB - TB
FLASH HDD
ms
1 – 2000 MiB/s
removable drives
MB - TB
FLASH HDD optical
>100 ms
variable
cloud storage
variable
HDD
high
low
archival storage
< EB
magnetic tape optical HDD
VOLATILE
NON-VOLATILE
BIT-RATE, COST-PER-BIT
CAPACITY, LATENCY

Figure 1.1. Memory hierarchy, based on latency time, showing the separation between typology of storage. Starting from the top (register class), each class offers a generally higher capacity – lower cost-per-bit, at the expense of higher latency and lower bit-rate. Features such as volatility, capacity, bit-rate and commonly found technology are also reported. Numerical data retrieved from Ref. [23].

Here it is possible to see how the lowest latency is found in memory devices within the processing unit, the so-called registers. The *registry* class memory sees the maximum optimization towards low latency (ns scale), and thus, the performance over energy consumption, ease and cost of fabrication, and footprint are poor.

Cache class memory comes below the registry class and has the scope to build a reservoir of information between the memory unit and processing unit. It is also very expensive to fabricate but allows higher capacity than the registers. In today's processing units, it is usually possible to find a sub-division of the cache memory into further levels, from 0 to *n*, with level 0 cache being the closest to the processor (and with *n* lower than 4 in most commercialized devices).

As aforementioned, the vN *main memory* contains data and instructions which need to be accessed by the processing unit. All relevant data can not be stored in the registers, nor in the later introduced cache-level memory. Therefore, such data needs to be stored in a device which balances affordability, sufficiently low latency values, sufficiently high bandwidth and capacity.

Above the main memory class (with main memory included), devices are optimized for low latency, balancing memory density, fabrication cost, and energy consumption, and can be directly accessed by the processing unit (hence the optimization for low lag and high bandwidth). The currently available technologies allow the independent storage/retrieval on each unit state (i.e. *random access memory*, or RAM) with fast access/write time, even though the capabilities are counterbalanced by the energy-requirement drawback. In particular, traditional RAM does not allow the retention of each memory unit state, unless energy is constantly provided (volatile memory). For instance, SRAM memory (usually implemented for cache-level devices) maintains its state in the bistable configuration of a transistor-based flip-flop circuit; and thus, after bias removal, such state is quickly lost. DRAM memory (common technology for RAM main storage) instead retains the information within transistor-controlled capacitors but needs constant refreshing (with a kHz – scale rate) to withhold such charge.

Below the main memory class lays the secondary storage superclass. This contains devices whose technology is balanced towards much higher capacity, low cost-per-bit, and most of all, longer data retention. Indeed, the large majority of mass-storage devices are capable to retain the stored information without bias current (non-volatile memory technology). Usually, the lowest latency (of this class) is found in the *mass-storage* unit, which resides within the computational device. *Removable devices and on-line storage* (cloud storage) have large latency but allow to access to an even larger pool of data.

Examples of commonly found technologies for the secondary storage class are the magnetic storage, where information is encoded onto the local magnetic orientation of a polycrystal alloy disc (magnetic discs) or magnetic material dispersion on a polymeric tape (magnetic tapes); optical storage, with the binary information written as non-reflective marks on an otherwise reflective alloy substrate; and the aforementioned Flash (NAND) storage.

The tertiary storage class is usually adopted for archival purposes and usually involves a robotic arm taking the memory devices to a reading unit. Examples of these are tape or optical discs libraries, whose collective size exceeds the Tera- or Peta-byte scale.

Whilst SRAM, DRAM, and Flash memory/Magnetic disc are the most widespread technologies, improvement of any of the qualities characterizing the memory is always beneficial. New technologies, usually referred to as emerging technologies, are being investigated, with the view of delivering low lag, write voltage and energy usage; high memory density, and possibly the long-sought-after non-volatility for main unit applications. Table 1.1 shows some of the most promising emerging technologies to replace the current standards.

Table 1.1. Shortlist of some promising emerging memory technologies exhibiting non-volatile data retention (Data retrieved from Refs. [24]–[29])

	FeRAM	STT-MRAM	ReRAM
Materials adopted	Ferroelectric materials, such as SrBiTiO or PbZrTaO	Ferromagnetic materials such as the CoFeB alloy	Metal oxides, PC materials
Storage mechanism	Inversion of the polarization of a ferroelectric material, exhibits two resistive states	Tunnelling magnetoresistance (read) and Spin-transfer torque effect (write)	Resistivity levels switching by formation of conductive filaments or phase switching
Unit cell area [F^2]	6-40	6-50	4-30
Read latency [ns]	20 - 80	3 - 50	10 - 60
Write latency [ns]	50 - 100 ns	<10 ns	10 - 200 ns
Switching current density [A/cm^2]	---	10^6	10^6-10^7
Switching energy [pJ]	1000	10-100	10-1000
Number of cycles before failure	10^{14} - 10^{15}	10^{12} - 10^{15}	10^8 - 10^{13}
Data retention	> 10 years	10 years	10 – 100 years

1.3. Phase-change materials and phase-change memory devices

Amongst the emerging technologies, a sub-set of the resistive RAM memory class, the *phase-change memory* (PCM), has recently found commercial success in niche applications such as high-performance mass-storage [30] but is currently being investigated to further spread its range of applicability. The mechanism behind the functioning of PCM devices is the storage of the binary information within the resistivity of the unit cell, allowed by the peculiar material properties of the *phase-change* (PC) materials. In general terms, any material included in the PC materials category satisfies the following conditions [31]:

- Possesses at least 2 distinct phases
- Each phase is characterized by a unique set of optical and electrical properties
- Phase transition is reversible
- Phase transition is allowed on a short timescale (with short being purposely left unspecific, as it might depend on the application).

Many PC materials possess 2 stable or metastable phases and thus allow the phase retention in a non-volatile manner. An example of these is the ternary $Ge_xSb_yTe_z$ [31] chalcogenide alloy, which can be usually found in amorphous or crystal phases, depending on the thermal history. Other materials, such as V_xO_y alloy [32] (belonging to the *transition metal oxides* (TMOs) class) are instead characterized by the volatile nature of the phase transition (i.e. removal of the driving force is quickly followed by a return to the stable phase).

PC materials find a rather wide range of applications, in the electronics, optical, and photonics fields. The reason lays mostly on the tunability (either volatile or non-volatile) of the material properties, which allows one to actively control the response of the device.

Since the first pioneering works from Ovshinsky and collaborators [33]–[36], the main behavioural characteristics of the chalcogenide PC materials have been described: i) reversible switching with electrical or optical pulses, ii) two distinctly different sets of material parameters, iii) threshold conduction mechanism for the amorphous phase, and iii) the “memory effect“ (i.e. non-volatile switching). Those characteristics were exploited to devise at least two commercially successful types of memory devices.

The first example is found in the rewritable optical storage technology, patented by Matsushita/Panasonic at the end of the 1980 decade. Non-rewritable optical storage, as mentioned above, relies on the reflection of an aluminium layer encapsulated between (transparent) polycarbonate substrates. Whilst the unmarked layer is highly reflective, the fabrication of μm-scale pits on such layer allows a reading laser head to detect an appreciable difference in reflectivity between marked and unmarked zones. Thus, the binary information would then be encoded in the presence or absence of such pits, which would then be revealed by use of a laser. First devised at the end of the 1950s, this architecture was later developed in a variety of formats [37], amongst which one may find successful examples such as *Compact Discs* (CD), *Digital Versatile Discs* (DVD), and Blu-ray discs. But, regardless of the success, such technology did not allow the end-user to write information (until the introduction of the CD-R format, using a chemical dye to write *once* [38]): it was thus limited to the *read-only memory* (ROM) class.

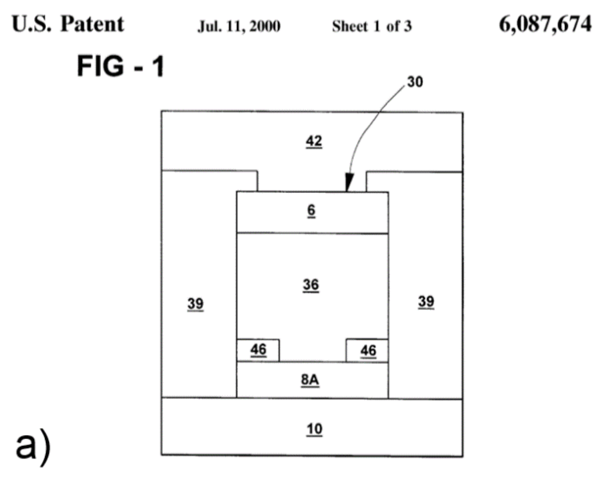
Yamada and collaborators [39] introduced the adoption of PC materials (the Ag/In doped Sb_2Te alloy[2]) in place of the aluminium reflective layer, with the view to leverage on its phase bi-stability (exhibiting different optical properties), and optically induced reversible phase switching. This allows one to write information as amorphous marks on a crystal layer (characterized by distinct reflection coefficients). As a result of the phase transition capability of PC alloys, both write and erase processes could be carried out with a laser, allowing the end-user to perform multiple write-erase cycles. As a consequence, the rewritable optical storage technology rapidly diffused through the market.

As for integrated solutions, the idea of relating the PC material's resistance to a memory state goes back to older works of Ovshinsky and collaborators [34] but saw a first commercial application much later. In the year 2000, Ovshinsky patented a device called Ovonic EEPROM [40] (*electrically erasable programmable read-only memory*), where the information is encoded in the crystal fraction of a PC material inclusion between two electrical contacts (see Figure 1.2).

---

[2] In contrast to other alloys, AgInSbTe exhibits a strongly growth-dominated crystallization mechanics [99], suitable for faster erasure kinetics of small marks, thence its choice by Yamada and collaborators.

This device exhibits 2 or more distinct resistivity states, depending on the crystal fraction value. The crystal fraction would be changed by delivery of tailored high-energy electrical pulse sequences, after which the device would then achieve the desired resistance level. Indeed, this is a noteworthy example of integrated electrically rewritable memory device, whose concept would then be the foundation of other successful implementations (e.g. the Samsung PRAM [41] and Intel-Micron 3D XPoint [30]).



30 – memory element
42 – electrode grid structure
6 – first contact
36 – memory material
46 – insulation layer
8A – second contact
39 – insulating material
10 – silicon substrate

b)

Figure 1.2. Ovonic EEPROM device, reprinted from [40]. a) Schematics of the memory device as reported on the patent application. The original caption reads: ”Cross-sectional view of a memory element of the present invention having a first and second contact each adjoining the volume of memory material”. b) Provided captions for the indices in (a).

The PC EEPROM concept offers many advantages with respect to NAND Flash memory technology, such as i) a write typical time (hundreds of ns) about 3 orders of magnitude shorter, ii) an estimated number of cycles (10^8 to 10^9) for unit cell 4-6 order of magnitude higher, iii) a data retention duration (decades at 85 ^{o}C) unaffected by charge leak as Flash memory. Nevertheless, this architecture doesn’t come without negative points, such as a required current density (10^7 A/cm^2 [42]) 1-2 orders of magnitude higher than Flash [24], and different wear mechanisms such as phase segregation which leads to memory failure and are yet to be fully addressed [31], [43]. The potential of the PC memory is yet to be completely explored, with an unclear classification between mass storage and main memory; but it has the potential to bridge the gap between the two classes, offering a non-volatile memory platform with above-average cost-per-bit but lower write time than Flash technology, retaining CMOS compatibility [44].

1.4. N-vN phase-change computational devices

It is noteworthy that Ovshinsky himself started his research on PC materials with the view to design an artificial neuron [45]; in particular, the investigated PC alloy demonstrated threshold conductivity effect and accumulation effect, which mimic some of the then understood biological neuron functionalities (i.e. the neuronal input-based signalling and plasticity [46]).

The silicon technology steady improvement, combined with theoretical work on training algorithms and advances on the understanding of the biological brain, led to a more recent (and still growing – see Figure 1.3) increase in the scientific community interest for artificial adaptation and exploitation of the brain functionalities.

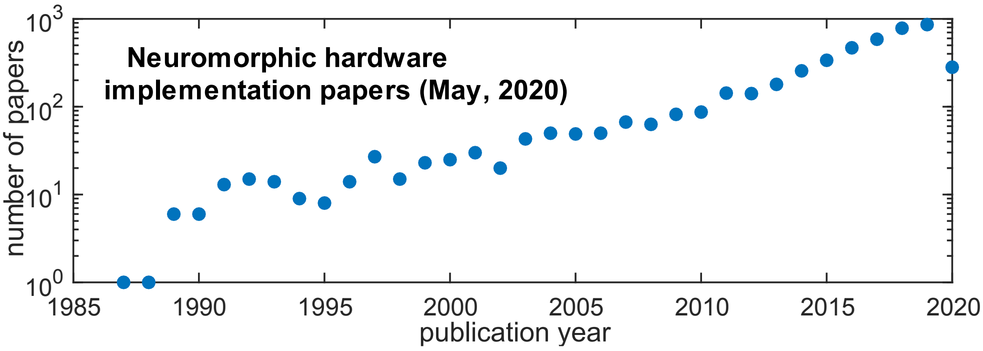


Figure 1.3. Number of scientific publications (per year) inherent to the hardware implementation of neuromorphic computing concepts, showing the exponential growth of such interest since the definition and dissemination of the concept from Carver Mead [12] (Data retrieved from Ref. [47]).

The development of resistive-based multistate memories (or memristors as they are sometimes known) has also pushed forward the research of the potential of neuromorphic devices and, in general, of the in-memory computing paradigms [15]. The reason is found in the memristor’s inherent continuous variation of the resistance level between two extremal values, and the accumulative behaviour demonstrated in PC materials makes in particular PCM devices much suitable for scopes demanding such degree of plasticity, as in the case of brain-inspired computing devices. Logic primitives [48], accumulative behaviour-based computing [49], matrix-vector multiplications [50], mixed-precision computing [51], and full-scale neural network implementations [52], [53] have all been successfully explored on PCM enabled IC platforms.

Nevertheless, the question over the suitability of electrical interconnects in a field requiring potentially millions of connections remains, as it is one with an uncertain scalability future, due to the CMOS circuitry energy requirement for data shuttling from gate to gate. It is possible to avoid this question steering away from the common CMOS technology, adopting photonics primitives to inherit the native advantages of photonics in terms of interconnects efficiency. Indeed, several interesting examples of photonic neuromorphic computing devices can be found in the scientific literature [54]–[58]. Nevertheless, a key element always hindering the practical application and commercial appeal of such proposals is found in the memory device. All-optical storage solutions are usually volatile, due to the photons being the main information carriers. All-optical memory devices then needs constant refresh of the memory state, and thus nullify one of the main advantages of the optical approach: energy efficiency. The physical mechanism for all-optical memory storage usually involves the refractive index variation, caused by carrier-induced nonlinearity. Such a phenomenon can then be further exploited and enhanced by adoption of high Q-factor structures, which i) locally magnify the electric field and consequently the light-matter interaction, and ii) increase the lifetime of the resonant state. Examples of these are found with photonic crystals [59] micro-ring resonators [60], [61] and nanocavities [62], which are characterized by a state maximum lifetime in the ns-scale. Other approaches are also being explored, such as the photon-phonon information encoding [63], but yet the information lifetime does not exceed the ns scale. Longer retention or non-volatility can be accessed by use of electro-optical implementation which would still require EOC[3], or by exploiting thermal deformation [64] or MEMS[4] [65] approaches, which suffer from increased energy consumption and does not offer the same operation temporal scale allowed by carrier-driven physical phenomena.

An avenue of work which has the potential to combine non-volatile optical domain memory storage and processing relies on the work carried out by Pernice, Bhaskaran, Wright and their groups, namely, the *phase-change photonic memory* [66]–[68] (depicted in Figure 1.4, as reported from [67]). This device shifts the storage paradigm from photon-encoding to crystal fraction encoding, and transfers the memristor primitive into the all-photonic domain: it modulates the waveguide optical attenuation in a non-volatile fashion by use of all-optical pulsing schemes, reliant on the implementation of a thin layer of PC material on the waveguide top surface.

---

[3] Electro-optical converter

[4] Micro-electro-mechanical system

This memory architecture efficiently utilises an otherwise feeble interaction between matter and light, while exploiting the wide bandwidth and potentially lossless signalling advantages of the optical domain. Experimental investigation resulted in the demonstration of multi-level storage [68] and logic gating [69], highlighting the potential to supply for traditional computing primitives. Additionally, this device has also been exploited for in-memory computing [70], [71], synaptic weighting [72] and as one of the building blocks for the photonic implementation of a spiking neural network [73], thus paving the way for n-vN light-fast processing at a fraction of the energy requirement of its electrical counterpart [74].

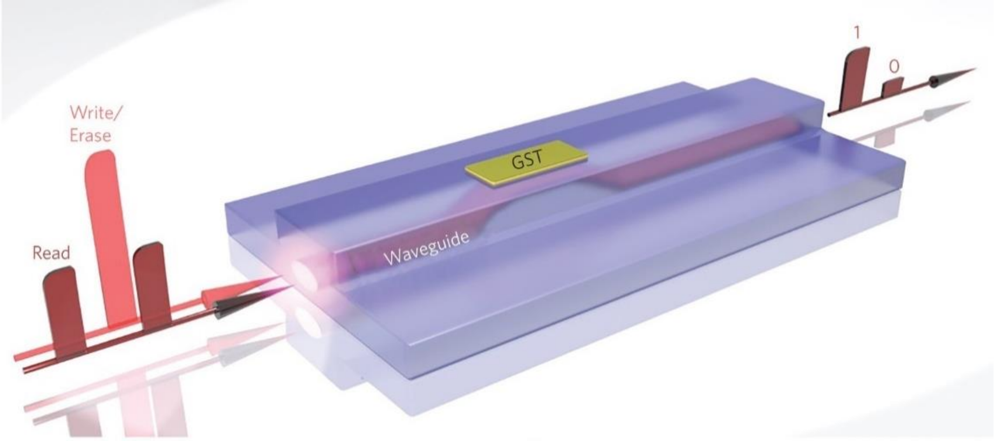


Figure 1.4. Concept of the phase-change integrated photonic memory (reprinted with permission from [67]). Original caption: “Information is stored in the phase state of the GST section on top of the nanophotonic waveguide. Both reading and writing of the memory can be performed with ultrashort optical pulses because the guided light interacts with the GST via its evanescent field. In the readout, data are encoded in the amount of optical transmission through (along) the waveguide because the two crystallographic states of the GST exhibit a high contrast in optical absorption”

1.5. Thesis objectives

This thesis is focused on the theoretical and numerical analysis of the phase-change integrated all-optical memory, with the view to unravelling its underpinning physical mechanisms. An advanced computational framework is developed for physically-realistic simulation of the optical, thermal, and phase-switching processes that lie at the heart of the operation of the all-optical photonic phase-change memory. This enables the design of novel device configurations, with improved performance capabilities (in particular low energy and fast switching). The thesis also proposes novel characterization methods to determine critical material parameters, namely the complex refractive index of phase-change materials, that in turn affect key memory performance parameters.

1.6. Thesis outline

This section provides a more focused picture of the thesis structure.

Chapter 2 provides the essential theoretical elements necessary to understand the phase-change photonic memory architecture, enabling the pursuit of this study.

Chapter 3 describes the experimental and computational methodologies used in this work. Section 3.1 shows the integrated photonic platform framework, and the elements necessary to carry out experimental verification of the analysed device. Section 3.2 describes the finite-element methodology for the device numerical analysis. Section 3.3 provides a wider perspective on the numerical techniques employed to investigate the phase transition of phase-change materials.

Chapter 4 focuses on the adopted computational techniques and illustrates the development carried out on the modelling tools. Section 4.1 provides the description of the advancement of the 3-dimensional cellular automata phase-change modelling. Section 4.2 describes the simulation framework realisation and the computational techniques to integrate the finite-element method solver onto a fully self-consistent calculator.

Chapter 5 illustrates the characterization techniques and results for the optical properties of the investigated material, specifically GST alloys. It describes a novel technique of optical property estimation, rooted in reflectivity measurements in combination with an elaborate fitting technique based on literature datasets.

Chapter 6 portrays the numerical analysis carried out on the conventional phase-change integrated photonic memory architecture. It describes the simulation results and the comparison with the current experimental data, in view of consolidating the adopted physical model and underlining its advantages, capabilities and limitations.

Chapter 7 describes a novel plasmonically-enhanced integrated phase-change photonic memory architecture, and the numerical and experimental analysis performed to explore its performance characteristics. The first section describes the fundamentals of the proposed device concept and functioning. The following sections seek a detailed analysis of the proposed device physical behaviour, operations and performance potential.

Chapter 8 summarizes the thesis content and results and provides an outlook on further development of the proposed numerical techniques and memory concept.

Chapter 2. Theoretical background

2.1. Physical mechanisms in integrated photonics platforms

Waveguides are pseudo-unidimensional devices capable of confining light within a cross-sectional area whose size scales roughly as λ^2 (with λ being the wavelength). In the majority of the waveguide designs and configurations, at least two distinct volumes can be identified: the high-index core and the low-index cladding. Usually, in integrated photonic devices, a third element can be identified, the substrate, situated at the bottom of and mechanically supports the core and the cladding (see Figure 2.1).

The following sub-sections will explore various physical mechanisms involved in the operation of photonic and optoelectronic devices relevant to this thesis, to provide the necessary background to understand the analysis and further developments presented in the following chapters.

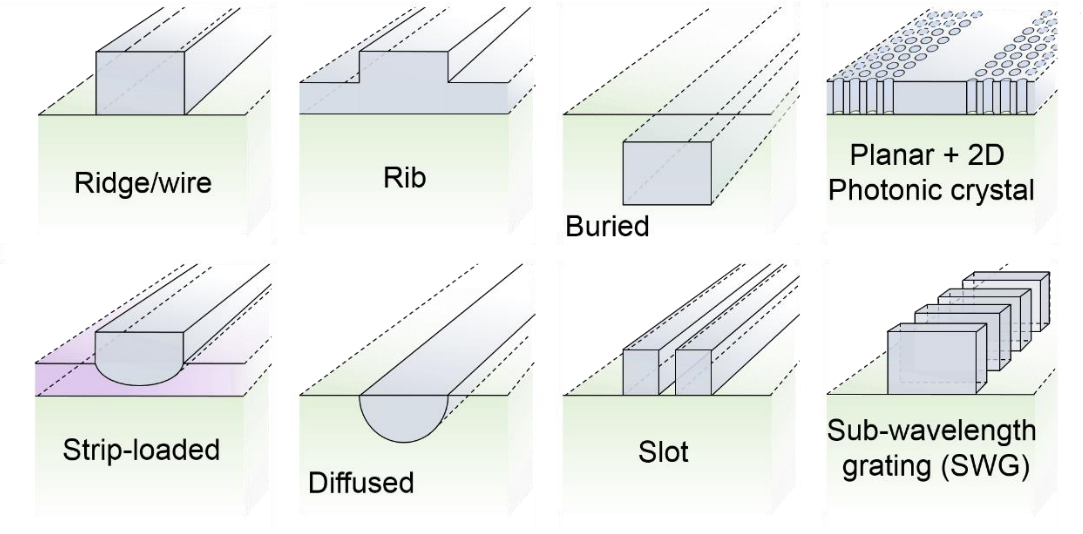


Figure 2.1. Examples of integrated platform waveguide configurations. To provide the necessary visual contrast, the substrate is coloured in green, and the core is presented in grey. The cladding is not portrayed.

2.1.1. Electromagnetic wave propagation and optical waveguide design

In a first approximation, the optical confinement within the waveguide core can be described by total internal reflection. A more thorough analysis through the Maxwell's equations is needed however to reveal beam transport properties such as the guided mode or the evanescent field, which are of fundamental importance for the investigations carried out in this work.

Maxwell’s equations relate the macroscopic spatial and temporal evolution of the electric field vector[5] $\boldsymbol{E}$ and magnetic field vector $\boldsymbol{B}$, as follows (here presented in the differential form) [75]:

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon} \qquad \text{Eq. 2.1a}$$

$$\nabla \cdot \boldsymbol{B} = 0 \qquad \text{Eq. 2.1b}$$

$$\nabla \times \boldsymbol{E} = -\mu \frac{\partial \boldsymbol{B}}{\partial t} \qquad \text{Eq. 2.1c}$$

$$\nabla \times \boldsymbol{B} = -\frac{1}{\varepsilon\mu}\frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{J} \qquad \text{Eq. 2.1d}$$

with ρ being the free electric charge density in the propagation medium and $\boldsymbol{J}$ the current density. ε and μ are the dielectric permittivity and magnetic permeability respectively, which describe the medium optical properties (which will be specifically addressed in the following section). Application of the curl on both sides of Eq. 2.1c leads to:

$$\nabla \times (\nabla \times \boldsymbol{E}) = -\nabla \times \left(\frac{\partial \boldsymbol{B}}{\partial t}\right) \qquad \text{Eq. 2.2}$$

$$\nabla \cdot (\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E} = -\frac{\partial}{\partial t}(\nabla \times \boldsymbol{B}) \qquad \text{Eq. 2.3}$$

and eventually, substituting from Eq. 2.1d and Eq. 2.1a, to:

$$\frac{1}{\varepsilon\mu}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} - \nabla^2 \boldsymbol{E} = \left(\frac{1}{\varepsilon}\nabla\rho + \mu\frac{\partial \boldsymbol{J}}{\partial t}\right) \qquad \text{Eq. 2.4}$$

which defines the general form of a 3-dimensional plane wave equation[6]. The left-hand side of Eq. 2.4 corresponds to the D’Alembert operator, acting on the electric field and conveying the wave motion; the right-hand side instead describes the sources of the electric field. For sake of simplification, the source term is here dropped, as the description will focus on the wave term. Eq. 2.4 can then be rewritten as [75]:

---

[5] In this work, the vectorial notation adopts the bold font; the vector modulus uses the roman font.

[6] An explicit choice is made here to pursue the electric field - dependent description, as in the investigated devices the materials involved usually have weak magnetic response at the spectral range of interest. A parallel formulation of Eq. 2.4 for the magnetic field can be obtained starting from Eq. 2.1d.

$$\nabla^2 \boldsymbol{E} = \frac{1}{\varepsilon\mu}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} = \frac{\tilde{n}^2}{c^2}\frac{\partial^2 \boldsymbol{E}}{\partial t^2} \qquad \text{Eq. 2.5}$$

which also includes the following additional information on the relationship between electric permittivity ε, magnetic permeability μ, refractive index n, extinction coefficient κ, and speed of light c (for isotropic materials with weak magnetic response):

$$\varepsilon = \varepsilon_0 \varepsilon_r \qquad \text{Eq. 2.6a}^{7}$$

$$\mu = \mu_0 \mu_r \qquad \text{Eq. 2.6b}$$

$$\sqrt{\varepsilon_0 \mu_0} = c \qquad \text{Eq. 2.6c}$$

$$\tilde{n} = n - i\kappa \qquad \text{Eq. 2.6d}$$

$$\varepsilon_r = \varepsilon_r' + i\varepsilon_r'' = (n^2 - \kappa^2) + i2n\kappa \qquad \text{Eq. 2.6e}$$

The general solution of Eq. 2.5 is [75]:

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{e}_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{r})} e^{i(-\omega t)} \qquad \text{Eq. 2.7}$$

where $\boldsymbol{e}_0$ is the wave polarization vector, $\boldsymbol{r} = (x \cdot \hat{x}, y \cdot \hat{y}, z \cdot \hat{z})$ is the spatial coordinate vector, t is the temporal coordinate, and $\boldsymbol{k}$ and ω are the wave propagation vector and wave frequency, which take the following values:

$$|\boldsymbol{k}| = k = \omega / c \qquad \text{Eq. 2.8a}$$

$$\omega = 2\pi c / \lambda \qquad \text{Eq. 2.8b}$$

It is possible to further simplify Eq. 2.7, by separating the propagation axis variable (here defined as z) from the remaining ones (x,y). As a result, the solution reported in Eq. 2.7 takes the following form [75]:

$$\boldsymbol{E}(x, y, z, t) = \boldsymbol{U}(x, y) e^{-\gamma z} e^{i(-\omega t)} \qquad \text{Eq. 2.9}$$

This formulation describes a distinct electric field distribution which can be applied to any waveguiding device. The transverse vectors (along the x and y directions)

---

[7] ε and μ are defined as absolute permittivity and permeability respectively. For sake of simplification, the values of those properties in void are expressed as ε_0 and μ_0; the relative permittivity and permeability, ε_r and μ_r, are then characterizing any material property in relation to the void properties.

are collected under the term $\boldsymbol{U}$, which will be analysed later in this section. The initial focus is then brought to the exponential terms.

The electric field temporal variation is trivial, in which the field varies harmonically. Also, the propagation vector terms vary according to a simple exponential function, $e^{-\gamma z}$, with γ being the complex propagation constant [75]:

$$\gamma = \alpha + i\beta \qquad \text{Eq. 2.10}$$

where α is the wave Napierian attenuation constant, which appears whenever the waveguide introduces scattering or absorption loss contributions. In particular, it is possible to estimate the intensity of the travelling wave as follows, by substituting Eq. 2.10 into Eq. 2.9:

$$\begin{aligned} I(z) \propto |\boldsymbol{E}(x,y,z,t)|^2 &= |\boldsymbol{U}(x,y)|^2 \cdot \left|e^{-2\gamma z} e^{i(-2\omega t)}\right| \\ &= |\boldsymbol{U}(x,y)|^2 \cdot \left|e^{-i2\beta z} \cdot e^{-2i\omega t}\right| \cdot e^{-2\alpha z} \\ &= |\boldsymbol{U}(x,y)|^2 \cdot e^{-2\alpha z} \qquad \text{Eq. 2.11} \end{aligned}$$

Thus, the intensity dependency on the propagation axis is an exponential decay, similar to the Beer-Lambert’s law.

The second right-hand term of Eq. 2.10, β, is the wave phase constant, representing the phase shift per unit length along the waveguide propagation direction. Whilst this term allows one to derive the phase velocity, the propagation speed of the travelling wave corresponds to the group velocity, $v_g = \frac{\partial \omega}{\partial k}$.

It is also useful to introduce an alternative notation which characterizes the wave propagation, namely the *effective refractive index*, $\tilde{n}_{eff}$. The $\tilde{n}_{eff}$ can be treated as the complex index for free-propagating waves, but includes in its formulation the effect of the materials and geometry of the waveguiding device. Specifically, the following relation between γ and $\tilde{n}_{eff}$ is found [75]:

$$\gamma = ik\tilde{n}_{eff} = ik(n_{eff} - i\kappa_{eff}) \qquad \text{Eq. 2.12}$$

with the terms n_{eff} and κ_{eff} determining once more the travelling wave phase velocity and loss coefficient respectively. By combining Eq. 2.6 with Eq. 2.12, and deriving with respect to $k = 2\pi/\lambda$, one can write the group refractive index as [75]:

$$n_g(\lambda) = n_{eff}(\lambda) - \lambda \frac{\partial n_{eff}}{\partial \lambda} \qquad \text{Eq. 2.13}$$

which consequently determines the group velocity as $v_g = c/n_g$.

Returning to Eq. 2.9, the cross-sectional field distribution is fully described by the term $\boldsymbol{U}(x, y)$, which may contain both transverse field and longitudinal field components. $\boldsymbol{U}(x, y)$ is found by means of the Helmholtz equation solution [75]:

$$(\nabla^2 + k^2 - \gamma_\nu^2)\boldsymbol{U}_\nu(x, y) = 0 \qquad \text{Eq. 2.14}$$

specifying that $k = \tilde{n}k_0$, with k_0 being the wave propagation constant in vacuum. Usually, for integrated waveguides, the Helmholtz equation has no analytical solutions [75]; but approximated or numerical methods are routinely applied to find these, such as the effective index method [76], [77], the Marcatili method [77], [78], or finite element solvers. The index ν applied to $\boldsymbol{U}$ (and γ) labels represents each of the allowed solutions for the specific waveguide, the so-called *propagation modes* (in general, for channel waveguides ν is an integer number, equal to the cardinality of the intensity peaks set of the transversal plane). For any waveguide, a finite number of modes allow for a harmonic solution of the $\boldsymbol{U}_\nu$ term, which then identify a confined, or *guided mode*. Any non-guided mode, or *leaky mode*, eventually dissipates all its optical energy outside the waveguide, but the guided modes allows for the propagation of the optical energy along the waveguide optical axis, with the intensity of the propagating beam driven by Eq. 2.11.

In planar waveguides, i.e. devices which confine the propagation in one direction only, two cases are distinguished, depending on the vanishing of the z-direction field terms. The TE (*transverse electric*) mode sees $E_z = 0$, whereas the TM (*transverse magnetic*) mode has $H_z = 0$. In integrated channel waveguides such as the ones depicted in Figure 2.1, which bind the propagation on two directions, this condition is not always satisfied. The distinction is then applied to the relative magnitude of the fields in the x and y directions[8]; TE modes have $E_x \gg E_y$, whereas TM modes have $E_x \ll E_y$.

---

[8] Since the condition on the vanishing E_z and H_z are not satisfied in channel waveguides, literature adopts the notation of quasi-TE and quasi-TM mode in place of TE and TM mode respectively, to address these unequivocally. For sake of conciseness the ‘quasi-’ prefix is here omitted but implied.

An example of numerically obtained TE and TM modes for a silicon nitride on silica rib waveguide is reported in Figure 2.2. The reported waveguide structure is specifically optimized for TE-mode C-Band transmission (1525 nm - 1570 nm), but the concept applies to any optical waveguide. In Figure 2.2(a, b) the 2D plots represent the transversal (normalized) electric field. Interpolation of the data calculated on the horizontal and vertical cut lines passing through the waveguide centre (also reported as blue and red dashed lines in Figure 2.2(a)) is reported in Figure 2.2(c,d) and Figure 2.2(e, f) respectively. In all cases, it is notable how the solutions see a non-zero electrical field outside the waveguide region. Figure 2.2(c-f) confirms this observation. More specifically, the electric field decay outside the core region is found to closely follow an exponential function (fits by

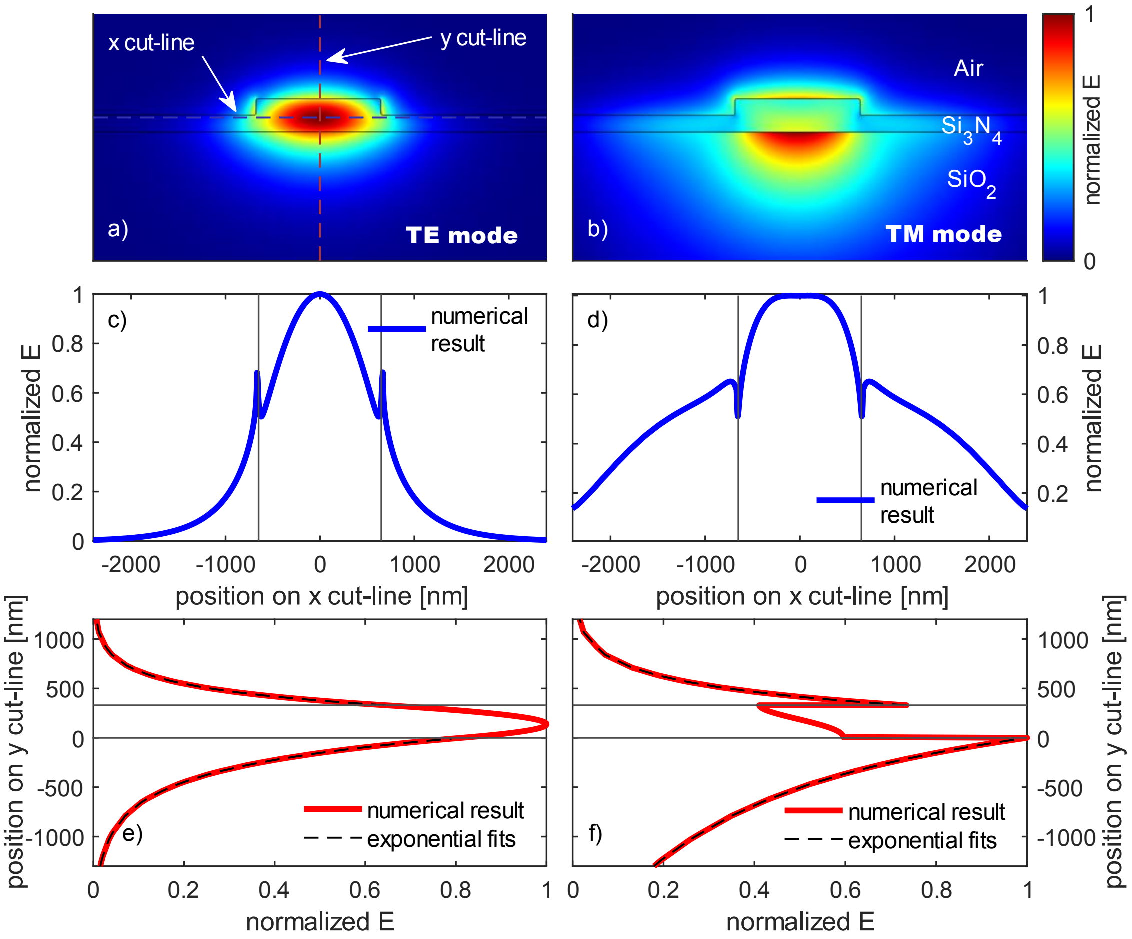


Figure 2.2. TE and TM modes profiles of a SiN-on-silica rib waveguide, optimized for C-Band transmission (width of 1300 nm, height of 330 nm, etch step of 165 nm). TE mode solutions are reported on the left (a,c,e); TM mode solutions are reported on the right (b,d,f). (a,b) 2D representation of the transversal electric field solution (normalized), as calculated with finite-element methods (colour bar reported on the right). (c,d) 1D plot of the solution, as extrapolated at the waveguide mid-hight cut-line (represented in (a) as “x cut-line”). (e,f) 1D plot of the solution, as extrapolated at the waveguide vertical symmetry axis (represented in (a) as “y cut-line”). These also include the exponential fit of the field components outside the waveguides (black dashed lines).

$y = \exp\big((x - x_0)/\sigma\big)$ are also plotted in Figure 2.2(e, f)). This region of the electric field is referred to as *evanescent field*, and accounts for a significant portion of the mode electric field, albeit rapidly decaying in a few hundreds of nm (see Figure 2.2(c-f)).

2.1.2. Material optical properties

Maxwell's equations link the electric and magnetic field variation to the material properties, via the dielectric permittivity ε and magnetic permeability μ. Once more, the magnetic properties are here neglected, to focus on the description of materials with a weak magnetic response (i.e. $\mu_0 \approx 1$). The information about the medium-radiation interaction is then fully determined by the dielectric permittivity $\varepsilon = \varepsilon_0\varepsilon_r$. Starting from this consideration, the definition of displacement field D, polarization density P, and electric susceptibility χ_e are provided in the following constitutive relations [79]:

$$\boldsymbol{D} = \varepsilon\boldsymbol{E} = \varepsilon_0\varepsilon_r\boldsymbol{E} = \varepsilon_0\boldsymbol{E} + \varepsilon_r\boldsymbol{E} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P} = \varepsilon_0(1 + \chi_e)\boldsymbol{E} \qquad \text{Eq. 2.15a}$$

$$\boldsymbol{P} = \varepsilon_r\boldsymbol{E} = \varepsilon_0\chi_e\boldsymbol{E} \qquad \text{Eq. 2.15b}$$

$$\chi_e = \frac{\boldsymbol{P}}{\varepsilon_0\boldsymbol{E}} = \varepsilon_r - 1 \qquad \text{Eq. 2.15c}$$

The polarization $\boldsymbol{P}$ describes the density of the material permanent and induced electric dipole moments. The susceptibility χ_e quantifies the magnitude of the polarization vector[9]. As visible from Eq. 2.15, the calculation of ε_r depends on the statement of a model for the susceptibility, which is indeed the most convenient natural variable for the description of a medium behaviour when subjected to *electromagnetic* (e.m.) radiation. What follows is then the derivation of the susceptibility from a semi-classical description.

An electric dipole can be formed whenever a distortion of the charge distribution in an atom or molecule occurs. In solid media, this is strictly related to the displacement of the electron distribution with respect to the nucleus (distortional polarization). This phenomenon follows the stimulated damped harmonic oscillator equation. An example for the simple case of one electron moving around a nucleus, stimulated by an external harmonic field (see Eq. 2.7), is provided below, by means of the Drude-Lorentz description [79]:

---

[9] Whilst in the example case susceptibility and dielectric permittivity are portrayed as scalar values, in full generality these are second rank tensors [79], which for ease of computation and notation can be reduced to scalar values in case such tensors are scalar matrices (in homogeneous, isotropic materials).

$$m_e \frac{d^2\boldsymbol{r}}{dt^2} + m_e \Gamma \frac{d\boldsymbol{r}}{dt} + m_e \omega_0^2 = -q_e \boldsymbol{E}(\omega) \qquad \text{Eq. 2.16}$$

with m_e and q_e being the electron mass and charge respectively, Γ the oscillator damping factor, and ω_0 the system eigenmode (the natural oscillator frequency). An illustrative description is provided in Figure 2.3.

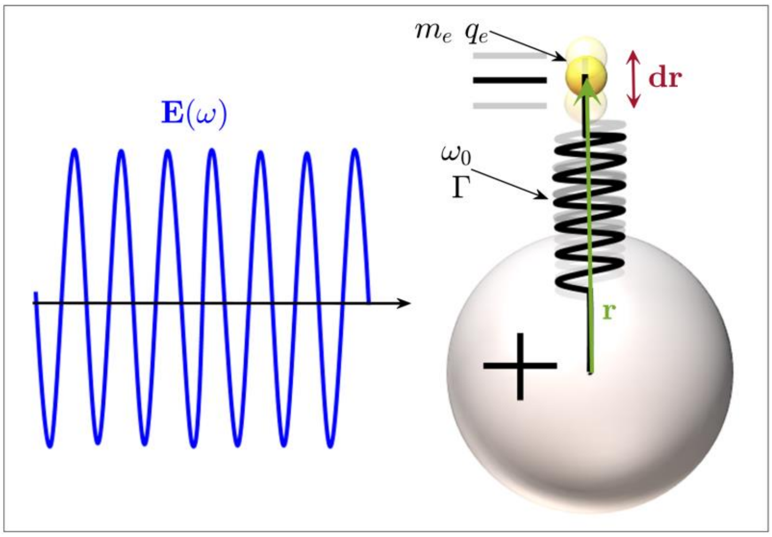


Figure 2.3. Representation of the nucleus - electron schematics, upon which the Drude-Lorentz model is based.

The $m_e \frac{d^2 r}{dt^2}$ (and $m_e \Gamma \frac{dr}{dt}$) terms describe the (damped) motion of free electrons subjected to the external field, whereas the $m_e \omega_0^2$ term includes the contribution of the bound electrons, which are involved in interband transitions (i.e. absorption and emission processes due to a transition between states $\nu_a \rightarrow \nu_b$ of characteristic energy $\hbar\omega_0$). Remembering the time dependency of the electric field, the solution of Eq. 2.16 for the spatial distribution $\boldsymbol{r}(\omega)$ becomes [79]:

$$\boldsymbol{r}(\omega) = -\frac{q_e}{m_e} \frac{\boldsymbol{E}}{\omega_0^2 - \omega^2 - i\Gamma\omega} \qquad \text{Eq. 2.17}$$

Multiplication by $-q_e$ both the left and right hands of Eq. 2.17 yields the local description of the polarization:

$$-q_e\boldsymbol{r}(\omega) = \frac{q_e^2}{m_e}\frac{\boldsymbol{E}}{\omega_0^2 - \omega^2 - i\Gamma\omega} = \boldsymbol{p}(\omega) \qquad \text{Eq. 2.18}$$

The latter is a useful analytical model for the local phenomenon (i.e. taking place on a single medium constituent unit, such as ions or molecules), which can then be related to the macroscopic polarization field $\boldsymbol{P}(\omega)$. Specifically, to obtain the emerging behaviour of $\boldsymbol{P}(\omega)$ from $\boldsymbol{p}(\omega)$, it is necessary to consider the statistical average of the widely varying polarization field generated by atomic charges. Thus, it is possible to write the following relation [79]:

$$\boldsymbol{P}(\omega) = N\langle\boldsymbol{p}(\omega)\rangle \qquad \text{Eq. 2.19}$$

with N being the density of atoms or molecules in the medium. Substitution of Eq. 2.18 in Eq. 2.19 provides the following description for the macroscopic polarization derived from the local electronic distribution arrangement:

$$\boldsymbol{P}(\omega) = N\frac{q_e^2}{m_e}\frac{\boldsymbol{E}}{\omega_0^2 - \omega^2 - i\Gamma\omega} \qquad \text{Eq. 2.20}$$

One last substitution of Eq. 2.20 in Eq. 2.15c eventually yields the desired formulation for the (complex-valued) susceptibility [79]:

$$\begin{aligned}\chi_e(\omega) = \frac{\boldsymbol{P}(\omega)}{\varepsilon_0\boldsymbol{E}} &= \frac{1}{\varepsilon_0}\frac{Nq_e^2}{m_e}\frac{1}{\omega_0^2 - \omega^2 - i\Gamma\omega} \\ &= \frac{1}{\varepsilon_0}\frac{Nq_e^2}{m_e}\frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} \\ &\quad + i\frac{1}{\varepsilon_0}\frac{Nq_e^2}{m_e}\frac{\Gamma\omega}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} \qquad \text{Eq. 2.21}\end{aligned}$$

which is here also separated in its real and imaginary components. The dielectric permittivity is then straightforwardly obtained combining Eq. 2.21 and Eq. 2.15c:

$$\begin{aligned}&\varepsilon_r(\omega) = \varepsilon_r'(\omega) + i\varepsilon_r''(\omega) \\ &= 1 - \omega_p^2\frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} + i\omega_p^2\frac{\Gamma\omega}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} \qquad \text{Eq. 2.22}\end{aligned}$$

which includes the additional definition of *plasma frequency* ω_p [79]:

$$\omega_p = \sqrt{\frac{Nq_e^2}{\varepsilon_0 m_e}}. \tag{Eq. 2.23}$$

which denotes the frequency at which the electron ensemble of the medium oscillates at the same phase (thus leading to infinite phase velocity). Many other extensions to this model are also found in literature, to include factors such as multiple damping terms, inhomogeneous broadening, multiple-level transitions, and behaviour towards the limits [79], [80]. Additionally, metals usually do not display the damped system mechanism with radiation interacting with free electrons, thus the ω_0 term in Eq. 2.22 is vanishingly small. This leads to the following rearrangement of Eq. 2.22 [79], useful to mimic a metal-like response:

$$\varepsilon_{free}(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\Gamma\omega} \tag{Eq. 2.24}$$

As a visual example, two electric permittivity datasets are reported in Figure 2.4.

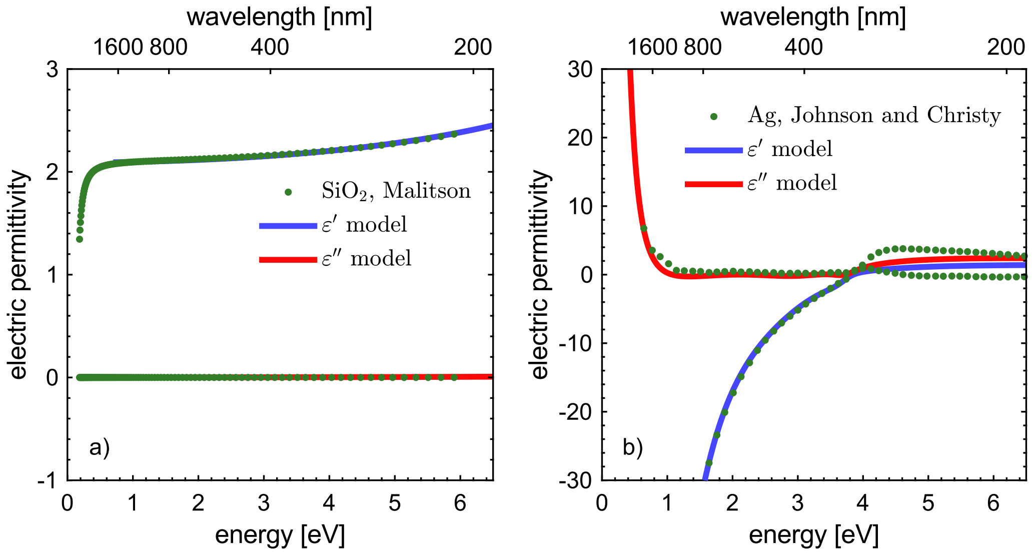


Figure 2.4. Electric permittivity and fit models. a) SiO$_2$: experimental data from [81] (green dots), fit model adapted from [82] (continuous lines). b) Ag: experimental data from [83] (green dots); model from [84] (continuous lines).

The first is a typical and well-studied insulator, fused-silica (SiO$_2$). The fit model [82], derived from Eq. 2.22 and adapted for the reference data [81], uses a single Lorentzian resonator ($\omega_0 = 13$ eV, out of the plotted spectral range; $\Gamma = 0.1$, and $\varepsilon_s = 1$, where ε_s is the static permittivity). Indeed, the model is capable of following the electric permittivity trend across a large portion of the spectrum, from the near UV to the near IR, although further from this range the reliability decreases. The second case (Figure 2.4b) reports the data of an archetypical conductor, polycrystalline silver [83]. The reported model uses 4 distinct Lorentzian oscillators [84], which is purposely built to reproduce the electric permittivity trend at low frequencies, but can follow the general features also at higher energies (although, once more, with lower reliability, see mismatch above 4 eV; models tailored for the UV-visible range can also be found in literature, see e.g. [85]).

2.1.3. Plasmonic effects

Another solution of Eq. 2.5 is found for systems comprising an interface between an insulator and a conductor. Let the case be of an e.m. wave, travelling along the z direction, with the electric field varying only on the y direction (TM mode); let the insulator-conductor interface be on the xz plane, with optical properties ε_1 and ε_e for the insulator and conductor respectively (see Figure 2.5 for a visual representation of the system). The associated wave equation takes the following form [86]:

$$\boldsymbol{E}(x, y, z, t) = \boldsymbol{U}(y) e^{i(k_z z - \omega t)} \qquad \text{Eq. 2.25}$$

which is conceptually similar to Eq. 2.9. Inserting the Helmholtz equation, and solving for the propagation constant γ by considering the appropriate boundary conditions such as the continuity of the fields ($\varepsilon_1 E_y = \varepsilon_2 E_y$, and $k_{2,y}/k_{1,y} = -\varepsilon_2/\varepsilon_1$) along with the counterpart for the H field, it is possible to obtain the dispersion relation of the so-called *Surface Plasmon Polariton* (SPP) [86]:

$$\gamma_z = k_z' + i k_z'' = k_0 \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}} \qquad \text{Eq. 2.26a}$$

$$\gamma_{j,y} = i k_{j,y}'' = k_0 \sqrt{\frac{\varepsilon_j^2}{\varepsilon_1 + \varepsilon_2}} \qquad \text{Eq. 2.26b}$$

with the label j being an index identifying either the insulator (1) or the metal (2). The SPP mechanism describes the collective surface-bound electron excitations, propagating via Eq. 2.26 across the interface. The field in both the propagation direction and the out-of-plane direction is evanescently decaying, with three different 1/e characteristic lengths ($L_z = \frac{1}{2k_z''}$ and $L_{i,y} = \frac{\lambda}{2\pi}\sqrt{\frac{|\varepsilon_1'|+\varepsilon_2}{\varepsilon_j^2}}$ [86]). It is important to underline that such mechanism finds a real-valued $\gamma_z = \beta = k_z'$ when $\varepsilon_2' < 0$ and $|\varepsilon_2'| > |\varepsilon_1'|$.

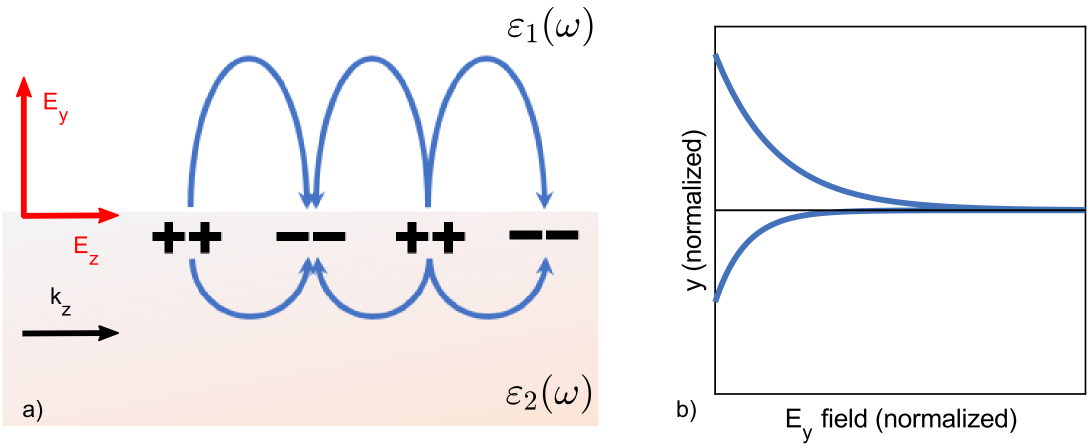


Figure 2.5. SPP mechanism schematics. a) prospect of the metal-insulator interface, and electric field lines (red: TM input field, blue: SPP field). b) Plot of the field magnitude as a function of the vertical position, showing the relative higher magnitude on the insulator side (arising from the field continuity condition when $|\varepsilon_2'| > |\varepsilon_1'|$) and the exponential decay trend.

Similarly to the plasma frequency condition, it is found that for surface modes, $\beta \rightarrow \infty$ when the following is satisfied [86]:

$$\omega_{sp} = \frac{\omega_p}{\sqrt{1+\varepsilon_1}} \qquad \text{Eq. 2.27}$$

which can be found when $\varepsilon_2(\omega) = -\varepsilon_1(\omega) = 1 - \omega_p^2/\omega^2$. SPPs are two-dimensional leaky e.m. waves, propagating at the insulator-conductor interface. *Surface Plasmon Resonance* (SPR), found when $\omega \rightarrow \omega_{sp}$, is in contrast a non-propagating collective oscillation. Use of SPPs for mode guiding is generally not feasible, yet plasmonic resonances possess a peculiar optical behaviour, as the e.m. field enhancement at the surface reaches its peak, as well as the absorbed and diffused components of the e.m. field.

In finite-sized metal particles surrounded by a dielectric medium, the wave propagation description is not applicable. Nevertheless, also such particles have alluring optical properties, arising from both the geometrical configuration and the material optical properties, which may support *localized surface plasmon* *resonances* (LSPR). In particular, resonant nanoparticles possess a rather large extinction cross-section, high absorption, scattering and diffusion coefficients, and a large e.m. field enhancement in proximity of the metallic surfaces. The exploitation of such effects in this thesis is discussed in Chapter 7.

2.2. Thermal diffusion and thermal boundary resistance

The thermal diffusion model enables the tracking of the temperature (and thermal energy) evolution, as a function of time and stimuli, of a given system. As it will be clearer in the following chapters, the temperature variation in space and time has dramatic influence over the optical behaviour of the devices investigated in this thesis, thus the focus of this section.

Fourier [80] first reported the analytical formulation of the heat diffusion by conduction, for an isotropic homogeneous medium, as follows (here in differential form, for a 3-dimensional object):

$$\boldsymbol{q} = -k_T \nabla T(t) \qquad \text{Eq. 2.28}$$

where $\boldsymbol{q}$ is the heat flux, k_T[10] is the thermal conductivity, and $\boldsymbol{T} = \boldsymbol{T}(r,t)$ the time-dependent temperature field.

By the 1st law of thermodynamics [80] (energy conservation), any infinitesimal variation of a closed system internal energy dU must be compensated by the net heat flow dQ crossing the system boundary, and the additional mechanical work $dW = -PdV$. Assuming constant pressure, and by use of the specific enthalpy of the closed system $H(T,P,V) = U(T,P,V) + PV$, the following arises [87]:

$$dU = dQ - PdV \qquad \text{Eq. 2.29a}$$

$$dH = dU + PdV + VdP = dQ - PdV + PdV + VdP = dQ \qquad \text{Eq. 2.29b}$$

with the VdP term neglected due to the constant pressure assumption. Deriving Eq. 2.29b with respect to temperature follows [87]:

---

[10] The □$_T$ subscript is here arbitrarily applied, to distinguish the thermal conductivity from the k wavevector notation.

$$\frac{dQ}{dT} = \frac{dH}{dT} = \rho c_P \qquad \text{Eq. 2.30}$$

with ρ being the medium density and c_P the heat capacity. The integration of Eq. 2.30 in dT provides a tool to relate the thermal energy with the medium temperature variation; and the following derivation by time describes the rate of change of the thermal energy [87]:

$$\frac{dQ}{dt} = \rho c_P \frac{dT}{dt} \qquad \text{Eq. 2.31}$$

For any point in the system, the infinitesimal variation $\frac{\partial Q}{\partial t}$ must be matched by the variation of the heat flux in space $\frac{\partial q}{\partial r}$, once more due to energy conservation considerations. It is then possible to write the following, by combining Eq. 2.28 and Eq. 2.31 [87]:

$$\frac{dQ}{dt} = -\nabla q \qquad \text{Eq. 2.32a}$$

$$\rho c_P \frac{dT}{dt} = -\nabla(-k_T \nabla T) \qquad \text{Eq. 2.32b}$$

To provide full flexibility of application, Eq. 2.32a usually includes the source of external perturbation, as follows [87]:

$$\rho c_P \frac{dT}{dt} + \nabla(-k_T \nabla T) = Q_v \qquad \text{Eq. 2.33}^{11}$$

with Q_v being the *volumetric heat source*, which can be obtained as the spatial and temporal integral of the dissipated power. Eq. 2.33 is the heat transfer equation, mostly applied throughout the investigations reported in this work. Other contributions, such as the temperature spatial first derivative (which models material transfer, i.e. convective heating), radiative heat flux (which can be introduced with an additional q contribution) , and stress work (appearing as an additional energy term on the right-hand side) are not applied in this work. Heat

---

[11] The right-hand term of Eq. 2.33 would also need to contain the latent heat of transformation in case of phase transition (as well as the parameters dependencies from crystal fraction and temperature). The phase-transition process is thoroughly discussed in Section 2.3. It is here anticipated though that the latent heat of transformation contribution brings negligible variation of the here portrayed experimental results, and therefore it is omitted in Eq. 2.33 for sake of simplification.

by conduction is then the predominant mechanism, which is the macro-scale average behaviour emerging from two main mechanisms: phonon propagation, and electronic heat conduction. Whilst this is a good approximation in most of the cases, a specific aspect linked to the meso-scale and nano-scale features is responsible for a considerable divergence from such law, and thus must be taken into account separately. Specifically, the presence of surface roughness or reticular defects (e.g. lattice mismatch, dangling bonds, heteroatoms, missing atoms, substitutions, etc.) may lead to a substantial deviation from the expected thermal behaviour. This element is prevailing at interfaces between solids, where the conduction mechanism can not easily calculate the correct heat flux (see Figure 2.6 for a visual representation of the examined case). To account for this contribution, a simple and yet effective method relies on the introduction of the *thermal boundary resistance* (TBR), which enforces a limit towards the heat flux crossing an a, b interface as follows [87]:

$$\boldsymbol{u} \cdot q_{a,b} = \frac{T_b - T_a}{R_{TBR}} \qquad \text{Eq. 2.34}$$

where R_{TBR} is the phenomenological variable describing such heat flux limitation ($\boldsymbol{u}$ being the vector normal to the a, b interface).

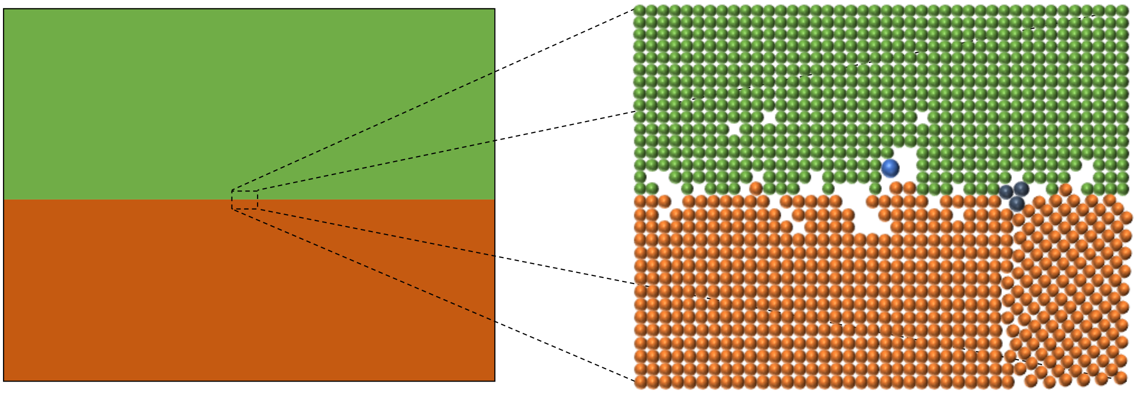


Figure 2.6. Visual representation of the nano-scale defects at the interface between solids, here depicted in two distinct colours (left: macroscopic picture, right: nano-scale representation), giving rise to the localized decrease of the thermal conductivity phenomenologically described as thermal boundary resistance (TBR).

The TBR method indeed simplifies the theoretical analysis of nano-scale devices, allowing one to adopt the macroscopic heat conduction description in presence of non-classical elements.

2.3. Phase-change materials

All materials undergo phase changes, but in the fields of electronics and photonics the term phase-change material is usually reserved for materials in which the following is satisfied [31]:

- Enable the reversible phase transition between 2 or more phases
- Each phase possesses a distinct set of optical/electrical properties
- The timescale of such transition belongs to the sub-second regime, with particular accent on the ps to µs scale.

An additional characteristic, not strictly required to define the PC material class but quite important from the technological perspective, is the switching cyclability; depending on the application, an endurance of 10^6 to 10^{12} cycles might be required.

The switching in PC materials can be immediately followed by a return to the equilibrium phase (*volatile*), or on the contrary having long-lasting effects, with no immediate return to the original state (non-*volatile*). The archetypal volatile PC material is vanadium oxide V_xO_y, whose *insulator-to-metal transition* (IMT) mechanism was first reported by Morin in 1958 [88]. More specifically, VO_2 is characterized by a hysteretic behaviour of the optical properties (refractive index) which is directly correlated to the monoclinic-to-rutile phase transition around 340 K. Other examples of materials exhibiting the IMT transition are Ti_2O_3 [88], and rare-earth nickelates such as SmNiO3 [89].

A different class of PC materials relies on the *glass-forming* capability, which is inherently a non-volatile transition. Chalcogenides are a typical example of this case. Chalcogenides are, in full generality, alloys containing elements of the 6th group (sulphur, selenium, and tellurium[12]) combined with electropositive elements (usually, elements from the 3rd, 4th and 5th groups).

This section is an overview of the non-volatile phase-change materials technological history, physical mechanisms, and applications, in particular oriented to the archetypical chalcogenide PC alloy, $Ge_2Sb_2Te_5$ (GST).

---

[12] Oxygen compounds, whilst technically chalcogenides, are traditionally referred to as oxides.

2.3.1. Phase transition mechanisms

In general terms, for a closed system in equilibrium, composed by C components and P phases, the Gibbs phase rule [87] determines the number of thermodynamic degrees of freedom F, i.e. the number of independent, intensive, state variables (e.g. pressure, temperature, crystal fraction, chemical potential, etc). Assuming constant pressure (i.e. further removing a degree of freedom), the Gibbs phase rule is stated as:

$$\mathrm{F} = \mathrm{C} - \mathrm{P} + 1 \qquad \text{Eq. 2.35}$$

For a single-component mixture of phases, only $\mathrm{P} = 2$ is an allowed solution (with $\mathrm{F} = 0$). Henceforth, only two phases can coexist in equilibrium, with a single transition temperature T_{tr} that determines the zero-dimensional boundary between the two phases: below or above T_{tr}, only one phase exists; at T_{tr}, a mixture of the two phases can be found. The single-valuedness of T_{tr} is also ensured by the Gibbs phase rule, as T_{tr} is an intensive state variable, and by assumption, $\mathrm{F} = 0$. Whilst at equilibrium (by definition) the Gibbs free energy of the two phases is equal, there is a non-zero variation of the enthalpy that determines the reaction direction. Thus, during a phase transition, the enthalpy of transition ΔH_{tr} determines the fraction of the relative ratio of phases f as a function of the provided energy Q as follows:

$$f = \frac{Q}{\Delta H_{tr}} \qquad \text{Eq. 2.36}$$

Additionally, Eq. 2.36 implies that the temperature of the closed system does not change during the phase transition. Temperature invariance has a drastic impact on the specific heat. In first-order transitions, the following applies [87]:

$$\lim_{\partial T \to \infty} c_P = \lim_{\partial T \to \infty} \left(\frac{\partial H}{\partial T}\right)_P = \infty \qquad \text{Eq. 2.37}$$

Therefore, in the simplistic example case, a phase transition is sourced by an energy source Q, and driven by the chemical potential, which determines at a given temperature the most stable of the phases, or the ratio f thereof at the transition temperature T_{tr}. The process is reported in Figure 2.7(a).

Turning the attention once more to the Gibbs phase rule, but in multi-component systems, it is possible to see that for $C > 1$, $\mathrm{F} > 0$. Henceforth, T_{tr} varies over a range delimited by two boundaries. In case of the solid-liquid transition (which is usually incurring in PC materials), these boundaries are the *solidus temperature* and the *liquidus temperature*, which can be analytically determined via application of the Gibbs-Duhem law [87] (which states the zero-sum of all the derivatives of the partial molar properties). The derivation won't be here reported, as the intent is to provide an intuitive vision of a generic phase transition for binary compounds. As reported in Figure 2.7(b), such boundaries are usually distinct, which implies that the transition process undergoes the formation of a mixture of phases, whose composition can drastically differ from the original one[13].

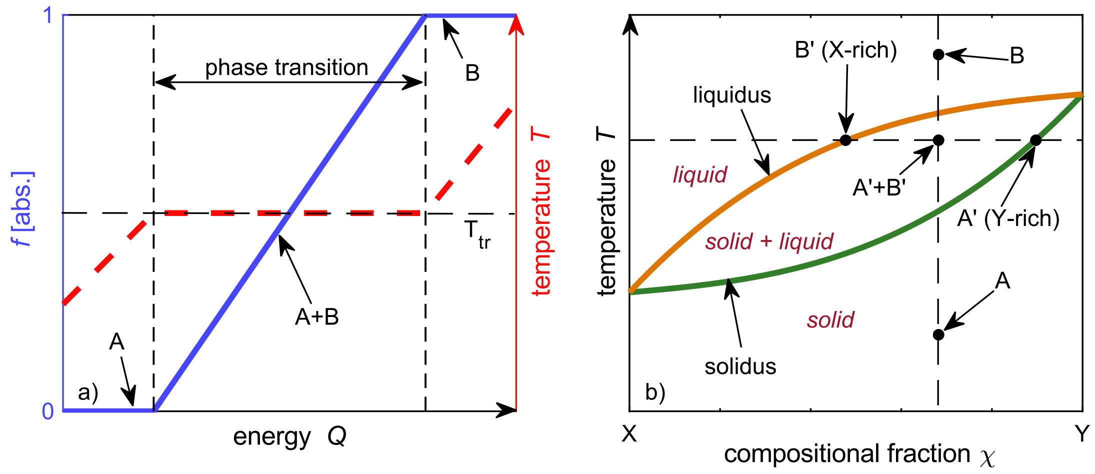


Figure 2.7. Phase transition process. a) Phase fraction f (blue, continuous line) and temperature T (red, dashed line) as a function of the provided energy Q. Notably, temperature remains constant at the transition temperature T_{tr} for the duration of the process, where instead f shifts linearly from 0 (*A* phase) to 1 (*B* phase). b) Phase diagram for an XY binary compound (stoichiometry vs temperature plot). An example composition ($\chi = 0.66$) is highlighted, whose transition from the A phase (solid) towards the B phase (liquid) takes place via the intermediate mixture of Y-rich solid phase and X-rich liquid phase, when between the liquidus line (orange) and solidus line (green).

Whilst the thermodynamic description above is sufficiently accurate in certain scenarios, a macroscopic description can not easily unravel the nature of the phase transition at the spatial scale of its constituents, such as the formation of crystal nuclei and growth of these within an amorphous matrix. Furthermore, there are many additional considerations to the reported thermodynamic description which steer the phase transition mechanism away from the example ideal cases, such as the kinetics of reaction. These will be explored in the following subsections.

---

[13] This is also a well-known PC material-implemented device source of degradation [334], along with the reactivity with other elements (see. e.g. [335]–[337]).

2.3.1.1. Homogeneous nucleation

Preliminarily, it is useful to provide the definition of *cluster*, which is a solid aggregate of unit atoms or molecules. Let it be a medium of phase *A*, within which the formation of a spherical cluster of phase *B* is observed. This phenomenon can be described in thermodynamic terms [87] by definition of the free energy of transition, ΔG_{nu}, whose value is determined by the energy required to form a new *AB* surface, the energy dissipated by the *A* volume loss, and the energy required by the *B* volume appearance. Volume and surface values associated with a spherical cluster (of radius r) are trivially calculated as $\frac{4}{3}\pi r^3$ and $4\pi r^2$ respectively. The intrinsic properties which need to be considered are the following:

- The volume free energy term, which can be summarized as the difference in free energy between the *B* and *A* phases, here labelled ΔG_v.
- Also dependent on the volume, the misfit strain energy ε, arising from the different densities of *A* and *B*.
- The surface energy term, σ, which includes the effect of the surface formation between *A* and *B*.

These can be combined to define the free energy variation ΔG_{nu} as follows [31], [90]:

$$\Delta G_{nu} = \Delta G_{nu}(r) = \frac{4}{3}\pi r^3(\Delta G_v + \varepsilon) + 4\pi r^2\sigma \qquad \text{Eq. 2.38}$$

with the underlying assumption that the surface energy term is independent of the cluster volume, and that there is no intermixing layer between *A* and *B*. By its definition, a negative ΔG_{nu} value leads to the spontaneous nuclei formation. It is then important to highlight the effect of each of the contributions in Eq. 2.38. For such an equation to be negative, at least one of the two contributions (volume and surface) must be negative. By definition, $\sigma = \left(\frac{\partial \mathrm{G}}{\partial a}\right)_{p,T}$ [87], [90] with a being the surface of the spherical cluster. If negative, any closed system would see a decrease of the free energy by creation of surface, hence being unphysical. Thus, the driving force of the spontaneous formation of a spherical cluster is a negative value of ΔG_v (note that strain contribution can either hinder or promote such effect depending on the individual cases). Figure 2.8 reports the variation of the volume terms and surface term as a function of the radius, as well as the nucleation free energy variation.

For real positive values of the radius, both volume terms and the surface term are monotonic. As foreseeable, the surface term hinders the nucleation mechanism for lower values of the radius, whereas at higher values the nucleation is favourable (i.e. a negative ΔG_{nu} is found). It is possible to identify a critical point, where the interaction between volume and surface contributions lead to the sign change of the 1st derivative of ΔG_{nu}. Such point is referred to as *critical radius* r_c, associated with a *critical energy* $\Delta G(r_c)$*.* These are calculated as:

$$r_c = -\frac{2\sigma}{(\Delta G_v - \varepsilon)} \qquad \text{Eq. 2.39}$$

$$\Delta G(r_c), = \frac{16}{3}\frac{\pi\sigma^3}{(\Delta G_v - \varepsilon)^2} \qquad \text{Eq. 2.40}$$

The r_c value finds the cluster with equal probability to dissolve (return to the *A* phase, or going left on the graph) or to grow (leading towards the *B* phase, or going right on the graph). The critical energy $\Delta G(r_c)$ then corresponds to the energy barrier a system needs to overcome in order to set the nucleation process.

It Is possible to infer the effect of the temperature on $\Delta G(r_c)$, by making explicit its formulation as a function of temperature-dependent parameters. As a first approximation, ε and σ are considered to be temperature-independent. It is then possible to turn the attention to the free energy volume contribution, $\Delta G_v(T)$ [31], [87]:

$$\Delta G_v(T) = \frac{4}{3}\pi r^3 \frac{1}{v_m}\frac{L}{T_c}\Delta T \qquad \text{Eq. 2.41}$$

with v_m being the molar volume, L the molar latent heat of transition, T_c the transition temperature and $\Delta T = (T - T_c)$ (with T being the system temperature). Eq. 2.41 underlines that $\Delta G_v(T)$ is a linear function of the temperature, assuming a zero value at the transitional temperature (i.e with a $\Delta T = 0$).

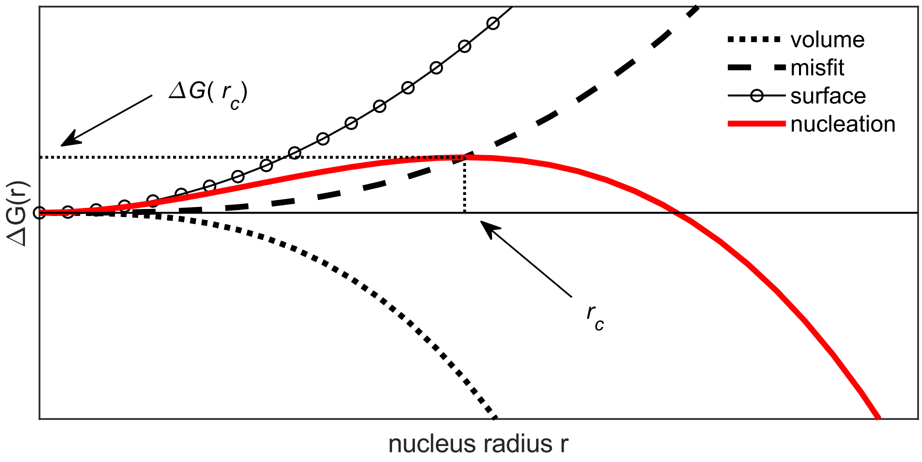


Figure 2.8. Free energy difference plots for a nucleation event, as a function of the spherical cluster radius. The black line depicts the contributions as appearing in Eq. 2.38, whereas the red line depicts the free energy variation during the nucleation process (i.e. increase of the spherical cluster radius), and which identifies a critical radius r_c and critical energy $\Delta G(r_c)$.

2.3.1.2. Heterogeneous nucleation

The process described above is the basis of the homogeneous nucleation, a process taking place within the boundary of an ideal (defect-free) medium with invariant composition across its spatial dimensions. Whilst being devised for liquid-to-solid transitions, it is also applicable to the solid-to-solid transitions with complete miscibility between phases.

An additional process leading to cluster formation is the heterogeneous nucleation, which sees the negative contribution to the energy barrier of a medium defect (e.g. the medium surface, grain boundaries, fractures, dislocations, heteroatoms). This is due to defect relaxation (defects exhibit higher free energy than the ideal crystal), which is then followed by the corresponding Gibbs energy decrease. The process can be seen in Figure 2.9, which is a bidimensional representation of the formation of a *B*-phase cluster at the interface between two *A*-phase grains.

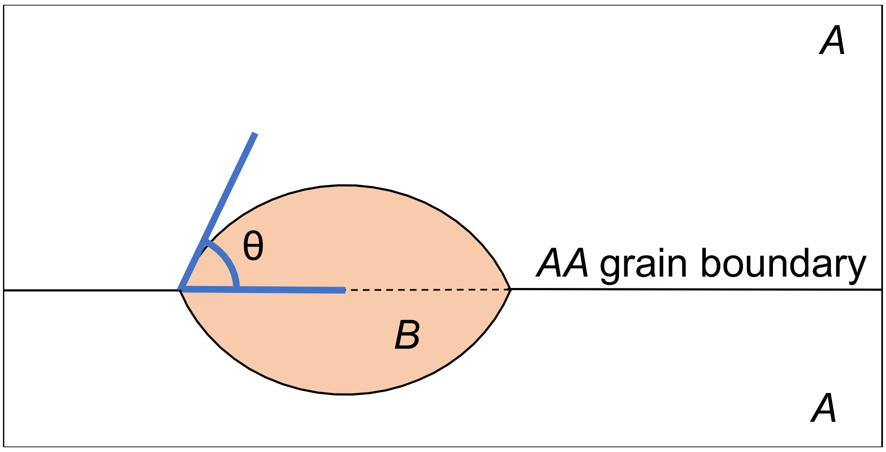


Figure 2.9. Heterogeneous nucleation, taking place on a grain boundary. The angle θ results from the equilibrium condition found in the release of the surface energy σ_{AA} and in the formation of the *AB* surfaces with energy σ_{AB}, thus leading to the formation of a distorted spherical shape (lenticular).

To modify the homogeneous nucleation model to allow for such heterogeneous nucleation, the following contributions need to be considered:

- Creation of the AB interface area, $\alpha_{AB} = 4\pi r^2(1 - co\,s\,\theta)$
- Loss of the AA interface area, $\alpha_{AA} = \pi r^2 \ sin^2 \ \theta$
- Cluster volume, $V = \pi r^3 \ (2 + co\,s\,\theta)(1 - co\,s\,\theta)^2 = 2\pi r^3 S(\theta)$
- with the shape factor, $S(\theta) = \frac{1}{2} \ (2 + co\,s\,\theta)(1 - co\,s\,\theta)^2$

The shape factor also represents the ratio between the volume of the lenticular cluster over the volume of the spherical cluster with equal curvature radius. The so-called *wetting angle* θ results from the balance of the involved surface energies [31], [91]:

$$\cos\theta = \frac{\sigma_{AA}}{2\sigma_{AB}} \qquad \text{Eq. 2.42}$$

Therefore, combining Eq. 2.38 with the aforementioned contributions, it is possible to rewrite the variation of the Gibbs free energy for the heterogeneous mechanism of cluster formation [31], [91], as:

$$\begin{aligned} \Delta G_{nu}^{hetero}(r) = {} & \Delta G_v(r) + \Delta G_v^{misfit}(r) + \Delta G_{s,AA}(r) + \Delta G_{s,AB}(r) \\ = {} & \Delta G_{nu}^{homo}(r) \ \cdot S(\theta) \qquad \text{Eq. 2.43} \end{aligned}$$

It is interesting to note how the shape factor has a scalar value spanning from 0 to 2, and thus it can either promote or hinder the heterogeneous nucleation mechanism, depending on the value of the involved surface energy (see Eq.

2.42). More importantly, the heterogeneous mechanism does not influence the critical radius, but rescales the value of the critical energy, as visible in Figure 2.10.

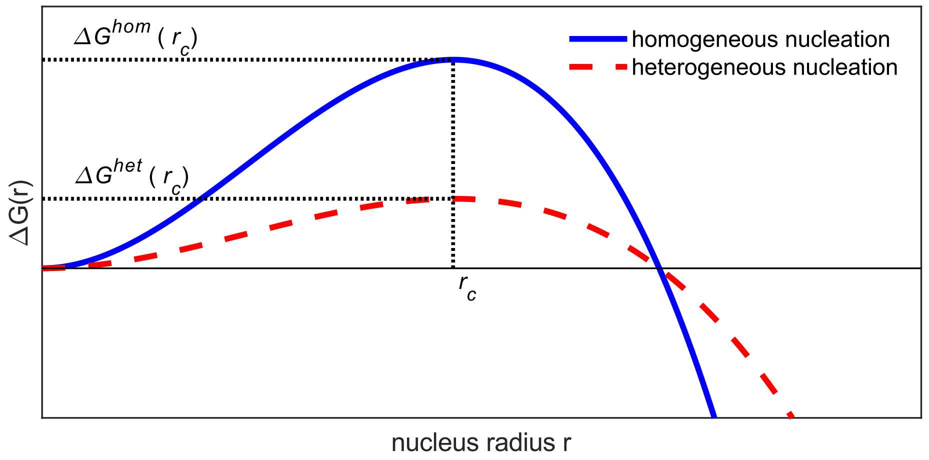


Figure 2.10. Comparison between the Gibbs energy variation resulting from the homogeneous and heterogeneous nucleation mechanisms (with a favourable surface energy balance, i.e. sub-unit shape factor). It is visible how the energy barrier is rescaled in the heterogeneous mechanism, but the related critical radius value does not vary.

2.3.1.3. Growth mechanism

Whilst the nucleation process explains the formation of stable clusters, their increase of volume (i.e. growth) follows different mechanisms. There are principally two different cases: growth at the crystal-liquid interface, and growth by local change of composition. Once more, the present section would consider a system in equilibrium, i.e. an $A \rightarrow B$ transition with $G_A = G_B$.

During a phase transition, each cluster C_n (with n being the number of constitutive elements, i.e. atoms) aggregates a time-variable number of atoms, which then grows and decreases in size following a series of additive and subtractive reactions [90], [92]:

$$C_{n-1} + a \underset{k_n^-}{\overset{k_{n-1}^+}{\rightleftarrows}} C_n \qquad \text{Eq. 2.44}$$

with each chemical reaction characterized by a (size dependent) kinetic coefficient. Approximating this expression, by dropping the size dependency of the kinetic coefficients, it is possible to derive a simplified model for crystal growth. The *net kinetic coefficient* k_0 which models the net transfer of constituent units (atoms) across the two phases boundary, then takes the following form:

$$k_0 = k^+ - k^- \qquad \text{Eq. 2.45}$$

It is also useful to report that the kinetic coefficients, for the reaction reported in Eq. 2.44, can be expressed via statistical considerations [87], [90] as an Arrhenius term:

$$k^{\pm} = k_0^{\pm} e^{-\frac{\Delta G^{\pm}}{k_B T}} \qquad \text{Eq. 2.46}$$

which relates the kinetic coefficient with the Gibbs energy. Growth at the crystal-liquid interface, referred to as *interface-controlled growth*, is described with a velocity v at which the crystal-liquid interface moves, dependent on two configurational parameters and on the kinetic coefficients [90]:

$$v \propto \lambda_D \cdot \left(k_0^+ e^{-\frac{\Delta G_n^+}{k_B T}} - k_0^- e^{-\frac{\Delta G_{n-1}^-}{k_B T}} \right) \qquad \text{Eq. 2.47}$$

Here, growth is the process by which the components in the liquid phase bond with the crystal sites interface, crossing the average interatomic distance λ_D. At equilibrium, the growth and dissolution kinetic factors have equal magnitude: $k^+ = k^- = k_0$. Also, the Gibbs free energy difference after chemical transformation can be expressed in terms of sum of the contribution from the cluster and the adjoining atom, by Hess's law. Thus, it is possible to reformulate Eq. 2.47 as [90]:

$$v \propto \lambda_D \cdot k_0 \left(e^{-\frac{\Delta G_n^+}{k_B T}} - e^{-\frac{\Delta G_n^+ + \Delta G_{atom}^-}{k_B T}} \right) = \lambda_D \cdot k_0 \left(1 - e^{-\frac{\Delta G_{atom}^-}{k_B T}} \right) \qquad \text{Eq. 2.48}$$

This description may be insufficient though to model PC materials growth, as these instead are often successfully modelled by use of the second of the mentioned growth mechanisms: the diffusion-limited growth[14]. This mechanism arises with the diffusion of species being the limiting factor on the growth velocity, a condition previously unaccounted. Stokes-Einstein equation provides a statistical definition of the diffusivity D, by mean of the viscosity η [93]:

---

[14] PC materials fast phase-switch dynamics is also explained in terms of bonds reorientation, in place of the components diffusion. The microscopic picture of the phase change in PC materials is an ongoing matter of research.

$$D = \frac{k_B T}{3\pi\lambda_D\eta} \qquad \text{Eq. 2.49}$$

Another key element by relating diffusivity and the kinetic coefficient [93]:

$$k_0 = \frac{6D}{\lambda^2} \qquad \text{Eq. 2.50}$$

Substituting Eq. 2.50 and Eq. 2.49 into Eq. 2.48, the growth velocity for the diffusion-limited process is then obtained [90], [93]:

$$v \propto \frac{2k_B T}{\eta\pi\lambda^2}\left(1 - e^{-\frac{\Delta G_{atom}^{-}}{k_B T}}\right) \qquad \text{Eq. 2.51}$$

which, similarly to Eq. 2.48, hints to an increasing crystal growth as temperature raises, but with the additional influence of the viscosity term.

2.3.1.4. Viscosity in phase-change materials

Interestingly, the dynamics of viscosity alone is a parameter of high importance to fully understand and characterize the glass-forming alloys, not only for crystal nucleation (the viscosity enters in nucleation rate formulations not here reported [93]–[95]) and growth, but also determines the glass-formation behaviour. Particularly in several PC materials, it is possible to find an interesting dependency of the viscosity η from the temperature [95]–[97].

As seen in Eq. 2.49 and Eq. 2.50, the kinetics of the reaction is tightly bound with viscosity. Solid and stable amorphous materials are characterized by high viscosity, which prevents the fast crystallization progress, hence the stability at room temperature. Heated sufficiently though, amorphous materials recrystallize following the processes described in the previous sections, due to the viscosity dependence on temperature. An empiric parameter, the glass temperature T_g [96], is defined as the temperature at which $\eta = 10^{12}\ [\mathrm{Pa \cdot s}]$, although its definition might be sometime differing as it is also dependent on the thermal history [98]. Such temperature generally corresponds to the onset of the crystallization process. At temperatures lower than T_g, viscosity can be expressed as an Arrhenius function of temperature as follows [95], [98]:

$$\eta(T \leq T_g) = k_\eta e^{\frac{E_a}{k_B T}} \qquad \text{Eq. 2.52}$$

At high temperatures, the MYEGA model (Mauro-Yue-Ellison-Gupta-Allan) is instead more suitable to reproduce the experimental estimation of the viscosity, as it relies on the *fragility* parameter m which quantifies the deviation from the Arrhenius behaviour [95], [99]:

$$\begin{aligned} \log_{10} \eta(T > T_g) \\ &= \log_{10} \eta_\infty + (12 - \log_{10} \eta_\infty) \\ &\cdot \frac{T_g}{T} \exp\left(\left(\frac{m}{12 - \log_{10} \eta_\infty} - 1\right)\left(\frac{T_g}{T} - 1\right)\right) \end{aligned} \qquad \text{Eq. 2.53}$$

The plot of the viscosity as a function of temperature is reported in Figure 2.11.

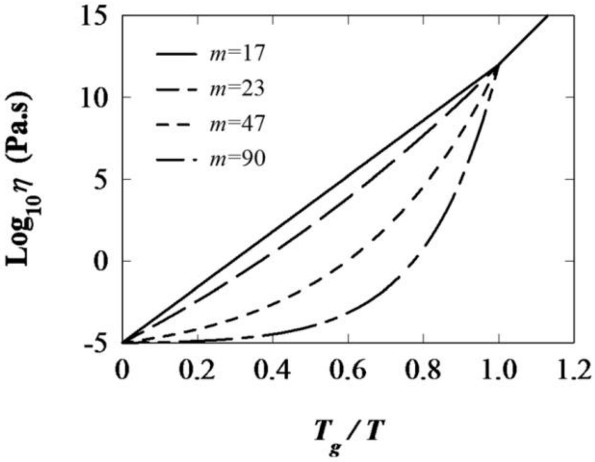


Figure 2.11. Angell plot of the viscosity η as a function of the temperature T, for four different values of the fragility parameter m. Reprinted from [95].

Amorphous materials account for a missing long-range order of its components, which is not the expected arrangement for the lowest energy state (i.e. the amorphous phase chemical potential is higher than the crystal phase chemical potential of the same material). It follows that the amorphous phase is *metastable*, meaning that the glassy phase is not thermodynamically stable, but kinetically prevented to move towards the lowest enthalpy state (i.e. crystallize). This hindrance can be regarded as the sudden decrease in viscosity experienced when cooling the system below glass temperature. Remembering also the kinetic description of the growth velocity (see Eq. 2.52), it follows that to obtain amorphization, a viable method consists of the rapid cooling of the melt phase, which does not leave sufficient time for the crystal front to expand.

2.3.2. A short overview of chalcogenides phase-change alloys

In chalcogenides, PC behaviour arises from the capability of such alloys, which find their room-temperature lowest enthalpy state in a crystalline phase, to form a metastable glassy (or amorphous) phase when subjected to quick temperature decrease (e.g. for GST, with temperature gradients of > 10^9 K/s [31]). Non-thermal phase switching has also been reported [100]–[102], although its mechanism is still a matter of debate. Whilst the crystal phase exhibits metallic or semi-metal properties, the amorphous phase usually exhibits insulator-semiconductor properties, so the phase transition is followed by an evident change in optical properties as well as a variation of the band diagram. Other than the change of the static conductivity, another major effect of phase switching is observed on the shift of the Lorentzian resonant frequency, usually referred to as change of the band gap (i.e. energy associated to the electronic transition from the valence band to the conduction band). Such variation, as outlined in Section 2.1.2, has a direct effect on both the real and imaginary component of the refractive index, thus the phase switching phenomenon in these materials have a determining effect on the resulting light-matter interaction.

Before the work of Ovshinsky, mention of the peculiar behavioural characteristic ascribed to PC materials can be dated back to 1917 [103], [104], from the research of Waterman on molybdenite (MoS_2). Waterman reported the molybdenite non-ohmic conduction behaviour, and later correlated such phenomenon to the material temperature, but his findings were not then followed by significant progress. In 1962 [105], the non-ohmic, reversible, bi-stable conduction mechanism is also reported for the AsTeI and AsTeBr glasses, but again with no attempt to explain the nature of such a mechanism. Ovshinsky started his investigation on PC material in the same period, and saw in his 1968 publication [34] the detailing of both the reversible switching and threshold conduction of the $Si_{1.2}GeAs_3Te_{4.8}$ alloy, as well as a model proposal based on the nucleation theory to explain the empirical findings.

In the 1960 to 1980 period, many other chalcogenide compositions were investigated, showing the characteristic phase-dependent optical properties. For example, GeTe investigations started in the 1960's, and analysis of its divergent material properties depending on the material phase is already reported in 1970 [106]. In the same period, various antimonides, selenides and tellurides were the subject of interest for their phase-dependent optical properties and/or mechanism of reversible switching, such as Sb_2Se_3 [107], [108], Sb_2S_3 [109], Sb_2Te_3 [110], GeSe [111], GaSe [112], [113], AsSe [114], InSb [115], [116] and GaSb [115], [116] to cite only a few.

However, after Ovshinsky's work, the following major development on the implementation of PC material is found in the investigations carried out in the 1980s decade, with regards to the rewritable optical disc mentioned in Section 1.3. The most promising results, merited by the high speed of the phase transition, was initially obtained with the GeTe alloy, which exhibits a stable glassy state below 440 K [106]. In 1986 its adoption as the rewritable optical disc reflective layer was demonstrated [117], with phase switching obtained via laser pulsing in timescales between 30-100 ns. Improvement for optical storage was found in the InSe alloy [118], which allowed direct optical overwrite in place of using an erasure procedure, thus further simplifying the device operativity. Yamada et al. [39] then reported for the first time the use of $Ge_2Sb_2Te_5$, obtained by co-sputtering of the GeTe and Sb_2Te_3 compositions.

$Ge_2Sb_2Te_5$ provided a rather suitable platform for tuning PC material properties. Importantly, the 225 composition is not the single stoichiometry available to pursue PC applications. Research reveals that for this ternary alloy, it is possible to map several of the phenomenological and intrinsic properties as a function of the stoichiometry, such as static electrical conductivity, optical properties, crystallization temperature, melting point, and various others. An example, reported from [119], is here illustrated in Figure 2.12 where the (complex) refractive index variation $\Delta n + i\Delta\kappa$ (at 405 nm) is mapped, for a few selected compositions.

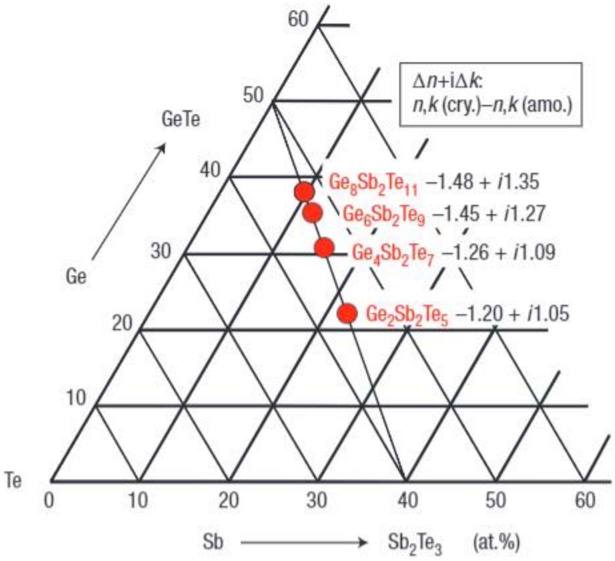


Figure 2.12. $Ge_xSb_yTe_z$ stoichiometry diagram, reporting some of the available compositions and their variation of the complex index at the wavelength $\lambda = 405$ nm. Reprinted from [119]

Additionally, extensive investigations of the doping PC materials, with the aim of tuning thermal, electrical, or optical properties, or to vary the PC material transition temperature and switching kinetics have been carried out. A non-exhaustive list of dopants is here provided (ordered by date of reference publication): nitrogen [120]–[124], oxygen [121], [122], fluorine [122], tin [125], bismuth [125], silicon [126], molybdenum [127], titanium [128], carbon [129], aluminium [130], zinc [131], silver [132], tungsten [133], indium [134], copper [135], chromium [136], selenium [137], scandium [138], arsenic [139].

More recently, other compositions have been further investigated, targeting lossless or low-loss tunable optical properties. Indeed, such compositions would find a wide range of applications, in particular the ones where lossless phase modulation is beneficial (e.g. beam steering, phase shifters, etc.), or in case of single-photon manipulation such as in integrated quantum photonics. One of these PC alloys is the heavily selenium-substituted GST alloy, specifically the $Ge_2Sb_2Se_xTe_{5-x}$ (GSST) [140]. Other examples are the GaLaS (GLS) alloys [141], as well as the previously mentioned Sb_2S_3 [142], [143], and Sb_2Se_3 [143].

2.3.3. $Ge_2Sb_2Te_5$ phase-change material

The $Ge_2Sb_2Te_5$ composition (hereby simply identified as GST) is indeed the archetypal PC alloy. Investigated for optical storage during the 1980s decade [39], it accounts for a series of beneficial characteristics for memory purposes. A

GST *time-temperature transformation* (T-T-T) diagram (shown in Figure 2.13) outlines the coexistence of 4 phases: *amorphous*, *crystalline face-centred cubic*, *crystalline hexagonal*, and *liquid*, of which the amorphous and cubic phases are metastable [31], with a retention of the amorphous phase at room temperature which is indicated in the decades (also depicted in the T-T-T diagram). The melting point, above which only the liquid phase is found, is generally reported to be between 873 K [144] and 930 K [145]. Below such temperature, GST can be found in any mixture of the three solid-state phases, although generally one or two are found. The cubic crystal phase is generally obtained in laboratory conditions upon heating of the amorphous phase above 450 K - 500 K for a few minutes; as for the hexagonal phase, this is obtained upon recrystallization above 500 K - 600 K for few minutes [146]–[148]. From experimental characterization [96], and ab-initio calculation [149], a peak growth velocity of the crystal phase is found in the m/s range. Given the typically adopted PC material volumes (i.e. with features size in the nm to μm scale) such growth velocity translates in a minimum crystallization time between the ns scale and the μs scale.

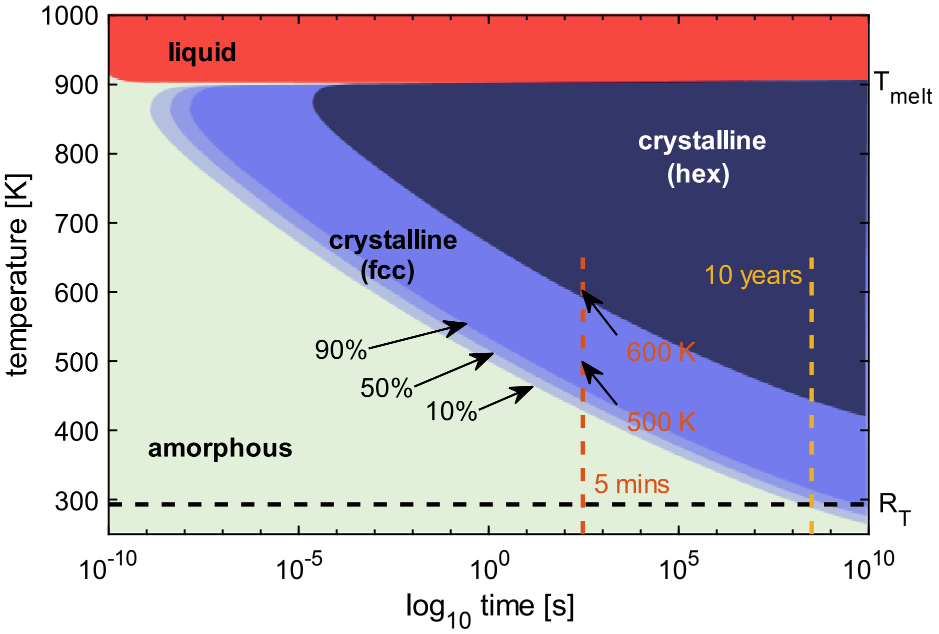


Figure 2.13. T-T-T diagram for $Ge_2Sb_2Te_5$ (adapted from [150]). 4 GST phases (amorphous, crystalline fcc, crystalline hexagonal and liquid) are here depicted. A 5 minutes vertical line and a 10 years vertical line are also reported, to highlight the annealing process conventionally carried out in lab conditions and the amorphous GST long-time phase retention.

In device applications, switching from amorphous, or as-deposited phase, is also obtained via optical or electrical stimuli, instead of direct provision of thermal energy exploiting either Joule heating or optical dissipation mechanisms. In general terms, crystallization is carried out by delivery of a relatively low power, long-duration stimulus (i.e. from hundreds of ns to tens of µs). The principle behind this method is to allow the temperature increase necessary to trigger both nucleation and growth, particularly evident in the 600 K to 800 K range (i.e. above an empirically determined *crystallization temperature* T_c). Avoidance of increasing the temperature above melting point is also necessary. Thus, optimization of the pulsing scheme is mandatory for each device. Switching from crystal to amorphous phase is instead carried out through delivery of a relatively high energy, low width, pulse. A rapid increase of the temperature locally melts GST. At the stimulus withdrawal, since heat is confined within a low effective volume (as a result of the short pulse duration), a high temperature gradient is expected, which quickly dissipates the locally stored heat, avoiding crystallisation. A schematic of the methodology is reported in Figure 2.14.

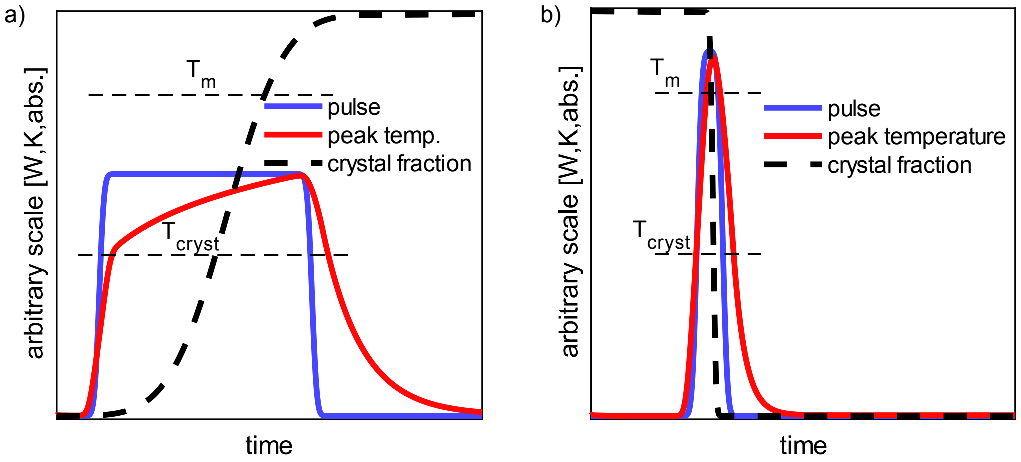


Figure 2.14. Amorphization and recrystallization processes schematic. a) Recrystallization from the amorphous phase, via a long-duration, relatively low-power, pulse. The temperature increase is steadily increasing throughout the pulse duration, allowing to reach and surpass T_{cryst} but remaining below T_m, thus triggering the crystallization process. b) Melt-quench amorphization process. The relatively high power pulse, of short duration, allows increasing the PC material temperature above melting point. At the pulse withdrawal, the temperature rapidly decays, quenching the molten phase into the amorphous matrix.

Whilst some differences in the thermal and optical properties between cubic and hexagonal GST are reported (see e.g. [151], [152]), in device-oriented literature only cubic GST is generally addressed, for the increased energy and time requirements to obtain the hexagonal phase (see Figure 2.13). Therefore, in this work, *crystalline* GST will henceforth refer to the cubic phase, unless explicitly stated.

Amorphous and crystal phases possess remarkably different optical properties. Below, the refractive index n and extinction coefficient κ are reported, from characterisations carried out in our laboratories or reported in the literature. Notably, GST optical properties, as well as many compositions from the $Ge_xSb_yTe_z$ family, are quite susceptible to the subtle variation of fabrication techniques and conditions which affect the material stoichiometry, as well as material thickness, and surface properties. Figure 2.15 shows a few optical property datasets of GST taken from literature, to provide an intuitive picture of the dataset variation.

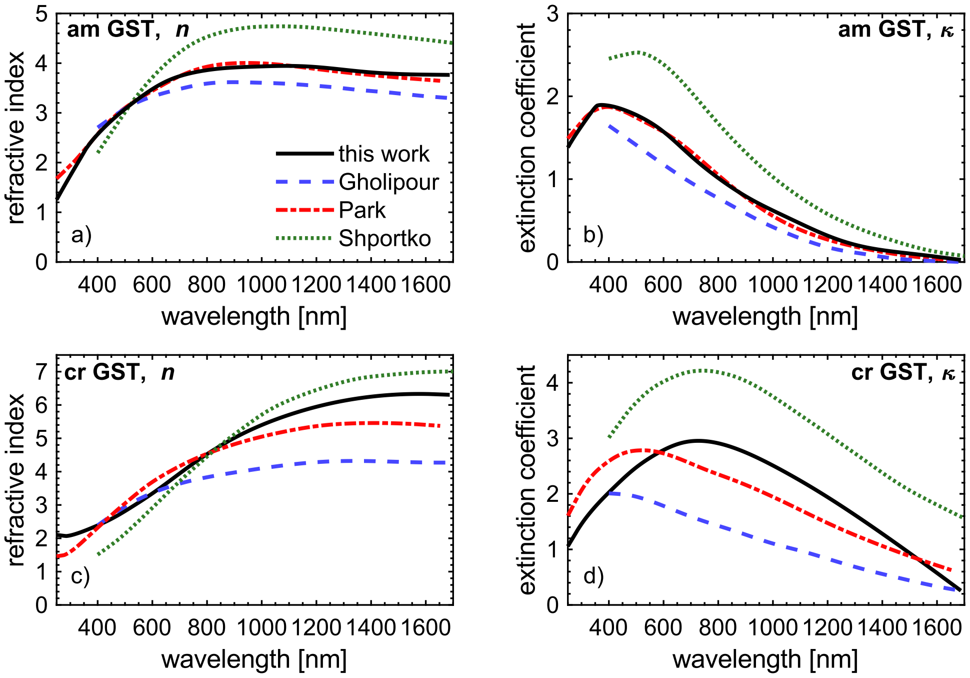


Figure 2.15. GST n and κ optical properties (a, c: refractive index; b, d: extinction coefficient), for amorphous (a,b) and crystalline (c,d) phases. To provide an intuitive picture of the variability of the GST optical properties depending on the composition, a few other datasets have been reported from the works of Gholipour [153], Park [154] and Shportko [155].

Whilst the optical properties of a phase change material are easily accessible via well-established characterization methods, and most are reported in literature for the amorphous and crystal phase, the optical properties of a partially crystallized medium need a separate analysis. Specifically, the *effective medium approximation* (or *effective medium theory*) provides a mean to describe the macroscopically emerging behaviour of a mixture of multiple phases or compositions. With respect to the dielectric function, one finds that, historically, two of this theories gained a broad agreement, each with its own boundaries of applicabilities, advantages and disadvantages: the Maxwell Garnett approximation [156], and the Bruggeman model [157]. Several other works [158] also modify those theories to develop more advanced models, which however encompass a narrower range of applicability. For the scope of this work, we employ the Bruggeman model, as (in contrast with the limitations of the Maxwell Garnett model [158]) it is applicable for crystal inclusions in an amorphous matrix on the whole crystal fraction range. The effective medium electric permittivity, ε_{eff}, is thus calculated by the following equations [158], starting from the crystal fraction χ, and the amorphous and crystal phases electric permittivity (ε_{am} and ε_{cr} respectively):

$$\varepsilon(\chi, \varepsilon_{am}, \varepsilon_{cr}) = \frac{1}{4}\left(\xi + \sqrt{\xi^2 + 8\varepsilon_{am}\varepsilon_{cr}}\right) \qquad \text{Eq. 2.54a}$$

$$\xi(\chi, \varepsilon_{am}, \varepsilon_{cr}) = (2 - 3\chi)\varepsilon_{am} + (3\chi - 1)\varepsilon_{cr} \qquad \text{Eq. 2.54b}$$

The resulting optical properties can be viewed in Figure 2.16, where the evolution of both n and κ can be followed as a function of the crystal fraction. This plot points out the wide potential in tunability GST possesses, particularly in (but not limited to) the near-IR range.

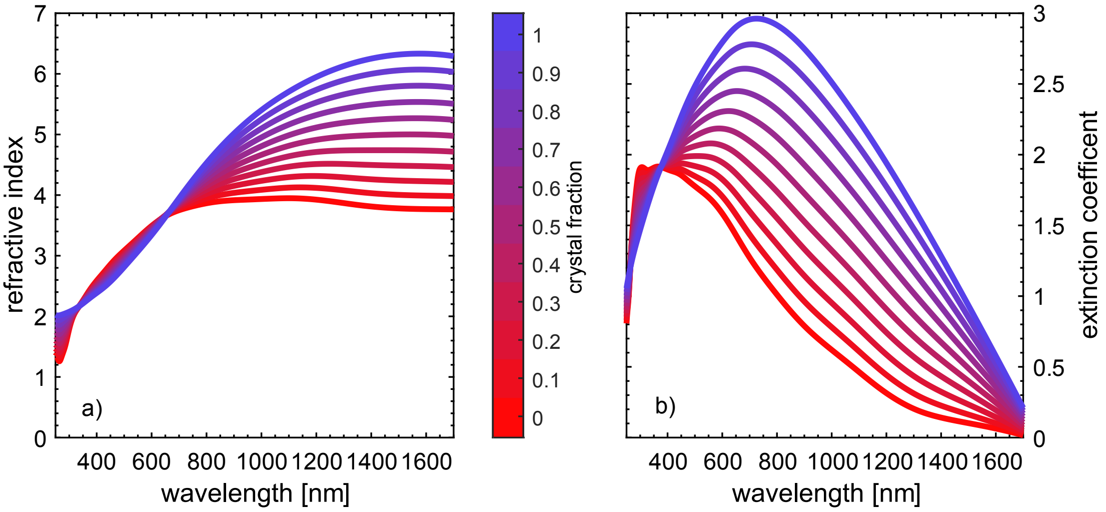


Figure 2.16. GST optical properties as a function of the crystal fraction. a) Refractive index; b) extinction coefficient. The crystal fraction χ value is color-coded for each data plot (colour bar at the centre). The calculation has been performed via Eq. 2.54, from the experimentally determined data reported in Figure 2.15 (black line).

Molten (or liquid) GST optical properties are not clearly reported in literature. In previous numerical works [159], molten GST was assumed to exhibit amorphous phase optical properties, due to the shared disordered atomic arrangement. However, the increase of both thermal and static electrical conductivity of molten

GST [160]–[162] hint to an enhancement of its metallic behaviour. Therefore, in this work, molten GST is assumed to exhibit crystal-like optical properties[15].

Temperature does not play a role only in the determination of the phase (through phase switching processes), Several works [67], [163]–[165] point to a strong dependency of the optical properties on the temperature, independently from the crystal fraction variation. This behaviour is referred to as *thermo-optical effect*, which can be viewed as the influence on the refractive index from the temperature-induced lattice distortion, variation of the electronic population of the density of states, and of the band gap. Here, the following complex index values for GST are adopted[16] from [163] (in *refractive index units*, $\mathrm{R.I.U.}$):

$$n_{cr} = 6.11[RIU] - 2.2 \cdot 10^{-4} \left[\frac{\mathrm{R.I.U.}}{°C}\right] * (T[°C]) \qquad \text{Eq. 2.55a}$$

$$\kappa_{cr} = 0.83[RIU] + 1.56 \cdot 10^{-3} \left[\frac{\mathrm{R.I.U.}}{°C}\right] * (T[°C]) \qquad \text{Eq. 2.55b}$$

$$n_{am} = 3.94[RIU] + 1.1 \cdot 10^{-3} \left[\frac{\mathrm{R.I.U.}}{°C}\right] * (T[°C]) \qquad \text{Eq. 2.55c}$$

$$\kappa_{am} = 0.045[RIU] + 4.1 \cdot 10^{-4} \left[\frac{\mathrm{R.I.U.}}{°C}\right] * (T[°C]) \qquad \text{Eq. 2.55d}$$

Thus, for each wavelength the complex index of GST variates with both temperature and crystal fraction. A visual representation of the refractive index variation as a function of crystal fraction variation is reported in Figure 2.17, arising from the combination of Eq. 2.54 and Eq. 2.55, and including the assumption of crystalline phase-like behaviour of the liquid phase.

An additional factor which is experimentally confirmed to affect the optical properties is the radiation intensity. *Non-linear optical* (NLO) phenomena can be described in terms of the electric susceptibility (see Eq. 2.15), which can be redefined via Taylor series expansion of the electric field model (i.e. $\boldsymbol{P} = \sum_i \chi_i \boldsymbol{E}^i$). The characterization of the crystalline GST nonlinear optical properties has been carried out in the visible spectral range by Liu, Wei and Gan [166], but no data in the NIR range, nor specifically for the amorphous phase, is available; therefore, such effects are not here integrated in the proposed GST refractive index.

---

[15] This approximation can be further justified in view of the optical response of the phase-change photonic memory, whose transmission does not show discontinuity after melting (see Section 6.5.1 for further details).
[16] The investigation Stegmaier et al. [163] has been carried out from room temperature to 350 K. The data here presented linearly extrapolate the found trend, aware that further experimental confirmation is needed.

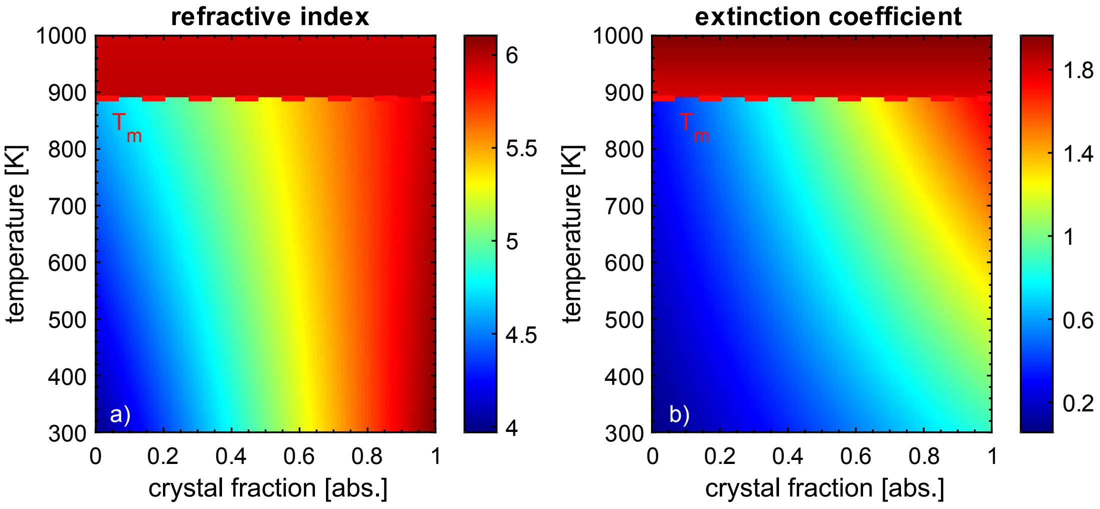
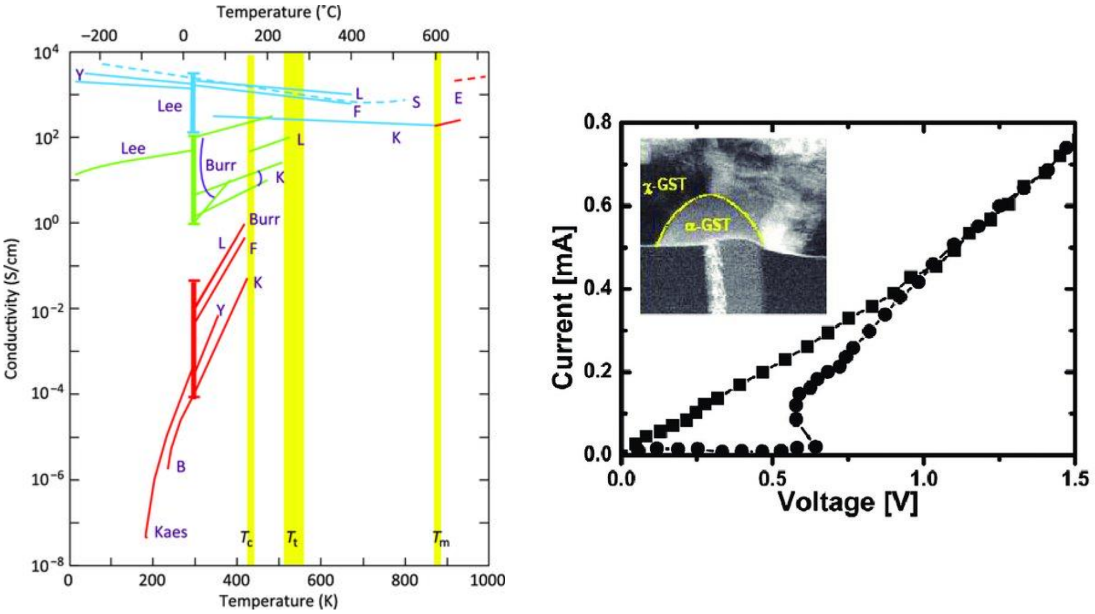


Figure 2.17. 2D plot of temperature- and crystal fraction-dependent GST optical properties at 1550 nm. Data is derived from Eq. 2.54 and Eq. 2.55, assuming crystal-like optical properties for the liquid phase. a) Refractive index, n (ranging from 3.97 to 6.11). b) Extinction coefficient, κ (ranging from 0.045 to 1.96).

Static electrical conductivity is also a rather important characteristic of GST, as its phase-dependent conduction determines the suitability for memory applications. GST exhibits a rather large contrast in conductivity after switching from amorphous to crystal phase, with a reported conductivity delta between 1.5 to 6 orders of magnitudes, depending on the references (see Figure 2.18(a)). Crystal GST conductivity σ_{cr} experimental data can be conveniently modelled by use of a simple Arrhenius equation [167], which hints to a thermally-activated mechanism of carrier generation:

$$\sigma_{cr}(T) = \sigma_{cr,0} \cdot e^{-\frac{E_{a,cr}}{k_B T}} \qquad \text{Eq. 2.56}$$

The case of the amorphous phase, and in general, of amorphous chalcogenides' conductivity, has been a matter long debated since its early characterization [34]. Figure 2.18(b) reports the current-voltage diagram for amorphous GST when a ramped voltage is applied. At low biases, it conforms to an ohmic behaviour with low conductivity (~1 S/m), whereas at higher bias (above a threshold voltage V_{th}), the conductivity rapidly increases by various orders of magnitude. Experimental data points to the Poole-Frenkel conduction mechanism, which explains conductivity of insulators as the combined effect of temperature and external field on the promotion to the conduction band of electrons which would otherwise be bound to localized states, and to which they return by electron-hole recombination following a temperature- and electric field-dependent rate. Sub-threshold conductivity has been successfully modelled [168] by use of a Poole model, whereas at higher values of the electric field a Poole-Frenkel mechanism is indeed reported [168]. Nevertheless, research is still active in this fundamental-oriented area, consequently to the variability of the amorphous phase experimental conductivity (see Figure 2.18(a)).



a) b)

Figure 2.18. Static electrical conductivity of GST. a) Dependency of the conductivity on temperature and phase, from various references (reprinted with permission from [169]). Red lines correspond to the amorphous and liquid phase; green lines to the fcc phase; blue lines to the hexagonal phase. b) Experimental IV plot of amorphous GST in a mushroom cell (reprinted with permission from [170]). Circles represent the data collected during the current increase; squares represent the data collected during the current decrease.

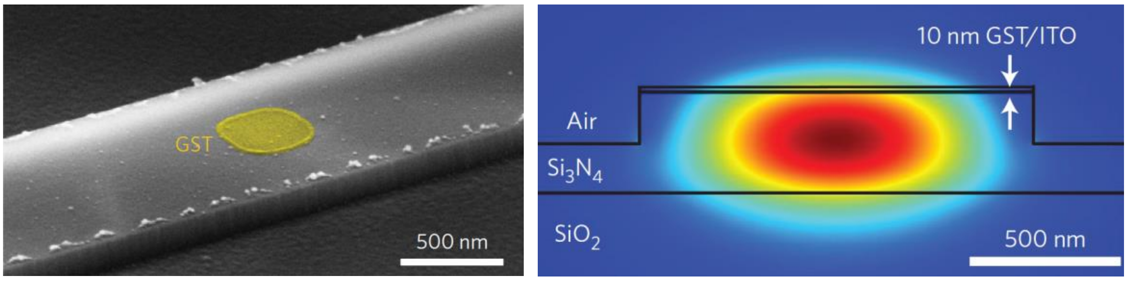
2.4. Integrated all-photonic memory

As briefly reported in Chapter 1, one of the critical elements of the development for optical and all-optical computing devices is the information storage, with particular accent on RAM-class or cache-class. Whilst PC materials have been used for decades for optical storage in optical disk formats, their use with more modern integrated photonic devices has, until very recently, received much less attention. A particular advantage of PC based integrated photonic memories is that storage in them is truly non-volatile, with states able to persist on very long timescales (years). This is in contrast to most other approaches to integrated optical memories, where only volatile storage, or non-volatile storage for durations in the microsecond range, has generally been achieved [171].

As presented in Section 2.3.3, the PC material GST possesses the following characteristics: i) retention of the phase structure for long time duration in absence of energy supply, ii) phase transition potential characteristic time and energy in the orders of ns – pJ respectively, iii) potential of all-optical switch through optical absorption, iv) phase-dependent complex index. Therefore, the GST platform allows one to encode the information as crystal fraction, assuming the amorphous and crystal phases to be promptly accessible by shaping its thermal history. Indeed, PC materials are an established solution for information storage in a non-volatile fashion. In Section 1.3, a few examples were reported utilizing PC material as a platform for electrically-accessible memory cells. The all-photonic phase-change memory brings the same principles in the optical domain, relying upon the PC material property variation after optically-induced phase-switching for non-volatile information storage.

2.4.1. Memory configuration

In 2012, Bhaskaran and Pernice [66] proposed the all-photonic PC memory architecture, consisting of a TE-single mode integrated SiN waveguide, on whose top surface a thin layer of GST is deposited (usually, 10 nm, then capped for protection purposes by ITO[17] or SiO_2), referred to as *optical unit cell* (see Figure 2.19). The unit cell is generally considered to totally cover the waveguide surface, with a uniform thickness, and variable length which determines the total absorption and insertion loss IL. Other works dedicated to the experimental investigation of the unit cell memory's potential and operation have subsequently followed [67], [68], [172].



a) b)

Figure 2.19. Phase-change photonic memory standard configuration (reprinted with permission from [67]). a) False-coloured *scanning electron microscopy* (SEM) image of the unit cell, fabricated on top of a rib waveguide. b) Cross-sectional representation of the device, with a superimposed simulated waveguide electric field TE mode profile.

---

[17] Indium Tin Oxide.

In this architecture (of Fig. 23), the propagating optical mode, as demonstrated in Section 2.1.1, possesses a non-zero component of the e.m. field outside the waveguide boundaries, which then evanescently couples with the optical unit cell on top of the waveguide. Resulting from the field distribution properties of the propagating mode, and the lossy optical characteristics of GST[18], a certain fraction of the optical power is dissipated via absorption and scattering mechanisms. The amount of energy dissipated translates in the modulation of the transmitted power T, which then takes the role of readable value of the stored information. With the unit cell in the amorphous phase, merited by the relatively low extinction coefficient (0.045 at 1550 nm, see Section 2.3.3), T is relatively high (close to unity for a fully amorphous unit cell). Instead, with a crystalline unit cell, the extinction coefficient becomes roughly 20-fold the previous value, resulting in a much higher optical power dissipation. Consequently, the transmission T is lowered by an appreciable quantity, which is proportional to the unit cell size [67].

2.4.2. Information storage and retrieval

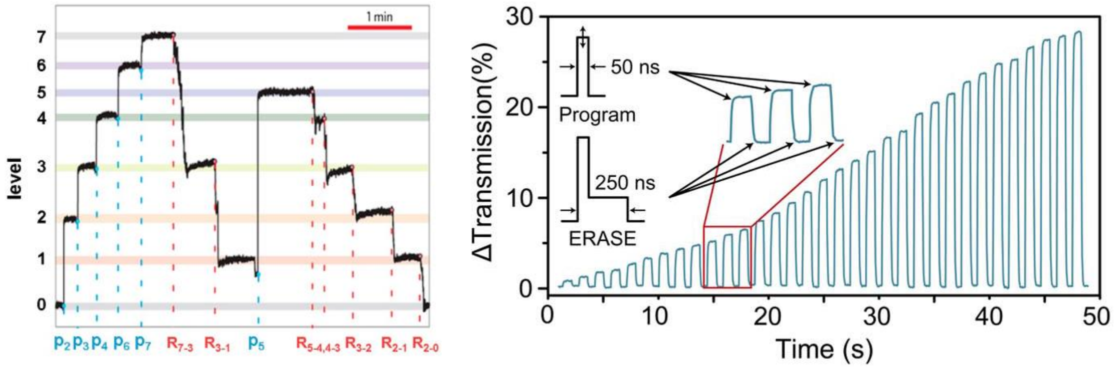
The read operation is performed by delivery of a low-intensity optical signal along the waveguide, which once collected and compared to the input level allows one to determine the transmission T level. A unit cell of length of 250 nm has been proven [67] to be sufficient to determine a measurable modulation of the transmitted power (with an optical contrast $(\mathrm{T}_{max} - \mathrm{T}_{min})/\mathrm{T}_{max}$ of 0.6%). Longer unit cell lengths of course determine a wider maximum optical contrast, e.g. a 28% for the 5 μm cell is reported in the same work.

Information writing and erasing is performed via the same optical path, by delivery of opportunely tailored high-energy optical pulses. The methodology is the same as previously reported and illustrated in Figure 2.14. Amorphization is obtained via a single, high power pulse. Recrystallization via appropriate shaping of the pulse to accrue complete recrystallization aiming for energy-efficiency and reproducibility of results. In Ref. [67], a train of decreasing energy pulses is used for recrystallization, which in place of the single pulse avoids unwanted melting due to the increased absorption during recrystallisation process. The same result, with added flexibility value, is demonstrated via use of a double step pulse [68] (i.e. a high power-short duration pulse, immediately followed by a lower power-longer duration pulse, or “tail”). The additional value lies in the reproducibility of the final crystallization, regardless of the initial level. This is of fundamental importance when aiming for multi-level storage. A third technique relies on pulse-width modulation, reported in [69], which also allows one to address arbitrary levels independently from the unit cell state.

---

[18] The operation has been conceptualized for the TC band, nevertheless other wavelengths can be adopted following a dedicated optimization of the waveguide and unit-cell configuration (see e.g. [338]).

Multi-level storage is attained by reaching partial crystal fraction levels, in contrast to the simpler 2-level mechanism for binary storage. For each crystal fraction level l, a biunivocal correspondence is found with a transmission level T_l, allowing for information retrieval. The density of these levels is then a function of the total contrast between lowest and highest transmission levels, added to the write error which the operations involve (assuming a sufficient resolving power of the detector stage). Single-cell 8-level storage (3 bits) has been initially demonstrated on a 5 μm long unit cell in [67], in which any of the desired levels is reached by a unique pulse sequence (see Figure 2.20(a)). In Li et al. [68], by use of a 4 μm unit cell, and employing the double step pulse technique, access to up to 34 unique levels (>5 bits) has been demonstrated (see Figure 2.20(b)). Additionally, each level is accessible independently from the original level, with the final crystallization degree uniquely determined by the pulse duration. As a demonstration of the scalability of the device concept, Feldmann et al. [173] realized a single memory chip comprising 256 unit cells, addressed by use of a 9 word line – 9 bit line optical waveguide configuration.

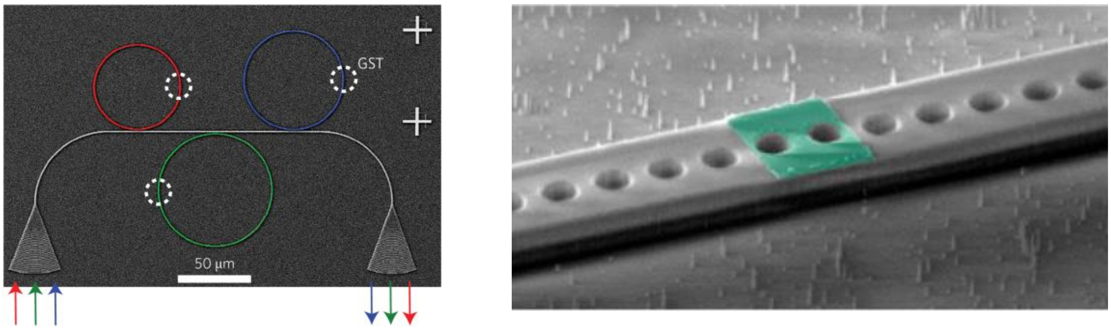


a) b)

Figure 2.20. Multi-level storage demonstration on a PC photonic memory conventional configuration. a) 8-level (3 bit) memory addressing on a 5 μm unit cell, by independently optimized pulsing schemes (reprinted with permission from [67]). b) 34 levels (>5 bit) storage, achieved by double-step programming pulse (shown in inset) on a 4 μm unit cell (reprinted with permission from [68]).

2.4.3. Additional functionalities

PC photonic memory devices can easily be implemented on the top surface of micro-ring resonators. One of the straightforward advantages this configuration brings is the WDM/WDDM standard technology functionality, which allows one to address individually multiple unit cells accessible through a single optical bus (see Figure 2.21(a)). This has been successfully demonstrated in [67], which shows the independent read-write operativity on three unit cells accessed via the same optical path. The slight phase shift induced by the unit cell refractive index variation induced by the PC cell, leads to a shift of the resonant structure coupling strength. In high Q-factor structures, this leads to significant variation of the transmitted power: in fact, when a GST unit cell is embedded in a micro-ring resonator, modulation depth greater than 10 dB is reported [66]. An additional advantage brought by resonant structures is the decrease of the energy requirement for write and erase operations. This is also exploited with a different kind of resonant cavity, a unidimensional photonic crystal (see Figure 2.21(b)), proposed in Von Keitz et al [174], which reports the decrease of the switching energy by around an order of magnitude when compared to previous work on the conventional device.



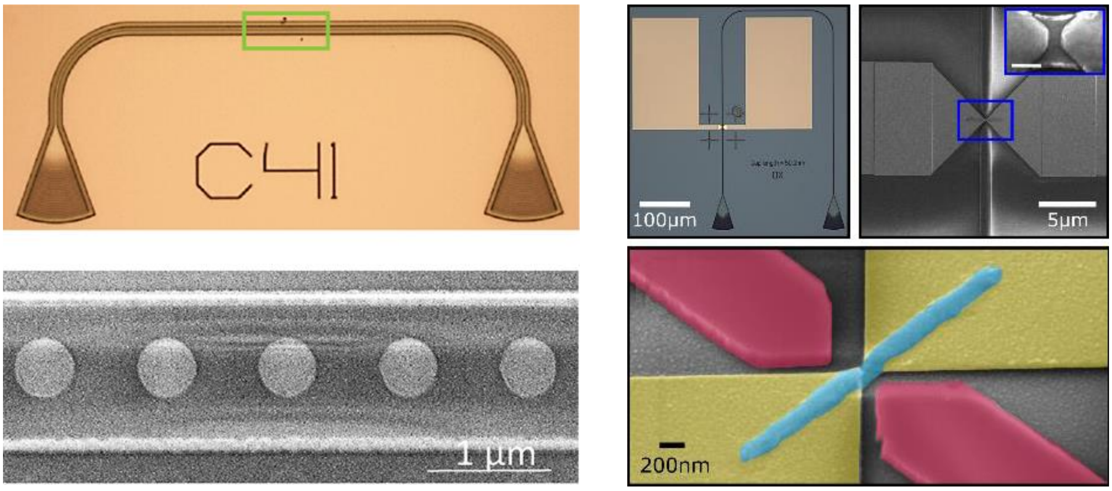
a)

b)

Figure 2.21. False-coloured SEM images of a) micro-ring implemented and b) 1D-photonic cavity implemented variants of the PC photonic memory conventional configuration, aiming for added functionality and increased energy performance. (a) is reprinted with permission from [67]. (b) is reprinted with permission from [174].

One of the challenges faced since the integrated phase-change memory architecture was devised is the full exploitation of the totality of the PC volume, i.e. the complete amorphization of the unit cell and the subsequent reversible return to the fully crystalline state. Access to the fully amorphous state invariably led to device failure, as the deformation of the unit cell or ablation of the layers induced by the high-energy write pulses compromises irreversibly the unit cell optical characteristics [172]. A solution was proposed in Wu et al. [175], where the authors report a unit cell split into sub-wavelength sized sub-units (see Figure 2.22(a)), and encapsulated with a robust Al_2O_3 60 nm layer added to the ITO 10 nm proposed in [67]. Whilst the unit cell splitting has been previously reported in Cheng et al. [72], the combination of unit cell splitting and restriction on the central region of the waveguide, added to a more solid cap layer, enabled all-optical access to the fully amorphous state without damaging the device.

Mixed-mode operativity, i.e. read-write operations arbitrarily performed in both the optical and electrical domain on the same unit-cell, has been pursued by design in Farmakidis et al. [176]. The authors propose the implementation of the unit cell within a nano-sized gap between two Au electrical contacts, which interjects a rib waveguide (see Figure 2.22(b)). In this architecture, switching of the unit cell is influencing both the electrical path resistivity and the waveguide transmissivity, merited by the field enhancement effect due to the proximity of the metallic contacts.



a) b)

Figure 2.22. Additional variants of the PC photonic memory device conventional configuration a) Wu et Al. split unit cell configuration (top: chip-scale microscopy image; bottom: SEM image of the cell sub-units). Reprinted with permission from [175]. b) Mixed-mode device optical and SEM images (reprinted with permission from [176]), which allows both electrical and optical read-write operativity.

2.4.4. Computation with PC photonic memory devices

Developing on the basis of the PC photonic memory, several works [69], [70], [72]–[74] explored the possibility to adopt this platform to perform tasks conventionally addressed by the processing unit.

In Feldmann et al. [70], the authors proposed a crossed waveguide grid configuration, where a unit cell is located at the intersection of each crossing. In such configuration, a below-threshold write signal can be sent through a waveguide, leading to the unit cell temperature increase but without sourcing any crystal fraction variation. If a similar signal is sent through the orthogonal waveguide, and the two pulses are temporally correlated, then the pulses additive effect is sufficient to induce a permanent change of the unit cell crystal fraction. The authors use this configuration and exploit the pulse correlation to perform simple arithmetical calculations, i.e. additions and multiplications, directly on-chip without the need to read the unit cell stored values and perform the calculation externally, i.e. demonstrating in-memory computing on a photonic platform.

Cheng et al. [69] instead used the conventional device configuration to perform logic operations. The authors attain to this functionality by encoding the inputs as pulse widths, which determine an arbitrary variation of the unit cell crystal fraction. Delivery of both signals onto the same unit cell yields a predictable crystal fraction variation. The following comparison of the output signal (i.e. read of the logic operation result) with an appropriate reference, allows performing NAND and OR logic gating.

A more deliberate attempt to perform one of the neuromorphic computing primitives is found in a following work, Cheng et al. [72]. This work reports the combination of the unit-cell split configuration, whose interaction with the propagating mode is further enhanced via waveguide tapering; and a switching pulse scheme reliant on the variation of the number of consecutive delivered pulses, rather than PAM or PWM techniques[19]. The authors demonstrate that merging these elements, the device transmission can be precisely modulated unbound from the original state. The technique was proposed as a solution to realize a photonic synapse since the weight (i.e. transmissivity) could easily be tuned without previous knowledge of the weight, and with the added simplicity of the pulsing scheme.

Combination of the accrued experimental and theoretical knowledge on the behaviour of the unit cell in the straight waveguide, micro-ring resonator, and its crystal fraction variation under optical stimuli (along with other standard photonic platform device technology) led to the realization of the first, non-volatile photonic spiking neural network [73], counting of 4 neurons and 60 synapses. In this device, each neuron soma consists of a single waveguide crossing a micro-ring resonator, upon which a unit cell is fabricated at one of the intersections, enabling neuronal weighting. 15 drop-in micro-ring resonators are collocated along the same waveguide, each one also modulated by use of a unit cell acting as synaptic weighting. Delivery of wavelength-multiplexed synaptic signals through the drop-in ports can modulate the neuronal weighting after sufficient applied power. Neuronal functionality is then obtained probing a waveguide coupled with the soma micro-ring, which exhibits a large modulation of the transmission after the neuronal weight variation (due to the temporal correlation of the stimuli). The device has been demonstrated capability for both supervised and unsupervised learning, in particular applied for pattern recognition.

---

[19] Pulse amplitude modulation, or pulse width modulation.

Work is now in progress to expand this technology application range and test its potential. A recent publication [74] reports the experimental investigation on a photonic *convolutional neural network* (CNN), whose memory states are once more stored in photonic unit cells. The measured test device performance (i.e. multiply-accumulate operations rate) highlighted how this photonic platform can challenge the device electronic counterpart state-of-the-art technology (though in specific applications such as the CNN, which benefits from the massive parallelism which is intrinsic of the optical domain signalling).

Chapter 3. Experimental methodologies and computational techniques

3.1. Experimental techniques for photonic device test and characterization

3.1.1. Integrated photonic chip test and characterization

This subsection briefly describes the experimental setup from which the investigated *all-Photonic Phase-Change Memory* (PPCM) device experimental data can be gathered (see Figure 3.1). As this setup is readily adaptable to other photonic devices, a general perspective is here maintained.

A typical photonic chip consists of a substrate, upon which the photonic devices are fabricated via various methods, such as *Chemical Vapour Deposition* (CVD), evaporation, magnetron sputtering techniques, etching, and lithography techniques such as laser writing or electron-beam lithography. Waveguides are the most basic and yet essential elements on a photonic device, serving as a signal bus to shuttle the propagating mode from point to point on the photonic chip.

If a light source is not embedded within the device (as in the general case of the PPCMs reported in the literature and investigated in this thesis), the platform needs an element which serves to couple an external optical source (generally, via optical fibres) to the waveguide. The most commonly found are *prism coupling*, *butt coupling*, *end-fire coupling,* and *grating coupling* [177]. Regardless of the implementation, all these accomplish the task of coupling the fibre mode (desirably, but not uniquely, a Gaussian profile) to the waveguide mode, thus enabling the signal propagation across the photonic network. The same mechanism is implemented at the output of the photonic chip, where no detector is implemented on-chip.

Thus, the test bench usually includes a *laser source*, delivering the optical signal towards the chip via optical fibres. Generally, a polarization control device is also interjected between light source and coupling device, to maximize the coupling efficiency, and a single-mode optical fibre is used after the polarization control to maintain the selected mode.

An optical chip may require multiple sources, such as in the testing of the PPCM. A *continuous laser* is usually adopted to deliver a constant low-power signal to detect the attenuation or spectral response of the device, without perturbing the system. This optical source will hereby be referred to as *optical probe*, or *probe*. A second source, of different wavelength, is usually used with the PPCM device to trigger the phase-change process. Its temporal profile is shaped via the *Electro-Optic Modulator* (EOM), driven in turn by an electrical pulse generator. In this work, the high-power pulses from the EOM will be generally addressed to as *optical pump*, *pump,* or simply *pulse*. To augment the delivered power to the required magnitude, an *Erbium-Doped Fibre Amplifier* (EDFA) can be used, if the source is not sufficiently intense. Another convenient stage is the *optical bandpass* filter, which reduces the intensity of undesired wavelengths, which may arise from the laser source.

Optical circulators are also placed next to the input and output of the photonic chip. These are multiple-port (3 or more) devices that allow the delivery from one port to the successive, in a circular manner. Optical circulators allow bi-directional transmission on the test device, which simplifies the process of separation and detection of the different probe and pulse wavelengths. Additionally, an optical circulator functions as *optical isolator*, which acts as the optical-domain

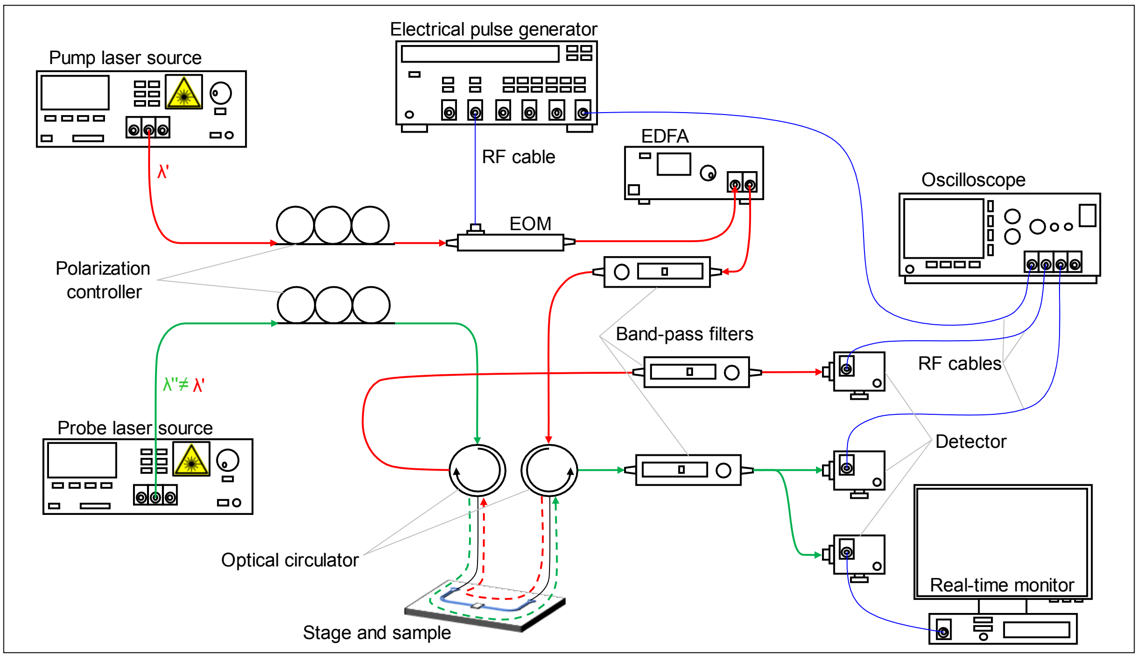


Figure 3.1. Optical bench setup schematics. The probe laser operates in CW mode, and is guided to the device via polarization controller and an optical circulator. The probe output is then collected and delivered to the detector, passing through a second optical circulator and a band-pass filter. The probe is here also used for real-time monitoring, to compensate for misalignment. The pump laser is guided towards a polarization controller first, then modulated via EOM, and amplified through EDFA. The pump is then filtered via band-pass filter, and through the optical circulators is sent towards and collected from the photonic chip. The output is then collected from the oscilloscope. The oscilloscope also drives the electrical pulse generator, which provides the necessary signal shape to the EOM.

counterpart of a diode. Thus, an optical circulator effectively separates the laser source from the test device and avoids any harmful feedback due to reflection.

The optical chip stage can include additional element to ensure the measurement temporal stability, avoid cross-talk between active elements, and maximise the efficiency of the signal injection and collection. The most commonly found are the moving stage (allowing one to control translation and rotation of the tested optical chip), and the thermal stage, a highly-thermal conductive (or active) element which physically supports and controls the temperature of the photonic chip. It is noted though that, for the PPCM device, the thermal control stage is generally not adopted [67], [68], [71], in view to avoid undesired thermal gradients (or variation thereof) during the high-temperature stages of Write and Erase (more details on those operations can be found in Chapter 6).

After collection from the photonic platform, the optical signal is detected and converted to a time-resolved electrical signal for post-processing. A real-time monitor is also beneficial for the adjustment of all-optical elements, in particular for the alignment of the optical fibres on the coupling stage. One important part of post-processing requires the calibration of the signal after the inevitable loss faced at each of the off-chip stages (e.g. optical circulators, optical fibres, eventual splitters, coupler devices, detector wavelength- and power-dependent efficiency, etc.). This then enables the effect of any phase-change cell integrated with the waveguide to be quantitatively evaluated.

3.1.2. Characterization techniques for optical material properties

As underlined in Section 2.1, the photonics field heavily relies on the exploitation or manipulation of the material optical properties. Thus, awareness of the material properties of the investigated or adopted materials is of extreme importance, either for the experimental validation of a theoretically predicted behaviour, or for the confirmation of the assumed optical characteristic leading to the design, simulation, and optimization of an experimental device.

A method allowing to access to the optical properties, in particular of thin films, is ellipsometry, which is nowadays the de facto standard methodology for the investigation of a material layer optical properties, due to this technique being non-destructive and extremely sensitive. Ellipsometry [178], [179] determines the polarisation ratio $\tan \Psi(\lambda)$ and phase $\Delta(\lambda)$ variation of an impinging (coherent and focused) beam of wavelength λ, shone at near-Brewster angle, upon reflection from the investigated layer. Specifically, the fundamental equation of the ellipsometry correlates the ratio between s and r polarization of the incident beam with the parameters Ψ and Δ as follows [179], [180]:

$$\tan \Psi e^{i\Delta} = \mathrm{r_p}/\mathrm{r_s} \qquad \text{Eq. 3.1}$$

As visible from Eq. 3.1, the ratio between the two orthogonal polarization states is a complex number (i.e. contains both phase and amplitude information), and as for such, retains the information over the real and imaginary components of the reflective layer optical properties. The Fresnel coefficients (or, more properly, the ratio thereof) that Eq. 3.1 allows to retrieve can be used to reconstruct the Jones matrix of the investigated layer stack [179]. Therefore, after measurement of the Ψ and Δ angles, and provided the knowledge of the layer thickness, it is possible to fit such wavelength-dependent data, and uniquely relate the dielectric function to the measured quantities.

Such operation though comes at the cost of the fairly involved analysis following the data collection, particularly with regards to the selection most suitable mathematical model retrieving the optical properties information from the measured quantities. The analysis is further involuted due to the additional dependency of Ψ and Δ from many of the layer properties (beside the layer thickness), such as surficial properties, composition, phase, contaminants, conductivity, porosity and density [181]. Therefore, the ellipsometrical determination of the optical properties is, in general, a non-trivial operation, and to a certain degree, its outcome may depend on the user experience and training, as well as on the determination (or inference) of the additional parameters required by the adopted model.

Alternatively to the ellipsometry technique, optical spectroscopy [182] can be adopted to obtain information on the emission, reflection, transmission and absorption of the investigated sample, which once more can be correlated with a unique set of n and κ values via analytical techniques (see e.g. the Swanepoel method [183], [184]). Optical spectroscopy measures the amplitude variation of the reflected (or refracted) optical beam upon interaction with the investigated system. This technique is widely employed to characterize the optical characteristics of the analyte, and as the measured quantity carries only the amplitude information, methodologies adopting optical spectroscopy techniques need to retrieve two sets of wavelength-dependent parameters (or rely on hard assumptions) to correlate the measured data with the analyte optical properties. In this work, a novel technique reliant on spectrophotometry data is illustrated in Chapter 5, which does not involve hard assumption but instead converges towards a consistent match with ellipsometric data via modification of the analyte configuration.

3.2. Finite Element modelling

The design and numerical testing of photonic circuits and nano-photonic components require the solution of an analytical model aimed to mimic the system’s investigated quantities. As an example, the investigation of the transmittance of a simple optical transmission wire (i.e. an integrated waveguide) requires the solution of Maxwell’s equations, along with the solution of the Helmholtz's equation to identify the waveguide propagating mode. In case of the PPCM, it is necessary to concurrently emulate a large number of linked physical phenomena, amongst which the temperature distribution across the device, and phase-state distribution across the unit cell volume, have the largest impact. Therefore, one would be bound to find the time-dependent solution of the heat diffusion equation, as PC materials (in particular GST) exhibit a rather strong thermo-optical effect (as underlined in the previous chapter). One also needs to calculate the time and temperature-dependent phase-transition process of the unit-cell (a partial list of suitable models is found in Section 3.3). Indeed, the system of equations here briefly summarized can’t be trivially solved, not even considering that some of the mentioned components (e.g. the Helmholtz's equation) do not possess an analytical solution.

The *Finite Element Method* (FEM) is a mathematical method purposely devised for the solution of partial differential equation systems expressed in the weak formulation. This technique is widely adopted to obtain the numerical approximation of many challenging non-analytical solutions in the field of engineering, mathematics and physics [185]. Along with other methodologies [186], most prominently the *Finite Difference Time Domain* (FDTD), it is widely employed for the analysis and numerical characterization of photonic circuits component. The basis of the FEM technique relies on the subdivision of the analysed system into *finite elements* (also called *meshes*), which permits one to find the mesh-related numerical approximation of the localized equation system. For a more accurate description of the mathematical principles that FEM techniques (in particular, oriented to the COMSOL environment), the reader is referred to Appendix I.

3.2.1. COMSOL environment

COMSOL Multiphysics is a FEM simulation environment, which allows one to solve conventional physics-based systems of partial differential equations. This software tool has been adopted to carry out the analysis illustrated in this thesis, as it allows one to perform calculations on the coupled e.m. wave field and the temperature fields, via the RF (*radio-frequency*) and Heat modules. It also allows the incorporation of external algorithms, solvers, and data-handling tools, via use of the Livelink module. In particular, in the work of this thesis, the MATLAB Livelink module is extensively used to implement calculations currently unavailable in COMSOL.

A COMSOL model generally consists of a finite *geometry*, which depicts a physical device subject. The software environment allows the definition of several *domains*, which describe continuous entities of the geometry (i.e. the air domain, the substrate domain, the device domain, etc.). As well as the 3D domains, 2D *boundaries*, 1D *edges*, and 0D *points* can also be defined, according to the dimensionality of the investigated model.

Material properties need to be assigned to each physical entity. COMSOL allows the definition of a list of *materials*, which can then be attributed to each independent geometrical entity. Each material is a container for the several material properties needed for the aimed calculation, such as optical or thermal parameters. These can assume scalar values, tensorial values, or can be defined as functions of the system variables (and, as for such, designating variable-dependent material parameters, e.g. temperature-dependent thermal conductivity).

To easily define the *Partial Differential Equation* (PDE) system which needs to be solved, COMSOL requires the user to select or build one or more *Physics*, which are base sets of equations and constitutive relations that pertain to a physical problem. The previously mentioned RF and Heat modules are indeed ensembles of available COMSOL Physics, purposely built for the solution of radio-frequency and heat diffusion related equation systems respectively. In this work, the *Electro-Magnetic Wave* (emw) and *Heat Transfer* (ht) are implemented, which support respectively i) the calculation of the Maxwell equations in the high-frequency spectral range (suited for calculations on geometrical features with size comparable to the radiation wavelength); ii) the calculation of the heat diffusion equation in non-specific systems. One of the advantages of COMSOL Multiphysics is its native support of coupled physical phenomena, which is indeed extremely useful in view of modelling the mechanism of the PPCM this work focuses on. Thus, the solution of two coupled PDE systems, dependent on time, the e.m. field and the temperature variables (as in the aforementioned case scenario), can be pursued.

To define the specific behaviour which needs to be investigated, COMSOL requires the user to set up one or more *Studies*, which are software features that pinpoint the desired subset of equations, and the *Solver*, which collects the algorithmic tools and their parameters to pursue the studies solutions.

3.2.2. Approximations, initial conditions and boundary conditions

To adapt the physics to the model context, each COMSOL Physics implements a set of *features*, which are terms or conditions applicable to the underlying equations (e.g. boundary conditions, initial conditions, or further constitutive relations). An example can be found in the initial temperature T_{init} of a time-resolved system, which at $(t = 0)$ forces $(T = T_{init})$ where specified. Generally, unless particular conditions are expected, the software environment sets conventional and agreeable values to most of the involved variables, such as initial temperature, heat flux, input power / voltage / current, solar irradiation, and electric or magnetic field.

In both the thermal and the electromagnetic models though, it is necessary to provide the appropriate boundary condition to the geometrical features enclosing the model (for a 3D model, those would be its external surfaces). This necessity arises from the truncation of the domains, enforced to allow the calculation over a computationally affordable number of nodes. The nature and location of the boundary conditions can dramatically change the device behaviour; thus, a careful model setup is compulsory to avoid unphysical behaviour or calculation artefacts.

For the thermal model, there are two main boundary conditions applied throughout this work: the *perfect insulator* ($\boldsymbol{u} \times Q = 0$, with $\boldsymbol{u}$ being the unit vector normal to the boundary surface), and the *constant temperature* ($T = T_{user}$). The first is usually adopted where negligible or no heat diffusion (across the surface) is expected, or where a symmetric behaviour is an acceptable approximation (e.g. along the physical device symmetry plane).

One more crucial element to the numerical emulation of heat distribution in photonic devices is the thermal contact, or thermal boundary condition, which is emulated through the *Thin Layer* feature, allowing one to follow the thermal boundary formalism explored in Section 2.2.

The RF module offers a wide set of boundary conditions, but in this work the model boundaries can be tackled by use of four of the RF module features. Let the case be of a waveguide model, for sake of clarity. First, the model needs the definition of *ports*, i.e. surfaces upon which the guided mode can be calculated, excited, and absorbed. The other external surfaces can be assigned as *perfect electric conductor* (PEC, $\boldsymbol{u} \times E = 0$) and *perfect magnetic conductor* (PMC, $\boldsymbol{u} \times H = 0$). These nullify the component of the electric field (or magnetic field, in the second case) parallel to the surface. Such conditions are generally used to reflect in a lossless manner the electric (or magnetic) field components which would otherwise cross the surface. Conversely, to mimic an infinite domain (i.e. obtain a reflectionless surface) there are two methods. One is a Robin boundary condition (which imposes the value of both the local solution and derivatives to the PDE), called *scattering boundary condition*, devised for ignoring the reflection of normally-incident e.m. field, but still sufficiently effective (i.e. R = 0.01) for incidence angles below 60° (see Ref. [187]). The second method obtains a more realistic simulation of non-reflective surfaces (i.e. generally having a much lower reflection coefficient R for most of the incidence angles), at the expense of increased computational costs. It is referred to as *Perfectly Matched Layer* (PML), and relies on the adiabatic decrease of the e.m. field magnitude travelling across a purposely defined domain (of course set contiguously to the model external surfaces of interest). PMLs are the conventionally adopted method to rid of undesired reflections; but due to the size of the models involved in this work, and consequently to the much increased calculation time arising from the use of PMLs, the scattering boundary condition is here preferred, specifically opting to use second-order approximation for the plane-wave type [187]).

3.2.3. Meshing techniques

It is useful to remind the reader that the meshing operation tessellates the n-dimensional environment with sub-domains, called meshes, and consequently defines the spatial distribution of nodes (meshes corners) upon which the FEM environment calculates the PDE solution. As briefly mentioned in Section 3.2.1, the reliability of the calculation results heavily depends on meshes shape, size, and number. This subsection explains the principles upon which the most suitable meshing technique is determined.

The mesh building methods adopted in this work usually rely on variants of the Delaunay triangulation, creating triangular patterns on 2D surfaces and prisms on 3D surfaces. Sometimes it is more convenient to exploit the model geometrical symmetry. The translational symmetry of the waveguide is a useful example: a waveguide model can be meshed starting from the port 2D surface, and then “extruding” such a pattern along the surface normal axis (*swept meshing*). Alternatively, Cartesian grids can also be adopted for rectangular areas, in conjugation with the swept meshing techniques in case of parallelepipeds. The combination of all the above techniques is often encountered in the models devised for the work of this thesis.

As previously hinted, and overlooking mesh shape, the density of nodes (correlated to the mesh size) is the most impactful parameter on the solution stability and reliability. Trivially, with a higher number of nodes, the solution improves both reliability and spatial resolution. Nevertheless, a higher number of nodes also implies a higher computational requirement (more details are provided in the following section). A finely tuned meshing operation balances the need to obtain useful and reliable solutions within reasonable time, or within the barriers posed by the computational resources.

One approach for balanced meshing is found in the *Nyquist-Shannon sampling theorem* [188], which establishes that the reconstruction of a one-dimensional signal requires a sampling rate equal (or greater than) twice the signal's highest frequency. Thus, to solve for harmonically varying fields (e.g. the e.m. field) meshes size Δx must not exceed $\lambda/2n$ (n being the refractive index). This limit retains validity for far-field propagation, but becomes insufficient in case of nanophotonic devices, for the following reasons:

i. the non-trivial near-field propagation is especially important, and requires a higher sampling rate to reliably define propagating and scattered e.m. fields;
ii. nanophotonic devices usually possess sub-wavelength features, thus needing again higher sampling resolution to avoid computational artefacts;
iii. metals (and other materials with high extinction coefficients) need a separate treatment based on the resolution of the 1/e field penetration length (skin depth) $\delta \approx \sqrt{\lambda/\pi c\mu\sigma}$ [189], which at optical frequencies is of the order of a few nanometres.

Therefore, to better mesh this class of devices, the mesh density is accurately tuned to satisfy the following:

$$\Delta x \leq \lambda/\mathrm{k}'n \qquad \text{Eq. 3.2}$$

with k' being an empirically determined non-dimensional parameter, of value usually found between 4 and 10, which retains the stability of the model solution.

The above solution is suitable to reliably emulate the e.m. propagation, but the solution of the heat transfer problem PDE requires a different analysis, as it involves the non-harmonic temporal variation of the dependent variables. The *Courant-Friedrichs-Lewy* (CFL) *theorem* [190] states the necessary condition upon whose satisfaction the PDE system can converge towards a stable solution. Specifically, the CLF theorem establishes that, for a PDE system, the Courant number Co must be equal or lower than 1. For a d-dimensional systems, the Courant number is defined as follows:

$$Co = \Delta t \cdot \left(\sum_{j=1}^{d} \frac{u_{x,j}}{\Delta x} \right) \leq Co_{max} \qquad \text{Eq. 3.3}$$

where $u_{x,i}$ is the magnitude of the system length/time velocity variation, Δt is the temporal variation and Δx is the spatial variation. The CFL theorem defines Co_{max} as unary, albeit the solver method and the specific investigated problem may require lower values [191]. Using the definition of thermal diffusivity[20] $\alpha_{th} = k_{th}/\rho C_p$ [m^2/s], it is possible to relate the maximum size of the mesh based on the (minimum) time-step duration [190]:

$$\Delta x = \sqrt{\alpha_{th} \Delta t \cdot Co_{max}} \qquad \text{Eq. 3.4}$$

As for the implementation of the Nyquist-Shannon theorem, is possible to empirically determine an arbitrary value $\mathrm{k}'' \approx \Delta t \cdot Co_{max}$ via explorative simplified models.

Alternatively, an independent estimation of the time step value Δt helps to identify possible mesh criticalities. Time step generation is a complex problem, and a suitable description of the available methods can be found in Ref. [185]. An empirical approximation finds an effective Δt value 2 orders of magnitude lower than the characteristic time duration of the input transient; this approach has been indeed used during this investigation, to infer the time-step lower boundary.

In full generality, a quick method to probe the meshing quality is to determine the solution variation as a function of k' and k'', leading to the estimation of the best compromise between calculation time and satisfactory results when such variation is arbitrarily small.

3.3. Modelling of the phase-switching mechanism

In chalcogenides, especially in GST alloys, the numerical emulation of the phase transition is a non-trivial part of the modelling task. The reason lies in the nature of the phase transition mechanism, which sees the concurrent action of several self-influencing mechanisms, e.g. the growth and dissolution of crystallites within the glassy domain described in Section 2.3.1.

---

[20] The $\square_{th}$ suffix applied to the thermal diffusivity symbol α_{th} is an arbitrary choice, to distinguish this parameter from the previously defined Napierian attenuation constant α (see Eq. 2.10).

A possible approach that provides elements of understanding over the phenomena taking place during the phase transition is found in *first-principle* approaches. These, known also as *ab-initio* or *atomistic* methods. There are several possible implementations (e.g. Hartree-Fock, Density Functional Theory, Quantum Monte-Carlo, etc.), which aim to determine the solution of the system wave function. Indeed, such methods have been successfully applied to various PC material alloys, providing an interpretation of the complex dynamics of the phase transition mechanisms, as well as the material properties (see e.g. [149], [192]–[196] for the GST alloy). Nevertheless, these imply an extremely high computational requirement, which delegates their use to model small clusters (relatively small numbers of atoms). Thence, ab-initio methods application is oriented towards the analysis of fundamental properties, and is not encountered in device-oriented works such as in the current thesis. There are indeed semi-empirical methods, or techniques which use particle-like descriptions (e.g. Molecular Dynamics), which ease the burden of the computational requirement. But whilst their application is successful for specific research areas (e.g. protein folding), these are still not practical to model the volumes of PC material usually involved in device applications.

Thus, it is necessary to devise low-computational intensity methods to describe the phenomenological behaviour of the phase-change process, when oriented towards the simulation of device-scale processes. Additionally, whilst the PDE definition may be feasible, the practical application within the FEM environment can be challenging, and thus the following methods are usually implemented within a separate software environment (e.g. MATLAB).

3.3.1. JMAK model

The cumulative work and contributions of A. Kolmogorov [197], W. Johnson and R. Mehl [198], and M. Avrami [199]–[201] led to the definition of the so-called JMAK phase-change model (named after the contributors’ initials) between the years 1937 and 1941. This model describes the variation of the crystal fraction parameter of an amorphous material as a consequence of the annealing process. Despite the following years and intensity of research on the topic, this model finds still widespread application, due to the very simplistic formulation and ease of calculation:

$$\chi = 1 - e^{-(K(T)t)^{m+1}} \qquad \text{Eq. 3.5}$$

with χ being the crystal fraction, varying as a logistic function of time t, a phenomenological exponent m, and temperature, here incorporated in the effective rate term K, accounting of both nucleation and growth processes, which can be expressed as an Arrhenius term:

$$K(T) = \nu e^{-\frac{E_A}{k_B T}} \qquad \text{Eq. 3.6}$$

where ν is the frequency factor, E_A is the activation energy, and T the system temperature. The above equations have been formulated following a few key assumptions, reported below [202], [203].

i. The annealing process is carried out under isothermal conditions
ii. The reaction proceeds via nucleation and growth processes
iii. The nucleation is a process which sees the formation of spherical clusters, whose shape can deviate only after impingement with other clusters or surfaces
iv. The nucleus critical size is zero
v. The nucleation is an otherwise stochastic process, not influenced by the matrix structure
vi. The nucleation is a spatially uniform process advancing in time according to a time-independent rate
vii. The growth process is driven by a purely interface-controlled mechanism
viii. The growth velocity of the crystal phase is independent of the cluster size.

Unfortunately, most of these conditions are not met when considering the phase switching of PC materials in a device context [31]. Additionally, the JMAK model lacks the capability to distinguish the grain size distribution, deriving from the convolution of nucleation and growth mechanisms, which may be of extreme importance in understanding PCM devices behaviour.

A deviation from the JMAK behaviour is then expected, and several works have attempted to refine the results in order to meet a broader set of nucleation and growth mechanisms and geometries (see e.g. [167], [204]–[207]). A list of parameters and model adaptations suitable for emulating the GST alloys phase transition with the JMAK model can be found in Refs. [94], [167], [208].

3.3.2. Classical Nucleation-Growth (CNG) model

The *Classical Nucleation-Growth* (CNG) model, in contrast with the JMAK model, separates the nucleation and growth mechanism. The CNG model is built upon the classical theory of steady-state crystal nucleation, based on the work of Gibbs, but developed and contributed by M. Volmer and A. Weber [209], R. Becker and W. Doring [210], D. Turnbull [211] and J. C. Fisher [92] between 1926 and 1969. The theory proposes that the crystallization of an amorphous material starts with the formation of small, unstable clusters which become stable once a critical size is reached (i.e. clusters become nuclei which can grow rather than being dissolved). The driving force for crystallization is the difference in the Gibbs free energy of the amorphous and crystalline phases. The CNG approach also proposes that nucleation can occur in two different ways, i.e. via i) homogeneous nucleation, in which nucleation can take place in any region of the amorphous material, and via ii) heterogeneous nucleation, in which nucleation can preferentially occur at surfaces, interfaces, and impurities. The steady-state nucleation and growth rates in the CNG theory are typically expressed as follows:

$$\mathrm{I}_{\mathrm{nuc}}^{\mathrm{ss,hom}} = \mathrm{k}_{\mathrm{nuc}}^{\mathrm{hom}} \exp\left({}^{-E_a}/_{k_B T}\right) \exp\left({}^{-\Delta\mathrm{G}_{\mathrm{cluster}}}/_{k_B T}\right) \qquad \text{Eq. 3.7a}$$

$$\mathrm{I}_{\mathrm{nuc}}^{\mathrm{ss,het}} = \mathrm{k}_{\mathrm{nuc}}^{\mathrm{het}} \exp\left({}^{-E_a}/_{k_B T}\right) \exp\left({}^{-\Delta\mathrm{G}_{\mathrm{cluster}} f(\vartheta)}/_{k_B T}\right) \qquad \text{Eq. 3.7b}$$

$$\mathrm{I}_{\mathrm{gro}}^{\mathrm{ss}} = \mathrm{k}_{\mathrm{gro}} \exp\left({}^{-E_a}/_{k_B T}\right) \phi(T) \left(1 - \exp\left({}^{-\Delta\mathrm{G}_{\mathrm{cluster}}}/_{k_B T}\right)\right) \qquad \text{Eq. 3.7c}$$

with

$$\Delta\mathrm{G}_{\mathrm{cluster}} = \frac{16}{3}\pi\frac{\sigma^3}{\mathrm{g}^2} \qquad \text{Eq. 3.7d}$$

$$\mathrm{g} = \Delta\mathrm{H}_{\mathrm{f}}\frac{T_m - T}{T_m} \qquad \text{Eq. 3.7e}$$

$$f(\vartheta) = \frac{(2-\cos\vartheta)(1-\cos\vartheta)^2}{4} \qquad \text{Eq. 3.7f}$$

Eq. 3.7d describes the critical energy barrier for cluster formation, as discussed in Section 2.3.1. Eq. 3.7e defines the bulk free energy difference between crystalline and amorphous phase (with the free energy contribution dependency on temperature here further simplifying the approximation proposed by Sing and Holtz [212]). Eq. 3.7f defines the shape factor for heterogeneous processes, with ϑ being the contact angle between crystallite and the external surface.

Turning the attention back to the growth rate (Eq. 3.7c), a term not belonging to the original description [92], [209]–[211] but later proposed in Peng et al. [213] is here adopted, precisely $\varphi(\mathrm{T})$, whose analytical definition is:

$$\varphi(\mathrm{T}) = \exp\left(-\frac{0.8}{1-\mathrm{T}/\mathrm{T_m}}\right) \qquad \text{Eq. 3.7g}$$

This is a phenomenological factor related to the growth mode on thin PCM layers, and whose parameters arise from experimental data fitting [213]. It is hence a suitable addition for the simulation analysis carried out in the work of this thesis.

As the JMAK model, also the CNG model is based on a few rigid assumptions which for convenience are here highlighted:

i. The formation of critical clusters is counterbalanced by the loss of part of the previously formed one, independently from time or crystal fraction
ii. All clusters larger than the critical size are assumed to grow

Of course, such assumptions are useful to simplify the model derivation, but i) is met only at the early stage of the crystallisation, and fails as the amorphous regions are depleted; ii) is a restrictive assumption as it does not include the possible dissolution of larger nuclei. Additionally, the CNG model ignores sub-critical clusters[21], whose number and distribution may significantly affect the crystallization process [214].

Nevertheless, compared to the JMAK, the CNG model offers the opportunity to better comprehend the non-trivial self-influencing dynamics between nucleation and growth processes. With an opportune algorithmic implementation, it is possible to successfully adapt the CNG theory to track the formation and dissolution of individual crystals, a feature missing in the JMAK formulation since it tethers nucleation and growth mechanisms into a single equation. The algorithmic steps necessary to be performed for the CNG emulation are also computationally affordable. Therefore, like the JMAK model, the CNG model has been successfully adopted to model PCM devices (see e.g. GST devices investigation in [215] and [216]). A list of material parameters which can be adopted to model GST alloys with the CNG model is found in Refs. [94], [215].

---

[21] It is here reminded that a sub-critical cluster is the result of a (spherical) cluster formation of radius below critical radius, and that would therefore more promptly dissolve in the amorphous matrix than grow (see Section 2.3.1.1).

3.3.3. Master Rate equation model

In contrast with the JMAK and CNG models, the *Master Rate Equation* model (MRE) instead has been devised to emulate the behaviour of PC materials based on a solid grounding on the underlying chemical reactions.

The MRE model, as applied to phase-change materials, was first proposed by S. Senkader and C. D. Wright [94] in 2004, and further publications from the same research group [95], [214], [217], [218] elaborated on the model refinement and its application. The MRE focuses on the determination of the crystalline clusters concentration (within the amorphous matrix), with a set of interdependent equations for the cluster size n [monomer units], as a function of temperature T, time t, and geometry (via explicit use of the hemispherical cap model, see Section 2.3.1.2). Specifically, the MRE defines the *ordinary differential equation* (ODE) of the cluster concentration $f(n,t)$ as follows:

$$\frac{\partial f(n,t)}{\partial t} = g(n-1,t)f(n-1,t) - d(n,t)f(n,t) - g(n,t)f(n,t) + d(n+1,t)f(n+1,t) \qquad \text{Eq. 3.8a}$$

where the terms $g(n,t)$ and $d(n,t)$ describe the growth (monomer attachment) and dissolution (monomer detachment) rates respectively, as:

$$g(n,t) = 4\pi r^2 \lambda f(1,t) \gamma \mathrm{e}^{\frac{\Delta G_{n\rightarrow n+1}}{2k_B T}} \qquad \text{Eq. 3.8b}$$

$$d(n+1,t) = 4\pi r^2 \lambda f(1,t) \gamma \mathrm{e}^{\frac{\Delta G_{n+1\rightarrow n}}{2k_B T}} \qquad \text{Eq. 3.8c}$$

Here, r is the cluster radius, λ is the average interatomic distance, and $\gamma = D/\lambda^2$ is the molecular jump frequency, obtained as a ratio between the atomic diffusivity D (see Eq. 2.49) and the square of the interatomic distance. The values of

$\Delta G(r,T)$ can instead be obtained as previously specified (see e.g. Eq. 3.6d-e), with the additional layer of refinement included using the spherical cap model.

Whilst the MRE model shares a common ground with the CNG theory, it is characterized by a far simpler implementation, shown in Eq. 3.8a. The material parameters can indeed be tuned to reproduce experimental results (see e.g. [94]), with the same drawbacks of the previously illustrated models (e.g. a possible arbitrariness in the identification of the correct material parameters, a concept already explored in Ref. [219]). Regardless, whilst the MRE is so far the most realistic phase-change model, it also poses a significant challenge on its practical implementation. Being an intrinsically adimensional model, the MRE loses the spatial information on the crystal distribution, which is detrimental for the simulation of devices heavily reliant on such characteristic (e.g. PC memristive devices). Additionally, the solution of very big ODE systems (as the MRE requires) is a computationally-intense calculation, prone to numeric errors due to the large variability of the orders of magnitude of the involved variables (e.g. the distribution function f). This approach thus suffers from a prohibitive computational cost when evaluating the phase transition of a large number of meshes, an eventuality faced with large volumes, but also when involving high temperature spatial non-uniformity and/or gradients. Concluding, this method can be extremely efficient and accurate for systems not reliant on a highly resolved crystal spatial information and involving a large degree of symmetry, whose calculation can be performed on a simplified geometry which can then be unfolded to cover the original investigated volume.

3.3.4. Gillespie Cellular Automata (GCA) model

The Gillespie-Cellular Automata model (GCA) [220] has been derived from the MRE, integrating concepts from the phase-field theory [221], and implemented onto a Gillespie-cellular automata framework [222]. This model was initially devised with the view to reintroducing the spatial information lost with the MRE, easing its computational cost and yet retaining its formalism. The result bridges atomistic and empirical models, operating with a nanometre-scale resolution but reliant on a classical description.

The GCA model considers a homogeneous, isotropic material in a square lattice of size 0.82 nm (equivalent to the diameter of a GST spherical monomer, as calculated from the density and molar weight of GST [94]). The state of each site (i,j) in the (2D) lattice is outlined by two quantities: the phase, r_{ij}, and the crystal orientation, φ_{ij} (with two adjacent crystalline sites belonging to same crystallite if they have the same orientation). The local changes that can occur are defined by three events: i) nucleation, where two neighbouring amorphous sites become a single crystalline cluster; ii) growth, where an amorphous site (i,j) becomes attached to an adjacent crystal; iii) dissociation, where crystalline site (i,j) detaches from the original crystal to become amorphous. The occurrence rate of these are defined as follows:

$$\mathrm{C}_{ij}^{\mathrm{nuc}} = \begin{cases} \text{if } r_{ij} = 0, & k_0\, exp\left(-E_a/k_B T\right) \dfrac{n_{ij}^{am}}{n_{ij}} \xi(T, S_m) \\ \text{else}, & 0 \end{cases} \qquad \text{Eq. 3.9a}$$

$$\mathrm{C}_{ij,\psi}^{\mathrm{gro}} = \begin{cases} \text{if } r_{ij} = 0, & k_0\, exp\left(-E_a/k_B T\right) \xi\left(T, S_m\, \theta\left(n_{ij}^{\psi}\right)\right) \\ \text{else}, & 0 \end{cases} \qquad \text{Eq. 3.9b}$$

$$\mathrm{C}_{ij}^{\mathrm{dis}} = \begin{cases} \text{if } r_{ij} = 0, & 0 \\ \text{else}, & k_0\, exp\left(-E_a/k_B T\right) \xi\left(T, S_m\, \theta\left(n_{ij}^{\psi}\right)\right)^{-1} \end{cases} \qquad \text{Eq. 3.9c}$$

with

$$\xi(T, A) = exp\left({}^{\Delta g(T)}\!/_{2k_B T} - {}^{\sigma A}\!/_{2k_B T}\right) \qquad \text{Eq. 3.9d}$$

$$\Delta g(T) = v_m \Delta \mathrm{H}_f \frac{T_m - T}{T_m}\left[\frac{7T}{T_m + 6T}\right] \qquad \text{Eq. 3.9e}$$

$$\theta\left(n_{ij}\right) = \frac{n_{ij} - 2n_{ij\varphi}^{or}}{n_{ij}} \qquad \text{Eq. 3.9f}$$

where k_0 is the molecular jump frequency, E_a is the activation energy of nucleation and growth processes [94], T is the temperature, S_m is the surface area of a single site. The n_{ij} notation represents the Moore set of neighbours of a (i,j) site, with the subscript $\square_{am}$ referring to neighbouring amorphous monomers, and the superscript $\square^{\psi}$ referring to the neighbouring crystalline monomers with same crystal orientation φ_{ij}. Eq. 3.9d defines $\xi(T,A)$, the thermodynamic contribution to the (i,j) site interaction. Here, σ is the interfacial energy at the crystalline-amorphous boundary, and the temperature-dependent bulk free energy difference term (per unit molecule) $\Delta g(T)$ is further illustrated in

Eq. 3.9e [94], [212]. $\Delta\mathrm{g}$ is a function of the enthalpy of fusion, $\Delta\mathrm{H}_f$, the monomer volume, v_m, and the melting point T_m, with the assumption of local equilibrium for the enthalpy of transition (i.e. $\Delta\mathrm{H}_{\mathrm{fusion}} = \Delta\mathrm{H}_{\mathrm{solidification}}$). Eq. 3.9f defines the coordination factor $\theta(n_{ij})$, which scales the value of $\mathrm{S_m}$ according to the $\mathrm{n_{ij}}$ neighbours phases and orientations.

This set of calculations needs to be performed for each (i,j) site, and updated after any variance of the phase and orientation of the (i,j) site or its neighbourhood n_{ij} (since the rates are also affected by the neighbourhood state). Such update occurrence is referred to as *event*. The process that performs the (i,j) phase and orientations update instead follows the strategy dictated by the Gillespie method.

The first step of the GCA method requires the calculation of the activity a_0, defined as follows:

$$a_0 = \sum_{i,j} \mathrm{C}_{ij}^{\mathrm{nuc}} + \sum_{i,j} \sum_{\psi \in \Psi_{ij}} \mathrm{C}_{ij,\psi}^{\mathrm{gro}} + \sum_{i,j} \mathrm{C}_{ij}^{\mathrm{dis}} \qquad \text{Eq. 3.10}$$

specifying that the Ψ_{ij} notation includes all the possible ψ orientations of the crystal sites in n_{ij}. The time $d\tau$ required for an event to take place is generated as $d\tau = 1/a_0 \cdot \log(1/\eta')$, with η' being randomly chosen from a (0,1) uniform distribution; the type of event (i.e. nucleation, growth, dissociation) is instead selected via *Inverse Transform Sampling* (ITS), from the *Cumulative Distribution Function* (CDF) array calculated as the cumulative summation of the activity contribution from each site and type of event. As these operations are carried out, a site and event are uniquely selected, and the relative operation is carried out onto the phase and orientation fields, to which follows the update of the rates and the CDF array.

This algorithm results in the emulation of a stochastic process, previously unavailable, driven by the same principles underlying the MRE method. The complex dynamics of crystal formation and dissolution emerges from this relatively simple set of equations, with the advantage of requiring a relatively low computational expense. An example of its results can be found in Fig. 3.2. This method has been developed for a 2D surface, and has been successfully adopted for the simulation of various devices requiring the investigation of the crystal conformation at the mesoscale, which understanding is essential for electrical devices. The parameters used for the portrayed simulation are listed in Table 3.1.

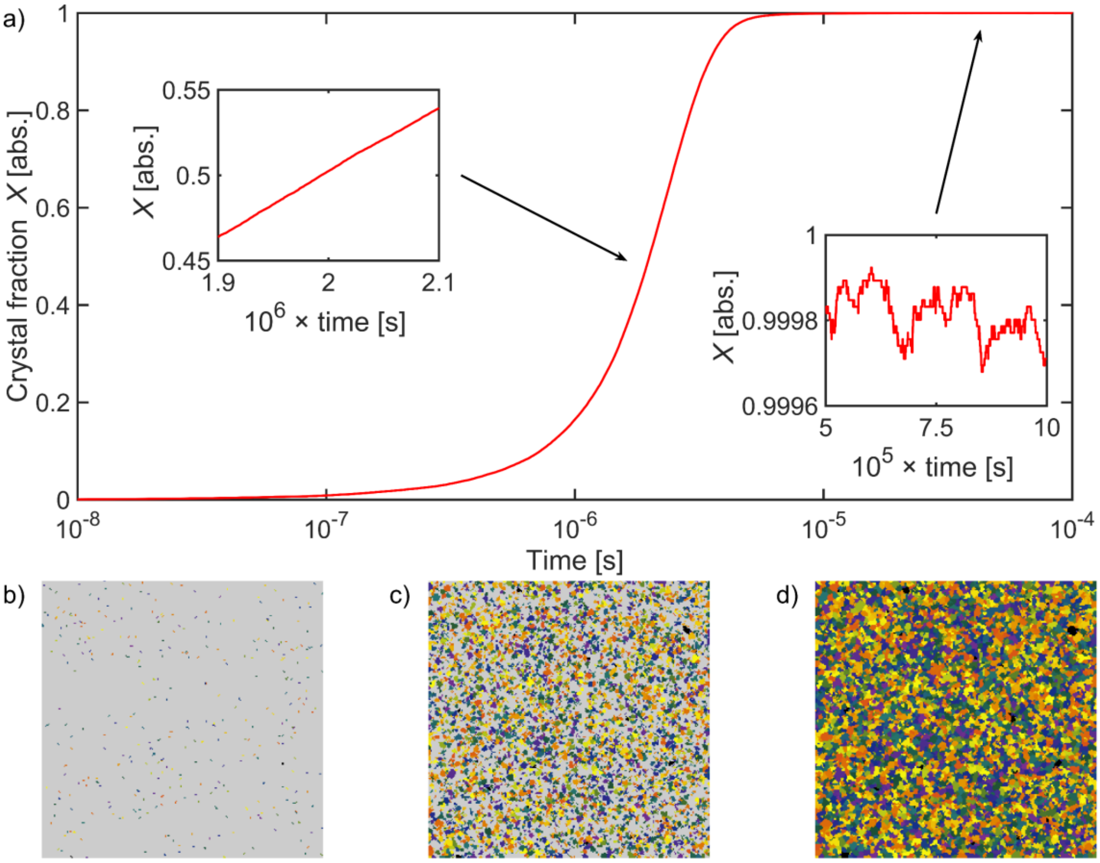


Fig. 3.2. Gillespie Cellular Automata, simulation results of the annealing performed at 573 K (300 °C). (a) Crystal fraction χ versus time t (s) during the annealing process. Details of the progress are shown as insets, during the growth phase (top-left) exhibiting a smooth behaviour, and at completed recrystallization (bottom-right), showing the stochastic attachment and detachment of monomers at $\chi \cong 1$. (b-d) Images showing the GST crystallization evolution, starting with a pure amorphous material (grey), and ending with the full recrystallization (coloured crystallites). Figures were obtained at (b) $t = 5 \times 10^{-7}$ s, $\sim 3.1 \times 10^3$ steps, $\chi = 0.05$; (c) $t = 2 \times 10^{-6}$ s, $\sim 2.8 \times 10^4$ steps, $\chi = 0.5$; (d) $t = 10^{-4}$ s, $\sim 6.1 \times 10^5$ steps, $\chi = 0.9997$.

Table 3.1. Parameters adopted for the GCA phase transition simulation demonstrated in Fig. 3.2.

Symbol	Parameter	Value	Reference
σ	Interfacial energy	0.066 [J/m^2]	[219]
S_m	Molecular surface area	2.112×10^{-18} [m^2]	[220]
E_a	Activation Energy (nuc., gro.)	2.1 [eV]	[220]
k_0	Molecular jump frequency	10^{22} [Hz]	[220]
ΔH_f	Enthalpy of fusion	1.121×10^{-9} [J/m^3]	[219]
v_m	Molecular volume	2.887×10^{-28} [m^3]	[220]
T_m	Melting temperature	889 [K]	[220]

To date (i.e. prior to this thesis), only the 2D formalism for the GCA had been developed. However, in many device configurations (e.g. where rotational symmetry does not exist) a 3D form may be needed. Moreover, the scalability of the pre-existing model is poor, as the computational requirement, albeit low, grows as $\mathcal{O}(n^2)$. Such a method thus requires further adaptations to be useful for simulations of many real-world device configurations and larger PC material volumes (as found in photonic devices). The extension of the model to 3D and improvement in its scalability has therefore been carried out as part of this thesis (and is discussed in Chapter 4).

Chapter 4. Advanced modelling development

The research work which is reported in this thesis has been initially focused on the simulation and analysis of the phase-change photonic memory (which is detailed in Chapter 6 and Chapter 7). However, the development of the modelling framework revealed to be a compulsory prerequisite to pursue such a scope. This Chapter illustrates the progress towards the realisation of the required simulation framework, which touches a diverse range of aspects. Specifically, the contribution to the scientific knowledge is here outlined in the definition of a fast and reliable 3D phase-change model for photonic devices, of the FEM modelling development for the phase-change photonic memory, and the integration of those elements onto an automated algorithm that i) transfers the required variables between software environments and ii) determines a suitable in-model temporal progress to retain solution stability whilst reducing the computational requirement.

4.1. 3D phase-change modelling

In device-oriented applications (and relative literature), the PC model is often restricted two-dimensional simulations; the reason is mostly based on the decrease of the computational requirement, but also due to the complexity of a practical implementation of the 3D model. The applicability of this approach is limited to domains with high symmetry, allowing the folding of a 3D domain onto a 2D representation.

However, the device symmetry is not the only factor to consider when pursuing such strategy: another fundamental condition is the symmetry of the various fields involved in the phase transition (e.g. temperature). It is easy to picture scenarios facing significant challenges when trying to bend the shape of both the PC material volume and the variables field onto a low-dimensional surface[22]. Thus, since a 2D model is not always suitable, the full support of 3D environments is a highly desirable feature in phase-change models.

---

[22] A single off-center heat point source within a rectangular parallelepiped already presents challenges when trying to simulate the phase transition on a single 2D area. In this example, approximations can be made simulating the phase transition on three orthogonal cut planes (and then extrapolating the phase information outside the planes), but the single-point heat source is still conveniently simple, as it lays completely within the proposed orthogonal planes.

The JMAK and MRE are natively 0-dimensional models, which can then be easily adapted for such scope. In contrast, the GCA usually finds only an explicit 2D implementation, due to computational intensity reasons. It would be extremely beneficial to tackle the optimization issues and missing algorithmic elements that have, to date, prevented the application of the GCA to 3D environments. This is the driving force for the realization, in the work of this thesis, of an expanded GCA framework for the modelling of the phase transition (specifically, for GST alloys, but trivially adaptable for any PC material) in a fully 3D environment[23]. Branching from the proposed model, a simplified CNG model based on the developed 3D cellular automata framework is also demonstrated.

4.1.1. 3D cellular automata

The previously illustrated implementation of the 2D cellular automata, summarized in Section 3.3.4 and more thoroughly described in Ashwin et al. [220], requires a rectangular PC material area made up from square cells, each containing a monomer of PC material and uniquely identified by their cartesian coordinates (i, j). In the 2D domain, the GCA algorithm assigns a set of variables (e.g. phase, orientation, rates, etc.), having same dimensionality (2D) and shape (i rows, j columns). For convenience, those variables will hereby be referred to as *GCA terms*. It is here underlined that the cartesian notation has a dual function:

i. it allows the algorithm to fetch the monomer GCA terms stored in the 2D arrays
ii. it correlates the monomers by their spatial location, giving access to a trivial retrieval of the (Moore[24]) neighbourhood coordinates by procedural addition (or subtraction) of a unit to the (i, j) coordinates

Thus, there is an implicit biunivocal correspondence between the monomers' spatial arrangement and their algorithmic GCA term representation. Moreover, the rectangular section eases the algorithm implementation (as the 2D array notation is supported in both C and MATLAB), and the investigated area is still convenient for examining the cross-sectional phase-distribution in devices with cylindrical symmetry [159], [223], [224].

---

[23] This Section describes the algorithmic routines performed on MATLAB, as this software facilitates the data exchange with the COMSOL software environment (see section 4.2). MATLAB does not explicitly support the *pointer* data type, which would suggest different (and possibly more efficient) coding strategies. Therefore, those strategies are here purposely avoided.

[24] A 2D Moore neighbourhood considers both edge-sharing and corner-sharing vicinal monomers as part of the neighbourhood set. See Ref. [339] for further information.

Another reason for the missing development of a fully 3D GCA framework lies in the computational requirement increase. One of the most demanding tasks of the GCA implementation is the *Inverse Transform Sampling* (ITS), which for each iteration requires the calculation of the cumulative sum of the $3M$-sized[25] CDF array (M being the total number of monomers, rapidly increasing with the modelled domain dimensionality)[26]. Also the Moore neighbourhood set cardinality of a D-dimensional region scales as $3^D - 1$, and so does the rates-update algorithm complexity as $\mathcal{O}(n^\alpha)$ (with $\alpha \propto D$).

Due to the intrinsically asymmetric nature of the PPCM devices analysed in this work, the original GCA model of Ashwin et al. [220] is unsuitable and too computationally demanding to manage for 3D computations[27], requiring the development of a more flexible and optimized algorithm. The following issues need to be addressed:

i. lack of the flexibility required to model asymmetric PC material regions
ii. lack of a 3D framework, which is found to be necessary for a more accurate description of asymmetric temperature (or, generally, driving force) fields;
iii. fast growth of the GCA computational demands as the number of monomer increases.

The method developed in this thesis thus modifies the GCA implementation as follows:

i. explicitly detangles the spatial correspondence between the investigated region and variable arrays, by use of a flexible 1D notation for any arbitrary D-dimensional array
ii. improves the ITS method via use of nested arrays calculation.

The scope, advantage and results of each operation are illustrated below.

---

[25] The factor 3 arises from the three different mechanisms (nucleation, growth, dissociation).
[26] The Inverse Transform Sampling, and the associated Cumulative Distribution Function concepts, are integral part of the GCA algorithm, serving the function to reconstruct the probability distribution of each interaction (event) within the model. See Ashwin et al. [213] for more details.
[27] Amusingly, the first test model attempt (counting $\sim 23 \cdot 10^6$ monomers) reported an estimated calculation time of 10 years to completion.

4.1.1.1. Explicit unidimensional coordinate system

As described above, for each (i,j) monomer of the modelled region, the GCA algorithm must store the following GCA terms: phase r_{ij}, orientation φ_{ij}, rates $\mathrm{C}_{ij}^{\mathrm{nuc}}$, $\mathrm{C}_{ij,\psi}^{\mathrm{gro}}$, $\mathrm{C}_{ij}^{\mathrm{dis}}$, activity a_{ij}, CDF, as well as partially solved components of the GCA equations. The original GCA algorithm stored each of the GCA terms in a set of 2D arrays, whose size matched the investigated area, with the cartesian coordinate providing burden-less access to any of the monomer properties, and to the address of the monomer neighbourhood. It is here reiterated that in the GCA model (as in any cellular automaton) the spatial relation between monomers is a fundamental element, whose lack of optimization would impact the functioning and performance of the algorithm.

The boundaries of the investigated area are also critical elements, as the neighbourhood set of the surficial monomers differ from those within the bulk, potentially leading to calculation artefacts or unpredictable code behaviour. This was originally solved expanding (by 1 monomer unit) the size of each GCA term array in any direction, limiting the calculation to the inner area (ignoring any effect on the boundary monomers).

Explicitly addressing the mathematical expression of the monomers' location, instead of relying on a practical but rigid algorithmic device, becomes extremely advantageous for flexibility purposes. This allows one to i) release the restriction to rectangular 2D surfaces without added computational costs, and ii) significantly reduces the number of out-of-boundary monomers, thus reducing the overall computational cost.

It is here proposed to store each variable on a unidimensional array, obtained by column-wise decomposition of the 2D monomer array. The loss of the spatial information can then be restored by dropping the cartesian notation in favour of the unidimensional (m) notation, combined with a pre-calculated look-up table (referred to as *Pointer matrix* P) which seamlessly finds the needed unidimensional coordinates. In more details, the (m) coordinates of each of the monomer neighbours are stored in P, at the column coordinate corresponding to the monomer (m). The algorithm then fetches any required coordinate via a read instruction on P. The concept is illustrated in Figure 4.1.

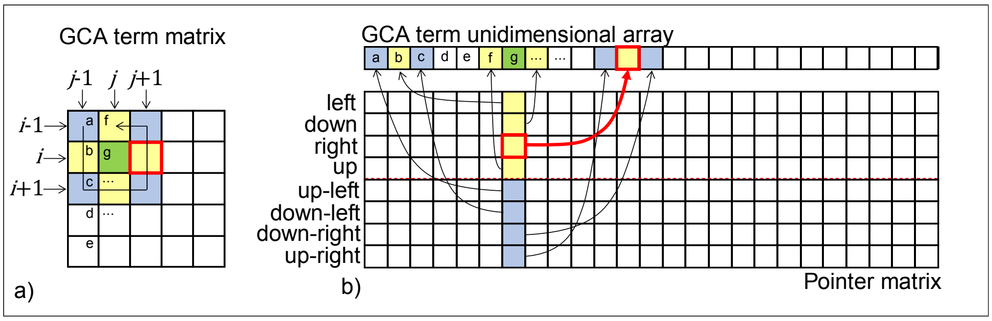


Figure 4.1. Schematics of the original and proposed GCA neighbourhood indexing strategies. a) GCA term matrix and coordinate system of the Moore neighbourhood, colour-coded so to distinguish first neighbours (yellow) and second neighbours (blue) of the highlighted monomer (green). b) Corresponding representation via *unidimensional array*, conjugated with the correlated *Pointer matrix*. The proposed column-wise decomposition is inferable observing the monomer letter indices. The coordinate system here considers a single-valued notation, explicitly stored in the pointer matrix. The example shows the process for the same neighbouring monomer identification as in (a) ("right" monomer). Here, the identification retrieves the coordinate from the Pointer matrix at the position dictated by the (m) coordinate (green square).

The boundary regions can then be explicitly addressed by storing in the P matrix a univocal coordinate of the unidimensional array (conveniently taken as $(M+1)$), referred to as nP or *nil-pointer*, redirecting to a dummy neighbour. Hence, a single boundary monomer is adopted to represent any out-of-boundary position of an M-sized region; and each GCA term would be stored on a $M+1$ sized array. This strategy is illustrated in Figure 4.2. It is here added that each GCA term (nP) property would be ideally updated with a value negating any effect on the calculation, preventing non-realistic outputs driven by the surface elements. Alternatively, this method could be exploited to apply localized surface-bound properties (e.g. surface energy), also adopting multiple dummy pointers.

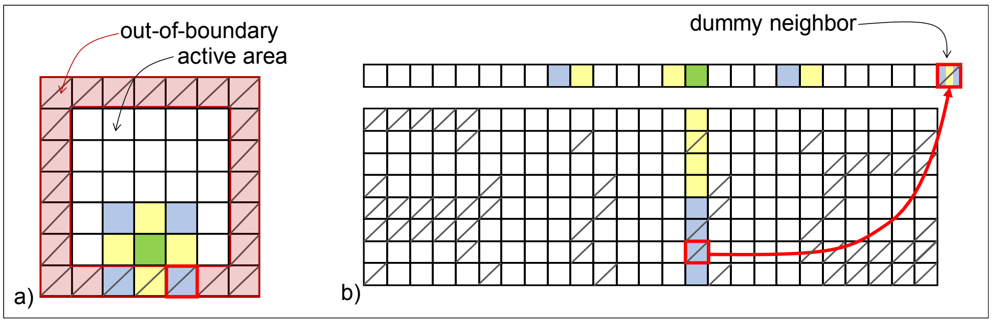


Figure 4.2. Surficial monomers representation. a) Original implementation, negating the surfaces influences by expanding the matrix by one monomer unit on each direction, and restricting the calculation onto the inner area. b) Proposed solution, which assigns a predefined dummy neighbour where needed via Pointer matrix. The red features in a) and b) depict the out-of-boundary neighbour selection process.

Not being tied by the investigated area, the data structure outlined in the previous section can then be not only optimized for the calculation algorithm but also used to represent any non-symmetric or disjointed investigated area. Conveniently, this strategy applies to any region dimensionality $D > 1$. It is here underlined that the proposed unidimensional notation does not introduce any additional computational complexity during the neighbour lookup, as in place of simple procedural calculations, a one-step read operation is performed. This comes at the expense of higher memory allocation for the P matrix, and of a one-time call to the procedure generating such matrix. The methodology to calculate the P matrix and the details on its 3D implementation are illustrated in Appendix II.

4.1.1.2. Inverse Transform Sampling algorithm enhancement

The ITS is one of the most computationally demanding elements of the GCA: it can be a relatively slow calculation, which must be performed during each GCA sequential iteration. Specifically, the ITS performs the following:

- calculation (or update) of each rate for every monomer
- calculation of the CDF array for every rate
- selection of a random location on the CDF array, via bisection method, returning both value and coordinate (the latter is needed in the GCA)

The CDF calculation was originally not optimised [220], being carried out each iteration from index (1) to index $(3 \times M)$. Yet, the CDF variation takes place in a predictable manner, as after a rate update at the index (x) (with $1 \geq x \geq 3M$) the summations at $m > x$ reiterate the previous results, and the summations at $m \geq x$ update the CDF subtracting the former rate in (x) and adding the new one.

Upon this knowledge, and building on the explicit 1D array notation of the CDF, it is possible to envisage a new algorithm which potentially lowers the computational complexity from $\mathcal{O}(n)$ to $\mathcal{O}(\sqrt{n})$. The technique, referred to as *split method*, is conceptually represented in Figure 4.3. The idea is based on the splitting of the CDF array, of size $3M$, into a sequence of sub-arrays (*chunks*). The update operation will then be carried out only on the updated chunks, thus excluding from the CDF recalculation a large portion of the rates summation.

A total of c chunks is required, each of size c_a, such as $c \cdot c_a = 3M$. A reasonable choice for the value of both c and c_a is $\sqrt{3M}$, with trivial adjustments required in case of $\sqrt{3M}$ being irrational. An additional array, *summary*, is needed to contain the cumulative sum of the chunks last element. The worst-case scenario for the split method sees a maximum of 28 cumulative summations[28], which translates into $28 \times c_a$ summations; one additional cumulative summation over the last elements of the chunks (to update the summary array) is also required. The total number of summations counts to $(28+1) \times \sqrt{3M}$; therefore, the split method becomes advantageous over the single-shot technique when $3M > 29^2$.

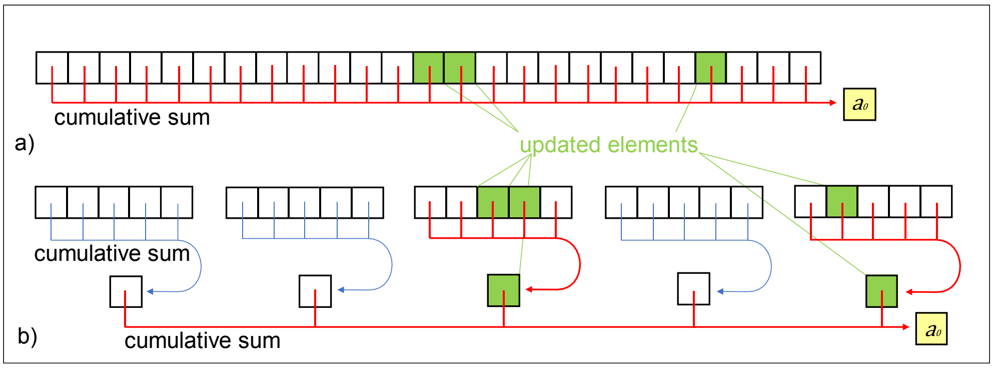


Figure 4.3. Schematics of the CDF update method, and ITS activity a_0 calculation. a) Previous implementation, which recalculated the CDF array after variation of any of its elements. b) Proposed method, which splits the CDF array and performs the cumulative sum operations only where any variation is observed (red lines), added to one further cumulative sum to yield the aimed activity value.

A test of the proposed algorithm has been carried out, with the view to quantitatively compare it to the conventional (single-shot) methods. The results are reported in Figure 4.4, as a log-log plot. As foreseen, a clear advantage of the split configuration is found for large-sized arrays. In the investigated array size range, the split method results to be up to 3 orders of magnitude faster than the single-shot method. The expected complexity of the two methods is also confirmed and evidenced by plotting the same dataset on the linear scale (see insets). Figure 4.4 also show that the proposed method becomes favourable at sizes $3M > 10^4$, differing from the theoretically calculated value (29^2); this has no theoretical grounding, but is due to the sum subroutines that MATLAB involves.

---

[28] The corner-sharing nucleation region is the worst case scenario, with a total of 14 vicinal monomer columns (belonging to the shared neighbourhood set). If the chunks correspond to each column, and see an additional boundary crossing the shared neighbourhood set, then the total number of chunks is twice the number of vicinal columns.

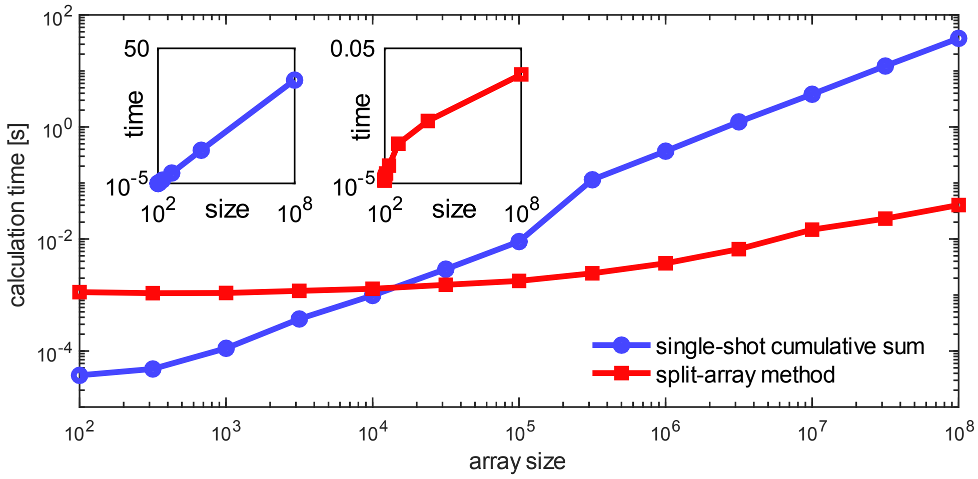


Figure 4.4. Performance comparison (log-log plot) of the inverse transform sampling methods. The blue line corresponds to the former single-shot implementation; the red line reports the data for the proposed split-array method. Insets reports the same datasets plotted on cartesian axis. It is clearly visible that, for larger sized arrays, the proposed methodology can reduce the calculation time by up to 3 orders of magnitude, whereas for smaller arrays the single-shot method is still the most suitable choice.

4.1.2. 3D Gillespie Cellular Automata

The 3D GCA developed in this thesis combines the elements described in Section 4.1.1 with the former GCA algorithm implementation of Ashwin et al [220].

The functionalities of the 3D GCA model are illustrated in the sub-sections below, as well as a comparison with the previously developed 2D implementation.

4.1.2.1. Degenerate case: 2D cellular automata

This sub-section compares the results of a simple 2D test case with the output of the model reported in Ashwin et al. [220]. The aim is to confirm the consistency with the previous algorithm. For this task, the same equation set and simulation parameters reported in Ashwin et al. are chosen, except the interfacial energy σ which is here assumed to be σ = 0.044 [J/m^2] (the mismatch is possibly due to a typo in the manuscript of Ashwin et al.).

In the first case, a 2D region of 256 × 256 monomers is initially set in the amorphous state, subjected to an annealing temperature (i.e. assuming a thermal bath maintaining the temperature constant across the simulated region) of 404 K (131 °C). The phase spatial distribution is monitored at crystal fractions χ = 0.05, χ = 0.2, and χ = 0.97 (see Figure 4.5(a-c)) The simulation result is compared to Ashwin et al. (Figure 4.5(d-f)). There is a quite good agreement between the outputs of the current model and the previous model. The set crystallization thresholds are reached at similar times in a-d and b-e; and similar characteristics are found in the grain size and distribution, for each plot. The total annealing time is different, although the value of Ashwin et al. has been obtained for a crystal fraction value of 0.999, in place of 0.97 used here. Such difference contributes to the temporal divergence. It is also noted that any variation of crytical model parameters (such as the interfacial energy), now difficult to track, could also play a role on the found simulations divergence; however, such variation is difficult to quantify.

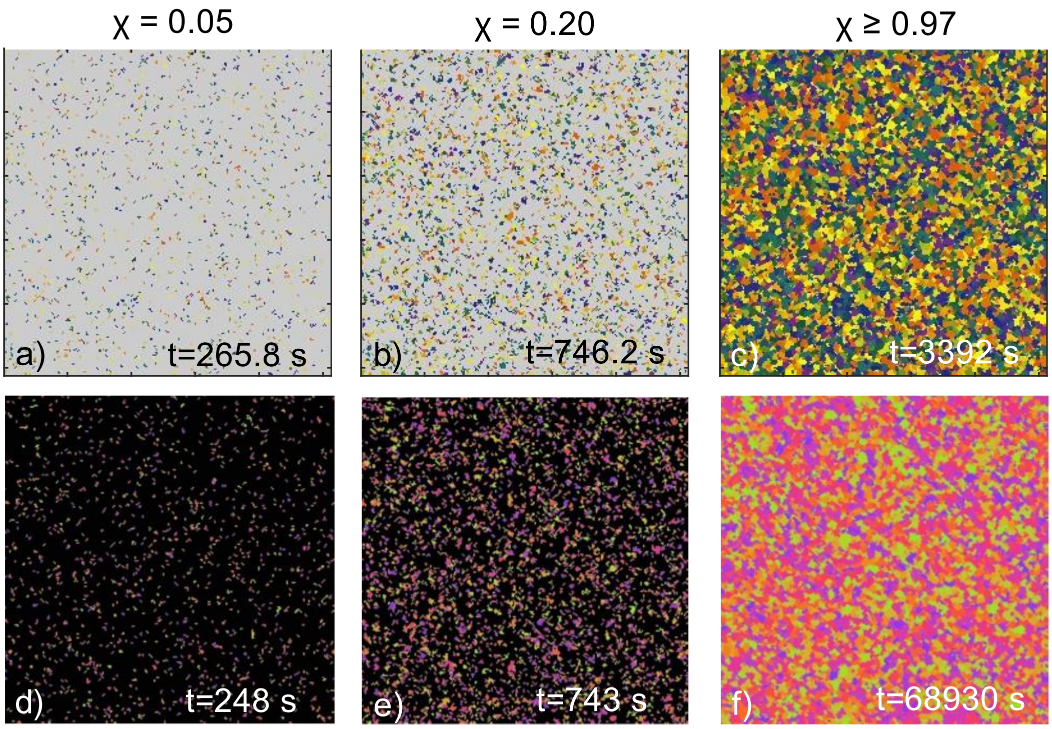


Figure 4.5. 3D GCA, simulation results of the 2D degenerate case (a-c), and comparison to the previous model results (d-f) (reprinted from [220]). The process here illustrated is the annealing of an amorphous region at a constant temperature of 404 K. Each panel represents the phase spatial distribution at different crystal fraction (a,d: 0.05, b,e: 0.2, c: 0.97 (0.999 for f)); crystal grains are visible as coloured areas over the uniform background. A good agreement of timings, grain size and grain distribution is reached.

A second experiment is carried out, with the only parameter variation being the annealing temperature now set to 680 K (407 °C). Results and comparisons are reported in Figure 4.6. At higher temperatures, larger crystal grains are expected, due to the temperature dependency of the size of stable grains, a phenomenon experimentally verified for GST [148] (more details are provided later in this section). The original model conveyed such a result (see Figure 4.6(d-f)); and the current model finds similar solutions for both timings and grains size distribution (see Figure 4.6(a-c)). As in the previous case, a deviation of the grains size is found for the last time frame (see Figure 4.6 (c-f)). The annealing at such high temperature leads to Ostwald ripening-like effects, i.e. promoting the growth of larger grains at the expense of the disappearance of smaller grains. Such effect is stronger in Ashwin et al., due to the larger number of algorithmic steps (as the monitored crystal fraction is larger). This case also highlights a larger divergence of the timings at lower crystal fraction values, possibly due to minor variations from the originally adopted parameters, but otherwise not clearly identified.

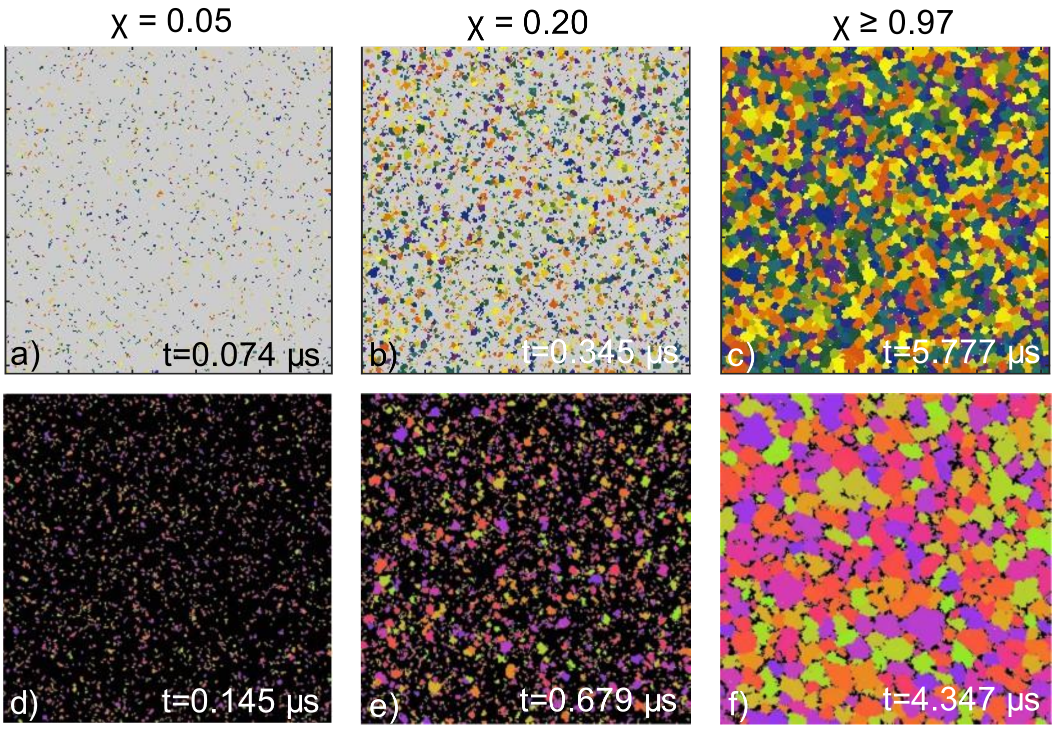


Figure 4.6. 3D GCA simulation results of the 2D degenerate case annealing process at 680 K (a-c), and comparison to the previous model results (d-f) (reprinted from [220]). As in Figure 4.5, each panel row identifies a unique crystal fraction threshold (a,d: 0.05, b,e: 0.2, c: 0.97 (0.99 for f)). Once more, timings, grain size and grain distribution of the current model and Ashwin's model find a good correlation, except for the last panel (due to the different monitored crystal fraction).

In Ashwin et al., a quantitative comparison of the simulated and experimentally measured grain size distribution is absent. Here, this additional element is provided, with a focused comparsion between the simulator output and the previously mentioned data reported in Xu et al. [148]. Specifically, average grain sizes of 4.4 nm, 5.5 nm, and 6.1 nm are reported for the GST annealing at the temperatures of 220 °C, 250 °C, and 300 °C respectively, for a simulated area of 256×256 monomers. Figure 4.7(a-c) report the crystallization pattern obtained reaching a crystal fraction of $\chi = 0.97$, starting with a completely amorphous 2D region, and setting the aforementioned temperatures for the three simulated annealings. An interfacial energy parameter $\sigma = 0.066$ [J/m^2] has been used for this task (adopted from the later work of Diosdado et al. [223]), as a lower value leads to much decreased grain sizes. The histogram data (Figure 4.7(d-f)) collects the grain sizes, calculated from the shown outputs as the diameter of the grain-fitting circle. The median diameter is eventually calculated from these datasets: as visible, the three median values have a remarkable match with the experimental data, with values of of 4.45 nm for 220 °C, 5.31 nm for 250°C, and 6.62 nm for 300°C are calculated. This experiment further confirms the suitability of the developed model to pursue the numerical investigation of the crystallization dynamics for the modelled material, with the proviso of careful parameter selection.

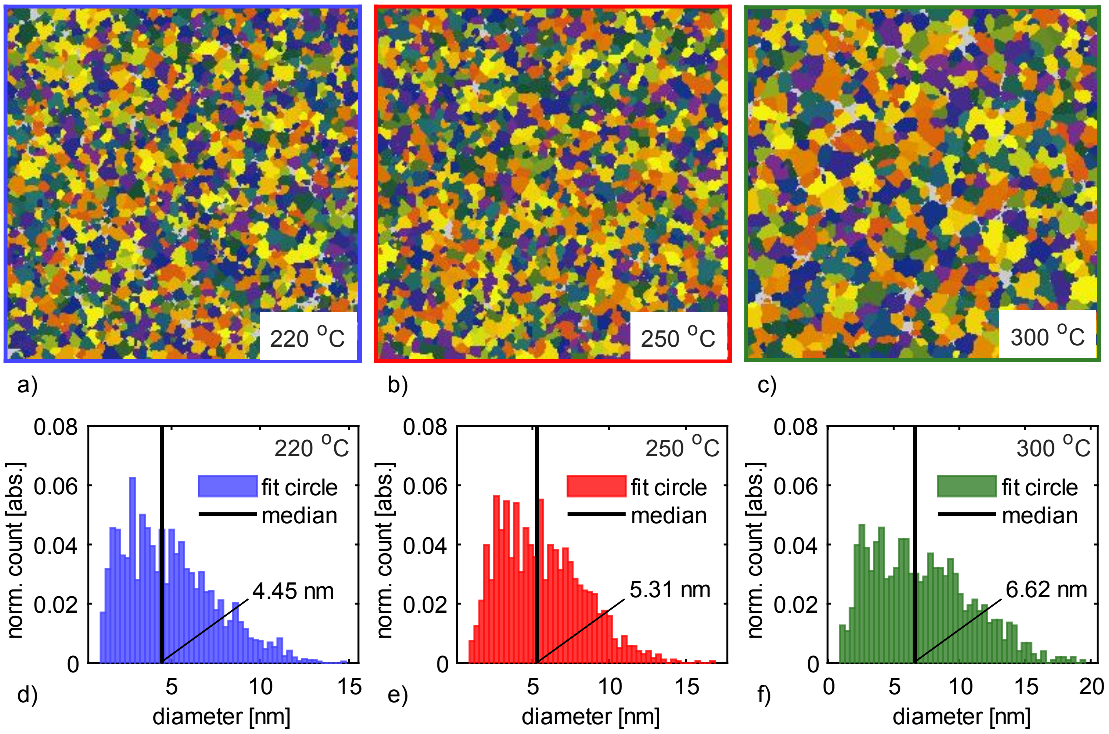


Figure 4.7. Determination of the average (median) grain size from the simulation output at the annealing temperatures of 220 °C (a), 250 °C (b), and 300 °C (c). (d) Histogram reporting the distribution of the grain sizes from the simulated experiment, and comparison of the median data with the experimental one reported in Xu et al. [148].

4.1.2.2. Comparison between 2D and 3D model outputs

With the previous section validating the consistency between the outputs, it is now possible to compare the 2D output with the 3D output by use of simple test models. The aim is here to identify any dependency of the model output from the model dimensionality. Similar results are expected, as the 2D case is assumed to be a cut-plane representation of a 3D volume (assuming translational symmetry). Nevertheless, the original model has a few degrees of freedom, tuned to reproduce experimental results (e.g. the surface energy term), so a minor mismatch on the dynamics of the crystal growth may be present.

The annealing of an initially amorphous test model is carried out with a fixed temperature of 500 K. The results are reported in Figure 4.8, for the crystal fraction values of 0.2, 0.5, 0.97. The first row (a-c) illustrates the result for a 64 × 64 monomers 2D area; the second row illustrates the result for a simulated 64 × 64 × 10 monomers 3D volume[29].

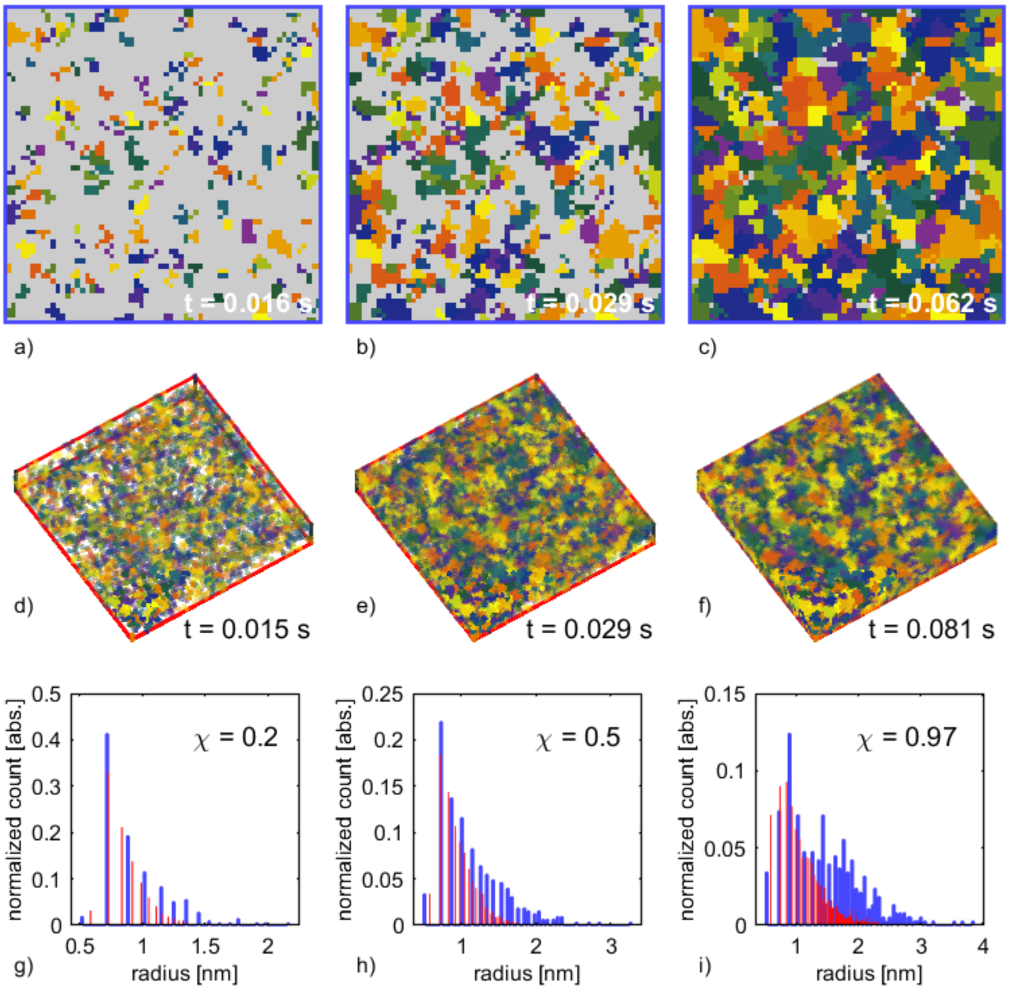


Figure 4.8. Comparison of the annealing simulation result at fixed temperature (500 K) with the 2D and 3D GCA models. The 2D simulation (a-c) is carried out on a 64 × 64 monomers surface, and the 3D simulation (d-f) emulates a 64 × 64 × 10 monomers volume. The analysis of the grain size distribution is reported in the histogram plots in (g-i), with the blue dataset reporting the 2D case analysis, and the red dataset for the 3D case. Data is reported for the crystal fraction values of $\chi = 0.2$ (a,d,g), $\chi = 0.5$ (b,e,h), and $\chi = 0.97$ (c,f,i).

---

[29] It is here remarked that only the crystal sites are displayed in any of the following 3D plots.

The results show a clear match of both the crystallization time and its kinetics. The analysis of the grain sizes (for each crystal fraction value) is reported in a histogram on the third row (g-i), which plots the normalized count of the crystal grains vs the approximating spherical or circular particle radius. These also show and quantify the expected deviation of the crystal formation.

The same experiment is then carried out imposing a temperature gradient, where the model left boundary is set at 500 K, and the right boundary is set at 700 K. A diverse distribution of grain size is expected across the investigated region, from one boundary to the other, as well as a different incubation time (e.g. nuclei appearance). Indeed, such a result is obtained: data is reported in Figure 4.9. Here, is observable how the very same kinetics of the crystal formation and evolution is again maintained (see reported times on each frame).

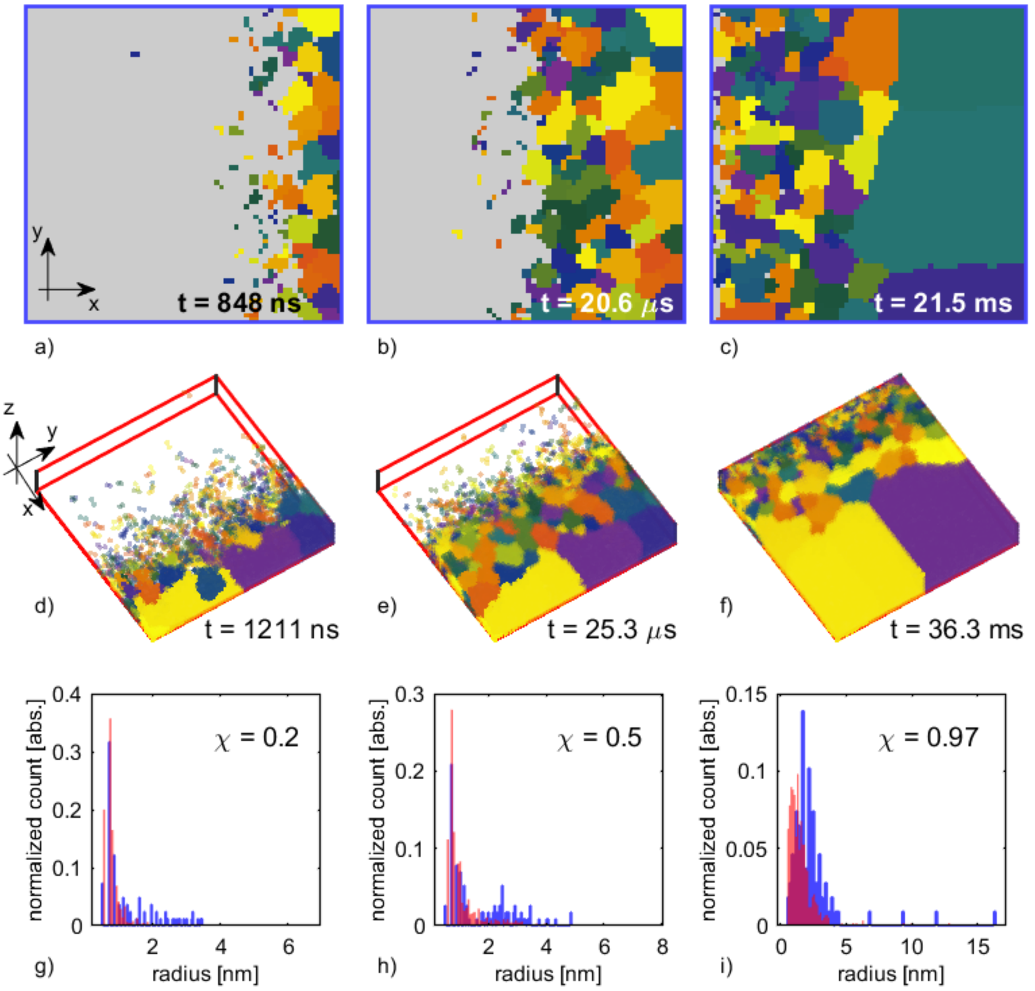


Figure 4.9. Comparison of the simulation results for a 2D and 3D GCA models, applying a temperature gradient between the left (500 K) and right (700 K) boundaries. The experiment and figure follow otherwise the same steps described in Figure 4.8.

A quantitative analysis of the grain size (see histograms in Figure 4.9(g-i)) shows once more the same difference in the grain size distribution previously observed (growing as the crystal fraction increases), which sees the 2D model having a broader size distribution slightly shifted towards larger grains.

The results obtained confirm that the proposed 3D GCA model reproduces the crystal growth kinetics of the 2D model, with minor variations intrinsically due to the model stochasticity, but also attributed to the larger surface area present in the 3D model with respect to the 2D model. However, the grain size distribution slightly differs from the 2D case, as it seems that the 3D model outputs a larger number of small-sized grains. In addition to the surface variation, this is possibly due to the need for revision of the equations driving the neighbourhood influence on the rates. An alternative explanation finds the culprit on the free parameters of the GCA model (above all the surface energy), which may need fine-tuning to correctly apply for the 3D case. Nevertheless, since the kinetics remains largely similar, the above data prove that the proposed model can be used to predict the crystal fraction variation as a function of temperature in a wider number of cases.

4.1.2.3. Computational requirement of the GCA 3D model

The computational requirement of the 3D GCA model is here compared to the 2D model. A series of annealing simulations has been carried out, from 300 K to 850 K, with a 50 K step. A fixed number of monomers is used for both the 2D model (with a 64 × 64 monomers area) and the 3D model (with an 8 × 8 × 64 monomers block). The calculation time vs temperature data is reported in Figure 4.10.

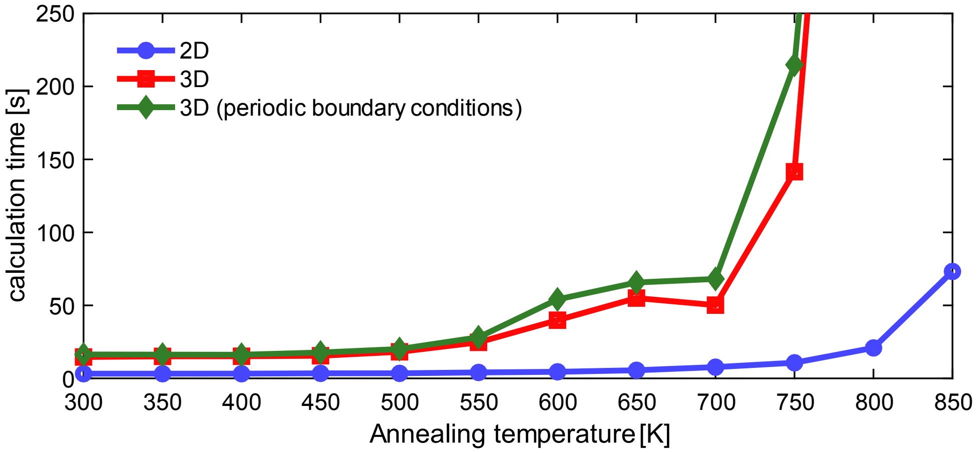


Figure 4.10. GCA calculation time, comparison between 2D and 3D cases as a function of temperature. All simulations account for the same number of monomers: 64 × 64 monomers for the 2D case; 8 × 8 × 64 for the 3D case. For the GCA model, datapoints above 800 K are not present; the depicted line trend is based on an estimated time of 10^3 seconds for 800 K.

The data evidence an exponential increase of the calculation time with the model temperature. The reason is found in the competing nucleation/growth and dissociation mechanisms: as the temperature approaches melting point, a larger share of the nucleation and growth events is immediately counteracted by dissociation. This mechanism would become predominant as temperature increases even further, not resulting in any meaningful variation to the crystal fraction (though changing the grain size distribution as aforementioned). Figure 4.10 also compares the 3D model output with or without periodic boundary conditions. The results confirm the above assumption, with the surfaces facilitating the stabilization of small nuclei (as the total number of neighbours decreases at the surfaces, see Eq. 3.9f)[30].

A set of test simulations at increasing temperature have been carried out to track the count of events per simulation time. The data is reported in Figure 4.11, in which it is evident how the count of dissociation events quickly overcome nucleation and growth as temperature increases. Additionally, dissociation seems to be strongly correlated with the nucleation events (seemingly finding dissociation counts to be twice the nucleation counts; an additional fraction of the dissociation events also contrasts the growth mechanism). Henceforth, the plots quantitatively corroborate the previous explanation, pointing to the aforementioned mechanism as the source of the increased calculation time.

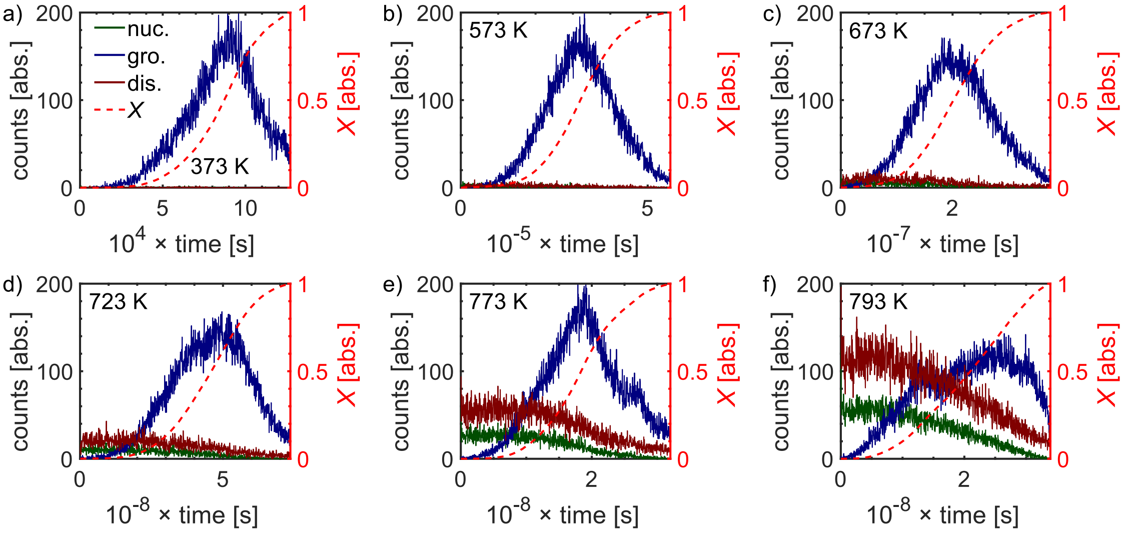


Figure 4.11. Number of events (nucleation, growth and dissociation) per simulation time, during a simulated annealing process of an initially amorphous domain at different temperatures. Each of the plots represents a separate annealing process at constant temperature (temperature is reported in plot). The crystal fraction plot is also plotted (red, dashed) with its relative axis on the right of each plot. Legend reported in (a).

---

[30] This effect does not arise by design and was not influential in the previous GCA model. In future, this peculiarity might require further analysis, to either vary the mathematical expression or to adjust the surface energy effect of the surfaces.

A simple method to remove the issue whereby at high temperatures nucleation and growth are immediately ‘undone’ by dissociation, would be not to allow nucleation and growth events above a certain threshold temperature, although the kinetics of the melting process then becomes less physically realistic. A partial solution sees the conditional classification of the nucleation mechanism above or below melting point (or any arbitrary threshold temperature). If nucleation is observed above such a threshold, the algorithm would then not enact the corresponding algorithmic routines, but still increases the GCA time as if the three mechanisms took place. This method is less invasive on the melting kinetics, as it does not prevent growth and dissociation to take place, but still requires the continuous reiteration of the ITS (or specifically, of the bisection search method), which remains a demanding task. For the PPCM device, which can easily require the calculation over 20×10^7 monomers, the continuous formation and dissolution of unstable nuclei becomes a prominent issue. Thus, a suitable solution can not be easily found within the boundary of the GCA model definition, and a different strategy is needed.

4.1.3. 3D Cellular Automata based on the Classical Nucleation-Growth

The CNG model states the kinetics of the phase-change process via three coupled equations, modelling homogeneous nucleation, heterogeneous nucleation, and growth rates (see Eq. 3.7(a-g)). The solutions of these can be calculated directly providing only temperature as a driving force, besides a handful of parameters; and thus, such a model is characterized by a very low computational demand. It is possible to further manipulate the CNG equation set to be used within a cellular automata framework, yielding a binary-encoded phase matrix as in the GCA model. This section illustrates the steps taken for such an approach, and shows the results of the 3D CNG cellular automata model, discussing its calculation output and comparing it with the previously reported GCA. As it will be demonstrated, the CNG model is indeed a suitable option to replace the GCA for the calculation over large PC material volumes, such as in a PPCM device.

4.1.3.1. From rates to probability: definition of the CNG cellular automata core equation set

A simple method to introduce the CNG equation set within a binary cellular automata framework consists of the calculation of the time-step (τ) dependent probability $P_{event}(\tau)$ for a certain mechanism to take place. Here, the notation event refers to either nucleation or growth, as used for the GCA model.

It is possible to write the probability terms as follows [225]:

$$P_{nuc,hom}(\tau) = I_{hom}^{ss} \times \tau \times v_m \qquad \text{Eq. 4.1a}$$

$$P_{nuc,het}(\tau) = I_{het}^{ss} \times \tau \times a_m \qquad \text{Eq. 4.1b}$$

$$P_{gro}(\tau) = I_{gro}^{ss} \times \tau \times l_m \qquad \text{Eq. 4.1c}$$

Here, the terms v_m, a_m, and l_m refers to monomer volume, surface, and diameter respectively. I_{hom}^{ss}, I_{het}^{ss} and I_{gro}^{ss} are instead defined in eq. Eq. 3.7(a-c).

Now, from this new set of equations, it is possible to obtain the probability of each event to take place (on each monomer / for a given time-step). At every time-step τ_i, the algorithm calculates each of the portrayed probabilities P_i, and then generates a set of uniformly distributed random numbers $\rho_i = (0,1)$. The algorithm then performs the routines associated with the events if $P_i > \rho_i$, i.e. nucleus formation (having a unique phase φ_i) or crystal adjoining. Both events see the amendment of the binary phase variable from amorphous (0) to crystalline (1).

Of course, the calculation is limited to the monomer sites that can perform such an event. For homogeneous nucleation, the sites are limited using the crystal phase as a vector mask, to get rid of the already crystallized sites. For heterogeneous nucleation, an additional vector mask limits the calculation to the surficial sites (the surfaces are easily determined through the Pointer matrix, as the row addresses containing any nil-pointer value). As for the growth mechanism, an additional mask is created (and easily updated after each event), which tracks the amorphous sites’ neighbouring crystallized monomers.

4.1.3.2. Dissociation rate and probability

The CNG model does not natively support the dissociation event, and therefore, any crystal monomer can not return to the amorphous phase. To overcome this limitation, a bespoke[31] dissociation rate function is here built, with the only scope to mimic the amorphization kinetics of the GCA model at high temperatures. The dissociation rate is devised as the advancement of the melt front (analogously to the growth rate), modelled as an error function of temperature, with a set velocity of 0.5 m/s at melting temperature, and a standard deviation σ value of 4. The dissociation velocity (and its corresponding probability) analytical expressions reported below:

$$\mathrm{u}_{dis} = \mathrm{u}_{dis}^{\max} \cdot \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{T - T_m - F_{dis}\left(\mathrm{n}_i^{\varphi}\right)}{\sigma\sqrt{2}}\right)\right) \qquad \text{Eq. 4.2a}$$

$$F_{dis}\left(\mathrm{n}_i^{\varphi}\right) = \mathrm{k_u} - \mathrm{k_n}\frac{\sum_j w_j \mathrm{n}_j - \sum_i w_i \mathrm{n}_i^{\varphi}}{\sum_j w_j \mathrm{n}_j} \qquad \text{Eq. 4.2b}$$

$$P_{dis}(\tau) = \mathrm{u}_{dis}(T) \times \tau \times l_m \qquad \text{Eq. 4.2c}$$

Here, F_{dis} is a factor which sets the effective temperature for the error function. The parameters $\mathrm{k_u}$ [K] and $\mathrm{k_n}$ [K] are numerically determined. The first one sets the desired dissociation velocity (0.5 m/s) at melting temperature; the second one subtracts a $\mathrm{k_n}$ value, rescaled by the (weighted) number of neighbours $\left(\mathrm{n}_i - \mathrm{n}_i^{\varphi}\right)$ of different orientation (n_i^{φ} being the neighbours with same phase and orientation as the dissociating monomer; n the total number of neighbours; w the inverse distance normalizing factor, as previously discussed in the GCA sections). $\mathrm{k_n}$ corresponds to 2σ in the current implementation. $\mathrm{u}_{dis}^{\max}$ is set as the maximum growth velocity.

The dissociation mechanism is by choice evaluated only at $T > \left(T_m - F_{dis}(\mathrm{n}_i)\right)$, so to avoid unrealistic amorphization at temperatures below melting point.

---

[31] The proposed function does not carry any physical interpretation, and it has been derived from the progressive refinement of its results with the view to match the generic GCA output approaching T_{melt}.

4.1.3.3. Algorithmic steps of the CNG cellular automata

The simulation time t of the CNG cellular automata is split in further time-steps τ. The reason for this choice is to maintain each of the calculated probabilities ($P_{nuc,hom}$, $P_{nuc,het}$, P_{gro}, and P_{dis}) values below unity[32], and thus avoiding calculation artefacts (i.e. time-step dependent output).

The time-step generation routine limits the time-step value to the one which yields a monomer maximum growth probability of 0.5; above the temperature at which the peak growth velocity is reached, such minimum time-step is maintained (273 ps in the current implementation)[33]. Of course, the time-step is also adjusted according to the phase configuration of the investigated domain, so to avoid useless iterations in case of no possible event (e.g. a completely crystallized domain below dissociation threshold would always have each probability below 0, and thus a single iteration is performed with $\tau = t$). The time-step is recalculated every arbitrary number of iterations $i \geq 1$.

After obtaining the time-step τ, the algorithm has all the necessary elements to calculate the events' probability and perform the following actions (for nucleation, growth, and dissociation, in such order).

i. For each event, the algorithm observes if such event is available (i.e. nucleation and growth in a fully crystal domain is prevented, as is dissociation in a fully amorphous domain or with a maximum temperature $T_{max} > \left(T_m - F_{dis}(\mathrm{n}_i)\right)$).
ii. If the outcome is positive, the algorithm selects from the phase array all locations upon which such event is available, and calculates the related probability P_{event}.
iii. It then calculates an equally-sized array R of uniformly distributed random values between 0 and 1.
iv. It calculates a coordinate array $C = P_{event} > R$.

---

[32] In the GCA model, avoiding a probability value above 1 is not necessary, as the time advancement is determined by the event probability. The proposed method instead sets no correlation between probability and time-step, and by allowing such probability to excede unity, the algorithm may detect a single event, whereas the rate could potentially lead to multiple events.

[33] The calculation of a more appropriate time-step at high temperatures would require the evaluation of each combination of phase and temperature across the investigated domain. It is instead more effective to reiterate the CNG calculation for shorter-than-necessary time-steps.

v. Where such C value is true, the algorithm performs such event, in the following manner:
    a. Nucleation sees the formation of a nucleus, i.e. update of the phase to $r = 1$.
    b. Growth sees the adjoining of a monomer to a neighbouring crystal grain, i.e. update of the phase to $r = 1$ and orientation to $\varphi = \rho$ with ρ being a random value between the orientation values of the surrounding crystal grains
    c. Dissociation sees the melting of a monomer, i.e. update of the phase to $r = 0$ and orientation to $\varphi = \rho$ (with ρ being a random value $0 < \rho < 1$).

The algorithm reiterates the sequence i-v, updating the simulation time after each time-step, until the calculated time corresponds to the input simulation time.

4.1.3.4. CNG cellular automata: simulation results

This section illustrates how the crystallization results of the CNG model described above has very similar characteristics (crystallization time and grain size distribution) to the GCA model. The parameters adopted are the same used for the simulation carried out with the GCA model, except a 1.5 multiplicative factor for the surface energy, which has been found to better fit the GCA behaviour (although unsupported by a clear theoretical explanation). Table 4.1, listing the CNG model still unreported parameter values, is shown below. These values were initially taken from Fan et al. [225], and adjusted to compensate for the variation of the other parameters (e.g. enthalpy of fusion) so to maintain the peak values of the calculated rates reported by Fan. These were then further varied to mimic the kinetics of the GCA model. To further illustrate the specifically implemented CNG model, the probabilities of homogeneous nucleation, heterogeneous nucleation, growth, and dissociation are plotted as a function of temperature in Figure 4.12 (per second per monomer).

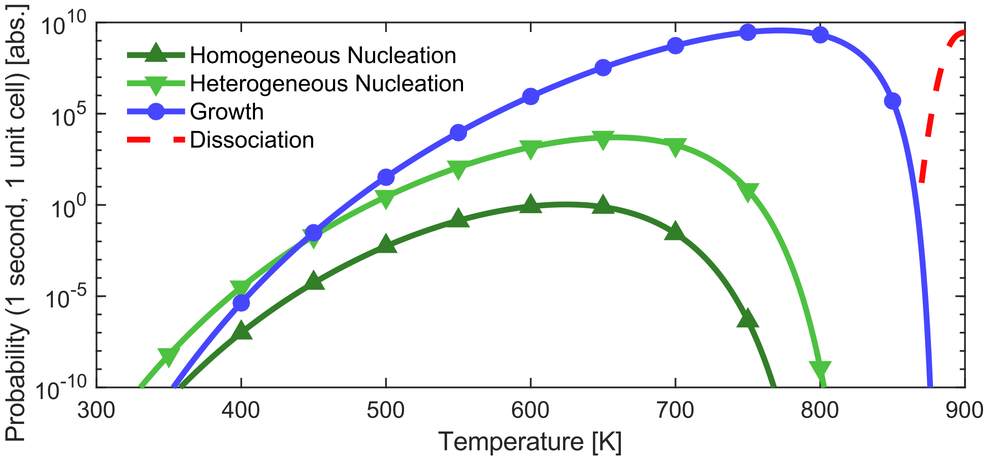


Figure 4.12. CNG model, events probability (per second – unit cell), as calculated with Eq. 4.1(a-c), Eq. 4.2(a-c) and the simulation parameters illustrated in Table 3.1 and Table 4.1 (the dissociation rate is here averaged over the possible neighbourhood configurations.).

Table 4.1. Parameters adopted for the CNG phase transition simulation.

Symbol	Parameter	Value
$\mathrm{k}_{\mathrm{nuc}}^{\mathrm{hom}}$	Homogeneous nucleation, kinetic parameter	1.2076×10^{49} [(m·(*monomer unit*))^{3}/s]
$\mathrm{k}_{\mathrm{nuc}}^{\mathrm{het}}$	Heterogeneous nucleation, kinetic parameter	7.4643×10^{40} [(m·(*monomer unit*))^{2}/s]
$\mathrm{k}_{\mathrm{gro}}$	Growth, kinetic parameter	2.0205×10^{8} [m·(*monomer unit)*/s]
$E_{a,\mathrm{gro}}$	Activation energy (growth mechanism only)	2.8 eV
$f(\vartheta)$	Contact angle	0.6 [abs.]

The first simulation example, reported in Figure 4.13, reports the phase plot of the crystallization obtained at a constant temperature of 407 K, starting from a completely amorphous domain. As expected, the crystallization starts with the random formation of small nuclei, to which little growth follows, and eventually bring to the near-complete crystallization after 88610 s. Importantly, similar kinetics and grain size distribution is observed in the 2D and 3D cases, confirming once more the reliability of the 3D framework. Compared to the result obtained with the GCA model (see Figure 4.8), it is possible to observe a rather similar grain size narrow distribution, and the monitored annealing times are again quite close. As for the grain size, the CNG here outputs smaller grains, showing a more marked nucleation-dominated growth.

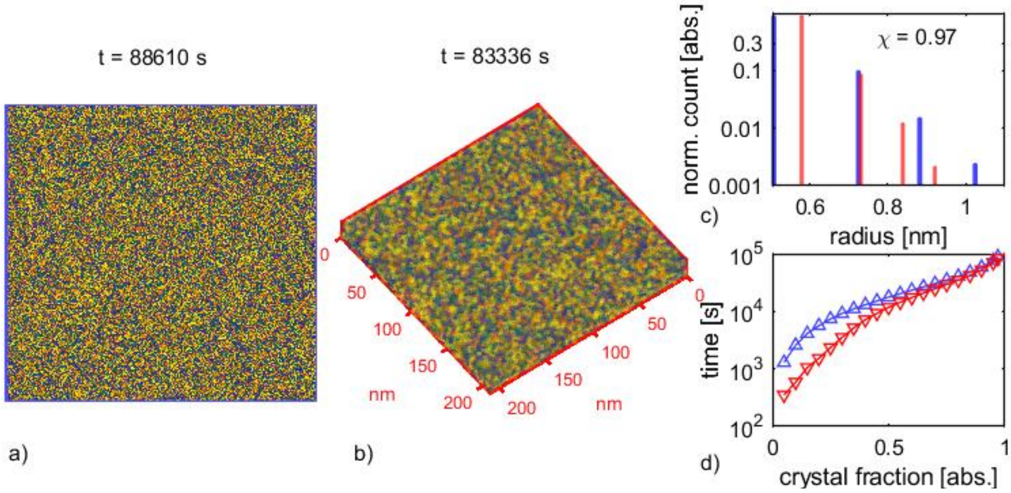


Figure 4.13. CNG model, annealing result at a constant temperature of 407 K, for the 2D (a) and 3D (b) cases. c) Comparison of the grain size distributions, for the 2D (blue) and 3D (red) cases. Note the log scale on the vertical axis. d) Comparison of the crystal fraction variation as a function of time, same colour code as for (c).

The second simulation explores and highlights the different behaviours of the 2D and 3D cases at high temperatures, due to the strong influence of the heterogeneous nucleation. Both cases see a (linear) temperature gradient, with the left and right domain boundaries set at 500 K and 700 K respectively. The results are displayed in Figure 4.14. In the 2D case, the nuclei originated at high temperatures expand until the middle of the simulated region, leading to 2 rather large nuclei occupying most of the simulated space. In the 3D case, such effect is still present but moderated by the co-presence of many initial nuclei, which compete with each other during the growth phase, thus i) growing to much smaller sizes, and ii) initiating the crystallization at an earlier time. Both effects are evident in the two reported charts (see Figure 4.14(c-d))

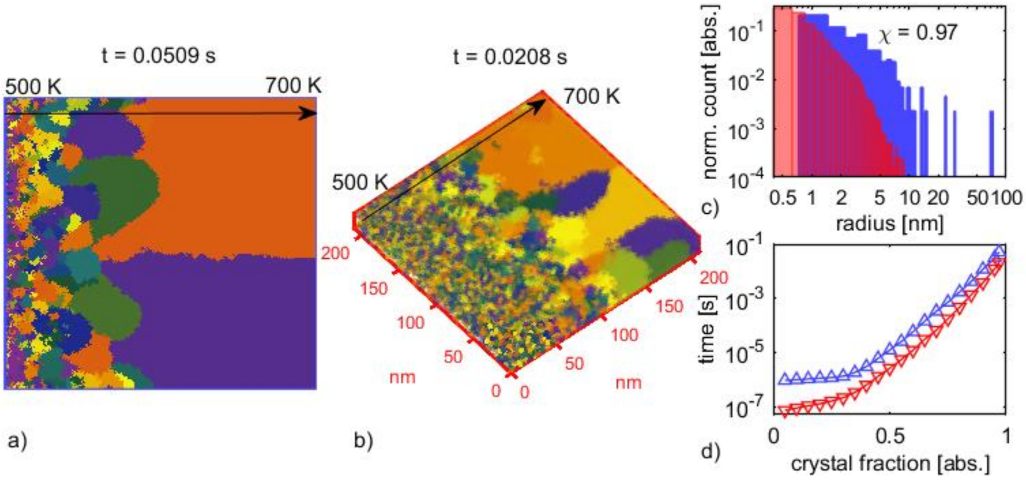


Figure 4.14. CNG model, annealing result with a linear increasing temperature from the left boundary (500 K) to the right boundary (700 K), for the 2D (a) and 3D (b) cases. The temperature (linear) gradient is visually represented as a black arrow in (a). c) Comparison of the grain size distributions, for the 2D (blue) and 3D (red) cases. d) Comparison of the crystal fraction variation as a function of time (same colour code as in (c)).

It is clear from the above that similar results are generated by the 3D CNG and 3D GCA methods, both in terms of dynamics and kinetics. The following section elaborates on this aspect, quantitatively underlining any differences between the results from each method, but also demonstrating how the CNG model is indeed capable of capturing the fundamental traits of the more complex GCA model.

4.1.3.5. Comparison of 3D CNG and GCA models

A set of different simulations is carried out using a variable combination of geometries and temperature gradients, similar to what has been illustrated in Figure 4.9 with the GCA model. The aim is here to observe the CNG model behaviour under conditions non-reproducible with a 2D model, to identify any deviation from the GCA model, and test the realism of the simulation output.

The first case (see Figure 4.15) is the annealing of a 256 × 256 × 10 [monomers] rectangular parallelepiped, with a temperature set at 500 K and 700 K on the left and right boundaries respectively. The result shows how a close match with the GCA model is obtained, both for the simulation time (to complete annealing) and, qualitatively, on the grain size distribution (shown in the histogram). The latter shows a mismatch at both tails of the distribution, which will be discussed in Section 4.1.3.6.

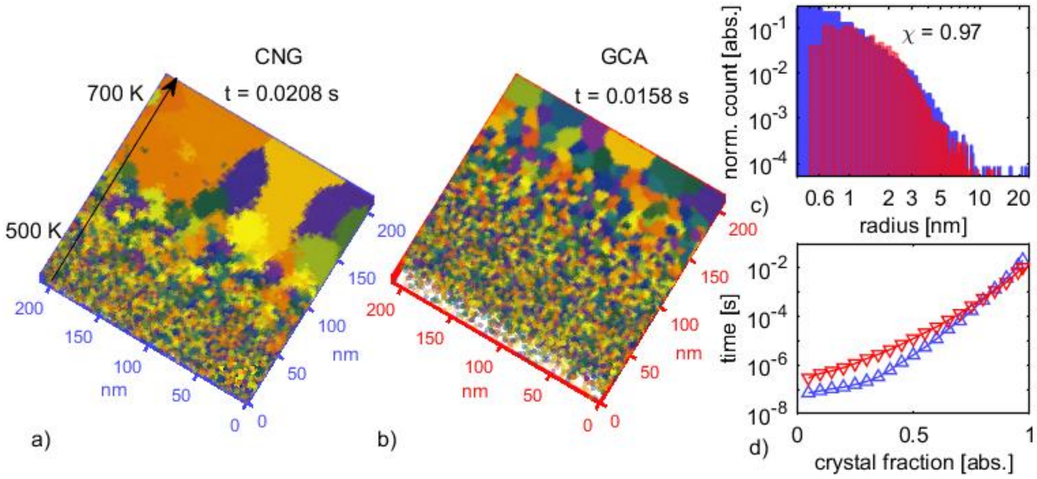


Figure 4.15. GNG and GCA simulation comparison, on a (256 × 256 × 64 monomers) thin layer, with a linearly increasing temperature from the left xz boundary (500K) to the right xz boundary (700 K). Temperature is represented in (a) as a black arrow. a-b) CNG and GCA simulation outputs respectively, at $\chi = 0.97$. c) Grain size distribution, for (blue) CNG and (red) GCA outputs. d) Crystallization time as a function of the crystal fraction, for (blue) CNG and (red) GCA models.

A second simulation set is carried out on a highly asymmetric geometry, composed by the intersection of a cube, a sphere and a cone (see Figure 4.16). The temperature is set to linearly increase along the z axis, from 473 K to 573 K. Again, it is possible to observe how the crystallization kinetics is conserved, with a qualitatively comparable grain size found across the model (larger on top, much smaller on the bottom). The histogram here highlights a reasonable quantitative match of both the grain average size and size distribution for the two methods, although evidencing the larger grain maximum size. This simulation also highlights the flexibility of the 3D environment, adapted to perform the calculation onto an otherwise challenging volume to discretize for the 2D representation.

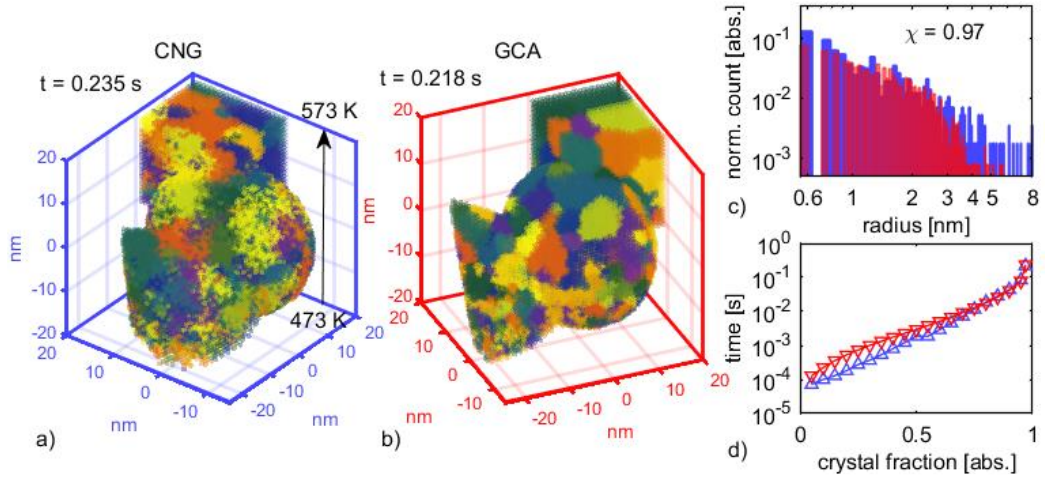


Figure 4.16. GNG and GCA simulation comparison, on a highly asymmetric geometry, combined with a z-dependent linear temperature gradient (shown in (a) as a black arrow). a-b) CNG and GCA simulation outputs respectively, at $\chi = 0.97$. c) Grain size distribution, for (blue) CNG and (red) GCA outputs. d) Crystallization time as a function of the crystal fraction, for (blue) CNG and (red) GCA models.

One last simulation is carried out on a 64 × 64 × 10 [monomers] rectangular parallelepiped, and applies an arbitrary and asymmetric temperature distribution, originating from a de-centred point heat source (of temperature T = 573 K), from which the temperature decays linearly along the x and y axis, and with a quadratic dependency along the z axis. The simulation is here carried out until reaching a crystal fraction value of $\chi = 0.3$. Once more, both the GCA, and the CNG reach the aimed χ value at the same time, conserving the same kinetics (see crystal fraction plot). The crystallized portion of the simulated volume reaches the same height in both the GCA and CNG plots. A reasonably comparable grain size distribution is also obtained (see histogram).

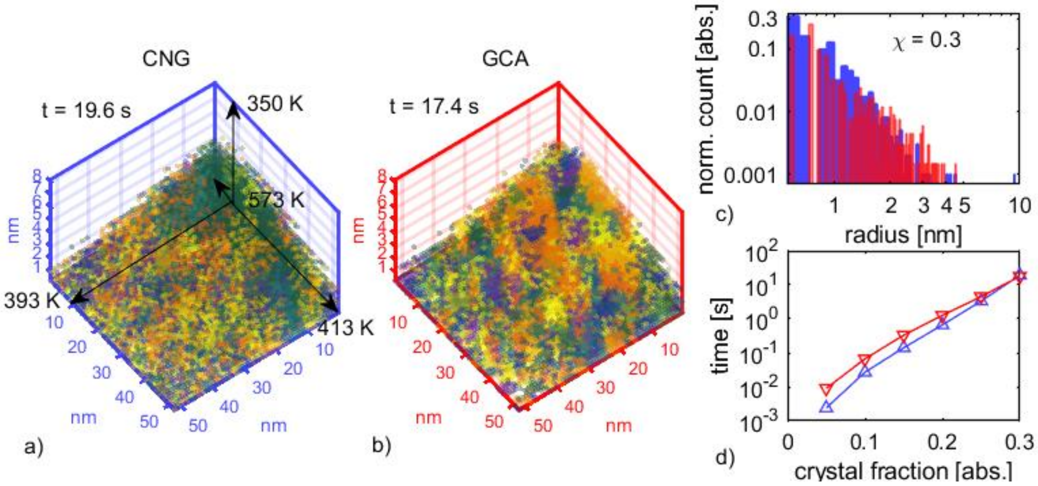


Figure 4.17. GNG and GCA simulation comparison, on a thin layer with a de-centred 0D heat source. Temperature boundaries and gradients are represented in (a) as black arrows. a-b) CNG and GCA simulation outputs at $\chi = 0.3$. c) Grain size distribution, for (blue) CNG and (red) GCA outputs. d) Crystallization time as a function of the crystal fraction, for (blue) CNG and (red) GCA models.

4.1.3.6. Discussion and conclusion

The scope of Section 4.1.3 is to illustrate the key concepts, implementation, and results of the developed CNG model are illustrated. The physics behind the CNG model is largerly unchanged from the model statement found in literature [92], [209]–[211]. However, the fundamental innovation is found in its implementation, which employs a probabilistic picture of the nucleation and growth rates [213] in combination with the cellular automata numerical technique (developed to improve the GCA model and shown in Section 4.1.1). This development was carried out to embed a 3-dimensional and flexible spatial framework to the CNG model, which was previously only partially fulfilled by Peng et al. [213]. In their work, Peng et al. adopt a square lattice pseudo-3D representation of the modelled volume, in which the depth direction is not directly tackled, but emulated via optimsation of a dimensionless free parameter, which modulates the cystal growth velocity. Their strategy is rather efficient to model the crystallization kinetics of a rewritable disc device, in which i) the laser illumination (and consequential optical heating) follows predictable patterns (in fact, Peng's manuscript details the equations driving the thermal spatial-temporal evolution induced by the impinging radiation), and ii) the change of the surficial phase-state of the PCM layer heralds the assumed crystallisation kinetics along the depth direction. However, as we have here demonstrated, detangling the CNG model from the device specific features allows to reach the aimed flexibility of implementation for a much wider set of devices and operative conditions.

As for the specifics of the CNG model behaviour, the above simulation results evidence how the GCA and CNG models develop a comparable kinetics of crystallization. The grain size distributions show qualitative similarities (e.g. similar median size value, and similarly shaped distribution), but the quantitative analysis evidences the following differences, more pronounced at higher temperatures. First, the CNG allows small nuclei (down to one monomer) to form even close to melting point (which would be a non-realistic behaviour). The reason lies in the nucleation mechanism integrated within the cellular automata framework, which sees the formation of a single-monomer nucleus which would eventually grow to super-critical size, and thus only approximating the dependency of the stable nuclei size with temperature. This effect could be tackled by coding suitable routines modelling the nucleation on multiple sites simultaneously, better representing the correctly-sized nucleus[34] derivable from the original model definition. Secondly, the CNG model usually shows larger grains, which may be fixed by using a better approximating set of material parameters. Moreover, in the CNG output, the smoother grain borders found with the GCA model are not obtained. Both effects are influenced by the implemented growth mechanism, which sees no rate compensation based on the neighbourhood distribution, a characteristic embedded only in the GCA model. Possible corrections (beyond the eventual fine-tuning of the simulation parameters) can be envisaged introducing such mechanism in the rate equations.

Nevertheless, for the simulation of large PC material volumes in photonic or optical devices, the presented aberrations lead to negligible or no calculation artefacts. Standing by a classical perspective of the light propagation, light-matter interaction takes place with a scale dictated by the impinging beam wavelength λ. Due to the diffraction limit, any feature well-below $\lambda/2$ becomes practically non-influential, and the effective medium theory (which is here used to feed the phase data to the FEM solver) instead provides an averaged optical property of the PC material, negating any substantial influence of the differences between GCA and CNG outputs.

---

[34] One of the main and standard features of the CNG model is the definition of the correctly-sized stable nucleus. Since (in this cellular automata algorithm implementation) there is no mechanism preventing the growth of any nucleus below melting point, the characteristic CNG reliability for the nuclei stability and size is maintained, although in an approximate vision which sees its formation as a single monomer, which eventually grows to super-critical size.

Concluding, the CNG model is capable of capturing the essential traits of a physically-realistic GCA model, at a fraction of the computational cost[35], thus allowing fast calculation for large 3D volumes encountered in photonic devices. Differences in the crystal grain distribution are observed, but these won't affect the material properties calculated from the resulting phase matrix. A bespoke dissociation mechanism is introduced to model melt-quench processes. Thus, the CNG model here illustrated is the method of choice to model the phase transition in the self-consistent framework (discussed in the next section) required for the simulation of photonic phase-change devices.

4.2. Integration of FEM software and phase-change modelling

COMSOL is a very useful tool for the evaluation of the tensorial fields as a function of time. Nevertheless, as reported in Section 3.2, it is not natively capable to adopt the phase-change model of choice. Moreover, COMSOL can not easily share the same variables workspace with a different software environment during the calculation, but for a few explicit exceptions. Thus, a solution is needed to merge the two software environments: COMSOL, and the phase-transition algorithm, here implemented in MATLAB language. This approach will need to solve also a second problem: the systematic evaluation as a function of (simulated) time of the FEM results and phase-transition results, without the requirement of constant user-provided input.

This section illustrates the methodologies here adopted to build the full simulation framework, capable of automatically carrying out the self-consistent calculation of a realistic phase-transition mechanism without algorithm fine-tuning or user input requirement.

4.2.1. COMSOL framework

Any COMSOL model for the investigation of a PPCM device needs to contain the following elements: Geometry, Meshes, Physics, Materials, Study, and Solvers.

---

[35] For instance, a 256×256×10 3D GCA model (6.5×10^5 monomers) may take several days to complete the annealing at $T > 800$ K; the same result can be obtained in seconds to minutes with the 3D CNG model. Spanning from room temperature to melting point, a time reduction factor of 10 to 10^6 is estimated, assuming a monomer count of 10^6.

Below follows a detailed description of the practical solutions adopted within the models used in this work.

4.2.1.1. Interfaces

Since the interest is focused on the numerical emulation of an optical feature of typical size s, such as $\lambda/10 > s > 10 \cdot \lambda$, the RF interface in COMSOL is the most suitable module to adopt. Specifically, the *Electromagnetic Waves, Frequency Domain* (emw) Physics interface is selected. From a merely technical perspective, this module also contains a subset of the features contained in the Heat Transfer module. Thus, the RF module allows one to perform the needed time-resolved calculation of the heat diffusion PDEs.

4.2.1.2. Geometry

The Geometry of the 3D PPCM model consists of a ridge (or rib) waveguide, extruded over a certain length, and includes a portion of the substrate and a portion of the cladding[36]. The unit cell resides at the centre of the waveguide, and roughly at the centre of the model.

The model characteristic sizes are $x = 4\lambda$, $y = 4\lambda$, and $z \leq 3\lambda$ (the latter depending also on the unit cell length), with λ being the wavelength of choice (usually, $1550\ nm$).

Considering the volume shape in the xy plane, a rectangle is expected. To reduce to the essential the volume of substrate and cladding required to be included in the simulation, the extrusion of two semi-ellipses (also on the xy plane) is employed to trim the rectangle corners. The two semi-ellipses meet orthogonally at the waveguide xz plane, and their semi-axes are adjusted to fit within the rectangular area. The two xy surfaces of the model volume are instead left perpendicular to the z axis. The picture of a representative PPCM model Geometry is provided in Figure 4.18(a).

---

[36] For sake of clarity, the cartesian axis system considers the waveguide propagation axis to be z, and the (second) in-plane and out-of-plane axis to be x and y respectively.

4.2.1.3. Mesh

The meshing operation is carried out as described in Chapter 3. The model meshing is usually refined in three blocks: the region containing the element of interest (the unit cell), and the contiguous regions containing input and output branches of the waveguide. The unit cell mesh size is usually controlled via triangular meshing of its surface, followed by tetragonal meshing of the included volume.

The waveguide is instead meshed via triangular meshing on an external xy surface, followed by swept mesh technique across the model towards the opposite xy plane, again refining the mesh size in proximity of the unit cell. Since swept mesh requires the continuity of the domain shape on the extruded xy plane along the z axis, in case such condition is not met, an enclosing region (or set of enclosing regions) is programmed so as to allow swept mesh where possible; eventually, conjoining volume slices are meshed via tetragonal elements.

Substrate and cladding meshing is controlled to obtain a slow increase of the tetragonal mesh size irradiating from the unit cell, which also results in a suitably finer mesh region close to the unit cell ($\pm$ 100 nm).

A typical meshing result is displayed in Figure 4.18(b).

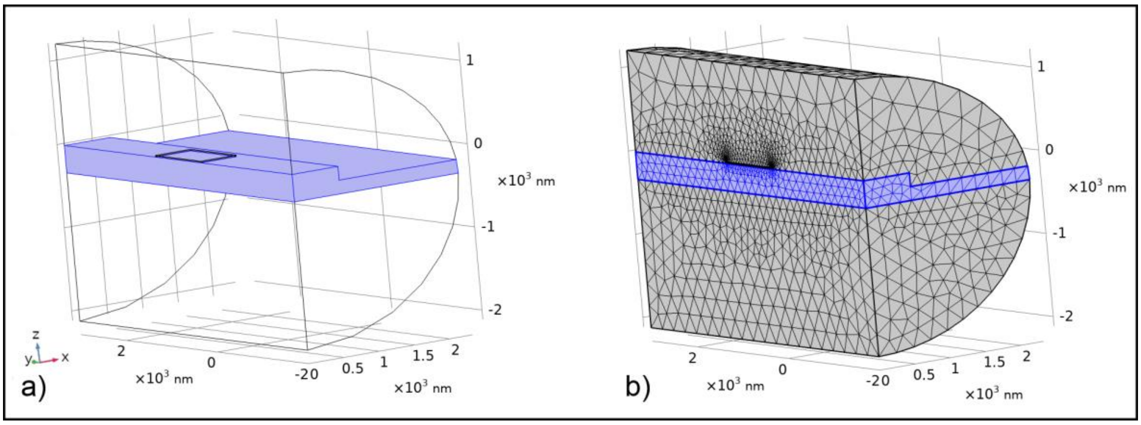


Figure 4.18. Typical PPCM device COMSOL model. To provide a visual aid, the core of the waveguide is highlighted in blue. a) Geometry. b) Meshing results

4.2.1.4. Physics and features

For the *emw* interface, the model xy external surfaces are indicated as user-defined *Ports*, where the propagating mode is excited and absorbed. The remaining external surfaces are instead assigned to the 2nd order Scattering Boundary Condition.

The *ht* interface considers all external surfaces as fixed temperature, since these are distant enough from the unit cell, and therefore do not induce any meaningful unphysical behaviour (such as rapid temperature decrease). TBR layers (previously described in Section 2.2) are placed along the surfaces between the waveguide domain and the substrate domain, and along the GST surfaces as well. The Heat Source feature is assigned to the optically absorptive domains, generally within the unit cell, possibly within its cap layer (ITO has indeed a positive extinction coefficient in the NIR spectral region, see e.g. [172]), and on any metal domain if present. The expression used within the Heat Source feature is emw.Qh (Joule losses, [W/m^3])[37], determined as the integral of the optical power P loss:

$$Q_h = \int_z \frac{P_z}{P} = \int_z \exp\left(-\frac{4\pi\kappa z}{n_z\lambda}\right) \qquad \text{Eq. 4.3}$$

The model described so far consists of the full 3D geometry. Nevertheless, since a clear yz symmetry plane is identifiable along the centre of the waveguide[38], the model is further refined to exclude one of the two halves, so as to half also the number of *Degrees Of Freedom* (DOF), and hence the calculation time (see Appendix I for more details). The *emw* places a PEC boundary condition on such a plane, allowing only the propagation of TE modes. The *ht* interface considers instead a *symmetry boundary* condition, which acts as a thermal insulator (i.e. negating heat flux across such surface). The result of this size-reduction operation is also visible in Figure 4.18.

---

[37] In COMSOL, exchange of variables between Physics is not generally allowed, unless explicitly set up via Multiphysics branch, an additional COMSOL module specifically built to perform such task. In this case, the use of the output of the *emw* interface (the optical loss) is thus possible within the ht interface (as volumetric heat source).

[38] For the specific purpose of model volume reduction, the symmetry element must be found on the variable field. This usually also correspond to the geometry symmetry element, but any geometrical symmetry element does not constitute a sufficient condition.

4.2.1.5. COMSOL Study setup

To evaluate the optical propagation mode, the *Boundary Mode Analysis* (B.M.A.) Study is used, once for each of the two ports (in case of a simplified 2D model, the *Mode ana*lysis (M.A.) Study is instead employed). To find the desired TE mode, which for the PPCM device is usually TE_1, the study is set to search for modes with an effective mode index close to the waveguide core refractive index. Unless the investigated surface contains any other feature than the waveguide (e.g. metal surfaces, other dielectric structures which extruded could support propagating modes, etc.), the mode with the highest index corresponds to the desired one.

To emulate the mode propagation, a time-independent analysis is usually adopted, selecting the *Frequency Domain* (F.D.) Study. The B.M.A. solution needs to be provided as a *non-solved-for solution* within the F.D. Study set up (to assign the calculated effective mode index). This operation is sufficient to mimic probe operativity (i.e. stimulation via continuous, low-power laser), as it assumes an unperturbed stationary-state system.

The heat diffusion problem can be independently solved using *Time Dependent* (T.D.) Study. In this case, the *non-solved-for solution* must be set up as the previously determined solution of the F.D. Study (so to provide the aforementioned heat source value).

To evaluate the evolution of the system for long time durations under high-power excitation, where a variation of the optical properties is expected, an iterative approach can be adopted. Thus, to reiterate the process via separate studies, the F.D. Study must be set up with its previous solution as a *solved-for solution* (which also carries over the B.M.A. solution), and the T.D. Study as a *non-solved for solution*. Similarly, the T.D. Study needs to carry over its previous state (e.g. temperature distribution) setting its previous solution as a *solved-for solution*.

This method allows tight control over the output calculation, although it is prone to computational artefacts as the T.D. Study time-step must be kept relatively short (since the *emw* related variables remain constant during the time-step).

An alternative method uses the *Frequency Transient* (F.T.) Study, which combines the F.D. and T.D. capability, with a slightly increased computational cost. The F.T. Study is less prone to calculation artefacts, as COMSOL accommodates over (simulated) time the *emw* related variables, as well as the ht variables (thus avoiding deleterious discontinuities).

4.2.2. MATLAB framework

Since the phase-change process is emulated outside the COMSOL environment, an iterative approach is needed to emulate the model behaviour for long time durations. In this work, MATLAB is the essential environment to manage some of the otherwise user-driven operations carried out by COMSOL, to yield the phase transition calculation and link the input and output of PC model and COMSOL model. More specifically, a bespoke MATLAB algorithm performs the following, for each iteration:

i. calculation and application of the time-dependent and time-step dependent variables (e.g. optical power, heat of transition, etc.)
ii. FEM iteration (COMSOL model setup and evaluation)
iii. withdrawing and processing the desired data from COMSOL
iv. emulation of the phase transition
v. injection of the phase variable in COMSOL
vi. and driving the time-stepping and total simulation time evolution by determination of the time step

The simulation framework is further illustrated in Figure 4.19, with the provision of a data flow diagram, and an algorithm flowchart. As shown, the approach here pursued is the coding of independent objects, whose inner working is hidden to the outer layer, and whose higher-level management requires only the input of the required variables and the use of a minimal number of functions, namely a “run” function which advances a time-step, and an “initiate” function which creates the ad-hoc object instances. The following objects are used:

a. TIME, which drives the temporal sequence of all the remaining objects
b. TIME-STEP GENERATION, which determines the time-step size (more details in Section 4.2.2.3)
c. TIME-DEPENDENT INPUT, which provides the analysed phenomenon driving force (i.e. optical pulse profile, see Section 4.2.2.1)
d. PHASE-CHANGE MODEL, which determines the phase transition as described in the section above

e. COMSOL MODEL, which manages the FEM environment

Further details[39] are provided in the subsections below.

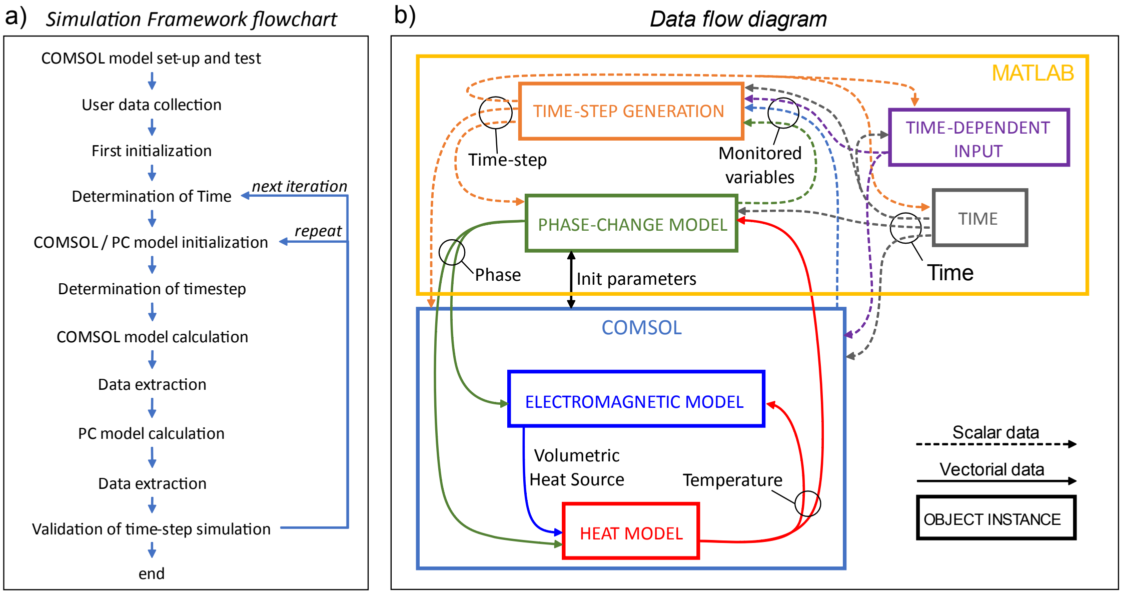

a)
Simulation Framework flowchart
COMSOL model set-up and test
User data collection
First initialization
Determination of Time
next iteration
COMSOL / PC model initialization
repeat
Determination of timestep
COMSOL model calculation
Data extraction
PC model calculation
Data extraction
Validation of time-step simulation
end
b)
Data flow diagram
MATLAB
TIME-STEP GENERATION
TIME-DEPENDENT INPUT
Time-step
Monitored variables
PHASE-CHANGE MODEL
TIME
Time
Phase
Init parameters
COMSOL
ELECTROMAGNETIC MODEL
Volumetric Heat Source
Temperature
HEAT MODEL
Scalar data
Vectorial data
OBJECT INSTANCE

Figure 4.19. Simulation framework schematics. a) Simplified logical flowchart. b) Data flow diagram (user data is not reported for sake of clarity). Legend on the bottom right.

4.2.2.1. Time-dependent user-provided variables

A time-dependent input is often needed, see e.g. write and erase operations in the PPCM device (in the form of the optical pump power). To allow full flexibility, but avoiding overburdening the simulation set up process, the algorithm builds a piecewise temporal profile of the user-provided input (i.e. the optical pulse). The piecewise curve is composed of two Heaviside functions, connected via a straight segment. This piecewise function can be uniquely defined by 5 variables (e.g. start time, end time, rise time, fall time, and peak value). An additional variable defines the slope for the connecting segments between the two Heaviside ramps.

4.2.2.2. Data exchange between COMSOL and PC models

The output of the phase-change model, i.e. the spatial phase map, is saved after each phase-change model iteration on a unique text file, as a spreadsheet dataset, where to each i coordinate set $(x, y, z)_i$ a single-valued phase information f_i is associated[40]. The COMSOL model is set up to withdraw the phase information on such file as needed. The phase information f_i is then processed internally within the COMSOL environment, via bespoke functions, to yield meaningful material parameters or physical behaviour.

---

[39] The code will be made available after debugging, on https://github.com/EmanueleGemo/CoMa-Interface.

The phase information (from the cellular automata models) is calculated on a raster defined by a regular cubic lattice, of lattice step $l = 0.82\,nm$. As aforementioned, for a $1\,\mu m$ PPCM unit cell (half model), this results in a quite large dataset, containing more than 11 million entries, each being a set of four double-precision numbers (x, y, z, f). The size of such dataset has a large computational cost for both the file exchange routine and the COMSOL evaluation time. To reduce the task complexity, an alternative approach implemented here for large PC material volumes consists of the calculation of the effective medium theory value for each model mesh. Each COMSOL mesh may contain a relatively large subset of those cubic lattice unit cells (e.g. a square pyramid mesh of size $5\,nm \times 5\,nm \times 5\,nm$ would contain 91 unit cells). Since COMSOL performs an analogous operation, averaging the phase information value of the unit cells to obtain the mesh information[41], it is convenient to pre-calculate such values to thus exchange a much reduced dataset. This operation is indeed performed after a one-time interrogation to the COMSOL model of the location of each mesh coordinates, and assignation of the respective lattice cells via distance minimization.

The coordinate set upon which the phase-change model is evaluated (thus yielding the phase information) corresponds to the same coordinate set from which the temperature (or any other needed variable) is extrapolated to drive the phase-change model. The determination of such a coordinate set for the cellular automata models strictly follows the one illustrated in Appendix II.

4.2.2.3. Adaptive time-step determination method

The determination of a suitable simulation time for the FEM model is a critical factor which needs to be explicitly addressed. Within the time-resolved FEM simulation, the time-stepping can be adjusted according to the mesh refinement and nature of the investigated phenomena (see Section 3.2.3). But assignment of the same time-step value to the total FEM simulation time (to then exchange the data with the phase-change model) is not feasible for the type of simulations carried out in this thesis, since it would result in excessively long computation times.

---

[40] At the current moment, an efficient method to inject time-stepped large datasets without use of external files has not been found. It might be possible to implement a routine which edits a COMSOL internal table; this alternative would indeed be faster, yet retaining most of the limitations illustrated in this sub-section.

[41] This is only an empirical observation. Alternatively, COMSOL may query random locations within the mesh, although a quick test seemed to exclude such possibility.

Both FEM and CNG solvers rely on the subdivision of the simulation time on time-steps, here identified as *inner time-steps* τ_i. Each FEM (or CNG) run simulation time $\tau_o = \sum \tau_i$ will be referred to as *outer time-step*. For sake of clarity, the total simulation time t would be $t = \sum \tau_o$. The outer time-step generation is the issue here addressed.

A simplistic and often effective method consists of the assignment of a single value τ_o, which splits the total simulation time in $(1 + \frac{t_{total}}{\tau_o})$ time-steps. This method is suitable when the single FEM calculation time is approximatively linear to the time step size, when certain events are expected at a predictable frequency, or also when such events would not have strong repercussion on the calculation efficiency/results reliability. In the case of the PPCM device, neither of these conditions are met. Therefore, a variable time-step is here desirable, as it drastically improves the solution quality and/or decreases the calculation time.

There is a rich literature focused on adaptive time-stepping methods (or, more properly, on the numerical solution of ODEs and PDEs), having different ranges of application, stability, and suitability. The simplest one, and the one used here in a modified version, is the Euler method, which predicts the solution via linear extrapolation, and entails a (global) error proportional to the time-step. The method monitors a predetermined selection of the models' input and output variables, obtained after each iteration. It then infers the following time-step output via extrapolation, by use of the Akima spline function [226]. The necessary/allowed range of variability of certain parameters can be estimated or known a priori: it is then possible to force the time-step duration so as to maintain the tracked variables time-difference within a user-provided range (delta), to avoid computational artefacts and maintain the desired output resolution. A graphical schematic of this method is also provided in Figure 4.20.

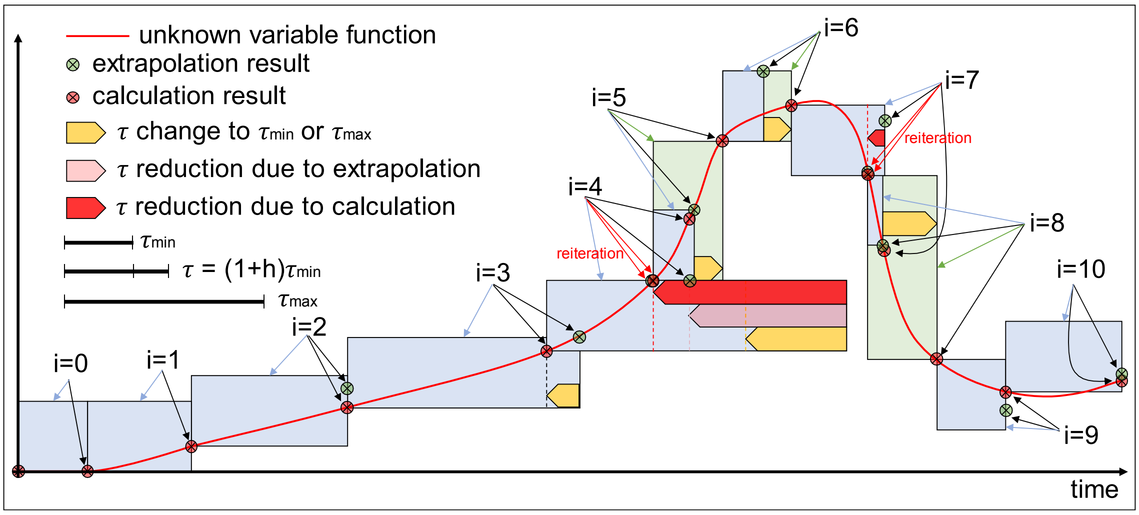

unknown variable function
extrapolation result
calculation result
τ change to τ_{min} or τ_{max}
τ reduction due to extrapolation
τ reduction due to calculation
τ_{min}
τ = (1+h)τ_{min}
τ_{max}
i=0
i=1
i=2
i=3
i=4
i=5
i=6
i=7
i=8
i=9
i=10
reiteration
reiteration
time

Figure 4.20. Time-step generation, example and schematics. The red line corresponds to the ideal variable function, unknown to the algorithm but employed to generate the most suitable time-step. Each iteration, the algorithm extrapolates the variable value at a set of test locations, with the view to not allow the function to increase over a user-defined delta (height of the rectangles). Time-step increase with respect to the previous iteration is limited as $\tau_i = (1+h)\tau_{i-1}$. Time-step values below or above user-defined boundaries are also adjusted accordingly. In case that the calculated value differs from the previous one (above the predefined delta), the time-step is further adjusted and the value is recalculated or interpolated. This particular example also shows the limitation of the algorithm, as the minimum time step is not well suitable to follow the particularly steep variable dynamics.

The method's main parameters are the minimum time-step value τ_{min}, a maximum one τ_{max}, and a time-step magnification rate h. Additionally, the method requires the provision of the n tracked variables - variation tolerance, Δ_n.

Before each simulation iteration i, the algorithm performs the following:

a. Calculates the new time-step based on the magnification rate and the previous time-step as $\tau_i = (1+h)\tau_{i-1}$.
b. Extrapolates the expected value of the tracked variables as a function of time ($v_{n,sol}(t_{0\to i})$), via Akima spline function. The extrapolation temporal coordinate is $t_i + \tau_i$, which yields $v_{n,ext}(t_i + \tau_i)$.
c. The difference between the last result and the extrapolated one is calculated, as $\Delta_{n,i} = v_{n,ext}(t_i + \tau_i) - v_{n,sol}(t_i)$.
d. If $|\Delta_{n,i}| \leq \Delta_{n,ref}$ (with $\Delta_{n,ref}$ being a user-provided tolerance), the time-step is determined.
e. Alternatively, the algorithm calculates a more suitable time-step interpolating the available datapoints $\left(v_{n,sol}(t_{0\to i}) \cup v_{n,ext}(t_i + \tau_i)\right)$, again using the Akima spline function.
f. If the tested time-step is lower than τ_{min}, or higher than τ_{max}, the time-step is adjusted accordingly to fit between the boundaries.

The requirement of the calculation repetition is considered after the simulation task is completed. Step c is repeated, with the use of $v_{n,sol}(t_i + \tau_i)$ in place of $v_{n,ext}(t_i + \tau_i)$. If a divergence is found above the $\Delta_{n,ref}$ tolerance, then the solution is possibly interpolated on the time-consuming FEM solver and recalculated on the faster phase-change model.

The control of the algorithm source code adds much beneficial flexibility of implementation, which would be lacking with the adoption of already published methods. One useful feature is the selection of user-provided parameters as a tracked variable, which allows interpolation in place of extrapolation. This solution allows to tailor the time-step generation according to user-provided data, but without the need for active user intervention. Yet, the user might additionally specify the time location of certain known events (e.g. rectangular pulses edges), forcing the algorithm to replace the time-step with a predetermined one.

4.3. Conclusion

This chapter illustrates the algorithmic innovations brought about to provide the needed capabilities for the emulation of the behaviour of a PPCM device via FEM methods. The new findings are then implemented onto a fully automated software environment, which removes complexity from the simulation task and importantly, speeds up the computations, so allowing physically-realistic simulation of real-world photonic phase-change devices.

In the scientific literature dedicated to the simulation of phase-change devices, most solutions adopted a reduction of the FEM model dimensionality, or similar approximations in order to simulate the phase-change at a reduced level of complexity. This approximation relies on assumptions not always applicable, such as the translational or rotational symmetry of the device and the temperature distribution. Moreover, the most physically-realistic models available to this date have high computational cost, thus further justifying the spatial restriction over the simulated volume or area. To overcome this limitation, a full – 3D cellular automata simulation platform has been developed, which i) solves the issues related to the spatial arrangement of irregular regions via a unidimensional array description coupled with the management of an explicit pointer matrix; ii) proposes an innovative solution which reduces the computational requirement of the ITS algorithm from $\mathcal{O}(n)$ to $\mathcal{O}(\sqrt{n})$, via use of sub-arrays over which the involved summations have much lower computational costs.

These solutions have been brought into the Gillespie Cellular Automata, to define a previously-missing 3D version of this phase-change model. The consistency (of the degenerate 2D case) with the previous model has been fully maintained, with the further advantage of a rebated computational cost, as predicted. The 3D capability has been demonstrated to realistically emulate asymmetric 3D distribution previously dealt only with zero-dimensional models.

A separate model, based on the more simplistic Classical Nucleation-Growth model, has been developed, with the view to mimic the functionalities of the GCA model at a fraction of its computational cost. This comes at the partial loss of the GCA physically realistic output (in particular, the grain size and distribution, as well as the amorphization process); nevertheless, such loss not only is compensated by allowing calculations on very large volumes in reasonable times but would not be particularly significant for optical devices (due to the almost linear variation of optical properties with crystal fraction), making this the model of choice for the investigation carried out in this work (see also Section 4.3.1).

A software environment has been implemented, which couples the capabilities of the FEM environment with the chosen phase-change model. The FEM model has been thoroughly described, showing the techniques adopted to reduce the volume size and thus lower its computational cost. A detailed description of the elements composing the algorithm which manages the iterative calculation of FEM model and PC model is eventually reported.

4.3.1. Note on the applicability and suitability of the 3D-CNG model

It is worth stating that the developed CNG model is not necessarily the optimal choice regardless of the intended application, as other models also have distinct advantages. For instance, the simpler JMAK model has been recently employed in G.C. Carrillo et al. [227] to rapidly extrapolate the crystal fraction of large volumes, drawing advantage from its non-spatial crystal fraction definition and extremely fast calculation. The MRE produces much more detailed solutions, and its use can be almost as efficient where spatial and temporal temperature transients are slow and predictable. The GCA model is instead optimal for applications seeking a reliable and detailed description of the crystal distribution in confined regions, for instance in nano-electronics [159].

It is also fundamental to reiterate that the introduction of the third dimension in the mathematical model does not come at a negligible computational cost, and the use of suitable 2D approximations would be of certain advantage in case of translational or rotational symmetry of the PCM volume and of the temperature field. Also, 2D approximations are always convenient where semianalitical models can be applied (see Peng et al. [213]).

Howeever, due to the peculiar absence of both symmetries and applicable approximations in the subject of this study (the photonic phase-change memory), and the desire to obtain a well-defined crystal fraction spatial distribution, the 3D-CNG becomes the model of choice for the investigation of this work, combined with the FEM environment as described in Section 4.2. Chapter 6 and Chapter 7 will illustrate the results obtained with such a simulation environment.

Chapter 5. A novel method to estimate the complex index parameters from reflection measurements

This Chapter reports the development of a novel complex refractive index estimation technique, based on the amplitude modulation of the reflection from an artificial cavity. As here demonstrated, the simple experimental setup, combined with numerical techniques and drawing hints from literature data, can provide a quick and effective way of determining refractive index when direct measurement techniques (such as ellipsometry) are not readily available.

5.1. Why is a quick method for the complex index estimation needed?

The functioning of nanophotonic devices heavily relies upon the optical properties of their constitutive material, i.e. the complex refractive index. This is in particular true for optical resonators characterized by high Q-factor, see e.g. nano-ring resonators [228], plasmonic resonators [229] or photonic crystals [174]. Thus, the certainty over the optical properties of key materials for nanophotonics is of paramount importance.

For the case of many phase-change materials, an approximate idea of the n and κ values can be quickly obtained after literature research, but the variability of such values can significant (see Figure 2.15 for the GST case). This uncertainty is due to a variety of reasons (see e.g. [230]–[234]) but most can be traced back to differences in fabrication conditions and as-fabricated material composition. The fabrication process entails the tight control of several factors, to obtain precise on-demand composition: deposition rates and temperatures, contaminants in the deposition chamber or substrate, and gas composition being the most predominant. Similarly, the surface properties are determined by various factors, such as temperature, deposited film thickness, substrate composition, and roughness. And whilst the optical properties of some materials (e.g. SiO_2) are generally consistent within literature, even with minor variations of the fabrication process, it is evident that this is not the case for PC materials.

Therefore, the numerical design and optimization of any optical or photonic device employing PC materials can not be realistically pursued without a precise estimation of the PC material's complex refractive index. The primary tool for optical properties investigation in the UV-Vis-NIR spectral range is ellipsometry

[179] (see also Section 3.1.2). This technique generally provides an accurate determination of the complex index; it nevertheless requires relatively expensive equipment that is not always available, and a quick, cheaper alternative would thus be a beneficial addition.

5.2. Technique description

The proposed technique, which finds its first reported use in Ruíz de Galarreta et al. [235], consists of the determination of the reflection spectrum amplitude of a *Gires–Tournois* (GT) etalon, composed of the investigated material laying above a highly reflective layer. In a little more detail, the GT etalon is a singular case of a (asymmetric) *Fabry–Pérot* (FP) cavity, where the Fresnel coefficient of the top layer is relatively lower than the Fresnel coefficient of the bottom layer, which is instead close to unity. It operates similarly to the FP cavity, although compulsorily in reflection. For the sake of conciseness, this structure is here referred to simply as a *cavity*. A schematic representation is shown in Figure 5.1(a), which depicts the possible reflected and refracted pathways of the impinging radiation.

As it is here demonstrated, the configuration of the cavity shapes a unique reflection pattern where the cavity resonant mode is established, a condition determined solely by the impinging wavelength, the involved materials thicknesses and refractive indices. The experimental method implies the measurement of the reflection amplitude only, which does not carry over the optical phase information: thus, reliable calculation of both n and κ values is not generally allowed. But the absorbing GT cavity maps the phase information onto the reflected amplitude and helps to solve for two unknowns amongst the constitutive parameters of the underlying mathematical system.

5.2.1. Theoretical elements and uniqueness of the solution in FP resonance condition

The following paragraph refers to the schematics proposed in Figure 5.1(a), which consists of a three-layer stack (metal, dielectric, air). The analytical determination of the reflection pattern of the GT cavity entails the calculation of the Fresnel reflection coefficients, and the solution for the surfaces-induced multiple-beam interference. In normal incidence condition (as in a common reflectometry measurement), for isotropic and non-magnetic materials, the layer stack reflection coefficient $\mathrm{R}(\lambda)$ can be obtained from the Fresnel-Airy formula as follows [236]:

$$\mathrm{R}(\lambda) = |\tilde{r}(\lambda)|^2 = \left|\frac{r_{12}(\lambda) + r_{23}(\lambda)e^{i\tilde{\beta}(\lambda)}}{1 + r_{12(\lambda)}r_{23}(\lambda)e^{i\tilde{\beta}(\lambda)}}\right|^2 \qquad \text{Eq. 5.1}$$

with $\tilde{\beta}(\lambda)$ being the roundtrip propagation phase and loss accumulation inside the cavity, calculated as:

$$\tilde{\beta}(\lambda) = \frac{4\pi}{\lambda}\tilde{n}_2(\lambda)t = \frac{4\pi}{\lambda}n_2(\lambda)t \ + \ \frac{4\pi}{\lambda}i\kappa_2(\lambda)t = \delta(\lambda) + i\alpha(\lambda) \qquad \text{Eq. 5.2}$$

Here, $\tilde{n}_2$ refers to the complex index of the cavity material, which in the proposed method is the investigated unknown. The separation into its real and imaginary components $\delta(\lambda) + i\alpha(\lambda)$ simplifies the following description. But before continuing, and returning momentarily to Eq. 5.1, the Fresnel coefficients r_{ij} of each ij surface can instead be calculated as follows, for lossy materials [237]:

$$r_{ij}(\lambda) = \frac{\eta_j(\lambda) - \eta_i(\lambda)}{\eta_j(\lambda) + \eta_i(\lambda)} \qquad \text{Eq. 5.3}$$

with the complex impedance $\eta(\lambda) = \sqrt{\mu(\lambda)/\varepsilon(\lambda)}$. For lossless materials, this equation reduces to the more commonly known form, which sees the refractive index in place of the complex impedance.

For lossless materials, it is easy to determine that Eq. 5.1 sees its minimum when $\beta(\lambda)$ is equal or close to an odd multiple of π, originated by propagative phase accumulation inside the cavity (see Figure 5.1 (a) for a schematic graphical representation). This condition is known as Fabry-Perot resonance, and its satisfaction can be found in the minimization of $e^{i\tilde{\beta}(\lambda)}$. Thus, the resonance condition can be here visualized by rewriting Eq. 5.2 (for lossless materials):

$$\frac{4\pi}{\lambda}n_2(\lambda)t = \tilde{\beta}(\lambda) = m\pi \qquad \text{Eq. 5.4}$$

with m being an odd integer number. This expression also states how there is a unique relation between the wavelength of the reflection minimum λ_{res}, the cavity investigated layer thickness t, and its optical property, here n_2. But also in case of the complex index $\tilde{n}_2$, both the resonance modulation as a function of the wavelength and its absolute value $\mathrm{R}(\lambda_{res})$, are uniquely determined. More specifically, the quality factor of the cavity Q is calculated as follows [238]:

$$\mathrm{Q} = \frac{2\pi t n(\lambda)}{\lambda} \cdot \frac{1}{\log\left(r_{12}\left|1 - e^{i\tilde{\beta}(\lambda)}\right|\right)} \qquad \text{Eq. 5.5}$$

Knowing that the Q-factor is also related to the reflectance modulation by the well-known expression (although in the limit for weakly damped systems)

$$\mathrm{Q} = \frac{\lambda_{res}}{\Delta\lambda} \qquad \text{Eq. 5.6}$$

it is possible to see how the reflectance amplitude around resonance is determined by both real and imaginary components of the refractive index of the cavity layer. Henceforth, the determination of Q and $\mathrm{R}(\lambda_{res})$ allows one to calculate the complex value of $\tilde{\beta}(\lambda_{res})$, and consequently, to determine the values of $n(\lambda_{res})$ and $\kappa(\lambda_{res})$. An example of the GT etalon reflection pattern, which further exemplifies the discussion above, is provided in Figure 5.1(b)

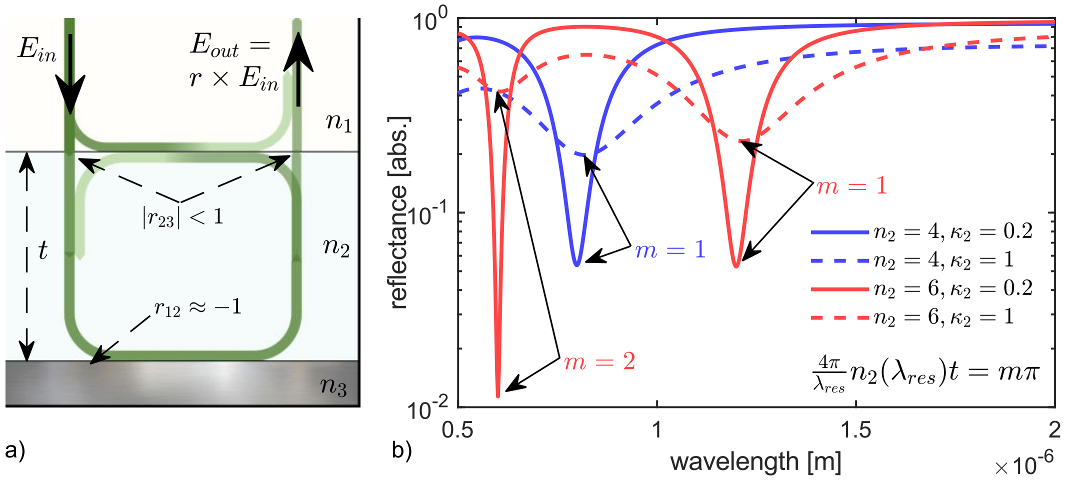


Figure 5.1. Schematics of the experimental device setup and method principles. (a) Gires-Tournois etalon schematics, and e.m. field reflection and refraction patterns. (b) Reflectance spectra for 4 sets of cavity material optical properties, which highlight the optical properties unique impact on the reflection modulation.

5.2.2. Methodology

Of course, certainty in the complex refractive index over a single wavelength is not particularly beneficial, additionally so considering that a perfect match between numerical and experimental reflection in resonant cavities is notoriously difficult (see e.g. [142], [239]). Therefore, the method relies on the fit of the whole reflectance spectrum with an analytical function, determined by the available experimental data: impinging wavelength, cavity structure, and optical properties. Specifically, the cavity reflection spectrum can be calculated via use of the Transfer Matrix Method (TMM) [240]. This approach facilitates the calculation extension to structures which may differ from the ideal condition (e.g. stacks with additional layers to the mirror and those of the cavity) and evens out any eventual mismatch due to experimental errors. To summarize, the fit procedure finds the cavity layer n and κ values by minimizing the *mean squared error* (MSE) between the TMM simulated reflectance and the experimental reflectance. This is performed upon knowledge of the invariant parameters (i.e. the FP cavity structure, comprising the layer stack order, the thickness and the (known) optical properties).

As it is so far formulated, the fitting process has two main limitations: i) the TMM function is not invertible, and ii) there are two unknown parameters (the cavity layer n/κ values), which need to be found by the fitting of one variable (the experimental reflectance). This limitation can be overcome by assuming that the the $n(\lambda)/\kappa(\lambda)$ dispersion does not stray too far from previously determined $n(\lambda)/\kappa(\lambda)$ sets, i.e. the dispersion plots of (supposedly) similar materials are correlated. To mathematically exploit such assumption, the investigated material optical properties are allowed to assume complex index values derived from the interpolation (or weighted average) of already existing datasets. This approach provides the necessary boundary condition to find the best approximation of $n(\lambda)/\kappa(\lambda)$ that fit the experimental reflectance using the TMM function, via evolutionary computation techniques.

To provide a visual representation of the problem, an example is reported in Figure 5.2. Here, the colour-coded area represents the squared error between experimental reflectance (at a fixed wavelength λ_i) and the numerically obtained reflectance for the optical properties reported on the horizontal and vertical axis (the error increases from red to blue). One first notable detail is the presence of several combinations of n/κ values leading to the desired experimental reflectance. Of course, most of these are rather distant from the experimental optical properties, and the example usefully underlines so. The crosses report the literature data obtained at the same λ_i. Restriction of the test n/κ values to the ones obtainable as weighted average of the literature data, determines a much smaller, and physically meaningful, range of test values. A parallel calculation process (for all the investigated wavelengths) is carried out to find the $n(\lambda)/\kappa(\lambda)$ dispersions that minimizes the simulated-experimental squared error sum over λ.

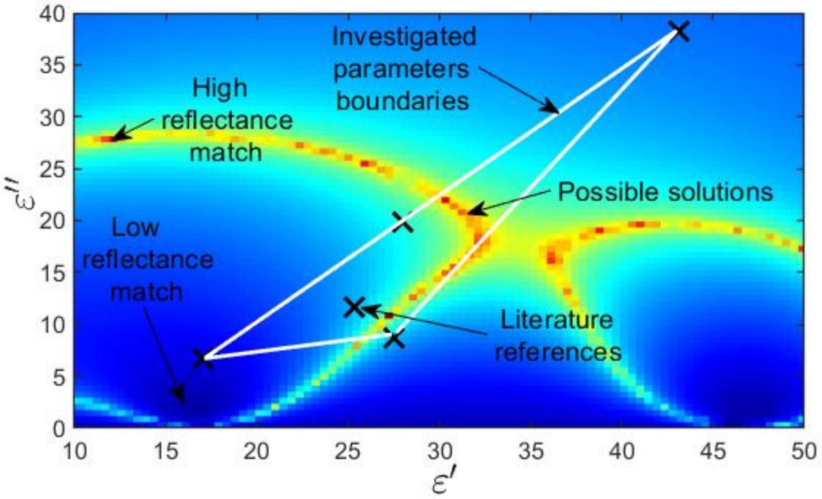


Figure 5.2. Schematics of the methodology principles, based on the simulated *vs* experimental squared error calculation for any possible cavity optical property value (inferable from the x and y axis). The colour-coded area depicts the aforementioned error, with the lower values found in red, and the highest ones plotted in blue. The fit procedure is restricted to the optical property values obtainable via weighted average of the literature data (here reported as black crosses). The search space area is here highlighted by its white boundaries. This frame reports the process for a single wavelength: the fit procedure is nevertheless carried out on each of the investigated wavelengths at the same time.

5.2.3. Numerical solver

A suitable n/κ solution set is found by finding appropriate weights w for the weighted average operation. It is here noted that each weight is applied to the whole investigated range (i.e. w parameters are not λ-dependent). Thus, the test solution $\tilde{n}(\lambda) = n(\lambda) + i\kappa(\lambda)$ is determined as follows:

$$\tilde{n}(\lambda) = \sum_{j}^{r} w_j \, n_j(\lambda) + i \sum_{j}^{r} w_j \, \kappa_j(\lambda) \qquad \text{Eq. 5.7}$$

with r being here the total number of references (literature values used to limit the search space). It is here also implied that the sum of the weights is set to 1:

$$\sum_{i}^{r} w_i = 1 \qquad \text{Eq. 5.8}$$

This assumption does not allow the code to find any n/κ point outside the allowed parameters determined by the reference set (e.g. white lines in Figure 5.2). This restriction may be relaxed permitting a reasoned extension of the weights search space, which may be useful in certain case scenarios (but as this analysis is focused on the general case with minimal hands-on requirements, such an approach won't be here covered).

The validation of the tested solution is obtained from the MSE, calculated from the analytically obtained reflectance spectrum, R_{calc} and the experimental one, R_{exp}:

$$MSE = \frac{1}{m}\sum_{j}^{m}\left(R_{calc}(\lambda_j) - R_{exp}(\lambda_j)\right)^2 \qquad \text{Eq. 5.9}$$

with m being the discrete number of wavelengths upon which the calculation is carried out. The selection of the interpolation weights after each iteration is carried out via a pattern-search algorithm. This method has been chosen due to its suitable convergence time, its quick adaptability to multi-dimensional search spaces (with the dimensionality dictated by the number of adopted references), and its capability to avoid shallow minima, combined with the convex objective function [241] (here corresponding to the minimization of MSE($\sum w_i$)).

For every iteration of the pattern-search algorithm, the method calculates the MSE output for each investigated wavelength, based on the same subset of weights; it thence improves upon the weight set, continuously refining the previous solution, until reaching a predetermined tolerance. Of course, doing so may find a good match for certain wavelengths, whilst losing confidence on other portions of the spectrum. A possibility is here to force the code to find the exact solution at the resonant wavelength, where supposedly the solution is unique; this method has nevertheless been found to be unsuitable, as minor deviations of the calculated reflectance from the experimental one might force the code output towards worse fit of the total reflectance. Thus, the average MSE is applied, so as to not amplify local calculation errors. Each step of the proposed method is here summarized:

1. Generation of a random weight set w_i, with $\sum_i w_i = 1$ (as many weights as references)
2. Construction of a test index $\tilde{n}(\lambda)$, as $\tilde{n}(\lambda) = \sum_i w_i \tilde{n}_i(\lambda)$
3. Calculation of the cavity reflectance R via TMM, including such test index $\tilde{n}(\lambda)$ as the optical parameter of the investigated material
4. Comparison of the calculated and experimental reflectances after calculation of the mean square error: $MSE = \frac{1}{m}\sum_j^m \left(R_{calc}(\lambda_j) - R_{exp}(\lambda_j)\right)^2$
5. Whereas such MSE is arbitrarily small (or after an arbitrary number of steps 1-5 iterations), the solution is found; otherwise, the algorithm returns to step 1 and generates an improved w_i set via pattern-search method, which for each iteration strengthens the choice of w_i sets lowering the resulting MSE.

An algorithmic flowchart is provided in Figure 5.3 below.

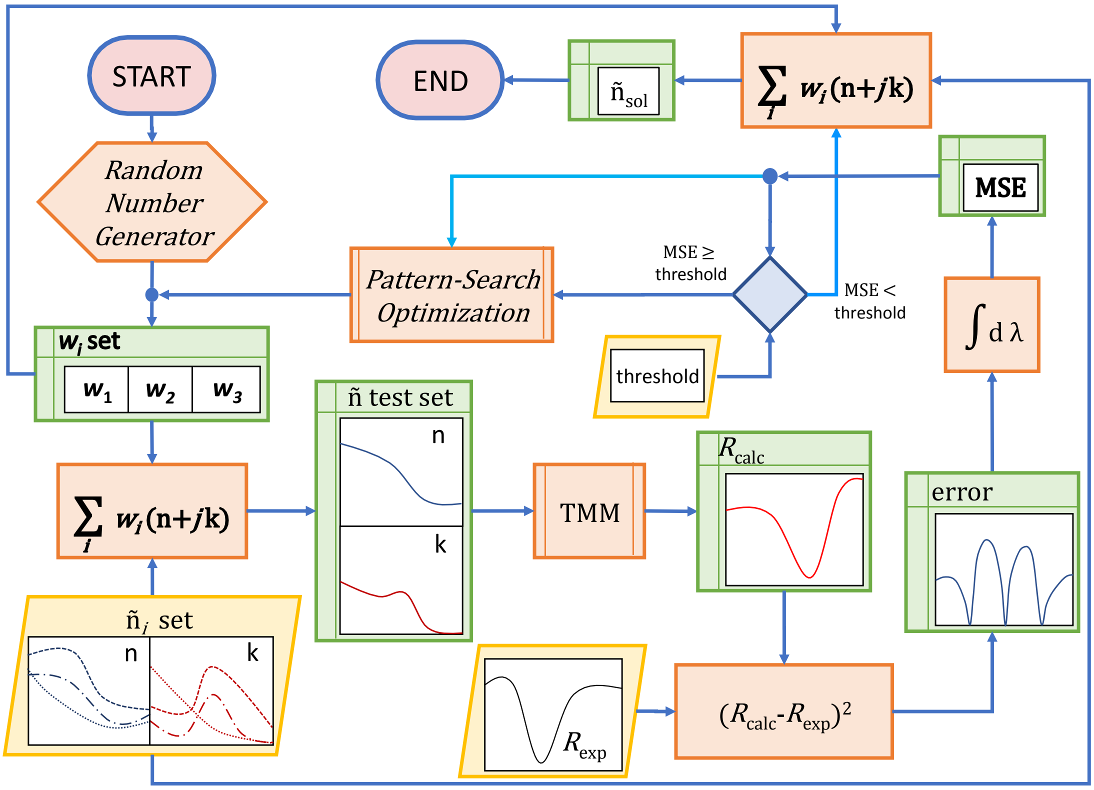

START
END
$\tilde{n}_{sol}$
$\sum_i w_i(\mathrm{n}+j\mathrm{k})$
Random Number Generator
MSE
Pattern-Search Optimization
MSE ≥ threshold
MSE < threshold
$\int \mathrm{d}\lambda$
w_i set
w_1
w_2
w_3
threshold
ñ test set
n
k
TMM
R_{calc}
error
$\sum_i w_i(\mathrm{n}+j\mathrm{k})$
$\tilde{n}_i$ set
n
k
R_{exp}
$(R_{calc}-R_{exp})^2$

Figure 5.3. Reflectance fit – n/κ determination method, algorithmic flowchart

5.3. Experimental confirmation

Tri-layer cavities, where here the three layers comprise an optically thick (100 nm) aluminium reflective layer, a PC material layer, and a 10 nm SiO_2 capping layer (to prevent the PC material layer from oxidation) were fabricated on Si substrates using magnetron sputtering (in Ar atmosphere at a base pressure of 10^{-7} mbar).

X-ray diffraction measurements (XRD) were made after fabrication (to corroborate that the as-sputtered films were in the amorphous phase) and contact *atomic force microscopy* (AFM) scans were carried out to obtain the cavity thicknesses. Crystallization of the PC material films was thermally induced via hot plate annealing at 200 ˚C for 10 minutes, and the phase transition (crystallinity of the layers) was confirmed afterwards via further XRD measurements. Film thicknesses were re-measured after crystallization to account for any volume reduction on switching from the amorphous to the crystalline phase, a well-known phenomenon [242].

The reflectance spectra of the FP cavities were then measured over a wavelength range from 800 nm to 1600 nm using a JASCO 5300 micro-spectrophotometer[42]. These measurements were normalized against a calibrated aluminium mirror of known reflectance, with the detector noise subtracted. The n and κ values of the investigated PC material cavities were then obtained by the aforementioned numerical technique. The optical properties of the non-investigated materials were taken from [81] for SiO_2, and [243] for Al, as these show marginal variations in literature and are thus considered reliable sources (maximum values of Δn = 1.5% and $\Delta\kappa$ = 0.6% have been calculated through the investigated range). It is also noted that the error derived from these parameters would induce a proportional, yet negligible, error on the aimed n/κ estimation.

Finally, the reliability and accuracy of the proposed Fabry-Perot reflectance spectroscopy method were assessed by comparison of the obtained refractive indices with those obtained from ellipsometry measurements over the same samples. Ellipsometry data was gathered using a J.A. Woollam M.2000 ellipsometer at 55°, 65°, and 75°.

---

[42] This apparatus enable the use of the confocal configuration. In case of a different angle of incidence, modifications of the TMM as a function of incidence angle and numerical aperture need to be applied.

5.3.1. $Ge_2Sb_2Te_5$

The first investigation is carried out upon GST, as it has a quantitatively high variance in terms of literature-provided optical material properties (as an example, Figure 2.19 reported a selected list of datasets in the NIR range, which highlights δn and $\delta\kappa$ uncertainty values up to 2.9 [R.I.U.] and 2.6 [R.I.U.] respectively for the crystal phase; or 1.6 [R.I.U.] and 1.2 [R.I.U.] respectively for the amorphous phase). The GST layer thickness, determined after use of the optical properties found in literature data with Eq. 5.4, was measured to be 55 nm for the as-deposited phase (amorphous). After the annealing process, the layer thickness was again measured via AFM, which outlined a shrinkage to 47 nm for the crystal GST. The fit procedure has been carried out for both the amorphous and the crystal phase by use of literature values in references [66], [153], [155], [175], [244], [245].

5.3.1.1. Amorphous GST

The experimentally determined reflectance, R_{exp}, is shown in Figure 5.4(a), as a black dotted line. As visible, the resonance dip is found around 1100 nm, well within the investigated range boundaries as desired. Notably, the calculated reflectance based on the ellipsometrically determined optical properties (blue dashed line, reported as *comparison* for sake of simplicity) is very close to the experimental one, confirming that the model parameters have been correctly determined. The reflectance calculated assuming the optical parameters taken from the various references used are also reported in Figure 5.4, as grey lines (coloured in progressively more vivid green, in proportion to the weight assigned by the fit algorithm).

The fit result (red line) has a close match with the experimental one, with a MSE of 7.21×10^{-4}. The closest match with the reflectances calculated using only literature data is obtained from [175], which yields a MSE of 1.33×10^{-2}, roughly 20 times higher. Other relative error values span from 2.26×10^{-2} to 9.47×10^{-2}, confirming how the reliance upon literature data may be highly misleading.

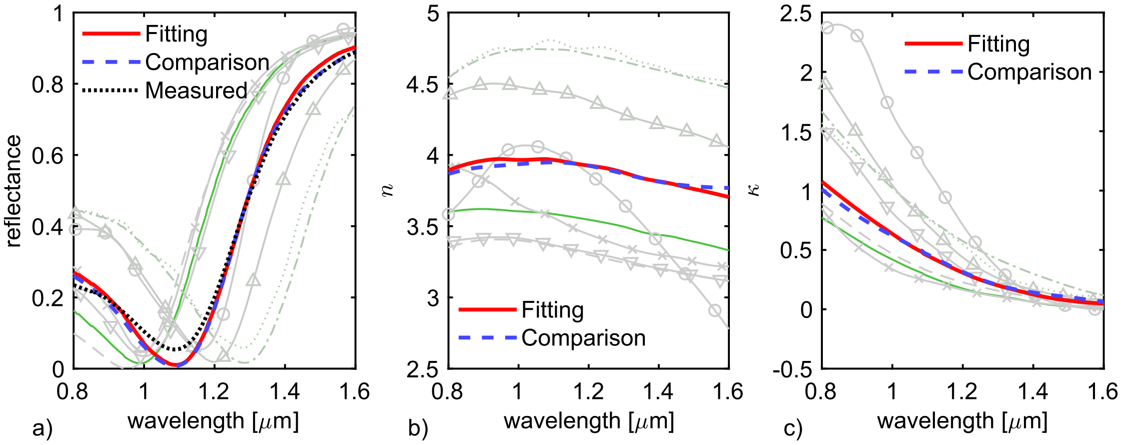


Figure 5.4. Amorphous GST, reflectance fit results and n/κ data extraction. In each panel, literature data (here from Refs. [66], [153], [155], [175], [244], [245]) is reported in un-saturated colours, grey to green as the fit method weight increases. a) Experimental (black dotted line) and calculated data. The comparison line (blue dashed line) is calculated based on the ellipsometrical estimation of the n/κ parameters. b-c) Refractive index and extinction coefficient respectively, with the ellipsometric data (blue dashed line) superimposed for confirmation of the fit quality.

Now the focus is shifted to the n/κ dispersions, plotted in Figure 5.4(b-c) respectively. Here, the dispersion data obtained from the fit procedure (red lines) can be compared to the ellipsometrically obtained optical properties of the same material (blue dashed line). The reference data is again reported as grey lines, to provide a graphical representation of the variation of parameters reported in the literature.

In line with the reflectance measurements, a close match between the algorithm output and the experimental n/κ values is evident. Average relative errors $\bar{\zeta}$ here provide a quantitative representation of the fit confidence, and values of $\bar{\zeta}$ were $\bar{\zeta}(n) = 0.55\%$, and $\bar{\zeta}(\kappa) = 6.85\%$. Again, literature data finds much higher relative errors (e.g. minimum values of $\bar{\zeta}(n) = 7.35\%$ and $\bar{\zeta}(\kappa) = 23.7\%$ are here found).

5.3.1.2. Crystalline GST

The same reference set is used to fit the data of crystalline GST. These, and the fit results for the refractive index, extinction coefficient, and reflection spectrum are plotted in Figure 5.5. The resonance peak is located at 1420 nm, still within the range boundaries (of the spectrophotometer) even after annealing. As before, the calculated reflectance yields a remarkably low MSE = 2.08×10^{-4}. The MSE obtained from the literature data-calculated reflectances span from 8.22×10^{-3} to 0.160 (obtained from [175] and [66] respectively).

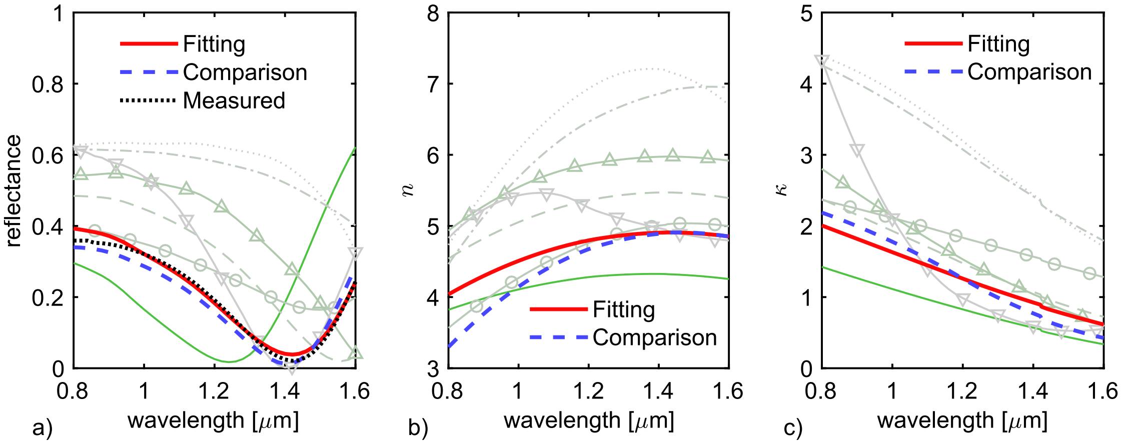


Figure 5.5. Fit results for the crystalline GST sample, reported in analogy with the previous case. (a) Reflection spectra. (b) Refractive index. (c) Extinction coefficient.

In this case, a more evident error is noted at wavelengths distant from the resonance peak, especially for the refractive index n. Nevertheless, from the quantitative perspective, a good match with the ellipsometrically determined data is still obtained, more so when compared to most of the reference set. Average relative errors of $\bar{\zeta}(n)$ = 5.39% and $\bar{\zeta}(\kappa)$ = 15.3% are obtained; the whole reference pool has instead minimum $\bar{\zeta}$ values of $\bar{\zeta}(n)$ = 3.07% and $\bar{\zeta}(\kappa)$ = 22.9% (obtained from [175] and [245] respectively). Note that here a lower error can be obtained using the refractive index of reference [175], which is although correlated to a much higher error for the extinction coefficient ($\bar{\zeta}(\kappa)$ = 64.8%).

5.3.1.3. Comparison of the crystal and amorphous phases fit results

The investigations carried out on the amorphous phase consistently yield a lower $\bar{\zeta}$ error than the ones carried out on the crystal phase (see also GeTe in Section 5.3.2). It is here supposed that minor changes of the material composition may cause appreciable variation of the crystal reticular structure, and thence of the corresponding band-gap and optical properties. Consequently, the reference pool exhibits wider divergence, worsening the fit quality. This also correlates with the better agreement here found for the amorphous material fit, as it would implicitly be more defect-tolerant with respect to the composition, exhibiting a lower Δn and $\Delta\kappa$ divergence from sample to sample (this also agrees with the lower Δn and $\Delta\kappa$ of the reference pool). Additionally, the crystal phase is strongly absorptive in the investigated spectral range, and therefore the fit process would see a wider solution range, again decreasing the fit quality (see Appendix III for more details).

5.3.2. GeTe

GeTe is another well-known PC material that sees lack of consistency between published datasets of its optical properties. In this case, the selected reference pool is obtained from Refs. [155], [244], [246]–[248]. Data is visible in Figure 5.6(b-c) and Figure 5.7(b-c), and as in the previous case, it reiterates the necessity to confirm self-fabricated film optical properties, since (as for GST) uncertainty delta values well above 1 [R.I.U.] can be calculated in all cases (n and κ for both amorphous and crystal GeTe).

For the fabrication of the GeTe cavity device, to avoid possible issues derived from the fast deposition rate, the aimed GeTe layer thickness was 200 nm, hitting a higher resonant mode order ($m = 4$ in Eq. 5.4). AFM scan yielded a thickness of 203 nm for the as-deposited phase, and after annealing, of 177 nm for the recrystallized phase.

5.3.2.1. Amorphous GeTe

As with the previously reported examples, the amorphous GeTe reflectance shows the characteristic cavity feature, here centred around 1200 nm. Reflectance data is reported in Figure 5.6(a). A good agreement is obtained after calculating the reflectance from the ellipsometrically estimated optical properties (blue dashed line), albeit noting a slight shift of the resonance wavelength.

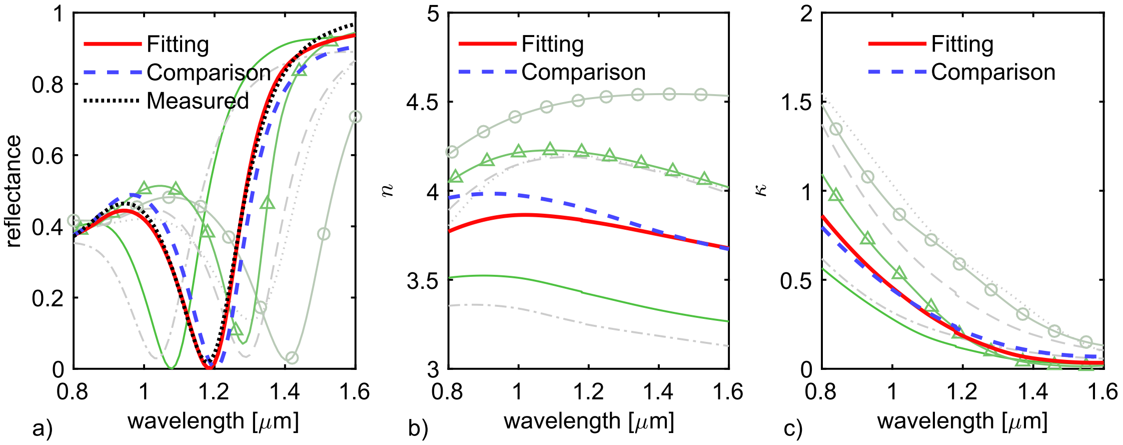


Figure 5.6. Fit results for the amorphous GeTe cavity sample, reported in analogy with the previous cases. (a) Reflection spectra. (b) Refractive index. (c) Extinction coefficient. Literature data is taken from Refs. [155], [244], [246]–[248].

The fit result has also an excellent match with the measured spectrum, with an MSE of 2.93×10^{-4}; the lowest MSE calculated from the literature datasets is instead 3.95×10^{-2} (from [248]). As for the optical properties, a corresponding good result is obtained for both the refractive index ($\bar{\zeta}(n)$ = 3.57%) and the extinction coefficient ($\bar{\zeta}(\kappa)$ = 13.7%). Minimum values of $\bar{\zeta}(n)$ = 6.29% and $\bar{\zeta}(\kappa)$ = 19.2% are obtained from the reference pool (from [155] and [246] respectively), which highlights the remarkable improvement in the optical properties' accuracy.

5.3.2.2. Crystalline GeTe

In the case of crystalline GeTe, the resonance dip (λ_{res} = 1576 nm) is found close to the range boundary. This is a limit case for the applicability of the technique, as it does not contain the whole resonance feature, and might lead to difficulty to obtain a good fit. Data is reported in Figure 5.7.

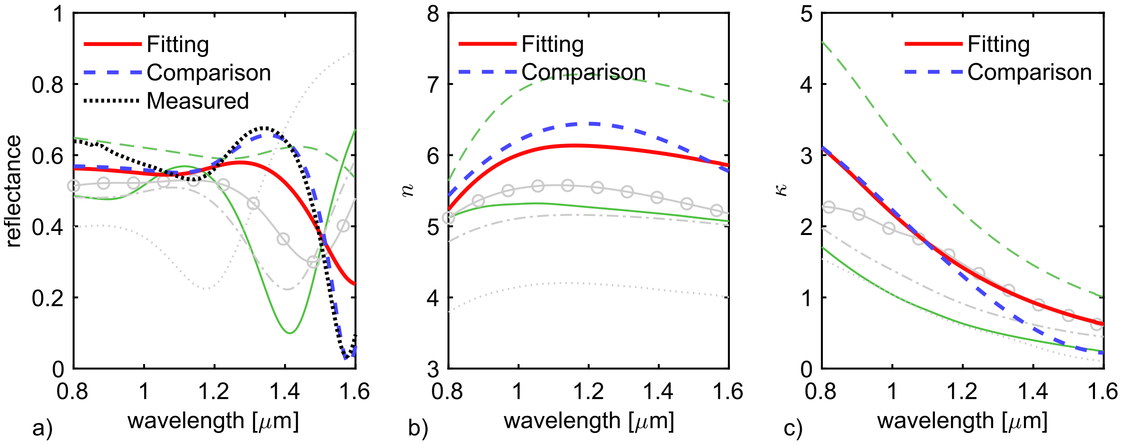


Figure 5.7. Fit results for the crystalline GeTe cavity sample, reported in analogy with the previous cases. (a) Reflection spectra. (b) Refractive index. (c) Extinction coefficient. Literature data is taken from Refs. [155], [244], [246]–[248].

As predicted, the reflectance fit is here much poorer than in any of the previously reported cases, with a MSE = 7.31×10^{-3}, opposed to the minimum MSE value of 2.94×10^{-2} for the reference pool (obtained using [248]). Predictably, this is also paralleled by a lower match of the refractive index with the experimentally determined one. Average relative errors of $\bar{\zeta}(n)$ = 3.22% and $\bar{\zeta}(\kappa)$ = 42.1% are found, which still compares well with the values found taking any arbitrary value from literature: the minimum values of $\bar{\zeta}(n)$ = 11.4% and $\bar{\zeta}(\kappa)$ = 37.2% are obtained from [155] and [244] respectively. Whilst the algorithm improves the reliability over any arbitrary reference choice, the relatively poor reflectance fit suggests that the adopted reference pool is not suitable for this specific problem.

5.3.3. Discussion

With the exception of the crystalline GeTe, a good fit of the reflectance corresponds to a much-improved estimation of the n and κ values in all the remaining cases. It is here noted though that the correlation between low error on the reflectance fit, and low error on the n/κ solutions, is not generally ensured. In principle, an infinite number of n/κ sets yields an acceptable reflectance value; the element that supports the reliability on this correlation is the restriction of the determination of n/κ solutions to the ones close to literature data. This approach, extended and mediated over the whole investigated range, ensures a physically realistic output, and thus allows to find accurate n/κ solutions by minimizing the reflectance fit error.

With respect to the fit operation, the retention of a good agreement around λ_{res} (and the deterioration of the fit quality further from such wavelength) is an expected behaviour of the algorithm (see e.g. the crystalline GST case, Figure 5.5(b)). During the fit procedure, the MSE is calculated averaging the error of the squared values across the whole investigated range: henceforth, a mismatch at the resonant wavelength (where the reflection rapidly decreases) produces a higher increase of the MSE than the same mismatch where the dispersion is near-flat. Thus, the algorithm inherently optimizes for the resonant wavelength output and henceforth loses accuracy further from the resonant location.

The possible mismatch between fit result and experimental data can be due to several reasons, including i) insufficient resolution of the resonance feature on the reflectance spectrum, ii) inherent measurement uncertainties of reflectance and film thicknesses, ii) surface roughness effects (on reflectance), iv) the effects of any ‘interfacial’ layers. The last element has indeed some degree of influence over the reported cases, where it is likely that a thin layer of aluminium oxide forms on top of the Al layer, especially after annealing, as the addition of such an oxide layer improves the algorithm output (decreases the n/κ mismatch). This helps to emphasize the fact that that, although the technique is quite simple in its implementation, it still relies on the accurate determination of the non-investigated model parameters (i.e. optical properties and thicknesses of the dielectric and metal layers in the stack).

But, even after confirmation of these parameters, one further source of mismatch is the reference pool. The fit operation is heavily reliant on the provided reference data. More than the lack of resolution on the cavity feature on the reflectance, the crystalline GeTe case suffers from an insufficiently extended n/κ search space. This consideration urged a further elaboration of the proposed method and led to the design of a numerical technique improving on the available reference pool; the methods and results of such an alternative technique (which combines the method here illustrated with a complementary numerical routine) are reported in details in Appendix III.

5.4. Conclusion

The technique described in this chapter is designed to estimate the complex refractive index without the need for expensive equipment and/or complicated post-processing algorithms. The data here presented confirm how a strict control of the optical parameters of the fabricated materials (especially for PC materials) is of paramount importance, due to the wide divergence of the characterization reported in literature assumed to describe the same material, and the proposed method supply a fast and cost-effective tool to accomplish so. This method requires the measurement of the reflectance from a simple layer stack forming an asymmetric Fabry-Perot absorbing cavity. Such device maps both amplitude and phase information onto the reflected spectrum with a unique configuration, comprising the reflectance damping at the resonant wavelength λ_{res}, and the Q-factor of such damped oscillator. The reflectance spectrum uniqueness allows the retrieval of the n/κ information at resonant wavelength, here obtained by fit of the reflectance feature using a TMM function, combined with evolutionary computation techniques. Specifically, the optical properties are determined as the weighted average of a reference pool, which limits the possible values of n/κ to those close to the aimed values; the fit process refines on the applied weights in order to minimize the MSE between experimental and simulated reflectance spectra. As the fit process is carried out on the whole investigated spectrum, the fit result yields the $\tilde{n}(\lambda)$ dispersion as $\tilde{n}(\lambda) = \sum_j w_j \tilde{n}_j(\lambda)$.

Chapter 6. Understanding the Read and Write processes in the conventional phase-change photonic memory architecture

6.1. Phase-change photonic memory: state of the research

As illustrated in Section 2.1, the field propagation within a rib or ridge type photonic waveguide is extremely sensitive to any disturbance of the dielectric function in proximity to the waveguide external surfaces. The evanescent field coupling is indeed one of the most commonly engineered solutions aiming to maximize the interaction between the guided mode and any external element. This arrangement is indeed used for the *phase-change photonic memory* (PPCM), with a thin layer of GST (hereby referred to as the *unit cell*) responsible for the tunable and non-volatile modulation of the transmitted signal. Here, the design driving force is to conjugate the advantages and flexibility brought by the photonic platform, with the intrinsic functionality of PC materials. For sake of clarity, Figure 6.1 reports a schematics of the device concept and functioning.

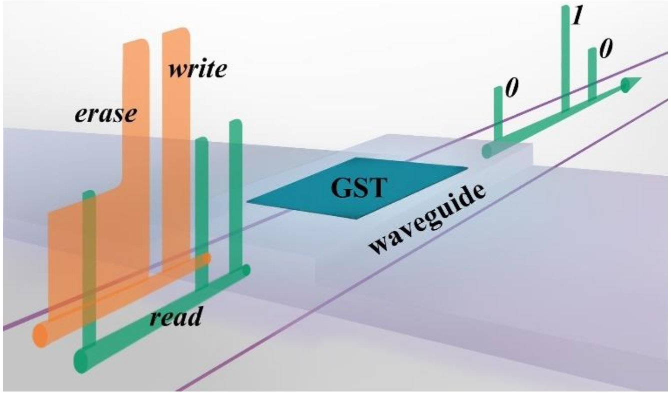


Figure 6.1. PPCM device schematics and functioning principles. A section of the waveguide is portrayed (indigo), comprising the GST unit cell on its top surface (emerald). A Write-Erase pulse sequence (orange) is delivered trough the waveguide, and a set of read pulses (green) highlights the consequent change of the device transmissivity, here enabling a binary memory functionality.

The PPCM device was initially proposed by Pernice and Bhaskaran in 2012 [66], implementing a 10 nm layer[43] of PC material (GST) on a micro-ring resonator[44]. It was there demonstrated that, because of the widely diverging complex indices of GST amorphous and crystal phases (see Figure 2.16), the propagating signal would undergo very different attenuations and phase shifts with the unit cell in its different states.

---

[43] The thickness of the layer, usually 10 nm, is selected to balance the optical attenuation, the achieved optical contrast (i.e. read signal) after Write (i.e. amorphization) or Erase (i.e. crystallization) operations, and the energy efficiency of those [67], [251].

[44] Whilst the analysis here reported focuses on the straight waveguide configuration, the interaction mechanism remains the same as with a unit cell fabricated on the top surface of a micro-ring resonator.

Switching from one phase to the other was obtained optically, by Joule heating after delivery of optical pulses (sent down the waveguide) of appropriate power. The increase of the temperature of amorphous GST above a crystallization threshold temperature (for sufficiently long duration) induces fast recrystallization. Conversely, optical pulses with relatively higher power and shorter durations allow the rapid heating of GST above its melting point, whilst reducing the heating of the surrounding volume; this process allows for a fast temperature decrease once the pulse terminates, hence inducing glass formation (amorphization) of the molten GST phase (melt-quench process). Readout is carried out with low power optical signals, that do not result in any significant temperature increase; thus, as the crystal fraction is retained unless temperature increases, the information can be firmly encoded onto the unit cell phase-state. Moreover, it was reported that the read signal modulation is univocally related to the thin layer crystal fraction, therefore highlighting the potential for multi-level storage further developed in the following years [68]. Indeed, after the first manuscript on PPCM devices from Pernice and Bhaskaran [66], several publications from the same research groups [67]–[70], [73], [74], [163], [172]–[174], [249]–[255] have reported subsequent advances in PPCM development and application.

However, a complete, physically realistic model encompassing optical, thermal and phase-change processes involved in the write, read and erase operations for PPCM devices was absent at the commencement of the work of this thesis. Such a model is developed in this chapter, with the aim of discovering the parameters and mechanisms that play a determinant role in the PPCM behaviour. To build a reliable simulation model, based on the available data, the geometry of the PPCM device, the material parameters and the physical mechanisms are here thoroughly analysed, with the view to pinpoint the elements which have a dominant effect on the PPCM optical, thermal and switching behaviour.

6.2. Simulation of the Read operation

The first step is the consolidation of a FEM model via simulation of the Read operation for a test experimental configuration. A suitable starting point is the analysis of the results reported in Ríos et al. [67]. This is a particularly useful piece of work, as it encompasses several different configurations, and brings additional pieces of information which become essential for the construction of a solid and well-defined simulation model.

6.2.1. PPCM configuration

In Ríos et al. [67], each PPCM device is fabricated on a rib SiN waveguide (330 nm × 1300 nm, with an etching step of 165 nm), placed on top of a SiO_2/Si wafer. The unit cell is reported to be 10 nm thick, with an equally-thick cap layer of ITO; the unit cell is placed only on the top surface of the waveguide, with a width of 1 µm.

The SiN refractive index is $n_{SiN} = 1.99$ at 1550 nm[45] at room temperature. For the GST complex index, the initial assumption is based on the data reported from Stegmaier et al. [163] (as also highlighted in Section 2.3.3), thus yielding $\tilde{n}_{cr} = 6.11 - 0.83i$. The ITO capping layer has a refractive index of $\tilde{n}_{ITO} = 0.6 - 0.04i$ [172]. Due to the consistency of the published optical parameters for SiO_2 and Si, these values are taken from literature (Refs. [256] and [257] respectively).

The initial simulation is performed on a 1 µm-long unit cell, for which Ref. [67] reports an optical attenuation of -0.97 dB, which translates into an absolute value of the transmission of $\mathrm{T_{exp}} = 0.798$.

6.2.2. Crystalline unit cell simulation results and analysis

A 3D simulation model matching the geometrical parameters above mentioned is constructed (see Section 3.2 for further details). To mimic the Read operation, a test optical pulse of 1 µW[46] is delivered through the waveguide on a crystalline unit cell. Within the COMSOL environment, for any model containing 2 or more ports being solved for the wave propagation, the S-parameters[47] are automatically calculated, allowing for the easy retrieval of the transmittance information. In the photonic memory two-port model, the absolute transmission $\mathrm{T_{sim}}$ is obtained as follows:

---

[45] In most of the works concerning the PPCM device, pump and probe wavelengths are often varied, due to the experimental setup. For example, in Ríos et al. [67] the pump laser wavelength is 1559.87 nm; the probe one is 1560 nm. Other works use wavelengths between 1530 and 1610 nm. To simplify the analysis, it is here assumed (unless explicitly stated) that the propagation wavelength can be approximated at 1550 nm, and the derived material optical properties are assumed accordingly.

[46] The power of the read pulse operation is non-influential, as the read operation is simulated as a time-independent phenomenon.

[47] The S-parameter approach considers a device as a black-box, whose interaction with any external components takes place through ports. More specifically, the S-parameters are the unit-less representation of the magnitude and angle of the reflected and scattered waves emitted and collected from the model ports, generally represented in a matricial N×N form, where N is the number of ports.

$$\mathrm{T_{sim}} = |\mathrm{S}_{21}|^2 \qquad \text{Eq. 6.1}$$

The initial result yields a transmission value of $\mathrm{T_{sim}} = 0.581$. The mismatch with the experimental transmission value ($\Delta\mathrm{T} = \mathrm{T_{exp}} - \mathrm{T_{sim}} = 0.218$) is too sizeable to be attributed to experimental data fluctuations (i.e. uncertainty in the experimental measurements, which can be estimated to be 2.5% in Rios et al. [67]); therefore, the origin of such a mismatch must lie elsewhere. A possible source is certainly found in the optical properties of GST, since i) for the initial simulation these were are assumed a priori, as no quantification is pursued in [67]; ii) as outlined in Chapter 5, reported GST optical properties are notoriously divergent in the literature. Nevertheless, to reliably pinpoint all the possible sources of the found mismatch, it is of paramount importance to systematically determine the magnitude of the influence that each of the model parameters has on the transmission output. To accomplish this task, a series of simulations are performed, varying only one amongst the most important geometrical and material parameters: unit cell thickness, width and length; waveguide thickness, width, and etch step depth; unit cell n/κ, cap layer n/κ, and waveguide n. These variables are swept to cover a (roughly) 20% variation above and below the initially used ones (with the exception of the optical parameters, which are here allowed to vary over a wider set of values). The results are reported in Figure 6.2.

These results highlight the effect of each variable on the transmission modulation. It is here reiterated that, due to fabrication tolerances, uncertainty of the measurements, and arbitrariness of the presumed (otherwise missing) variables, any of the investigated parameters can indeed deviate from the expected ones. Assuming that the most likely scenario sees a small relative error attributed to each variable, it follows that the variables more strongly correlated to the PPCM signal modulation would also be the reasonable source of the found mismatch.

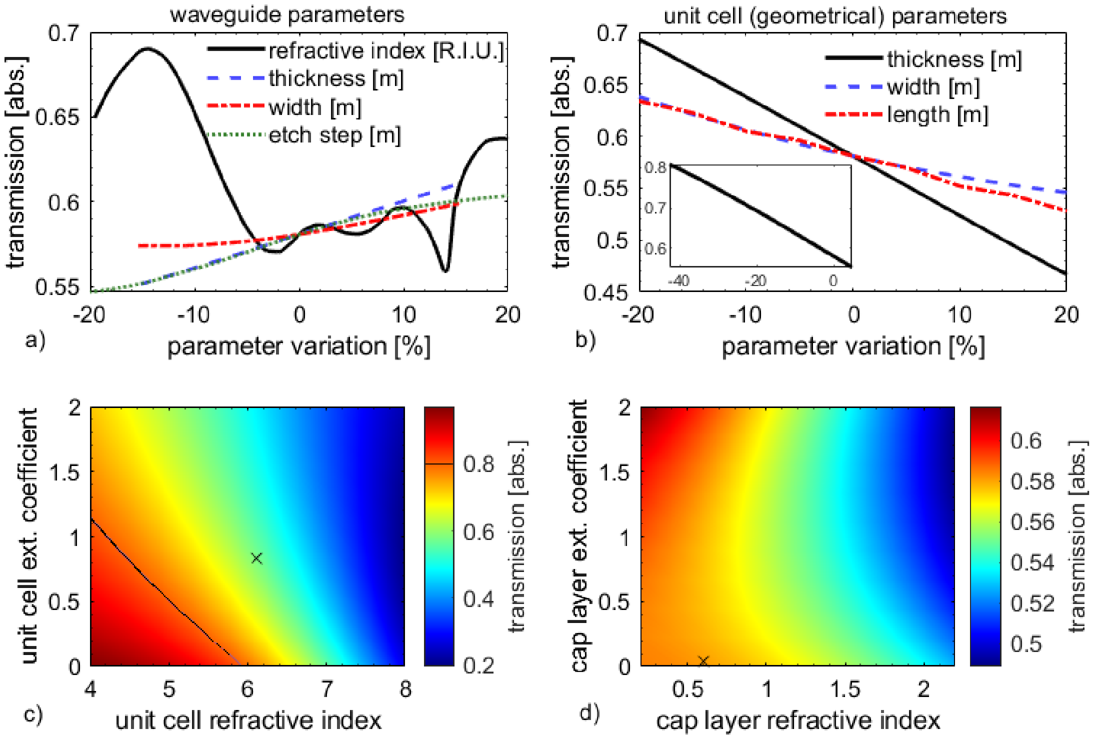


Figure 6.2. Simulation results for the Read operation, obtained upon variation of a single parameter amongst the ones initially assumed. a) Transmission upon variation of the waveguide-related parameters (thickness, width, etching step, and refractive index). b) Transmission calculated varying one of the geometrical parameters of the unit cell: thickness, width, length. The inset shows the extended range calculation for the unit cell thickness. c) Results upon variation of both real and imaginary components of the unit cell refractive index (visible on the x and y axis respectively). The experimental transmission value ($T = 0.798$) is highlighted as a black line. d) Results upon variation of both real and imaginary components of the cap layer refractive index (visible on the x and y axis respectively). In (c,d), the initially adopted refractive indices are marked with a black cross.

6.2.2.1. Influence of the PPCM geometry

It is immediately visible how the system is remarkably tolerant over the estimation error on most of the geometrical parameters (see Figure 6.2(a-b)), with transmission variations ΔT (between the highest and lowest calculated value) below 10%, thus not justifying the divergence between the initial simulation results and the experimental results reported by Ríos et al. [67]. This holds true for all waveguide-related parameters, and for the unit cell width and length. The case of the unit cell thickness, due to its value being in the order of few nanometers, will be discussed separately in Section 6.2.2.3.

6.2.2.2. Influence of the optical properties

Now the attention is turned to the optical properties. The waveguide refractive index (black line in Figure 6.2(a)) has a non-trivial effect on the transmission, as it modifies the e.m. field distribution of the device (waveguide + unit cell) propagating mode, and therefore it also changes the extent to which the waveguide mode is accommodated onto the device mode along the propagation direction. This results in the non-monotonic variation of both scattering and unit cell absorption (major contributors to the signal loss), and henceforth, transmitted power. Regardless, even by assuming a (rather unphysical) value of $n_{SiN} = 1.7$ (where the peak transmission is calculated, see Figure 6.2(a)), the transmission only increases up to $\mathrm{T} = 0.690$, thus not yet justifying the mismatch between calculated and experimental value. Therefore, in all subsequent simulations, the SiN waveguide refractive index (of $n_{SiN} = 1.99$) is also assumed to be correct, with its variation being mostly uninfluential.

ITO, the material here composing the capping layer fabricated on top of the unit cell, sees a wide range of published optical properties, varying after both different fabrication processes and thermal history, thus the expanded investigated range for unit cell thickness shown in Figure 6.2(b). Yet, the effect that the capping layer refractive index has on the overall transmission is rather negligible (see Figure 6.2(d)). A minor variation of the capping layer complex index has a very small impact on the transmission; additionally, a large variation of the same has a higher probability to lead to transmission decrease, in contrast to the aimed simulation optimization. Henceforth, the capping layer optical properties are here not considered to be the source of the found mismatch[48]. In the following simulations ITO retains the optical properties aforementioned: $\tilde{n}_{ITO} = 0.6 - 0.04i$.

As for the unit cell refractive index, Figure 6.2(c) demonstrates how the transmitted power on readout is susceptible to significant variation upon change of the unit cell optical properties. The same plot is reported in Figure 6.3, with additional marks showing the reported literature data for crystalline GST, at 1550 nm. Stegmaier [163] reports the only data point obtained from a device close in nature to the PPCM device (using unit cell-implemented MZI[49] devices), whereas the remaining data points are obtained via ellipsometry carried out onto thin GST layers.

---

[48] Upon this result, it is here assumed that a variation of the cap layer thickness should also have comparatively negligible impact (see e.g. the case for an absent (or transparent) cap layer, where the ITO assumes the complex index value of $\tilde{n} = 1 + 0i$ for a $\Delta\mathrm{T}$ variation of 0.01)

[49] Mach–Zehnder Interferometer.

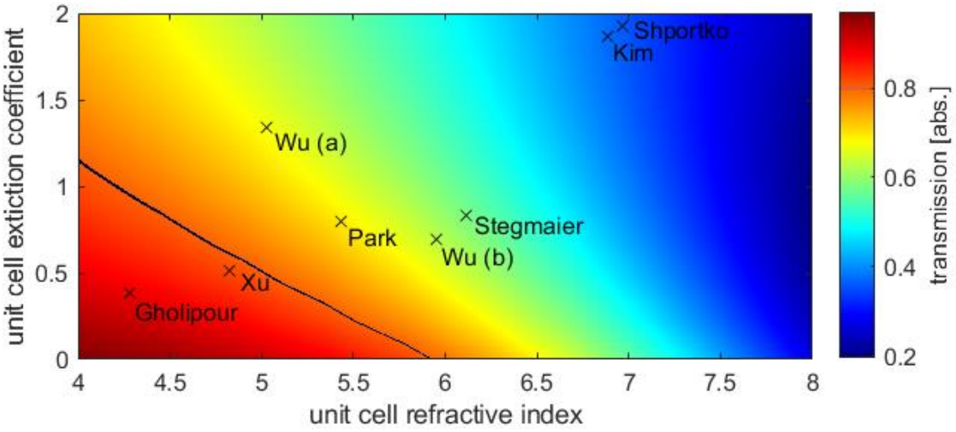


Figure 6.3. Transmission as a function of the unit cell refractive index, data from Figure 6.2(c) with additional markers showing the available literature data for the unit cell material (GST) at 1550 nm. Literature data are taken from Refs. [153], [155], [163], [175], [244], [245], [258]; markers labels report the manuscripts first authors.

Amongst the reported values, one data point shows that GST may indeed be fabricated with optical properties justifying the found mismatch (Xu et al. [245], close to the black line plotted at $\mathrm{T} = 0.798$). Nevertheless, the Xu dataset is in some measure reflecting the optical properties of mixed-phase GST (i.e. partially crystallized GST), as it corresponds to the investigation of progressively annealed GST, hence with uncertainty over the investigated layer crystal fraction. There is only one dataset, from Gholipour et al. [153], which reportedly measures fully crystalline GST and yields optical properties capable of increasing the (simulated) PPCM transmission towards that reported experimentally.

Assuming that the unit cell optical properties are the only source of calculation mismatch to the data reported in Ríos et al. [67], there are three case scenarios: i) the PPCM device operates with a partially crystallized GST unit cell, ii) the fabricated GST has a singularly low refractive index and extinction coefficient (as found in Ref. [153]), or iii) a combination of both.

In Ríos et al. [67], the authors report that the crystal phase is obtained after initial hot-plate annealing (though not specifying temperature nor time), followed by further crystallization via a series of melt-quench/recrystallization operations (until a post-recrystallization constant baseline is reached). As the process does not rely on a single operation, but instead uses a sequence of steps until a transmission minimum is achieved, it is reasonable to assume that the reported data corresponds to a completely crystal unit cell, thus preventing to blindly adopt crystal GST properties which may be attributed to partially-crystalline phase (hence negating points (i) and (iii)).

As for point (ii), there is no reason to exclude a priori the fabricated GST with low refractive index and extinction coefficient; nevertheless, as most of the literature data point to higher values of both (yielding transmission values ($\mathrm{T} < 0.7$)), such probability is, at least, low. Thus, it is most likely the case that, the unit cell optical properties are probably different from the ones adopted for the initial simulation, as a small variation of such properties leads to significant variation of the transmitted power. Likewise, the unit cell optical properties are probably not the only source of mismatch, for the reasons discussed above.

6.2.2.3. Influence of the unit cell thickness

Returning to the PPCM geometry, the factor which more strongly determines the PPCM transmission is the unit cell thickness, here responsible for a variation in transmission above 23% over the calculated range (see Figure 6.2(b)). Indeed, if an extended thickness range is considered (see inset in Figure 6.2(b)), the data shows that as the unit cell thickness decreases, a near-linear increase of the transmission is obtained. This analysis also shows that a unit cell of 6 nm would be necessary to obtain the aimed transmission value of $\mathrm{T} \cong 0.8$. Such a thickness corresponds to a decrease in the estimated GST volume by 40%. Indeed, it is well known that the PC material layer thickness reduces upon recrystallization (as reported in Chapter 5, see also Ref. [259] which reports a volume reduction of up to 11.4%). The presence of voids on the as-fabricated layer may explain further shrinkage from the expected layer thickness. Thus, it is justifiable to presume that the unit cell thickness can be lower than the nominal 10 nm.

6.2.3. Adjusted simulation: results and conclusion

In conclusion, a likely source of mismatch between simulated and experimental data sees the combination of i) divergence of the GST optical properties from the assumed ones, and ii) deviation of the unit cell thickness from the nominal one. The separation of those effects can not be pursued unless devising an ad-hoc experiment. To pursue the simulation framework validation and limit the number of degrees of freedom, the GST refractive indices are fixed to the ones precedently portrayed (see Figure 2.17); instead, the unit cell thickness is adjusted to an *effective* value, to mimic the experimental results and to test the simulation framework response. The Read operation simulation for the 1 μm unit cell and 5 µm unit cell (the latter having a reported attenuation of -7.14 dB, corresponding to $\mathrm{T} = 0.193$) are then reiterated, varying the unit cell thickness.

Results are reported in Table 6.1 and these show that close match of simulated and experimental transmission can indeed be obtained in such manner, using 6 nm and 8 nm respectively for the unit cell thickness. For sake of completeness, the e.m. field distribution is also reported in the two cases, in Figure 6.4. It is noted that the top row of Figure 6.4 (a,c) shows unusual patterns which hints to cavity effects. However, those patterns are due to the unit cell mode accommodation from and to the waveguide mode, as clearly visible in Figure 6.4(b,d).

Obviously, the two effective thicknesses do not correspond to each other. This can be due to i) differences during the fabrication processes for the 1 and 5 µm cells or ii) degradation of the unit cell surface uniformity due to iterated Write-Erase cycles (as discussed in Refs. [69], [175]), which also suggests that the longitudinal modal profile (widely different in the 1 µm and 5 µm unit cells, as visible in Figure 6.4) may determine slightly different patterns of heat-induced thickness deformation (and henceforth minor deviations from the expected transmission level). It is here noted though that, assuming an experimental error over the transmission data of 2.5%, it would be possible to approximate the thickness of the investigated unit cells to 6.5 and 7.5 nm, approaching an average value of 7 nm and possibly further justifying the experimental deviation sources previously mentioned.

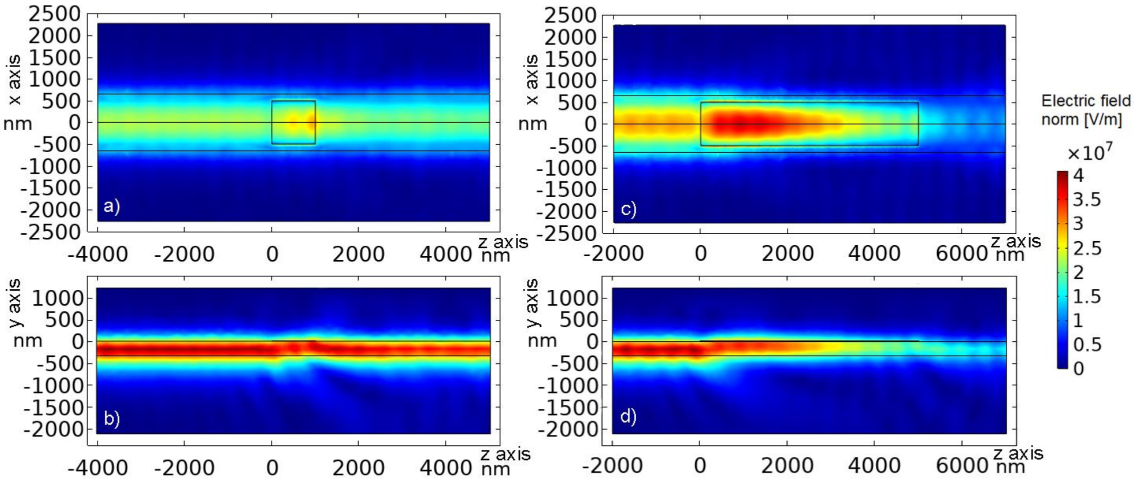


Figure 6.4. Read simulation, electric field norm distribution across the two orthogonal cut planes oriented along the waveguide propagation axis (colour bar on the right). The propagation vector is oriented from left to right. (a,c) zx waveguide top-surface cut plane; (b,d) zy symmetry cut plane. (a,b) 1 µm unit cell, thickness of 6 nm, transmission $\mathrm{T} = 0.796$. (c,d) 5 µm unit cell, thickness of 8 nm, transmission $\mathrm{T} = 0.189$.

Table 6.1. Simulated Read operations of the devices reported in Ríos et al. [67]. Matching unit cell (effective) thickness values are highlighted in red.

Unit cell length	6 nm	6.5 nm	7 nm	7.5 nm	8 nm	experimental
1 µm	0.796	0.772	0.746	0.720	0.693	0.798 ± 0.025
5 µm	0.349	0.302	0.260	0.221	0.189	0.193 ± 0.025

6.3. Simulation of the Write operation

Having clarified the effect of the variables for the Read operation, it is now appropriate to evaluate the influence of the temperature variation caused by the optical write (or pump) pulse. The investigation is now performed on the data reported in Ríos et al. [71]; the change of experimental reference is due to a more accurate characterization of the PPCM device for the purpose of investigating the Write-Erase cycle, with respect to the earlier work referenced above.

The device configuration consists of a 2 µm-long unit cell, distributed across the whole width of the waveguide top surface; the waveguide is a rib SiN waveguide (333 nm, 161 nm etching step), fabricated on a SiO_2 substrate. The measured waveguide width was 1250 nm, and the intended unit cell and cap layer thicknesses were 10 nm each as before, though AFM measurements on the fabricated device yielded a thickness of the unit cell – cap layer stack (after crystallization) of 17.8 nm, with an estimation error of 1 nm.

The aim of this section is to demonstrate successful simulation of the Write-Erase experimental data from [15], performed by delivery of a 14.1 mW / 25 ns Write pulse, and a 14.1 mW / 25 ns – 5.64 mW / 100 ns Erase pulse, as reported in Figure 6.5 (data obtained from Ref. [71]).

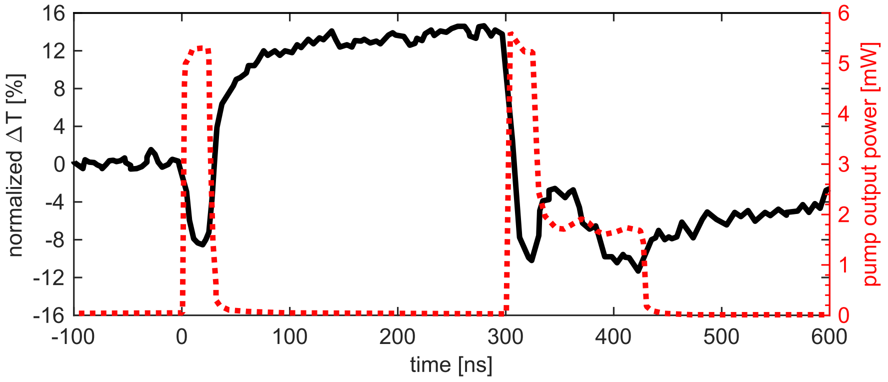


Figure 6.5. Experimental (normalized) transmission (left axis, continuous black line) and output pump power (right axis, dotted red line) measured during Write / Erase on a 2 µm unit cell. Image adapted from Ref. [71]

6.3.1. Read operation and confirmation of the model reliability

The simulation model was adapted to the new device configuration from Ref. [15], with the unit cell refractive index being again assumed to be the one reported in Stegmaier et al. [163]. To determine the effective thickness of the unit cell (i.e. the GST layer) and cap layer, a parametric sweep of the Read operation simulation is performed, varying the ratio between unit cell and cap layer thicknesses; the results are shown in Figure 6.6. The aim is to find the ratio value yielding the experimentally determined transmission value of $\mathrm{T} = 0.46$, for a fully crystalline unit cell.

The results show a matching value of the transmission adopting a ratio value of 0.78, which corresponds to a cap layer thickness of 10 nm and a unit cell thickness of 7.8 nm. Reminding ourselves about the volume reduction upon GST recrystallization (which follows the deposition of amorphous GST), the obtained result is in very good agreement with the experimental data. This result also further confirms the reliability of the Read operation simulation.

By establishing both room-temperature parameters and the PPCM device geometrical configuration, it is now possible to evaluate the PPCM behaviour after delivery of a high-power pump (write) pulse.

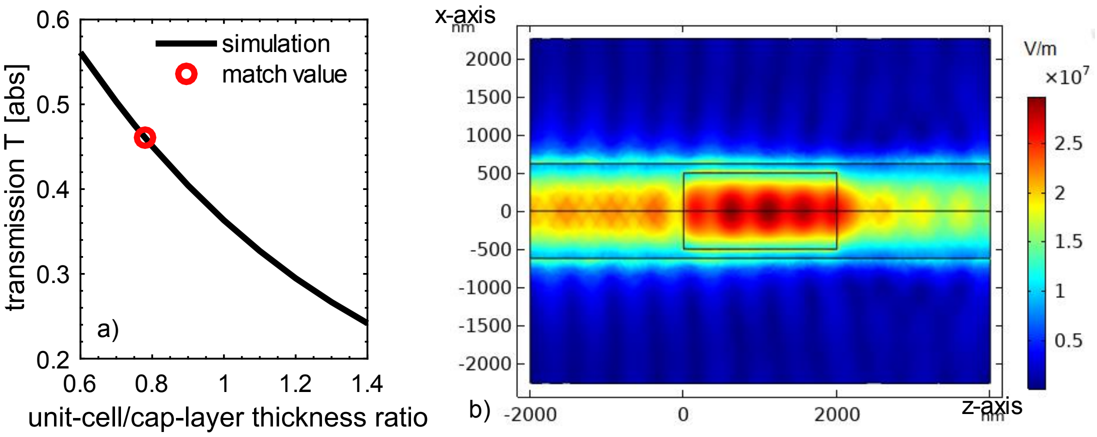


Figure 6.6. Simulation of the Read operation for a 2 μm unit cell. a) Parametric sweep simulated transmission as a function of the ratio between unit cell and cap layer thickness. A matching value to the experimental data is found with a unit cell / cap layer thickness ratio R = 0.79, corresponding to a unit cell thickness of 7.79 nm. b) e.m. field representation (on the waveguide top surface, zx orientation) corresponding to the unit cell thickness of 7.79 nm.

6.3.2. Temperature-dependent optical properties

With the introduction of the temperature field, the temperature dependence of many of the simulation variables (material properties) is ideally required (for a realistic simulation). Unfortunately, several of these dependencies are yet to be experimentally determined. To solve this issue, as previously done it is necessary to rely on an initial set of assumptions, which can then be further refined.

The first assumption on the (complex) refractive index has already been illustrated in Section 2.3.3, which sees the linear extrapolation of the dataset provided by Stegmaier et al. [163] outside the investigated temperature range. The temperature dependency of SiN is also considered [260], seeing a (real-valued) thermo-optical coefficient $\beta = 2.45 \times 10^{-5} [1/\mathrm{C}°]$ (also linearly extrapolated where data is missing). The thermal dependencies of the remaining optical parameters (ITO, SiO_2, Si, air) are instead assumed to be negligible.

6.3.3. Material properties for the thermal model

Having considered the optical and thermo-optical parameters as above, attention is now turned to the definition of the thermal parameters (thermal conductivity, heat capacity, and the thermal boundary resistance, alongside with their thermal dependencies). Density also plays an important role in the determination of heat diffusion and temperature distribution (see Eq. 2.33). Following the analysis of the Read operation, an initial investigation departs from the assumption of reasonable values for the undefined parameters, to then observe the effect on the transmission and temperature upon variation of each of those.

The GST thermal conductivity is generally found to be between 0.16 W/Km and 0.97 W/Km for amorphous and crystalline GST respectively. Thickness dependent effects also see a decrease of the thermal conductivity of very thin crystal GST layers to values close to the ones reported for the amorphous phase (see e.g. Ref. [261]). Those factors hint to a strong dependency upon fabrication conditions (and thermal history) of the thermal conductivity. As no experimentally measured thermal conductivity data is available for the material used on the investigated PPCM device in [15], the values of 0.2 W/Km and 0.58 W/Km are initially assumed, taken from literature [262]. The heat capacity is usually reported to be close to 210 J/K [262], regardless of the phase. The crystalline GST density is reported to be between 6150 Kg/m^3 [215] and 6350 Kg/m^3 [172]; taking the lower value, and assuming a 6.5% decrease of the density upon amorphization [259], the amorphous phase density is determined as 5750 Kg/m^3.

The thermal boundary resistance between GST and surrounding materials is another effect that needs to be taken into consideration (see Section 2.2) but lacks a thorough investigation for the GST/SiN interaction and ITO/GST. The initial investigation here assumes a temperature-independent value of $\mathrm{TBR} = 3.5 \times 10^{-8}$ W/m^{2}K [263], [264] for the GST interfaces. A value of $\mathrm{TBR} = 2 \times 10^{-9}$ W/m^{2}K [265] is also postulated for the SiN/SiO$_2$ interface.

The capping layer (ITO) thermal parameters are also not experimentally measured, but median values from literature are here initially adopted (ρ = 7100 Kg/m^3; c_P = 1290 Kg/m^3; k_{th} = 11 W/Km) assuming that such a thin layer plays a minor role in the overall temperature distribution.

The SiN parameter definition is somewhat more problematic. Particularly for the thermal conductivity, literature data reports values spanning from 2 W/Km to 120 W/Km [266]. The density is another parameter highly dependent on phase and fabrication method, with values reported from 2170 Kg/m^3 [267] to 3170 Kg/m^3 [268] and above. As for the heat capacity, works such as Refs. [269]–[271] highlight quite similar values, (between 650 and 780 J/Kg at room temperature), but also imply a strong correlation to the temperature. Thus, defining the thermal parameters for the waveguide material is quite a challenging task, considering the crucial role of heat sink that the waveguide encompasses. To reduce the complexity of this issue, density and heat capacity are defined here by choice of compatible works investigating SiN platforms for photonics applications. The heat capacity is taken from Ref. [269]. This reference also provides a temperature-dependent description of the thermal conductivity, which sits around 18.4 W/Km at room temperature and reaches 26.5 W/Km at the GST melting point. As for the density, Ref. [272] is chosen, yielding a value of 2750 Kg/m^3, intermediate between the described literature range.

6.3.4. Crystalline unit cell: optical heating and cooling simulations

The effect of the optical heating is here investigated, without focusing on the switching behaviour. A single (rectangular) pulse of 10 ns /14.1 mW is delivered through the waveguide, and a crystalline unit cell is assumed, to mimic the effect of the Write pulse seen in Figure 6.5, and below magnified in Figure 6.7 (from 0 to 25 ns).

The duration of 10 ns is here adopted as the transient is sufficient to flatten the experimental transmission data (see the low variation of ΔT_{norm}, after 10 ns, in Figure 6.7(a)), hence suggesting initiation of the GST melting, which is a convenient indication for the expected simulated temperature profile. The 1/e cooldown transient of 9 ns, also highlighted in Figure 6.7(a), will be discussed in Section 6.5.1. For sake of completeness, a plot from Rios et al. [67] is reported in Figure 6.7(b), corresponding to the application of a 10 ns pulse (although the unit cell size is not specified). This dataset shows remarkable consistency with the dynamics reported in Figure 6.7(a), justifying the focused investigation below.

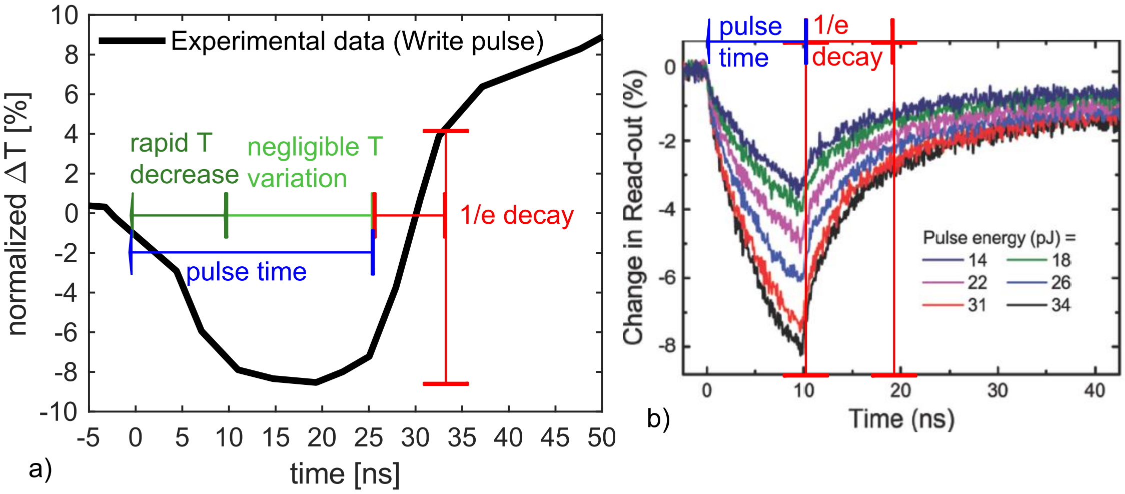


Figure 6.7. Experimentally determined (normalized) transmission data, Write operation transient and cool-down 1/e decay time estimation. (a) Data from Ref. [71]. (b) Data from Ref. [67]

6.3.4.1. Simulation results: thermal parameters influence

A set of simulations was performed by variation of one parameter among the thermal conductivity of silicon nitride (momentarily dropping the temperature dependency), the thermal conductivity of GST, and the thermal boundary resistance. The parameter values are chosen from the various references listed above, or expanding their range to obtain a more thorough insight. The calculation results are reported in Figure 6.8.

The first row of Figure 6.8 shows the PPCM device transmission using different parameters. The red dotted line is common amongst all panels, reiterated to provide a visual reference for sake of comparison. An initial observation outlines the higher impact of a change of the GST TBR value (Figure 6.8(c)), with respect to either the GST thermal conductivity (Figure 6.8(a)) or the SiN thermal conductivity (Figure 6.8(b)). More specifically, the GST thermal conductivity seems to have a rather negligible impact on the transmission, with a maximum variation (amongst the calculated values) $\Delta T < 0.01$. The SiN thermal conductivity seems to play a more important role, particularly as its value decreases. Most of these plots (except the one using low TBR in Figure 6.8(c)) share a common divergence from the experimental data, with the transmission value generally lower than expected.

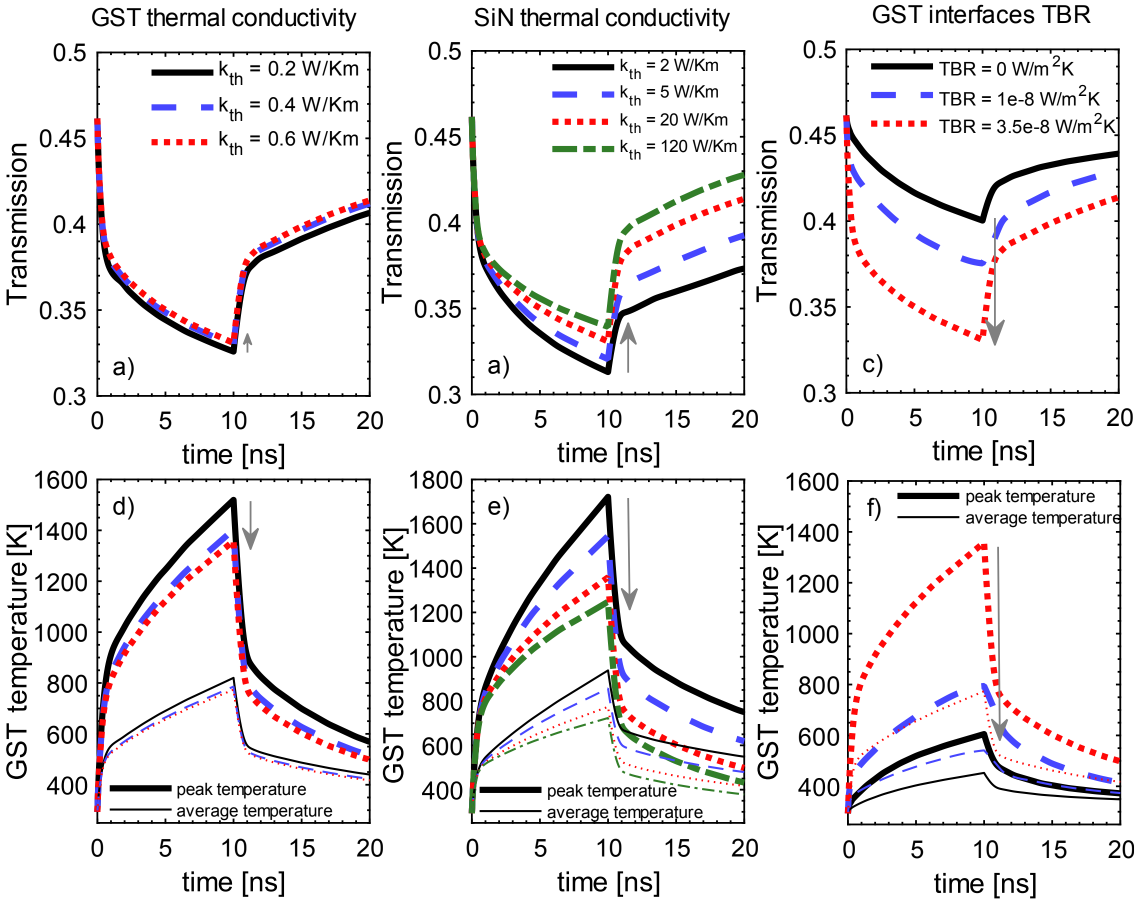


Figure 6.8. Simulation results for the high-energy pump stimulation (14.1 mW, 10 ns) on the 2 μm unit cell (legends on top panels, column-wise). Each column focuses on the variation of a single parameter, specified on top. a-c): Device transmission profiles. d-f): Peak unit cell temperature (thick line), average unit cell temperature (thin lines). In (a-c) and (d-f), the red dotted lines represent the same dataset, to provide a visual clue for sake of comparison. The grey arrows guide the eye from the results obtained with the lowest parameter value to the results obtained with the highest ones, also quantifying the impact of the variation of those.

The attention is now turned to the temperature plots (bottom row), where the unit cell maximum and average temperature are reported (noting that the coordinates of the peak temperature position is not necessarily fixed or shared amongst the various simulations). These show that all configurations not lowering the TBR (Figure 6.8(d-e)) can reach the GST melting point. It is also visible that, once more, the thermal conductivity of GST (Figure 6.8(d)) plays a marginal role; the same is valid for SiN, unless strongly reduced (Figure 6.8(e)), in a similar fashion to what has been observed for the transmission data. The average temperature data provides the additional information that a strong reduction of the SiN thermal conductivity may not be suitable. Using values of 5 W/Km (or 2 W/Km), the maximum average temperature is close to (or above) melting point. By the previous assumption (see introduction of Section 6.3.4), the delivery of a 10 ns / 14.1 mW optical pulse would only initiate melting; therefore, as the average unit cell temperature goes above melting point, the simulation indicates that a low value of SiN thermal conductivity would not fit the expected behaviour.

6.3.4.2. Simulation results: absence of thermal boundary resistance

To further investigate the thermal conductivity and TBR parameters' interplay, and exclude artificially induced effects on the transmission by use of the TBR, a new set of simulations is performed excluding such mechanism (i.e. TBR = 0), along the lines of the experiment shown in Figure 6.8. The results are reported in Figure 6.9.

The results indeed show that, removing the GST TBR from the model, both the transmission modulation and the temperature increase are reduced. Additionally, it is evident how the influence of the variation of the thermal conductivity values plays here a more important role than the one observed in combination with TBR. As an example, the GST thermal conductivity is here responsible for a variation of the transmission modulation up to 3% (Figure 6.9(a)), which is three-fold the one previously observed with a TBR = 3.5 W/m^{2}K (c.f. Figure 6.8(a), where the maximum variation reaches a peak value below 1%). Likewise, the maximum transmission divergence induced by the variation of SiN thermal conductivity jumps from the previous value of 3% (see Figure 6.8(b)) to the current of 8% (Figure 6.9(b)). Similar considerations can be made by observation of the peak and average temperature plots in Figure 6.9(c,d).

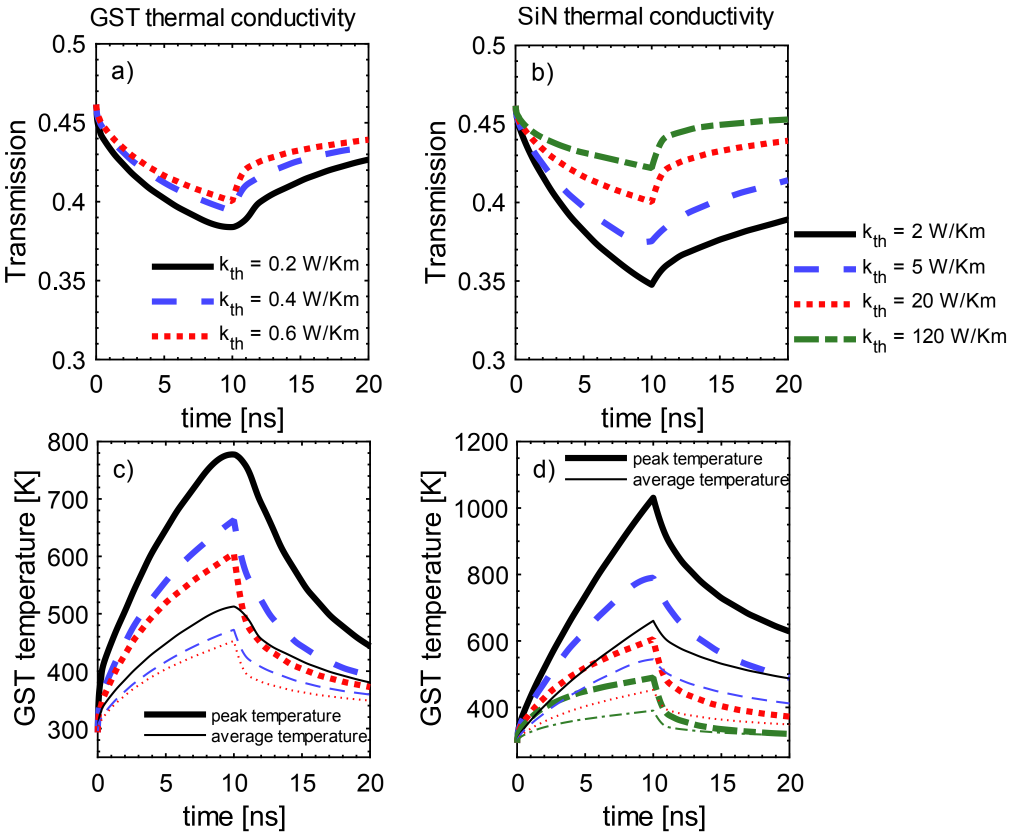


Figure 6.9. Simulation results, assuming absence of thermal boundary resistance on the GST interfaces, using a 14.1 mW, 10 ns pump pulse on the 2 μm unit cell (legends on top panels, column-wise). Each column focuses on the variation of a single parameter, specified on top. A,b): Device transmission profiles. c,d): Peak unit cell temperature (thick line), average unit cell temperature (thin lines). In (a,b) and (c,d), the red dotted lines represent the same dataset, to provide a visual clue for sake of comparison.

This experiment indeed confirms that the simulated system may achieve GST melting temperature also in absence of TBR, by decrease of the thermal conductivities (in particular, the one of silicon nitride), underlining a complex interplay between the thermal conductivity and TBR parameters.

6.3.4.3. Analysis of the optical heating simulation results

However, a comparison between the simulated behaviours with and without TBR also evidences how those parameters shape the heat transient, helping towards a precise choice for the model configuration. To facilitate the comparison, two datasets (one with and one without TBR) are reported in, chosen because of the similar calculated minimum transmission.

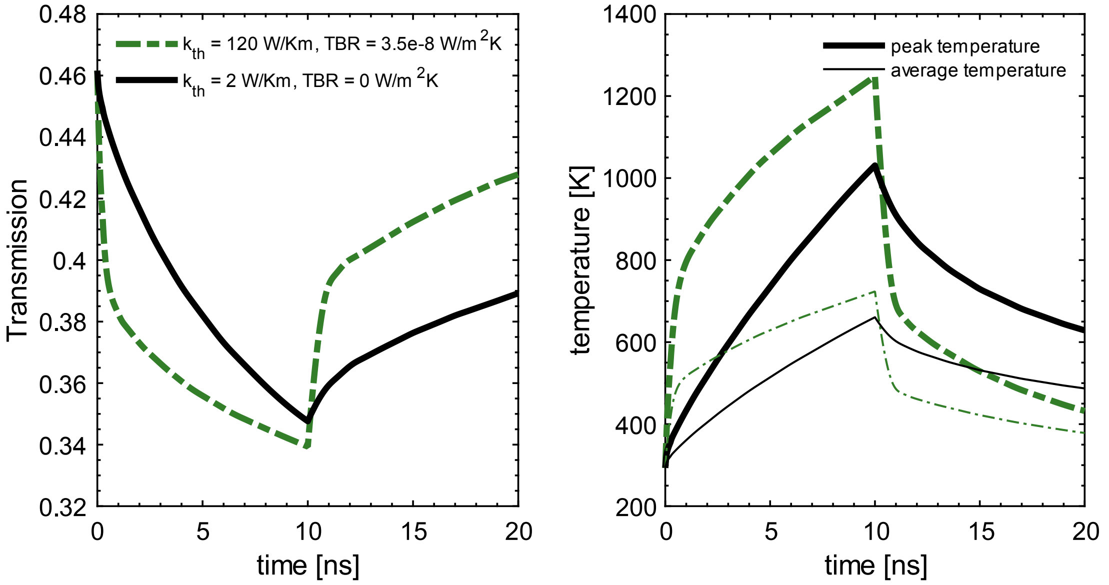


Figure 6.10. Comparison of the transmission (a) and temperature (b) transient after delivery of a 10 ns / 14.1 mW optical pulse to a crystalline, 2 µm – long unit cell.

Despite the similar transmission minima and the reach of melting temperature, the two simulations show remarkably different dynamics. The simulation adopting TBR show i) a remarkably faster transient (both in temperature and transmission, which show an initial fast ramp followed by a shallower variation), ii) higher peak and average unit cell temperatures, and iii) a remarkably similar 1/e transmission decay (equivalent roughly to a 1/e thermal time constant) to the one estimated from the experimental data: specifically, the experimental 1/e decay time is estimated to be $\tau \cong 9$ ns (see Figure 6.7); the TBR dataset in Figure 6.10 (green line) exhibits a similar decay time $\tau = 8$ ns, wheras for the "no TBR" dataset (black line) the same value is $\tau \cong 25$ ns. The characteristics here listed show that the inclusion of TBR in the simulation model should bring significantly closer results to the experimental data, which in Figure 6.7 shows a fast phase-switching transient.

It is found that the use of the SiN thermal conductivity of 20 W/mK results in a 1/e decay time of 9 ns, matching the experimental data of Figure 6.7(a) and Figure 6.7(b), and for such these simulation results orient towards the choice to implement the TBR to fit the PPCM device behaviour.

6.3.5. Full simulation of Write and Erase processes

In discussions thus far, the analysis was concentrated on the unravelling of the influence that the material properties have on both the PPCM transmission and the unit cell temperature. The results indeed capture the expected mechanics of the PPCM device, with the rapid temperature transient associated with the delivered optical energy. However, the quantitative results also point to the mismatch between simulated and experimental transmission at higher temperatures. It is nevertheless possible to perform the phase-change simulation of the full Write / Erase process, to evaluate whether the induced temperature spatial and temporal profile justifies the remaining components of the simulation framework, i.e. material properties, physical mechanisms, and interaction thereof.

6.3.5.1. Numerical experiment description

The integrated FEM-phase-change simulation is carried out on the same device configuration analysed in Section 6.3.4 and described in Ref. [71] (i.e. the 2 μm – long unit cell). The simulation starts at time zero with a fully crystalline unit cell. The Write pulse temporal profile is determined by the nominal 14.1 mW power, and pulse width of 25 ns, with both rise-time and fall-time of 5 ns. At 300 ns, the Erase sequence is delivered, consisting of the combination of the previous Write pulse (14.1 mW / 25 ns) immediately followed by a 5.64 mW / 100 ns pulse, including rise and fall times of 5 ns (with the combined curve smoothed to remove rapid and unphysical variation of the pulse temporal profile).

6.3.5.2. Results and analysis

The results are reported in Figure 6.11, with the transmission data superimposed on the experimentally determined results one for the sake of comparison.

As foreseen, and mentioned in the previous section, the simulated transmission shown in Figure 6.11(a) is not fully capable to fit the experimental data. The simulation yields a consistently lower transmission during the high-energy pulsing, hinting to a deviation of the optical properties (at high-temperature or optical energies) from the assumed ones.

Figure 6.11(b) reports the absorption coefficient of the PPCM device, which shows how roughly half of the propagating optical energy is converted into thermal energy. The second loss contribution is scattering, which accounts for less than -1 dB, and is seemingly not sensitive to the temperature or phase variation. The absorption coefficient provides a clear indication that the simulation model needs to be refined with the provision of further experimental evidence. During the Write operation, the experimental transmission reaches a plateau at 0.42, and assuming a scattering coefficient of 0.08 (confirmed by the separate analysis in Ref. [172]), it points to a maximum absorption of 0.5. The simulated data instead reach an absorption coefficient peak of 0.58. Yet, the simulation model can not produce the same transmission modulation after the Write pulse. An increase of the model temperature by means of raised TBR (or similar mechanisms) would also lead to an increase of the unit cell extinction coefficient, which would consequently lead to a larger transmission discrepancy. Possible answers to this apparent anomaly will be discussed in Section 6.5.1.

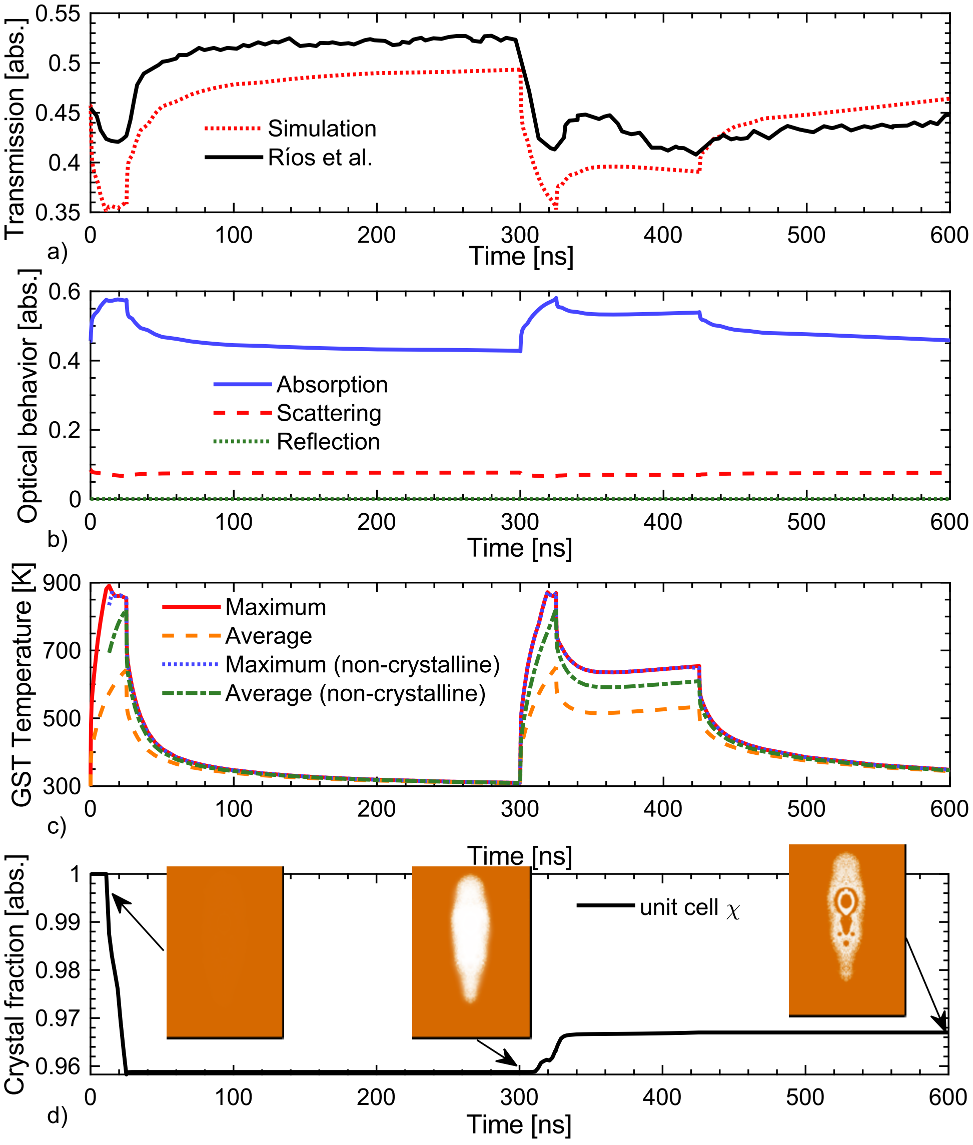


Figure 6.11. Simulation results for the Write / Erase process on a 2 μm unit cell, by use of a 14.1 mW / 25 ns Write pulse and a 14.1 mW / 25 ns, 5.64 / 100 ns Erase pulse, as reported in [71]. a) Transmission, simulated (red dotted line) and experimental (black line, reported from Figure 6.5). b) PPCM reflection, scattering and absorption coefficients. c) Temperature profiles calculated on the unit cell: peak temperature, average temperature, and peak / average temperature of the non-crystal phase (amorphous or molten GST). d) Crystal fraction. Insets report the phase distribution (orange, crystalline; white, amorphous) as viewed from the unit cell top surface.

Whilst simulation and experimental transmission do not agree in absolute terms, it is nevertheless possible to capture the key traits of the PPCM optical behaviour. A nice agreement between experimental and simulated thermal decay is found after both Write and Erase pulses, in the ranges between 50 ns – 300 ns and 450 ns – 600 ns. This illustrates that the optical behaviour at lower temperatures (reported in Figure 6.11(c)) is well captured, and consequently, the assumed thermal parameters are suitable for the simulation task.

The slope mismatch between the simulated and experimental transmission observed during the tail of the Erase pulse (340 ns – 425 ns) can be explained as a lack of optimization of the phase-change model at the temperatures encountered in such a range. By fine-tuning the phase-change model parameters, a faster recrystallization rate can be simulated, which consequently increases the local crystal fraction and therefore the optical absorption. It is nevertheless not possible to exclude the contribution of other factors, as the aforementioned deviation from the assumed GST optical properties, which would influence the PPCM temperature. Therefore, such fit operation would remain a mere exercise without adding to the knowledge of the PPCM behaviour, and the resolution of this simulation mismatch can only be obtained after a thorough investigation over the GST optical properties.

It is interesting to note that a crystal fraction change of 4% is responsible for a relatively large change of the transmission. Whilst the model needs to be further optimized to better follow the experimental data, such a behaviour indeed carry a significant piece of information. The PPCM device finds its sensitivity to the crystal fraction largely concentrated at the centre of the unit cell, because of the modal electric field distribution, as also analysed in Ref. [172]. Therefore, it is reasonable to state that a relatively minor change of the crystal fraction along the symmetry line of the unit cell (as indeed shown in Figure 6.11(d)) leads to a measurable transmission modulation, and the unit cell portions further from such symmetry line only have a minor weight on the device optical behaviour. This is also in agreement with the initial investigation of the unit cell width influence (portrayed in Figure 6.2(b)), which shows a relatively low variation of the transmission as a function of the unit cell width.

6.4. Effect of high-energy probing: tunable volatility

The simulation of the Write and Erase processes above provided a clearer picture of the PPCM physical mechanisms and over some of the material parameters. Whilst there are still unanswered questions over the involved material properties, it is possible to consolidate the reliability of the constructed model via simulation of the PPCM device under different operative conditions.

The PPCM device has been conceptualized to hold its phase information for as long as thermodynamically allowed by the GST unit cell. Nevertheless, novel memory and computing concepts, such as the ones explored in the non-von Neumann architectures, may benefit from reintroduction of a volatility element into the otherwise non-volatile memory concept (i.e. generating short-term memory devices). An example of these concepts is the leaky-integrate-and-fire neuron model, which sees the soma element continuously de-exciting after stimulation, only to trigger the neuronal spike after reaching a predefined temporally-integrated accumulation threshold.

One option to perform such operation with a PPCM device is to use the pump pulse to provide the temporal stimuli, and the probe (or any other continuous signal) to tune the unit cell temperature above room temperature, thus hastening the natural recrystallization process, and yielding the desired tunability. This section explores the simulation of such tunability (or short-term memory effects), attempting to reproduce the experimental results portrayed in Youngblood et al. [254].

6.4.1. PPCM configuration

The device here is fabricated on a rib waveguide of width 1300 nm, thickness 335 nm, and etching step 165 nm. The unit cell configuration consists of a full-waveguide-width unit cell, of length 1 μm, and nominal thickness of 10 nm, with an ITO capping layer of 10 nm. The GST deposition is extended outside the waveguide width; whilst the outside area covered has no role in the PPCM memory functionality, it contributes to the device IL, and thus should be included in the model. The initial transmission T_0 of 0.40 (with a fully crystalline unit cell) is provided [273]. The model is set to the nominal thickness value, yielding a matching transmission value T_0 of 0.395, referred to as $level\ 0$.

6.4.2. Unit cell: amorphous mark definition

The unit cell is subjected to an amorphization process via delivery of a single Write pulse of 486 pJ / 20 ns (combined with the high energy-probe), setting the PPCM state to $level\ 1$ (T_1 of 0.465 at room temperature). As the definition of the amorphous mark via phase-change simulation may yield different transmission outputs, a different strategy is here adopted to define the $level\ 1$ phase distribution. A simple test of the simulation model is performed, by delivery of a 486 pJ / 20 ns = 24.3 mW pulse, added to the continuous probe. The temperature distribution is observed, and the amorphous mark is determined by taking as a crystalline-amorphous boundary the iso-temperature contour which determines the aimed transmission variation $\Delta\mathrm{T} = \mathrm{T}_1 - \mathrm{T}_0$ when tested under continuous probing only. As there is no analytical method to pursue such output, the result is obtained by iterative refinement of the previously determined solution. An illustrative example of the technique and result is reported in Figure 6.12.

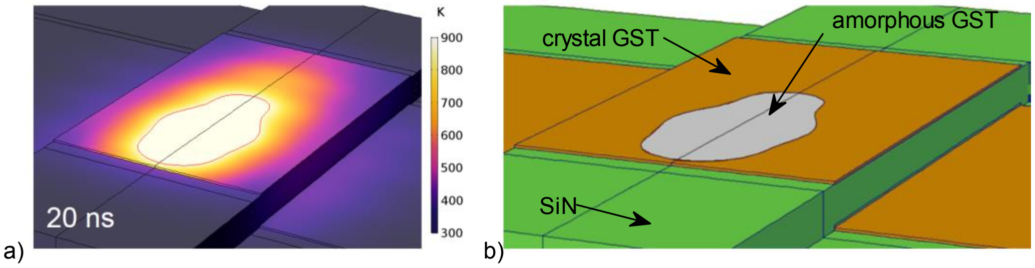


Figure 6.12. Tunable volatility experiment: determination of the unit cell amorphous mark. a) Simulated temperature distribution after delivery of the combined pump and probe signals on the 1 µm unit cell. The melting temperature isothermal line is highlighted in red. b) Definition of the model domains, based on the determined isothermal line which yields the pre-determined transmission $level\ 1$.

6.4.3. Simulation results

The unit cell is subjected to a continuous probe, of constant power, for 3 different simulations, corresponding to the experimental probe powers used in Ref. [254] (i.e. 1.29 mW, 1.84 mW, and 2.36 mW). The aim is to observe the crystallization dynamics which returns the device state from $level\ 1$ to $level\ 0$, and find a possible correlation to the transmission experimental data. Since the thermo-optical effect is shifting the transmission levels (due to the probe-induced temperature increase), it is more convenient to refer to the normalized optical contrast, defined as the normalized variation between T_1 (initial level, partially amorphized unit cell) and T_0 (final level, fully recrystallized), see Eq. 6.2:

$$Contrast = \frac{T - T_0}{T_1 - T_0} \qquad \text{Eq. 6.2}$$

The experimental data are reported in Figure 6.13, as continuous lines. The simulated data points are superimposed for ease of comparison.

The data underlines how the simulation framework is indeed capable of correctly predicting the device optical dynamics at the low temperatures induced by the continuous probe. Each of the emulated data sets is in very good agreement with the respective experimental data, albeit minor deviations are indeed present, as expected from the various elements of uncertainty faced during the simulation framework building. Remarkably, the phase-change simulation parameters did not need any tweak (from those used in previous sections) to fit the experimental data, which further substantiate the use of the proposed techniques and model parameters in these experimental conditions (i.e. relatively low-power stimuli).

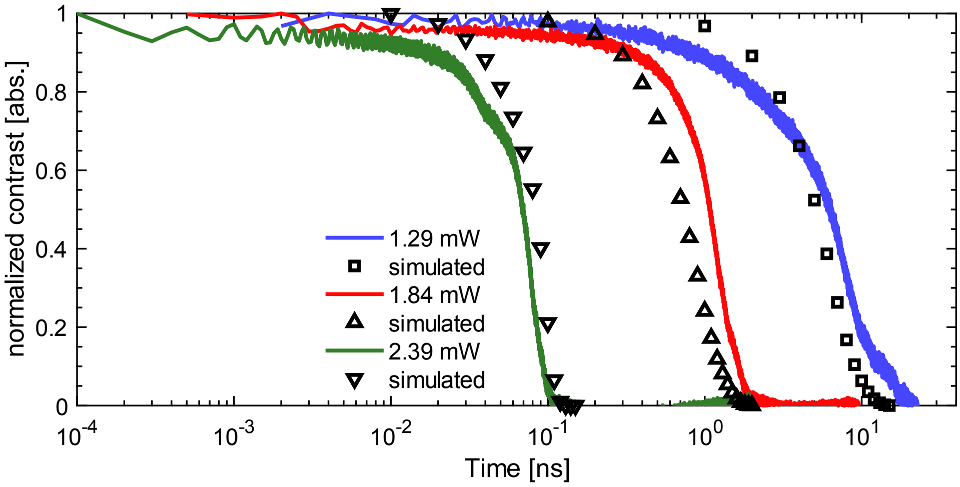


Figure 6.13. Experimental and simulated normalized optical contrast as a function of time, for a 1 μm unit cell under stimulation of a continuous high-energy probe.

6.4.4. Simulation data analysis

Supported by the physically-realistic output, it is then possible to analyse the temperature data in detail. Figure 6.14(a) reports the unit cell peak temperature as a function of time. This figure provides a figure of merit for the temperatures encompassed by the unit cell during the recrystallization time. Each simulated experiment finds a well-defined range of temperatures, justifying the unique crystallization time experimentally found. Specifically, the crystallization time t_χ is defined as the time required for the normalized transmission to return to the (arbitrary) value of 0.1. This value is calculated from both simulated and experimental data. The simulated t_χ is plotted against the inverse of the (simulated) maximum temperature on an Arrhenius plot (Figure 6.14(b)), to highlight any eventual correlation. The relation between simulated and experimental t_χ and probe power is also investigated, and reported on the same figure (Figure 6.14(b)).

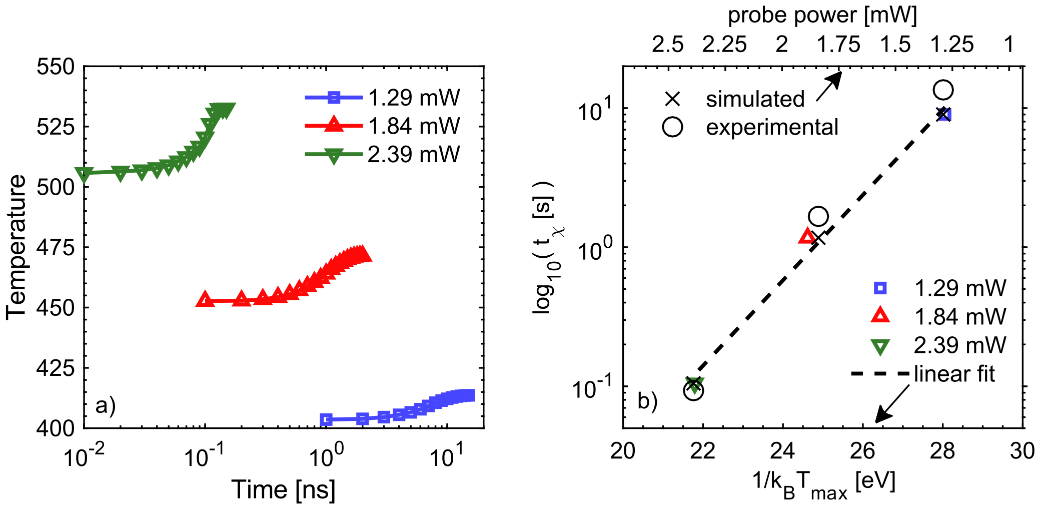


Figure 6.14. Thermal behaviour of the PPCM during high-power probe induced recrystallization. a) Simulated unit cell simulated peak temperature as a function of time, for the three adopted probe powers. b) Analysis of the dependency of the crystallization time t_χ as a function of the inverse of the temperature (bottom axis, coloured markers) and probe power (top axis, black markers). The linear fit is calculated from the inverse-temperature plot.

The plot of (the logarithm of) t_χ versus the inverse of the peak temperature $1/k_B T$ reveals the exponential correlation between the two quantities. This is an expected behaviour, due to the demonstrated Arrhenius crystallization kinetics of GST at temperatures below 650 K [274]. The linear fit of $\log t_\chi$ reveals an activation energy of 0.70 eV. Remarkably, such linear relation is retained when plotting the same data points ($\log_{10} t_\chi$) against the probe power (top axis; data reported as black crosses). Figure 6.13(a) already reported the qualitative match between simulated and experimental optical contrast. The t_χ values calculated from the experimental data fit in very good agreement with the simulated ones (Figure 6.14(b), top axis, data reported as black circles). The additional exponential correlation between probe power and recrystallization time nicely interlaces with the view of the cited Arrhenius dependency and supports the linear correlation between probe power and unit cell temperature. This insight of the PPCM mechanisms allows one to finely tune the retention time of the unit cell information (i.e. crystal fraction), as such behavior is retained outside the investigated range, as demonstrated from additional experiments [254]. Nevertheless, a deviation is expected as temperature further increases, due to the intrinsic non-Arrhenius behaviour of GST at higher temperatures [274].

6.5. Discussion of simulation results and conclusion

As demonstrated, building a suitable PPCM simulation model requires an attentive analysis of its many components. The reported investigation aided towards the definition of a suitable model capable to capture the essential traits of the PPCM optical behaviour, using both high-energy and low-energy signals. An agreeable match with the experimental data is obtained, particularly at low temperatures. Such agreement validates the physics of the simulation framework and validates the ohmic loss nature of the temperature increase and phase transition underpinning the PPCM thermal and optical behaviour. The model also provided a picture of an otherwise impenetrable temperature and phase (time-resolved) distribution, which would help for both optimization and design of improved PPCM architectures. Yet, this chapter outlines also the quantitative mismatch between experimental and simulated data, particularly pronounced under high energy pulsing. Thus, there are still open questions over the full reliability of the implemented model during Write and Erase operations, which is discussed in more detail below.

6.5.1. Hypothesis: non-linear optical effects

The Write and Erase simulated processes deviate from the experimental data in a peculiar way, which may offer an insight into understanding the PPCM optical behaviour in more detail. Under the Write process, the simulation yields a much lower transmission than the experimental data. Yet, after the Write process, the experimental data reports a larger optical contrast, albeit both processes share a similar 1/e thermal decay. The same phenomenon is found again in the Erase process. The simulation output can not be adjusted to the experimental data by tweaking the phase-change model parameters, as increasing or decreasing the crystallization rate would impact both transients, albeit in different manners. Artificially reducing the crystallization rate would initially induce a larger optical contrast, as it would prevent the fast recrystallization observed during the simulation process. At the same time, a larger amorphous mark prevents the inducement of a higher temperature increase, and thus, would also slow the Erase process. Conversely, increasing the crystallization rate would be beneficial to fit the Erase process but, possibly, reduce the amorphous mark size, and hence, lower the already low optical contrast. In simplified terms, it is as if the experimental GST switch operation is "more efficient" than the simulated one.

A possible explanation is found by implying optical nonlinearity of the GST. This material, as well as other PC materials, are known to show quite large third-order optical nonlinear coefficients [275], responsible in general terms for processes such as self-focusing, saturable absorption (SA), and reverse saturable absorption (RSA). Indeed, evidence of some of these effects is published in literature for crystalline GST for the visible spectral range [166]. Yet, there is no direct characterization of such property at 1550 nm. Moreover, quantification of such phenomena for amorphous GST may be difficult to access, as the characterization would be hindered from the fast recrystallization process.

Nevertheless, a piece of evidence which further supports the possibility of SA/RSA behaviour is found in Li et al. [68] (which contains the here analysed data previously published in Ríos et al. [71]). In this work, the authors demonstrate the capability of the programming pulse, or double-step pulse, composed by an initial (high power, low width) optical "spike" followed by a (low power, longer duration) pulse "tail". Such a pulsing scheme is capable of arbitrarily shifting the memory level to any predetermined value, regardless of the initial condition. This implies that the spike component of the programming pulse sets the unit cell state to a "virtual" level, ready to undergo the predetermined crystallization under the programming pulse tail. But the programming pulse can also be used without the second pulse step, and thus, the "virtual level" enabled by the pulse spike corresponds to the high-temperature state (possibly molten phase) of the highest accessible level (here referred to as $level\, n$), as no recrystallization is enacted.

Since the level accessibility is experimentally demonstrated with the same programming pulse regardless of the initial state, it follows that the unit cell temperature can be increased so to access the "virtual level" also when amorphized, i.e. where the memory state is set to $level\, n$. If only optical linear phenomena are present, the unit cell would absorb different fractions of optical energy, depending on the initial condition. Ignoring the thermo-optical effect, the

2 μm fully crystalline cell ($level\ 0$) presented in Section 6.3.4 absorbs 42.3% of the propagating optical energy. The amorphized unit cell ($level\ n$) instead has an estimated[50] absorption coefficient of ~36%. Yet, it is experimentally proven that the "virtual level" is accessible through both initial conditions, without damaging the crystalline unit cell, and leading to extremely similar transition temporal transients (the Write and Erase normalized contrast is reported in Figure 6.15). Indeed, NLO effects on GST can help explain these experimental findings.

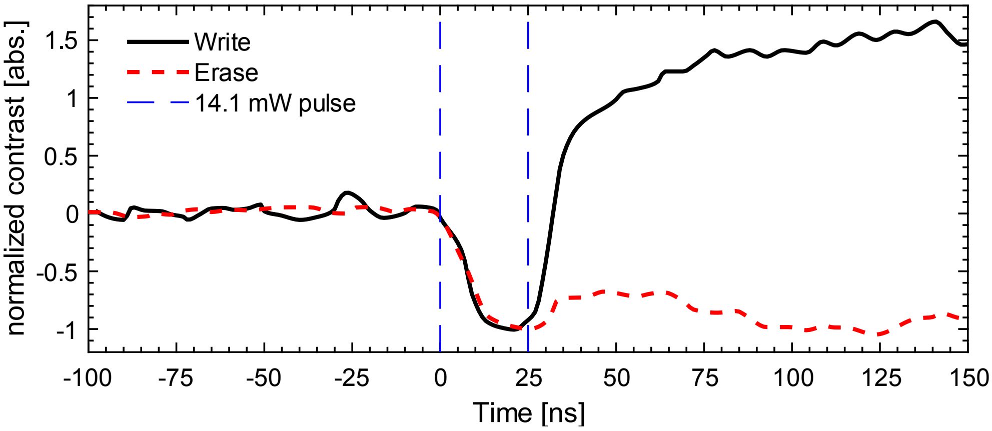


Figure 6.15. Comparison of the normalized contrast data (from Refs. [68], [71]) during Write and Erase sequences, both delivering an initial 14.1 mW pulse. The transient indicates very similar dynamics, regardless of the different initial conditions (c.f. Figure 6.5).

At this moment, the presented simulation framework is not capable to correctly reproduce these peculiar mechanisms. It is here proposed that either SA takes place with crystalline GST, and/or RSA phenomena are found in amorphous GST, possibly alongside other non-linear effects boosting the optical-to-thermal energy conversion efficiency at higher energies (so making such amorphous phase optical behaviour close to the one of the crystal phase). This idea is also explored in García-Cuevas et al. [227], proposing a behavioural simulation framework in which (at higher temperatures and optical energies) the device effective refractive index converges towards a unique value, regardless of the device crystallization level. Indeed, the behavioural model correctly predicts the PPCM behaviour, albeit such a model is built to specifically fit the experimental data without enforcing a physical explanation. Future research may focus on the testing of the GST NLO properties in the infrared, thus providing the needed information to consolidate and quantify, or disprove, such hypothesis.

---

[50] Determined from the experimental data, considering negligible variation of the PPCM scattering and reflection coefficients.

6.5.2. Conclusion

The simulation techniques and results for the operations carried out on the PPCM device have been illustrated.

The Read operation simulation revealed the necessity to precisely quantify the geometrical and material parameters necessary to build the simulation model, highlighting the strong effect of the unit cell optical properties and thickness. Additional pieces of evidence found by literature analysis justify the arbitrary variation of those parameters to fit the experimental data, in particular with respect to the unit cell thickness, which may drastically change upon PPCM use upon crystallization, and due potentially to heat-induced deformation –see Ref. [172]). By reasonable adjustment of the thickness to an inferred effective value, experimental data fit can be successfully and consistently obtained.

The simulation of the Write operation highlighted the requirement of experimental verification for GST and SiN thermal conductivity, the GST interfaces' TBR, and the GST optical properties. The thermal parameters cohesively concur to define the temperature spatial and temporal profile of the unit cell. As for the optical parameters, the adopted parameters do not allow one to correctly reproduce the experimental data during Write and Erase process, hinting to the presence of NLO effects; the literature data reports such effects for GST and other PC materials, and the PPCM experimental data analysis also directly points to such effects.

The simulation model is nevertheless capable of capturing the essential traits of the temperature transient, confirming the assumed physical mechanisms, and producing remarkably accurate results when the PPCM device is operated by use of low power signals. The progressive recrystallization induced by a relatively high power continuous probe (useful potentially for the realization of short-term memory and 'leaky' neuronal type devices) has also been simulated, producing the expected results and further unveiling the relation between provided optical energy and unit cell temperature.

Chapter 7. A novel architecture: plasmonically-enhanced unit cell

Chapter 6 illustrated the progress on the simulation of the performance of the conventional PPCM device, encompassing the elements which drive its functioning and providing a simulation framework capable of emulating its optical and thermal behaviour.

This Chapter illustrates a possible avenue of work pursuing the enhancement of the PPCM performance, reliant on both the acquired knowledge demonstrated in the previous chapter, and the benchmarking tools provided by the simulation framework. The conventional architecture suffers from a physical limitation on the maximum fluence that the unit cell perceives, dictated by the propagation mode and the evanescent field coupling. As it will be here demonstrated, the use of an optimized plasmonic dimer nanoantenna, in conjunction with a PC material inclusion, can in contrast far exceed the conventional PPCM physical limitations, drastically increasing the interaction between propagating mode and a much-reduced volume of PC material, while maintaining comparable optical contrast. This in turn leads to the extreme decrease of both the temporal and energy requirement of Write and Erase operations.

7.1. Introduction: use of optical resonant structures

The research on the PPCM device demonstrated it to be a viable and flexible solution for fast[51], non-volatile, all-optical memory purposes [67], [68], [173]. Yet, a pursuit of memory performance enhancement is always beneficial. A possible strategy consists of the implementation of the unit cell within optical resonant structures, whose optical behaviour can benefit from the additional tunability offered by PC materials, and in turn, may magnify light-matter interaction via the exploitation of resonant modes, improving the switching energy efficiency.

7.1.1. Dielectric resonant structures

Experimental works merging resonating photonic structures and PC materials technology have been carried out since the original PPCM concept proposal, mostly using dielectric resonant structures (micro-ring resonators [66], [67], [175], [276], [277] or photonic crystals [174], [278], [279]). The micro-ring resonator in particular is a consolidated technology, which can be easily combined with PC materials and offers several potential advantages:

---

[51] Write and Erase operations on conventional PPCMs operate with pulse widths of tens and hundreds of nanoseconds respectively; Read operations usually requires timings in the ps range.

- Potential to achieve low IL (when correctly optimized)
- Absence of additional or challenging fabrication processes
- Owing to the micro-ring spectral selectivity, it adds WDM functionality to the PPCM concept
- Increase the achievable optical contrasts between phases
- Increase of the Write/Erase energy efficiency.

These advantages arise in micro-ring resonators (and, in general, in photonic resonant structures) due to i) the e.m. field spatial confinement, and ii) the increased life-time of the resonant mode, which can be estimated from the Q-factor as $\mathrm{Q} = \omega \, \Delta\tau \approx \lambda/\Delta\lambda$ (see also Section 5.2.1).

However, an intrinsic drawback of dielectric integrated photonic structures lays in their footprint, which usually finds its feature size $\mathcal{F}$ in the order of tens to hundreds of micrometres, so as to not induce high IL. Such a large footprint is detrimental for high density integration, and thus increases costs and/or layout challenges faced during the PIC design. Moreover, the coupling between unit cell mode and propagating mode still exploits the evanescent field mechanism, and for such, does not deliver to the unit cell optical intensity values larger than the peak input intensity.

7.1.2. Hybrid configuration: implementation of metallic nanoantennas

A bypass to these limitations is the use of plasmonic resonant structures, capable of squeezing light further down to sub-wavelength $\mathcal{F}$-sizes. Despite being generally ruled out due to CMOS compatibility concerns, recent publications do in fact demonstrate the feasibility of the integration of plasmon-supporting metals with the integrated photonic platform at a proof-of-concept stage [280]. It is also noted that such structures are often characterized by their sub-wavelength sizes, which can be tackled via use of electron-beam lithography techniques but do constitute a barrier towards scalability and cost control. However, advanced optical projection techniques such as immersion EUV lithography would eventually be capable to supply for the required minimum feature size and crytical dimensions [281], [282], with an even better outline with the forecoming developments [283].

Therefore, this work proposes the implementation of a plasmonic nanotantenna, being initially fabricated on the waveguide top surface. The combination of plasmonic resonant nanoparticles with the conventional photonic integrated platform has already been numerically [284]–[288] and experimentally [289]–[296] explored, with the view to analyse and turn to applicative advantage the nanoantennas capability to i) couple efficiently with the propagating mode to establish the nanoantenna eigenmode; ii) induce local e.m. field “hot-spots”, i.e. spatially confined increases of the e.m. field magnitude; iii) control the optical energy emission directionality, and thereby, potentially modulate the waveguide transmission. However, the tunability of nanoantennas in PIC devices remains a particularly under-explored concept.

A physical trait that can be turned to advantage is the sensitivity of the resonant structure eigenfrequency to its chemical/material environment[52]. Specifically, PC materials placed in proximity of nanoantennas can intrinsically provide the tunability functionality, owing to their phase-switching mechanism. Thus, this Chapter explores the unprecedented combination of such class of devices with PC materials in PIC devices, with the scope to pursue the non-volatile and tunable modulation of the plasmonic resonator coupling strength (and therefore the device transmissivity), also drawing advantage from the device capability to magnify the local e.m. field within a highly-confined region. This strategy permits one to reach the intended main objectives: the magnification of the interaction between optical mode and PC material, the reduction of the device $\mathcal{F}$-size well below the diffraction limit, the unit-cell volume shrinkage, and, importantly, the consequent increase of the memory speed and energy performance.

7.2. Investigation of the plasmonic device concept optical behaviour

It is convenient to focus first on a simplified nanoantenna test design, to characterize the desired modulation of the waveguide transmission and then infer the effect of the nanoantenna configuration on its optical behaviour. To simplify the comparison with the works illustrated in Chapter 6, a SiN 330 nm × 1300 nm rib waveguide (etching step of 165 nm) is here chosen, so as to maintain the same photonic platform and focus attention to the device optical behaviour only. The wavelength of choice is also 1550 nm, unless otherwise specified.

---

[52] The nanoparticles dielectric function, along with the nanostructure geometry and symmetry, defines eigenfrequencies, eigenmodes, and of course resonance Q-factors (see e.g. Ref. [86])

7.2.1. Test device: single bar nanoantenna

An initial test device consists of a single bar, of section 40 nm (width) × 30 nm (thickness), and of variable length, placed on the waveguide top surface with its variable length orthogonal to the waveguide main axis[53], see Figure 7.1 The material of choice for the nanoantenna is Ag, as this material has excellent plasmon-supporting properties, and is an already consolidated solution in the plasmonics field using wavelengths in the IR spectral range (see e.g. [297]). The optical properties for Ag are initially taken from Johnson et al. [83], for a complex index value of $\tilde{n} = 0.14447 - 11.366i$.

The numerical experiment here focuses on the investigation of the scattering parameters and nanoantenna absorption. The results are illustrated in Figure 7.1(a), as a function of the bar length, whose investigated range spans from 50 nm to 1300 nm (the latter corresponding to the waveguide width). The investigation here carried out will also focus on an additional element, the *Enhancement Factor* EF, defined as the ratio between the squared spatial average of the electric field norm $|E_{loc}|$ (calculated at the edges of the bars[54]), and the square of the peak input electric field norm $|E_{in}|$ (calculated at the model input port):

$$\mathrm{EF} = \frac{1}{V_{gap}} \int_V |E_{loc}|^2 dV \cdot \frac{1}{|E_{in}|^2} \qquad \text{Eq. 7.1}$$

It is immediately visible from Figure 7.1(a) how the device transmission T is (predictably [285]) highly modulated as a function of the bar length. It is possible to identify two distinct dips of the transmission at 272 nm and 928 nm, whose major contribution is the scattered energy S, also plotted in Figure 7.1(a) (green dash-dotted line). This confirms that the optical energy is indeed harvested from the nanoantenna, to be redirected out-of-plane. The absorbed component of the optical energy A (red dashed line) plays a secondary role on the device optical transmission, with two minor A peaks also placed in correspondence of the scattering ones. The reflection coefficient R (blue dotted line) finds similar characteristics, at a further reduced magnitude.

---

[53] It is here noted that for this configuration, as well as for all the following, the vertical edges of the investigated structures are smoothed so to provide a minimum curvature radius of 10 nm, as to avoid to focus on theoretically available but physically inaccessible optical performance. Whilst the choice is arbitrary, it relates to the tolerance over the feature size conceptually approachable during fabrication with the available facilities, and negates singularities of the emulated results.

[54] Note that the volume over which the retrieval of E_{loc} is performed is defined as a would-be disc of radius 25 nm and thickness 30 nm, centered at the bar edge, excluding the overlapping volume between disc and bar

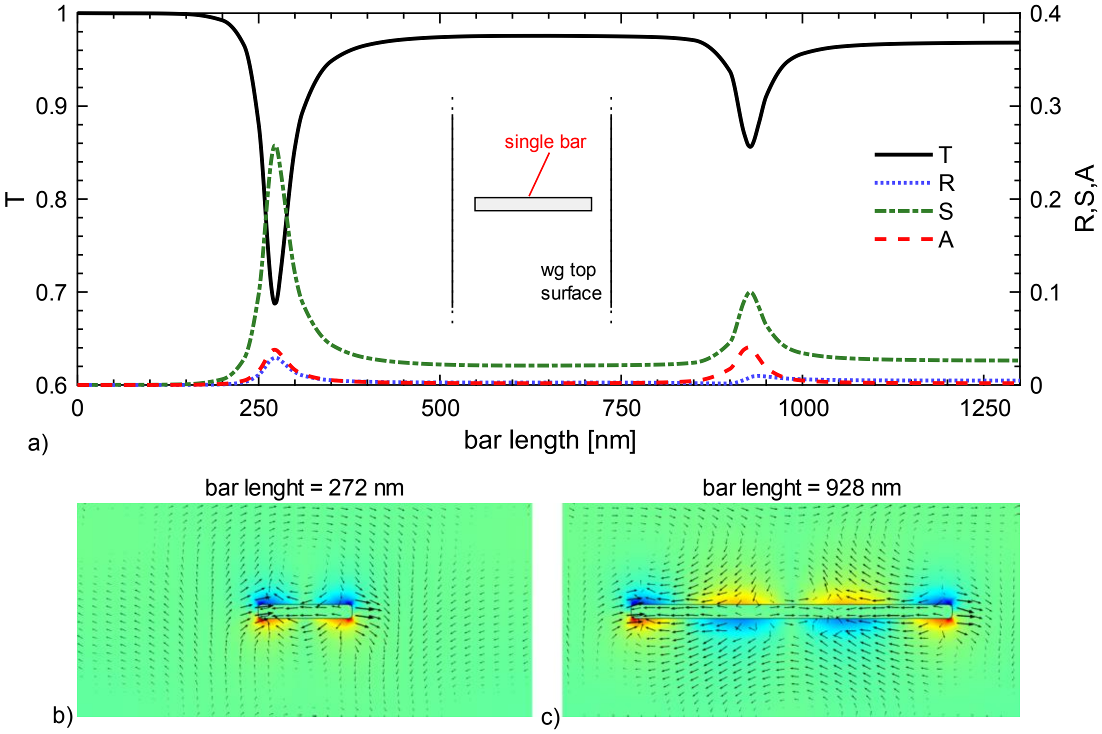


Figure 7.1. Numerical analysis of the single bar nanoantenna structure (width of 40 nm, thickness of 30 nm, variable length). a) device optical behaviour, as a function of the bar length: transmission T (left axis, black continuous line); reflection R, scattering S, Ag absorption A (right axis). Inset shows a schematic of the analysed configuration. b,c) electric field on the orthogonal cut plane passing through the bars mid-point (the width of the plot corresponds to the waveguide width). The z component of the field is colour-mapped (blue representing negative values; red for positive values); arrows represent the x and y components (log scale).

The electric field distribution on the bar mid-point cut plane (parallel to the SiN platform) is reported in Figure 7.1(b-c), for the bar lengths corresponding to the two peak locations. The first solution (Figure 7.1(b)), retrieved at 272 nm, identifies the expected distribution of a dipolar resonance, with localized electric field peaks at the bar short edges (see arrows directions). The value of the EF here spikes to 209, proving the efficiency of such a nanoantenna to harness the optical energy and magnify the e.m. field in proximity of its edges. The second solution (retrieved at 928 nm, see Figure 7.1(c)) corresponds instead to a quadrupolar resonance mode, and sees the formation of additional e. field nodes within the bar length. At this configuration, an EF value of 82 is calculated, confirming that also the quadrupolar resonant mode is capable of beneficially harnessing the optical energy (albeit at reduced efficiency).

7.2.2. Test device: dimer bar nanoantenna

A second configuration is here tested, proposing a dimer nanoantenna consisting of two bars placed symmetrically with respect to the waveguide main axis, maintaining the same orientation as the single bar case (orthogonal to the waveguide main axis). A schematic representation of the proposed design is reported in Figure 7.2(a) inset. The numerical investigation of a similar dimer configuration [288] highlighted the extreme increase (up to 3 orders of magnitude) of the irradiance within the nanogap location. The tested gap size g is fixed as $g = 10$ nm. The numerical results are displayed in Figure 7.2.

The transmission data shows that the optical behaviour of the dimer configuration is quite similar to that of the single bar configuration (reported as a grey line in the figure). Scattering and absorption also show similar trends to the previously illustrated configuration. The first resonant mode, at a bar length of 227 nm, finds a transmission minimum value of $\mathrm{T} = 65.25\%$. The related e.m. field distribution is reported in Figure 7.2(c) (c), which confirms the same nature of the dipolar resonance observed on the single bar. The quadrupolar resonant mode is again confirmed from the e.m. field distribution retrieved at 560 nm, and reported in Figure 7.2(d). Figure 7.2 also reports the plot of the norm of the fields (Figure 7.2 (e) and (f) for the dipolar and quadrupolar modes, respectively). These plots help to visualise the formation of the e. field hot-spot within the nanogap, in both the dipolar and quadrupolar resonant mode; as the plots share the same normalized colour scale, the plots also underline the higher efficiency of the process for the dipolar resonance mode.

The EF is here reported as an inset (Figure 7.2(b)). This plot quantifies how the e.m. field value within the nanogap is remarkably higher than the one previously calculated at the single bar edges, corroborating the findings published in Jin et al. [288]. A maximum EF value of 1340 is found at 229 nm, at the dipolar resonant peak. Whilst the quadrupolar resonant mode has a minor effect on the transmission (c.f. Figure 7.2), the EF value here retains a remarkably high value

($\mathrm{EF} =$ 114 at 560 nm), with a magnitude comparable to the dipolar peak EF calculated for the single bar.

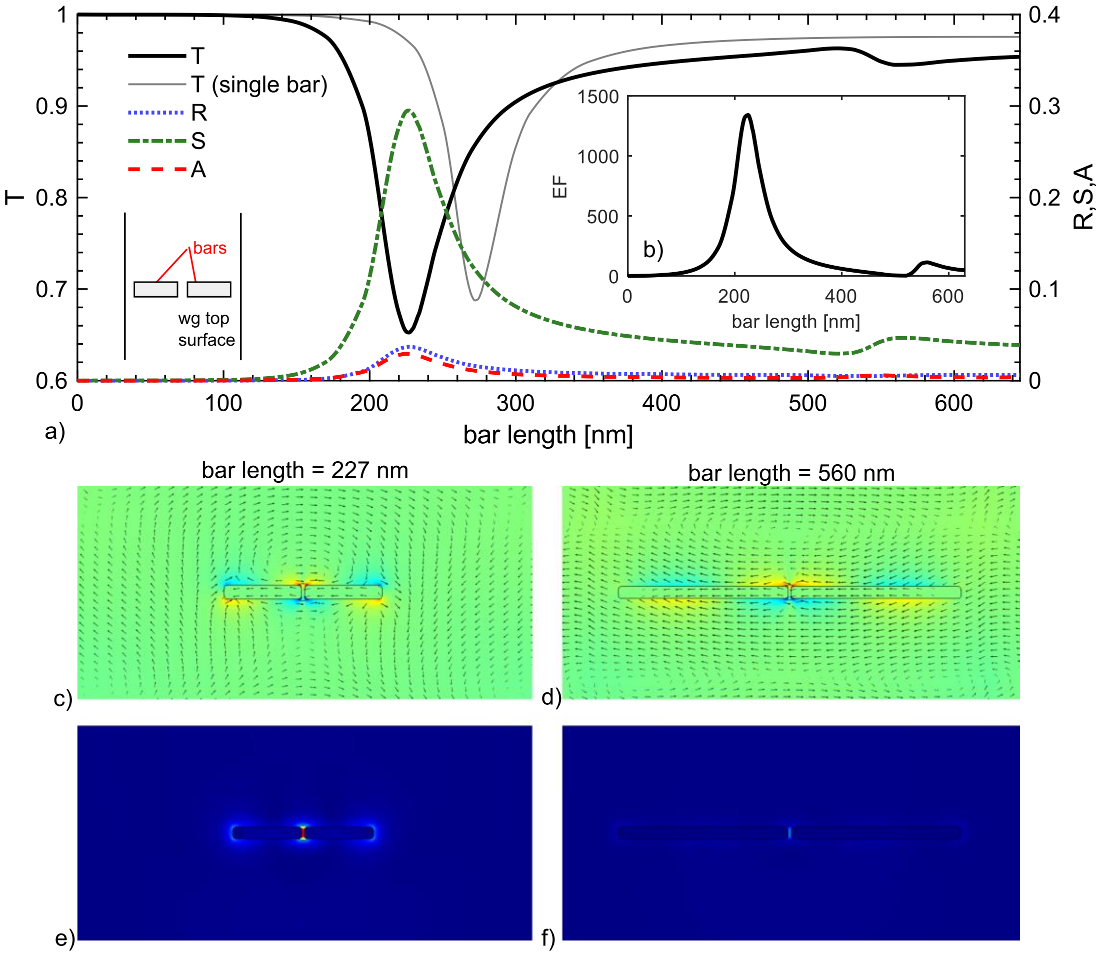


Figure 7.2. Numerical analysis of the dimer bar nanoantenna structure (width of 40 nm, thickness of 30 nm, gap distance 10 nm, variable length). a) device optical behaviour, as a function of the bar length: transmission T (left axis, black continuous line), compared to the previous result (grey line); reflection R, scattering S, Ag absorption A (right axis). Inset illustrates the device schematics. b,c) electric field on the orthogonal cut plane passing through the bars mid-point (the width of the plot corresponds to the waveguide width). The z component of the field is colour-mapped (blue representing negative values; red for positive values); arrows represent the x and y components (log scale). d,e) norm of the electric field (blue to red, normalized colour scale) corresponding to the solutions plotted in (b) and (c) respectively.

7.2.3. Nanoantenna cross-section and nanogap sizes

The effect of the gap size is here analysed, with the scope to quantify its influence over the device optical behaviour. To limit the complexity of the task, the analysis is here focused on the dipolar resonance[55].

---

[55] The following results use a bar width of 30 nm in place of the previously adopted $w = 40$ nm, hence the deviation from the results reported above.

7.2.3.1. Gap distance

An initial simulated parametric sweep was carried out varying the gap distance g parameter (between 5 nm and 40 nm); the results are reported in Figure 7.3, which evidences how the increase of g leads to an increment of the transmission modulation, although the increment reduces as g increases. The EF sees instead an inverse proportionality to the gap size, with the peak EF value strongly and monotonically decreasing as g increases. This observation well relates to the picture of the EF being inversely proportional to the resonance modal volume [86], [298]. Thus, it is expected that such trend maintains validity for extrapolated values at higher gap sizes g.

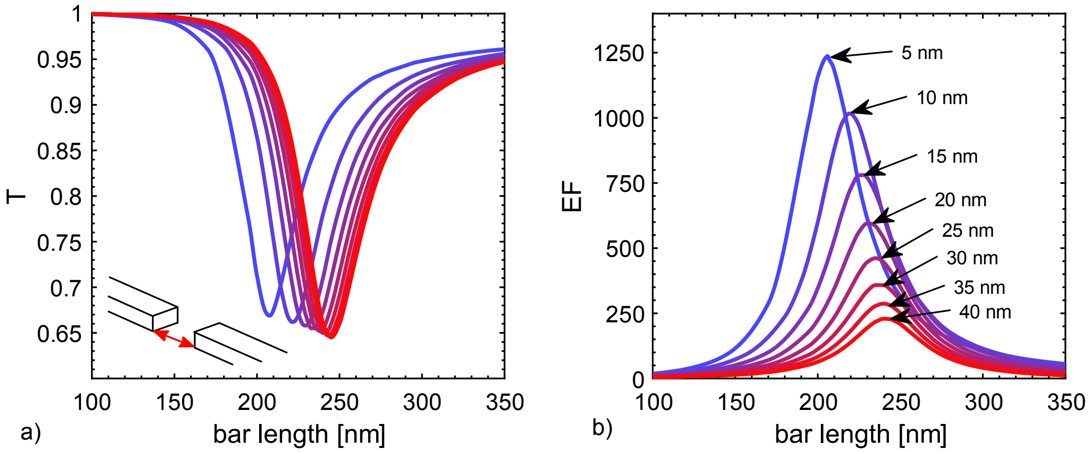


Figure 7.3. Effect of the gap distance variation upon device transmission (a) and enhancement factor (b). Gap distance g values are indicated in (b).

To further confirm such effect, two additional parametric sweeps are performed, varying instead the effect of the bars' height and bars' width, in place of the gap size. It is necessary though to highlight how those two parameters not only vary the optimum resonant configuration (i.e. bars' length) but also affect the nanoantenna scattering cross-section, thus a less trivial relation between parameter variation and transmission is expected.

7.2.3.2. Bars' thickness

Results related to the thickness of the bars are reported in Figure 7.4. The investigated thickness range spans from 10 nm to 50 nm. As expected, it is here visible that the increase of the nanoantenna thickness correlates with a variation of the transmission. It is here noted that also the gap vertical size follows the same variation as the thickness of the bars.

At the shortest thickness, the shallowest T modulation is found. Such modulation increases up to $t = 40$ nm, after which the device T modulation again decreases. As for the EF value, a thickness increase corresponds with a marked EF decrease.

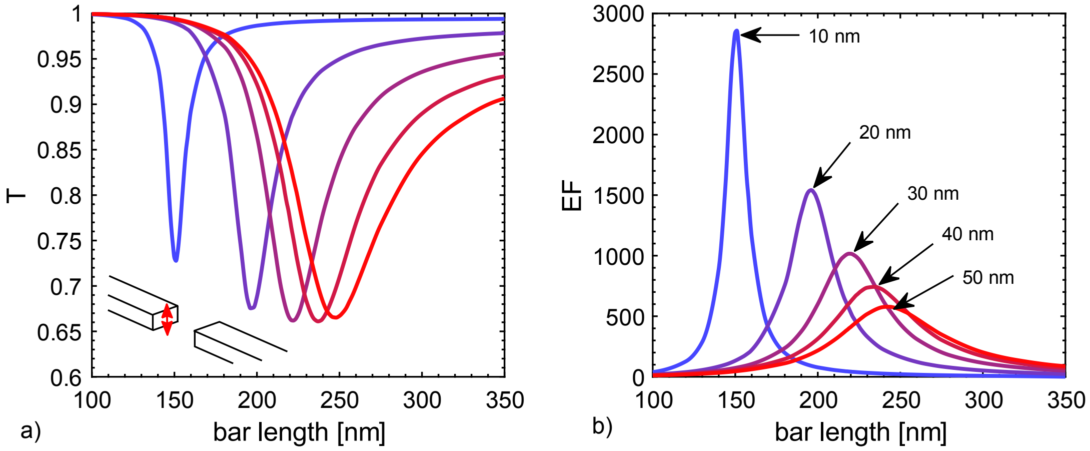


Figure 7.4. Effect of the bars thickness variation upon device transmission (a) and enhancement factor (b). Bars thickness t values are indicated in (b).

7.2.3.3. Bars’ width

The simulation results for the bar width variation effect are reported in Figure 7.5. These plots again show a strong correlation between achievable transmission modulation and bar width w parameter.

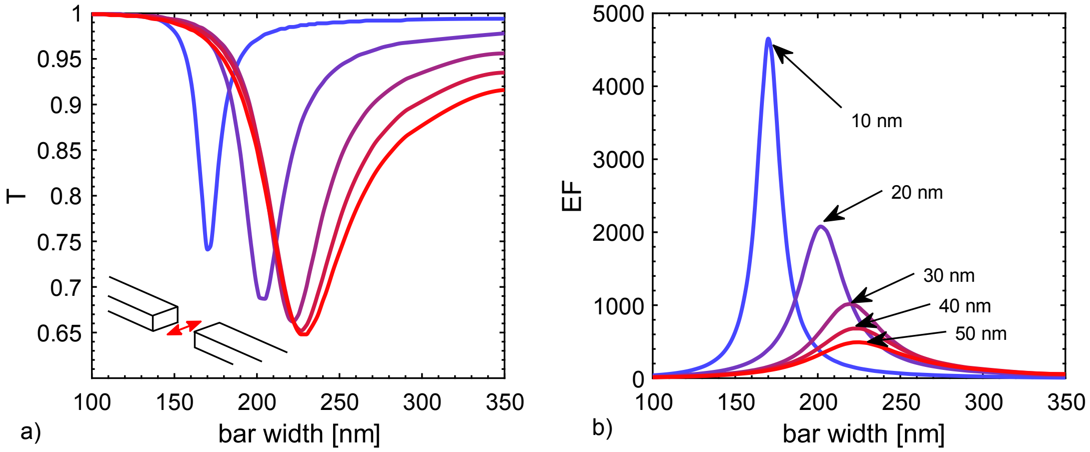


Figure 7.5. Effect of the bars width variation upon device transmission (a) and enhancement factor (b). Bars width w values are indicated in (b).

As w is increased, the transmission minimum also decreases in a monotonic fashion. As for the enhancement factor, the same trend is observed, in a rather similar fashion to what has been already demonstrated in Figure 7.4(b). It is worth nothing though that the calculation of EF in this case needs to account for the accommodation of the modal volume onto larger longitunidal sections, as the bar width is increased.

7.2.3.4. Results analysis and discussion

Here the attention is brought to the minimum T values and maximum EF values calculated before, as the device optical modulation and energy harnessing efficiency are the two most beneficial characteristics (in terms of device performance) pursued in this work. These values are plotted onto a normalized parameter scale in Figure 7.7, with the normalization here effected by the common values of $g = 10$ nm, $t = 30$ nm, and $w = 30$ nm.

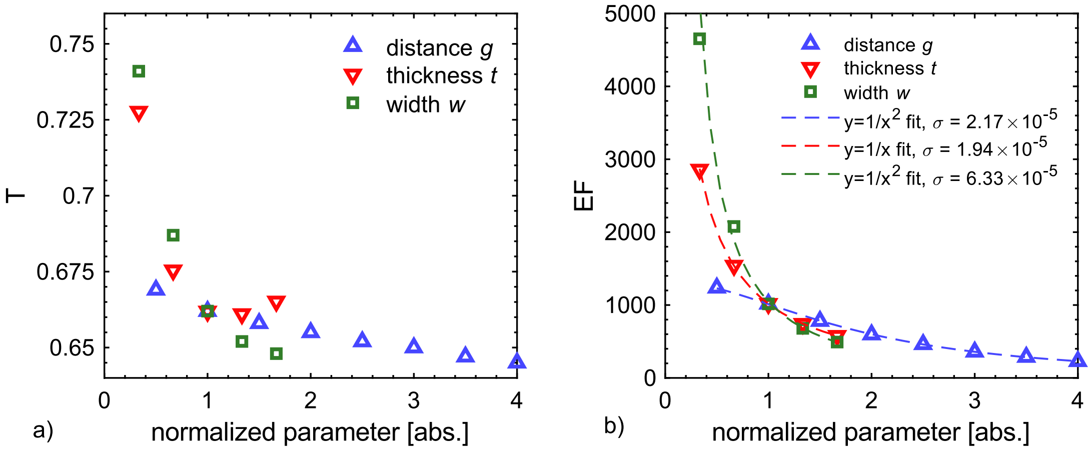


Figure 7.6. Summary of the variation of minimum transmission T (a) and (inverse of) maximum EF (b), calculated at dipolar resonant configuration, as a function of the investigated parameter (normalized scale): gap distance g, bars thickness t, and bars width w.

With respect to the optical attenuation, the results show a general decrease of T as the parameters increase (Figure 7.6(a)). Gap distance and width follow this trend throughout the whole investigated range, inducing a monotonic decrease of the calculated T_{min}. The thickness however finds a minimum value around $t = 40$ nm, after which T_{min} increases. These results show that, to obtain the maximum T modulation, the gap size reduction is generally beneficial, but a blind decrease without informed optimization is not an inherently successful option.

The inverse value of EF is plotted against the normalized parameter scale in Figure 7.6(b). EF clearly follows an inverse quadratic dependence on the gap distance g and width w, and an inverse linear dependence on bar thickness. The three fit lines are also plotted in Figure 7.6(b), reporting also the low value of the fit standard error σ. With respect to the width dependency, it is here again noted that the calculation took into account the rescaling of the integration volume (along the width variation direction) proportional to the width variation. In case of fixed integration volume, a minimum of EF would be located halfway through the investigated range, as EF would increase either using narrower nanoantenna features (at low w) or restricting the integration volume at its center (at high w). This underlines how the EF estimation here carried out would only provide a qualitative figure-of-merit of the nanoantenna energy-harnessing efficiency.

7.2.4. Effect of a GST inclusion

The dimer configuration demonstrated a higher e. field peak (located at the nanogap) than the single bar configuration. It is then a natural choice to test the device optical behaviour with a GST inclusion within the device nanogap[56]. The assumption is that the nanoantenna optical response would display considerable differences for amorphous and crystal GST, as the different nanogap refractive index n_{gap} would alter the plasmonic nanoantenna resonant response. Thus, the phase transition of GST would provide the tunability element previously missing in the nanoantenna-only configuration.

The numerical analysis previously illustrated in Section 7.2.2 is here repeated, separately introducing the optical properties of amorphous and crystal GST, to test the proposed hypothesis. The gap distance g and bar thickness t are here initially maintained to $g = 10$ nm and $t = 30$ nm. The model includes GST within the volume previously considered to evaluate the EF parameter, i.e. a disc centred at the nanogap of radius r_{gap} = 25 nm, excluding the volume occupied by the Ag dimers[57]. A schematic representation of the proposed configuration is reported in Figure 7.7. The analysis is here again carried out focusing on the dipolar resonance, and the simulation results are reported in Figure 7.8.

---

[56] The suggested fabrication strategy adopts ad-hoc collimators for the GST sputtering deposition [340].

[57] Of course, the GST deposition would occur on the top surface of the Ag resonators as well. Since the field enhancement effect occurs predominantly at the nanogap location, the GST material on the top surface of the resonant nanostructure has a largely negligible effect on the device optical performances, and is therefore here excluded for sake of simplification.

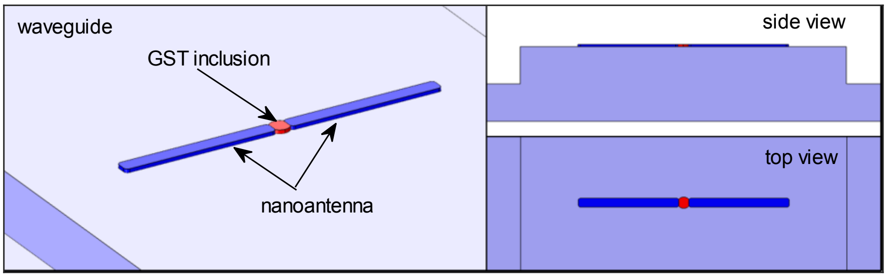


Figure 7.7. Schematics of the GST-implemented dimer bar nanoantenna.

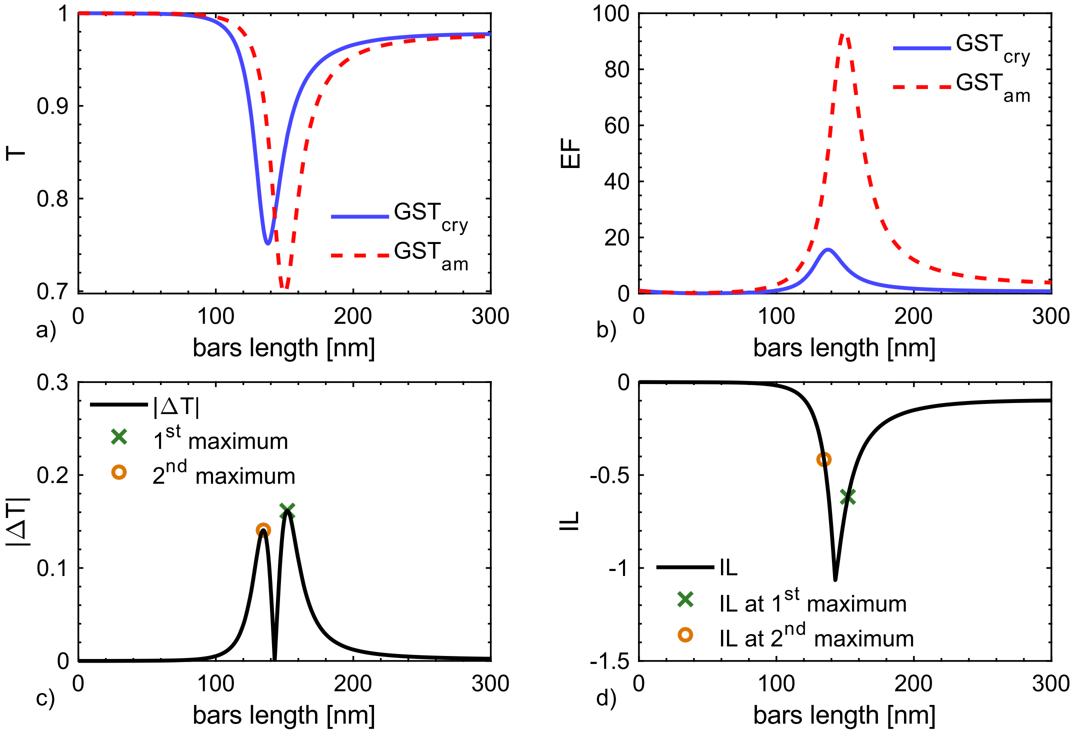


Figure 7.8. Simulation results for the amorphous and crystalline GST-implemented dimer bar nanoantenna, as a function of the length of the bars. (a) Transmission T. (b) Enhancement factor EF. (c) Transmission difference $|\Delta\mathrm{T}|$ calculated from the data reported in (a). (d) Device Minimum insertion loss IL; the two marks highlight the values calculated at the two peaks locations found in (c).

The comparison of the transmission of the amorphous and crystalline GST configurations (Figure 7.8(a)) confirms the expected notable variation. Whilst the transmission minimum for crystalline GST is found at a bar length $b = 135$ nm, the transmission minimum for the amorphous GST is found at $b = 151$ nm. The differences in the optical response of the two configurations provide the desired optical modulation and enables the optical memory functionality by tailoring the device optical attenuation via phase switching.

As for the EF (Figure 7.8(b)), the GST inclusion diminishes its previously observed exceptionally high value, which is indeed an expected behaviour[58]. Nevertheless, EF still reaches remarkably high values, up to 15.64 for the crystal GST, and 93.61 for the amorphous GST. To underline the importance of this result, it is here reiterated that the calculated EF correlates with the ratio between the average irradiance in the nanogap I_{loc}, and the input irradiance I_{in}. Thus, the GST inclusion sees an effective power $P_{eff} \propto I_{loc} = \mathrm{EF} \times I_{in}$, and consequently, suggests a proportional reduction of the optical switching energy requirement.

The difference $|\Delta\mathrm{T}|$ between the transmission values in the two configurations is plotted in Figure 7.8(c). The data shows that two possible solutions maximize the theoretical $|\Delta\mathrm{T}|$, each one close to the aforementioned transmission minima. A maximum value of $|\Delta\mathrm{T}| = 16.16\%$ is found at the bar length of $b = 152$ nm, involving a IL contribution of -0.62 dB, calculated from the crystal GST T data. A second maximum of $|\Delta\mathrm{T}| = 14.02\%$ is found for $b = 135$ nm, which decreases the theoretical IL value to -0.45 dB (calculated from the amorphous GST T data).

Concluding, the proposed device is indeed sensitive to the optical properties of the material introduced within the nanogap. After switching of the GST phase, the device sees a marked modulation of its optical attenuation. Consequently, this device can encode the optical memory information onto the GST crystal fraction, comparably to the conventional PPCM device. The sizeable e.m. field increase within the GST-filled nanogap has a dual benefit, as it magnifies the resonant mode sensitivity to the change of the nanogap optical properties, and increases the effective optical power delivered to the GST inclusion.

7.3. Nanoantenna design optimization

Before investigating the proposed plasmonic device's thermal and switching performance in memory applications, a more thorough investigation of the device design and configuration need to be pursued. Thus, here various nanoantenna materials, designs, and configurations are investigated, to i) clarify the picture of the investigated phenomena, ii) provide a perspective of the available optimization parameters, and iii) gain a deeper understanding of the boundaries of feasibility and applicability of the proposed plasmonic memory concept. To provide a consistent basis for the comparison of the results, in this section the nanoantenna thickness is fixed to $t = 30$ nm, whilst the gap distance is increased to $g = 40$ nm, (aiming towards more feasible fabrication process). For ease of comparison purposes, the optical characteristics of the dimer-bar configuration are reiterated in Figure 7.9 (for a bar width of 40 nm).

---

[58] This effect can be understood in terms of the constitutive relation of the electric field continuity across the interfaces, which rescales the displacement vector by the dielectric function, along with the e.field suppression dictated by the higher κ of GST when compared to the air-filled gap.

It is also here remarked that the analysis here pursued considers a single nanoantenna element on a waveguide. Appendix IV instead expands the analysis to include the results obtained via simulation of multiple nanostructures along the same waveguide, with the view to outline an alternative and viable investigation route to pursue further memory performances improvement.

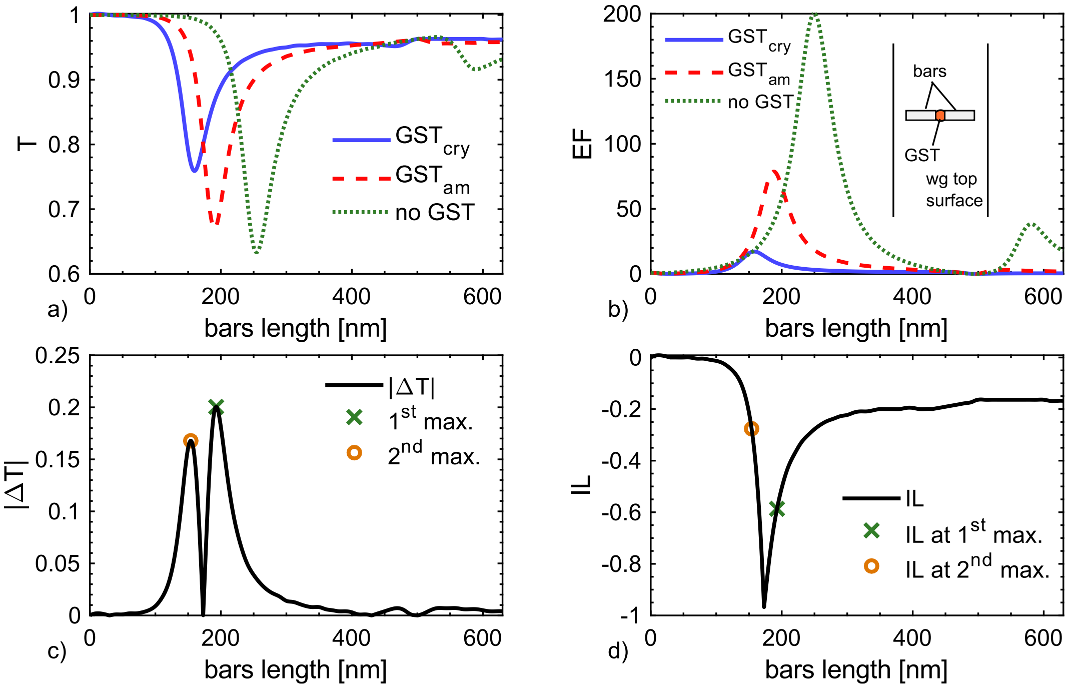


Figure 7.9. Simulated optical properties of a dimer-bars nanoantenna, of width 30 nm, and thickness 30 nm, as a function of the length of the bar. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $|\Delta\mathrm{T}| = |\mathrm{T}_{cry} - \mathrm{T}_{am}|$. (d) Theoretical minimum insertion loss, $\mathrm{IL} = 10 \cdot \log_{10} \max(\mathrm{T}_{GST})$. The markers underline the values calculated at the $|\Delta\mathrm{T}|$ maxima found in (c).

7.3.1. Alternative nanoantenna geometries

7.3.1.1. Bowtie nanoantenna

The optical properties of an Ag bow-tie nanoantenna are here analysed (see Figure 7.10 for the graphical representation). The structure consists of two prisms (with an equilateral triangle base) mirroring each other along the waveguide main axis, with their corners closing towards the central nanogap. A smoothing radius of 10 nm is also taken into consideration, and the resulting gap change is compensated accordingly. The simulation results are presented in Figure 7.11.

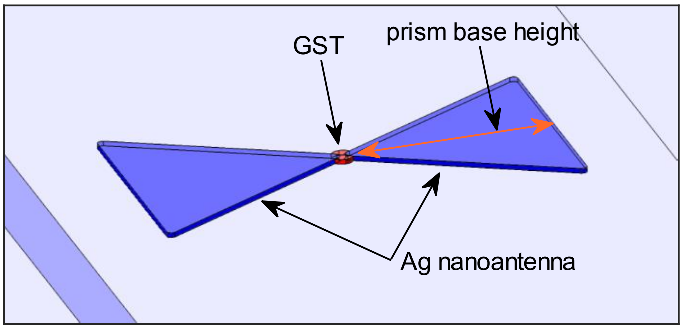


Figure 7.10. Schematics of the bow-tie nanoantenna device, fabricated on the waveguide top surface.

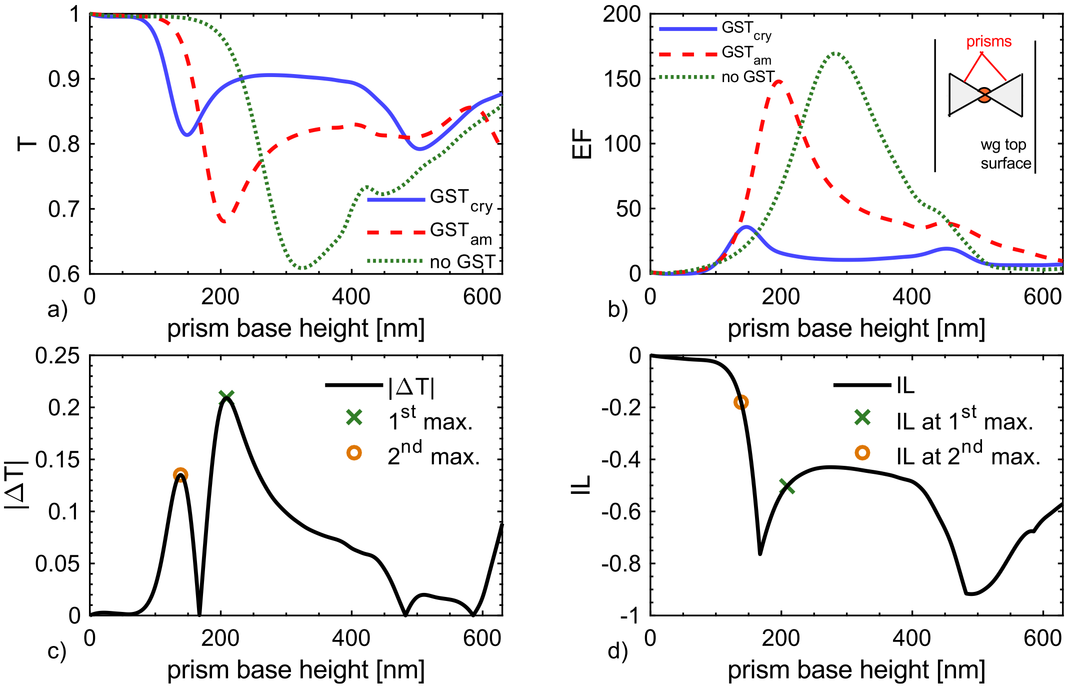


Figure 7.11. Simulated optical properties of a bow-tie nanoantenna, of 30 nm, as a function of the prism base (equilateral triangle) height. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $|\Delta \mathrm{T}| = |\mathrm{T}_{cry} - \mathrm{T}_{am}|$. (d) Theoretical minimum insertion loss, $\mathrm{IL} = 10 \cdot \log_{10} \max(\mathrm{T}_{GST})$. The markers underline the values calculated at the $|\Delta \mathrm{T}|$ maxima found in (c).

The simulation yields an optical behaviour similar to the dimer-bar configuration. Figure 7.11(a) reports the transmission data. It is here observable that the transmission modulation reaches similar values to the bars design, both in the non-GST implemented and GST-implemented configurations. The EF data (Figure 7.11(b)) underlines a beneficial increase of the e.m. field in the nanogap, possibly due to the narrow features of the nanostructures at the nanogap location. As a result, the EF values are generally higher than in the bar configuration (although this is valid only for the GST-implemented configuration). Yet, the resonance data features have a narrower separation than in the bars configuration, and the calculation of $\Delta\mathrm{T}$ (Figure 7.11(c)) yields a generally worse result, additionally inducing an increased IL (Figure 7.11(d)). Moreover, after the resonance feature (i.e. at larger sizes values), the scattering component remains higher than previously observed, possibly due to the larger scattering cross-section of this nanoantenna design.

Whilst the optical performances are reduced, this nanoantenna design may find suitable application where the maximization of the EF is required. Optimization of its additional configurational degree of freedom (i.e. the nanogap bow angle) may also bring beneficial results.

7.3.1.2. Dimer-disc nanoantenna

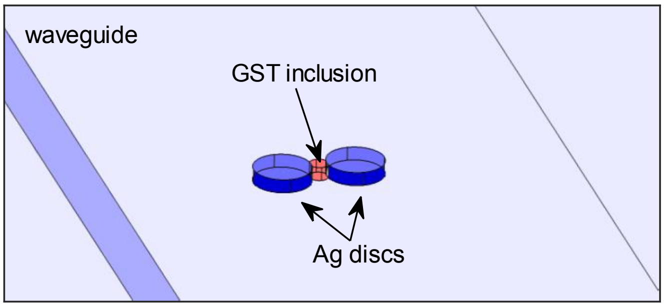


Figure 7.12. Dimer-discs nanoantenna schematics.

The dimer-disc nanoantenna (schematics shown in Figure 7.12) presents a far-simpler design than the previously considered ones (i.e. removing all high-curvature corners from the device geometry). Yet, by appropriate design of disc radii, while fixing the nanogap distance to the value used before, this nanoantenna design demonstrates comparably beneficial optical characteristics. The transmission data (see Figure 7.13(a)) of the nanoantenna without a GST inclusion shows a marked modulation as the diameter increases. The transmission rapidly decreases beyond a diameter value of 450 nm, supporting the increased scattering cross-section explanation provided for the bow-tie nanoantenna. After GST inclusion, the resonance optimal diameter again shifts towards smaller diameters, here in a particularly efficient fashion. The lack of the smaller nanoantenna features of the bow-tie design is possibly responsible for the lower EF here estimated (see Figure 7.13(b)). Yet, even in this case, peak values of $\mathrm{EF}_{cry} = 20$ and $\mathrm{EF}_{am} = 56$ are calculated, achieving the aimed increase of the optical fluence. Although the transmission modulation is also decreased in comparison to bars and bow-tie designs, the relatively larger shift of the resonance feature produces a comparable $|\Delta\mathrm{T}|$ profile (see Figure 7.13(c)), demonstrating its suitability for the proposed memory architecture even in this simplified configuration.

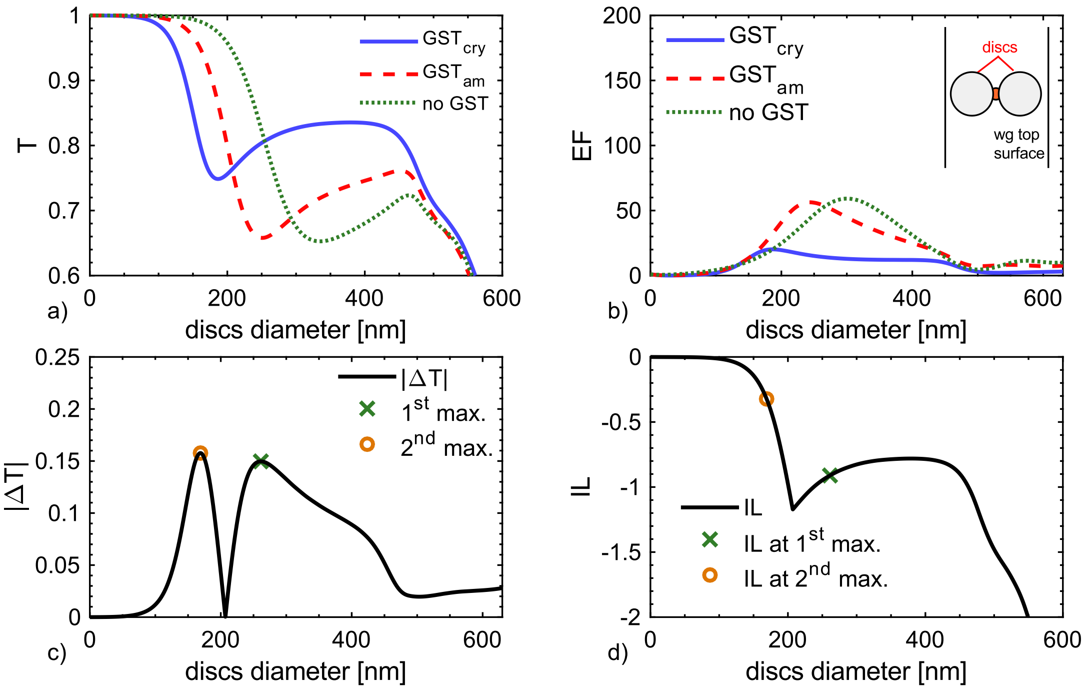


Figure 7.13. Simulated optical properties of a dimer-discs nanoantenna, of 30 nm, as a function of the diameter of the discs. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $|\Delta\mathrm{T}|$. (d) Minimum insertion loss IL. The markers underline the values calculated at the $|\Delta\mathrm{T}|$ maxima found in (c).

7.3.2. Encapsulated configuration

The memory Write operation leads to GST melting. In the proposed configuration, no mechanism would impede the molten phase flow outside the preconfigured nanogap, if melting takes place throughout the whole unit cell. As for such, it is here analysed a configuration adding an encapsulation layer onto the whole nanostructure, to prevent this type of unit cell damage. The scope is to determine the optical behaviour of the plasmonically-enhanced unit cell after a change of its chemical environment, allowing an informed estimation of the most suitable cap material.

The proposed cap structure is a 60 nm thick rectangular parallelepiped, whose orthogonal section adds 30 nm to any of the nanoantenna boundaries (see Figure 7.14 for a schematic representation). The nanoantenna geometry is fixed to the dimer bars configuration, with $g = 40$ nm, $w = 40$ nm, and $t = 30$ nm.

The investigated capping materials are SiO_2, Si_3N_4, and Si, selected so as to vary the cap structure refractive index.

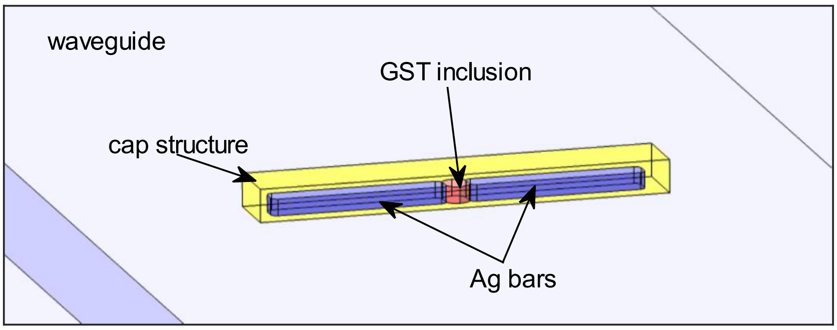


Figure 7.14. Schematics of the encapsulated dimer-bar unit cell.

The results of the device simulated optical properties, as a function of the length of the bars, are reported in Figure 7.15. It is here possible to note how the nanostructure encapsulation has a mildly detrimental effect on the device optical performance, with a decrease of the transmission T modulation (Figure 7.15(a)), the EF (see Figure 7.15(b)), and also on the achievable theoretical $|\Delta\mathrm{T}|$ (Figure 7.15(c)).

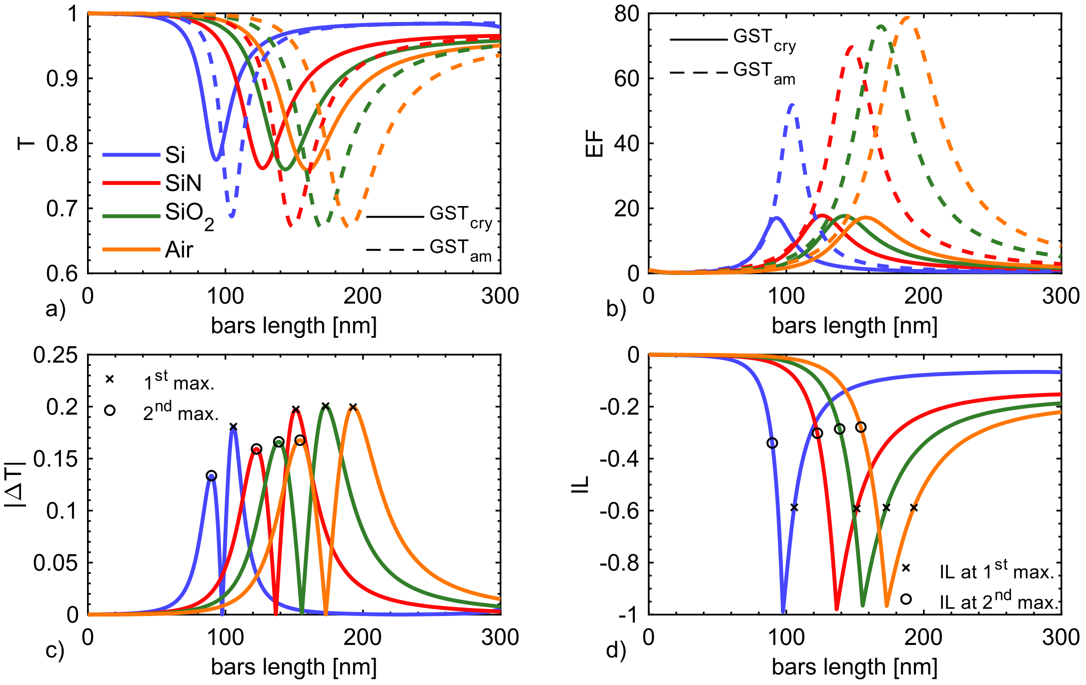


Figure 7.15. Optical characteristics of the encapsulated plasmonically-enhanced unit cell (dimer-bars configuration), as a function of the bars' length. The color legend is reported in (a) only. (a) Transmission for crystalline (continuous line) and amorphous (dashed line) GST inclusion. (b) EF values, calculated at the GST inclusion volume. (c) Transmission difference $|\Delta\mathrm{T}|$, with the two peaks of each individual line highlighted by a cross and a circle. (d) Minimum IL, with the IL values calculated at the $|\Delta\mathrm{T}|$ peaks highlighted as in (c).

I The IL data is reported in Figure 7.15(d). This plot also underlines that the best performance is obtained with the non-encapsulated configuration, and increasing values of IL are then obtained using SiN, SiO_2, and eventually Si.

n conclusion, unless the case of Si is considered, the $|\Delta\mathrm{T}|$ and IL values at resonance vary within 5% from the non-encapsulated configuration. Therefore, the simulation reveals that this strategy can indeed be effective to negate the underlined source of unit cell damage, without interfering with the device operativity. As for the EF, the crystal case sees negligible variation of its peak value even in the Si configuration, whereas the amorphous GST case sees an appreciable decrease of the EF even after use of low-index capping materials (albeit still reaching peak values above 55). This observation suggests a decreased recrystallization energy efficiency of the encapsulated configuration, yet not negating the superior optical performance in comparison to the conventional PPCM architecture.

7.3.3. Waveguide-embedded configuration

Insofar, the investigation assumed the fabrication of the plasmonically-enhanced unit cell on the top surface of an already existing photonic platform, as for the conventional PPCM. The initial results evidenced how this is indeed an interesting avenue of work; yet, the proposed device still relies on the evanescent field coupling to excite its resonant mode, which was previously underlined as one of the PPCM limiting factors towards improved energy efficiency.

Here the optical response of the device is investigated after bringing the nanostructure closer to the otherwise natural coordinate of the peak e.m. field, i.e. by embedding the plasmonic resonator within the waveguide, as previously performed for a simplified nanoantenna in Castro-Lopez et al. [285]. The hypothesized fabrication process would see the localized etching of the waveguide, followed by fabrication of the plasmonically-enhanced unit cell, and re-deposition of the core (and cladding) material. Alternatively, one may proceed with the fabrication of the plasmonically-enhanced unit cell as an intermediate stage during the SiN layer deposition, to only then proceed with the remaining stages of the PIC fabrication (we note that SiN CVD techniques temperatures can be compatible with the incorporation of Ag or GST inclusions, see e.g. Kou et al. [299]).

The results below are calculated fixing the gap distance $g = 40$ nm, and the thickness $t = 30$ nm. The device bottom surface is assumed to be positioned at the etching step height. The investigation here focuses on the optical performance of the dimer-bars configuration; an illustrative representation of the bars-embedded device is provided in Figure 7.16

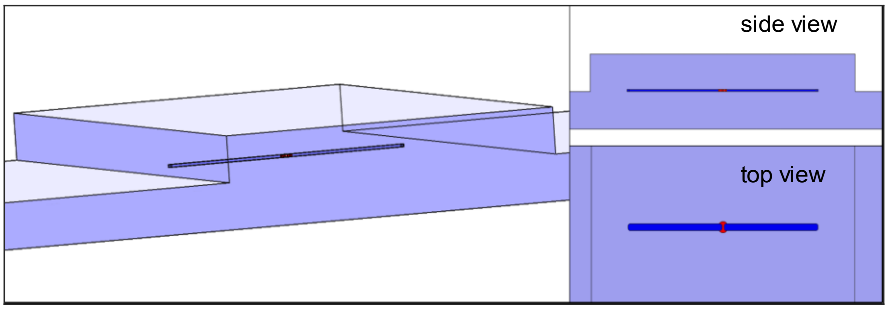


Figure 7.16. Schematics of the embedded dimer-bars device design.

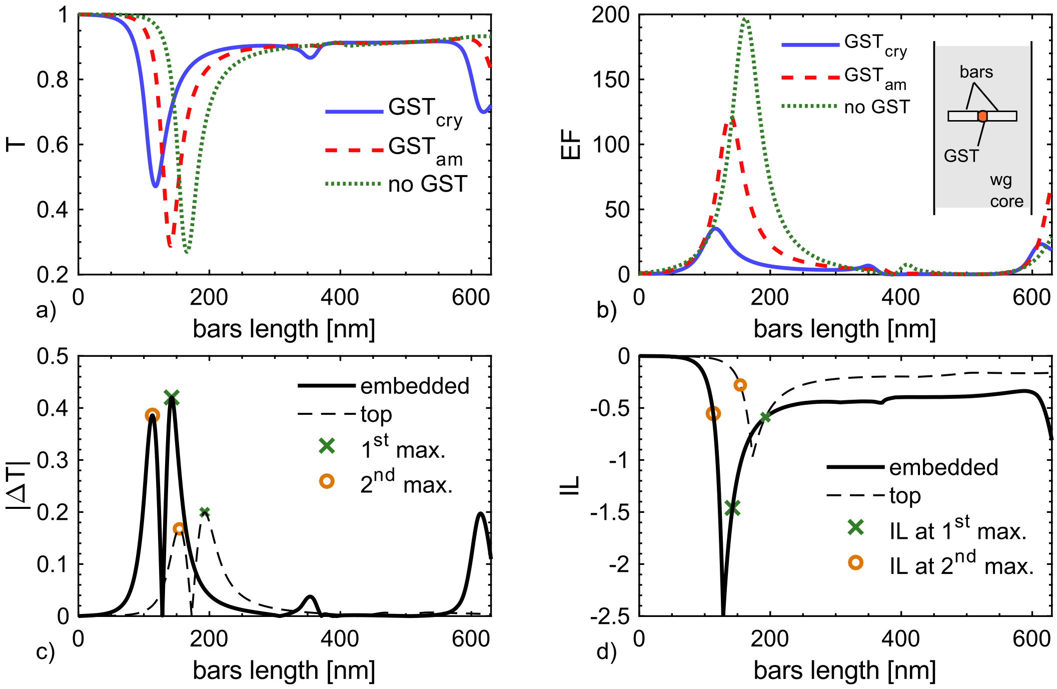


Figure 7.17. Simulated optical properties of a dimer-bars nanoantenna, of width 30 nm, and thickness 30 nm, as a function of the length of the bars. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $|\Delta\mathrm{T}|$; a comparison with the on-top configuration is provided (d) Theoretical minimum insertion loss, $\mathrm{IL} = 10 \cdot \log_{10} \max(\mathrm{T}_{GST})$. The markers underline the values calculated at the $|\Delta\mathrm{T}|$ maxima found in (c). Data of the on-top configuration is also provided in (c) and (d), for sake of comparison.

The bars' design simulation results (Figure 7.17) shows a dramatic improvement of the optical performance in the embedded configuration. The modulation of the transmission (Figure 7.17(a)) reaches for the nanoantenna-only configuration a minimum of 0.27, maintaining the previously observed narrow resonance feature.

The same is observed after inclusion of GST, with minima of 0.48 and 0.28 for the crystalline and amorphous phase respectively. The EF plot (Figure 7.17(b)) also shows a beneficial increase of the peak values at resonance with respect to the on-top configuration, when GST is included. The combination of the transmission data highlights a much improved theoretical optical contrast (Figure 7.17(c)) , with a peak value of 38.6% at the bars' length of 114 nm, for which an IL value of -0.55 dB is calculated. An even higher $|\Delta\mathrm{T}|$ value of 42.2% is calculated at 142 nm, although at the cost of an increased $\mathrm{IL} =$ -1.46 dB. Thus, at the expense of a slightly increased IL, the embedded configuration would theoretically drive the optical attenuation up to values of ~40%, doubling the best result obtained with the on-top configuration (c.f. data in Figure 7.17(c)).

The effect of the nanoantenna vertical displacement on the optical behaviour is here further investigated. The $\Delta\mathrm{T}$ and IL values are reported in Figure 7.18, as a function of bar length and vertical displacement. The $\Delta\mathrm{T}$ plot (Figure 7.18. (a)) confirms that the previously selected vertical location is indeed beneficial, although a ~3.5% $\Delta\mathrm{T}$ improvement can be obtained by further optimization of its configuration. The IL plot (Figure 7.18. (b)) also highlights that a minor variation should be expected as the nanoantenna configuration is optimized, allowing an (although marginal) improvement.

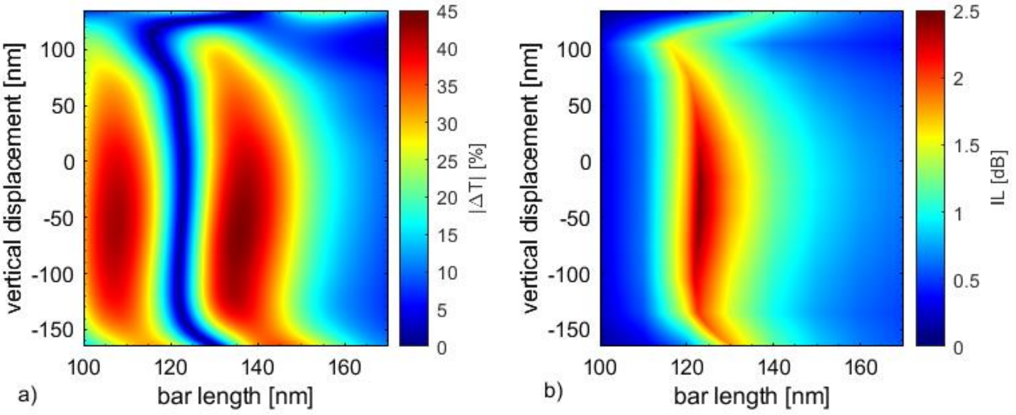


Figure 7.18. Optical behaviour of the dimer-bars embedded configuration, as a function of bars length and vertical displacement. The vertical axis is centred at the etching step location (waveguide mid-height); the nanoantenna develops vertically starting from the position indicated on the vertical axis. (a) $\Delta\mathrm{T}$, calculated between crystalline and amorphous configurations. (b) Insertion loss [dB].

The results above focus on the optical performance of the dimer-bars configuration. However, the device optical characteristics are improved also in the bowtie and dimer-discs designs, with the results being shown in Appendix V.

7.3.4. Alternative nanoantennas materials

The choice of silver as the material supporting the plasmonic resonance was driven from already published works in the same field, and the widespread agreement that Ag is a ‘good’ plasmonic metal. However, Ag is quite a diffusive metal, and in particular is known to diffuse rather easily into phase-change chalcogenides (see e.g. Ref. [124]). Thus, for real-world applications, alternatives to Ag may be required. Possible alternatives are investigated in this section, focusing on the peak $\Delta\mathrm{T}$ and EF performance of a dimer-bars nanoantenna embedded configuration (selected as it shows the most feature-less transmission plot), as a function of the nanoantenna material.

As the range of available materials is rather large, the choice is here limited to elemental materials. This selection includes also unsuitable materials such as Hg or alkali metals, as to broaden the picture of the device optical behaviour in response to the nanoantenna optical properties. The list of materials, optical properties, and corresponding references is provided in Table 7.1.

Table 7.1. Optical properties of the nanoantenna investigated materials.

Element	n	κ	Reference	Element	n	κ	Reference
Ag	0.1388	11.31	[300]	Nb	1.435	8.6	[301]
Al	1.3474	14.133	[302]	Ni	3.0902	7.7167	[83]
Au	0.1898	10.35	[300]	Os(e)	1.78	8.08	[303]
Ca	0.80789	7.2432	[304]	Os(o)	0.84	7.19	[303]
Co	3.5575	7.1245	[83]	Pb	1.9782	9.83	[305]
Cr	3.6683	4.188	[83]	Pd	3.7139	8.6409	[306]
Cs	0.28451	3.6527	[307]	Pt	5.3678	7.0663	[243]
Cu	0.1863	10.95	[300]	Rb	0.098203	4.2876	[308]
Fe	3.1045	5.2352	[83]	Rh	3.63	10.3	[309]
Hg	6.0879	8.3127	[310]	Sn	3.2475	9.02	[311]
K	0.06733	4.8185	[307]	Ta	0.85995	8.4309	[312]
Li	0.53634	7.2217	[313]	Ti	3.4323	3.1162	[306]
Mg	6.7735	12.523	[306]	V	3.8901	2.999	[306]
Mn	3.7441	5.3022	[83]	W	2.2299	4.8291	[312]
Mo	2.591	9.0405	[314]	Zn	3.8045	13.192	[314]
Na	0.12579	6.75599	[307]	Zr	3.5166	2.6957	[306]

The simulation results are plotted in Figure 7.19. These plots confirm that the initial choice of Ag as a plasmon-supporting material is indeed the most suitable one (from a purely optical performance point of view, as per its larger optical contrast and reasonable IL). In practice though, all of the tested 11th group metals (Ag, Au, Cu) perform similarly in both ΔT and IL parameters, with minor deviations one from another. Interestingly, Na also shows remarkable optical characteristics, although its use as a nanoantenna material is of course heavily prevented from both the low melting point and the high chemical reactivity.

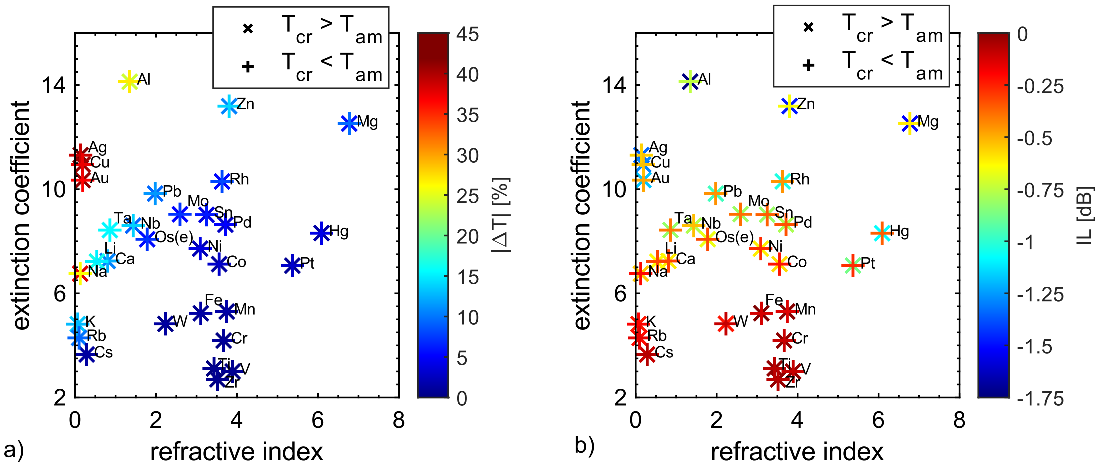


Figure 7.19. Embedded dimer-bars device optical behaviour as a function of the nanoantenna material. Two markers (vertical and diagonal crosses) are used, reporting the two values corresponding to the first and second $|\Delta T|$ peaks respectively. (a) Peak $|\Delta T|$ values. (b) IL calculated at the peak ΔT configuration.

Amongst the remaining tested elemental materials, Al shows the best optical performance, with a maximum $\Delta T = 26.8\%$ and related IL = -0.754 dB. Although Al does not easily diffuse into GST, due to the strong nature of the covalent bonds with Sb or Te [315], its melting point of 933 K is close to the one of GST. This solution might then suffer from GST-metal molten phase mixing, leading to device failure after repeated high-temperature ramps. Yet, as it will be demonstrated in Section 7.4.2, the temperature increase of the nanoantenna is drastically reduced with respect to the one calculated on GST, thus possibly preventing such a mechanism.

Two other suitable materials exhibiting $\Delta T > 10\%$ are found, selected avoiding unsuitable choices such as alkali metals or other elements with $T_{melt} < T_{melt}^{\mathrm{GST}}$:

- Ta ($\Delta T = 16.3\%$, IL = -0.41 dB)
- Nb ($\Delta T = 13.6\%$, IL = -0.53 dB)

Concluding, this preliminary investigation expectedly highlights as viable materials the ones with low refractive index ($0 < n < 1.5$) and high extinction coefficient ($\kappa > 10$). The numerical emulation confirms Ag as the most suitable choice, but also evidences that CMOS-compatible material options are available.

7.4. Investigation of the memory functionality

7.4.1. Device design and Read operation

Arising from the previous analyses, the investigation here carried out focuses on the memory performance of a feasible nanoantenna device design. The configuration of choice is the dimer disc nanoantenna, consisting of two symmetrically placed discs of thickness $t =$ 30 nm, diameter $d =$ 150 nm, and gap size $g =$ 40 nm. The parameters have been chosen from a simulated parametric sweep, as to balance the obtainable optical contrast whilst keeping the IL at a minimum. Such a configuration is close to the peak dipolar resonant mode found for the crystalline GST. The GST volume arises from the intersection of the discs' volume with a centrally-located disc of radius 30 nm, thus yielding a GST volume of roughly 6.4×10^4 nm^3. The device also implements a 5 nm cap layer of SiO_2, deposited on top of the Ag and GST surfaces, as this material was found to have the lowest impact on the device performance[59]. A schematic of the device geometry is reported in Figure 7.20(a-c).

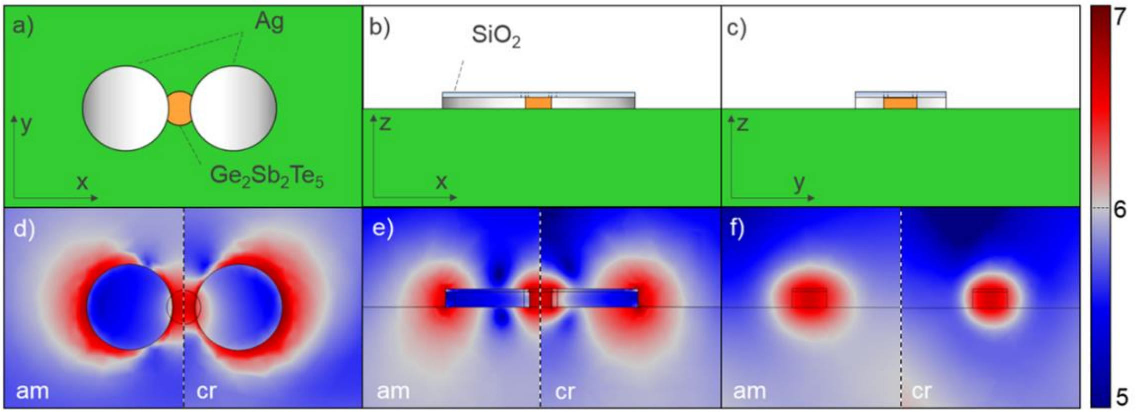


Figure 7.20. Schematics of the nanoantenna design (a-c), observed from the three orthogonal cross sections cutting through the nanogap centre, and corresponding simulated norm of the electric field (d-f) for the crystalline GST and amorphous GST inclusion (colour bar on the right, $\log_{10}$ values in V/m units).

---

[59] It is here noted that the residual GST eventually deposited on top of the Ag nanoantenna has a negligible influence on the device performance, and is thus neglected for sake of simplicity.

Both the amorphous and crystalline cases show the strongly localized increase of the electric field magnitude at the nanogap (see Figure 7.20(d-f)), as reported in the previous sections. The EF values of 20.2 and 15.2 are calculated for amorphous and crystalline GST respectively. As for the transmission parameter, this configuration yields a T value of 0.943 and 0.799 for amorphous and crystalline GST respectively, yielding a theoretical $|\Delta\mathrm{T}|$ of 14.3%. It is also noted that such configuration yields a normalized optical contrast $\Delta\mathrm{T}_{\mathrm{norm}} = \Delta\mathrm{T}/\mathrm{T}_{min} =$ 15.3%, not dissimilar from the reported experimental contrast of a conventional PPCM device using a 1 μm unit cell [67], [254], but obtained using only a fraction of the GST volume necessary for the conventional PPCM. As it will be demonstrated in the following section, the volume reduction is indeed a very beneficial element, as it reduces the effective heating volume.

A summary of the optical characteristic of the proposed device, revealed under Read operation, is reported below in Table 7.2. Appendix VI expands the analysis on the fabrication tolerance of such a nanostructure design.

Table 7.2 Summary of plasmonic phase-change optical cell optical behaviour at 1550 nm

	T	R	S	A (total)	A (GST)	EF	$\Delta\mathrm{T} = \mathrm{T}_{am} - \mathrm{T}_{cry}$
GST_{cry}	0.799	0.014	0.122	0.070	0.063	15.2	14.4%
GST_{am}	0.943	0.006	0.048	0.005	0.003	20.2	

The spectral response of the proposed device is also here investigated, to determine its suitability for optical wavelengths not originally intended for the device scope. Since the waveguide geometry is optimized for the C-band, the spectral range of choice is roughly limited between the O-Band and the L-Band (more specifically, between λ = 1350 nm and λ = 1700 nm). Figure 7.21(a) reports the device transmission, for the fully amorphous and fully crystalline unit cell. The data highlights a rather broad spectral range in terms of available optical contrast, with a useful ΔT value above 10% found between λ = 1390 nm and λ = 1650 nm (green area in Figure 7.21(b)). This optical behaviour is indeed well suited for WDM applications requiring a similar response on a broad spectral range, in parallel with the results obtained with the conventional PPCM devices for memory [67] and SNN computing [73] applications.

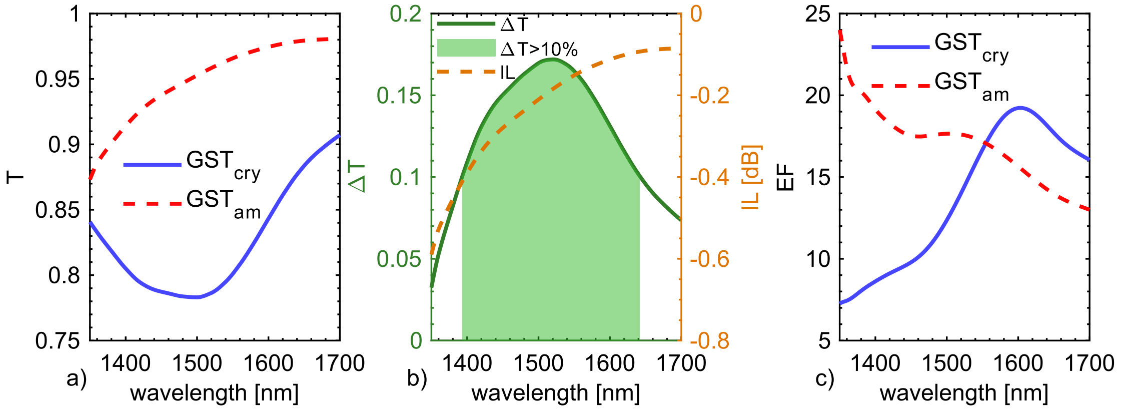


Figure 7.21. Spectral response of the proposed device configuration, in the 1350 nm – 1700 nm range. a) Device transmission for amorphous (red, dashed line) and crystalline (blue, continuous line) GST phases. b) derived $\Delta\mathrm{T}$ (green, left axis) and IL (orange, right axis). c) EF values, same colour scheme as in (a).

7.4.2. Write operation

This section illustrates the thermal and phase-switching performance of the device configuration illustrated above. As a starting point, it is necessary to determine the suitable pulse shape to obtain a large reamorphization of the completely crystalline unit cell (it is here reminded that the amorphization process is usually termed the ‘Write’ process, and re-crystallisation the ‘Erase’ process).

It is here underlined once more the low value of the ratio between the volumes of the conventional and plasmonically-enhanced unit cells, which for similar optical performance (i.e. normalized $\Delta\mathrm{T}$) results to be $6.4 \times 10^4/1.3 \times 10^7 = 1/180$. It is thus expected a proportional decrease in the required switching energy. Additionally, the EF value of 15.2 for the plasmonically-enhanced PPCM needs to be considered, as it would subject the GST to an effective optical power approximately equal to EF times the input power. The conventional architecture typically requires a Write pulse power of 10 mW to 20 mW for a duration of (roughly) 10 to 30 ns (see Refs. [67], [68], [172]).

Using the above observations as a guideline, it is possible to envisage a rectangular Write pulse of 1 mW, for a duration of 2 ns (2 pJ). The pulse shape is tested on a full simulation; the combined results of temperature and crystal fraction evolution, as a function of time, are reported in Figure 7.22.

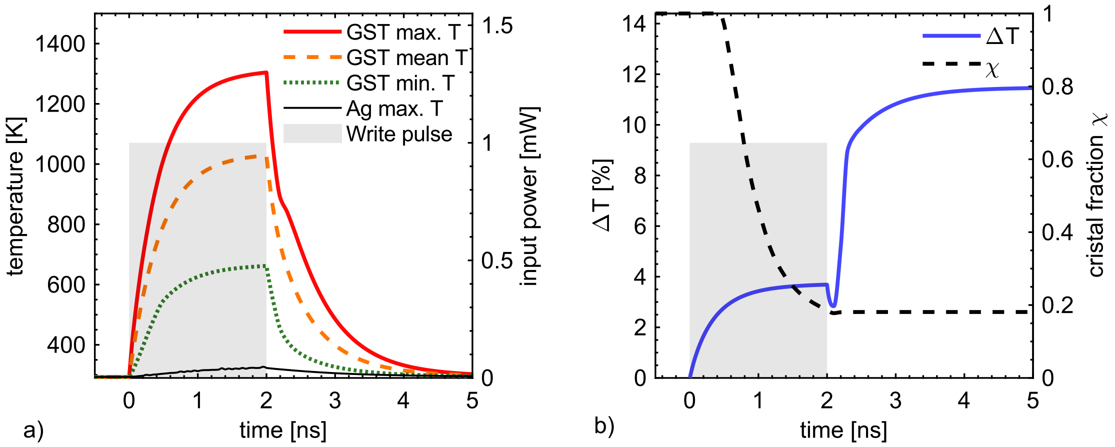


Figure 7.22. Switching dynamics during Write operation for the optimized plasmonically-enhanced unit cell described in Section 7.4.1. The Write pump is a rectangular pulse of 1 mW / 2 ns, reported in both plots as a grayed area to provide a visual guide (vertical scale in (a), right axis). (a) Temperature on the unit cell and Ag resonators. (b) Left: crystal fraction χ. Right: device transmission variation ΔT.

The unit cell temperature plot can be found in Figure 7.22(a), where a summary of the peak, average, and minimum temperature is reported. These show the exceptionally rapid temperature transient of the unit cell. The onset of the melting process, which begins at 0.4 ns, is also a noteworthy detail, which hints to the possibility of easily accessible sub-ns write processes with optimized Write pulses. Both peak and average unit cell temperatures rapidly increase above melting point: yet, both temperature profiles rapidly decreases (with a 1/e thermal decay of 0.6 ns) at the pulse end.

The numerical analysis reveals that the induced temperature profile lowers the GST crystal fraction from a value of 1 (fully crystalline) to a value of 0.18. The crystal fraction plot as a function of time is reported in Figure 7.22(b), right axis. On the same panel, left axis, the transmission variation $\Delta \mathrm{T}$ is reported, which reveals a post-pulse shift of $\Delta \mathrm{T} = 11.5\%$. It is here noted that, from Li et al. [68], the same variation in transmission requires a programming energy of ~98 pJ, thus revealing an improved Write efficiency (i.e. ratio of the Write pulse energies) by a factor of 49. The calculated maximum optical fluence within the unit cell is 0.012 J/cm^2, which is around two orders of magnitude below the damage threshold [316]. To prevent thermal damage effect [316], the peak temperature is here also maintained below an arbitrary threshold of 1500 K.

The FEM 3D solution provides a clearer picture of the spatial distribution of the optical loss and temperature distribution. Figure 7.23(a) underlines how the majority of the ohmic loss is found close to the Ag resonators, which is an expected result considering the higher field magnitude closer to the plasmonic nanoparticles. Nevertheless, the GST TBR results to be sufficient to prevent the quick heat diffusion towards the metal volume, thus locating the temperature peak close to the centre of the GST inclusion (see Figure 7.23(b)). Specifically, a nanoantenna peak temperature of 327 K has been calculated during the Write process (see also Figure 7.22(a) for a representation of the Ag peak temperature dynamics). A TBR at the Ag/SiN interface (TBR = 5×10^{-9} m^2K/W) is also assumed [317], which may further contribute to limit the heat diffusion, across Ag, towards the substrate.

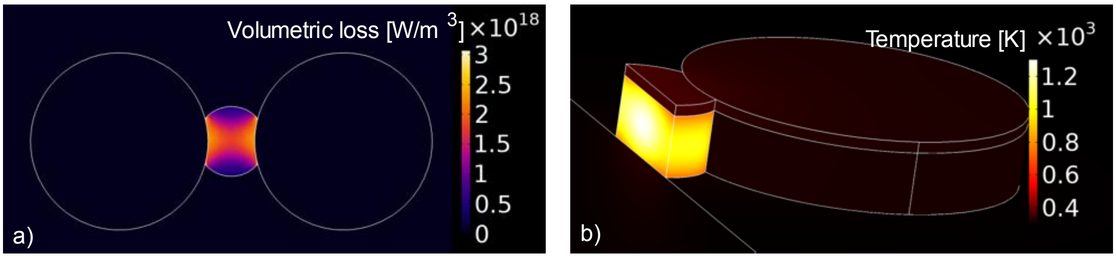


Figure 7.23. Write operation loss (a) and temperature (b) distribution plots; data taken at the end of the Write pulse (2 ns). The optical loss density 2D plot (a) is calculated on the cut plane passing through the unit cell central point.

The melting process starts from the GST volume center, then quickly spreads towards the top region (where a barrier to the heat diffusion path is found), to then further expand towards the bottom region when approaching the pulse end. Barely any recrystallization post-pulse is detected. An illustration of the achieved phase distribution across the unit cell is shown in Figure 7.24.

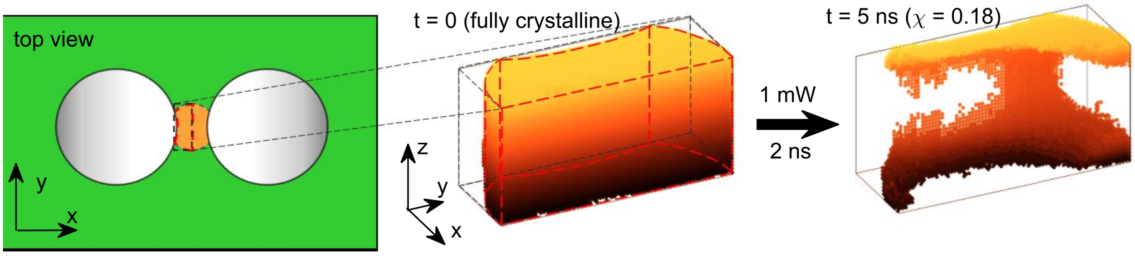


Figure 7.24. Spatial GST phase distribution across the unit cell (the plots report the crystalline phase only).

The fundamental limit of the Write process energy requirement is the energy necessary to bring GST from room temperature to melting point, added to the enthalpy of fusion (melting). Thus, to a first approximation, and using the previously reported parameters, the minimum theoretical Reset (i.e. amorphization) energy E_{min} is estimated as follows:

$$
\begin{aligned}
E_{min} &= c_P \Delta T + \Delta H_{fus} \\
&= \left(6150\ \left[\frac{\mathrm{Kg}}{\mathrm{m}^3}\right] \times 210 \left[\frac{\mathrm{J}}{\mathrm{KgK}}\right]\right) \times (889[\mathrm{K}] - 293[\mathrm{K}]) \\
&+ 1.121 \times 10^9 \left[\frac{\mathrm{J}}{\mathrm{m}^3}\right] \cong 1.9[\mathrm{aJ/nm}^3] \qquad \text{Eq. 7.2}
\end{aligned}
$$

With the proposed Write pulse, the calculated optical energy absorbed and scattered during the process totals 340 fJ. This value, divided by the amorphized unit cell volume, yields an energy requirement of 4.36 [aJ/nm^3], which is remarkably close to the theoretical limit of 1.9 [aJ/nm^3] calculated in Eq. 7.2. In comparison, using the data retrieved from the simulation task on the 2 µm conventional PPCM (illustrated in 6.3.5), an energy requirement of 21.35 [aJ/nm^3] is estimated, 4.9 times higher than the plasmonically-enhanced architecture. This observation consolidates the fact that not only the required GST volume to switch is here much lower, but also, the switching operation is inherently more efficient.

In place of the energy efficiency, it is here proposed a different *Figure Of Merit* (FOM) as a tool to rapidly evaluate the energy and optical performance of a non-volatile PPCM. The suggested FOM considers the ratio between the $\Delta\mathrm{T}$ contrast (in dB) and the delivered input energy (in pJ). For the Write process, a FOM of 263 [dB/nJ] is here calculated.

7.4.3. Erase operation

The complete recrystallization[60] process requires that the crystal fraction shifts from its initial value to a predetermined one ($\chi = 1$). The design of a suitable pulse sequence is here pursued by observation of the spatial phase-state distribution after the Write process, combined with the notion of the Erase sequences adopted for the conventional device.

In the conventional architecture, both a pulse sequence of decreasing energy and a double-step pulse have been experimentally demonstrated to restore the full crystal phase of the GST unit cell [67], [68]. Both approaches lower the power as time advances, so as to account for the increase in optical absorption (and the consequent heat generation) with the increase of the crystal fraction. Thus, a pulse sequence applying a decaying power is necessary.

---

[60] This section focuses on the recrystallization after the previously described Write operation. In contrast, the recrystallization from the completely amorphous unit cell ($\chi = 0$) would certainly yield a different result. The energy and time requirement would substantially increase, due to the decreased absorption of the amorphized unit cell; thus, this pulse optimization it is not here pursued, albeit implying the cost of a reduced $\Delta\mathrm{T}$ achievable modulation of the proposed device.

Another useful consideration arises from the observation of the recrystallization kinetics, which sees its fastest pace at a temperature around 650 K to 800 K, according to the adopted CNG phase-change model (see Figure 4.12). Therefore, the pulse sequence needs to induce such a temperature at the boundary between amorphous and residual crystal phases, inducing the fastest crystal growth.

For the plasmonically-enhanced unit cell, the peak temperature is reached at the unit cell centre (see Figure 7.23(b)). Thus, starting from the amorphized unit cell, and attempting to initiate the recrystallization from the edges of the amorphous mark (see Figure 7.24) may lead to the uncontrolled melting of the central volume. It is therefore necessary to carefully control the temperature decay so as to obtain the optimal temperature distribution on the GST phases boundary.

As a last remark, it is necessary to bring to attention the fast temperature decay, due to the very high temperature gradient reached during heating. Such fast dynamics of the temperature distribution works against the efficacy of the multi-pulse scheme, which ostensibly relies on the cumulative effect of the temperature increase. The combination of the above observations drives one towards the design of an erase pulse shape suited to the plasmonically-enhanced composed of two subsequent elements:

i. an initial single spike in energy, capable of rapidly increasing the amorphous GST temperature to melting point, so to set the optimal initial condition for the recrystallization process, already close to or above the peak crystal growth temperature in the whole unit cell;
ii. a following linearly decaying pulse (i.e. a triangular pulse) with the scope to control the temperature decrease, moving the front of the optimum recrystallization temperature from the outer edges to the central position of the GST inclusion, also preventing the melt of the already recrystallized volume.

With this scheme as a guideline, it has been possible to reach a remarkably effective Erase pulse design. This consists of a first 1.5 mW / 1.5 ns rectangular pulse, followed by a 1.2 mW to 0.5 mW / 15 ns linearly decaying triangular pulse, for a total delivered energy of 15 pJ. Interestingly, this pulse sequence finds a remarkable similarity with the electrical Erase sequence independently proposed in Farmakidis et al. [176].

The simulation of the Erase process starts from the previously obtained memory level ($\chi = 0.18$) from the Write simulation, to demonstrate the suitability of the proposed PPCM architecture for rewritable binary storage. The results are illustrated in Figure 7.25.

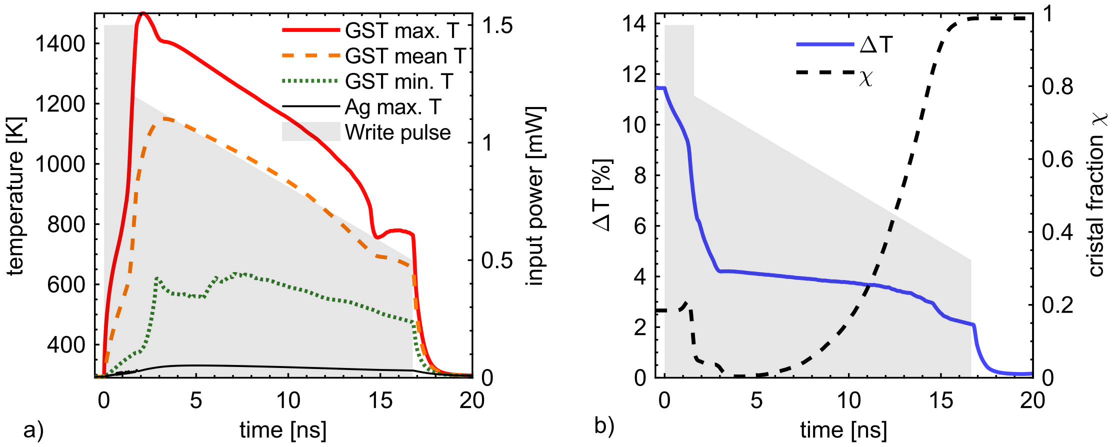


Figure 7.25. Switching dynamics during the Erase operation, starting from the crystal fraction value of $\chi =$ 0.18 obtained after the Write operation. The Erase pulse profile is reported in both plots as a greyed area to provide a visual guide (vertical scale in (a), right axis). (a) Peak, average and minimum temperature on the unit cell and resonators. (b) Left: crystal fraction χ. Right: device transmission variation ΔT.

The above results show that the initial spike in temperature (Figure 7.25(a), red continuous line) is responsible for the shift of the crystal fraction to a near-zero value (Figure 7.25(b), black dashed line), with the remaining crystal GST located at the edges of the unit cell in proximity with the resonators. Then, the purposely designed triangular pulse controls the temperature decay as intended. The temperature plots describe the initial temperature decrease that sets the edges of the amorphous mark to a convenient 750 K value, allowing the aimed for fast crystal growth. As time flows, the accumulated heat diffuses from the central volume to the outer surfaces, driving the recrystallization from the outer surfaces to the centre in a progressive fashion, until the end of the pulse, with an achieved value of the crystal fraction of $\mathrm{X} = 0.98$, indeed very close to unity, and recovering a $\Delta T \cong 0.2\%$, close to the original crystal state. The 1/e decay time estimation returns a value of 0.4 ns, which is even lower than the one estimated during the Write operation; the reason lays in the fact that the temperature at the end of the Erase process is much lower than the one found at the Write process end. An Erase performance FOM of 35.6 [dB/nJ] is found. Once more, as in the case of the Write operation, the Ag temperature do not increase above 330 K, negating Ag melting (and in all likelihood any increase of GST-Ag interdiffusion, though further analysis is here required).

7.4.4. Multi-level storage

So far, the analysis has been restricted to the demonstration of the binary storage functionality. This subsection expands the previous investigation, pursuing the *Multi-Level* (ML) storage capability. More specifically, the analysis focuses on the achievement of predetermined optical contrast levels, from any arbitrary initial unit-cell condition. ML storage is indeed a much-desired capability, as it further expands the pool of addressable memory states, and thus the memory density. Also, as ML storage capability has already been demonstrated with the conventional architecture [68], a natural question would arise regarding such ability being found with the proposed plasmonically-enhanced architecture.

As a proof of concept, let 4 unique contrast levels be determined: $level\ 0$ (fully crystalline unit cell, X = 1; $\Delta\mathrm{T} =$ 0%); $level\ 1$ ($\Delta\mathrm{T} =$ 3.82%); $level\ 2$ ($\Delta\mathrm{T} =$ 7.64%); $level\ 3$ ($\Delta\mathrm{T} =$ 11.44%, with the previously illustrated crystal fraction of X = 0.18). A strategy to achieve such contrast levels is to apply the Erase sequence to the lowest memory state ($level\ 3$), with the additional condition to stop the recrystallization process (i.e. the Erase sequence) when the aimed for $\Delta\mathrm{T}$ level is achieved. The proposed pulse scheme will be henceforth addressed to as *Programming* pulse.

$Level\ 3$ can be reached interrupting the Programming sequence at 10.2 ns (10.8 pJ). $Level\ 2$ is reached at 12.4 ns (12.4 pJ, with a resulting χ = 0.44). $Level\ 1$ can be reached at 13.8 ns (13.4 pJ, χ = 0.72). Of course, $Level\ 0$ is reached with the use of the full Erase sequence (pulse duration of 15 ns, pulse energy of 15 pJ), as demonstrated in Section 7.4.3. The crystal fraction data is reported in Figure 7.27(b). The inset in Figure 7.27(b) reveals the quasi-linear dependency between unit cell crystal fraction and device $\Delta\mathrm{T}$, which confirms the feasibility of the proposed Programming sequence as a method to reach evenly separated $\Delta\mathrm{T}$ (and χ) memory levels. Figure 7.26 reports the obtained crystal phase distributions in a pseudo-tridimensional fashion within the half-cell.

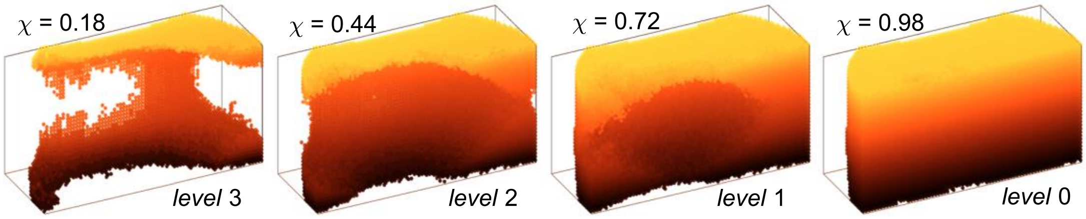


Figure 7.26. Crystalline GST spatial phase distribution throughout the unit cell, for the proposed 4-levels memory storage, as calculated from the Programming pulse sequences illustrated in Figure 7.27.

The programming operation is optimized for the increase of the crystal fraction from $level\ 0$. In order to perform the Programming operation from any arbitrary level, it is necessary to tackle the instability of the heated system, which above melting temperatures sees a rapid increase of the optical power absorption (due to the molten GST phase, which by assumption retains the material properties of the crystalline phase). Owing to this premise, the delivery of a high-energy pulse as the Programming pulse to low memory level may damage the unit cell. Henceforth, it is here demonstrated a complementary operation to the Programming sequence, referred to as *Overwrite*. It is found that the application of the initially described Write pulse (1 mW / 2 ns), to any of the previously considered memory levels, lowers the crystal fraction value to χ = 0.18 ($level\ 3$), within an error of ± 0.02. The numerical emulation also finds that a delay of 1 ns between Overwrite and Programming pulse is sufficient to thermalize the GST volume, removing any unexpected variation from the expected temperature profile during and after the application of the Programming pulse and thus negating any divergence from the aimed memory level. The crystal fraction values plotted in Figure 7.27(a) as a function of time show such functionality, which paired to the illustrated Programming sequence (see Figure 7.27(b)), fully certifies the 4-level storage capability of the plasmonically-enhanced PPCM device.

By combining the energy of Overwrite and Programming sequences, the average energy-per-level input requirement (for the proposed 4-level storage) is estimated to be 12.2 [pJ/$level$]. The FOM derived from all the possible combinations of initial and final levels is 42.06. As a comparison, Li et al. work on the conventional device yields a ML FOM of 5.20.

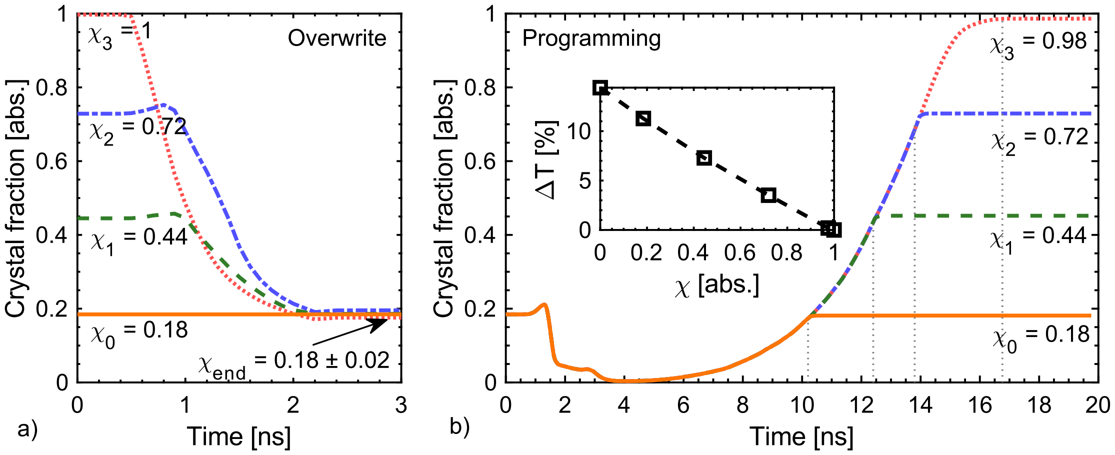


Figure 7.27. Multi-level storage demonstration, crystal fraction data as a function of time. The pulse delivered power as a function of time is also reported as a greyed area, to guide the eye along the operations process. a) Reset operation, shifting any arbitrary initial state to the value of $\chi = 0.18$. b) Programming operation, returning the previously determined state to any of the four predetermined crystal fraction values, within a tolerance of ± 0.02. Inset demonstrates the quasi-linear correlation between crystal fraction and device ΔT.

The ML analysis is here restricted to a proof-of-concept 4 levels memory density. Yet, the use of a smaller $\Delta\mathrm{T}$ separation would yield a larger memory capacity. The repeated simulation of the Programming sequence yields a crystal fraction error $\delta\chi$ = 0.03 (arising from the stochastic nature of the adopted PC model). Assuming a safe crystal fraction level separation of $\Delta\chi = 2\delta\chi$, it would be possible to obtain 14 memory levels, separated by an average $\Delta\mathrm{T}$ value of $\Delta\mathrm{T} = 0.88\%$ This level optical separation is also compatible with the experimental setup sensitivity of ~0.67% reported in previous works [67]. An extrapolated ML FOM of 20.05 is obtained in this case. Farmakidis et al. [176], in a later work, use a lower $\Delta\mathrm{T} = 0.2\%$ separation on a conceptually similar plasmonic-enhanced architecture and suggests potentially denser memory levels.

7.4.5. Performance comparison with the conventional architecture

The above results evidence the low energy requirement of the proposed plasmonically-enhanced device. The nature of such increased performance lies in the initially assumed device capability to harness the optical energy and focus onto a narrow GST region, which thus sees enhanced irradiance and increased thermal insulation. A comparison of the optical and energy performance of the plasmonically-enhanced PPCM device, with the conventional architecture, helps to quantify such improvement; Table 7.3 reports such data comparison, which also extends to include the comparison with GST-enabled optical switches.

Where data was not directly provided, a calculated value was instead supplied when possible.

Table 7.3. Comparison of proposed and experimental PPCM and GST-enabled switch devices performance. The worst and best performance values are highlighted in green and red respectively, for each column. *W* refers to the Write process; *E* refers to the Erase process.

Unit cell size [μm]	# Levels	Max. contrast ΔT_{norm} [%]	Shortest pulse width [ns]	1/e thermal decay time [ns]	Lowest switching energy [pJ]	Min. IL [dB]	**FOM [dB/nJ]**
Conventional configuration, multi-pulse [67]							
0.25×1×0.01	2	0.67	*W*:10	*W*: 10	n/s	n/s (≪ 1)	***W*: 0.523**[a]
5×1×0.01	8[b]	58	*W*: 10 *E*:>1300	*W*: 10 *E*: n/s	*W*: 52 *E*: 3095	n/s (~1)	***W*: 5.13** ***E*: 0.461**
Conventional configuration, double step pulse [68]							
4×1.3×0.01	34	28	*W*: 25 *E*:125	*W*: 200 *E*: 500	*W*: 68 *E*: 680	n/s (~1)	***W*: 10.5** ***E*: 2.09**
Conventional configuration on a micro-ring resonator [276]							
0.6×0.75×0.02 (30[d])	n/s	70	*W*: 25 *E*: 2500	n/s	*W*: 180 *E*: >1900	~4	***W*: 16.7** ***E*: 1.58**
Split discs configuration on a micro-ring resonator [175]							
0.5×0.5×0.01 × 5 units (~25[d])	n/s	87	*W*: 50 *E*: 2500	n/s	W: 220 *E*: n/s	<1	***W*: 40** *E*: n/a
1D Photonic Crystal (PhC) [174]							
0.91×1×0.01 (0.51 × PhC cavities[d])	n/s	96	*W*: 100 *E*:>2500	n/s	*W*: 150 *E*: 2160	>18	***W*: 10.6** *E*: n/a
Plasmonically-enhanced Mixed-mode device [176]							
10^{-2}×7.5×5×1 (~20[d])	20	~33	*W*: 5 *E*: 408	n/s	*W*: 16 *E*: 1260	~9	***W*: 3.33** ***E*: 0.43**
Plasmonically-enhanced all-optical unit cell (this work)							
10^{-2} × 6×4×3 (0.15×0.34[d])	4 (14[c])	12.5	*W*: 2 *E*: 16.5	*W*: 0.6 *E*: 0.4	*W*: 2 *E*: 15	0.38	***W*: 264** ***E*: 35.6**

[a] Calculated assuming a 50 pJ pulse, as reported on the same source albeit for different configurations.

[b] More recent works ([68], [172]) demonstrated much-improved memory density for the same device's class.

[c] Extrapolated value, based on a level separation of $\Delta T \cong 0.88\%$.

[d] Resonant structure size.

This comparison clearly outlines and quantifies the intrinsic advantages of the proposed plasmonically-enhanced PPCM. Taking as a direct comparison the work of Li et al. [68], the data yields a possible improvement of the memory switch operations of up to a factor of 7.6. Additionally, not only the Write and Erase pulse durations are much shorter, but the reduced thermal decay suggests a further improved memory operation frequency (increased speed). The combined Write/Erase pulse durations and 1/e thermal decay yields a possible memory speed improvement up to the staggering factors of 36 and 86, for Write and Erase operations respectively. Likewise, the minimum energy requirement is decreased by a factor of 34 and 45 for Write and Erase respectively. It is noted how both speed and energy are simultaneously improved by double-digit factors. It is also important to remark that the proposed pulse sequence is not yet optimized, and further improvements by adoption of more suitable combinations of pulse power and duration are foreseeable.

One last consideration is made upon the obtained memory density, which strongly correlates to the achievable tunable modulation $\Delta\mathrm{T}$. Indeed, the investigated device is capable of a relatively low modulation of 11.4%, being this 2.8 times lower than the one found in [68]. It is worth noting though that the proposed structure has been selected assuming the eventual fabrication complexity. As suggested before, the same concept can yield much-increased $\Delta\mathrm{T}$ values if so optimized. Additionally, possible routes to increase such optical contrast have already been suggested within this work. As an example, the embedded dimer-bars configuration (see Figure 7.17) is capable to yield a theoretical $\Delta\mathrm{T} =$ 38.6%, with an IL of -0.55 dB. The higher EF values there found also hint to further improvement of the switching energy and speed performance.

7.5. Summary

This Chapter illustrates a novel phase-change photonic memory architecture, which in contrast with the conventional configuration, tunes the resonance strength of a dimer plasmonic nanoantenna, to be fabricated on the waveguide top surface.

The nanoantenna optical behaviour has been numerically analysed, revealing a characteristic scattering cross-section which can be optimized to pursue maximum interaction, and consequently, maximum optical attenuation. Additionally, the dimer nanoantenna focuses the electric field onto the dimer’s nanogap (of characteristic size in the range of tens of nm), due to the well-known near-field enhancement, with an efficiency drastically superior to the evanescent field coupling of the conventional PPCM.

The insertion of GST within the nanoantenna gap shifts the device optimal resonant configuration, deriving from the modification of the nanoantenna environment dielectric function; consequently, the resonance strength (and the optical attenuation) can be tuned via phase switching. As the GST volume is subjected to high thermal confinement and perceives an effective irradiance EF times that of the input, it derives that the switching energy performance is also enhanced, compared the conventional architecture.

A device configuration has been proposed, specifically optimized in view of future proof-of-concept fabrication process. Write and Erase operations simulation have been performed, outlining the potential to induce non-volatile $\Delta\mathrm{T}$ variation by delivery of 2 pJ and 15 pJ respectively. A similar variation of the $\Delta\mathrm{T}$ via use of the conventional architecture would require energies 1 to 2 orders of magnitude larger, confirming the predicted superior energy performance. Multi-level functionality has also been demonstrated, via simulation of memory storage up to 4 levels (albeit data extrapolation suggests a memory density of 14 levels).

Additionally to the simulated demonstration of the device optical behaviour and memory functionalities, a larger characterization of the nanoantenna design, materials and configuration has also been presented. Strategies to improve the otherwise limited optical contrast have also been proposed, such as device embedding within the waveguide core, or the use of multiple interacting nanoantennas (detailed in Appendix IV). The simulation of those solutions indeed demonstrates the potential to access to much wider $\Delta\mathrm{T}$ values. These compare to those obtainable by much larger conventional unit cells, yet retaining the traits which allow for the sharply improved switching energy and speed performances.

The memory architecture here proposed is however not exempt from drawbacks. The sensitivity of its optical performance from nanogap size and dimers size needs to be properly addressed during either design (as demonstrated in this chapter and Appendix VI) and fabrication (see Gemo et al. [318]). However, the optimization of the fabrication process can be limited towards a few key parameters. The conventional lithography techniques already satisfy the critical dimension requirement of the proposed concept (see e.g. Thyagarajan et al. [319]). Additionally, plasmonic resonant elements are today at the centre of the attention of a large share of the scientific community, which in conjunction with the steady improvement of the fabrication techniques will realistically further fabrication reliability, scalability and cost-effectiveness of the proposed concept in the foreseeable future (see e.g. the EU research grant no. 899598, assigned for the Phemtronics project, which aims to adopt and expand the scalability of the plasmonic memory concept here proposed).

Chapter 8. Final Remarks

8.1. Conclusion

Fast and energy-efficient memory and computing solutions are nowadays of paramount importance. The design and fabrication of chip-to-chip and on-chip electrical interconnects face ever-growing challenges, that make it difficult to easily fulfil technology demands. Thus, research is nowadays shifting towards the photonic domain, aiming for computing devices inheriting the energy efficiency and wide bandwidth advantages of photonics interconnects Yet, there is an uncomfortable gap in the all-optical memory storage technology, which generally relies on energy-hungry active solutions. The phase-change photonic memory (PPCM) bridges this gFap, adding passive non-volatile storage functionality in the optical domain, with write and erase operations in the ns to µs scale. Thus, it is of great benefit to theoretically and numerically analyse the physical behaviour and operation PPCM devices, to achieve a full understanding of its potential, to improve its design, and to envisage its adoption and optimization for novel applications such as reconfigurable mode modulators [290], optical filters [320], 1×2 [321] or 2×2 [322] optical switches, directional couplers [289], sensors [323], and more.

This thesis aims to step towards such objectives. Chapter 2 lays the theoretical foundation of this work, covering the essential aspects of the physical mechanisms involved in the optical, thermal, and phase-switching behaviour of the PPCM. Chapter 3 portrays the technical aspects of this work, describing the tools adopted for this investigation; further theoretical details are also available in Appendix I.

The required improvement on the algorithmic side for realistic simulation of PPCM type devices is reported in Chapter 4 (with additional information included in Appendix II) There, the development of the 3D phase change model is outlined, starting with the determination of the 3D cellular automata elements, and then applying those to both the GCA model and the CNG model. A comparison of the two models is provided, supporting the choice of the latter (CNG) for photonic devices emulation, as it reliably pictures a similar kinetics to the more sophisticated GCA at a fraction of the computational cost. The chapter then follows with the definition of the simulation framework, which combines i) the ease of use of the FEM tool, ii) the purposely developed PC model, and iii) the flexibility of the time-advancement and data evaluation algorithm.

For this research, and in general in the optical and photonics engineering fields, the knowledge of the involved materials optical properties (spectroscopic refractive indices) is essential. Yet, obtaining such parameters from the scientific literature may be misleading (due to lab-lab differences in fabrication conditions, compositional variations etc.) and the ellipsometric determination of the optical properties requires rather specialist equipment and has non-negligible costs. Chapter 5 therefore develops a novel and viable method to quickly retrieve refractive index information, based on an innovative fitting technique of an easy-to-fabricate Girés-Tournouis cavity reflection spectrum. This technique requires minimal preparation and measurement costs, and remarkably improves on the n and κ estimation error without the necessity of specialized machinery or trained users. Additional development of this technique is detailed in Appendix III.

Chapter 6 takes the developments in the modelling of PPCM devices from previous chapters and applies them to a detailed numerical analysis of the conventional phase-change photonic memory. The results outlined a wide set of unknown variables, which have been systematically tuned to reproduce, as closely as possible, a range of published experimental results. It was shown to be possible to reproduce the key aspects of the device optical, thermal, and switching behaviour, with remarkable precision at low optical intensities and low temperatures. These results justify both the physical model assumptions and its algorithmic implementation in this thesis. As for the high temperature, high intensity regime, the simulation model underestimates the switching dynamics, underlining the requirement to thoroughly characterize the optical properties of PCMs (in this case GST) in such conditions.

The simulation framework developed in the thesis has also been adopted to explore a novel memory concept with a view to improving performance as compared to the ‘conventional’ PPCM device design. Chapter 7 thus illustrates the plasmonically-enhanced PPCM architecture, with additional information included in Appendices IV-VII. The design exploits an optimized plasmonic dimer nanoantenna, whose resonance strength is modulated by the optical properties (i.e. crystal or amorphous phase) of a GST inclusion within the dimer nanogap. The magnification of the near-field in the nanogap, the much-reduced volume of the devised memory unit cell, and the strong thermal confinement which such design entails, led to dramatic improvements in the switching efficiency of the GST element which tunes the nanoantenna scattering efficiency, and in turn, the device optical attenuation. The numerical results highlight drastically improved, by factors of one to two orders of magnitude, switching energy and switching speed performances when compared to the conventional architecture. Whilst the physical mechanism of this memory concept differs from the conventional architecture, the numerical investigation confirms its potential use as an all-optical multi-level memory element, also pinpointing additional avenues of investigation for the improvement of the device optical contrast, and consequently, memory density. The feasibility of the proposed memory concept is also addressed, in particular in Appendix VI. The memory performance is indeed sensitive to its configuration, but strategies to strengthen the fabrication tolerance have been subsequently outlined (see Gemo et al. [255] for further work on the subject), with the perspective that technological improvement on the very much active field of research on plasmonics would further subside the cost and scalability challenges that this memory concept would face.

8.2. Outlook

8.2.1. Development of a geometrically-optimized, tau-leaping, GCA model

The choice of the PC model for the simulation work of this thesis has been oriented towards the CNG model, due to its reduced computational cost. The price is the decreased accuracy on the recrystallization complex dynamics, which at the moment is assumed to not drastically influence the optical properties[61]. In contrast, both the GST electrical switching mechanism and the GST thermal conductivity are highly sensitive to such elements. For such applications it would be beneficial to further develop the GCA model by lowering the calculation time and resources requirement, to draw from its physical-realistic dynamics.

---

[61] The rationale behind this assumption relies on the fact that the impinging radiation would see an effective value, rather than responding to the sub-wavelength features of the grains sizes and shapes. It is to be noted though that recent works, such as Behrens et al. [341] or Yin et al. [342], point to the influence of the cluster size distribution on GST alloys optical properties (although also confirming how this effect is enhanced at shorter wavelengths than the one here adopted). Therefore, future work may also address this aspect of the numerical emulation framework.

One possible solution replaces the GCA cubic lattice with a rhombic dodecahedron automata cell arrangement, combined with the merging of the GCA and MRE models for the nucleation (and possibly any crystal fraction variation) via continuous automata framework [324] . The combination of MRE and GCA for the nuclei formation and growth would avoid the formation - dissociation iterating sequence mechanism at high temperatures, which would be incorporated in the incubation time and generation of an already stable nucleus of the correct size (or, in case of growth, locally adjusted variation of the grain size via growth). The rhombic dodecahedron is a space-filling polytope, which better approximates a sphere when compared to the GCA cubic lattice. Such a lattice allows calculating the locally fitting sphere (or spherical cap) approximation of an interrogated crystal orientation, for a neighbouring ensemble of automata cells. This would also yield a more precise estimation of enthalpy and surface energy contributions, in place of the quantized approach of the current model.

A complementary technique would develop on the time evolution of the automata, by implementing the tau-leaping mechanism [325]. This method would force the algorithm to advance in time by pre-determined time-steps, in place of the originally devised Gillespie time generation method. The kinetics of reaction can still fulfill the Gillespie kinetics, with a predeterminable error on the dynamics simulation. This method would require careful attention to avoid event collision; but, once this was achieved it would allow to massively parallelize an otherwise strictly sequential algorithm.

8.2.2. GST temperature- and intensity-dependent optical properties

As mentioned above, and throughout this work, the optical properties of GST are key components that need further study to fully unravel the complex dynamics of the optical, thermal and switching behavior of the PPCM device. Whilst the unperturbed GST can indeed be accurately measured, a setup capable to retrieve both amplitude and phase modulation of a GST layer, at a steady-state condition, at high temperatures and optical intensities would be greatly valuable.

A possible avenue of work would use the conventional PPCM device, in a similar fashion to the experiment portrayed in Stegmaier et al. [163]. By use of a sequence of differently-sized unit cells, on both a straight waveguide and on a MZI interferometer, it would be possible to accurately predict the optical properties of the unit cell material by use of the simulation software. Placement of such devices on top of a temperature-controlled heater would allow the aimed for steady-state temperature control, and consequently to fully characterize the temperature-dependent effective refractive index of (crystalline or liquid) GST. Guided by this knowledge, and performing the experiment via the use of a power-increasing sequence of continuous optical sources, it would then be possible to unravel the effect of the optical intensity on the effective refractive index. The FEM tool would then help to recover the GST optical properties from the estimated effective refractive indices values.

As for the amorphous GST, its fast recrystallization dynamics hinders the access to its optical properties at higher temperatures and optical intensities with the technique described above. It may be possible to overcome this limitation by use of short-duration, high-energy pulses; and combining the simulated results with i) the crystalline material characterization, ii) the effective medium theory, and iii) the knowledge on the crystallization kinetics.

8.2.3. Experimental verification of the plasmonically-enhanced unit cell

The work on the novel plasmonically-enhanced PPCM is still under the experimental verification stage. Since the numerical results show an exceptional increase of the memory speed and energy performance, completion of this experimental step would be very beneficial. A preliminary demonstration of the transmission dependency on the radii (of the dimer-disc on-top configuration) has been published (by the author and colleagues) in February 2020 [318]. These results confirmed that, as the radii of the fabricated discs increase, the predicted decrease of the waveguide transmittance is indeed observed, although with a larger IL contribution, and with a much broader transmission modulation. Future investigation would need to determine a suitable fabrication process which minimizes the fabrication defects, to support the outcome of the previous analysis, or confirm the predicted simulation results, as well as to demonstrate the predicted energy and speed performances. Possible NLO effects may also take place, as previously reported on similarly-sized nanogap-enhanced Ag optical structures (see e.g. [326]); thus, the memory concept would also profit by a focused theoretical analysis of the NLO phenomenon determined by the vicinal resonant nanostructures.

8.2.4. Adding the mixed-mode functionality to the all-photonic plasmonically-enhanced phase-change memory

Interestingly, the embedded dimer-bars configuration shows a particularly efficient scattering process at the quadrupolar resonance, which is also maintained after the GST inclusion is introduced into the gap region. The different transmission modulations for a crystalline and amorphous GST inclusion are then responsible for a remarkable optical contrast, even at the quadrupolar resonance. The bar length necessary to establish such mechanism is close to the waveguide half-width. It is then possible to envisage [255] the extension of the bar length outside the waveguide walls, combined with the interjection of a bow-tie like ITO electrical contact overlapping the out-thrusting bars (see Figure 8.1(a)). Figure 8.1(b) also evidences the abovementioned optical behaviour at bar lengths which exceed the waveguide walls. A preliminary simulation analysis of the ITO-implemented device is reported in Figure 8.1(c). The numerical data evidences that, adopting a bar length of 650 nm, one achieves a beneficial optical contrast of up to $|\Delta\mathrm{T}| \cong 20\%$, with a manageable IL close to 0.5 dB. Thus, it would be possible to adopt such a configuration to retain the same functionalities and performance scale of the presented all-optical plasmonically-enhanced PPCM, with the additional electrical domain operativity enabled by the lateral ITO contacts. A thorough simulation and experimental investigation of such concept would then be an interesting avenue of work, as, if the preliminary results are confirmed, it would lead to superior optical performance with respect to the mixed-mode device presented in Farmakidis et al. [176], at no additional fabrication complexity cost.

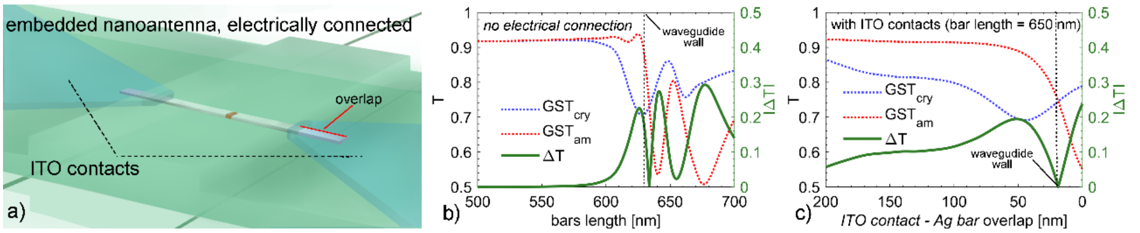


Figure 8.1. Electrically connected, embedded nanoantenna device concept and results. a) Schematic representation. b) Optical results of the un-connected device (phase-dependent transmission data, left axis; optical contrast $|\Delta\mathrm{T}|$, right axis), as a function of the bars length; the waveguide wall boundary is reported to guide the eye. c) Optical results of the connected device, for a bar length of 650 nm (the axis choice is the same as in (b)). The data is reported as a function of the ITO-bar overlap size.

Appendices

Appendix I. PDE formalism and solving methods

The FEM environment needs to provide the solver with an equation system in the form $Ax = b$. Albeit the equation set is defined (in COMSOL) by Physics and Features, the calculation can only be performed after the meshing process, as this truly defines the *Degrees Of Freedom* (DOF), i.e. the size, of the system matrices. To clarify the meaning of DOF, and the numerical processes leading to the objective solution x, the following explanation focuses on a specific example case, as a general description can not be successfully summarized within the space of an appendix. Additional information can be found in Refs. [185], [327]–[330].

I.1. Theoretical elements of the finite element method

Let the case be of a Helmholtz equation in the variable u defined in the spatial domain Ω:

$$-\nabla \cdot (\mathrm{c}\nabla u) + \mathrm{a}u = f, \mathrm{in}\ \Omega \qquad \text{Eq. A.1}$$

The following boundary conditions are also assumed:

$$\mathbf{n} \cdot (\mathrm{c}\nabla u + \mathrm{a}u) = \mathrm{g} - \mathrm{q}u + \mathrm{h}^{\mathrm{T}}\mu,\ \text{ on } \partial\Omega_{\mathrm{N}} \qquad \text{Eq. A.2a}$$

$$0 = \mathrm{R}(u),\ \mathrm{on}\ \partial\Omega_{\mathrm{D}} \qquad \text{Eq. A.2b}$$

$$u = \mathrm{r},\ \mathrm{on}\ \partial\Omega_{\mathrm{D}} \qquad \text{Eq. A.2c}$$

where Eq. A.2a is the Neumann boundary condition, acting on the u derivative, and Eq. A.2b-c are Dirichlet boundary conditions, imposing a value on u. To further understand these expressions, it is advantageous to specify in advance the meaning and roles of the involved parameters.

- $\mathbf{n}$ is the unitary vector normal to the $\partial\Omega$ boundary of Ω.
- R is an arbitrary expression to be evaluated on the $\partial\Omega$ boundary
- r is an arbitrary value that u needs to assume at the boundary
- the c, a, g, q terms are user-defined parameters.

The term h^{T} is the (transposed) *constraint force Jacobian*, which is calculated as:

$$\mathrm{h} = -\left(\frac{d\mathrm{R}(u)}{du}\right) \qquad \text{Eq. A.3}$$

The term μ is the *Lagrange multiplier*, an independent and unknown variable which needs to be adjusted so as to enforce the desired boundary condition.

With the necessary definitions being provided, it is possible to carry out the first needed mathematical operation, which is the conversion of the PDE strong formulation expressed in Eq. A.1 to the weak formulation. This can be attained via multiplication with a purposely chosen test function v, defined in the whole Ω domain but vanishing on the constrained portions of the $\partial\Omega$ boundary, and integration:

$$\int_{\Omega} (-\nabla \cdot (\mathrm{c}\nabla u) + \mathrm{a}u - f)v \, d\Omega = 0 \qquad \text{Eq. A.4}$$

Via integration by parts, it is possible to write:

$$\int_{\Omega} (\mathrm{c}\nabla u \nabla v + \mathrm{a}uv) d\Omega - \int_{\partial\Omega_{\mathrm{N}}} \mathbf{n} \cdot (\mathrm{c}\nabla u) v d\partial\Omega_{\mathrm{N}} + \int_{\partial\Omega_{\mathrm{D}}} \mathbf{n} \cdot (\mathrm{c}\nabla u) v d\partial\Omega_{\mathrm{D}}$$
$$= \int_{\Omega} f v d\Omega \qquad \text{Eq. A.5}$$

And after implementing the Neumann boundary conditions expressed in Eq. A.2a (also highlighting that $v = 0$ on $\partial\Omega_{\mathrm{D}}$):

$$\int_{\Omega} (\mathrm{c}\nabla u \nabla v + \mathrm{a}uv) d\Omega + \int_{\partial\Omega_{\mathrm{N}}} \mathrm{q}uv d\partial\Omega_{\mathrm{N}} = \int_{\partial\Omega_{\mathrm{N}}} \mathrm{g}v d\partial\Omega_{\mathrm{D}} + \int_{\Omega} f v d\Omega \qquad \text{Eq. A.6}$$

It is now possible to discretize the above equation. This operation allows the application of Eq. A.6 to every subdomain Ω_m, corresponding to the mesh element m. The formalism requires the trivial substitution (in Eq. A.6) of u_m and v_m to u and v respectively, along with the integral range redefinition (Ω_m and $\partial\Omega_{\mathrm{N}m}$), so it won’t be here reported.

To further exemplify the case in analysis, a unidimensional geometry is here assumed, consisting of *n* mesh elements, and *n+1* nodes. Consequently, the functions $u_m(x)$, of the spatial coordinate x, are the solutions sought after. The elements *linearity*[62] is also implied, and defines the form of the additional function by which u_m scales from node to node, via use of a linear *basis function* set $N(x)$. In this example, the function u_m is determined by its solutions at the nodes, which will be referred to as $U_i = u_m(x_i)$, and the *shape functions* $N_i(x)$, which are the non-null subset of the basis functions. In order to shape this definition to suite the discretized form of Eq. A.6, it is convenient to rewrite u_m as a function of a (normalized) dimensionless spatial variable, which is referred to as *local coordinate* ζ. ζ takes the value of 0 at the i^{th} element, and the value of 1 at the $(i \pm 1)^{th}$ element. The $u(\zeta)$ solution is then obtained as follows:

$$u_m(\zeta) = \sum_i U_i N_i(\zeta) \qquad \text{Eq. A.7a}$$

$$N_i(\zeta) = \begin{cases} \text{if } \zeta_{i-1} \le \zeta \le \zeta_{i+1}, & 1-\zeta \\ \text{else}, & 0 \end{cases} \qquad \text{Eq. A.7b}$$

The role of undetermined parameters is now being taken by the scalar values U_i, applied to each i^{th} node. Substituting Eq. A.7 in the discretized form of Eq. A.6, and using N_i as test functions v_m, yields a system of *n* equations and *n* unknown parameters U_i, which would then need to be integrated over each element m to obtain the numerical solution of Eq. A.1. The parameter *n*, defining the equation system size, is indeed the aforementioned DOF parameter.

I.2. Equation system solvers

To calculate the numerical solution, the FEM environment builds the needed equation system in the form $Ax = b$, where the A matrix and b vector are terms containing the integrals of the test functions and user-provided coefficients which define the specific problem, and x is the U_i vector. There are mainly two distinct classes of approaches that allow calculating the x solution vector: *iterative solvers*, and *direct solvers*.

---

[62] Polynomial elements can also be applied. These are more accurate to model curved surfaces, at the price of increased DOF number and hence computational requirements. This detail will not be here discussed for sake of simplicity.

Iterative solvers approach the solution via test of an initial guess x_g, and the iterative refinement of the guess solution via error control. The iterative process ends when the result reaches an arbitrary minimum error or maximum iteration threshold. These methods have the advantage to require a relatively low amount of computational resources, but require a good estimation of the initial guess, without which the calculation leads to slow or no convergence. In the COMSOL models realized in this work, this solver is seldomly used in the *Generalised Minimum RESidual* (GMRES) [331] implementation.

Direct solvers, on the other hand, calculate the solution after calculation of the inverse matrix A', such that $A'Ax = x = A'b$. These methods lead to better numerical results, as the matrix inversion is not afflicted by the systematic arbitrary error present in the iterative approach. Nevertheless, the amount of computational resources is generally much higher, which prevents the indiscriminate use of this approach. In the COMSOL models realized in this work, this approach is used in the *PARallel sparse DIrect linear SOlver* (PARDISO) [332] implementation, which has been empirically found to be the fastest method to be used with the available resources.

The above methods also usually benefit from the use of so-called *preconditioner* or *presmoother* algorithms [333], which aim to fasten the convergence towards the solution (for indirect solvers) or ease the inversion of the matrix (for the direct solvers).

Appendix II. Pointer matrix generation and 3D cellular automata implementation

This Appendix section provides further technical details on the elements needed for the construction of the 3D cellular automata model outlined in Section 4.1.1.

II.1. Generation of the Pointer matrix

It is useful to remind the reader that the construction of the Pointer matrix P is based on the cartesian coordinates (i, j, k) (or (i, j) in the 2D case) array, pointing to each unit cell of the cellular automata. In this work, this set of coordinates is obtained via FEM software, although it is fairly simple to calculate such a set for rectangular (2D) or rectangular parallelepiped (3D) shapes.

Before unravelling the proposed method, it is here noted that the specific arrangement of the neighbours' coordinates in the P matrix is not an important factor, as long as such arrangement remains constant during the construction of the P matrix. For instance, in Figure 4.1 (as indicated by the black arrows and the colour-coded elements), the arbitrary arrangement segregates the neighbourhood by type. This choice is only convenient for practical code implementation, and attributes (to the 1D notation) the highest ordinal value (observing the 2D representation) starting from the top-left position and going anti-clockwise around the (i, j) monomer (green in figure).

From the practical perspective, the construction of the Pointer matrix may be a mathematically trivial, but computationally challenging in the case of irregularly shaped regions; a direct method would have a complexity of $\mathcal{O}(n^a)$, with the parameter a $(1 > a > n)$ determined by the efficiency of the neighbour lookup function. It is here then suggested a fast and reliable algorithm (for a square or cubic unit cell arrangement), with a fixed computational complexity $\mathcal{O}(n)$. The following paragraphs illustrate the 3D concept, as the 2D case can be regarded as a degenerate case of the former. The proposed algorithm can be performed provided the knowledge of the cartesian coordinates of each monomer, as a $M \times 3$ sized – array referred to as A. The algorithm performs the following:

i. Calculation of $(x_{min}, y_{min}, z_{min})$ and $(x_{max}, y_{max}, z_{max})$, from the (x, y, z) monomers coordinates contained in A. These 2 coordinates determine the volume-enclosing rectangular parallelepiped (as the coordinates of its opposite corners)

ii. Determination of the monomer cubic lattice (and the monomer cartesian coordinates), also yielding the column size H, the number of columns per plane W, and the number of HW planes on the depth direction D

iii. Construction of the P matrix for the enclosing rectangular parallelepiped, where for every index (m), the neighbours have the following unidimensional coordinates (or combination of):
   a. $left \rightarrow (m - H)$.
   b. $down \rightarrow (m + 1)$
   c. $right \rightarrow (m + H)$
   d. $up \rightarrow (m - 1)$
   e. $top \rightarrow (m - HW)$
   f. $bottom \rightarrow (m + HW)$

iv. Correction of the P matrix at the surfaces of the rectangular parallelepiped, by overwriting the erroneous coordinates with the nil-pointer coordinate. More specifically, the erroneous coordinates are found at the neighbour locations whose cartesian coordinates contain the versor crossing the region surfaces:
   a. At the top surface (unidimensional coordinates from (1) to (HW)),
   b. At the $bottom$ surface (from $(HW \cdot (D - 1))$ to (HWD))

and then at the following locations, identified with a pivoting coordinate (i) spanning from (1) to (HWD), with a step of HW:

   c. At the $left$ surface (from (i) to $(i + H - 1)$)
   d. At the $down$ surface (from $(i + H - 1)$ to $(i + HW - 1)$, with a step of H)
   e. At the $right$ surface (from $(i + HW - H)$ to $(i + HW - 1)$)
   f. At the up surface (from (i) to $(i + HW - H)$, with a step of H)

v. Determination[63] of the coordinate set excluded from the investigated region, as a boolean array, a_e

---

[63] This operation is here quickly performed via use of the COMSOL Livelink environment, using the previously mentioned coordinate array A, via the `mphinterp` function after empty Study calculation:

```
bool_array = ~isnan( mphinterp( COMSOL_model_object, ...
'1', 'coord', A, 'selection', COMSOL_domain_indices ) );
```

Alternatively, the same result can be obtained by using the coordinate array A (in D-dimensional form) as a mask vector on the rectangular parallelepiped monomer coordinate array (also in D-dimensional form), to produce the desired boolean vector.

vi. Calculation of the array a_C, containing the ordinal number associated with the coordinate set included in the investigated region, obtained as follows: $a_C = (\sum_{cum} \neg a_e) \cdot \neg a_e$

vii. Determination of the new nP value as $nP = \sum \neg a_e + 1$

viii. Substitution with the nP value on the location of the excluded coordinates in a_C

ix. Substitution of the values in P with the ones pointed at in a_C

x. Reduction of P size by eliminating the a_e rows of P.

The proposed algorithm, although not being the best-optimized solution, is reasonably fast and extremely flexible, as it can yield the P matrix for any feasible combinations of monomer allocations in the raster determined by the volume-enclosing rectangular parallelepiped.

II.2. 3D cellular automata framework example

For an example case of a generic 3D PC material region, the P matrix would consist of a $M \times (3^3 - 1) = M \times 26$ array. Figure A. represents the case of a $3 \times 3 \times 3$ cube, which summarizes all the possible configurations of the P matrix for a rectangular parallelepiped volume, a widely applicable 3D shape.

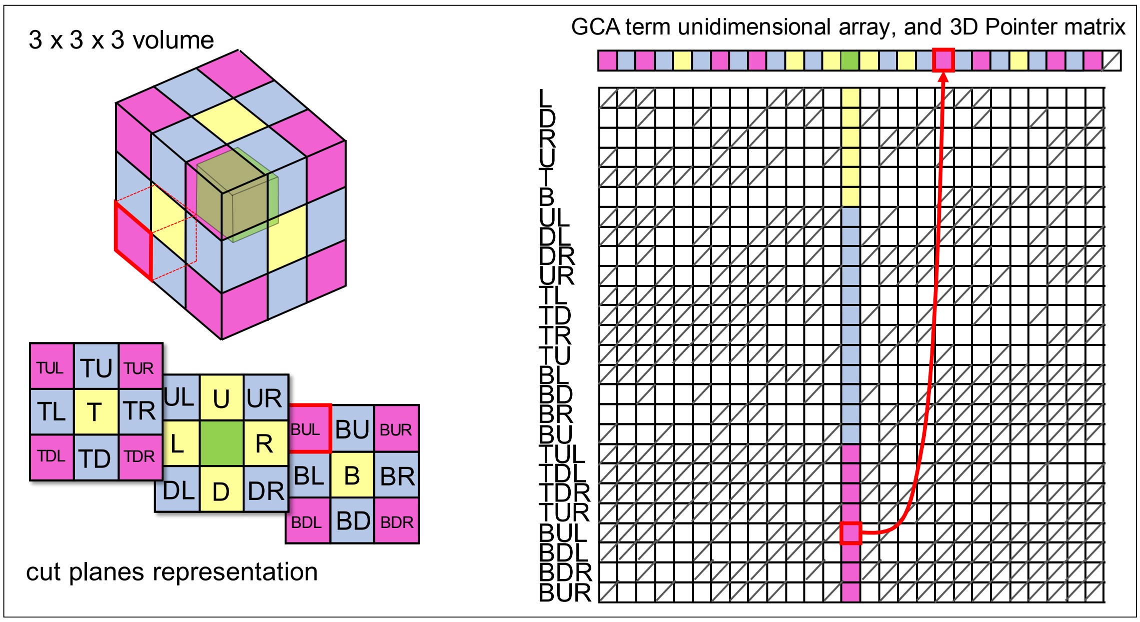

3 x 3 x 3 volume
GCA term unidimensional array, and 3D Pointer matrix
TUL
TU
TUR
TL
T
TR
TDL
TD
TDR
UL
U
UR
L
R
DL
D
DR
BUL
BU
BUR
BL
B
BR
BDL
BD
BDR
cut planes representation
L
D
R
U
T
B
UL
DL
DR
UR
TL
TD
TR
TU
BL
BD
BR
BU
TUL
TDL
TDR
TUR
BUL
BDL
BDR
BUR

Figure A.1. Representation via unidimensional array and Pointer matrix of a 3D volume. The abbreviated notations L (left), R (right), U (up), D (down), T (top), B (bottom) are here adopted for sake of conciseness. The red features, illustrating the selection of the *bottom-up-left* neighbour of the highlighted green monomer (at the centre of the volume), demonstrates how the procedure is carried out exactly as in the 2D cases illustrated in Figure 4.1(b) and Figure 4.2(b).

II.3. Periodic boundary conditions

On a different note, the explicit coordinate notation allows for easy implementation of the periodic boundary conditions, albeit only for a 2D rectangular or 3D rectangular parallelepiped region. In such a case, the algorithm would only perform tasks i-iv; during task iv, it would need to assign the indices corresponding to the opposite surface (to the one operating on) in place of nP. These coordinates are calculated as follows:

a. $left \rightarrow (m) = \big(m + H(W-1)\big)$.
a. $down \rightarrow (m) = (m + H - 1)$
b. $right \rightarrow (m) = \big(m - H(W-1)\big)$
c. $up \rightarrow (m) = (m - H + 1)$
d. $top \rightarrow (m) = (m + HW)$
e. $bottom \rightarrow (m) = (m - HW)$

II.4. 3D framework: unit cell relative distance

An additional element modifying the GCA 3D implementation with respect to the original 2D model is the distance between monomers. This factor has not been explicitly discussed before, but is a crucial element for the GCA rates calculation. As reported in Ashwin et al. [220], the distance is a parameter which (by assumption) is inversely proportional to the rate of interaction between vicinal sites, and acts as a weighting factor on the neighbour count sum in Eq. 3.9f. The 2D square lattice case considers two normalized distances w: unitary, for the 4 edge-sharing monomers, and $w = \sqrt{1^2 + 1^2} = \sqrt{2}$, for the 4 corner-sharing monomers (as derived from the calculation of the square lattice diagonal). In a cubic lattice, the following is encountered: there are 6 face-sharing monomers, having unitary normalized distance; $w = \sqrt{2}$ for 12 edge-sharing monomers; and $w = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$ for 8 corner-sharing monomers.

Appendix III. Artificial reference method

This Appendix section demonstrates how, combining the available reference set and the experimental data, it is possible to build an *artificial reference*, which allows one to drastically improve the algorithm reliability where the literature data does not allow a good fit around resonance. The main idea is to make use of two elements of knowledge which are not directly exploited from the previous method: the uniqueness of the solution at resonance wavelength, and the distribution of the low MSE solutions on the n/κ plane (as previously seen in Figure 5.2).

III.1. Method description

The artificial reference construction method starts by pinning the n/κ values at the FP resonance, found by minimizing the difference between simulated and experimental spectra; it then infers the n/κ wavelength dependency, by combining i) the integration of the n/κ coordinates after application of the $d/d\lambda$ derivative of the literature data average, and ii) the probability given by the reflectance delta $|\Delta\mathrm{R}|$ between experimental reflectance and any numerically determined reflectance value on the n/κ plane. In more detail, the proposed artificial reference construction method performs the following algorithmic steps[64]:

1. Builds the $\Delta\mathrm{R}(\lambda)$ planes[65] (see Figure A.2(a), top row) as the absolute value of the difference between i) experimental reflectance, and ii) simulated reflectance applying the $\varepsilon'/\varepsilon''$ coordinates to the cavity material
2. Finds the minimum $\Delta\mathrm{R}(\lambda_{\mathrm{res}})$ centroid $\tilde{\varepsilon}_{\Delta\mathrm{R}}$ coordinates, and $\Delta\tilde{\varepsilon}$ centroid radii
3. Calculates the references set average $\tilde{\varepsilon}_m(\lambda)$
4. Builds a translation vector $v(\lambda_{res}) = \delta\varepsilon_1 + i\delta\varepsilon_2$, which shifts $\tilde{\varepsilon}_m(\lambda_{res})$ towards $\tilde{\varepsilon}_{\Delta\mathrm{R}}$
5. Calculates the derivative $d\tilde{\varepsilon}_m(\lambda)/d\lambda$

---

[64] It is here highlighted once more that the code implementation is based on the dielectric function, in place of the complex index, as these are mutually interchangeable for non-magnetic materials.

[65] To avoid computational artifacts, over- and under-sampling, and high computational cost, any 2D area calculation is performed by progressive annealing of the coordinates of an initially random set of sampling points, driven by the objective of the calculation (the annealing is here aimed towards lower ΔR values). This subset of randomly distributed solutions is then upscaled to a high-resolved, regularly sampled solution plane, by use of Delaunay triangulation.

6. Calculates the wavelength-dependent $v(\lambda)$, starting from $v(\lambda_{res})$, by use of $v(\lambda + \Delta\lambda) = v(\lambda) \cdot d\tilde{\varepsilon}_m(\lambda + \Delta\lambda)/d\tilde{\varepsilon}_m(\lambda)$
7. Applies $v(\lambda)$ to $\tilde{\varepsilon}_m(\lambda)$, to find a *guess solution* $\tilde{\varepsilon}_{guess}(\lambda)$
8. Builds a 2D Gaussian distribution $\mathrm{G}(\lambda)$ (see Figure A.2(a), middle row), centred in $\tilde{\varepsilon}_{guess}(\lambda)$ and with standard deviation $\Delta\tilde{\varepsilon}$
9. Multiplies $\Delta\mathrm{R}(\lambda)$ and $\mathrm{G}(\lambda)$, yielding a combined probability space $\mathrm{C}(\lambda)$ (see Figure A.2(a), bottom row
10. Calculates the coordinates $\tilde{\varepsilon}_c(\lambda)$ of the combined space $\mathrm{C}(\lambda)$ weighted average, with the probability acting as the weighting factor (here, an arbitrary threshold can be applied to integrate only the relevant area)
11. Converts $\tilde{\varepsilon}_c(\lambda)$ in $\tilde{n}_c(\lambda)$, and fits the derived values set with a suitable function, to smooth its otherwise unphysical noise (see Figure A.2(b), where a 3rd order polynomial is used).

The resulting artificial $\tilde{n}_c(\lambda)$ reference has the potential to accrue a significantly closer match to the real material properties around the resonant wavelength, as it combines the TMM mathematical solutions with a guess based on the existing reference sets. An algorithmic flowchart is also provided in Figure A.3.

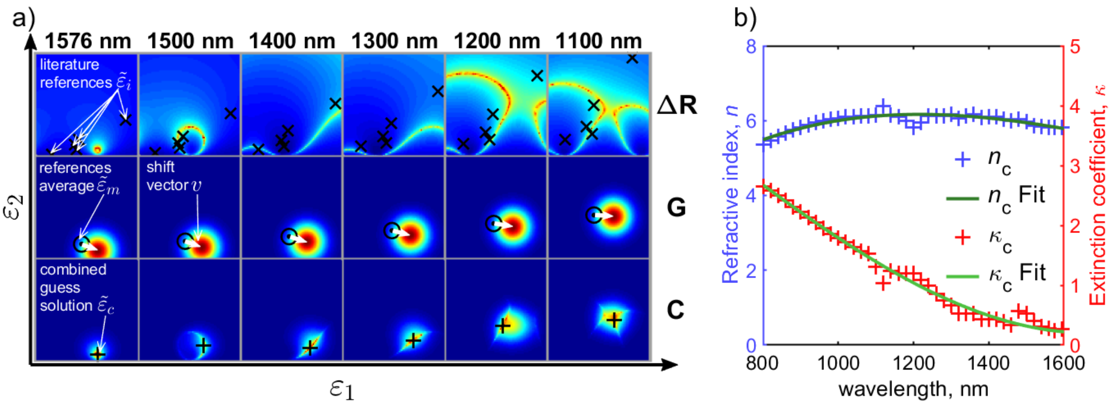


Figure A.2. Artificial reference calculation process, for the crystalline GeTe sample case. (a) Colour-coded data representation on the $\varepsilon_1/\varepsilon_2$ plane (ε_1 ranging from 10 to 50; ε_2 ranging from 0 to 40). The first column reports the data calculated at the resonant wavelength; columns 2 to 6 show additional frames as a function of the wavelength (values reported above). Top row: |ΔR| (blue-to-red pointing to higher-to-lower |ΔR| values), with crosses representing the selected Refs. [155], [244], [246]–[248]. Middle row: guess Gaussian distribution G (normalized probability data, blue-to-red pointing to lower-to-higher probability), with the circle representing the Refs. Average and the arrow corresponding to the shift vector v. Bottom row: combined probability C = ΔR×G (blue-to-red pointing to lower-to-higher probability), with the cross representing the C weighted average position. (b) Combined space n and κ values (crosses) calculated from the coordinates of the C weighted average values, as shown in (a)-bottom row. Continuous green lines are the 3rd order polynomial fits of the calculated data, and correspond to the resulting artificial reference.

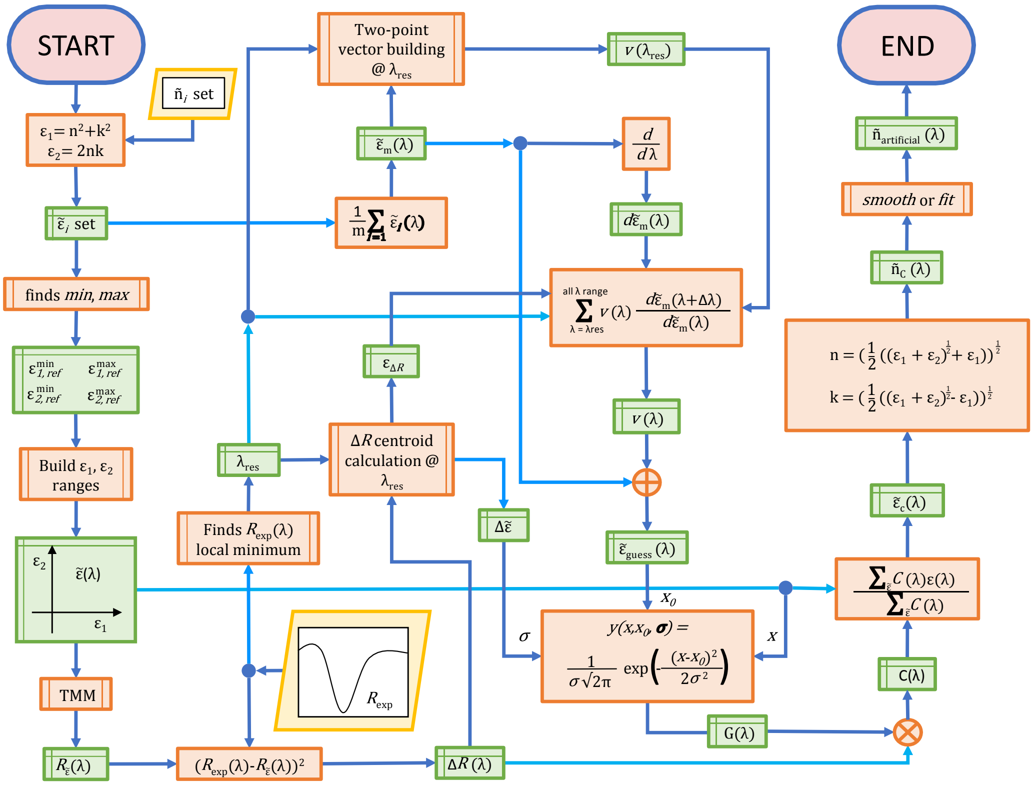

START
ñ_i set
ε1 = n²+k²
ε2 = 2nk
ε̃_i set
finds min, max
ε^min_1,ref ε^max_1,ref
ε^min_2,ref ε^max_2,ref
Build ε1, ε2 ranges
ε2
ε̃(λ)
ε1
TMM
R_ε̃(λ)
(R_exp(λ)-R_ε̃(λ))²
ΔR (λ)
Finds R_exp(λ) local minimum
R_exp
λ_res
ΔR centroid calculation @ λ_res
ε_ΔR
Δε̃
Two-point vector building @ λ_res
ε̃_m(λ)
1/m Σ ε̃_i(λ)
v(λ_res)
d/dλ
dε̃_m(λ)
all λ range
Σ v(λ) dε̃_m(λ+Δλ)/dε̃_m(λ)
λ = λres
v(λ)
ε̃_guess (λ)
x_0
σ
x
y(x,x_0,σ) = 1/(σ√2π) exp(-(x-x_0)²/2σ²)
G(λ)
END
ñ_artificial (λ)
smooth or fit
ñ_C (λ)
n = (½((ε1 + ε2)^½ + ε1))^½
k = (½((ε1 + ε2)^½ - ε1))^½
ε̃_c(λ)
Σ_ε C(λ)ε(λ) / Σ_ε C(λ)
C(λ)

Figure A.3. Algorithmic flowchart for the artificial reference determination.

III.2. Results and discussion

The results of the GeTe fit and GST fit, using the artificial reference (combined with the literature data) are proposed in Figure A.4 and Figure A.5 respectively. Both cases present an improvement of the fit quality for the reflectance spectrum and a correlated closer match of the derived n/κ values to the experimental ones.

The reflectance fit for the GeTe case (data shown in Figure A.4) reports a MSE value of 1.73×10^{-3}, reducing the former MSE value by a factor of 20. Also, the estimation of the optical properties improves, with a decrease of the relative error to $\bar{\zeta}(n)$ = 2.52% and $\bar{\zeta}(\kappa)$ = 14.5%. It is apparent how this method is particularly beneficial for crystalline GeTe, where the fit quality of the reflectance was not previously acceptable in resonance regime.

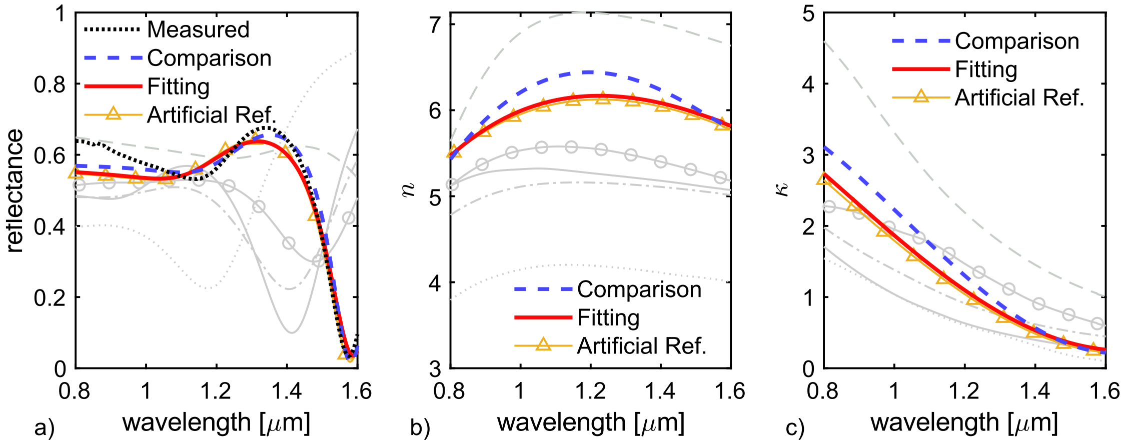


Figure A.4. Alternative fit of the crystalline GeTe data, after construction of the artificial reference based on the literature data reported in Figure 5.7. The black dotted line reports the measured reflectance spectrum. The blue dashed line reports the data related to the ellipsometric estimation of n/κ. The red line reports the fit results for each panel. The orange line highlights the artificial reference amongst the reference pool datasets. a) Reflectance. b) Refractive index. c) Extinction coefficient.

Also the case of crystalline GST (see Figure A.5) benefits from the use of this strategy, although not so significantly: the new reflectance finds a MSE = 9.76×10^{-5}, with a correlated decrease of the average relative errors to $\bar{\zeta}(n)$ = 5.13% and $\bar{\zeta}(\kappa)$ = 11.6%. It is here suggested that this method brings valuable advantage where the literature datasets do allow a suitable reflectance fit within the allowed n/κ values, as it expands the tested n/κ range to include the unique solution at λ_{res}.

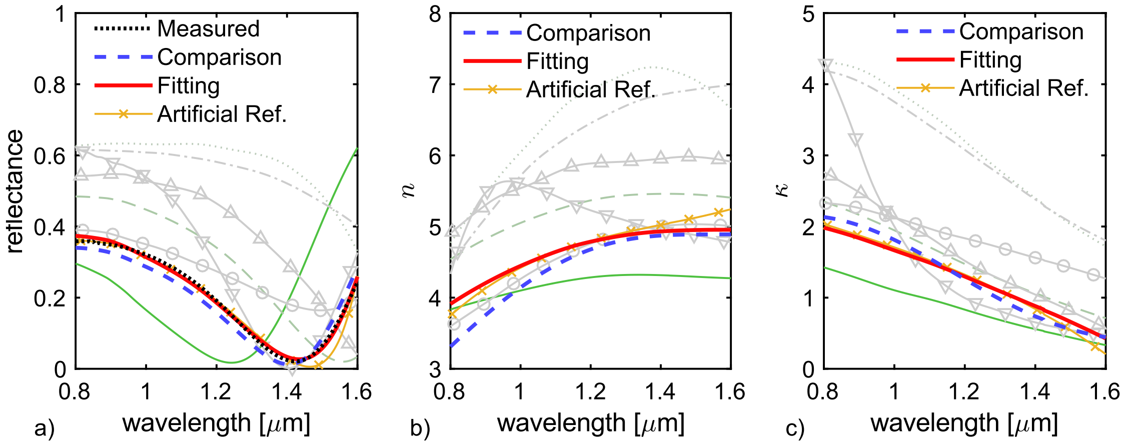


Figure A.5. Alternative fit of the crystalline GST data, after construction of the artificial reference based on the literature data reported in Figure 5.5, and reported with the format adopted in the previous (crystal GeTe) case. a) Reflectance. b) Refractive index. c) Extinction coefficient.

Appendix IV. Use of multiple nanostructures along the waveguide

The concept of the plasmonically-enhanced architecture has drawn inspiration from various publications in a large number of fields, but a notable case is the publication from Jin et al. [288]. There, the authors configure a metallic double-bowtie nanoantenna placed on the top surface of a Si waveguide. This structure efficiently suppresses the propagating mode of the resonant wavelength, with a simulated attenuation of up to 96%. The rationale of using two vicinal nanogaps lays in the interaction of the anti-symmetric resonant modes, which has also been numerically demonstrated to magnify the coupling efficiency between propagating mode and nanoantenna eigenmodes [288]. This avenue of work solves for the fabrication complexity presented from the embedded configuration, and as here demonstrated, has the potential to achieve similar, if not higher, optical modulation. Thus, this section shows the preliminary investigation of the optical effect of the use of multiple structures. It is here noted that, since these memory devices can’t be independently accessed, these are to be considered part of the same memory cell, and will henceforth be referred to as sub-units.

IV.1. Double nanoantennas

An initial simulation task follows the $\Delta\mathrm{T}$ and EF magnitude as a function of the centre-to-centre distance of two dimer-discs nanoantennas, for the amorphous and crystalline GST configuration. The results are portrayed in Figure A.6. All the degenerate case data (i.e. the values obtained with a single unit cell) are reported via markers, to provide a quick comparison reference. As visible, all data plots widely diverge from the behaviour obtained with only one unit cell, following a non-trivial dependency, arising from the mentioned anti-symmetric resonance of the two sub-units (easily inferable from the electric field plots, reported in Figure A.6(e-f)).

It is here noted that, close to the distance of 600 nm, the data plots seem to asymptotically reach towards the values belonging to non-interacting sub-units. Whilst this is evident for the transmission data, the same is expected to eventually take place for all datasets. Those marginal values are also remarkably lower than the ones obtainable by centre-to-centre distance optimization, confirming the results of Jin et al. In contrast, as the distance closes towards proximal sub-units, the arising interaction leads to detrimental optical characteristics (decreased EF, absorption, $\Delta\mathrm{T}$ and increased IL).

At the distance of 280 nm, the theoretical $\Delta\mathrm{T}$ (Figure A.6(b), green) sees a local peak of value $\Delta\mathrm{T} = 40.2\%$; and an interesting decrease of the IL is also observed at the same location (due to the increase of the transmission in the amorphous GST case, see Figure A.6(a)). At the same distance, the EF plot also shows increasing EF values (up to three times the ones calculated on a single structure), in both the amorphous and crystalline cases. As a direct and beneficial consequence, also the GST absorption is much increased (Figure A.6(d)), suggesting a more energy-efficient switching dynamics.

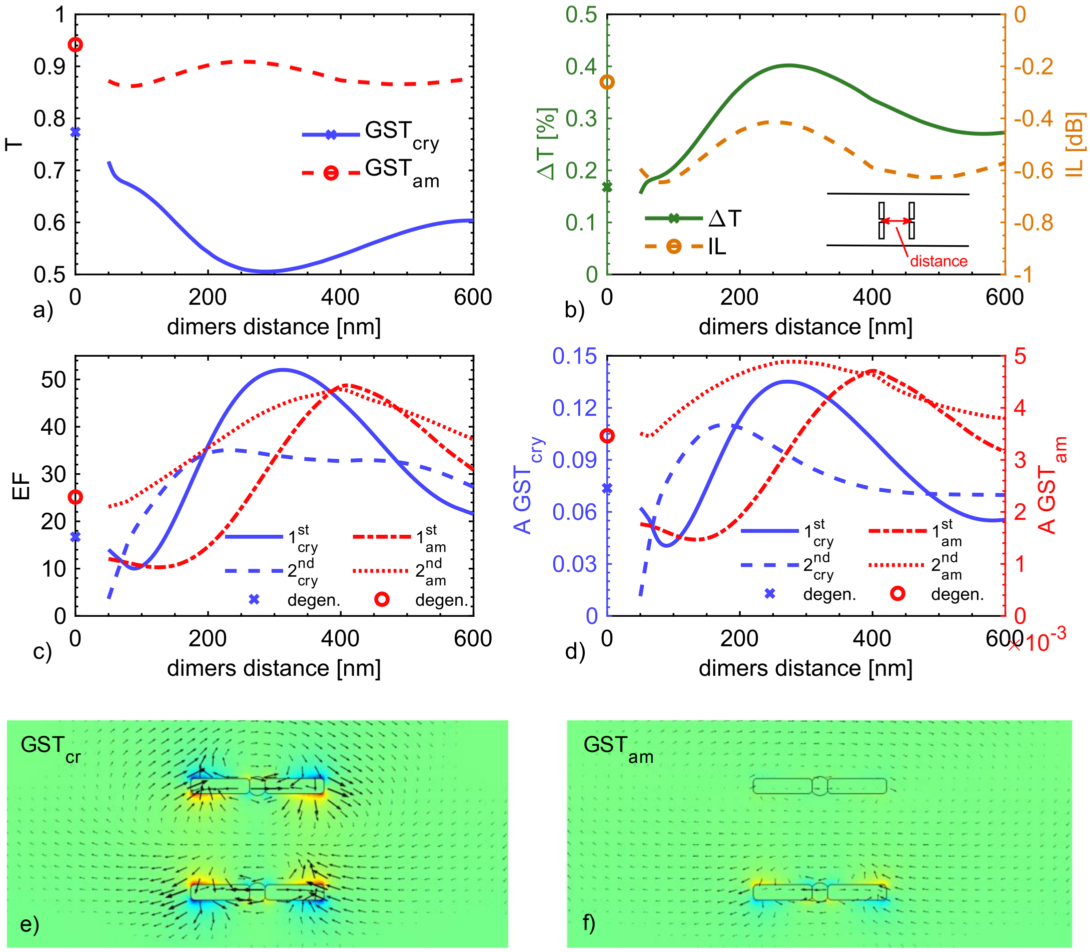


Figure A.6. Simulation results for two plasmonically-enhanced (dimer-bars) PPCM devices, fabricated on the waveguide top surface, as a function of the center-to-center distance (see inset in (b) for a graphical representation). The degenerate case data (corresponding to a single unit cell) is reported via markers at the $x = 0$ coordinate. a) Transmission data, for the amorphous (red, dashed) and crystalline (blue, continuous) GST. b) Left (green): theoretical $\Delta\mathrm{T}$. Right (orange): theoretical minimum IL. c) EF on the first and second sub-unit cells for the amorphous (red) and crystalline (blue) GST cases. d) GST absorption on the two sub-unit cells, for the amorphous (red) and crystalline (blue) cases. (e,f) Electric field distribution plot (on the sub-unit cells cut plane, top view), calculated at the distance of 280 nm.

These results confirm that the use of two interacting sub-units sharply increase the optical performance, albeit the switching dynamics and pulse scheme would require an attentive investigation to confirm the approach feasibility and avoid sub-units damage. Interestingly, the plots show that a solution in which the two nanostructures show equal behaviour exists, at the distance of 485 nm, where the resonant modes show symmetric (non-interacting) behaviour. This configuration does not show the remarkable advantage brought by the exploitation of the antisymmetric modes, yet it would simplify the determination of the switching pulse schemes.

IV.2. Multiple nanoantennas

In this section, the analysis of the use of multiple sub-units is reported. The aim is to infer the dependency of $\Delta\mathrm{T}$ and EF from the sub-units number $\#$. The results are reported in Figure A.7.

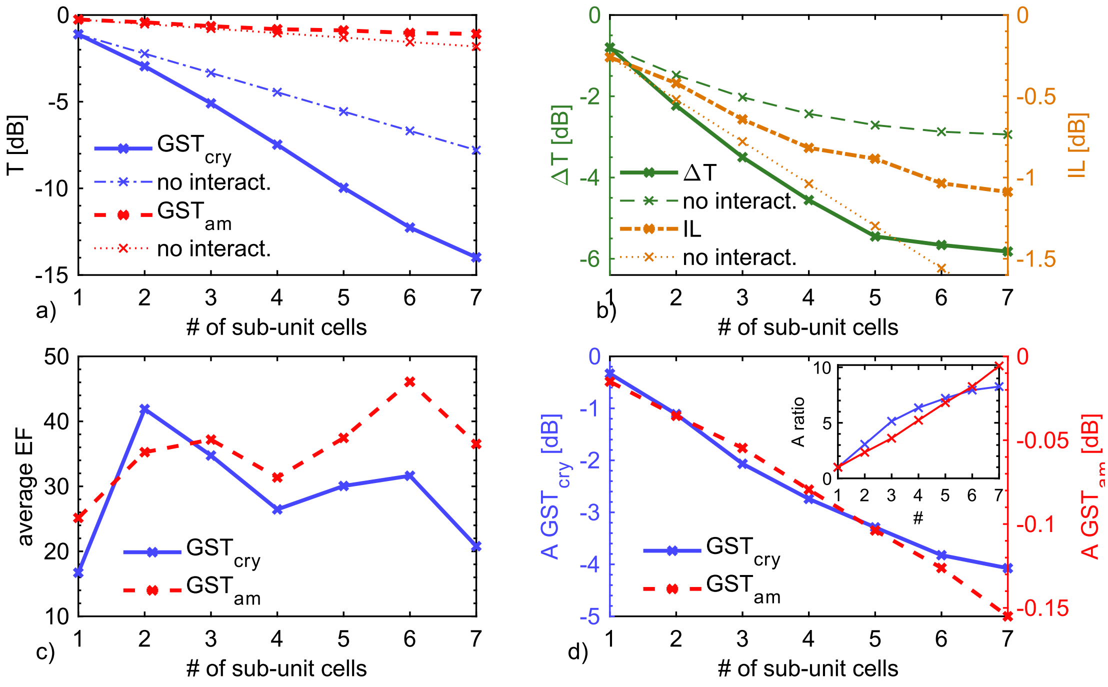


Figure A.7. Optical behaviour of interacting plasmonically-enhanced PPCM devices, with a centre-to-centre distance of 280 nm, as a function of the number of sub-units. In (a) and (b), the extrapolated behaviour of non-interacting sub-units is also reported (thin lines), as a source of comparison. a) Transmission data, for crystal (blue) and amorphous (red) GST. b) Left (green): $\Delta\mathrm{T}$. Right (orange): IL. c) Average EF for amorphous (red) and crystalline (blue) GST sub-unit cells. d) Average GST absorption for crystalline (blue, left axis) and amorphous (red, right axis) phase.

The transmission data (Figure A.7(a)) underlines that the achievable modulation is greatly benefiting from the presence of multiple sub-units under alternating anti-symmetric resonant modes. As a comparison, the plot also reports the extrapolated behaviour deriving from non-interacting sub-unit cells. It is immediately visible how both amorphous and crystalline interacting sub-units yield a transmission plot deviating from the theoretical exponential decrease expected by non-interacting sub-units. Specifically, the crystal case shows a much-reduced transmission data as the number of sub-units increases, whereas the amorphous case is remarkably characterized by a higher transmission than the non-interacting case. The combination of these datasets yields the $\Delta\mathrm{T}$ plots, reported in Figure A.7(b) (left, green). As inferable, the tunable optical modulation widely diverges from the non-interacting case, yielding a maximum $\Delta\mathrm{T}$ of 73.8% for 7 sub-units. It is worth noting though that the use of 5 sub-units already reaches a $\Delta\mathrm{T} =$ 71.4%, at a reduced IL (see Figure A.7(b), right axis); henceforth, 5 sub-units sets the upper limit for the beneficial number of sub-units from the modulation perspective.

The average EF values are reported in Figure A.7(c). These of course show the previously determined increase of the EF for the two sub-units, but also outline that such increase is not maintained for the crystal case, as the number of sub-units grows. This is indeed expected, as the propagating mode energy reduces along the waveguide axis. Of course, as the amorphous case is less optically dense than the crystal case, the average EF increases here more rapidly. It is noteworthy though that a common minimum is found at the value of 4 sub-units.

The average GST absorption is reported in Figure A.7(d). The data points to an increase of the GST absorption, which remarkably brings the crystal case from a minimum value of 7.4% to the maximum value of 60.8% (at $\# = 7$), roughly 8 times the single unit cell case. The amorphous case also shows a similar increase, pointing to the improved absorption efficiency arising from the interacting sub-units. Yet, the ratio of the calculated absorption coefficients over the absorption of a single unit cell shows the progressive reduction of such increase for the crystal case. A linear trend of this ratio would suggest a constant optical-to-thermal energy efficiency, and this is indeed true for amorphous GST. As for the crystal GST, this trend deviates from linearity after 3 sub-units. Therefore, this analysis suggests a decrease of the energy performance for aggregation of sub-units larger than 4, further lowering the previous indication.

IV.3. Conclusion

This analysis confirms that a much-improved optical behaviour arises from the optimized collocation of multiple plasmonically-enhanced PPCM devices, due to the nanoantennas resonant modes mutual influences. In the conventional architecture, a larger optical contrast can be naturally gained by extension of the unit cell length, which is not a viable option for the plasmonically-enhanced architecture. This finding confirms that the reiteration of the unit cell along the waveguide axis largely compensates for the lack of the abovementioned degree of freedom. As this method quantises the $\Delta\mathrm{T}$ increments, arbitrary $\Delta\mathrm{T}$ may be obtained by tuning of the nanoantennas configuration and distance.

The attention is here brought to the alternative dimer-discs and bowtie nanoantenna designs, combined with the embedded configuration.

The bow-tie design investigation reveals a generally marginal improvement with respect to the on-top configuration. Although the transmission modulation (Figure A.8(a)) reaches lower minima than what previously observed in the bars configuration (c.f. Figure 7.11(a)), a strong decrease of the EF values (Figure A.8(b)) is observed, and the separation of the transmission lines after GST inclusion does not allow an increased optical contrast (Figure A.8(c)). In particular, this case shows that the peak $|\Delta\mathrm{T}|$ values are broadly comparable to the ones obtained with the on-top configuration, paired with a decrease of the peak EF values (Figure A.8(b)) and an increase of the IL (Figure A.8(d)). Thus, the choice of this configuration is detrimental to the maximization of the optical contrast.

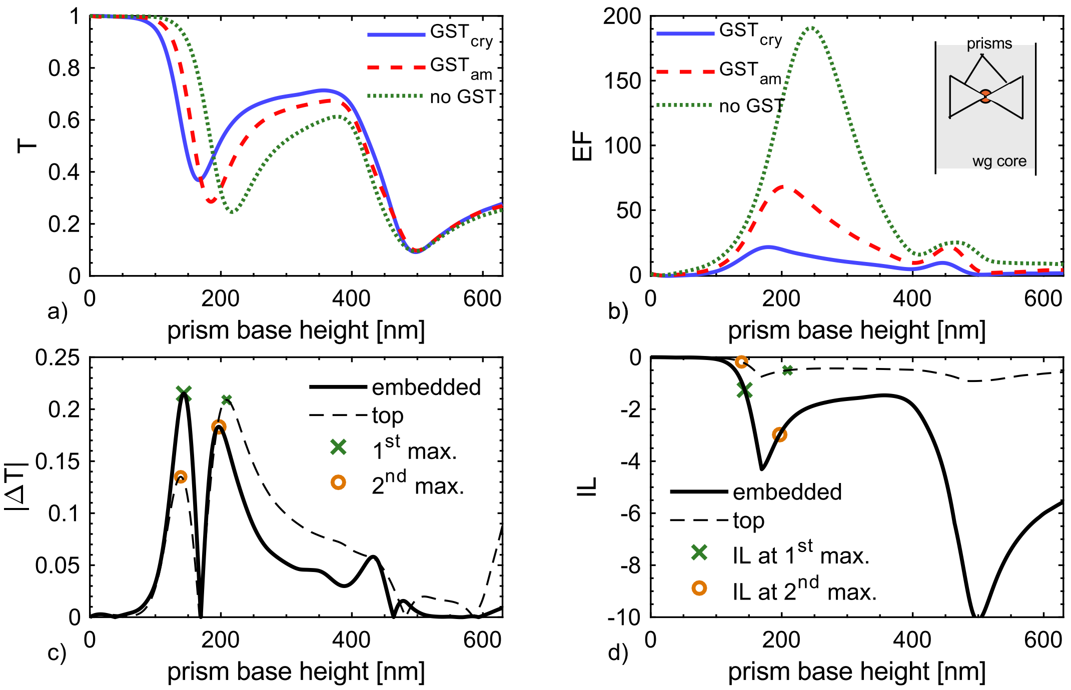


Figure A.8. Simulated optical properties of an embedded bow-tie nanoantenna, of thickness 30 nm, as a function of the prism base height. (a) Transmission data with and without GST inclusion. (b) EF, with and without GST inclusion. (c) Theoretical maximum $|\Delta\mathrm{T}|$ (d) Minimum insertion loss. The markers underline the values calculated at the $|\Delta\mathrm{T}|$ maxima found in (c). Data of the on-top configuration is also provided in (c) and (d).

The discs configuration (Figure A.9), shows the largest contribution to scattering at larger diameters, which hinders its viability for the intended scope. The EF is also increased, again to values comparable to the ones evaluated with the bars design. EF values of 37 and 41 are calculated for crystalline and amorphous GST respectively at the same $|\Delta \mathrm{T}|$ peak diameter.

Due to the larger and better-resolved (phase-dependent) resonance features, the $|\Delta \mathrm{T}|$ values (Figure A.9(c)) resulting from the combination of the transmission data show a peak value exceeding 36% (with an IL of -0.77 dB, see Figure A.9(d)) at the diameter of 136 nm. These are indeed comparable results to the ones observed with the bars design result.

Thus, whilst the quantitative comparison between the results of the bars and discs designs reveals an improved performance of the former case, the latter remains a suitable solution, with the advantage of a simpler fabrication design.

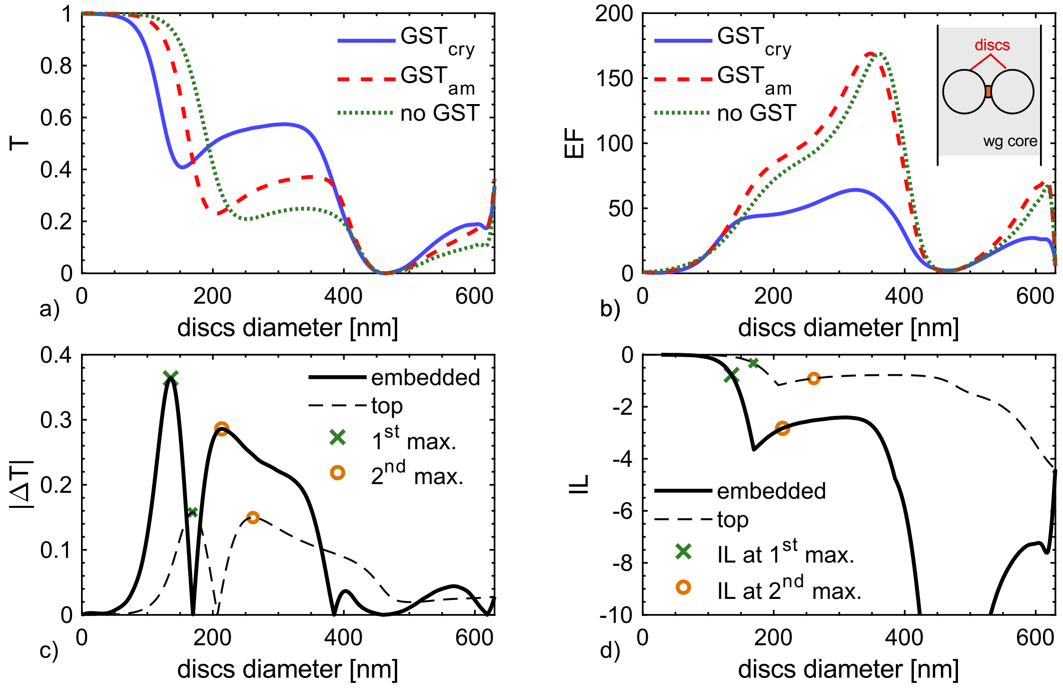


Figure A.9. Simulated optical properties of an embedded dimer-discs nanoantenna, of thickness 30 nm, as a function of the discs' diameter. (a) Transmission data with and without GS inclusion. (b) EF, with and without GS inclusion. (c) Theoretical maximum $|\Delta \mathrm{T}|$ (d) Minimum insertion loss. The markers underline the values calculated at the $|\Delta \mathrm{T}|$ maxima found in (c). Data of the on-top configuration (reported from Figure 7.13) is also provided in (c) and (d).

Appendix VI. Fabrication tolerances of the dimer-discs device

A preliminary analysis of the fabrication tolerance of the device analysed in Section 7.4 (specifically, for a disc diameter of 150 nm) is here performed. The displacement s from the waveguide symmetry axis, and the structure tilt angle θ from the aimed orthogonal orientation are here investigated. The tilt angle θ is here defined as the angle formed by the line cutting through the discs' centres and the orthogonal line to the waveguide axis. Data is reported in Figure A.10, and reveals that the device is relatively tolerant over both displacement and tilting. To maintain the device peak $\Delta\mathrm{T}$ within ~ 90% of its optimal value, it is sufficient to maintain a displacement $|\delta s| < 120$ nm, a tilt angle $|\delta\theta| < 18°$.

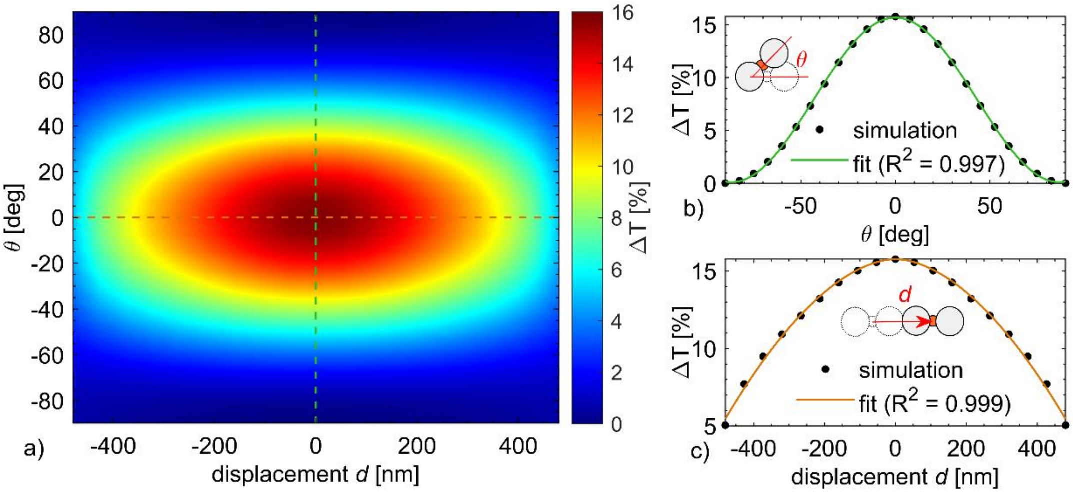


Figure A.10. Fabrication tolerance analysis of the proposed dimer-disc nanoantenna, with respect to displacement from the waveguide centre, and from tilting angle θ from the intended orthogonal orientation. a) Simulated nΔT, as a function of displacement and tilt angle. b) Fit of the data as a function of the tilt angle, calculated at displacement $s = 0$. c) Fit of the data as a function of the displacement, calculated at $\theta = 0$.

The transmission coefficient T strictly follows a $a \cdot \sin^2\theta + c$ dependency from the tilt angle θ, here defined as the angle between the line passing through the discs' centres and the orthogonal line to the waveguide axis (which would be its intended orientation). As the constant c is remarkably low for both amorphous and crystal phases, the $\Delta\mathrm{T}$ dependency can be approximated as follows:

$$\Delta\mathrm{T}(\theta) = \Delta\mathrm{T}(0) \cdot \sin^2\theta \qquad \text{Eq. A.8}$$

As for the displacement dependency, the analytical correlation is less clear, as it seemingly also depends on θ. It is possible though to phenomenologically describe the $\Delta\mathrm{T}$ dependency as a power function of the displacement s:

$$\Delta\mathrm{T}(s,\theta) = \Delta\mathrm{T}(\theta) \cdot b(\theta) \cdot s^{1.8} \qquad \text{Eq. A.9a}$$

with

$$b(\theta) = 10^{-6} \cdot (1 + \sin^2 \theta) \qquad \text{Eq. A.9b}$$

From the above analysis, it is possible to derive that an initial estimation of the device performance can also be obtained from the AFM or SEM measurement of the fabricated device. The device optical behaviour may be inferred by use of the equations above, following the calculation of the $\Delta\mathrm{T}(0)$ value via simulation of the ideal configuration.

Appendix VII. Insertion of an anti-diffusion barrier in the dimer-discs device

The implementation of a dielectric thin layer between GST and silver can be adopted to address the well-known GST-metal inter-diffusion problem. The proof-of-concept design illustrated above removes any element not strictly required to prove the feasibility of the approach, in view of following experimental verification. Yet, for sake of completeness, a simulation is performed after implementation of a 3 nm layer of Si between GST and the metallic nanoantenna (achievable by isotropic deposition techniques), to demonstrate how it affects the device performances and optical-thermal behaviour. Table A.1 reports a summary of such calculations.

Table A.1. Summary of the plasmonically-enhanced PPCM behaviour, after implementation of a 3 nm Si anti-diffusion barrier

	T	A	A (GST)	EF	$\Delta \mathrm{T} = \mathrm{T}_{am} - \mathrm{T}_{cry}$	Write operation (1 mW / 2 ns)
GST$_{cry}$	0.88	0.036	0.031	11.8	8.1%	Max. temperature: 969K Crystal fraction: 0.9
GST$_{am}$	0.96	0.004	0.002	16.8		

As visible, this solution quantitatively affects the plasmonic cell performance, but does not hinder the possibility to retain a useful contrast between the two GST phases. A decrease in the number of available memory levels is expected, as well as a decrease in the switching performances due to the lower GST absorption. The delivery of the previously described Write pulse here increases the crystalline GST temperature to a maximum of 969 K, with the resulting crystal fraction being X = 0.9. This proves how the structure retains the significant efficiency found for the device without the barriers implementation, although diminished in magnitude with respect to the write operation. As a comparison, the effect of a 1.5 mW pulse for 2 ns decreases the crystallinity to the value of X = 0.01 (with a peak temperature of 1500K), proving that the pulse optimization can overcome the efficiency decrease and retain a substantial improvement over the conventional architecture.

Based on this preliminary analysis, an increase in the energy requirement is estimated by a factor of 1.5. It is also worthwhile to highlight that the analysed configuration has a lower GST volume, due to the gap volume intended for the dielectric barrier. To maintain the previously found higher contrast, one would need to further optimize the structure configuration, with the possible choice to widen the nanoantenna gap or increase the structure thickness (to nullify the GST volume reduction), or pursue any of the illustrated performance-enhancing nanoantenna designs.

References

[1] J. Oliverio, “A Survey of Social Media, Big Data, Data Mining, and Analytics,” *J. Ind. Integr. Manag.*, vol. 03, no. 03, p. 1850003, Sep. 2018.

[2] A. Botta, W. de Donato, V. Persico, and A. Pescapé, “Integration of Cloud computing and Internet of Things: A survey,” *Futur. Gener. Comput. Syst.*, vol. 56, pp. 684–700, Mar. 2016.

[3] J. Burgess and J. Green, *YouTube: Online Video and Participatory Culture*. John Wiley & Sons, 2018.

[4] T. Böttger, F. Cuadrado, G. Tyson, I. Castro, and S. Uhlig, “Open Connect Everywhere,” *ACM SIGCOMM Comput. Commun. Rev.*, vol. 48, no. 1, pp. 28–34, Apr. 2018.

[5] J. R. Smith, “Riding the multimedia big data wave,” in *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, 2013, pp. 1–2.

[6] J. Gantz and D. Reinsel, “The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east,” *IDC iView IDC Anal. Futur.*, vol. 2007, no. 2012, pp. 1–16, Dec. 2012.

[7] H. Rong, H. Zhang, S. Xiao, C. Li, and C. Hu, “Optimizing energy consumption for data centers,” *Renew. Sustain. Energy Rev.*, vol. 58, pp. 674–691, May 2016.

[8] A. Andrae and T. Edler, “On Global Electricity Usage of Communication Technology: Trends to 2030,” *Challenges*, vol. 6, no. 1, pp. 117–157, Apr. 2015.

[9] J. von Neumann, “First draft of a report on the EDVAC,” *IEEE Ann. Hist. Comput.*, vol. 15, no. 4, pp. 27–75, 1993.

[10] J. Backus, “Can programming be liberated from the von Neumann style?,” *Commun. ACM*, vol. 21, no. 8, pp. 613–641, Aug. 1978.

[11] C. Sun *et al.*, “Single-chip microprocessor that communicates directly using light,” *Nature*, vol. 528, no. 7583, pp. 534–538, Dec. 2015.

[12] J. Lazzaro and C. Mead, “Circuit Models of Sensory Transduction in the Cochlea,” in *Analog VLSI Implementation of Neural Systems*, Springer,

1989, pp. 85–101.

[13] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85–117, Jan. 2015.

[14] J. Misra and I. Saha, “Artificial neural networks in hardware: A survey of two decades of progress,” *Neurocomputing*, vol. 74, no. 1–3, pp. 239–255, 2010.

[15] A. Sebastian, M. Le Gallo, and E. Eleftheriou, “Computational phase-change memory: beyond von Neumann computing,” *J. Phys. D. Appl. Phys.*, vol. 52, no. 44, p. 443002, Oct. 2019.

[16] G. Tanaka *et al.*, “Recent advances in physical reservoir computing: A review,” *Neural Networks*, vol. 115, pp. 100–123, Mar. 2019.

[17] W. A. Wulf and S. A. McKee, “Hitting the memory wall,” *ACM SIGARCH Comput. Archit. News*, vol. 23, no. 1, pp. 20–24, Mar. 1995.

[18] G. E. Moore, “Cramming more components onto integrated circuits,” *Electronics*, vol. 38, no. 8, Apr. 1965.

[19] L. B. Kish, “End of Moore’s law: thermal (noise) death of integration in micro and nano electronics,” *Phys. Lett. A*, vol. 305, no. 3–4, pp. 144–149, Dec. 2002.

[20] M. M. Waldrop, “The chips are down for Moore’s law,” *Nat. News*, vol. 530, no. 7589, p. 144, Feb. 2016.

[21] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, “Design of ion-implanted MOSFET’s with very small physical dimensions,” *IEEE J. Solid-State Circuits*, vol. 9, no. 5, pp. 256–268, Oct. 1974.

[22] M. Lenzlinger and E. H. Snow, “Fowler-Nordheim Tunneling into Thermally Grown SiO 2,” *J. Appl. Phys.*, vol. 40, no. 1, pp. 278–283, Jan. 1969.

[23] “https://www.sisoftware.co.uk/,” 2020. [Online]. Available: https://www.sisoftware.co.uk/. [Accessed: 01-Sep-2020].

[24] J. Boukhobza and P. Olivier, *Flash Memory Integration: Performance and Energy Issues*. Elsevier, 2017.

[25] H. Cai, W. Kang, Y. Wang, L. Naviner, J. Yang, and W. Zhao, “High

Performance MRAM with Spin-Transfer-Torque and Voltage-Controlled Magnetic Anisotropy Effects,” *Appl. Sci.*, vol. 7, no. 9, p. 929, Sep. 2017.

[26] D. Apalkov *et al.*, “Spin-transfer torque magnetic random access memory (STT-MRAM),” *ACM J. Emerg. Technol. Comput. Syst.*, vol. 9, no. 2, pp. 1–35, May 2013.

[27] H. Ishiwara, M. Okuyama, and Y. Arimoto, *Ferroelectric random access memories: fundamentals and applications*, vol. 93. Springer Science & Business Media, 2004.

[28] K. Kuan and T. Adegbija, “Energy-Efficient Runtime Adaptable L1 STT-RAM Cache Design,” *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 39, no. 6, pp. 1328–1339, Jun. 2020.

[29] A. Pirovano *et al.*, “Reliability Study of Phase-Change Nonvolatile Memories,” *IEEE Trans. Device Mater. Reliab.*, vol. 4, no. 3, pp. 422–427, Sep. 2004.

[30] P. Clarke, “Intel Micron launch ‘bulk-switching’ ReRAM,” *EE Times*, Jul. 2015.

[31] S. Raoux and M. Wuttig, *Phase Change Materials*. Springer Science & Business Media, 2009.

[32] Z. Shao, X. Cao, H. Luo, and P. Jin, “Recent progress in the phase-transition mechanism and modulation of vanadium dioxide materials,” *NPG Asia Mater.*, vol. 10, no. 7, pp. 581–605, Jul. 2018.

[33] S. R. Ovshinsky, “Symmetrical current controlling device,” 3,271,591, 06-Sep-1966.

[34] S. R. Ovshinsky, “Reversible Electrical Switching Phenomena in Disordered Structures,” *Phys. Rev. Lett.*, vol. 21, no. 20, pp. 1450–1453, Nov. 1968.

[35] E. J. Evans, J. H. Helbers, and S. R. Ovshinsky, “Reversible conductivity transformations in chalcogenide alloy films,” *J. Non. Cryst. Solids*, vol. 2, pp. 334–346, Jan. 1970.

[36] J. Feinleib, J. DeNeufville, S. C. Moss, and S. R. Ovshinsky, “Rapid reversible light-induced crystallization of amorphous semiconductors,” *Appl. Phys. Lett.*, vol. 18, no. 6, pp. 254–257, Mar. 1971.

[37] T. G. Brown, K. Creath, H. Kogelnik, M. A. Kriss, J. Schmit, and M. J. Weber, *The Optics Encyclopedia: Basic Foundations and Practical Applications 2 (GL)*. Weinheim: Wiley-VCH, 2004.

[38] E. Hamada, Y. Shin, and T. Ishiguro, “CD-Compatible Write-Once Disc With High Refrectivity,” in *Optical Data Storage Topical Meeting*, 1989, vol. 1078, p. 80.

[39] N. Yamada, E. Ohno, N. Akahira, K. Nishiuchi, K. Nagata, and M. Takao, “High Speed Overwritable Phase Change Optical Disk Material,” *Jpn. J. Appl. Phys.*, vol. 26, no. S4, p. 61, Jan. 1987.

[40] S. R. Ovshinsky, W. Czubatyj, D. A. Strand, P. J. Klersy, S. Kostylev, and B. Pashmakov, “Memory element with memory material comprising phase-change material and dielectric material,” 6,087,674, 11-Jul-2000.

[41] B.-H. Cho, D.-E. Kim, B.-G. Choi, and C. Kwak, “Phase change random access memory (PRAM) device,” 7,227,776, 05-Jun-2007.

[42] M. Joodaki, *Selected advances in nanoelectronic devices: logic, memory and RF*, vol. 175. Springer Science & Business Media, 2012.

[43] P. Noé, C. Vallée, F. F. F. Hippert, F. Fillot, and J.-Y. Raty, “Phase-change materials for non-volatile memory devices: from technological challenges to materials science issues,” *Semicond. Sci. Technol.*, vol. 33, no. 1, p. 013002, Jan. 2018.

[44] F. Bedeschi *et al.*, “4-Mb MOSFET-selected phase-change memory experimental chip,” in *Proceedings of the 30th European Solid-State Circuits Conference*, 2004, pp. 207–210.

[45] L. Hoddeson, P. Garrett, and G. Wicker, “Stanford Ovshinsky and the Genesis of the Cognitive Computer,” *Proc. IEEE*, vol. 107, no. 7, pp. 1457–1463, Jul. 2019.

[46] L. F. Abbott, “Lapicque’s introduction of the integrate-and-fire model neuron (1907),” *Brain Res. Bull.*, vol. 50, no. 5–6, pp. 303–304, Nov. 1999.

[47] “https://wcs.webofknowledge.com.” 2020.

[48] J. Borghetti, G. S. Snider, P. J. Kuekes, J. J. Yang, D. R. Stewart, and R. S. Williams, “‘Memristive’ switches enable ‘stateful’ logic operations via material implication,” *Nature*, vol. 464, no. 7290, pp. 873–876, Apr. 2010.

[49] C. D. Wright, Y. Liu, K. I. Kohary, M. M. Aziz, and R. J. Hicken, “Arithmetic and Biologically-Inspired Computing Using Phase-Change Materials,” *Adv. Mater.*, vol. 23, no. 30, pp. 3408–3413, Aug. 2011.

[50] L. Xia *et al.*, “Technological Exploration of RRAM Crossbar Array for Matrix-Vector Multiplication,” *J. Comput. Sci. Technol.*, vol. 31, no. 1, pp. 3–19, Jan. 2016.

[51] M. Le Gallo *et al.*, “Mixed-precision in-memory computing,” *Nat. Electron.*, vol. 1, no. 4, pp. 246–253, Apr. 2018.

[52] G. W. Burr *et al.*, “Experimental Demonstration and Tolerancing of a Large-Scale Neural Network (165 000 Synapses) Using Phase-Change Memory as the Synaptic Weight Element,” *IEEE Trans. Electron Devices*, vol. 62, no. 11, pp. 3498–3507, Nov. 2015.

[53] I. Boybat *et al.*, “Neuromorphic computing with multi-memristive synapses,” *Nat. Commun.*, vol. 9, no. 1, p. 2514, Dec. 2018.

[54] A. N. Tait *et al.*, “Neuromorphic photonic networks using silicon photonic weight banks,” *Sci. Rep.*, vol. 7, no. 1, p. 7430, Dec. 2017.

[55] T. W. Hughes, M. Minkov, Y. Shi, and S. Fan, “Training of photonic neural networks through in situ backpropagation and gradient measurement,” *Optica*, vol. 5, no. 7, p. 864, Jul. 2018.

[56] I. Chakraborty, G. Saha, and K. Roy, “Photonic In-Memory Computing Primitive for Spiking Neural Networks Using Phase-Change Materials,” *Phys. Rev. Appl.*, vol. 11, no. 1, p. 014063, Jan. 2019.

[57] T. Deng, J. Robertson, and A. Hurtado, “Controlled Propagation of Spiking Dynamics in Vertical-Cavity Surface-Emitting Lasers: Towards Neuromorphic Photonic Networks,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 23, no. 6, pp. 1–8, Nov. 2017.

[58] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, “Parallel photonic information processing at gigabyte per second data rates using transient states,” *Nat. Commun.*, vol. 4, no. 1, p. 1364, Jun. 2013.

[59] T. Tanabe, M. Notomi, E. Kuramochi, A. Shinya, and H. Taniyama, “Trapping and delaying photons for one nanosecond in an ultrasmall high-Q photonic-crystal nanocavity,” *Nat. Photonics*, vol. 1, no. 1, pp. 49–52,

Jan. 2007.

[60] B. Romeira, R. Avó, J. M. L. Figueiredo, S. Barland, and J. Javaloyes, “Regenerative memory in time-delayed neuromorphic photonic resonators,” *Sci. Rep.*, vol. 6, no. 1, p. 19510, May 2016.

[61] L. Liu *et al.*, “An ultra-small, low-power, all-optical flip-flop memory on a silicon chip,” *Nat. Photonics*, vol. 4, no. 3, pp. 182–187, Mar. 2010.

[62] M. T. Hill *et al.*, “A fast low-power optical memory based on coupled micro-ring lasers,” *Nature*, vol. 432, no. 7014, pp. 206–209, Nov. 2004.

[63] M. Merklein, B. Stiller, K. Vu, S. J. Madden, and B. J. Eggleton, “A chip-integrated coherent photonic-phononic memory,” *Nat. Commun.*, vol. 8, no. 1, p. 574, Dec. 2017.

[64] Y. Nagasaki *et al.*, “Optical bistability in shape-memory nanowire metamaterial array,” *Appl. Phys. Lett.*, vol. 113, no. 2, p. 021105, Jul. 2018.

[65] C. Errando-Herranz, F. Niklaus, G. Stemme, and K. B. Gylfason, “Low-power microelectromechanically tunable silicon photonic ring resonator add-drop filter,” *Opt. Lett.*, vol. 40, no. 15, p. 3556, Aug. 2015.

[66] W. H. P. Pernice and H. Bhaskaran, “Photonic non-volatile memories using phase change materials,” *Appl. Phys. Lett.*, vol. 101, no. 17, p. 171101, Oct. 2012.

[67] C. Ríos *et al.*, “Integrated all-photonic non-volatile multi-level memory,” *Nat. Photonics*, vol. 9, no. 11, pp. 725–732, Nov. 2015.

[68] X. Li *et al.*, “Fast and reliable storage using a 5 bit, nonvolatile photonic memory cell,” *Optica*, vol. 6, no. 1, p. 1, Jan. 2019.

[69] Z. Cheng, C. Ríos, N. Youngblood, C. D. Wright, W. H. P. P. Pernice, and H. Bhaskaran, “Device-Level Photonic Memories and Logic Applications Using Phase-Change Materials,” *Adv. Mater.*, vol. 30, no. 32, p. 1802435, Aug. 2018.

[70] J. Feldmann *et al.*, “Calculating with light using a chip-scale all-optical abacus,” *Nat. Commun.*, vol. 8, no. 1, p. 1256, Dec. 2017.

[71] C. Ríos *et al.*, “In-memory computing on a photonic platform,” *Sci. Adv.*, vol. 5, no. 2, p. eaau5759, Feb. 2019.

[72] Z. Cheng, C. Ríos, W. H. P. P. Pernice, C. D. Wright, and H. Bhaskaran, “On-chip photonic synapse,” *Sci. Adv.*, vol. 3, no. 9, p. e1700160, Sep. 2017.

[73] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, “All-optical spiking neurosynaptic networks with self-learning capabilities,” *Nature*, vol. 569, no. 7755, pp. 208–214, May 2019.

[74] J. Feldmann *et al.*, “Parallel convolutional processing using an integrated photonic tensor core,” *Nature*, vol. 589, no. 7840, pp. 52–58, Jan. 2021.

[75] B. E. A. Saleh and M. C. Teich, *Fundamentals of photonics*. John Wiley & Sons, Inc., 2019.

[76] K. Okamoto, *Fundamentals of optical waveguides*. Academic press, 2006.

[77] L. N. Binh, *Guided Wave Photonics: Fundamentals and Applications with MATLAB*. CRC Press, 2012.

[78] E. A. J. Marcatili, “Dielectric Rectangular Waveguide and Directional Coupler for Integrated Optics,” *Bell Syst. Tech. J.*, vol. 48, no. 7, pp. 2071–2102, Sep. 1969.

[79] R. W. Boyd, *Nonlinear optics*. Academic press, 2019.

[80] C. Kittel, *Introduction to Solid State Physics*. Wiley, 2004.

[81] I. H. Malitson, “Interspecimen Comparison of the Refractive Index of Fused Silica,” *J. Opt. Soc. Am.*, vol. 55, no. 10, p. 1205, Oct. 1965.

[82] “Lorentz Dispersion Model,” 2006. [Online]. Available: https://www.horiba.com/fileadmin/uploads/Scientific/Downloads/OpticalSchool_CN/TN/ellipsometer/Lorentz_Dispersion_Model.pdf. [Accessed: 21-Jul-2020].

[83] P. B. Johnson and R.-Wjp. W. Christy, “Optical Constants of the Noble Metals,” *Phys. Rev. B*, vol. 6, no. 12, pp. 4370–4379, Dec. 1972.

[84] F. Hao and P. Nordlander, “Efficient dielectric function for FDTD simulation of the optical properties of silver and gold nanoparticles,” *Chem. Phys. Lett.*, vol. 446, no. 1–3, pp. 115–118, Sep. 2007.

[85] S. O. Abdellatif, K. Kirah, R. Ghannam, A. Khalil, and W. Anis, “Modelling Various Solar Cells Materials Using Lorentzian-Drude Coefficients,” *Int. J.*

Recent Contrib. from Eng. Sci. IT, vol. 4, no. 4, pp. 63–67, Dec. 2016.

[86] S. A. Maier, *Plasmonics: fundamentals and applications*, vol. 677, no. 1. Springer Science & Business Media, 2007.

[87] Y. A. Cengel and M. A. Boles, *Thermodynamics: An Engineering Approach 6th Editon (SI Units)*. The McGraw-Hill Companies, Inc., New York, 2007.

[88] F. J. Morin, "Oxides Which Show a Metal-to-Insulator Transition at the Neel Temperature," *Phys. Rev. Lett.*, vol. 3, no. 1, pp. 34–36, Jul. 1959.

[89] N. Shukla, T. Joshi, S. Dasgupta, P. Borisov, D. Lederman, and S. Datta, "Electrically induced insulator to metal transition in epitaxial SmNiO 3 thin films," *Appl. Phys. Lett.*, vol. 105, no. 1, p. 012108, Jul. 2014.

[90] J. W. Christian, *The theory of transformations in metals and alloys*. Newnes, 2002.

[91] R. E. Smallman, *Modern physical metallurgy*. Elsevier, 2016.

[92] D. Turnbull and J. C. Fisher, "Rate of Nucleation in Condensed Systems," *J. Chem. Phys.*, vol. 17, no. 1, pp. 71–73, Jan. 1949.

[93] J. Orava and A. L. Greer, "Classical-nucleation-theory analysis of priming in chalcogenide phase-change memory," *Acta Mater.*, vol. 139, pp. 226–235, Oct. 2017.

[94] S. Senkader and C. D. Wright, "Models for phase-change of Ge2Sb2Te5 in optical and electrical memory devices," *J. Appl. Phys.*, vol. 95, no. 2, pp. 504–511, Jan. 2004.

[95] A. Aladool, M. M. Aziz, and C. D. Wright, "Understanding the importance of the temperature dependence of viscosity on the crystallization dynamics in the Ge 2 Sb 2 Te 5 phase-change material," *J. Appl. Phys.*, vol. 121, no. 22, p. 224504, Jun. 2017.

[96] J. Orava, A. L. Greer, B. Gholipour, D. W. Hewak, and C. E. Smith, "Characterization of supercooled liquid Ge2Sb2Te5 and its crystallization by ultrafast-heating calorimetry," *Nat. Mater.*, vol. 11, no. 4, pp. 279–283, Apr. 2012.

[97] P. Lucas, "Fragile-to-strong transitions in glass forming liquids," *J. Non. Cryst. Solids*, p. 119367, Jul. 2019.

[98] P. G. Debenedetti and F. H. Stillinger, “Supercooled liquids and the glass transition,” *Nature*, vol. 410, no. 6825, pp. 259–267, Mar. 2001.

[99] J. Orava, H. Weber, I. Kaban, and A. L. Greer, “Viscosity of liquid Ag–In–Sb–Te: Evidence of a fragile-to-strong crossover,” *J. Chem. Phys.*, vol. 144, no. 19, p. 194503, May 2016.

[100] D. Q. Huang *et al.*, “Nonthermal phase transition in phase change memory cells induced by picosecond electric pulse,” *Appl. Phys. Lett.*, vol. 98, no. 24, p. 242106, Jun. 2011.

[101] K. Makino, J. Tominaga, and M. Hase, “Ultrafast optical manipulation of atomic arrangements in chalcogenide alloy memory materials,” *Opt. Express*, vol. 19, no. 2, p. 1260, Jan. 2011.

[102] P. Fons *et al.*, “Photoassisted amorphization of the phase-change memory alloy Ge2Sb2Te5,” *Phys. Rev. B*, vol. 82, no. 4, p. 041203, Jul. 2010.

[103] A. T. Waterman, “The Electrical Conductivity of Molybdenite,” *Phys. Rev.*, vol. 21, no. 5, pp. 540–549, May 1923.

[104] A. T. Waterman, “XXI. On the positive ionization from certain hot salts, together with some observations on the electrical properties of molybdenite at high temperatures,” *London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 33, no. 195, pp. 225–247, Mar. 1917.

[105] A. D. Pearson, W. R. Northover, J. F. Dewald, and W. F. Peck Jr, “Chemical physical and electrical properties of some unusual inorganic glasses,” *Adv. Glas. Technol.*, vol. 2, pp. 357–365, Nov. 1962.

[106] S. K. Bahl and K. L. Chopra, “Amorphous versus Crystalline GeTe Films. III. Electrical Properties and Band Structure,” *J. Appl. Phys.*, vol. 41, no. 5, pp. 2196–2212, Apr. 1970.

[107] L. Gildart, “Hypothesis of Orbital Overlap Shifting,” *J. Appl. Phys.*, vol. 36, no. 1, pp. 335–336, Jan. 1965.

[108] C. Ghosh and B. P. Varma, “Some optical properties of amorphous and crystalline antimony trisulphide thin films,” *Solid State Commun.*, vol. 31, no. 9, pp. 683–686, Sep. 1979.

[109] C. Ghosh and B. P. Varma, “Optical properties of amorphous and crystalline Sb2S3 thin films,” *Thin Solid Films*, vol. 60, no. 1, pp. 61–65,

Jun. 1979.

[110] C. M. Garner, L. R. Gilbert, and C. Wood, “The amorphous Sb1−xTex system,” *J. Non. Cryst. Solids*, vol. 15, no. 1, pp. 63–73, Apr. 1974.

[111] V. R. Katti, P. A. Govindacharyulu, and D. N. Bose, “Electrical and optical properties of amorphous semiconducting GeSe and GeSbSe films,” *Thin Solid Films*, vol. 14, no. 1, pp. 143–148, Dec. 1972.

[112] B. H. Tredgold, R. H. Williams, and A. Clark, “Switching properties and photoconduction in GaSe,” *Phys. status solidi*, vol. 3, no. 2, pp. 407–410, 1970.

[113] B. H. Tredgold, R. H. Williams, and A. Clark, “Switching properties and photoconduction in GaSe,” *Phys. Status Solidi*, vol. 3, no. 2, pp. 407–410, Oct. 1970.

[114] R. F. Shaw, W. Y. Liang, and A. D. Yoffe, “Optical properties, photoconductivity, and energy levels in crystalline and amorphous arsenic triselenide,” *J. Non. Cryst. Solids*, vol. 4, pp. 29–42, Apr. 1970.

[115] C. E. Wickersham, G. Bajor, and J. E. Greene, “Impulse stimulated ‘explosive’ crystallization of sputter deposited amorphous (In,Ga)Sb films,” *Solid State Commun.*, vol. 27, no. 1, pp. 17–20, Jul. 1978.

[116] J. Stuke and G. Zimmerer, “Optical properties of amorphous III–V compounds. I. Experiment,” *Phys. Status Solidi*, vol. 49, no. 2, pp. 513–523, Feb. 1972.

[117] M. Chen, K. A. Rubin, and R. W. Barton, “Compound materials for reversible, phase-change optical data storage,” *Appl. Phys. Lett.*, vol. 49, no. 9, pp. 502–504, Sep. 1986.

[118] M. Terao, T. Nishida, Y. Miyauchi, S. Horigome, T. Kaku, and N. Ohta, “In-Se Based Phase Change Reversible Optical Recording Film,” in *Optical Mass Data Storage II*, 1987, vol. 695, pp. 105–111.

[119] M. Wuttig and N. Yamada, “Phase-change materials for rewriteable data storage,” *Nat. Mater.*, vol. 6, no. 11, pp. 824–832, Nov. 2007.

[120] H. Seo, T.-H. Jeong, J.-W. Park, C. Yeon, S.-J. Kim, and S.-Y. Kim, “Investigation of Crystallization Behavior of Sputter-Deposited Nitrogen-Doped Amorphous Ge 2 Sb 2 Te 5 Thin Films,” *Jpn. J. Appl. Phys.*, vol. 39,

no. Part 1, No. 2B, pp. 745–751, Feb. 2000.

[121] S. Privitera, E. Rimini, and R. Zonca, “Amorphous-to-crystal transition of nitrogen- and oxygen-doped Ge2Sb2Te5 films studied by in situ resistance measurements,” *Appl. Phys. Lett.*, vol. 85, no. 15, pp. 3044–3046, Oct. 2004.

[122] S. Privitera, E. Rimini, C. Bongiorno, A. Pirovano, and R. Bez, “Effects of dopants on the amorphous-to-fcc transition in Ge2Sb2Te5 thin films,” *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms*, vol. 257, no. 1–2, pp. 352–354, Apr. 2007.

[123] T.-Y. Yang, J.-Y. Cho, Yong-Jin Park, and Y.-C. Joo, “Effects of dopings on the electric-field-induced atomic migration and void formation in Ge2Sb2Te5,” in *18th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 2011, pp. 1–4.

[124] N. Raeis-Hosseini, S. Lim, H. Hwang, and J. Rho, “Reliable Ge2Sb2Te5-Integrated High-Density Nanoscale Conductive Bridge Random Access Memory using Facile Nitrogen-Doping Strategy,” *Adv. Electron. Mater.*, vol. 4, no. 11, p. 1800360, Nov. 2018.

[125] T.-J. Park, S.-Y. Choi, and M.-J. Kang, “Phase transition characteristics of Bi/Sn doped Ge2Sb2Te5 thin film for PRAM application,” *Thin Solid Films*, vol. 515, no. 12, pp. 5049–5053, Apr. 2007.

[126] J. Feng *et al.*, “Si doping in Ge2Sb2Te5 film to reduce the writing current of phase change memory,” *Appl. Phys. A*, vol. 87, no. 1, pp. 57–62, Feb. 2007.

[127] Y.-J. Huang, Y.-C. Chen, and T.-E. Hsieh, “Phase transition behaviors of Mo- and nitrogen-doped Ge2Sb2Te5 thin films investigated by in situ electrical measurements,” *J. Appl. Phys.*, vol. 106, no. 3, p. 034916, Aug. 2009.

[128] S. J. Wei *et al.*, “Phase change behavior in titanium-doped Ge2Sb2Te5 films,” *Appl. Phys. Lett.*, vol. 98, no. 23, p. 231910, Jun. 2011.

[129] X. Zhou *et al.*, “Carbon-doped Ge2Sb2Te5 phase change material: A candidate for high-density phase change memory application,” *Appl. Phys. Lett.*, vol. 101, no. 14, p. 142104, Oct. 2012.

[130] G. Wang *et al.*, “Improved thermal and electrical properties of Al-doped Ge2Sb2Te5 films for phase-change random access memory,” *J. Phys. D. Appl. Phys.*, vol. 45, no. 37, p. 375302, Sep. 2012.

[131] G. Wang *et al.*, “Phase change behaviors of Zn-doped Ge2Sb2Te5 films,” *Appl. Phys. Lett.*, vol. 101, no. 5, p. 051906, Jul. 2012.

[132] S. Kumar, D. Singh, and R. Thangaraj, “Electronic structure and optical band gap of silver photo-diffused Ge2Sb2Te5 thin film,” *Appl. Surf. Sci.*, vol. 273, pp. 437–443, May 2013.

[133] S. Guo *et al.*, “Temperature and concentration dependent crystallization behavior of Ge2Sb2Te5 phase change films: tungsten doping effects,” *RSC Adv.*, vol. 4, no. 100, pp. 57218–57222, Oct. 2014.

[134] G. Singh, A. Kaura, M. Mukul, J. Singh, and S. K. Tripathi, “Theoretical and experimental investigations of the properties of Ge2Sb2Te5 and indium-doped Ge2Sb2Te5 phase change material,” *Appl. Phys. A*, vol. 117, no. 3, pp. 1307–1314, Nov. 2014.

[135] K. Ding *et al.*, “Study on the Cu-doped Ge2Sb2Te5 for low-power phase change memory,” *Mater. Lett.*, vol. 125, pp. 143–146, Jun. 2014.

[136] Q. Wang *et al.*, “Cr-doped Ge2Sb2Te5 for ultra-long data retention phase change memory,” *Appl. Phys. Lett.*, vol. 107, no. 22, p. 222101, Nov. 2015.

[137] E. M. Vinod, K. Ramesh, and K. S. Sangunni, “Structural transition and enhanced phase transition properties of Se doped Ge2Sb2Te5 alloys,” *Sci. Rep.*, vol. 5, no. 1, p. 8050, Jul. 2015.

[138] Y. Wang *et al.*, “Scandium doped Ge2Sb2Te5 for high-speed and low-power-consumption phase change memory,” *Appl. Phys. Lett.*, vol. 112, no. 13, p. 133104, Mar. 2018.

[139] V. E. Madhavan, M. Carignano, A. Kachmar, and K. S. Sangunni, “Crystallization properties of arsenic doped GST alloys,” *Sci. Rep.*, vol. 9, no. 1, p. 12985, Dec. 2019.

[140] Y. Zhang *et al.*, “Broadband Transparent Optical Phase Change Materials,” in *Conference on Lasers and Electro-Optics*, 2017, no. c, p. JTh5C.4.

[141] Z. L. Sámson *et al.*, “Metamaterial electro-optic switch of nanoscale thickness,” *Appl. Phys. Lett.*, vol. 96, no. 14, p. 143105, Apr. 2010.

[142] W. Dong *et al.*, “Wide Bandgap Phase Change Material Tuned Visible Photonics,” *Adv. Funct. Mater.*, vol. 29, no. 6, p. 1806181, Feb. 2019.

[143] M. Delaney, I. Zeimpekis, D. Lawson, D. W. Hewak, and O. L. Muskens, “A New Family of Ultralow Loss Reversible Phase-Change Materials for Photonic Integrated Circuits: Sb 2 S 3 and Sb 2 Se 3,” *Adv. Funct. Mater.*, vol. 30, no. 36, p. 2002447, Sep. 2020.

[144] P. Guo, A. M. Sarangan, and I. Agha, “A review of germanium-antimony-telluride phase change materials for non-volatile memories and optical modulators,” *Appl. Sci.*, vol. 9, no. 3, p. 530, Feb. 2019.

[145] F. Dirisaglik *et al.*, “High speed, high temperature electrical characterization of phase change materials: metastable phases, crystallization dynamics, and resistance drift,” *Nanoscale*, vol. 7, no. 40, pp. 16625–16630, Sep. 2015.

[146] K.-K. Du *et al.*, “Control over emissivity of zero-static-power thermal emitters based on phase-changing material GST,” *Light Sci. Appl.*, vol. 6, no. 1, p. e16194, Jan. 2017.

[147] P. Singh, A. P. Singh, N. Kanda, M. Mishra, G. Gupta, and A. Thakur, “High transmittance contrast in amorphous to hexagonal phase of Ge2Sb2Te5: Reversible NIR-window,” *Appl. Phys. Lett.*, vol. 111, no. 26, p. 261102, 2017.

[148] J. Xu, C. Qi, L. Chen, L. Zheng, and Q. Xie, “The microstructural changes of Ge2Sb2Te5 thin film during crystallization process,” *AIP Adv.*, vol. 8, no. 5, p. 055006, May 2018.

[149] T. H. Lee and S. R. Elliott, “Ab initio computer simulation of the early stages of crystallization: Application to Ge2Sb2Te5 phase-change materials,” *Phys. Rev. Lett.*, vol. 107, no. 14, p. 145702, Sep. 2011.

[150] D. Lencer, “Design Rules, Local Structure and Lattice-Dynamics of Phase-Change Materials for Data Storage Applications,” RWTH Aachen University, 2010.

[151] H.-K. Lyeo *et al.*, “Thermal conductivity of phase-change material Ge2Sb2Te5,” *Appl. Phys. Lett.*, vol. 89, no. 15, p. 151904, Oct. 2006.

[152] M. Xu *et al.*, “Theoretical and experimental investigations of the optical

properties of Ge2Sb2Te5 for multi-state optical data storage," *J. Korean Phys. Soc.*, vol. 53, no. 4, pp. 2265–2269, Oct. 2008.

[153] B. Gholipour, J. Zhang, K. F. MacDonald, D. W. Hewak, and N. I. Zheludev, "An All-Optical, Non-volatile, Bidirectional, Phase-Change Meta-Switch," *Adv. Mater.*, vol. 25, no. 22, pp. 3050–3054, Jun. 2013.

[154] J.-W. Park *et al.*, "Optical properties of pseudobinary GeTe, Ge2Sb2Te5, GeSb2Te4, GeSb4Te7, and Sb2Te3 from ellipsometry and density functional theory," *Phys. Rev. B*, vol. 80, no. 11, p. 115209, Sep. 2009.

[155] K. Shportko, S. Kremers, M. Woda, D. Lencer, J. Robertson, and M. Wuttig, "Resonant bonding in crystalline phase-change materials," *Nat. Mater.*, vol. 7, no. 8, pp. 653–658, Aug. 2008.

[156] J. C. M. Garnett, "XII. Colours in metal glasses and in metallic films," *Philos. Trans. R. Soc. London. Ser. A, Contain. Pap. a Math. or Phys. Character*, vol. 203, no. 359–371, pp. 385–420, Jan. 1904.

[157] D. A. G. Bruggeman, "Berechnung verschiedener physikalischer Konstanten von heterogenen Substanzen. I. Dielektrizitätskonstanten und Leitfähigkeiten der Mischkörper aus isotropen Substanzen," *Ann. Phys.*, vol. 416, no. 7, pp. 636–664, 1935.

[158] T. C. Choy, *Effective medium theory: principles and applications*, vol. 165. Oxford University Press, 2015.

[159] H. Hayat, K. Kohary, and C. D. Wright, "Can conventional phase-change memory devices be scaled down to single-nanometre dimensions?," *Nanotechnology*, vol. 28, no. 3, p. 035202, Jan. 2017.

[160] K. Cil *et al.*, "Electrical Resistivity of Liquid Ge2Sb2Te5 Based on Thin-Film and Nanoscale Device Measurements," *IEEE Trans. Electron Devices*, vol. 60, no. 1, pp. 433–437, Jan. 2013.

[161] L. Crespi, A. Ghetti, M. Boniardi, and A. L. Lacaita, "Electrical Conductivity Discontinuity at Melt in Phase Change Memory," *IEEE Electron Device Lett.*, vol. 35, no. 7, pp. 747–749, Jul. 2014.

[162] A. Faraclas *et al.*, "Modeling of Thermoelectric Effects in Phase Change Memory Cells," *IEEE Trans. Electron Devices*, vol. 61, no. 2, pp. 372–378, Feb. 2014.

[163] M. Stegmaier *et al.*, “Thermo-optical Effect in Phase-Change Nanophotonics,” *ACS Photonics*, vol. 3, no. 5, pp. 828–835, May 2016.

[164] E. M. Vinod, R. Naik, A. P. A. Faiyas, R. Ganesan, and K. S. Sangunni, “Temperature dependent optical constants of amorphous Ge2Sb2Te5 thin films,” *J. Non. Cryst. Solids*, vol. 356, no. 41–42, pp. 2172–2174, Sep. 2010.

[165] M. Rütten, M. Kaes, A. Albert, M. Wuttig, and M. Salinga, “Relation between bandgap and resistance drift in amorphous phase change materials,” *Sci. Rep.*, vol. 5, no. 1, p. 17362, Dec. 2015.

[166] S. Liu, J. Wei, and F. Gan, “Optical nonlinear absorption characteristics of crystalline Ge 2 Sb 2 Te 5 thin films,” *J. Appl. Phys.*, vol. 110, no. 3, p. 033503, Aug. 2011.

[167] C. D. Wright, M. Armand, and M. M. Aziz, “Terabit-per-square-inch data storage using phase-change media and scanning electrical nanoprobes,” *IEEE Trans. Nanotechnol.*, vol. 5, no. 1, pp. 50–61, Jan. 2006.

[168] M. Le Gallo, A. Athmanathan, D. Krebs, and A. Sebastian, “Evidence for thermally assisted threshold switching behavior in nanoscale phase-change memory cells,” *J. Appl. Phys.*, vol. 119, no. 2, p. 025704, Jan. 2016.

[169] K. Tanaka, “Notes on Reproducibility of Ge2Sb2Te5 Properties,” *Phys. status solidi*, vol. 257, no. 6, p. 1900756, Jun. 2020.

[170] D. Ielmini and A. L. Lacaita, “Phase change materials in non-volatile storage,” *Mater. Today*, vol. 14, no. 12, pp. 600–607, Dec. 2011.

[171] T. Alexoudi, G. T. Kanellos, and N. Pleros, “Optical RAM and integrated optical memories: a survey,” *Light Sci. Appl.*, vol. 9, no. 1, pp. 1–16, 2020.

[172] C. Rios *et al.*, “Controlled switching of phase-change materials by evanescent-field coupling in integrated photonics [Invited],” *Opt. Mater. Express*, vol. 8, no. 9, p. 2455, Sep. 2018.

[173] J. Feldmann, N. Youngblood, X. Li, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, “Integrated 256 Cell Photonic Phase-Change Memory With 512-Bit Capacity,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 26, no. 2, pp. 1–7, Mar. 2020.

[174] J. von Keitz *et al.*, “Reconfigurable Nanophotonic Cavities with Nonvolatile

Response,” *ACS Photonics*, vol. 5, no. 11, pp. 4644–4649, Nov. 2018.

[175] C. Wu, H. Yu, H. Li, X. Zhang, I. Takeuchi, and M. Li, “Low-Loss Integrated Photonic Switch Using Subwavelength Patterned Phase Change Material,” *ACS Photonics*, vol. 6, no. 1, pp. 87–92, Jan. 2019.

[176] N. Farmakidis *et al.*, “Plasmonic nanogap enhanced phase-change devices with dual electrical-optical functionality,” *Sci. Adv.*, vol. 5, no. 11, p. eaaw2687, Nov. 2019.

[177] G. T. Reed and A. P. Knights, *Silicon Photonics*. Chichester, UK: John Wiley & Sons, Ltd, 2004.

[178] J. C. Lindon, G. E. Tranter, and D. W. Koppenaal, *Encyclopedia of Spectroscopy and Spectrometry*. Academic Press, 2016.

[179] R. M. A. Azzam and N. M. Bashara, *Ellipsometry and polarized light*. Amsterdam: North-Holland, 1977.

[180] D. W. Grainger and D. G. Castner, “Surface analysis and biointerfaces: Vacuum and ambient in situ techniques,” in *Comprehensive Biomaterials*, Elsevier, 2011.

[181] M. Tilli, T. Motooka, V. M. Airaksinen, S. Franssila, M. Paulasto-Kröckel, and V. Lindroos, *Handbook of Silicon Based MEMS Materials and Technologies: Second Edition*. Elsevier, 2015.

[182] J. G. Solé, L. E. Bausá, and D. Jaque, *An Introduction to the Optical Spectroscopy of Inorganic Solids*. Chichester, UK: John Wiley & Sons, Ltd, 2005.

[183] R. Swanepoel, “Determination of the thickness and optical constants of amorphous silicon,” *J. Phys. E.*, vol. 16, no. 12, pp. 1214–1222, Dec. 1983.

[184] R. Swanepoel, “Determining refractive index and thickness of thin films from wavelength measurements only,” *J. Opt. Soc. Am. A*, vol. 2, no. 8, p. 1339, Aug. 1985.

[185] O. C. Zienkiewicz, R. L. Taylor, and J. Z. Zhu, *The finite element method: its basis and fundamentals*. Elsevier, 2005.

[186] B. Gallinet, J. Butet, and O. J. F. Martin, “Numerical methods for nanophotonics: standard problems and future challenges,” *Laser Photon. Rev.*, vol. 9, no. 6, pp. 577–603, Nov. 2015.

[187] W. Frei, “Using Perfectly Matched Layers and Scattering Boundary Conditions for Wave Electromagnetics Problems,” *COMSOL Blog*, 2015. [Online]. Available: https://uk.comsol.com/blogs/using-perfectly-matched-layers-and-scattering-boundary-conditions-for-wave-electromagnetics-problems/. [Accessed: 05-Sep-2018].

[188] C. E. Shannon, “Communication in the presence of noise,” *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[189] D. J. Griffiths, *Introduction to Electrodynamics*. Cambridge University Press, 2017.

[190] R. Courant, K. Friedrichs, and H. Lewy, “On the partial difference equations of mathematical physics,” *IBM J. Res. Dev.*, vol. 11, no. 2, pp. 215–234, Mar. 1967.

[191] F. Denner and B. G. M. M. van Wachem, “Compressive VOF method with skewness correction to capture sharp interfaces on arbitrary meshes,” *J. Comput. Phys.*, vol. 279, pp. 127–144, Sep. 2014.

[192] W. Wełnic, S. Botti, L. Reining, and M. Wuttig, “Origin of the Optical Contrast in Phase-Change Materials,” *Phys. Rev. Lett.*, vol. 98, no. 23, p. 236403, Jun. 2007.

[193] C. Pauly *et al.*, “Evidence for topological band inversion of the phase change material Ge2Sb2Te5,” *Appl. Phys. Lett.*, vol. 103, no. 24, p. 243109, Dec. 2013.

[194] J. Akola and R. O. Jones, “Density functional study of amorphous, liquid and crystalline Ge2Sb2Te5 : homopolar bonds and/or AB alternation?,” *J. Phys. Condens. Matter*, vol. 20, no. 46, p. 465103, Nov. 2008.

[195] P. Jóvári, I. Kaban, J. Steiner, B. Beuneu, A. Schöps, and A. Webb, “‘Wrong bonds’ in sputtered amorphous Ge 2 Sb 2 Te 5,” *J. Phys. Condens. Matter*, vol. 19, no. 33, p. 335212, Aug. 2007.

[196] S. Caravati, M. Bernasconi, T. D. Kühne, M. Krack, and M. Parrinello, “First principles study of crystalline and amorphous Ge2Sb2Te5 and the effects of stoichiometric defects,” *J. Phys. Condens. Matter*, vol. 22, no. 39, pp. 399801–399801, Oct. 2010.

[197] A. N. Kolmogorov, “On the statistical theory of the crystallization of metals,”

Bull. Acad. Sci. USSR, Math. Ser, vol. 1, no. 3, pp. 355–359, Feb. 1937.

[198] W. A. Johnson and R. F. Mehl, "Reaction kinetics in processes of nucleation and growth," *Trans. Am. Inst. Min. Met. Eng.*, vol. 135, pp. 416–458, Feb. 1939.

[199] M. Avrami, "Kinetics of phase change. I General theory," *J. Chem. Phys.*, vol. 7, no. 12, pp. 1103–1112, Dec. 1939.

[200] M. Avrami, "Kinetics of phase change. II transformation-time relations for random distribution of nuclei," *J. Chem. Phys.*, vol. 8, no. 2, pp. 212–224, Feb. 1940.

[201] M. Avrami, "Granulation, phase change, and microstructure kinetics of phase change. III," *J. Chem. Phys.*, vol. 9, no. 2, pp. 177–184, Feb. 1941.

[202] J. Málek, "The applicability of Johnson-Mehl-Avrami model in the thermal analysis of the crystallization kinetics of glasses," *Thermochim. Acta*, vol. 267, pp. 61–73, Dec. 1995.

[203] M. Fanfoni and M. Tomellini, "The Johnson-Mehl-Avrami-Kolmogorov model: A brief review," *Nuovo Cim. D*, vol. 20, no. 7–8, pp. 1171–1182, Jul. 1998.

[204] K. Sekimoto, "Evolution of the domain structure during the nucleation-and-growth process with non-conserved order parameter," *Phys. A Stat. Mech. its Appl.*, vol. 135, no. 2–3, pp. 328–346, Apr. 1986.

[205] M. C. Weinberg, D. P. Birnie III, and V. A. Shneidman, "Crystallization kinetics and the JMAK equation," *J. Non. Cryst. Solids*, vol. 219, pp. 89–99, Oct. 1997.

[206] J. M. Rickman, W. S. Tong, and K. Barmak, "Impact of heterogeneous boundary nucleation on transformation kinetics and microstructure," *Acta Mater.*, vol. 45, no. 3, pp. 1153–1166, Mar. 1997.

[207] J. Farjas and P. Roura, "Modification of the Kolmogorov–Johnson–Mehl–Avrami rate equation for non-isothermal experiments and its analytical solution," *Acta Mater.*, vol. 54, no. 20, pp. 5573–5579, Dec. 2006.

[208] V. Weidenhof, I. Friedrich, S. Ziegler, and M. Wuttig, "Laser induced crystallization of amorphous Ge2Sb2Te5 films," *J. Appl. Phys.*, vol. 89, no. 6, pp. 3168–3176, Mar. 2001.

[209] M. Volmer and A. Weber, “Keimbildung in übersättigten Gebilden,” *Zeitschrift für Phys. Chemie*, vol. 119, no. 1, pp. 277–301, Jan. 1926.

[210] R. Becker and W. Döring, “Kinetische behandlung der keimbildung in übersättigten dämpfen,” *Ann. Phys.*, vol. 416, no. 8, pp. 719–752, 1935.

[211] D. Turnbull, “Under what conditions can a glass be formed?,” *Contemp. Phys.*, vol. 10, no. 5, pp. 473–488, Sep. 1969.

[212] H. B. Singh and A. Holz, “Stability limit of supercooled liquids,” *Solid State Commun.*, vol. 45, no. 11, pp. 985–988, Mar. 1983.

[213] C. Peng, L. Cheng, and M. Mansuripur, “Experimental and theoretical investigations of laser-induced crystallization and amorphization in phase-change optical recording media,” *J. Appl. Phys.*, vol. 82, no. 9, pp. 4183–4191, Nov. 1997.

[214] C. D. Wright, K. Blyuss, and P. Ashwin, “Master-equation approach to understanding multistate phase-change memories and processors,” *Appl. Phys. Lett.*, vol. 90, no. 6, p. 063113, Feb. 2007.

[215] C. D. Wright *et al.*, “The Design of Rewritable Ultrahigh Density Scanning-Probe Phase-Change Memories,” *IEEE Trans. Nanotechnol.*, vol. 10, no. 4, pp. 900–912, Jul. 2011.

[216] U. Russo, D. Ielmini, and A. L. Lacaita, “Analytical Modeling of Chalcogenide Crystallization for PCM Data-Retention Extrapolation,” *IEEE Trans. Electron Devices*, vol. 54, no. 10, pp. 2769–2777, Oct. 2007.

[217] K. B. Blyuss, P. Ashwin, A. P. Bassom, and C. D. Wright, “Master-equation approach to the study of phase-change processes in data storage media,” *Phys. Rev. E*, vol. 72, no. 1, p. 011607, Jul. 2005.

[218] A. Aladool, M. Aziz, and D. Wright, “Understanding the Influence of Initial Cluster Size Distribution On Crystallization Dynamics in The Ge 2 Sb 2 Te 5 Phase-Change Alloy,” *Phys. status solidi*, vol. 256, no. 7, p. 1800583, Jul. 2019.

[219] K. Kohary and C. D. Wright, “Modelling the phase-transition in phase-change materials,” *Phys. status solidi*, vol. 250, no. 5, pp. 944–948, May 2013.

[220] P. Ashwin, B. S. V Patnaik, and C. D. Wright, “Fast simulation of phase-

change processes in chalcogenide alloys using a Gillespie-type cellular automata approach," *J. Appl. Phys.*, vol. 104, no. 8, p. 084901, Oct. 2008.

[221] I. Steinbach *et al.*, "A phase field concept for multiphase systems," *Phys. D Nonlinear Phenom.*, vol. 94, no. 3, pp. 135–147, Jul. 1996.

[222] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *J. Comput. Phys.*, vol. 22, no. 4, pp. 403–434, Dec. 1976.

[223] J. A. Vázquez Diosdado, P. Ashwin, K. I. Kohary, and C. D. Wright, "Threshold switching via electric field induced crystallization in phase-change memory devices," *Appl. Phys. Lett.*, vol. 100, no. 25, p. 253105, Jun. 2012.

[224] C. D. Wright, P. Hosseini, and J. A. V. Diosdado, "Beyond von-Neumann computing with nanoscale phase-change memory devices," *Adv. Funct. Mater.*, vol. 23, no. 18, pp. 2248–2254, May 2013.

[225] Z. Fan and D. E. Laughlin, "Three Dimensional Crystallization Simulation and Recording Layer Thickness Effect in Phase Change Optical Recording," *Jpn. J. Appl. Phys.*, vol. 42, no. Part 1, No. 2B, pp. 800–803, Feb. 2003.

[226] H. Akima, "A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures," *J. ACM*, vol. 17, no. 4, pp. 589–602, Oct. 1970.

[227] S. G. C. Carrillo *et al.*, "Behavioral modeling of integrated phase-change photonic devices for neuromorphic computing applications," *APL Mater.*, vol. 7, no. 9, p. 091113, Sep. 2019.

[228] P. Dong *et al.*, "Low power and compact reconfigurable multiplexing devices based on silicon microring resonators," *Opt. Express*, vol. 18, no. 10, p. 9852, May 2010.

[229] M. Amin, M. Farhat, and H. Bağcı, "A nonlinear plasmonic resonator for three-state all-optical switching," *Opt. Express*, vol. 22, no. 6, p. 6966, Mar. 2014.

[230] H. Dieker and M. Wuttig, "Influence of deposition parameters on the properties of sputtered Ge2Sb2Te5 films," *Thin Solid Films*, vol. 478, no. 1–2, pp. 248–251, May 2005.

[231] T. Welzel and K. Ellmer, “The influence of the target age on laterally resolved ion distributions in reactive planar magnetron sputtering,” *Surf. Coatings Technol.*, vol. 205, pp. S294–S298, Jul. 2011.

[232] J. Aarik, A. Aidla, A.-A. Kiisler, T. Uustare, and V. Sammelselg, “Influence of substrate temperature on atomic layer growth and properties of HfO2 thin films,” *Thin Solid Films*, vol. 340, no. 1–2, pp. 110–116, Feb. 1999.

[233] W. Rieger, T. Metzger, H. Angerer, R. Dimitrov, O. Ambacher, and M. Stutzmann, “Influence of substrate-induced biaxial compressive stress on the optical properties of thin GaN films,” *Appl. Phys. Lett.*, vol. 68, no. 7, pp. 970–972, Feb. 1996.

[234] M. Bender, W. Seelig, C. Daube, H. Frankenberger, B. Ocker, and J. Stollenwerk, “Dependence of film composition and thicknesses on optical and electrical properties of ITO–metal–ITO multilayers,” *Thin Solid Films*, vol. 326, no. 1–2, pp. 67–71, Aug. 1998.

[235] C. R. de Galarreta *et al.*, “Nonvolatile Reconfigurable Phase-Change Metadevices for Beam Steering in the Near Infrared,” *Adv. Funct. Mater.*, vol. 28, no. 10, p. 1704993, Mar. 2018.

[236] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[237] K. E. Oughstun and C. L. Palombini, “Fresnel reflection and transmission coefficients for complex media,” in *2017 XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, 2017, pp. 1–2.

[238] S. Colburn, A. Zhan, and A. Majumdar, “Tunable metasurfaces via subwavelength phase shifters with uniform amplitude,” *Sci. Rep.*, vol. 7, no. 1, p. 40174, Mar. 2017.

[239] P. Hosseini, C. D. Wright, and H. Bhaskaran, “An optoelectronic framework enabled by low-dimensional phase-change films,” *Nature*, vol. 511, no. 7508, pp. 206–211, Jul. 2014.

[240] R. C. Rumpf, “Improved formulation of scattering matrices for semi-analytical methods that is consistent with convention,” *Prog. Electromagn. Res. B*, vol. 35, pp. 241–261, Oct. 2011.

[241] V. Torczon, “On the Convergence of Pattern Search Algorithms,” *SIAM J. Optim.*, vol. 7, no. 1, pp. 1–25, Feb. 1997.

[242] T. P. L. Pedersen, J. Kalb, W. K. Njoroge, D. Wamwangi, M. Wuttig, and F. Spaepen, “Mechanical stresses upon crystallization in phase change materials,” *Appl. Phys. Lett.*, vol. 79, no. 22, pp. 3597–3599, Nov. 2001.

[243] A. D. Rakić, “Algorithm for the determination of intrinsic optical constants of metal films: application to aluminum,” *Appl. Opt.*, vol. 34, no. 22, p. 4755, Aug. 1995.

[244] J.-W. Park *et al.*, “Optical properties of (GeTe, Sb2Te3) pseudobinary thin films studied with spectroscopic ellipsometry,” *Appl. Phys. Lett.*, vol. 93, no. 2, p. 021914, Jul. 2008.

[245] Z. Xu, C. Chen, Z. Wang, K. Wu, H. Chong, and H. Ye, “Optical constants acquisition and phase change properties of Ge2Sb2Te5 thin films based on spectroscopy,” *RSC Adv.*, vol. 8, no. 37, pp. 21040–21046, Jun. 2018.

[246] M. Jafari and M. Rais-Zadeh, “Zero-static-power phase-change optical modulator,” *Opt. Lett.*, vol. 41, no. 6, p. 1177, Mar. 2016.

[247] J. Robertson and B. Huang, “Bonding and optical contrast in phase change memory materials,” *Phys. status solidi*, vol. 249, no. 10, pp. 1867–1873, Oct. 2012.

[248] F. Peinado and I. S. Lerma, “Influence of the annealing on the optical properties of amorphous GeTe films,” *Rev. Phys. Appliquée*, vol. 12, no. 5, pp. 763–765, May 1977.

[249] C. Ríos *et al.*, “In-memory computing on a photonic platform,” *Sci. Adv.*, vol. 5, no. 2, p. eaau5759, Feb. 2019.

[250] Z. Cheng *et al.*, “On-chip photonic synapse,” *Sci. Adv.*, vol. 3, no. 9, p. e1700160, Sep. 2017.

[251] C. Rios, P. Hosseini, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, “On-Chip Photonic Memory Elements Employing Phase-Change Materials,” *Adv. Mater.*, vol. 26, no. 9, pp. 1372–1377, Mar. 2014.

[252] E. Gemo *et al.*, “Plasmonically-enhanced all-optical integrated phase-change memory,” *Opt. Express*, vol. 27, no. 17, p. 24724, Aug. 2019.

[253] X. Li *et al.*, “Experimental investigation of silicon and silicon nitride

platforms for phase-change photonic in-memory computing," *Optica*, vol. 7, no. 3, p. 218, Mar. 2020.

[254] N. Youngblood *et al.*, "Tunable Volatility of Ge2Sb2Te5 in Integrated Photonics," *Adv. Funct. Mater.*, vol. 29, no. 11, p. 1807571, Mar. 2019.

[255] E. Gemo *et al.*, "A plasmonically-enhanced route to faster and more energy-efficient phase-change integrated photonic memory and computing devices," *J. Appl. Phys.*, 2021.

[256] J. Kischkat *et al.*, "Mid-infrared optical properties of thin films of aluminum oxide, titanium dioxide, silicon dioxide, aluminum nitride, and silicon nitride," *Appl. Opt.*, vol. 51, no. 28, p. 6789, Oct. 2012.

[257] P. Dong *et al.*, "Novel integration technique for silicon/III-V hybrid laser," *Opt. Express*, vol. 22, no. 22, p. 26854, Nov. 2014.

[258] S.-Y. Kim, S. J. Kim, H. Seo, and M. R. Kim, "Variation of the complex refractive indices with Sb-addition in Ge-Sb-Te alloy and their wavelength dependence," in *Optical Data Storage '98*, 1998, pp. 112–115.

[259] W. K. Njoroge, H.-W. Wöltgens, and M. Wuttig, "Density changes upon crystallization of Ge2Sb2.04Te4.74 films," *J. Vac. Sci. Technol. A Vacuum, Surfaces, Film.*, vol. 20, no. 1, pp. 230–233, Jan. 2002.

[260] A. Arbabi and L. L. Goddard, "Measurements of the refractive indices and thermo-optic coefficients of Si3N4 and SiOx using microring resonances," *Opt. Lett.*, vol. 38, no. 19, p. 3878, Oct. 2013.

[261] J. P. Reifenberg *et al.*, "Thickness and stoichiometry dependence of the thermal conductivity of GeSbTe films," *Appl. Phys. Lett.*, vol. 91, no. 11, p. 111904, Sep. 2007.

[262] S. G. Carrillo *et al.*, "Design of practicable phase-change metadevices for near-infrared absorber and modulator applications," *Opt. Express*, vol. 24, no. 12, p. 13563, Jun. 2016.

[263] E. Yalon *et al.*, "Spatially Resolved Thermometry of Resistive Memory Devices," *Sci. Rep.*, vol. 7, no. 1, p. 15360, Dec. 2017.

[264] E. Bozorg-Grayeli, J. P. Reifenberg, M. A. Panzer, J. A. Rowlette, and K. E. Goodson, "Temperature-Dependent Thermal Properties of Phase-Change Memory Electrode Materials," *IEEE Electron Device Lett.*, vol. 32,

no. 9, pp. 1281–1283, Sep. 2011.

[265] J. Cho, K. K. Chu, P. C. Chao, C. McGray, M. Asheghi, and K. E. Goodson, “Thermal conduction normal to thin silicon nitride films on diamond and GaN,” *Thermomechanical Phenom. Electron. Syst. -Proceedings Intersoc. Conf.*, pp. 1186–1191, 2014.

[266] K. Hirao, K. Watari, M. E. Brito, M. Toriyama, and S. Kanzaki, “High Thermal Conductivity in Silicon Nitride with Anisotropie Microstructure,” *J. Am. Ceram. Soc.*, vol. 79, no. 9, pp. 2485–2488, Sep. 1996.

[267] G. Ziegler and D. P. H. Hasselman, “Effect of phase composition and microstructure on the thermal diffusivity of silicon nitride,” *J. Mater. Sci.*, vol. 16, no. 2, pp. 495–503, Feb. 1981.

[268] A. Karvounis *et al.*, “Mechanochromic Reconfigurable Metasurfaces,” *Adv. Sci.*, vol. 6, no. 21, p. 1900974, Nov. 2019.

[269] V. Palankovski and R. Quay, *Analysis and Simulation of Heterostructure Devices*. Vienna: Springer Vienna, 2004.

[270] A. Jain and K. E. Goodson, “Measurement of the Thermal Conductivity and Heat Capacity of Freestanding Shape Memory Thin Films Using the 3ω Method,” *J. Heat Transfer*, vol. 130, no. 10, Oct. 2008.

[271] P. Miranzo, E. García, C. Ramírez, J. González-Julián, M. Belmonte, and M. Isabel Osendi, “Anisotropic thermal conductivity of silicon nitride ceramics containing carbon nanostructures,” *J. Eur. Ceram. Soc.*, vol. 32, no. 8, pp. 1847–1854, Jul. 2012.

[272] M. B. Takeyama, M. Sato, Y. Nakata, Y. Kobayashi, T. Nakamura, and A. Noya, “Characterization of silicon nitride thin films deposited by reactive sputtering and plasma- enhanced CVD at low temperatures,” *Jpn. J. Appl. Phys.*, vol. 53, no. 5S2, p. 05GE01, May 2014.

[273] N. Youngblood *et al.*, “Tunable Volatility of Ge <inf>2</inf> Sb <inf>2</inf> Te <inf>5</inf> in Integrated Photonics,” *Adv. Funct. Mater.*, vol. 29, no. 11, 2019.

[274] N. Ciocchini, M. Cassinerio, D. Fugazza, and D. Ielmini, “Evidence for Non-Arrhenius Kinetics of Crystallization in Phase Change Memory Devices,” *IEEE Trans. Electron Devices*, vol. 60, no. 11, pp. 3767–3774, Nov. 2013.

[275] S. N. Yannopoulos, “Athermal photoelectronic effects in non-crystalline chalcogenides: Current status and beyond,” *arXiv Prepr. arXiv2001.04877*, Jan. 2020.

[276] M. Stegmaier, C. Ríos, H. Bhaskaran, C. D. Wright, and W. H. P. Pernice, “Nonvolatile All-Optical 1 × 2 Switch for Chipscale Photonic Networks,” *Adv. Opt. Mater.*, vol. 5, no. 1, p. 1600346, Jan. 2017.

[277] M. Rudé *et al.*, “Optical switching at 1.55 μm in silicon racetrack resonators using phase change materials,” *Appl. Phys. Lett.*, vol. 103, no. 14, p. 141119, Sep. 2013.

[278] R. Rajasekar, G. T. Raja, and S. Robinson, “Numerical Investigation of Reconfigurable Photonic Crystal Switch Based on Phase Change Nanomaterial,” *IEEE Trans. Nanotechnol.*, vol. 19, pp. 545–552, Jun. 2020.

[279] R. R. Ghosh, P. Bhardwaj, and A. Dhawan, “Numerical modeling of integrated electro-optic modulators based on mode-gap shifting in photonic crystal slab waveguides containing a phase change material,” *J. Opt. Soc. Am. B*, vol. 37, no. 8, p. 2287, Aug. 2020.

[280] B. Dong, Y. Ma, Z. Ren, and C. Lee, “Recent progress in nanoplasmonics-based integrated optical micro/nano-systems,” *J. Phys. D. Appl. Phys.*, vol. 53, no. 21, p. 213001, May 2020.

[281] L. Li, X. Liu, S. Pal, S. Wang, C. K. Ober, and E. P. Giannelis, “Extreme ultraviolet resist materials for sub-7 nm patterning,” *Chem. Soc. Rev.*, vol. 46, no. 16, pp. 4855–4866, 2017.

[282] I. Fomenkov *et al.*, “Light sources for high-volume manufacturing EUV lithography: technology, performance, and power scaling,” *Adv. Opt. Technol.*, vol. 6, no. 3–4, Jan. 2017.

[283] A. Wei *et al.*, “Advanced Node DTCO in the EUV Era,” in *2020 IEEE International Electron Devices Meeting (IEDM)*, 2020, pp. 41.2.1-41.2.4.

[284] N. Large, M. Abb, J. Aizpurua, and O. L. Muskens, “Photoconductively Loaded Plasmonic Nanoantenna as Building Block for Ultracompact Optical Switches,” *Nano Lett.*, vol. 10, no. 5, pp. 1741–1746, May 2010.

[285] M. Castro-Lopez, N. de Sousa, A. Garcia-Martin, F. Y. Gardes, and R. Sapienza, “Scattering of a plasmonic nanoantenna embedded in a silicon

waveguide," *Opt. Express*, vol. 23, no. 22, p. 28108, Nov. 2015.

[286] J. Ye, Y. Li, and S. Qu, "Energy-Efficiency Switchable Grating Coupler for Intra-Chip Wireless Optical Interconnection," *IEEE Photonics Technol. Lett.*, vol. 31, no. 17, pp. 1429–1432, Sep. 2019.

[287] R. Bruck and O. L. Muskens, "Plasmonic nanoantennas as integrated coherent perfect absorbers on SOI waveguides for modulators and all-optical switches," *Opt. Express*, vol. 21, no. 23, p. 27652, Nov. 2013.

[288] Y.-H. Jin, B. J. Park, and M.-K. Kim, "Extreme field enhancement in nano-gap plasmonic cavity via 90% efficient coupling with silicon waveguide," *Opt. Express*, vol. 24, no. 22, p. 25540, Oct. 2016.

[289] D. Vercruysse, P. Neutens, L. Lagae, N. Verellen, and P. Van Dorpe, "Single Asymmetric Plasmonic Antenna as a Directional Coupler to a Dielectric Waveguide," *ACS Photonics*, vol. 4, no. 6, pp. 1398–1402, Jun. 2017.

[290] Z. Li *et al.*, "Controlling propagation and coupling of waveguide modes using phase-gradient metasurfaces," *Nat. Nanotechnol.*, vol. 12, no. 7, pp. 675–683, Jul. 2017.

[291] F. Bernal Arango, A. Kwadrin, and A. F. Koenderink, "Plasmonic Antennas Hybridized with Dielectric Waveguides," *ACS Nano*, vol. 6, no. 11, pp. 10156–10167, Nov. 2012.

[292] A. Kriesch, S. P. Burgos, D. Ploss, H. Pfeifer, H. A. Atwater, and U. Peschel, "Functional Plasmonic Nanocircuits with Low Insertion and Propagation Losses," *Nano Lett.*, vol. 13, no. 9, pp. 4539–4545, Sep. 2013.

[293] F. Peyskens, A. Dhakal, P. Van Dorpe, N. Le Thomas, and R. Baets, "Surface Enhanced Raman Spectroscopy Using a Single Mode Nanophotonic-Plasmonic Platform," *ACS Photonics*, vol. 3, no. 1, pp. 102–108, Jan. 2016.

[294] Y. Luo *et al.*, "On-Chip Hybrid Photonic–Plasmonic Light Concentrator for Nanofocusing in an Integrated Silicon Photonics Platform," *Nano Lett.*, vol. 15, no. 2, pp. 849–856, Feb. 2015.

[295] A. Espinosa-Soria, E. Pinilla-Cienfuegos, F. J. Díaz-Fernández, A. Griol, J. Martí, and A. Martínez, "Coherent Control of a Plasmonic Nanoantenna

Integrated on a Silicon Chip," *ACS Photonics*, vol. 5, no. 7, pp. 2712–2717, Jul. 2018.

[296] A. Espinosa-Soria, A. Griol, and A. Martínez, "Experimental measurement of plasmonic nanostructures embedded in silicon waveguide gaps," *Opt. Express*, vol. 24, no. 9, p. 9592, May 2016.

[297] G. A. E. Vandenbosch and Z. Ma, "Upper bounds for the solar energy harvesting efficiency of nano-antennas," *Nano Energy*, vol. 1, no. 3, pp. 494–502, May 2012.

[298] S. A. Maier, "Plasmonic field enhancement and SERS in the effective mode volume picture," *Opt. Express*, vol. 14, no. 5, p. 1957, Mar. 2006.

[299] R. Kou *et al.*, "Spectrometric analysis of silicon nitride films deposited by low-temperature liquid-source CVD," *J. Appl. Phys.*, vol. 126, no. 13, p. 133101, Oct. 2019.

[300] S. Babar and J. H. Weaver, "Optical constants of Cu, Ag, and Au revisited," *Appl. Opt.*, vol. 54, no. 3, p. 477, Jan. 2015.

[301] A. Golovashkin, I. Leksina, G. Motulevich, and A. Shubin, "The Optical Properties of Niobium," *Sov. J. Exp. Theor. Phys.*, vol. 29, pp. 27–34, Jan. 1969.

[302] K. M. McPeak *et al.*, "Plasmonic Films Can Easily Be Better: Rules and Recipes," *ACS Photonics*, vol. 2, no. 3, pp. 326–333, Mar. 2015.

[303] V. V Nemoshkalenko, V. N. Antonov, and M. M. Kirillova, "The structure of the energy bands and optical absorption in osmium," *Zh. Eksp. Teor. Fiz.*, vol. 90, no. 1, pp. 201–208, Jan. 1986.

[304] A. G. Mathewson and H. P. Myers, "Absolute Values of the Optical Constants of Some Pure Metals," *Phys. Scr.*, vol. 4, no. 6, pp. 291–292, Dec. 1971.

[305] M. A. Ordal, R. J. Bell, R. W. Alexander, L. L. Long, and M. R. Querry, "Optical properties of Au, Ni, and Pb at submillimeter wavelengths," *Appl. Opt.*, vol. 26, no. 4, p. 744, Feb. 1987.

[306] K. J. Palm, J. B. Murray, T. C. Narayan, and J. N. Munday, "Dynamic Optical Properties of Metal Hydrides," *ACS Photonics*, vol. 5, no. 11, pp. 4677–4686, Nov. 2018.

[307] N. V. Smith, “Optical Constants of Sodium and Potassium from 0.5 to 4.0 eV by Split-Beam Ellipsometry,” *Phys. Rev.*, vol. 183, no. 3, pp. 634–644, Jul. 1969.

[308] N. V. Smith, “Optical Constants of Rubidium and Cesium from 0.5 to 4.0 eV,” *Phys. Rev. B*, vol. 2, no. 8, pp. 2840–2848, Oct. 1970.

[309] J. H. Weaver, C. G. Olson, and D. W. Lynch, “Optical investigation of the electronic structure of bulk Rh and Ir,” *Phys. Rev. B*, vol. 15, no. 8, pp. 4115–4118, Apr. 1977.

[310] T. Inagaki, E. T. Arakawa, and M. W. Williams, “Optical properties of liquid mercury,” *Phys. Rev. B*, vol. 23, no. 10, pp. 5246–5262, May 1981.

[311] A. I. Golovashkin and G. P. Motulevich, “Optical and electrical properties of tin,” *Sov. Phys. JETP*, vol. 19, no. 2, pp. 310–317, Feb. 1964.

[312] M. A. Ordal, R. J. Bell, R. W. Alexander, L. A. Newquist, and M. R. Querry, “Optical properties of Al, Fe, Ti, Ta, W, and Mo at submillimeter wavelengths,” *Appl. Opt.*, vol. 27, no. 6, p. 1203, Mar. 1988.

[313] M. Rasigni and G. Rasigni, “Optical constants of lithium deposits as determined from the Kramers-Kronig analysis,” *J. Opt. Soc. Am.*, vol. 67, no. 1, p. 54, Jan. 1977.

[314] M. Querry, *Optical Constants of Minerals and Other Materials from the Millimeter to the Ultraviolet*. Chemical Research, Development & Engineering Center, US Army Armament Munitions Chemical Command., 1987.

[315] C. Peng *et al.*, “Performance improvement of Sb2Te3 phase change material by Al doping,” *Appl. Surf. Sci.*, 2011.

[316] I. Hilmi, E. Thelander, P. Schumacher, J. W. Gerlach, and B. Rauschenbach, “Epitaxial Ge 2 Sb 2 Te 5 films on Si(111) prepared by pulsed laser deposition,” *Thin Solid Films*, vol. 619, pp. 81–85, Nov. 2016.

[317] S. Bai, Z. Tang, Z. Huang, and J. Wang, “Transient Thermoreflectance Measurement of Thermal Conductivity of Nanoscale Silicon Nitride Thin Films,” in *ASME 5th International Conference on Nanochannels, Microchannels, and Minichannels*, 2007, pp. 457–460.

[318] E. Gemo *et al.*, “Sub-wavelength plasmonic-enhanced phase-change

memory," in *Photonic and Phononic Properties of Engineered Nanostructures X*, 2020, p. 50.

[319] K. Thyagarajan, C. Santschi, P. Langlet, and O. J. F. Martin, "Highly Improved Fabrication of Ag and Al Nanostructures for UV and Nonlinear Plasmonics," *Adv. Opt. Mater.*, vol. 4, no. 6, pp. 871–876, 2016.

[320] S. H. Badri, M. M. Gilarlue, S. G. Farkoush, and S.-B. Rhee, "Reconfigurable bandpass optical filters based on subwavelength grating waveguides with a Ge 2 Sb 2 Te 5 cavity," *J. Opt. Soc. Am. B*, vol. 38, no. 4, p. 1283, Apr. 2021.

[321] M. Stegmaier, C. Ríos, H. Bhaskaran, C. D. Wright, and W. H. P. Pernice, "Nonvolatile All-Optical 1 × 2 Switch for Chipscale Photonic Networks," *Adv. Opt. Mater.*, vol. 5, no. 1, p. 1600346, Jan. 2017.

[322] P. Xu, J. Zheng, J. K. Doylend, and A. Majumdar, "Low-Loss and Broadband Nonvolatile Phase-Change Directional Coupler Switches," *ACS Photonics*, vol. 6, no. 2, pp. 553–557, Feb. 2019.

[323] C. Chen, D. A. Mohr, H. K. Choi, D. Yoo, M. Li, and S. H. Oh, "Waveguide-Integrated Compact Plasmonic Resonators for On-Chip Mid-Infrared Laser Spectroscopy," *Nano Lett.*, vol. 18, no. 12, pp. 7601–7608, 2018.

[324] H. Betel and P. Flocchini, "On the relationship between fuzzy and Boolean cellular automata," *Theor. Comput. Sci.*, vol. 412, no. 8–10, pp. 703–713, Mar. 2011.

[325] Y. Cao, D. T. Gillespie, and L. R. Petzold, "Efficient step size selection for the tau-leaping simulation method," *J. Chem. Phys.*, vol. 124, no. 4, p. 044109, Jan. 2006.

[326] H. Sánchez-Esquivel *et al.*, "Spectral dependence of nonlinear absorption in ordered silver metallic nanoprism arrays," *Sci. Rep.*, vol. 7, no. 1, p. 5307, Dec. 2017.

[327] J. Jin, *The Finite Element Method in Electromagnetics - 2nd edition*, 2nd ed. John Wiley & Sons, 2002.

[328] A. B. COMSOL, *COMSOL Multiphysics® v. 5.5 reference manual*. 2020.

[329] S. Brenner and R. Scott, *The mathematical theory of finite element methods*, vol. 15. Springer Science & Business Media, 2007.

[330] C. Johnson, *Numerical solution of partial differential equations by the finite element method*. Courier Corporation, 2012.

[331] Y. Saad and M. H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM J. Sci. Stat. Comput.*, vol. 7, no. 3, pp. 856–869, 1986.

[332] O. Schenk and K. Gärtner, "Solving unsymmetric sparse systems of linear equations with PARDISO," *Futur. Gener. Comput. Syst.*, vol. 20, no. 3, pp. 475–487, 2004.

[333] M. Benzi, "Preconditioning Techniques for Large Linear Systems: A Survey," *J. Comput. Phys.*, vol. 182, no. 2, pp. 418–477, Nov. 2002.

[334] A. Debunne *et al.*, "Evidence of Crystallization–Induced Segregation in the Phase Change Material Te-Rich GST," *J. Electrochem. Soc.*, vol. 158, no. 10, p. H965, Aug. 2011.

[335] H.-G. Boyen *et al.*, "Low-temperature interface reactions in layered Au/Sb films: In situ investigation of the formation of an amorphous phase," *Phys. Rev. B*, vol. 51, no. 3, pp. 1791–1802, Jan. 1995.

[336] P. Noé, C. Sabbione, N. Bernier, N. Castellani, F. Fillot, and F. Hippert, "Impact of interfaces on scenario of crystallization of phase change materials," *Acta Mater.*, vol. 110, pp. 142–148, May 2016.

[337] L. Lu, W. Dong, J. K. Behera, L. Chew, and R. E. Simpson, "Inter-diffusion of plasmonic metals and phase change materials," *J. Mater. Sci.*, vol. 54, no. 4, pp. 2814–2823, Feb. 2019.

[338] J. Faneca *et al.*, "Performance characteristics of phase-change integrated silicon nitride photonic devices in the O and C telecommunications bands," *Opt. Mater. Express*, vol. 10, no. 8, p. 1778, Aug. 2020.

[339] D. Das, "A Survey on Cellular Automata and Its Applications," in *Communications in Computer and Information Science*, 2012, pp. 753–762.

[340] J. Sarkar, *Sputtering Materials for VLSI and Thin Film Devices*. William Andrew, 2013.

[341] M. Behrens *et al.*, "Impact of disorder on optical reflectivity contrast of epitaxial Ge2Sb2Te5 thin films," *CrystEngComm*, vol. 20, no. 26, pp. 3688–

3695, Jun. 2018.

[342] Q. Yin and L. Chen, “Enhanced optical properties of Sn-doped Ge2Sb2Te5 thin film with structural evolution,” *J. Alloys Compd.*, vol. 770, pp. 692–700, Jan. 2019.