UNIVERSITE DE LILLE

Ecole doctorale : Sciences pour l'Ingénieur
Laboratoire : Institut d'Electronique, de Microélectronique et de Nanotechnologie - IEMN

THESE
Présentée par

Arun BHASKAR

Pour obtenir le grade de
Docteur de l'Université

Electronique, microélectronique, nanoélectronique et micro-ondes

Ingénierie de substrat par micro-usinage laser pour l'amélioration des performances de composants et fonctions RF intégrées en technologie SOI-CMOS

Soutenue le 07/10/2019 devant le jury composé de :

Mme Florence GARRELIE *Professeur, Université Jean Monnet, Saint-Etienne*	Rapporteur
Mr Pierre BLONDY *Professeur, Université de Limoges, XLIM*	Rapporteur
Mr Olivier UTEZA *Directeur de Recherche CNRS, LP3, Marseille*	Président
Mme Christine RAYNAUD *Marketing Manager RF Technologies, CEA-LETI, Grenoble*	Examinateur
Mr Jean-François ROBILLARD *Enseignant-Chercheur ISEN-IEMN, Lille*	Examinateur
Mr Emmanuel DUBOIS *Directeur de Recherche CNRS, IEMN, Villeneuve d'Ascq*	Directeur de thèse
Mr Christophe GAQUIERE *Professeur, Université de Lille*	Co-directeur de thèse
Mr Daniel GLORIA *Ingénieur R&D STMicroelectronics, Crolles*	Encadrant Industriel

UNIVERSITY OF LILLE

Doctoral school : Sciences for engineering
Laboratory : Institut d'Electronique, de Microélectronique et de Nanotechnologie - IEMN

THESIS
Presented by

Arun BHASKAR

For obtaining the grade of
Doctor of University

Electronics, Microelectronics, Nanoelectronics and Microwave

Substrate engineering using laser micromachining for improvement of RF devices and systems integrated in SOI-CMOS technology

Defended on 07/10/2019 in front of the jury composed of:

Mrs Florence GARRELIE *Professeur, Université Jean Monnet, Saint-Etienne*	Rapporteur
Mr Pierre BLONDY *Professeur, Université de Limoges, XLIM*	Rapporteur
Mr Olivier UTEZA *Directeur de Recherche CNRS, LP3, Marseille*	Président
Mme Christine RAYNAUD *Marketing Manager RF Technologies, CEA-LETI, Grenoble*	Examinateur
Mr Jean-François ROBILLARD *Enseignant-Chercheur ISEN-IEMN, Lille*	Examinateur
Mr Emmanuel DUBOIS *Directeur de Recherche CNRS, IEMN, Villeneuve d'Ascq*	Directeur de thèse
Mr Christophe GAQUIERE *Professeur, Université de Lille, IEMN*	Co-directeur de thèse
Mr Daniel GLORIA *Ingénieur R&D STMicroelectronics, Crolles*	Encadrant Industriel

Acknowledgements

I am very grateful for having got this opportunity to work on a challenging project which was multi-dimensional. I would like to thank my PhD supervisors Emmanuel, Daniel and Christophe for having constantly guided me. I have received good support and encouragement to try new ideas from all three of them. This has made the thesis a very memorable and cherished experience and I am thankful for that.

I would like to specially thank Emmanuel who had the oversight of my daily activities. His very approachable nature made it very easy for me to express my ideas and questions to him. He would always respond in an enthusiastic manner and spend adequate time to discuss the matter at hand. This constant support enabled me to stay motivated during the thesis.

It was great to work and also have social interactions with colleagues of my group over these years: Jean-François, Maciej, Valentina, Stanislav, Jean-Marc, Thierno, Quentin, Di, Tianqi, Justine, Jun, Pascale, Jean-Michel. I was also benefitted by a lot of technical help from other members of the lab. I would like to thank the cleanroom and RF characterization technical staff, Matthieu, Flavie, Etienne, and Vanessa for their valuable contributions at different phases of PhD.

Big thanks to all of my colleagues at ST Microelectronics and the RFSS team. I really liked the jovial morning coffee breaks and supportive working environment there and the team outings. A special mention of thanks to all the colleagues with their important contributions to the thesis: Cédric, Yohann, Simon, Vipin, Philippe, Jérome, Raphael.

I am in gratitude for my family who always encouraged me in my pursuits of education. Heartfelt thanks to my wife Deepthi, who has been a great source of motivation for me for well over two years now. She has enabled me to gain strength and help develop my overall personality. I would also like to express my thanks to all my good friends in Lille who I had a lot of memorable moments with: Gatien, Dev, Ragini, Ankita, Richa, Inder, Tanu, Naina, Shilpa, Piyush, Spurthi, and Ashwini.

I sincerely thank God who has given me knowledge that I have been blessed enough to receive. Heartfelt expressions of gratitude to my hero Swami Vivekananda and other great spiritual leaders whose words empower and embolden me to be who I am.

Table of contents

General Introduction 1

Chapter 1: Synergy of laser processing and SOI technology for More than Moore advancements 5

1.0 Introduction... 5

1.1 The semiconductor industry: An overview 7

1.2 *More than Moore*: Introduction and recent advances 9

1.2.1 Motivation...10

1.2.2 SoC integration: Die level optimization 11

1.2.3 Packaging: System level optimization 13

1.3 Laser processing: An emerging driver for *More than Moore* 17

1.3.1 The evolution of femtosecond laser 17

1.3.2 Application of femtosecond laser processing in *More than Moore*...18

1.3.2.1 Packaging and integration...18

1.3.2.2 MEMS...21

1.3.2.3 Microfluidics...23

1.3.2.4 Integrated photonics 25

1.4 SOI Technology: A platform for SoC integration 27

1.4.1 Memory and logic 29

1.4.2 RF/microwave circuits 30

1.4.3 Integrated photonic circuits 34

1.4.4 MEMS 36

Concluding Remarks...37

References 39

Chapter 2: Towards high-end RF electronics and laser material processing: State of the art 43

2.0 Introduction...43

2.1 Micromachining of silicon 45

2.1.1 Overview of methods for silicon etching 45

2.1.2 Femtosecond Laser processing: An unique tool for micro/nanostructuring of silicon...47

2.1.3 Laser micromachining: A potential candidate for high speed and large area silicon etching...50

2.2 Inductors 52
2.2.1 Factors affecting inductor performance 52
2.2.2 Inductors on SOI....................55
2.2.2.1 Q-factor using different substrate resistivities and BOX thickness 55
2.2.2.2 Inductor on commercial SOI and impact on device performance 57
2.2.3 Suspended inductors 58
2.3 RF Switch....................59
2.3.1 Overview of different implementations and technologies 61
2.3.2 Performance comparison: Bulk vs. SOI 62
2.4 Low noise amplifier (LNA)....................65
2.4.1 LNA FoM: Bulk vs. SOI....................65
2.4.2 LNA performance comparison by optimization of passives....................66
Concluding remarks....................69
References 71
Chapter 3: Laser processing applied to RF functions: Theoretical background and experimental methodology 77
3.0 Introduction....................77
3.1 Laser processing: Essential concepts....................79
3.1.1 Gaussian beam parameters....................79
3.1.2 Pulsed laser processing parameters....................81
3.2 Laser processing: Description of physical processes and timescales 85
3.3 Effect of different parameters on laser processing of silicon 89
3.3.1 Ambient conditions....................90
3.3.1.1 Medium of ablation 90
3.3.1.1 Substrate temperature....................91
3.3.2 Laser wavelength 92
3.3.3 Laser pulse width 94
3.3.3.1 Beam distortion effect....................94
3.3.3.2 Modification of ablation threshold....................95
3.3.4 Polarization....................96
3.4 Effect of ablation plume on laser processing....................97
3.5 Laser processing: Description of system....................100
3.5.1 Laser source and optics 100
3.5.2 Optical attenuator 101
3.5.3 Trepan head....................102

3.5.4 Galvanometric scanner 102
3.6 Laser processing: Experimental plan 102
3.7 RF circuits: Description and theory 104
3.7.1 Isolation structures 105
3.7.2 RF Switch 106
3.7.3 Inductors 108
3.7.4 Low Noise Amplifier (LNA) 109
Conclusion 112
References 113
Chapter 4: Femtosecond Laser Assisted Micromachining and Etch (FLAME) for SOI-CMOS technology 115
4.0 Introduction 115
4.1 Determination of laser beam waist and threshold fluence 117
4.2 Scribing of trenches in silicon 119
4.2.1 Galvo stability: 343 nm and 1030 nm line 120
4.2.2 Study of trenches: 343 nm source 121
4.2.2.1 Effect of speed 121
4.2.2.2 Effect of fluence 122
4.2.3 Study of trenches: 1030 nm source 123
4.2.3.1 Effect of speed 124
4.2.3.2 Effect of fluence 124
4.2.3 Detailed study of trench parameters for optimization of milling 126
4.2.3.1 Average trench depth (d_t) 127
4.2.3.2 Recast layer volume (V_r) 131
4.2.3.3 Recast layer average height (h_{avg}) 132
4.2.3.4 Average roughness of trench profile (R_a) 133
4.2.3.5 Trench width (W_t) 134
4.3 Laser milling 136
4.3.1 Choice of milling pitch 136
4.3.2 Determination of depth as a function of number of scans 138
4.3.3 Effects of ablation plume and redeposition on laser milling 139
4.3.3.1 Generation of particles 140
4.3.3.2 Screening of laser radiation 142
4.3.3 Analysis of milling quality 146
4.4 Femtosecond Laser Assisted Micromachining and Etch (FLAME) process 152

Concluding Remarks........154
References........155
Chapter 5: FLAME membranes for high-performance RF frontends 115
5.0 Introduction........157
5.1 Isolation structures........159
5.1.1 FLAME process parameters........159
5.1.2 S-parameter characterization........159
5.2 RF SP9T switch........163
5.2.1 FLAME process parameters........164
5.2.2 DC characterization........166
5.2.3 S-parameter characterization........167
5.2.4 Large signal characterization........170
5.2.4.1 Measurement setup........170
5.2.4.2 Results........172
5.3 Inductors........176
5.3.1 FLAME process parameters........176
5.3.2 S-parameter characterization........177
5.3.3 Modelling of inductors on membranes........185
5.4 Low noise amplifier (LNA)........189
5.4.1 FLAME process parameters........189
5.4.2 Noise figure measurement........190
5.4.3 Standalone input inductor measurement........191
5.4.4 Linearity measurements........192
Concluding Remarks........194
References........195

Conclusion and perspectives........197
References........200

List of Figures

Fig. 1.1: (a) Cumulative revenue generated as a result of continuous scaling of feature size of semiconductor chips [3] (b) The end products of scaling at different points of time depicting the shift in high volume markets opened up as a result of Moore scaling [4] 7
Fig. 1.2: The emergence of different paradigms of connected devices: Beginning from local area networks (LAN) connecting two or more desktop stations in a localized geographical area, the internet was born connecting computers worldwide. Advent of social media, entertainment and communication platforms enabled massive entry of people on the web. The next phase that is emerging rapidly is the Internet of Things [11]. 8
Fig. 1.3: Schematic comparing the different methods of system integration [12] 9
Fig. 1.4: A real-world analogy of a system which is compared to a city. Just as different bodies of the city need to interlink and communicate with each other, an electronic system needs to optimize its functioning between the interconnected components for efficiency. [13] 10
Fig. 1.5: A comparison of different integration schemes and the tradeoffs obtained between cost, performance and form factor for each type of technology.. 11
Fig. 1.6: Power management scheme on a portable electronic device depicting the typical range of voltages on the storage side and the load side [14]... 12
Fig. 1.7: Chip realized on BCD platform containing digital, analog, memory and High Voltage blocks [15] 12
Fig. 1.8: Different packaging integration schemes (a) SiP (b) PoP (c) PiP [16] 13
Fig. 1.9: 3D stacking of two dies using Through-Silicon Vias (TSVs) for chip to chip interconnections. The four arrows point to the locations where the reliability concerns emerge in such a package. The final package is ready for board level integration using solder bumps connected to the substrate. [19] 14
Fig. 1.10: (a) Array of TSVs realized using via-middle process flow (b) Close-up view of via (c) Cu interconnection by Cu-Cu TCB bonding (d) Cross section view of a 5 die stack [21] 15
Fig. 1.11: 2.5D packaging concept containing a silicon interposer for interconnection of dies. The demonstrated package here is a FPGA product developed by Xilinx [22]..................................... 16
Fig. 1.12: Comparison of different packaging technologies in terms of system performance metrics. [23] 16
Fig. 1.13: Comparison of 3 generations of femtosecond laser systems. Systems 1-7 have already been demonstrated. The systems 8 – 10 were under development as of the publication date of the article [28].. 18
Fig. 1.14: Use of temporary bonding in FO-WLP (a) For backside processing of wafer (b) For die placement and packaging (c) The release process of temporary bonding material using laser irradiation [31] 19
Fig. 1.15: The use of femtosecond laser for through silicon vias (a) Gaussian beam (b) Conventional Bessel beam (c) Tailored Bessel beam (d) Vias in 100 µm silicon (d) Front side of TSV array (b) Rear side of TSV array [33] 19
Fig. 1.16: Surface passive pressure sensor integrated on a thin film foil (a) Finished package (b) TGV for sensor electrical connection (c) Bonded sensor front and back view (d) Cavity in glass to create chamber for pressure reference [34] 20
Fig. 1.17: (a) Diaphragm in 4H-SiC substrate (b,c) SEM image of diaphragm milled with pulse energy 0.05 mJ and 0.15 mJ [35] 21
Fig. 1.18: (a) Dimensions of the comb drive: a=20 µm, b=10 µm, l=21 µm, d=15 µm (b) SEM micrograph of machined microgripper [36] 22
Fig. 1.19: (a) MEMS mold fabricated in silicon (b) MEMS structures realized using micromolding of polymeric precursors [37], [38] 22

Fig. 1.20: Micromachined MEMS cantilevers in silicon [39] 23
Fig. 1.21: (a) Micron sized particle filter using two photon polymerization (b) Filter matrix along the microfluidic channel in fused silica using selective etching of femtosecond laser irradiated areas in HF and KOH [40], [41] 23
Fig. 1.22: (a,b) Microfludic mixer schematic in fused silica (c,d) Microscope images of inscribed mixing channels (e,f) Fluorescence microscopy to quantify mixing [42] 24
Fig. 1.23: (a) Electrofluidic chip schematic with external supply (b) Fabricated electrofluidic chip containing two electrodes which are baised using external supply (c) The alignment of Euglena affected in the presence of external field. Turning off the field restores the random orientation of cells [43] 24
Fig. 1.24: (a) BFL fabricated on PMMA (b) 1D photonic crystal on sapphire [45], [46] 25
Fig. 1.25: Fabrication of 3D high-Q resonator in fused silica with SEM images (a) Sequence of steps (b) After laser direct write and HF etch (c) After laser annealing (d) Resonator with tilted angle (e) Vertical resonator with different heights [47] 26
Fig. 1.26: (a) Cross-sectional view of microring resonator (b) SEM image of the resonator (c) Transmission characteristics of the resonator [48] 26
Fig. 1.27: Large area microlens array fabricated on thin film silicon with average diameter (a) 20 µm (b) 30 µm with corresponding depth profiles shown on the left[49] 27
Fig. 1.28: Transistor configurations for PD-SOI and FD-SOI [51] 27
Fig. 1.29: Ultra thin body scaling routes for FDSOI technology [52] 29
Fig. 1.30: (a) SPICE simulation 16-bit adder circuit on bulk and FDSOI technology (b) Frequency gain measured for LVT (low threshold voltage) and RVT (regular threshold voltage) for regular bulk, strain engineered bulk and FDSOI technologies [52] 30
Fig. 1.31: Transistor f_t/f_{max} specification for different silicon based technology nodes [55] 31
Fig. 1.32: (a) Schematic of DSN measurement showing 8 inverter chains on digital side (b) Measured noise power on the analog side at two different supply voltages[60] 32
Fig. 1.33: Linearity of a coplanar waveguide measured for different substrate types with transmission line of length (a) 2.1 mm (b) 1 mm [61], [62] 33
Fig. 1.34: (a) Schematic showing the charge state at the silicon oxide interface at different bias conditions (b) Performance comparison between eSi and HR-SOI substrate [64] 34
Fig. 1.35: Optical functions on SOI substrate [67] 35
Fig. 1.36: Mach Zehnder Interferometer and optical switch [69] 36
Fig. 1.37: Pressure sensors on SOI membranes (a) N-MOSFET suspended on membrane of area 250 x 400 µm^2 (b) A current source suspended on the edge of the membrane [71] 36
Fig. 1.38: Self assembled 3D MEMS structures using CMOS compatible processes for temperature and flow sensing applications [71] 37
Fig. 1.39: A route for RF performance enhancement by making use of local substrate removal for alleviation of substrate effects 38

Fig. 2.1: Example of macro-sized cavities etched in silicon using (a) TMAH: Etch Rate – 1.2 µm min^{-1} [8] (b) Advanced silicon etching ,a variant of Bosch process: Etch Rate – 15 µm min^{-1} [9] . 46
Fig. 2.2: LIPSS on Titanium illustrated with SEM images taken after (a) Initial surface (b) 2 shots (c) 10 shots (d) 40 shots [10] 47
Fig. 2.3: Color printing on stainless steel using LIPSS [11] 48
Fig. 2.4: (a) LIPSS on silicon (b,c,d) Structured silicon views from different angles (e) Reflectance measurements of microstructured silicon 48
Fig. 2.5: (a) Method of fabrication of nanoparticles array by ejecting nanoparticles from a parent substrate to a host substrate (b) SEM images showing the formation of nanoparticle with the application of higher laser fluences (c) Nanoparticle array fabricated on glass substrate with the inset showing a single nanoparticles [15] 49

Fig. 2.6: Examples of milled cavities in silicon 50
Fig. 2.7 : (a) Inductance for varying number of turns (b) Q-factor for varying number of turns [24] 53
Fig. 2.8: (a) Characteristic dimensions of an inductor showing segments of two neighbouring turns (b) Variation of inductance, maximum quality factor and resonance frequency with variation in different dimensions and design of the inductor [24] 54
Fig. 2.9: (a) Inductance with varying substrate resistivity (indicated in legend in units of Ω.m) (b) Q-factor for varying substrate resistivity [24] 54
Fig. 2.10: Quality factor obtained for the use of different dielectrics [26] 55
Fig. 2.11: (a) Q-factor for different substrate resistivities (b) Self-resonance and inductance values for different substrate resistivities [27] 56
Fig. 2.12: (a) Q-factor obtained for varying thickness of oxide layer (b) Inductance values obtained on bulk Si wafers with a oxide thickness of 25 μm [26], [28] 56
Fig. 2.13: (a)Parallel stacking of metal lines to improve Q factor (b) Cross section of the inductor (c) Q-factor curves for varying number of parallelly stacked metal lines indicated in parantheses for RRS (10-20 Ω.cm) and HRS (>300 Ω.cm) substrate types. [29] 57
Fig. 2.14: (a) VCO circuit containing parallel LC tank circuit (b) Phase noise measured for different VCOs with varying Q-factor values (A) VCO-IV: Q = 52,4-metal HRS (B) VCO-III: Q = 22, 1-metal HRS (C) VCO-II: Q = 30, 4-metal RRS (D) VCO-I: Q = 11, 1-metal RRS 58
Fig. 2.15: SEM micrographs of micromachined inductors on silicon showing (a) Full removal of substrate (b) Partial removal of substrate [30], [31] 58
Fig. 2.16: Front end module components for LTE with switch module outlined 60
Fig. 2.17: Series shunt implementation of the switch showing transistor stacking in the series and shunt branches 60
Fig. 2.18: Examples of implementation of switch on SOI technology (a) S96T: GSM only (b) 2xSP6T: GSM and LTE support [47], [48] 61
Fig. 2.19: (a) R_{on} and C_{off} as a function of stack number (b) R_{on}*C_{off} product as a function of stack number. The switch is fabricated on 0.18 μm thin-film SOI technology [47] 62
Fig. 2.20: Switch characteristics comparison of bulk vs. SOI SPDT switch on 0.25 μm CMOS process (a) Insertion loss (b) Isolation [57] 64
Fig. 2.21: (a) Folded cascode architecture with outlined circuit components representing the different inductors tuned in the design (b) Noise figure as a function of quality factor [73] 67
Fig. 2.22: Effect of (a) source degeneration and (b) tail inductor on noise figure of LNA circuit [73] 68
Fig. 2.23: Effect of gate inductor quality factor on noise performance for (a) Constrained gain (b) Constrained power optimization [22] 69

Fig. 3.1: Gaussian beam description showing (a) the variation of beam width along the propagation direction (b) The radial distribution of intensity at an arbitrary z-position 79
Fig. 3.2: Overlap of pulse energy between successive pulses shown for displacement along a single direction of translation. The markers at $x = 0$ shows the fluence contributions from previous pulses and future pulses 82
Fig. 3.3: Dependence of accumulated fluence on the ratio of interpulse distance and beam waist showing regions of high sensitivity (region 1) and low sensitivity (region 2) 83
Fig. 3.4: Squared diameter of crater plotted as a function of pulse fluence to estimate threshold fluence for different number of pulses on TiO_2 at repetition rate of 1 kHz. The inset shows the estimated threshold values as a function of number of pulses. [3] 84
Fig. 3.5: Schematic representation of interaction between the electronic and lattice systems for (a) Femtosecond laser (b) Nanosecond laser (c) Normalized temperature profiles obtained for copper upon laser irradiation for different pulse durations using the TTM [6], [7] 86

Fig. 3.6: The summary of physical processes upon ultrashort laser irradiation and the typical timescales of these processes. The yellow region of the graph indicates the pulse duration [9] 87
Fig. 3.7: Graphical illustration of sequence of different physical processes due to laser irradiation on a semiconductor (a) Absorption (single and multiphoton) (b) Free carrier absorption/ Inverse bremsstrahlung (c) Impact/Avalanche ionization (d) Non equilibrium carrier distribution after excitation (e) Electron –electron scattering (f) Emission of phonons (g) Radiative recombination (h) Auger recombination (i) Carrier diffusion (j) Thermal diffusion (k) Material ablation (l) Condensation/resolidification [9] 88
Fig. 3.8: Ablation of silicon in air, alcohol and water (a) SEM micrograph of the obtained crater for different number of impulsions (b) Threshold fluence for Np=5000 for the 3 media by squared diameter regression method [19] 90
Fig. 3.9: Ablation threshold dependence on the substrate temperature for 3 wavelengths [20] 91
Fig. 3.10: Influence of substrate temperature on effective cutting speed of wafer grade silicon [21] 92
Fig. 3.11: Ablated volume and efficiency as a function of peak fluence normalized to the threshold fluence for three different wavelengths [23]. 93
Fig. 3.12: (a) Beam energy distribution taken after focusing of laser beam in air and measuring far-field intensity of the beam. The encircled region represents the outline of the original beam. (b) Different effects occurring at long and short pulse widths. At long pulse widths, melting is severe which causes redeposited slag and at short pulse widths non-linear effects results in non-gaussian energy deposition profiles and consequently enlarged ablation profiles. [17] 94
Fig. 3.13: Ablation threshold of silicon in the (a) femtosecond (b) picoseconds regime [23], [25]. In (a), two ablation thresholds are mentioned. Here F_{a1} refers to threshold for onset of ablation and F_{a2} refers to threshold for complete surface ablation 95
Fig. 3.14: Ablation efficiency of different materials studied by making grooves using a Yb: YAG thin disc laser wavelength of 1030 nm and different pulse widths [26].. 96
Fig. 3.15: Effect of polarization on the morphology of trenches micromachined in silicon (a) Polarization perpendicular to axis of the trench (b) Polarization parallel to axis of the trench (c) Circular polarization [2] 96
Fig. 3.16: Quality of exit holes during drilling (a) Using fixed direction of polarization showing elongation in direction perpendicular to direction of polarization (b) Dynamic polarization control with local polarization perpendicular to the tangent of the circular hole. [29] 97
Fig. 3.17: (a) Time evolution of different species of the ablation plume of silicon at a distance of 7 mm from the target surface and peak fluence of 0.5 J cm^{-2} (b) Yield of fast ions as a function of laser fluence (c) Yield of slow ions as a function of laser fluence [30] 98
Fig. 3.18: Velocity distributions during ablation of silicon using different laser parameters (a) Pulse width: 0.9 ps, Wavelength: 1055 nm, Fluence = 0.75 J cm^{-2}, Intensity: 8.3 x 10^{11} W cm^{-2} (b) Pulse width : 0.1 ps, Wavelength: 780 nm, Fluence = 0.5 J cm^{-2}, Intensity: 5.0 x 10^{12} W cm^{-2} [31] 98
Fig. 3.19: Evolution of laser vapour plume of aluminium for a burst with varying number of pulses in the burst showing accumulation of plume with laser conditions as follows: (i) λ = 800 nm (ii) Pulse width: 110 fs (iii) Burst frequency (fp): 1 kHz (iv) Peak fluence: 260 J cm^{-2}. Pictures take approximately 500 μs after the last pulse is applied [32] 99
Fig. 3.20: Schematic description of the laser system depicting the important optical components 101
Fig. 3.21: Variation of pulse energy and average power output as a function of repetition rate . 101
Fig. 3.22: Linear absorption coefficient of silicon as a function of wavelength [34] 103
Fig. 3.23: Schematic representation of an isolation structure with its characteristic dimensions shown in (a) Top view (b) Cross section showing the different SOI layers and path for signal coupling between two fingers 105

Fig. 3.24: Schematic representation of an SP2T switch showing (a) Transistor stacking and relative sizes of series and shunt branches (b) Equivalent circuit of the switch [36], [37] 106
Fig. 3.25: Layout view of the SP9T switch test structure used in the study having GSGSG access. The gate and body bias is tied for all transistor stacks. One DC bias is provided for the ON stack and the second one for other 8 stacks. The different tested bias conditions are shown on the right. 107
Fig. 3.26: (a) Cross sectional schematic of an inductor fabricated on SOI wafer showing the origin of different capacitances and resistances (b) Lumped circuit equivalent for TR-SOI (c) Lumped circuit equivalent for HR-SOI. Taken from [38] 109
Fig. 3.27: Typical receiver architecture showing the different components and associated gains and noise factors [39] 110
Fig. 3.28: (a) Transistor layout showing the implementation of the body contact (b) Circuit diagram for the common source LNA (c) Input block diagram of the LNA. [40] 110
Fig. 3.29: (a) Input stage of a common source LNA (b) Equivalent circuit for calculating the noise at the input stage of LNA [41].. 111

Fig. 4.1: (a) Squared diameter (x) plotted as a function of pulse energy for 1030 nm source. Green line is the least squares fit of the experimental data (b) Optical micrographs of ablated area shown for 4 different fluences (c) Measured diameter in x and y direction for different fluences 117
Fig. 4.2: (a) Squared diameter (x) plotted as a function of pulse energy for 343 nm source. Green line is the least squares fit of the experimental data (b) Optical micrographs of ablated area shown for 4 different fluences. Outlined region for 42.7 mJ cm^{-2} shows the distortion of beam and redistribution of incident energy (c) Measured diameter in x and y direction for different fluences 118
Fig. 4.3: Optical micrographs depicting evolution of laser spot with increasing laser fluence (a) 0.055 J cm^{-2} (b) 0.067 mJ cm^{-2} (c) 0.082 mJ cm^{-2} (d) 0.097 mJ cm^{-2} 119
Fig. 4.4: 2D laser scribing at different galvo speeds for 343 nm source 120
Fig. 4.5: 2D laser scribing at different galvo speeds for 1030 nm source 120
Fig. 4.6: Profile of trenches taken along the centre of the trench for 343 nm source at fluence of 4.2 J cm^{-2} and repetition rate 200 kHz and scan speeds of (a) 25 mm s^{-1} (b) 50 mm s^{-1} (c) 75 mm s^{-1} (d) 100 mm s^{-1} 121
Fig. 4.7: Profile of trenches along the centre of the trench for 343 nm source at scan speed of 100 mm s^{-1} and pulse repetition rate 200 kHz and fluence of (a) 3.7 J cm^{-2} (b) 6 J cm^{-2} (c) 9 J cm^{-2}. The microscope image shows the micro-holes at the bottom of the trench.. 122
Fig. 4.8: Profile of trench compared with the profile of recast layer for trench scribed at 6 J cm^{-2}. The red line is the profile at trench centre while blue line represents the recast layer profile.... 123
Fig. 4.9: Profile of trenches along the centre of the trench for 1030 nm source at fluence of 13.6 J cm^{-2} and repetition rate 30 kHz and scan speeds of (a) 25 mm s^{-1} (b) 50 mm s^{-1} (c) 75 mm s^{-1} (d) 100 mm s^{-1}.. 124
Fig. 4.10: Profile of trenches along the centre of the trench for 1030 nm source at scan speed of 100 mm s^{-1} repetition rate 30 kHz and fluence of (a) 6.6 J cm^{-2} (b) 19.9 J cm^{-2} (c) 33.1 J cm^{-2} (d) 46.1 J cm^{-2}.. 125
Fig. 4.11: Profile of trenches along the centre of the trench for 1030 nm source at scan speed of 30 mm s^{-1} repetition rate 30 kHz and fluence of (a) 6.6 J cm^{-2} (b) 19.9 J cm^{-2} (c) 33.1 J cm^{-2} (d) 46.1 J cm^{-2}.. 125
Fig. 4.12: Cross sectional profile of a trench 127
Fig. 4.13: Average trench depth plotted as a function of fluence for 3 cases at different scan speeds 128
Fig. 4.14: Axial trench profile showing narrowing of trench and possible presence of debris 130

Fig. 4.15: Recast layer volume plotted as a function of fluence for 3 cases at different scan speeds 131
Fig. 4.16: Recast layer average height plotted as a function of fluence for 3 cases at different scan speeds 132
Fig. 4.17: Average roughness of trench plotted as a function of fluence for 3 cases at different scan speeds 134
Fig. 4.18: Inner trench width plotted as a function of fluence for 3 cases at different scan speeds. The error bars represent symmetric error of one standard deviation 135
Fig. 4.19: (a) Cross-sectional profiles of single trench as a function of varying laser power at different scan speeds on 1030 nm source (b) Cross-sectional profiles for two lines scribed parallel to each other with varying pitch. Laser parameters: (i) Power: 0.26 W (ii) Scan speed: 20 mm s^{-1} (iii) Fluence: 6.4 J cm^{-2} 136
Fig. 4.20: (a) Cavity depth obtained for different squares plotted as a function of square side. Milling parameters (i) Source: 1030 nm (ii) Scan speed : 5 mm s^{-1} (iii) Repetition rate: 2 kHz (iv) Fluence: 56.8 J cm^{-2}(v) Number of passes: 15 (b) Plot of depth as a function of different milling parameters for the two laser sources having average power of 1 W for square of side 1 mm. 139
Fig. 4.21: Particle generation and redeposition as a function of laser fluence and pitch. Milling parameters (i) Source: 1030 nm (i) Area: 0.25 x 0.25 mm^2 (ii) Power: Column 1: 0.054 W, Column 2: 0.072 W, Column 3: 0.085 W, Column 4: 0.09 W) (iii) Scan speed: 10 mm s^{-1}........ 140
Fig. 4.22: Generation of particles during milling on 1030 nm source: (a) Case 1: Narrow opening, deep cavity (b) Case 2: Wide opening, deep cavity (c) Case 3: Narrow opening, shallow cavity. Large cavity milling parameters: Scan Speed - 50 mm s^{-1}, Fluence – 61.5 J cm^{-2}, No. of scans- 12. Small cavity milling parameters: Scan speed – 10 mm s^{-1}, Fluence – 7.1 J cm^{-2}, No. of scans – 80 141
Fig. 4.23: Demonstration of presence of ablation plume in the region of ablation (a) Milling process without air flow (b) Milling process with air flow (c) Depth profiling of particle aggregates along the white dotted line. Milling parameters (i) Scan speed: 50 mm s^{-1} (ii) Fluence: 9.8 J cm^{-2} (iii) Pulse repetition rate: 100 kHz (iv) No. of scans: 50 142
Fig. 4.24: Profilometer scans of milling with beam trajectory for each scan shown in the middle. Milling parameters (i) Scan speed: 100 mm s^{-1} (ii) Fluence: 9.9 J cm^{-2} (iii) Pulse repetition rate: 200 kHz (iv) No. of scans: 50 (v) Delay between successive scans: (a) 0 s (b) 1 s (c) 2 s (d) 3 s 144
Fig. 4.25: Profilometer scans of 2-way milling (i) Scan speed: 100 mm s^{-1} (ii) Fluence: 9.9 J cm^{-2} (iii) Pulse repetition rate: 200 kHz (iv) No of scans: 50 (v) Delay between successive scans: (a) 0 s (b) 1 s (c) Two-way milling trajectory 146
Fig. 4.26: Milling profiles taken for two different square cavities of side 1 mm milled on two different sources and different laser parameters (a) (i) Source: 343 nm (ii) Scan speed : 100 mm s^{-1} (iii) Repetition rate: 200 kHz (iv) Fluence: 9.6 J cm^{-2} (v) Number of scans: 40 (b) (i) Source: 1030 nm (ii) Scan speed : 30 mm s^{-1} (iii) Repetition rate: 20 kHz (iv) Fluence: 51.1 J cm^{-2} (v) Number of scans: 8 (c) Plot of depth distributions along with Gaussian fit 147
Fig. 4.27: Morphology of the bottom of the trench after cleaning in ultrasonic bath. Milling using 1030 nm source with parameters: Single scan; Scan speed - 20 mm s^{-1}; Repetition rate - 30 kHz; Pitch - 10 µm; Fluence (a) 2.8 J cm^{-2} (b) 4.7 J cm^{-2} (c) 12.7 J cm^{-2} (d) 17.4 J cm^{-2}........ 148
Fig. 4.28: Morphology of the bottom of the trench after cleaning in ultrasonic bath. Multipass milling (Depth ~ 100 µm) using 1030 nm source with parameters: Scan speed - 20 mm s^{-1}; Repetition rate - 30 kHz; Pitch - 10 µm Fluence/ Number of scans (a) 2.8 J cm^{-2}/65 (b) 4.7 J cm^{-2}/28 (c) 12.7 J cm^{-2}/17 (d) 17.4 J cm^{-2}/10........ 149
Fig. 4.29: Average depth of cavities milled on the 1030 nm source at repetition rate of 30 kHz for different fluences/ number of passes indicated in the legend. Dotted and solid lines represent data before and after ultrasonic cleaning respectively 150

Fig. 4.30: Cavity side walls after cleaning in ultrasonic bath. Milling parameters used are scan speed of 20 mm s^{-1} and repetition rate of 30 kHz with fluence/ number of scans (a) 2.8 J cm^{-2}/65 (b) 4.7 J cm^{-2}/28 (c) 12.7 J cm^{-2}/17 (d) 17.4 J cm^{-2}/10 151
Fig. 4.31: Reduction of holes by changing the focus for each pass. Milling parameters: (i) Speed: 30 mm s^{-1} (ii) Repetition rate: 20 kHz (iii) Fluence: 51.1 J cm^{-2} (iv) No. of scans: 8 (iv) Focus change after each scan (a) No focus change (b) 45 µm into the sample 152
Fig. 4.32:FLAME process steps (1) Lamination of Etch protect layer (2) Laser milling under SOI circuit (3) XeF2 dry etching (4) Stripping of etch protect layer. (a) Hatching of area ABCD to define trajectory of laser beam (b) View of cavity obtained in area ABCD 153
Fig. 4.33: Alignment methodology for creating reference axes on backside (a) Origin (b) x-axis (c) y-axis 153
Fig. 4.34: Morphology of cavity after 25 cycles of etching in XeF_2 154

Fig. 5.1: 2-port S-parameters measurement bench 160
Fig. 5.2: Isolation values before FLAME process (a) ISO-A (b) ISO-B (c) ISO-C (d) ISO-D 161
Fig. 5.3: Isolation values after FLAME process (a) ISO-A (b) ISO-B 162
Fig. 5.4: DLM images after FLAME process (a) ISO-A HR (b) ISO-A TR (c) ISO-B HR (d) ISO-B HR 162
Fig. 5.5: Difference in isolation values before and after FLAME process (a) ISO-A (b) ISO-B 163
Fig. 5.6: (a, c) Optical microscope images of cavities obtained for 2-step milling process of SP9T switch. (b,d) Cross sectional images of the fast and slow milling steps respectively. 164
Fig. 5.7: (a) Optical micrograph of switch cavity after laser ablation (b) Profilometer scan of the ablated area with outlined region showing locally deep regions (c) After 25 cycles of XeF_2 etch. Outlined area etched first suggesting a uniform vertical etch rate of cavity (d) After 50 cycles of XeF_2 etch (e) After 75 cycles of XeF_2 etch (f) After 100 cycles of XeF_2 etch (g) View of the switch front side using backlight illumination from the cavity side. 165
Fig. 5.8: DC characteristics of the RF switch on TR substrate for 2 channel lengths before and after substrate removal (a) 180 nm (b) 220 nm 166
Fig. 5.9: Self heating of RF switch seen in DC characterization on TR-SOI (a) 180 nm (b) 220 nm 167
Fig. 5.10: S_{21} plot for two different channel lengths and bias conditions before and after FLAME (a) HR-SOI substrate (b) TR-SOI substrate 169
Fig. 5.11: S_{11} plot for two different channel length and bias conditions before and after FLAME (a) HR-SOI substrate (b) TR-SOI substrate 170
Fig. 5.12: Schematic of 2nd and 3rd harmonic distortion measurement bench 171
Fig. 5.13: Linearity measurement of the bench using a thru line on Calkit 172
Fig. 5.14: Harmonic distortion dependence on area of silicon etched measured at bias B1 on HR substrate 173
Fig. 5.15: 2nd harmonic distortion before and after FLAME process (a) HR-SOI (b) TR-SOI. Symmetric error bars represent single standard deviation. 173
Fig. 5.16: 3rd harmonic distortion before and after FLAME process (a) HR substrate (b) TR substrate. Symmetric error bars represent standard deviation. 174
Fig. 5.17: Failure of ON transistor caused by excessive self heating at high input power 175
Fig. 5.18: Membrane area showing suitability for (a) Small and medium sized inductors (b) Large inductor 177
Fig. 5.19: Open pad deembedding methodology 178
Fig. 5.20: Measured values of self-resistance and Q-factor for the inductor on 4 different dies for HR substrate 178
Fig. 5.21: Corrected values of self-resistance and Q-factor for the inductor on 4 different dies for HR substrate. Here, the contact inductance is not removed from Z_{ind}. 179

Fig. 5.22: Inductor parameters for single turn inductors before and after FLAME (a) Self-inductance (b) Self resistance (c) Q-factor 180
Fig. 5.23: DLM images of single turn inductors taken close to the bond pad 181
Fig. 5.24: Inductor parameters for two turn inductors before and after FLAME (a) Self-inductance (b) Self resistance (c) Q-factor 182
Fig. 5.25: DLM images of two turn inductors taken close to the bond pad 182
Fig. 5.26: DLM images of single turn inductors 1T-S and 1T-M with different bond pad configurations 184
Fig. 5.27: Extracted parameters of inductors on partially removed TR pad without deembedding 185
Fig. 5.28: Pi-Model of integrated inductor on silicon technology 186
Fig. 5.29: Models of inductors on nearly lossless substrates (a) 3-elements frequency dependent model (b) 6-elements model (c) Modified 6-elements model 187
Fig. 5.30: Comparison of measured vs. modelled inductor parameters 188
Fig. 5.31: LNA layout showing the different inductors and corresponding drawn layout area for laser processing 190
Fig. 5.32: DLM images of suspended inductors for 3 test cases 190
Fig. 5.33: Comparison of (a) Noise figure with error bars representing min-max values (b) Gain before and after FLAME process 191
Fig. 5.34: Extracted parameters of input inductor from single port S-parameter characterization for HR substrate (a) Self Inductance (b) Self resistance (c) Q-factor 192
Fig. 5.35: Measurement setup for P_{1dB} and IIP3 193

Fig. C.1: Coupling of light from fibre to photonic IC using grating coupler [5] 199

List of Tables

Table 1.1: Comparison of different laser technologies and their relative capabilities [25] 17
Table 1.2: Comparison of PD-SOI and FD-SOI technologies [51] 28
Table 1.3: f_t/f_{max} values of recent silicon based commercial technologies 32

Table 2.1: Summary of standard silicon etching processes. Adapted from [1] and [2] 45
Table 2.2: High throughput volume ablation rates for different materials [20] 52
Table 2.3: Performance summary of suspended inductors on different substrates. Adapted from [33] 59
Table 2.4: Ron*Coff values for different process technologies [49] 62
Table 2.5: Performance specifications of switches realized on different technologies for GSM frequency bands 63
Table 2.6: Summary of LNAs realized in CMOS technologies. Adapted from [58] 65
Table 2.7: LNA designs on SOI technology. Taken from [69] 66

Table 3.1: Laser ablation crater study depicting ablation threshold, height and volume variation with change in laser wavelength [22] 92
Table 3.2: Summary of laser beam parameters for the 3 laser sources 100
Table 3.3: List of characterized isolation structures and their dimensions 106
Table 3.4: Description of inductors used in the study 108
Table 3.5: Summary of RF characterization performed on different RF circuits 112

Table 4.1: Parameters extracted from average trench depth using curve fitting 128
Table 4.2: Measured quantities at different values of delay for 1-way milling 144
Table 4.3: Measured quantities at different values of delay for 2-way milling 145
Table 4.4: Calculated quantities for milled cavities (single scan) on 1030 nm source at repetition rate of 30 kHz and scan speed of 20 mm s^{-1} 148
Table 4.5: Calculated quantities for milled cavities (multiple scans) on 1030 nm source at repetition rate of 30 kHz and scan speed of 20 mm s^{-1} 151

Table 5.1: FLAME process parameters for isolation structures 159
Table 5.2: Milling parameters of SP9T switch 164
Table 5.3: Milling parameters for inductor study 176
Table 5.4: Summary of inductor measurements 183
Table 5.5: Modified 6-element model parameters for 2-turn inductors 188
Table 5.6: Summary of LNA IIP3 measurements 193
Table 5.7: Summary of LNA P_{1dB} measurements 194

List of Abbreviations

BCD	-	Biploar-CMOS-DMOS
BER	-	Bit Error Rate
BFL	-	Binary Fresnel Lens
BOX	-	Buried Oxide
CAGR	-	Compound Annual Growth Rate
DIBL	-	Drain-Induced Barrier Lowering
DLM	-	Dual-Light Microscopy
DOF	-	Depth of Focus
DSN	-	Digital Substrate Noise
EMC	-	Encapsulant Mold Compound
eSi-SOI	-	enhanced Signal integrity Silicon-on-Insulator
FDD	-	Frequency Division Duplexing
FD-SOI	-	Fully Depleted Silicon-on-Insulator
FEM	-	Front-End Module
FIB	-	Focused Ion Beam
FLAME	-	Femtosecond Laser Assisted Micromachining and Etch
FoM	-	Figure of Merit
FO-WLP	-	Fan-Out Wafer-Level Package
GSG	-	Ground Signal Ground
HR-SOI	-	High-Resistivity Silicon-on-Insulator
IoT	-	Internet of Things
IP3	-	Third-order Intercept Point
ISM	-	Internal Stacking Module
KOZ	-	Keep-Out Zone
LAN	-	Local Area Network
LIPSS	-	Laser-Induced Periodic Surface Structures
LNA	-	Low Noise Amplifier
LoC	-	Lab-on-Chip
NA	-	Numerical Aperture
NF	-	Noise Figure
PD-SOI	-	Partially Depleted Silicon-on-Insulator
PiP	-	Package-in-Package
PoP	-	Package-on-Package
RDL	-	Redistribution layer
RIU	-	Refractive Index Unit
RO	-	Ring Oscillator
RR	-	Rayleigh Range
RRS	-	Regular Resistivity substrate
SiP	-	System-in-Package
SoB	-	System-on-Board
SoC	-	System-on-Chip
SOI	-	Silicon-on-Insulator
TGV	-	Through-Glass Via
TR-SOI	-	Trap-Rich Silicon-on-Insulator
TSV	-	Through-Silicon Via
TTM	-	Two Temperature Model
VCO	-	Voltage Controlled Oscillators
WLP	-	Wafer-Level Package

General Introduction

The invention of transistor and eventually of integrated circuits marks the beginning of the semiconductor industry. Gordon Moore penned the seminal paper on transistor scaling in 1965 predicting that the number of transistors in an integrated circuit doubles every year. A decade later, he revised his prediction that the number of transistors doubles every two years. This law has held remarkably true for several decades and has propelled the industry forward by enabling better performance through transistor scaling.

Since the early 2000s, it was found that transistor scaling alone will not suffice for pushing the performance limits of electronic systems. Also there was a need to improve the diversity of electronic systems in order to support complex functionalities in devices like cell phones. This led to the emergence of a whole new focus area of the semiconductor industry referred to as *More than Moore.* This keyword is an umbrella term for all technological innovations concerning improvement of system performance and its diversity. *More than Moore* deals with many different aspects such as electronic packaging, 3D integration, development of MEMS, photonic ICs, sensors, energy harvesters etc.

Within this broad context of *More than Moore*, the thesis will be focused on two specific areas: femtosecond laser processing and RF (Radio-Frequency) CMOS (Complementary Metal-Oxide-Semiconductor) technologies. The goal of the thesis is to apply laser processing on SOI (Silicon-on-Insulator) RF components like switches, inductors, LNA etc. The idea is to remove handler silicon under the active area of these components to obtain improved RF performance. Substrate coupling is an important factor in SOI RF circuits which results in degradation of performance. By removal of substrate, negative effects resulting from substrate coupling is eliminated which gives better performance.

The organisation of the manuscript is as follows. In Chapter 1, the contextual background of the thesis is given in a detailed manner. *More than Moore* developments over the past years are highlighted. Following this, the relevance of femtosecond laser processing and SOI technology for *More than Moore* are discussed using numerous examples. At the end of the chapter, a schematic is provided to elaborate the goal of the thesis.

In Chapter 2, the state of the art methods and technologies are reviewed within the framework of the thesis. At first, currently available methods for large area

etching of silicon are reviewed. Then, the scope of femtosecond laser processing of silicon is considered showing the versatility of processing. The potential of high speed processing of silicon using femtosecond laser is looked at for industrial competency. After this, state of the art performance of inductors, switches and LNA on SOI technology is discussed showing its advantage over bulk technology. Furthermore, the superior performance of inductors after partial/complete removal of substrate will be demonstrated. The impact of using higher Q-factor inductors for LNA performance is reported. In a nutshell, this chapter covers the benefits of using laser processing for etching silicon and importance of substrate removal.

Chapter 3 provides the necessary background theory and literature review relevant to the thesis. It begins with the definitions of different laser parameters. After this, the laser matter interaction in the femtosecond regime is illustrated. Then, the effect of choice of different chosen parameters on laser processing is discussed. The laser processing system that we use in the work is detailed and the experimental plan for achievement of substrate removal is laid out. Subsequently, the specific details of the RF components studied in this work are listed. The list of characterizations that are performed on different RF components is summarized at the end of the chapter.

The results and discussion for laser processing methodology is the subject matter of Chapter 4. Initially, experiments are performed to compare milling feasibility at two different laser wavelengths. A systematic study of scribing 1D trenches is used as a basis for determining optimal laser parameters. Using the optimal laser parameters, 2D milling is performed and quality is analyzed. Finally, the Femtosecond Laser Assisted Micromachining and Etch (FLAME) process is described which uses an additional step of XeF_2 gas etching to create SOI membranes.

The process described in Chapter 4 is used to create membranes of different circuits mentioned in Chapter 3. In Chapter 5, the results of RF characterization of these membranes are presented. The specific implementation of the FLAME process for each circuit is also presented. For all circuits, all measurements are made before and after FLAME process to quantify the improvement that is obtained by removing the substrate. Compact modelling is also presented for the inductors to understand the frequency dependent losses resulting from metal losses like skin effect and current crowding effect after substrate removal.

Finally, a general conclusion is presented recapitulating all the obtained results and providing an overall picture of the studies carried out in the thesis. A few perspectives are provided for future studies involving further optimization of laser process and newer avenues in substrate removal studies.

Chapter 1: Synergy of laser processing and SOI technology for *More than Moore* advancements

1.0 Introduction

This chapter provides the contextual background for the work that is carried out in the thesis. There are two main aspects in our study: laser material processing and Silicon-on-Insulator (SOI) technology. In a nutshell, the objective of the thesis is the utilization of laser processing for local substrate removal of SOI dies to augment the electrical performance of RF devices.

The framework being set, the organization of the chapter is as follows. Initially, a brief overview of the semiconductor industry is provided, tracing the key advancements. Proceeding further, *More than Moore* advancements and their relevance to today's industry requirements are discussed. This refers to the subset of technological innovations which are concerned with the overall improvement of system performance and extended capabilities like development of sensors, actuators, power harvesters, analog/RF modules, biochips etc.

We then demonstrate how laser processing can be an effective tool which offers several solutions for the *More than Moore* domain. The same discussion is extended to Silicon-on-Insulator (SOI) technology, which is relevant to both Moore scaling and More than Moore applications. It will be seen that both technologies can provide solutions for functional diversification and system performance improvements. At the end of the chapter, the potential impact of two worlds working together is highlighted.

Contents

1.1 The semiconductor industry: An overview 7
1.2 *More than Moore*: Introduction and recent advances 9
1.2.1 Motivation 10
1.2.2 SoC integration: Die level optimization 11
1.2.3 Packaging: System level optimization 13
1.3 Laser processing: An emerging driver for *More than Moore* 17
1.3.1 The evolution of femtosecond laser 17
1.3.2 Application of femtosecond laser processing in *More than Moore* 18
1.3.2.1 Packaging and integration 18
1.3.2.2 MEMS 21
1.3.2.3 Microfluidics 23
1.3.2.4 Integrated photonics 25
1.4 SOI Technology: A platform for SoC integration 27
1.4.1 Memory and logic 29
1.4.2 RF/microwave circuits 30
1.4.3 Integrated photonic circuits 34
1.4.4 MEMS 36
Concluding Remarks 37
References 39

1.1 The semiconductor industry: An overview

The buzzword in the electronics industry for many decades has been "Moore's law". This is a major technological foresight envisioned by Gordon Moore in 1965 [1]. The essence of this law is that the number of transistors doubles in every new technology node, which occurs about once every two years. This is because transistor dimensions scale as seen in Fig. 1.1a and allow higher density of transistors to be integrated on silicon. As a consequence, the speed of circuits and overall performance also improves in each generation [2]. The shrinking of feature size has enabled newer markets and thereby increasing the cumulative revenue of the semiconductor market year on year as seen in Fig. 1.1a. Even today, the transistor scaling continues to be relevant amidst multiple challenges related to channel scaling.

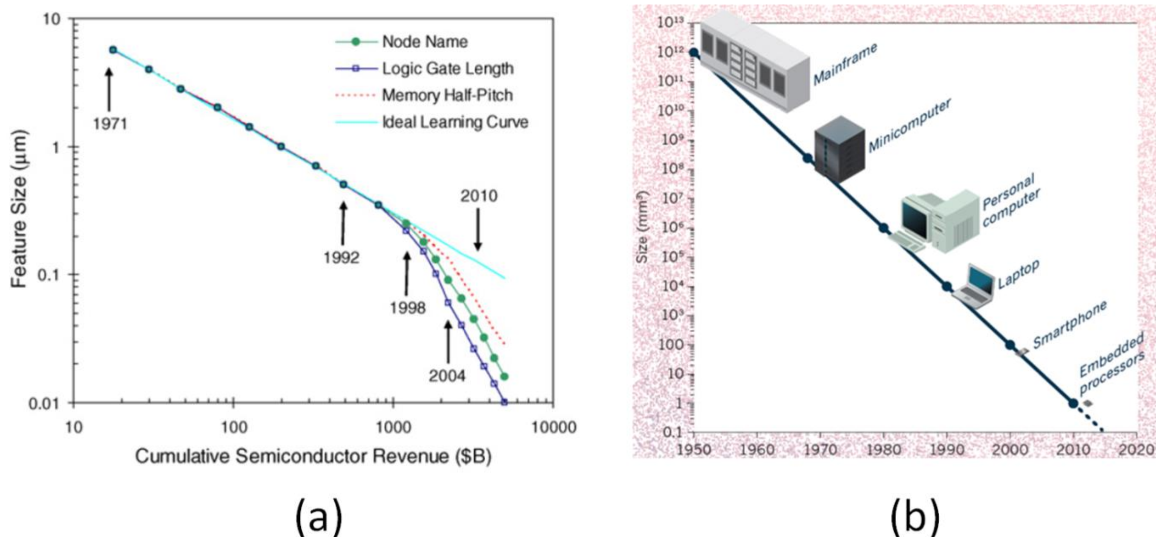


Fig. 1.1: (a) Cumulative revenue generated as a result of continuous scaling of feature size of semiconductor chips [3] (b) The end products of scaling at different points of time depicting the shift in high volume markets opened up as a result of Moore scaling [4]

Multiple systems have evolved over the past years and have grown in complexity and have shrunk drastically in size. To appreciate the complexity of today's technology, it is worthwhile to notice that today's smartphone packs more computing capability than a computer aboard the NASA Orion spaceship launched in 2014 [5]. Today's systems not only pack more computing power but are increasingly connected to the cloud.

It is instructive to look at what has driven the technological developments so far in terms of shifts in computing paradigms. As seen in Fig. 1.2, the internet revolutionized the application spectrum of connected devices. The sustained

growth of internet and internet-enabled devices has paved way for the latest gamut of applications which is commonly referred to as the Internet of Things (IoT). The purpose of IoT is the interconnection of multiple devices on the internet which sense their environments and communicate with other devices. Smart solutions can be arrived at by processing the large amounts of data generated by these devices. A large number of applications are envisioned such as efficient homes [6], smart cities [7], industrial productivity [8], and better healthcare [9] to name a few. The IoT revolution is predicted to put a staggering 212 billion devices online creating a $8.9 trillion market by the year 2020 [10].

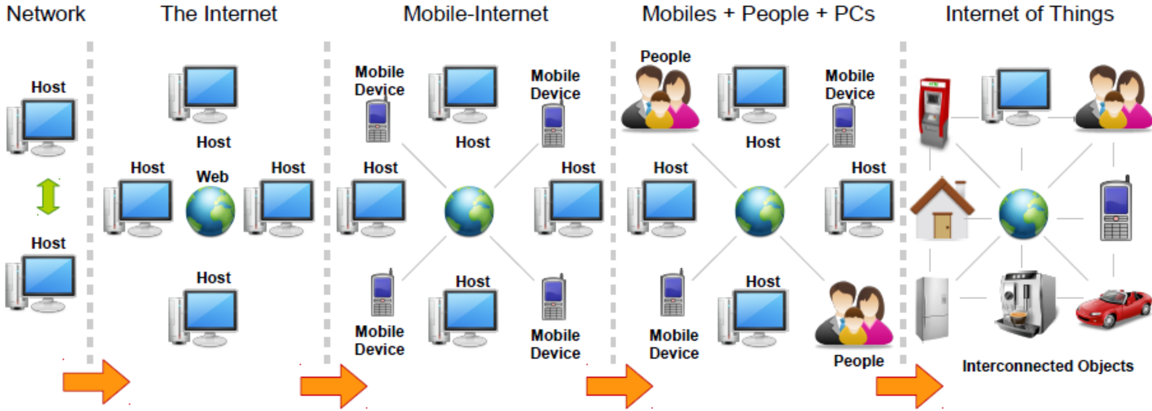


Fig. 1.2: The emergence of different paradigms of connected devices: Beginning from local area networks (LAN) connecting two or more desktop stations in a localized geographical area, the internet was born connecting computers worldwide. Advent of social media, entertainment and communication platforms enabled massive entry of people on the web. The next phase that is emerging rapidly is the Internet of Things [11].

In order to keep pace with the rapidly growing consumer market, the semiconductor industry has been a hub of multiple innovations. While Moore's law has been a consistent driver for faster operation of transistors coupled with reduced chip area and cost per function, it is no longer an adequate strategy by itself to meet the increasing demands of electronic systems. The integration of different system components becomes equally as important as the improvement of each of the individual component. This requires development of methods for seamless integration of multiple system components like displays, sensors, power controllers, analog/RF modules onto a single device. The collective efforts focused in this direction can be generalized as *More than Moore*.

Application requirements and technological innovations over the past years have driven the system design from mostly a System-on-Board (SoB) approach to System-on-Chip (SoC) and System-in-Package (SiP) approach as seen in Fig. 1.3 [12]. This approach has gained popularity since the early 2000s. In SoC based

design, the focus is on the integration of many different functional blocks on the same substrate which is optimized for power, performance and area (PPA). In SiP based design, the whole system is embedded in a single package containing multiple modules communicating and supporting each other. The additional advantage of SiP approach is that it allows integration of heterogeneous technologies which is not possible in SoC. These strategies are meant to improve the overall performance of the device. Improvements in overall system performance necessitate innovations in both SoC and SiP, which are discussed in subsequent sections.

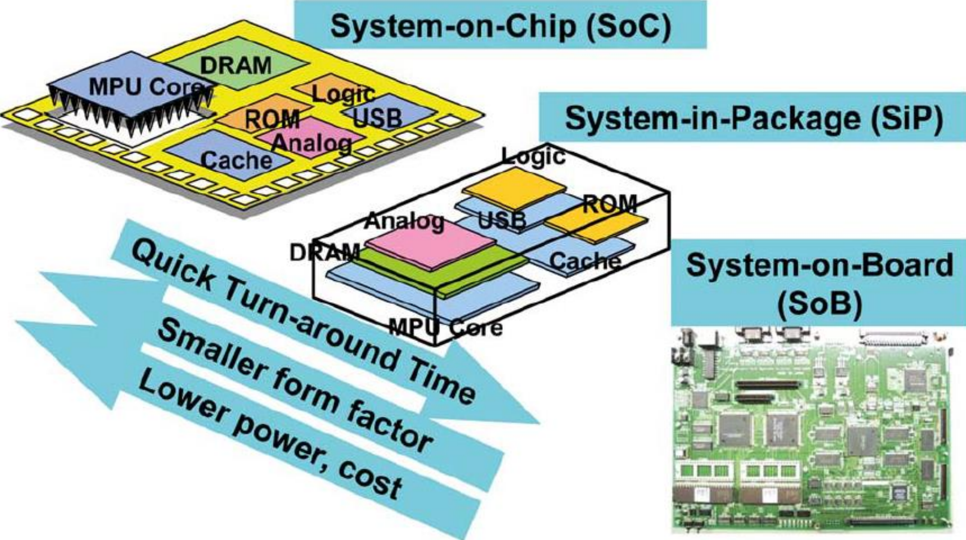


Fig. 1.3: Schematic comparing the different methods of system integration [12]

Thus far, the general trends of semiconductor industry with regards to the evolution of the consumer market and focus areas of technological advancement have been highlighted. *More than Moore* paradigm also has been briefly introduced. In the next section, the latter aspect is elaborated in more detail.

1.2 *More than Moore*: Introduction and recent advances

More than Moore refers to a broad spectrum of technological focus areas which supplements the classical Moore scaling to enable high-performance and highly-integrated systems. In addition to high-end applications, *More than Moore* addresses application areas such as IoT where there are several system constraints such as energy efficiency, low cost, self-powered, high volume etc.

In this section, the major developments in *More than Moore* are illustrated. More specifically, the different packaging methodologies are briefly discussed.

1.2.1 Motivation

The first motivation for *More than Moore* is efficient management of system complexity. The increasing complexity of systems needs to be managed on a holistic perspective rather than individualistic perspective of each component comprising the system. A good illustration of this idea is made in Fig. 1.4 where the analogy of a city is taken for a system with different components representing different units of a smart city. For a smart city, the individual units like storage, processing and external interfaces need to be optimized. In order to further improve the functioning of the city, its layout needs to be optimal. The planned layout should address the communication requirements between the different units. Units which are communicating more often with each other must be placed together for efficient operation. Also, smart ways must be implemented for the smooth flow of resources within the city.

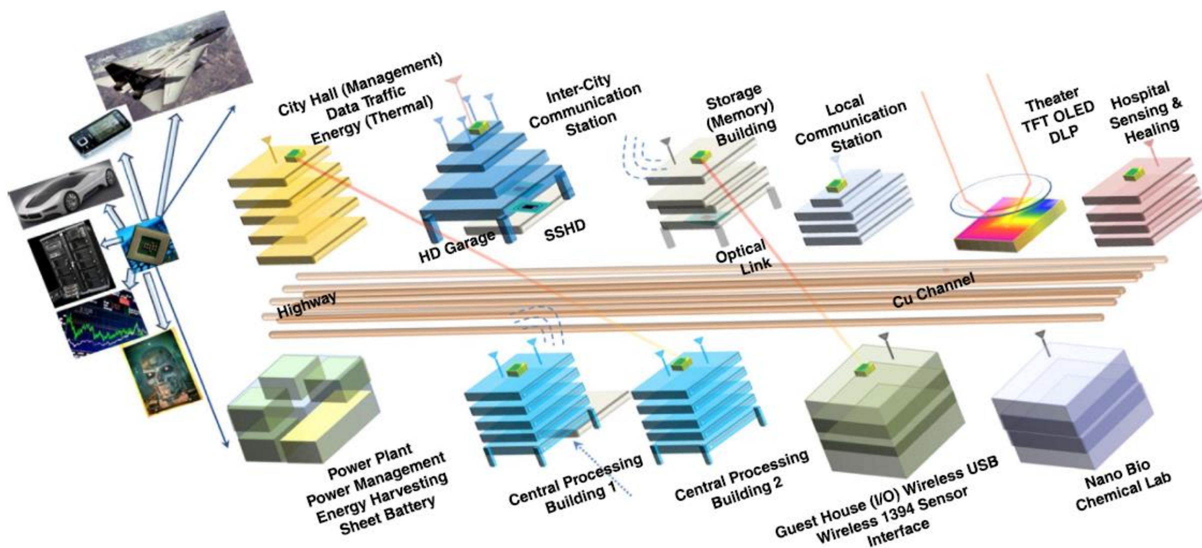


Fig. 1.4: A real-world analogy of a system which is compared to a city. Just as different bodies of the city need to interlink and communicate with each other, an electronic system needs to optimize its functioning between the interconnected components for efficiency. [13]

The second motivation for *More than Moore* is enhanced system diversity. For this, there needs to be sustained development of RF, analog, MEMS, sensors, energy harvesting and biosensing components on cost-efficient and robust technologies. This is essential for the upcoming generation of highly diversified smartphone and IoT devices. For specific application areas, one technology is dominant over the others. For instance, CMOS is dominant for memory and logic,

III-V materials for high performance RF and integrated photonics etc. Hence, efficient packaging solutions are vital for a diversified system composed of different kinds of components.

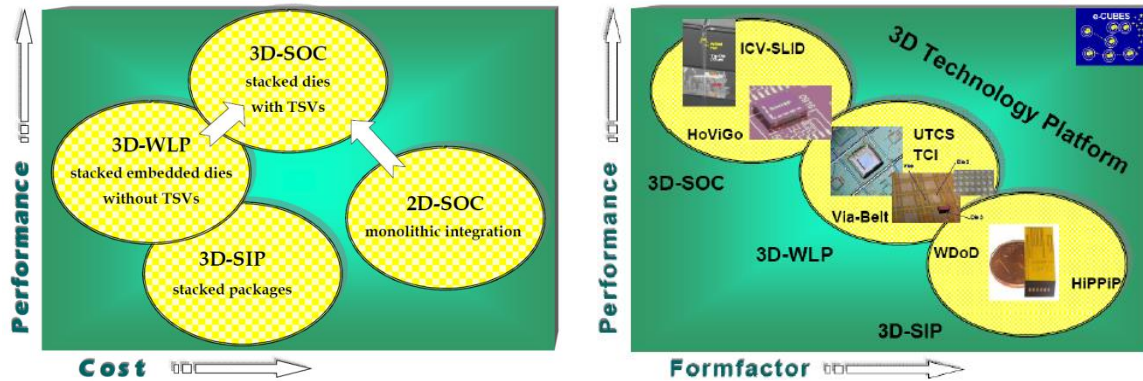


Fig. 1.5: A comparison of different integration schemes and the tradeoffs obtained between cost, performance and form factor for each type of technology

In order to address these two requirements of systems, several strategies have been adopted by the industry. They are briefly listed in Fig. 1.5. 2D-SOC scaling strategy involves inclusion of multiple functionalities on the same substrate which were traditionally fabricated using different platforms. 3D integration technologies like System in Package (SiP), Wafer-Level Package (WLP), and 3D-SoC helps better integration of multiple dies to improve performance and form factor. All these methods are elaborated in the coming sections. All of these packaging methods are relevant since the choice of method relies ultimately on the PPA and cost requirements of the system which can vary.

1.2.2 SoC integration: Die level optimization

SoC integration (monolithic integration) is the method of packing more components on the same silicon die. This is made possible by tailoring of fabrication process so that it is compatible for fabricating different types of devices to support extended functionalities. This is illustrated with the example of a commonly used fabrication process called BCD (Bipolar-CMOS-DMOS). This process is used in the fabrication of monolithic power ICs.

Consider a smartphone or any other portable electronic device. It is normally powered by a battery. The battery supply voltage can vary significantly depending on the battery technology as shown in Fig. 1.6. The power from the battery is delivered at the required voltage to the chip by using a power IC. Even on the chip, multiple voltage domains exist for digital, analog and general I/Os which also changes with each technology generation.

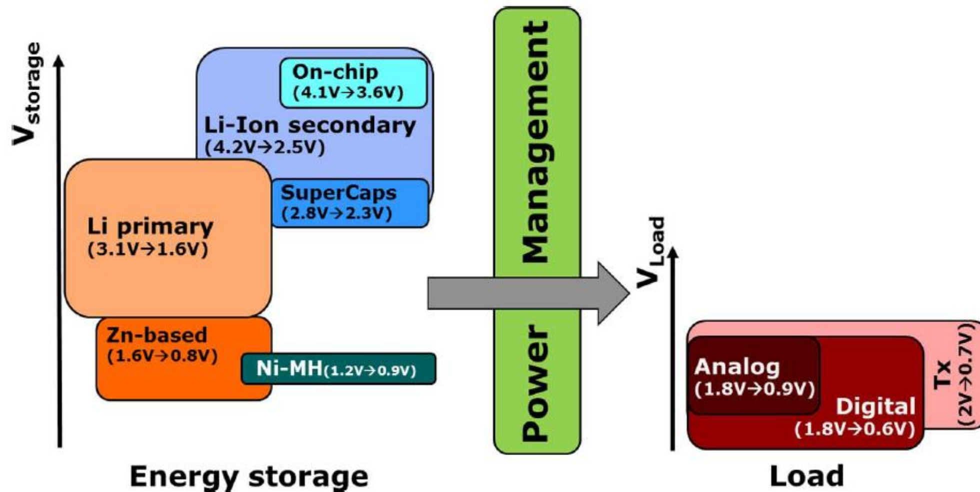


Fig. 1.6: Power management scheme on a portable electronic device depicting the typical range of voltages on the storage side and the load side [14]

Power management can be effected in multiple ways and a common choice is the use of switched-mode power supplies (DC/DC converters). The implementation of such converters consists of high power switches, reactive components and a controller. For power switches, CMOS or DMOS is used depending on the voltage requirement. For controller module, low power CMOS is preferred while bipolar transistors are used in the controller and circuit protection. CMOS, DMOS and Bipolar are traditionally different process technologies and if one has to implement the same power controller using separate process technologies, the integration would have to be done at the level of the package. By offering the combined processing capabilities in the BCD process, it is possible to integrate all these modules on the same die. An example of a chip realized using BCD process incorporating different functionalities is shown in Fig. 1.7.

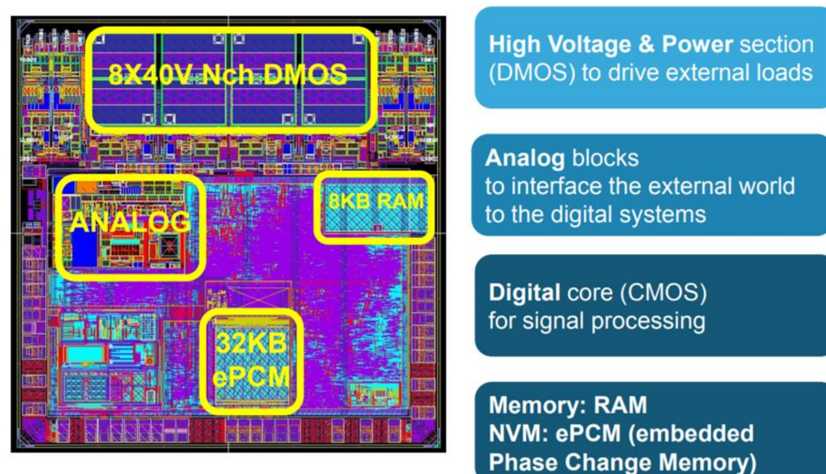


Fig. 1.7: Chip realized on BCD platform containing digital, analog, memory and High Voltage blocks [15]

1.2.3 Packaging: System level optimization

The advancement of packaging technologies have given rise to a host of schemes as to how each individual die within the package are integrated. There are packaging solutions both on the die-level and wafer level. Some of the prominent packaging methods on the die level are illustrated in Fig. 1.8 [16]. These are the earlier versions of use of 3D space for packaging. The 3 methods of packaging shown here are:

1. Single package SiP: This is the simplest concept of package where the different components of the system are integrated onto a single package which acts like a miniature circuit board. A typical example is that of a RF WLAN package where the different blocks are integrated on a package without the need for external passive components [17]

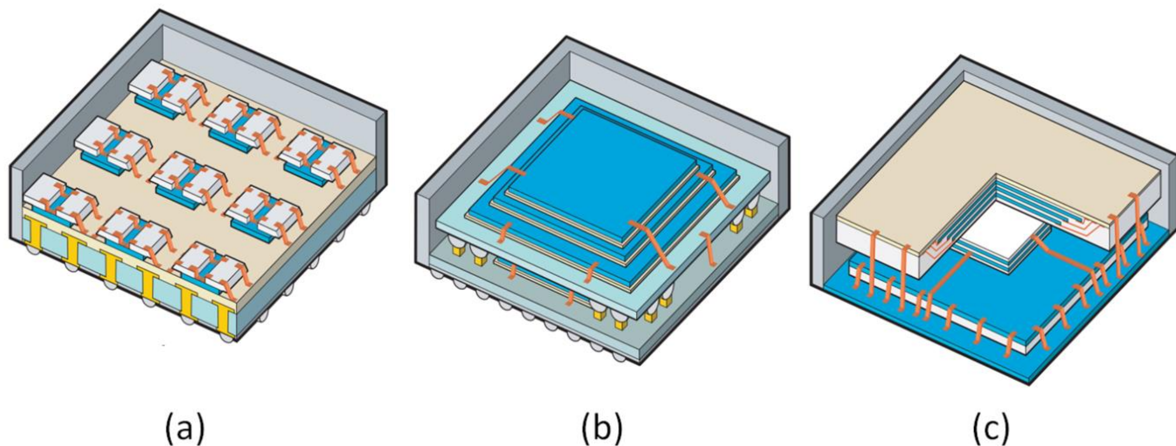


Fig. 1.8: Different packaging integration schemes (a) SiP (b) PoP (c) PiP [16]

2. Package-on-Package (PoP): This is a 3D packaging scheme to integrate already packaged modules onto a single package. The individual packages within the PoP package are stacked one on top of the other. A typical example of PoP package consists of a memory package stacked on top of logic package. The bottom side of the logic package has a high density pin count for connection onto the substrate and the top side contains land pads for logic at the periphery. The memory package again has a solder bumps matching the land pads on the logic package. To fix the package in place, the assembly is placed on the board and the solder reflow from bottom package to the board and the top package to the bottom package happens simultaneously resulting in a package on package structure. The same principle can be extended for stacking multiple packages.

3. Package-in-Package (PiP): This method is also 3D and involves stacking packages one on top of another like PoP but there is one major difference here. Considering the same example of memory and logic, in the PiP method, the memory package is stacked and molded onto the base package to obtain an Internal Stacking Module (ISM) which can be tested before final packaging [18]. This provides additional testability before final integration onto the package substrate. After testing, I/O pins of the logic package are connected to the substrate by either wire bonding or flip-chip. Similarly, the I/O connections of the memory are routed to the substrate by making use of the wire bonds from the top side. The final package is obtained by making use of an Encapsulant Mold Compound (EMC) to encapsulate the entire stack containing the two dies. The final product resembles a traditional FBGA package.

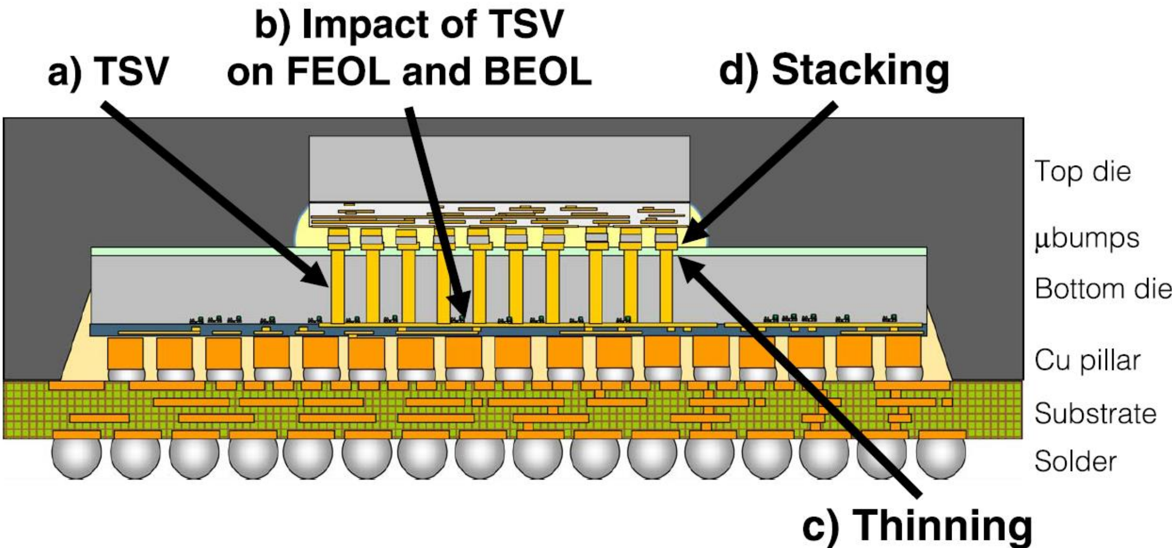


Fig. 1.9: 3D stacking of two dies using Through-Silicon Vias (TSVs) for chip to chip interconnections. The four arrows point to the locations where the reliability concerns emerge in such a package. The final package is ready for board level integration using solder bumps connected to the substrate. [19]

While, these examples illustrate evolution of packaging from a 2D regime to 3D regime using a combination of flip chip and wire bonding, there has been an additional driver for improvement of 3D integration: The use of Copper Through-Silicon Vias (TSVs). Fig. 1.9 illustrates a packaging concept based on the use of TSVs. The narrow TSVs made in the bottom die penetrate the entire thickness of the die and takes the BEOL interconnections on its back side. There are microbumps on the front side of the top die which is inverted. The microbumps of the two dies are aligned and attached together using different bonding methods. Fig. 1.9 also shows the important regions of reliability concern for copper TSV

implementation on silicon. Based on the hierarchy of the TSV interconnections, different 3D integration schemes are used like 3D-SoC, 3D-SiP and 3D-IC. A review of these methods has been made by Beyne [20].

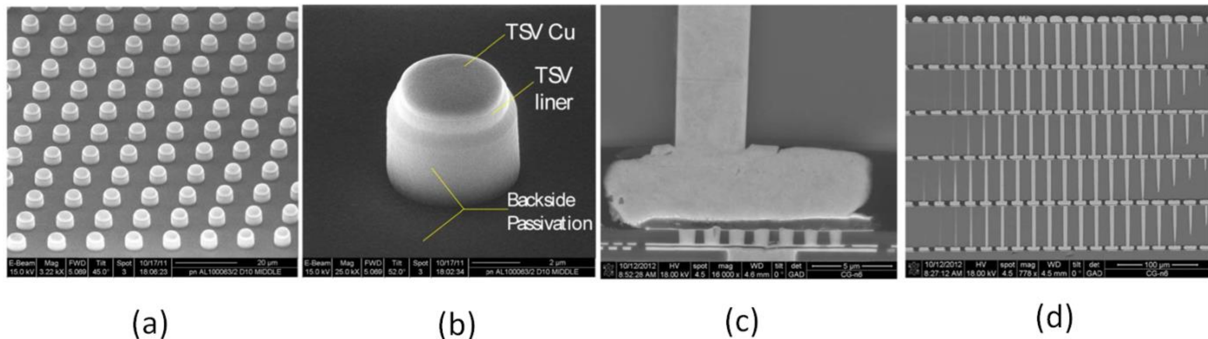


Fig. 1.10: (a) Array of TSVs realized using via-middle process flow (b) Close-up view of via (c) Cu interconnection by Cu-Cu TCB bonding (d) Cross section view of a 5 die stack [21]

Via-middle Copper TSV process flow is a commonly adopted method for 3D integration by most semiconductor companies [21]. It has been established that scaling of the TSV diameter is important for thermo-mechanical reliability. Using the via-middle approach, narrow via diameter of 2 μm has already been demonstrated as seen in Fig. 1.10. This reduces greatly the Keep-Out Zone (KOZ) which is the dead zone surrounding the TSV where no active device can be present. An example of 5 dies vertically stacked using this method (Fig. 1.10d) reveals the high interconnect density possible with use of TSV.

Packaging methods are continuously evolving and novel methods are emerging. Besides, 2D and 3D there is a new packaging technology methodology called 2.5D packaging which was originally developed by Xilinx. It makes use of the technology improvements seen in both 2D and 3D packaging. In this technique, the different dies which need to be integrated are connected by means of an interposer layer. The most common type of interposer is the passive silicon interposer. The example shown in Fig. 1.11 contains several FPGA dies which are interconnected using the passive silicon interposer. The silicon interposer contains 4 layers of metallization and TSVs which connect the different dies together. The advantage of this method is that the interposer layer itself does not contain any active devices and hence the reliability issues caused by the presence of TSVs on active devices are eliminated.

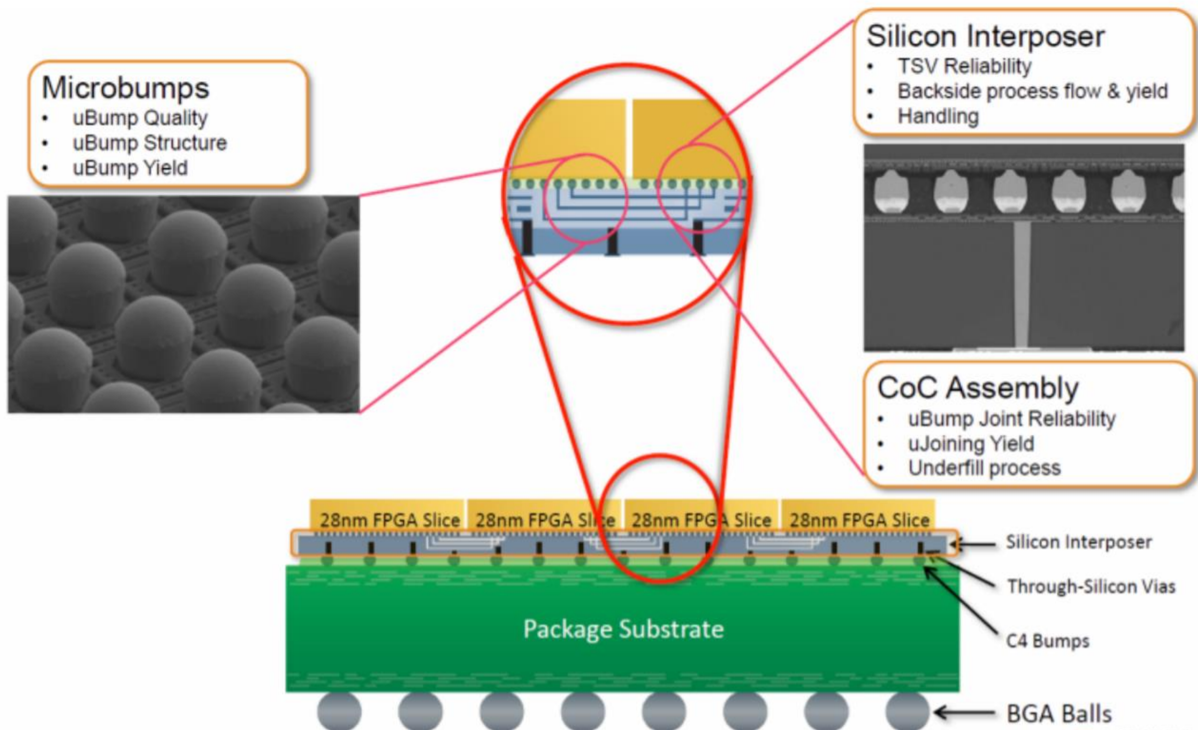
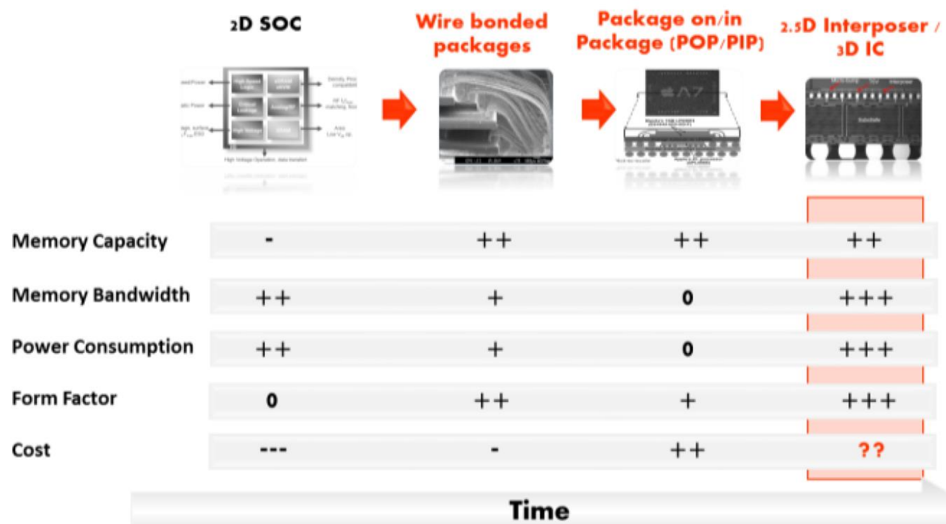


Fig. 1.11: 2.5D packaging concept containing a silicon interposer for interconnection of dies. The demonstrated package here is a FPGA product developed by Xilinx [22]

Finally, a comparison is made between the different packaging technologies as shown in Fig. 1.12. There is a trade-off between cost and obtained performance for each type of package. With further technological breakthroughs and volume production, the cost will go down enabling democratization of high-performance systems of the future. In the next section, the relevance of laser processing technology to *More than Moore* is discussed.



	2D SOC	Wire bonded packages	Package on/in Package (POP/PIP)	2.5D Interposer / 3D IC
Memory Capacity	-	++	++	++
Memory Bandwidth	++	+	0	+++
Power Consumption	++	+	0	+++
Form Factor	0	++	+	+++
Cost	---	-	++	??

Fig. 1.12: Comparison of different packaging technologies in terms of system performance metrics. [23]

1.3 Laser processing: An emerging driver for *More than Moore*

1.3.1 The evolution of femtosecond laser

The application of high power laser for conventional applications such as industrial drilling, welding and cutting has been know for many decades now. The entry of pulsed laser decades has provided more processing flexibility of lasers [24]. Pulse duration is one of the most important parameter of pulsed lasers, which decides the quality of laser machining (detailed in section 3.2). Table 1.1 shows a comparison of different types of lasers and corresponding pulse widths.

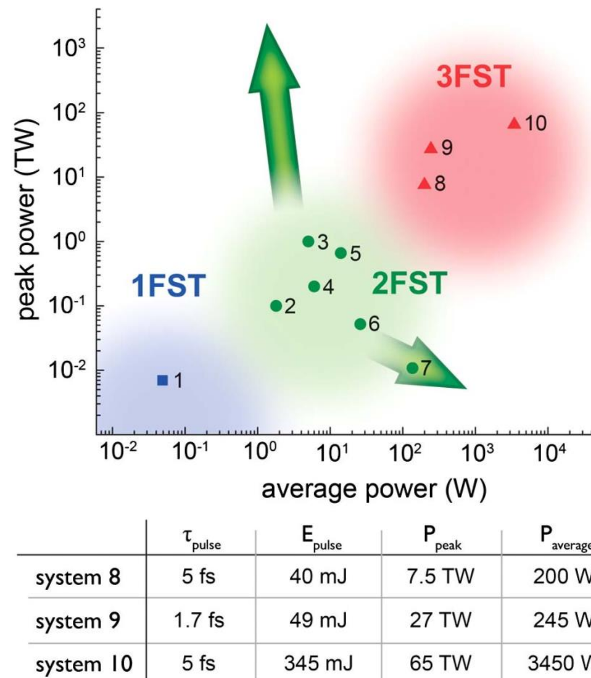
Table 1.1: Comparison of different laser technologies and their relative capabilities [25]

Laser type	Free running Nd:YAG	Q-switched Nd:YAG	CPA Ti:sapphire
Pulse width	1 ms	100 ns	100 fs
Intensity (W/cm^2)	10^6	10^8	10^{15}
Average Power (W)	100-1000	100-1000	1-10
Machining quality	Poor	Good	Excellent
Throughput/Speed	High	High	Low
Cost	Low	Medium	High

There has been tremendous interest especially in femtosecond lasers because of their excellent machining capabilities. The fs pulse length enables a whole new domain of laser material interactions which allows processing of different kinds of materials like metals, semiconductors, ceramics, dielectrics etc. The first femtosecond laser was developed in the 1970s and it had a poor peak power and average power output. Three generations of femtosecond lasers can be distinguished based on common architectural elements in the design. With each successive generation, the average and peak power output of such lasers has improved drastically as seen in Fig. 1.13.

Lower pulse width is required for not only improved machining quality but also for reduction of the heat affected zone which gives greater spatial control over the region of beam-material interaction [26]. Higher average and peak power are required to improve throughput. The average power output of fs lasers has exceeded the kW mark as of today [27]. In order to make use of such high power lasers for microstructuring applications, it is also crucial to develop optics that is

able to efficiently guide and modulate the beam and deliver the laser power in a controllable fashion.



	τ_{pulse}	E_{pulse}	P_{peak}	$P_{average}$
system 8	5 fs	40 mJ	7.5 TW	200 W
system 9	1.7 fs	49 mJ	27 TW	245 W
system 10	5 fs	345 mJ	65 TW	3450 W

Fig. 1.13: Comparison of 3 generations of femtosecond laser systems. Systems 1-7 have already been demonstrated. The systems 8 – 10 were under development as of the publication date of the article [28]

1.3.2 Application of femtosecond laser processing in *More than Moore*

1.3.2.1 Packaging and integration

The earliest use of lasers in microelectronics packaging was in dicing of silicon wafers. When this was first introduced, the formation of microcracks was an issue which resulted in reliability concerns. However, with the improvement in laser technology, especially with smaller pulse width lasers, the laser dicing processes have improved greatly. In today's commercially available tools, the use of nanosecond lasers is popular and the dicing thicknesses are typically < 100 µm [29]. The nanosecond laser has reduced chipping compared to mechanical dicing but the die strength is reduced because of thermally introduced degradation. The die strength can be recovered by using additional dry/wet etching techniques to remove thermally affected regions of the die.

The use of femtosecond lasers is another alternative to improve significantly the die strength and other figures of merit like surface roughness [30]. The maximum dicing throughput for femtosecond lasers remains < 100 mm s^{-1}. There is much scope for improvement of this figure and this point will be discussed in further detail in section 2.1.3.

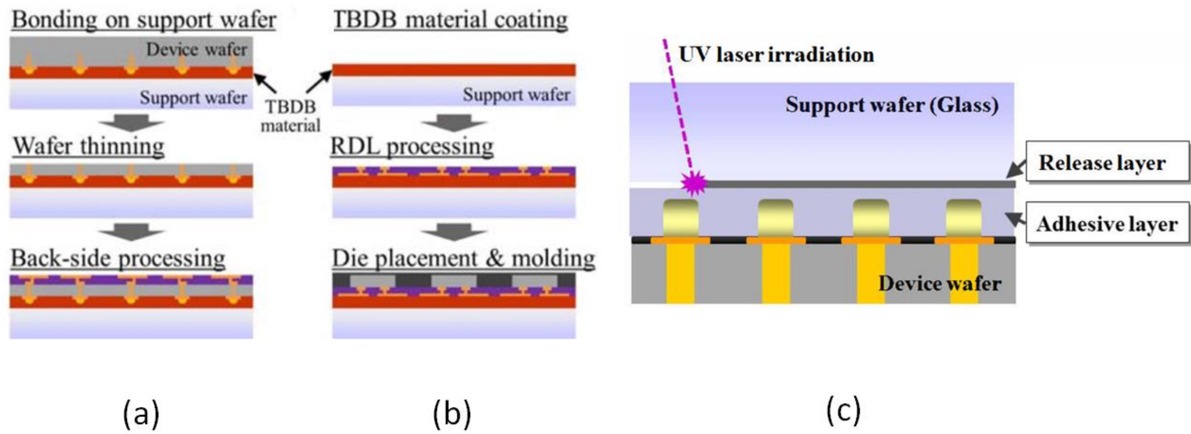


Fig. 1.14: Use of temporary bonding in FO-WLP (a) For backside processing of wafer (b) For die placement and packaging (c) The release process of temporary bonding material using laser irradiation [31]

The use of laser is also becoming popular in large-scale packaging applications like Fan-Out Wafer-Level Package (FO-WLP). FO-WLP process involves attachment of die/wafer onto a temporary carrier substrate for performing various steps like wafer thinning, backside processing, Redistribution Layer (RDL) processing etc. as shown in Fig. 1.14. The attachment is done by means of a release layer which can be removed by several methods like chemical, mechanical, thermal release etc. The use of lasers for removal of release layer has the potential for highest throughput (> 40 wafers per hour) as compared to other methods [31]. The use of pulsed excimer lasers is preferred in the industry because of their high pulse energy of the order 10s of mJ [32]. Femtosecond lasers can improve greatly the debonding process in terms of quality. For throughput, today's commercial femtosecond lasers still have energy in the µJ regime but evolution of 3rd generation femtosecond lasers can help augment the packaging market of femtosecond lasers.

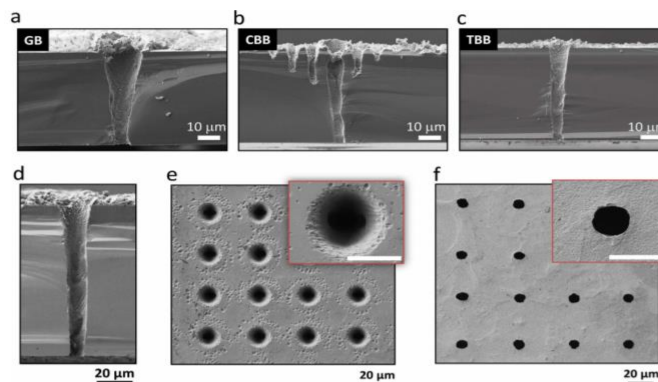


Fig. 1.15: The use of femtosecond laser for through silicon vias (a) Gaussian beam (b) Conventional Bessel beam (c) Tailored Bessel beam (d) Vias in 100 µm silicon (d) Front side of TSV array (b) Rear side of TSV array [33]

For 3D integration, the formation of TSVs of high reliability and high aspect ratio is important. In the work of He et al., the use of 780 nm femtosecond laser using 1.5 µm Bessel beam was demonstrated for the drilling of high quality TSV without any side-lobe damages [33]. As seen in Fig. 1.15, by making use of tailored Bessel beam, TSVs of diameter 10 µm is formed on a wafer of thickness 100 µm. The unique ability of femtosecond lasers to process all types of materials makes it a suitable candidate to drill vias in many packaging applications.

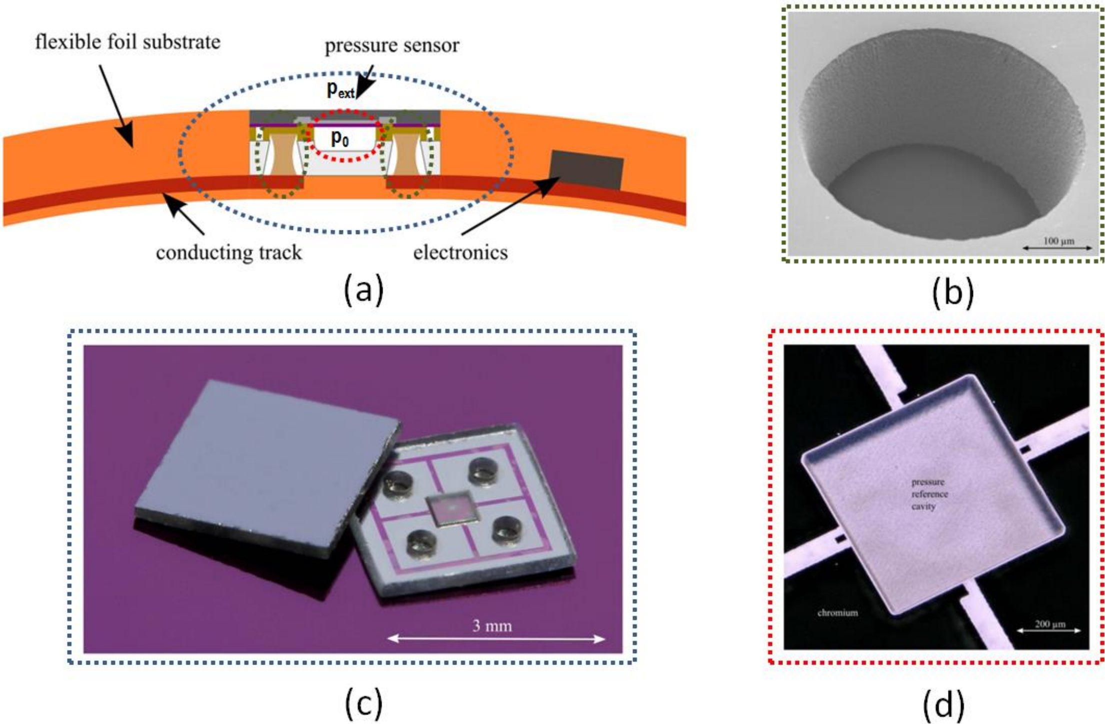


Fig. 1.16: Surface passive pressure sensor integrated on a thin film foil (a) Finished package (b) TGV for sensor electrical connection (c) Bonded sensor front and back view (d) Cavity in glass to create chamber for pressure reference [34]

Through-Glass Vias (TGVs) are becoming more relevant as several integration schemes like glass interposers are gaining popularity. Also, laser processing can potentially be a key component of both device fabrication and packaging in many applications. Fig. 1.16 shows a pressure sensor for aerospace application implemented using laser processing technology. It consists of a cavity for providing reference pressure (p_0). The sensing elements are piezoelectric membranes enclosed within this chamber connected electrically as a Wheatstone bridge. The electrical connections are routed to the package by means of a through glass vias and the final package is a thin flexible foil.

1.3.2.2 MEMS

Micro electromechanical systems are common in today's electronic products primarily as sensing elements. Fig. 1.17 shows the realization of diaphragms in 4H-SiC substrates of thickness 250 μm for pressure sensing applications. The appropriate choice of laser processing parameters can improve greatly the smoothness of the membrane. The use of femtosecond laser in this work greatly improves the reliability of the membrane because of reduced heat affected zone. The xyz spatial resolution is also very high with fast removal rates.

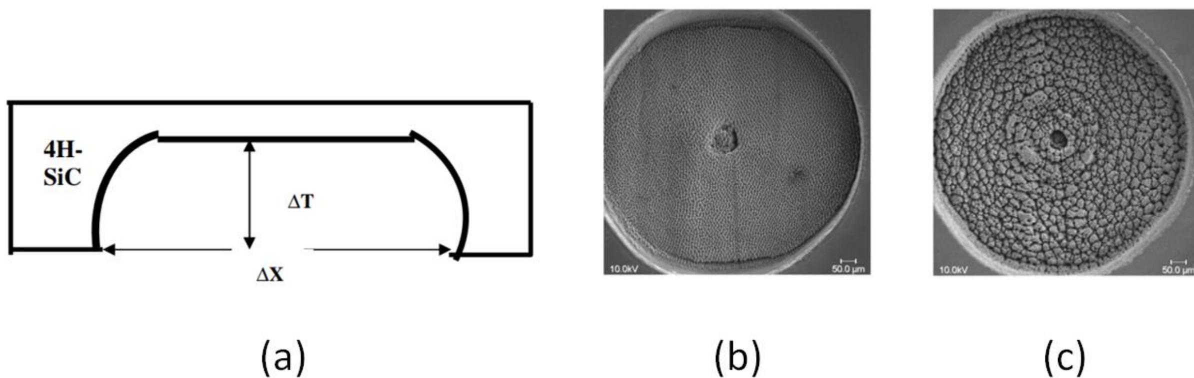


Fig. 1.17: (a) Diaphragm in 4H-SiC substrate (b,c) SEM image of diaphragm milled with pulse energy 0.05 mJ and 0.15 mJ [35]

Complex geometries can be realized in silicon to fabricate micromanipulators. These micromanipulators can be used to perform many functions. For instance, assembly of micromechanical parts like microgears and even handling of delicate test specimen. One common implementation of micromanipulator is the comb drive where an electrical stimulus is used for the actuation of the manipulator. Fig. 1.18 shows the comb drive implementation of MEMS micromanipulator on 3C-SiC thin film on top of silicon substrate. This film being difficult to chemically etch is easily removed with the use of femtosecond laser. After laser patterning, KOH etch is performed to release the structure. In this work, by using pulse energy of 1 μJ, the quality of the device is optimized.

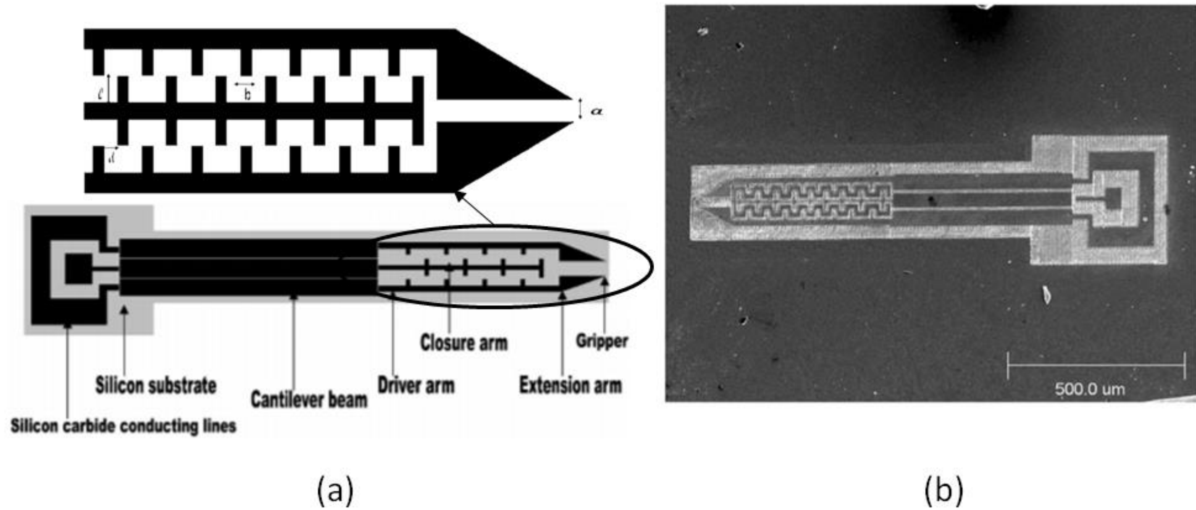


Fig. 1.18: (a) Dimensions of the comb drive: a=20 μm, b=10 μm, l=21 μm, d=15 μm (b) SEM micrograph of machined microgripper [36]

Laser ablation is also very suitable for making of molds in different kinds of materials suitable for fabrication of structures using techniques like micromolding and soft lithography. Fig. 1.19a depicts a mold made out of silicon using laser ablation. With the use of high pulse energy of 240 μJ, the fabrication of the mold in silicon wafer of thickness 500 μm takes only ~1 s. Fig. 1.19b shows how the molds can be used to fabricate different kinds of structures using the micromolding technique. In this technique, polymeric precursors are filled in the mold. Carbonization of the precursors results in densification and shrinkage of the liquid polymer into solid glassy carbon which function as MEMS structures.

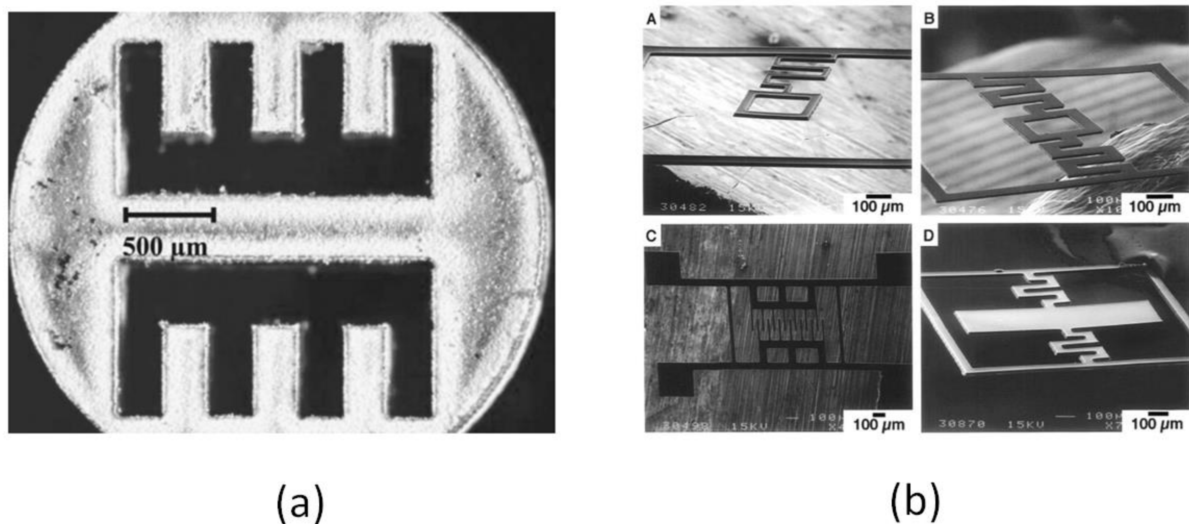


Fig. 1.19: (a) MEMS mold fabricated in silicon (b) MEMS structures realized using micromolding of polymeric precursors [37], [38]

Cantilevers are commonly used components in MEMS designs. Fig. 1.20 depicts 3 different types of cantilever machined in silicon: rectangular, T-shaped and triangular. Cantilevers are fundamental sensing elements in different types of chemical, biological and mechanical sensors. The sensing principle involves the detection of mechanical bending or the frequency of oscillation of the cantilevers. The sensing properties are a strong function of the cantilever geometry which affects the stiffness.

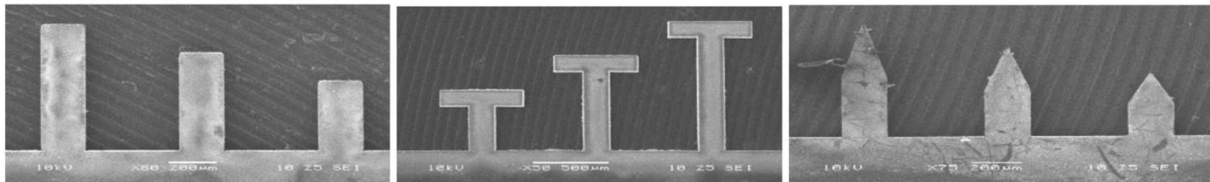


Fig. 1.20: Micromachined MEMS cantilevers in silicon [39]

1.3.2.3 Microfluidics

Laser micromachining is a very suitable method for fabrication of microfluidic devices/functions on different substrates. The suitability of laser processing for photonic structures enables seamless co-design and fabrication of optical and microfluidic channels for Lab-on-Chip (LoC) applications.

Fig. 1.21 shows two different types of filters implemented in microfluidic channels to filter out larger particles. In the first filter, the laser micromachined grids shows excellent patterning resolution of femtosecond laser close to submicron regime. Two-photon polymerization is used to fabricate the filters. The filter is implemented on prefabricated channels demonstrating the flexibility offered by laser processing for easy integration of features on existing design.

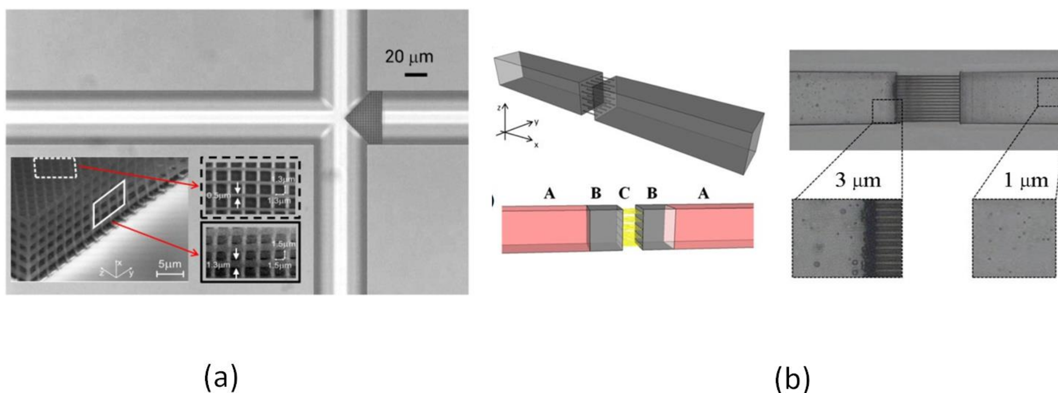


Fig. 1.21: (a) Micron sized particle filter using two photon polymerization (b) Filter matrix along the microfluidic channel in fused silica using selective etching of femtosecond laser irradiated areas in HF and KOH [40], [41]

In the second example, fused silica is patterned using femtosecond laser and the regions exposed to laser are etched away by wet chemistry. This method is in contrast to the first example, where the laser exposed areas are retained after development because polymerization occurs in these areas.

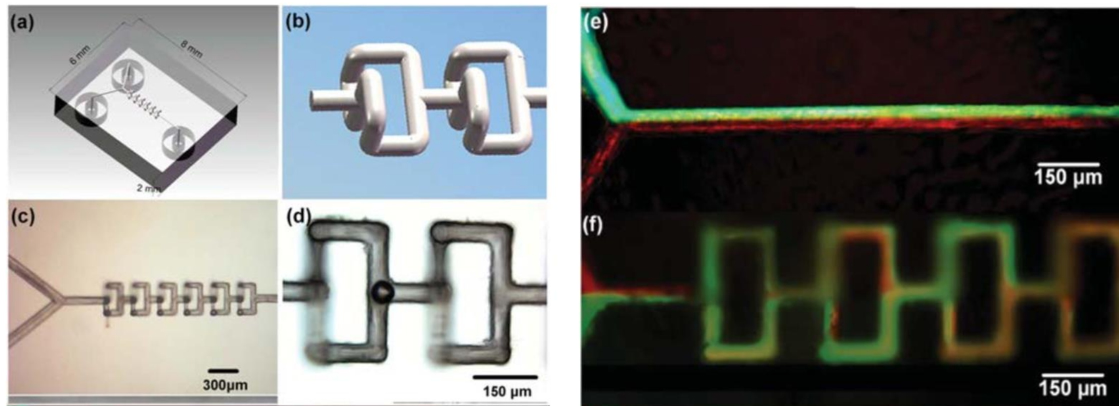


Fig. 1.22: (a,b) Microfluidic mixer schematic in fused silica (c,d) Microscope images of inscribed mixing channels (e,f) Fluorescence microscopy to quantify mixing [42]

A similar example of microfluidic mixing function realized on fused silica is shown in Fig. 1.22. 3D channel geometries help the efficient mixing of fluids as it passes along the length of the mixer. The mixing process can be visualized by introducing unmixed fluorescent green and red fluids and by capturing images of the liquids using fluorescence microscopy. It can be seen that fluids remain unmixed in a straight 1D channel while the 3D passive mixer achieves a good mixing between the fluids.

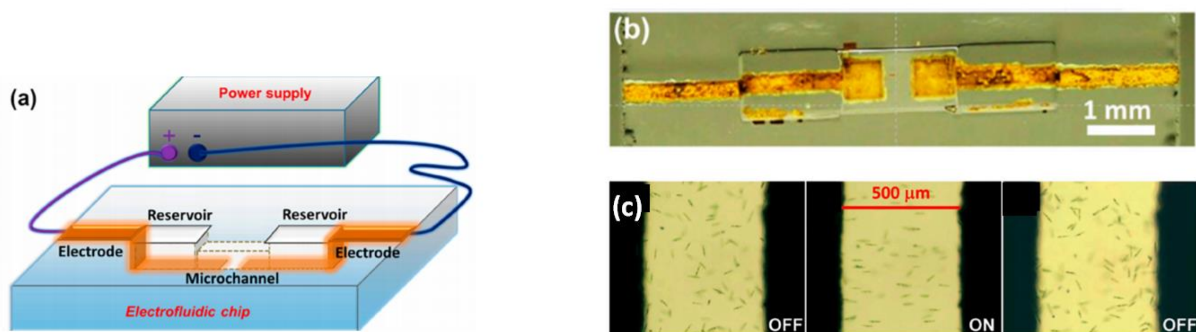


Fig. 1.23: (a) Electrofluidic chip schematic with external supply (b) Fabricated electrofluidic chip containing two electrodes which are biased using external supply (c) The alignment of Euglena affected in the presence of external field. Turning off the field restores the random orientation of cells [43]

Another interesting application is the use of electrostatic field to microfluidic application referred to as electrofluidics. By the application of electric field,

charged species can be deflected/oriented in the desired direction. Fig.1.23 depicts an electrofluidic channel which consists of a microfluidic channel surrounded on two sides by electrodes. When a potential is applied across these electrodes, an electric field is setup in this region. As seen in Fig. 1.23c, by application of an electric field, the euglena cells in the channel can be aligned perpendicular to the length of the channel.

1.3.2.4 Integrated photonics

Photonics is one of the main application areas envisioned for femtosecond laser micromachining. Integrated photonics is a topic of growing interest for short range applications like internet switches, high-data rate servers and supercomputers because of the limitations imposed by copper links [44]. By the use of active optical cables, high data rates of 120 Gb/s can be realized in such systems. But cost factor remains a challenge for such systems. Femtosecond laser could be a potential solution to high volume manufacturing because of high throughput capabilities and simple fabrication steps.

A variety of integrated photonic elements have been realized using laser processing. Binary Fresnel Lenses (BFLs) have been fabricated on the surface of PMMA as shown in Fig. 1.24a. These lenses are used in optofluidic applications where high Numerical Aperture (NA) lenses are needed to focalize light for excitation or detection. In this work, using highly focussed beam, a line width of 600 nm has been achieved which results in a high NA of 0.5.

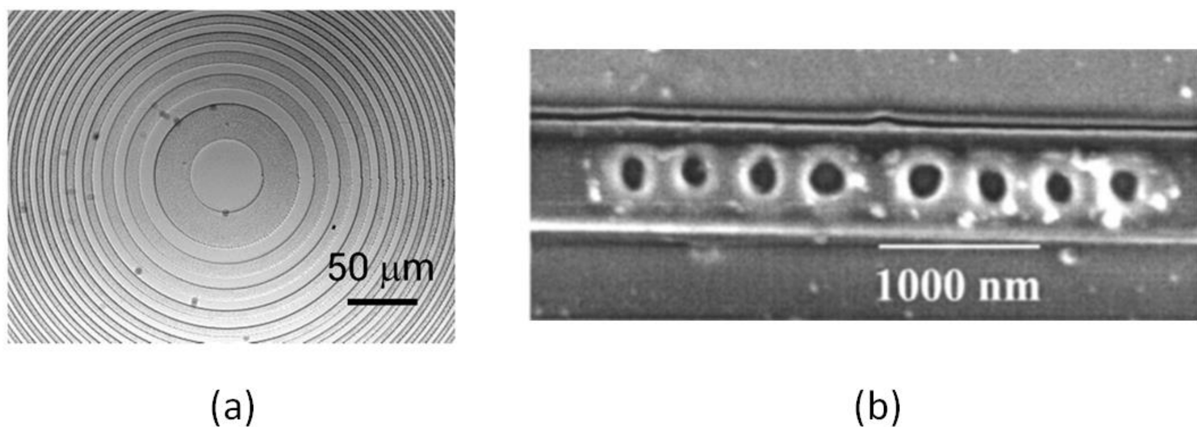


Fig. 1.24: (a) BFL fabricated on PMMA (b) 1D photonic crystal on sapphire [45], [46]

1D photonic crystal has been fabricated on Sapphire (Fig. 1.24b) which acts as band pass filter for 1550 nm with a bandwidth of ~3 nm. A high machining accuracy of ~15 nm has been reported in this study.

In addition to surface structures, femtosecond lasers are suited for 3D micromachining to obtain complex geometries. The fabrication of High-Q resonators in fused silica (Fig. 1.25) serves as a good example of 3D fabrication using a combination of laser and chemical etch. Such microresonators have been successfully integrated in microfluidic mixer applications as a refractive index sensor with a high sensitivity of 1.2 x 10^{-4} RIU (Refractive Index Unit).

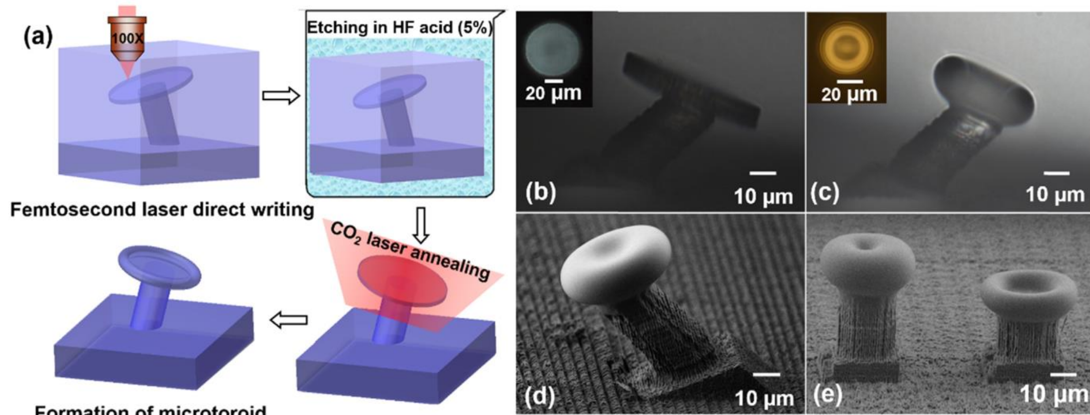


Fig. 1.25: Fabrication of 3D high-Q resonator in fused silica with SEM images (a) Sequence of steps (b) After laser direct write and HF etch (c) After laser annealing (d) Resonator with tilted angle (e) Vertical resonator with different heights [47]

While laser based fabrication methods are used for fabrication of photonic circuits, lasers can also be used as an auxiliary tool for enhancing existing fabrication techniques. Microring resonators fabricated on standard CMOS process (Fig. 1.26) are vulnerable to fabrication errors [48]. These fabrication errors lead to phase errors and sub-optimal performance of the resonator. By making use of femtosecond lasers, post fabrication amorphization of silicon can be performed to fine tune the microresonator to have the desired resonant behaviour. Fig. 1.26c shows the shift in resonance wavelength for the microresonator which is induced by making use of femtosecond laser processing.

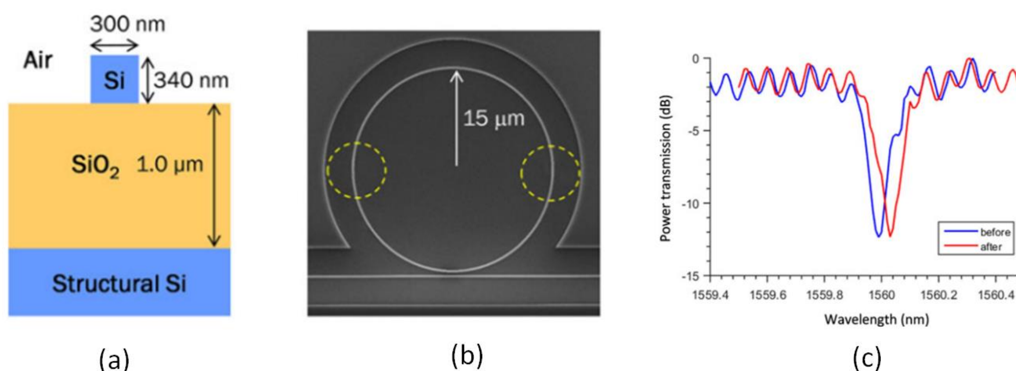


Fig. 1.26: (a) Cross-sectional view of microring resonator (b) SEM image of the resonator (c) Transmission characteristics of the resonator [48]

Femtosecond laser is also suitable for large area patterning for optical applications. As seen Fig. 1.27, large array of concave lenses has been fabricated on silicon film of area 1 cm x 1 cm by using laser processing followed by wet etching. The concave microlenses shown here have a diameter of 20 – 30 µm. The radius of curvature of the lens is tuned by varying the laser scanning speed. Such arrays can be used as reflective optical homogenizers.

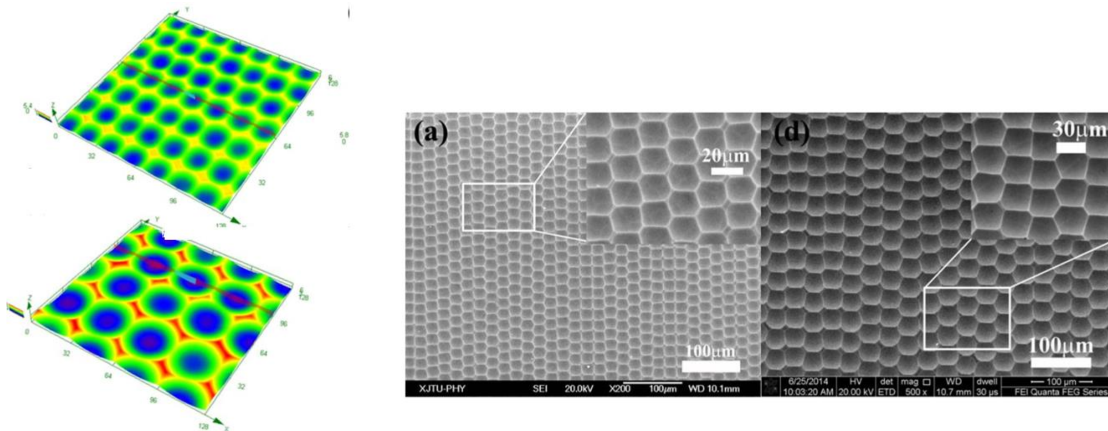


Fig. 1.27: Large area microlens array fabricated on thin film silicon with average diameter (a) 20 µm (b) 30 µm with corresponding depth profiles shown on the left[49]

1.4 SOI Technology: A platform for SoC integration

SOI technology has been in the mainstream semiconductor industry around for nearly 4 decades after IBM first started investigating this technology since 1989 [50]. The active device layer is isolated from the bulk (also known as handler) by a thin insulating silicon oxide layer known commonly as Buried Oxide (BOX). The insulation provided by the BOX results in reduced substrate parasitics. This results in a performance gain of 20 – 35% for SOI CMOS as compared to bulk CMOS node [50]. The commercialization of substrate technology and continuous improvements in process has led to different generations of SOI CMOS technology. Two variants of SOI technology are commonly identified: Partially Depleted SOI (PD-SOI) and Fully Depleted SOI (FD-SOI).

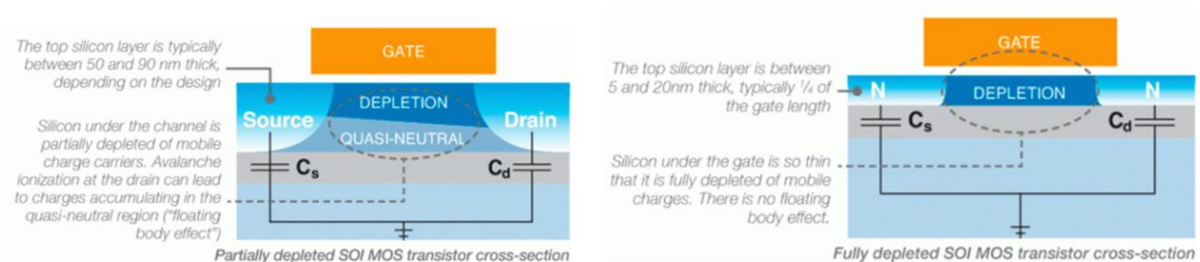


Fig. 1.28: Transistor configurations for PD-SOI and FD-SOI [51]

The transistor configuration in the two variants is shown in Fig. 1.28. The top silicon is thicker in PD-SOI and the depletion region does not extend the full thickness leaving a floating body. This negatively affects the functioning of the transistor. An example is the lowering of V_{th} at high drain source voltage, also referred to as Drain-Induced Barrier Lowering (DIBL). In FD-SOI, the very thin top silicon layer ensures that the full body is depleted. This eliminates the negative floating body effects seen in PD-SOI. The negative effects of floating body in PD-SOI can be countered by making use of body contacted devices which enables control of body voltage.

The two technologies are compared in Table 1.2. FD-SOI is suitable for ultra low power applications while PD-SOI is used for other applications like RF, analog, automotive etc. as it is more mature and robust. In the further sections, the compatibility of SOI for different applications is highlighted. It will be clear that in addition to high volume products, SOI technology offers prospects for integration of different functionalities on a chip and enabling SoC scaling.

Table 1.2: Comparison of PD-SOI and FD-SOI technologies [51]

Type	Structural Differences	Target Applications	Advantages	Challenges	Nodes
PD-SOI	-Doped channel -Top silicon 50 to 90 nm thick (or more for "thick SOI" applications"	-High performance microprocessors -Most others (embedded analog, RF, automotive, power, military, aerospace etc.)	-Well understood -Industrially proven -Easy to manufacture -Can leverage floating body for performance gain or memory applications	-Physical limits to scalability are approaching for high-performance	-180 nm to 22 nm
FD-SOI	-Often uses undoped or lightly doped channel -Top silicon 5 to 20 nm thick -Insulating BOX layer may also be ultra-thin: 5 – 50 nm	-High performance microprocessors -Low power electronics -Ultra-low power	-Leakage and power consumption are drastically reduced -For undoped channels, random fluctuations in Vth, are minimized -No floating body effect; easier to control short-channel effects	-New methodology needed for defect detection in very thin layers -Vth defined by gate work function and intrinsic body -Very thin body can be challenging to manufacture and implement performance boosters	-22 nm and beyond for high performance microprocessors and low power electronics -Ultra low power now at 150 nm

1.4.1 Memory and logic

For the high-volume logic and memory market, transistor scaling is the key. For bulk CMOS nodes beyond 32 nm, the variation of threshold voltages due to random dopant fluctuation is an increasing problem [52]. The use of SOI technology with a very thin body helps reducing the V_{th} fluctuation. Short channel effects can also be effectively controlled [53]. The use of Fully Depleted (FD) technology on SOI provides the following improvements: reduction of V_{dd}, area optimization, and low power consumption. The ultra thin body technology offers two routes to scaling as shown in Fig. 1.29: (a) Vertical FinFET (b) Planar FD-SOI

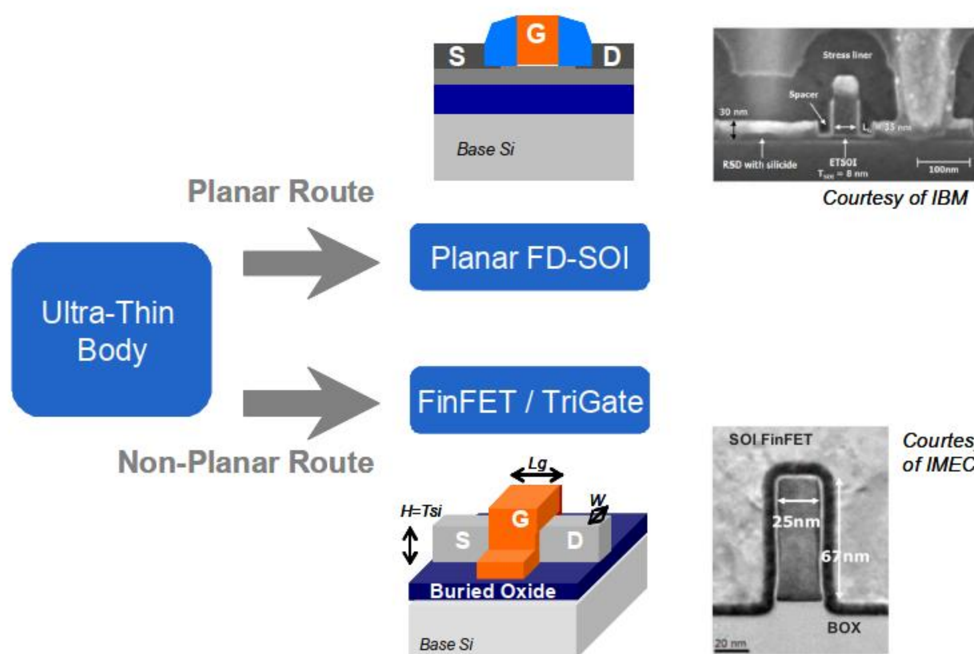


Fig. 1.29: Ultra thin body scaling routes for FDSOI technology [52]

The use of planar technology is less disruptive as it permits usage of existing EDA tools and methodologies for design. The planar technology can be easily adapted by designers. With the introduction of FinFET, better electrostatic control can be achieved which results in smaller leakage currents. Also, it is the leading technology for ultimate scaling of CMOS down to 5 nm [54]. However, FinFETs have much more complex topologies than planar counterparts. Also, modelling and parasitics extraction are much more challenging in the FinFET which results in an increased time of chip development. Hence, planar technologies will still be relevant where time to market factor is critical.

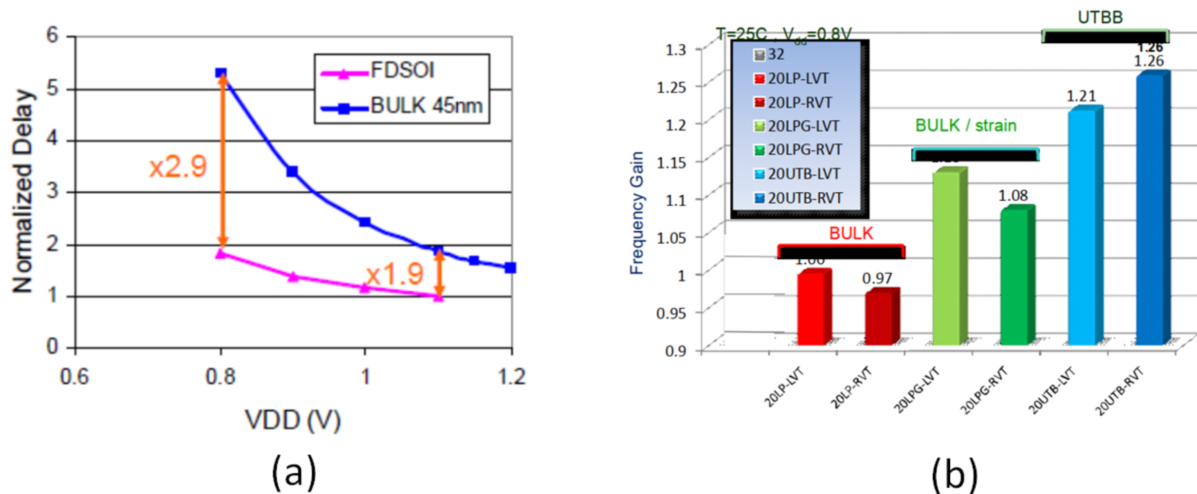


Fig. 1.30: (a) SPICE simulation 16-bit adder circuit on bulk and FDSOI technology (b) Frequency gain measured for LVT (low threshold voltage) and RVT (regular threshold voltage) for regular bulk, strain engineered bulk and FDSOI technologies [52]

The superior performance of FDSOI technology is highlighted in Fig. 1.30 for different technology nodes. It can be seen that low voltage drive provides nearly a 2.9x improvement in normalized delay for a 16-bit adder circuit. Also, the comparison between LVT and RVT libraries on FDSOI reveals higher gains as compared to bulk CMOS and also bulk CMOS with strain engineering.

For use in memories, FDSOI provides a better SRAM yield which can be used at lower V_{DDmin} [52]. This helps in active power saving for the SRAM. In bulk CMOS nodes beyond 32 nm, for good power maintenance, 8T or 10T configuration of the SRAM cell are preferred to avoid leakage. This increases the SRAM cell area and consequently reduces memory density. The other option in CMOS is to increase the operating V_{DD} of SRAM which leads to increased static power consumption.

A few examples of advantages of using FDSOI over bulk for logic and memory applications have been highlighted in this section. In the further sections, focus will be on the diversity of SOI for different applications.

1.4.2 RF/microwave circuits

The emergence of Internet of Things (IoT) with cloud based computing paradigms necessitates the development of power efficient wireless communication with a high data rates. With the data rate requirements continuously increasing, wireless designers need to be able to effectively handle problems like minimization of interference, reduction of power consumption etc. The use of frequency bands > 6 GHz also called millimeter waves will possibly be a key enabler for next generation communication systems in order to avoid spectrum

gridlock. Three important bands in the millimeter waves are Ka band (27.0–40.0 GHz), V band (57–64 GHz), and E band (71–76 and 81–86 GHz).

Silicon based technologies have been growing in relevance for use in both traditional RF/microwave as well as emerging millimeter wave applications. The scaling of RF transistors has been a leading driver for the increasing competiveness of silicon based RF circuits. The key RF figures of merit of a certain fabrication node are the unity current gain frequency (f_t) and the maximum oscillation frequency (f_{max}) for the transistor. The f_t and f_{max} values for different silicon based technologies are summarized in Fig. 1.31. The thumb rule for design recommends use of f_t and f_{max} values at least 3 times the operating frequency of the circuit [55]. These values can already be attained for millimeter wave operation by the existing fabrication processes using silicon. The device performances are comparable to traditional high-cost and high-performance platforms like GaAs and InP.

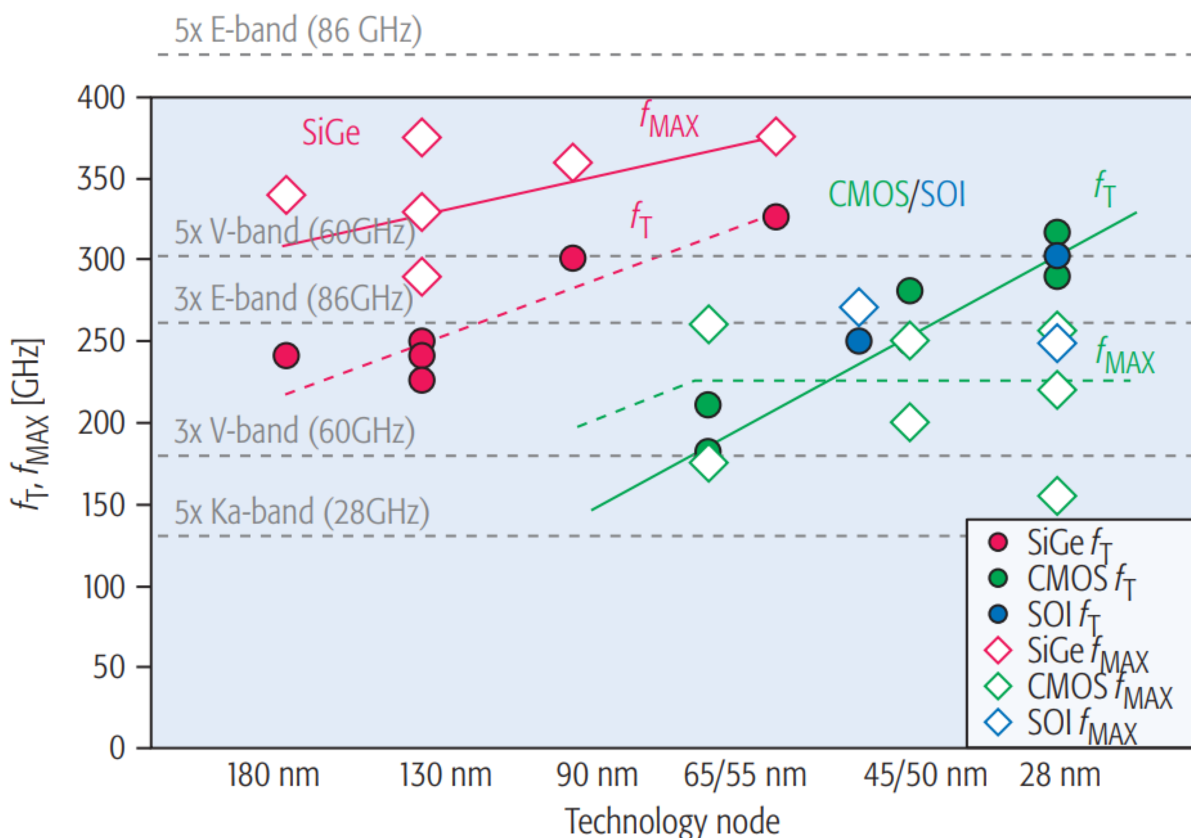


Fig. 1.31: Transistor f_t/f_{max} specification for different silicon based technology nodes [55]

While Fig. 1.31 is only indicative of the general trend of f_t/f_{max} scaling for different nodes, much higher values have already been demonstrated especially with the use of SOI substrate. By making use of channel strain for enhanced

mobility, f_{max} value of 500 GHz has been reported for 45-nm SOI nMOS by Lee et al. . Table 1.3 provides the state of the art f_t/f_{max} values on commercial silicon based technologies. The use of SOI substrate not only improves electrostatic control for transistors but have several other advantages as compared to bulk CMOS.

Table 1.3: f_t/f_{max} values of recent silicon based commercial technologies

Technology	ft (GHZ)	fmax (GHZ)	Reference
22NM FDSOI #	347 | 242 275 (mmW PFET)	371 | 288 299 (mmW PFET)	[56]
14NM FINFET %	314 | 285	180 | 140	[57]
28NM HIKMG BULK CMOS %	310 | 185	161 | 104	[58]
45NM PDSOI	296 (NFET)	342 (NFET)	[59]

- double sided gate contact, % - single sided gate contact

The first advantage is the possibility to integrate high quality factor passives on the substrate. An extensive discussion of this point is presented in section 2.2.2, where design of integrated inductors is discussed.

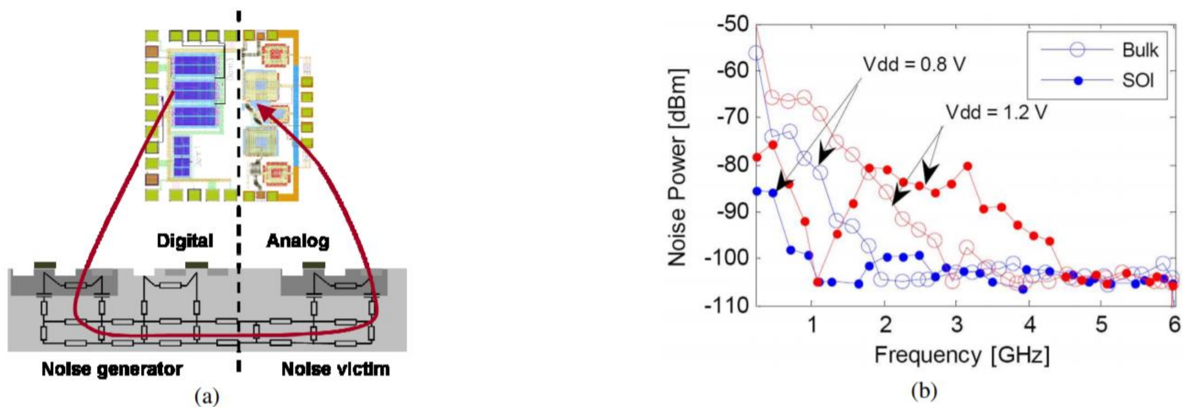


Fig. 1.32: (a) Schematic of DSN measurement showing 8 inverter chains on digital side (b) Measured noise power on the analog side at two different supply voltages[60]

For mixed circuit design, the digital blocks switching at high frequency can generate noise for analog circuits which can lead to poor performance of analog blocks. This coupling is quantified by the Digital Substrate Noise (DSN). Fig. 1.32 shows the results for measured DSN for 8 switching inverter trees at two bias voltages of 0.8 V and 1.2 V with a clock frequency of 225 MHz. It can be seen that the DSN generated for frequencies < 1 GHz at bias voltage of 1.2 V is

smaller in case of SOI as compared to bulk substrate. An improvement of ~20-25 dB as compared to bulk can provide the required substrate noise isolation for very sensitive analog circuits operating at MHz frequencies.

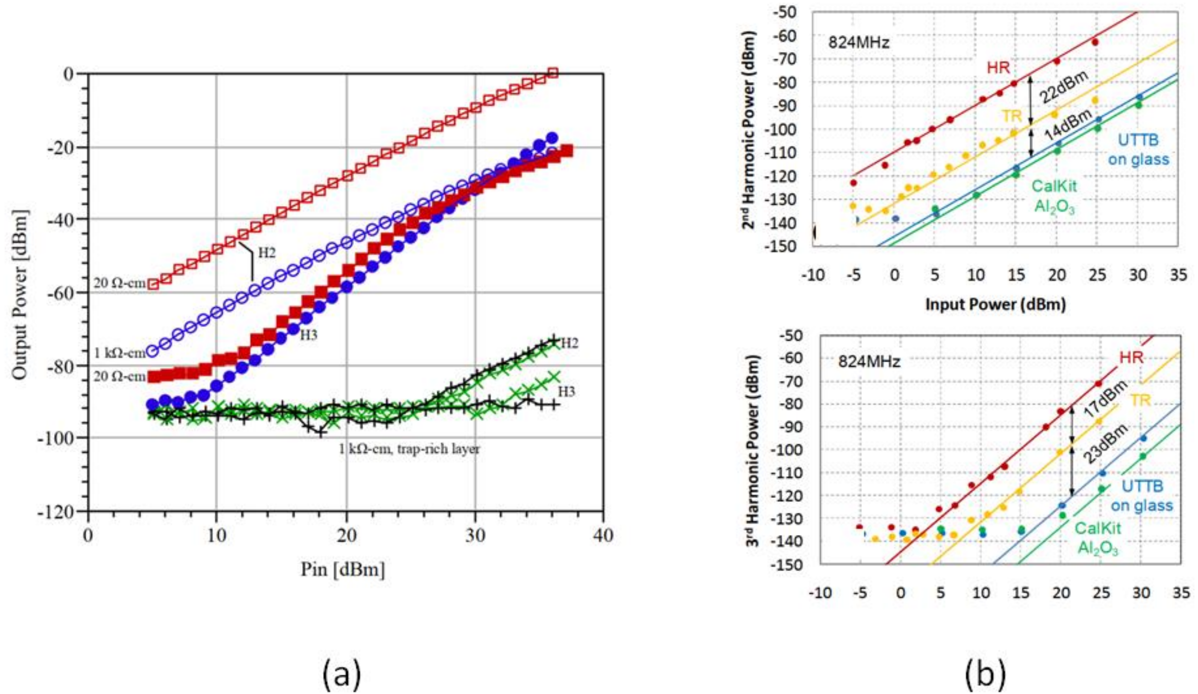


Fig. 1.33: Linearity of a coplanar waveguide measured for different substrate types with transmission line of length (a) 2.1 mm (b) 1 mm [61], [62]

With the advancement of communication standards and introduction of schemes like carrier aggregation enhances the data rate capabilities and also enables efficient utilization of spectrum. In carrier aggregation scheme, separate frequency bands (contiguous or non-contiguous) are aggregated in order to increase the overall bandwidth of the system. On the device level, in order to support carrier aggregation, one of the challenging aspects of the design is linearity. Non linear distortion is a big challenge for transceivers supporting aggregation of non-contiguous bands. This causes degradation of quality of the transmitted signal in Tx mode and also affect concurrent operation of Tx/Rx modes when using Frequency Division Duplexing (FDD) [63]. Substrate induced non-linear distortion is one of the main factors contributing to non-linear behaviour. Specialized SOI substrates provide an excellent solution for enhanced linearity performance.

The measurement of 2nd/3rd harmonic distortion on a coplanar waveguide serves as a good benchmark for substrate linearity. As shown in Fig. 1.33, the linearity improves with the increase in substrate resistivity to > 1 kΩ.cm. These

substrates are called High-Resistivity (HR) SOI substrates. While increase in substrate resistivity does improve linearity, the presence of the Parasitic Surface Conduction (PSC) layer at the silicon/oxide interface creates an inversion layer close to the interface (Fig. 1.34). This local region with higher conductivity leads to degradation of linearity. The introduction of a polysilicon layer at the interface helps trapping of charge carriers generated due to oxide charge. These substrates are called enhanced Signal Integrity (eSI) or Trap Rich (TR) SOI substrates. It can be seen in Fig. 1.33a that the use of trap rich layer greatly enhances the linearity for a transmission line by > 50 dB. The competences of eSI over HR-SOI are summarized in Fig. 1.34b.

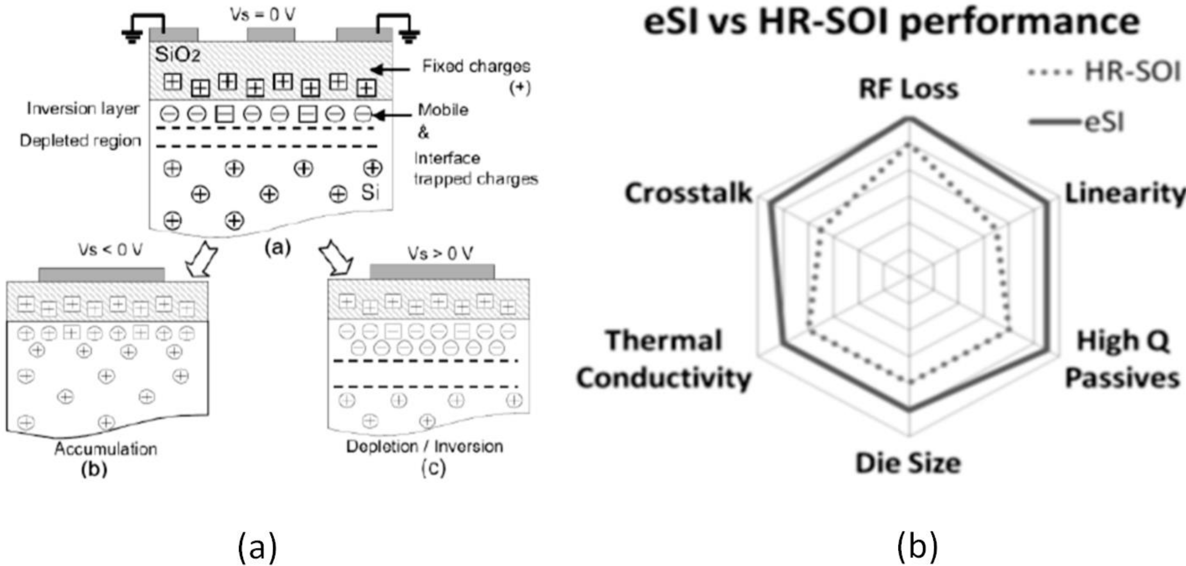


Fig. 1.34: (a) Schematic showing the charge state at the silicon oxide interface at different bias conditions (b) Performance comparison between eSi and HR-SOI substrate [64]

TR-SOI is currently the most competent commercial SOI technology for RF application. However, there is further scope for improvement. The removal of handler and transfer onto an insulating substrate results in even better linearity. This has been demonstrated in the work of Philippe et al. (Fig. 1.33b). There is significant improvement in the 2nd and 3rd harmonics of 1 mm coplanar waveguide on glass. The improvement of linearity by removal of handler substrate is one of the main motivating factors for this work.

1.4.3 Integrated photonic circuits

SOI technology is well suited also for the fabrication of integrated photonic devices. Integrated photonic circuits is a fast growing industry with a projected Compound Annual Growth Rate (CAGR) 26.4% for the next 5 years [65]. Photonic

ICs are important components of data centres and the continuous expansion of data centres provide opportunity for growth of photonic ICs.

By making use of a thicker BOX (~3 µm), photonic waveguides can be realized. The high contrast in refractive index between Si and SiO_2 provides the benefit of low losses. Fig. 1.35 shows the different optical functions that can be realized on SOI. A grating light coupler is formed by patterning of silicon or poly-Si which allows light from optical fiber to be coupled to the chip. Phase modification can be realized by injecting carriers on the silicon side of the waveguide. To interface optical signals to electrical signals, a photodiode is necessary. Germanium on silicon photodiodes are promising candidates for high speed communication applications. In the recently reported work of Huang et al, record sensitivities have been shown for different communication protocols like NRZ (up to 25 Gbps) and PAM4 (56 Gbps) using silicon on germanium avalanche photodiodes [66]. These promising results show the competency of SOI technology for future photonic applications where currently III-V materials are more popular.

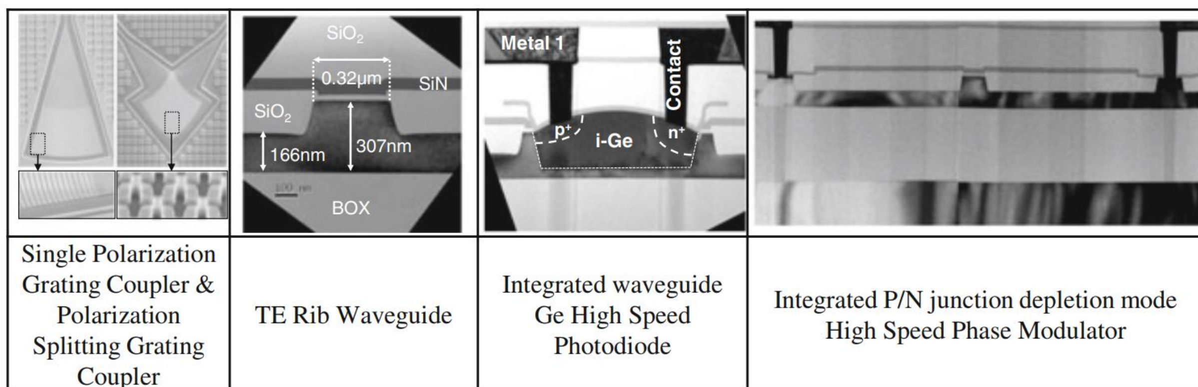


Fig. 1.35: Optical functions on SOI substrate [67]

The Mach-Zehnder Interferometer splits a single optical path to two, introduces phase modulations in one of the paths and then finally again recombines the two paths (Fig. 1.36). Based on this principle, 2x2 optical switches have been implemented by Rylyakov et al. with an ON/OFF ratio of >15 dB and a maximum transition time of 3.9 ns [68]. A much more area and power efficient switch design has been realized in the same work by making use of two ring resonators. With a low power consumption of 4 mW, a Bit Error Rate (BER) of 10^{-12} has been realized at optical power of -10 dBm using 40 Gbps random data.

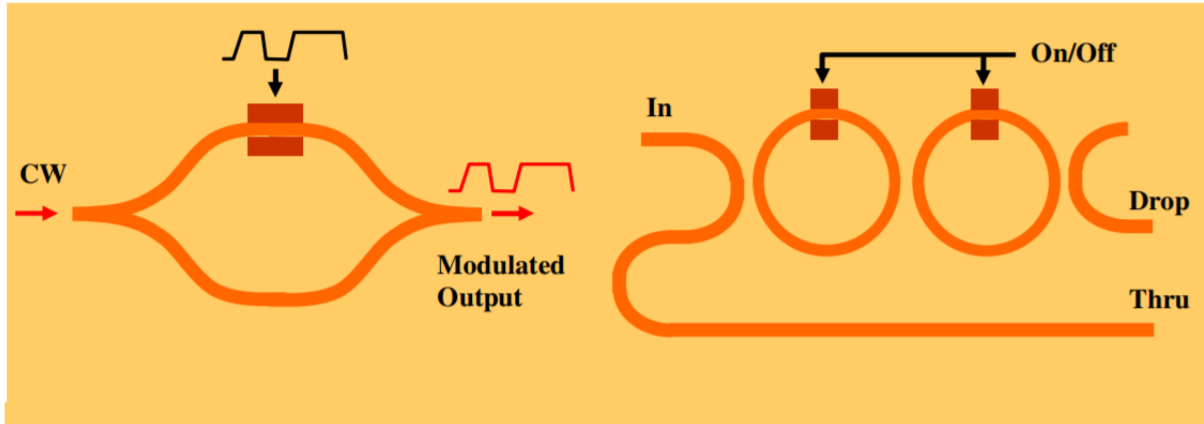


Fig. 1.36: Mach Zehnder Interferometer and optical switch [69]

1.4.4 MEMS

The BOX layer of the SOI substrate acts as an excellent etch barrier for etching of silicon. This makes it possible to fabricate MEMS structures using both back side and front side etching. Fig. 1.37 shows the implementation of pressure sensors by Olbrechts et al. using SOI transistors on membranes which act as electromechanical transducers [70]. Two configurations of the sensing element are reported. In the first case, an n-MOSFET is suspended on a membrane which drives the ring oscillator to provide a frequency output. The second implementation provides a current output and the sensing element is a current source placed at the borders of the membrane. For pressure change from 0.1 to 4.8 bars, there is a 12% change in frequency in the first implementation and a 20% change in current for the second implementation. This serves as an example of backside processing where the membrane is realized by means of backside etching of handler silicon using TMAH.

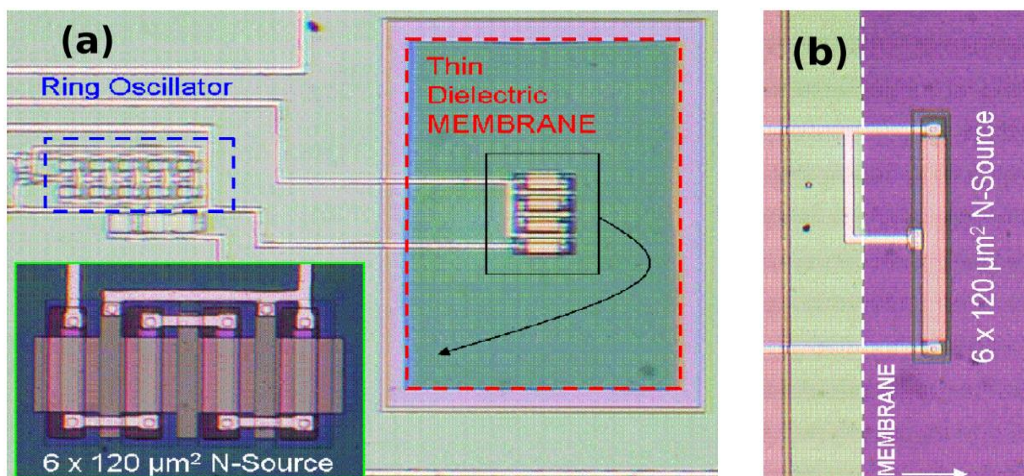


Fig. 1.37: Pressure sensors on SOI membranes (a) N-MOSFET suspended on membrane of area 250 x 400 μm^2 (b) A current source suspended on the edge of the membrane [71]

Fig.1.38 shows the implementation of self-assembled interdigitated sensors fabricated using a single photolithographic step [72]. The self assembly process is based on residual stress and plastic behaviour of multilayered structures. The sensing mechanism is based on the deflection of 3D movable cantilever beams under the presence of temperature/flow as the actuating stimuli. The most important fabrication steps are the thermal annealing and MEMS release. These steps have been successfully realized without affecting the supporting CMOS circuitry for the sensor which is a Ring Oscillator (RO). The cantilever beams are very thin with each beam having dimensions of length-200 μm, width- 10 μm and thickness – 1.25 μm. This serves as a good example of front side processing. In this case, the BOX does not act as a etch stop layer. Instead, it becomes the sacrificial layer which can be etched out to suspend MEMS structures.

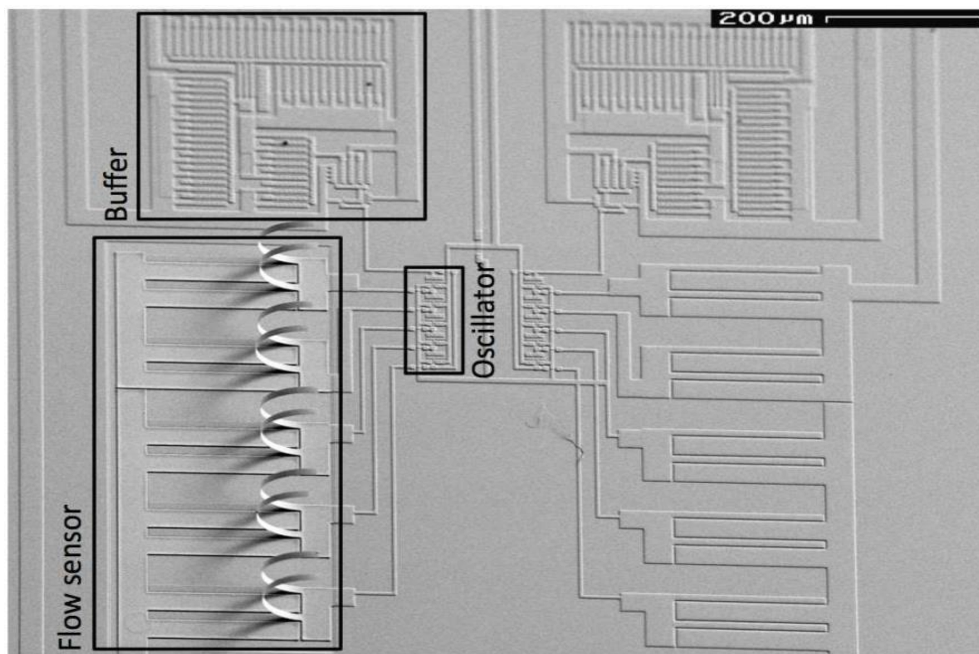


Fig. 1.38: Self assembled 3D MEMS structures using CMOS compatible processes for temperature and flow sensing applications [71]

Concluding Remarks

The general context of the semiconductor industry along with the current trends has been presented. The importance of *More than Moore* for the coming era of electronics has been highlighted. Efficient system integration and diversification are the cornerstones for *More than Moore*. The relevance of femtosecond laser processing and SOI technology as key enablers has been adequately highlighted by citing examples in current research.

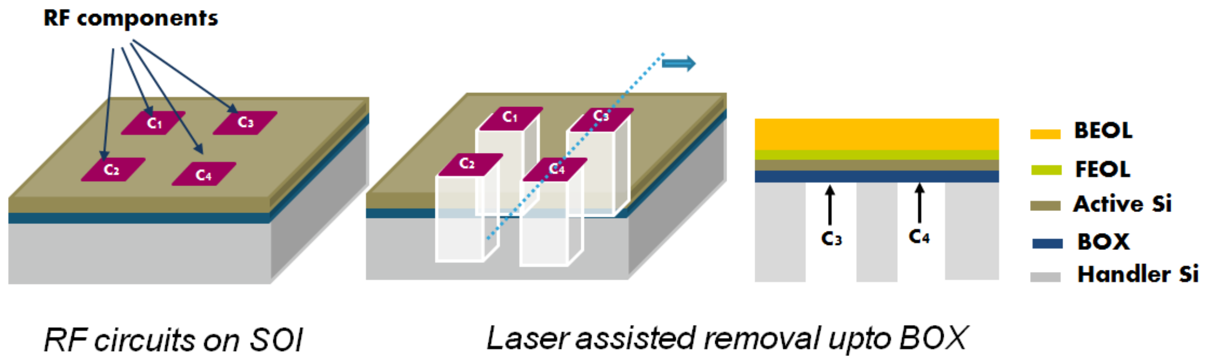


Fig. 1.39: A route for RF performance enhancement by making use of local substrate removal for alleviation of substrate effects

The vision of our research is to bring the two worlds together. We use laser removal of silicon to create locally freely suspended membranes of different RF functions on SOI substrate. The concept is illustrated in Fig. 1.39 which shows freely standing membranes of fabricated circuits realized using laser assisted processing. The objective is the improvement of RF losses and linearity in circuits which are degraded in the presence of handler substrate. The chapters that follow provide the state of the art (Chapter 2), literature review (Chapter 3), experimental description (Chapter 3) and results obtained in our study (Chapter 4 and Chapter 5).

References

[1] G. E. Moore, "Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff.," *IEEE Solid-State Circuits Soc. Newsl.*, vol. 11, no. 3, pp. 33–35, Sep. 2006.
[2] M. Bohr, "A 30 Year Retrospective on Dennard's MOSFET Scaling Paper," *IEEE Solid-State Circuits Soc. Newsl.*, vol. 12, no. 1, pp. 11–13, Winter 2007.
[3] C. A. Mack, "Fifty Years of Moore's Law," *IEEE Trans. Semicond. Manuf.*, vol. 24, no. 2, pp. 202–207, May 2011.
[4] M. M. Waldrop, "More than moore," *Nature*, vol. 530, no. 7589, pp. 144–148, 2016.
[5] M. Rosoff, "Your Phone Is More Powerful Than The Computer In The Spaceship NASA Launched This Week," *Business Insider*. [Online]. Available: http://www.businessinsider.com/your-phone-is-more-powerful-than-the-orion-computer-2014-12.
[6] M. Wang, G. Zhang, C. Zhang, J. Zhang, and C. Li, "An IoT-based appliance control system for smart homes," in *Intelligent Control and Information Processing (ICICIP), 2013 Fourth International Conference on*, 2013, pp. 744–747.
[7] D. Kyriazis, T. Varvarigou, D. White, A. Rossi, and J. Cooper, "Sustainable smart city IoT applications: Heat and electricity management & Eco-conscious cruise control for public transportation," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a*, 2013, pp. 1–5.
[8] K. Zhou, T. Liu, and L. Zhou, "Industry 4.0: Towards future industrial opportunities and challenges," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference on*, 2015, pp. 2147–2152.
[9] F. Fernandez and G. C. Pallis, "Opportunities and challenges of the Internet of Things for healthcare: Systems engineering perspective," in *2014 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*, 2014, pp. 263–266.
[10] L. Dignan, "Internet of things: $8.9 trillion market in 2020, 212 billion connected things," *ZDNet*. [Online]. Available: https://www.zdnet.com/article/internet-of-things-8-9-trillion-market-in-2020-212-billion-connected-things/.
[11] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context Aware Computing for The Internet of Things: A Survey," *IEEE Commun. Surv. Tutor.*, vol. 16, no. 1, pp. 414–454, 2014.
[12] R. Saleh *et al.*, "System-on-Chip: Reuse and Integration," *Proc. IEEE*, vol. 94, no. 6, pp. 1050–1069, Jun. 2006.
[13] K. Roy, B. Jung, D. Peroulis, and A. Raghunathan, "Integrated Systems in the More-Than-Moore Era: Designing Low-Cost Energy-Efficient Systems Using Heterogeneous Components," p. 10, 2016.
[14] M. Rose and H. J. Bergveld, "Integration Trends in Monolithic Power ICs: Application and Technology Challenges," *IEEE J. Solid-State Circuits*, vol. 51, no. 9, pp. 1965–1974, Sep. 2016.
[15] G. Ricotti, "Evolution of the Smart Power and High Voltage Technologies, Design and Simulations," presented at the 20th IEEE Workshop on Signal and Power Integrity (SPI).
[16] P. Apte, W. R. Bottoms, W. Chen, and G. Scalise, "Good things in small packages," *IEEE Spectr.*, vol. 48, no. 3, pp. 44–49, Mar. 2011.
[17] W. Diels *et al.*, "Single-package integration of RF blocks for a 5 GHz WLAN application," *IEEE Trans. Adv. Packag.*, vol. 24, no. 3, pp. 384–391, Aug. 2001.

[18] F. P. Carson, Y. C. Kim, and I. S. Yoon, “3-D Stacked Package Technology and Trends,” *Proc. IEEE*, vol. 97, no. 1, pp. 31–42, Jan. 2009.
[19] K. Croes *et al.*, “Reliability Challenges Related to TSV Integration and 3-D Stacking,” *IEEE Des. Test*, vol. 33, no. 3, pp. 37–45, Jun. 2016.
[20] E. Beyne, “3D interconnection and packaging: impending reality or still a dream?,” in *2004 IEEE International Solid-State Circuits Conference (IEEE Cat. No.04CH37519)*, San Francisco, CA, USA, 2004, pp. 138–139.
[21] E. Beyne, “Reliable Via-Middle Copper Through-Silicon Via Technology for 3-D Integration,” *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 6, no. 7, pp. 983–992, Jul. 2016.
[22] L. Madden *et al.*, “Advancing High Performance Heterogeneous Integration Through Die Stacking,” p. 8.
[23] R. Beica, “3D integration: Applications and market trends,” in *2015 International 3D Systems Integration Conference (3DIC)*, Sendai, Japan, 2015, p. TS5.1.1-TS5.1.7.
[24] M. D. Shirk and P. A. Molian, “A review of ultrashort pulsed laser ablation of materials,” *J. Laser Appl.*, vol. 10, no. 1, pp. 18–28, Feb. 1998.
[25] J. Meijer *et al.*, “Laser machining by short and ultrashort pulses, state of the art and new opportunities in the age of the photons,” *CIRP Ann.-Manuf. Technol.*, vol. 51, no. 2, pp. 531–550, 2002.
[26] R. Le Harzic *et al.*, “Comparison of heat-affected zones due to nanosecond and femtosecond laser pulses using transmission electronic microscopy,” *Appl. Phys. Lett.*, vol. 80, no. 21, pp. 3886–3888, May 2002.
[27] P. Wan, L.-M. Yang, and J. Liu, “All fiber-based Yb-doped high energy, high power femtosecond fiber lasers,” Opt. Express, vol. 21, no. 24, pp. 29854–29859, 2013.
[28] H. Fattahi *et al.*, “Third-generation femtosecond technology,” *Optica*, vol. 1, no. 1, p. 45, Jul. 2014.
[29] W.-S. Lei, A. Kumar, and R. Yalamanchili, “Die singulation technologies for advanced packaging: A critical review,” *J. Vac. Sci. Technol. B Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.*, vol. 30, no. 4, p. 040801, Jul. 2012.
[30] N. Sudani, K. Venkatakrishnan, and B. Tan, “Laser singulation of thin wafer: Die strength and surface roughness analysis of 80µm silicon dice,” *Opt. Lasers Eng.*, vol. 47, no. 7, pp. 850–854, Jul. 2009.
[31] K. Hasegawa, Y. Maruyama, and J. Corporation, “LASER RELEASABLE TEMPORARY BOND/DE-BOND MATERIALS FOR NEXT 3D PACKAGES,” p. 7.
[32] J. Fujimoto, M. Kobayashi, K. Kakizaki, H. Oizumi, T. Mimura, and H. Mizoguchi, “193 & 248 nm high power lasers for the micro and macro material processing,” p. 6.
[33] F. He *et al.*, “Tailoring femtosecond 1.5-µm Bessel beams for manufacturing high-aspect-ratio through-silicon vias,” *Sci. Rep.*, vol. 7, no. 1, p. 40785, Dec. 2017.
[34] M. Schwerter, D. Gräbner, L. Hecht, A. Vierheller, M. Leester-Schädel, and A. Dietzel, “Surface-Passive Pressure Sensor by Femtosecond Laser Glass Structuring for Flip-Chip-in-Foil Integration,” *J. Microelectromechanical Syst.*, vol. 25, no. 3, pp. 517–523, Jun. 2016.
[35] Y. Dong, R. Nair, R. Molian, and P. Molian, “Femtosecond-pulsed laser micromachining of a 4H–SiC wafer for MEMS pressure sensor diaphragms and via holes,” *J. Micromechanics Microengineering*, vol. 18, no. 3, p. 035022, Mar. 2008.
[36] M. Vendan and P. Molian, “Femtosecond pulsed laser microfabrication of SiC MEMS microgripper,” *J. Laser Appl.*, vol. 19, no. 3, pp. 149–154, Aug. 2007.
[37] Y. Wang, N. Dai, Y. Li, X. Wang, and P. Lu, “Ablation and cutting of silicon wafer and micro-mold fabrication using femtosecond laser pulses,” *J. Laser Appl.*, vol. 19, no. 4, pp. 240–244, Nov. 2007.

[38] O. J. A. Schueller, S. T. Brittain, C. Marzolin, and G. M. Whitesides, “Fabrication and Characterization of Glassy Carbon MEMS,” 1997.
[39] S. Johari, M. Z. Zainol, A. A. Azman, M. Mazalan, and Y. Wahab, “Fabrication of MEMS cantilever using laser micromachine,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 340, p. 012010, Mar. 2018.
[40] L. Amato, Y. Gu, N. Bellini, S. M. Eaton, G. Cerullo, and R. Osellame, “Integrated three-dimensional filter separates nanoscale from microscale elements in a microfluidic chip,” *Lab. Chip*, vol. 12, no. 6, pp. 1135–1142, 2012.
[41] S. LoTurco, R. Osellame, R. Ramponi, and K. C. Vishnubhatla, “Hybrid chemical etching of femtosecond laser irradiated structures for engineered microfluidic devices,” *J. Micromechanics Microengineering*, vol. 23, no. 8, p. 085002, 2013.
[42] Y. Liao *et al.*, “Rapid prototyping of three-dimensional microfluidic mixers in glass by femtosecond laser direct writing,” *Lab. Chip*, vol. 12, no. 4, pp. 746–749, 2012.
[43] J. Xu *et al.*, “Electrofluidics fabricated by space-selective metallization in glass microfluidic structures using femtosecond laser direct writing,” *Lab. Chip*, vol. 13, no. 23, pp. 4608–4616, 2013.
[44] C. Kopp *et al.*, “Silicon Photonic Circuits: On-CMOS Integration, Fiber Optical Coupling, and Packaging,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 17, no. 3, pp. 498–509, May 2011.
[45] R. M. Vázquez, S. M. Eaton, R. Ramponi, G. Cerullo, and R. Osellame, “Fabrication of binary Fresnel lenses in PMMA by femtosecond laser surface ablation,” *Opt. Express*, vol. 19, no. 12, pp. 11597–11604, 2011.
[46] M. Li *et al.*, “Photonic bandpass filter for 1550 nm fabricated by femtosecond direct laser ablation,” *Appl. Phys. Lett.*, vol. 83, no. 2, pp. 216–218, Jul. 2003.
[47] J. Lin *et al.*, “On-chip three-dimensional high-Q microcavities fabricated by femtosecond laser direct writing,” *Opt. Express*, vol. 20, no. 9, pp. 10212–10217, 2012.
[48] D. Bachman, Z. Chen, C. Wang, R. Fedosejevs, Y. Y. Tsui, and V. Van, “Postfabrication Phase Error Correction of Silicon Photonic Circuits by Single Femtosecond Laser Pulses,” *J. Light. Technol.*, vol. 35, no. 4, pp. 588–595, Feb. 2017.
[49] Z. Deng *et al.*, “Fabrication of large-area concave microlens array on silicon by femtosecond laser micromachining,” *Opt. Lett.*, vol. 40, no. 9, p. 1928, May 2015.
[50] G. G. Shahidi, “SOI technology for the GHz era,” *IBM J. Res. Dev.*, vol. 46, no. 2.3, pp. 121–131, 2002.
[51] “Fully Depleted (FD) vs. Partially Depleted (PD) SOI,” 15-May-2008. .
[52] C. Mazure, R. Ferrant, B.-Y. Nguyen, W. Schwarzenbach, and C. Moulin, “FDSOI: From substrate to devices and circuit applications,” in *2010 Proceedings of ESSCIRC*, Sevilla, Spain, 2010, pp. 45–51.
[53] Q. Liu *et al.*, “Ultra-thin-body and BOX (UTBB) fully depleted (FD) device integration for 22nm node and beyond,” in *2010 Symposium on VLSI Technology*, 2010, pp. 61–62.
[54] J. Kawa, “Designing with FinFETs: The opportunities and the challenges,” in *Synopsys Whitepaper*, 2012.
[55] A. D. Domenico *et al.*, “Making 5G Millimeter-Wave Communications a Reality,” *IEEE Wirel. Commun.*, p. 6, 2017.
[56] S. N. Ong *et al.*, “A 22nm FDSOI Technology Optimized for RF/mmWave Applications,” p. 4.
[57] J. Singh *et al.*, “14-nm FinFET technology for analog and RF applications,” *IEEE Trans. Electron Devices*, vol. 65, no. 1, pp. 31–37, 2018.

[58] K. W. J. Chew *et al.*, "RF performance of 28nm PolySiON and HKMG CMOS devices," in *2015 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, 2015, pp. 43–46.
[59] C. Li *et al.*, "A high-efficiency 5G K/Ka-band stacked power amplifier in 45nm CMOS SOI process supporting 9Gb/s 64-QAM modulation with 22.4% average PAE," in *2017 Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS)*, 2017, pp. 1–4.
[60] J. Raskin, "Is high resistivity SOI wafer the substrate solution for RF System-on-Chip?," in *2013 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, 2013, pp. 1–4.
[61] D. C. Kerr, J. M. Gering, T. G. McKay, M. S. Carroll, C. R. Neve, and J.-P. Raskin, "Identification of RF Harmonic Distortion on Si Substrates and its Reduction Using a Trap-Rich Layer," p. 4.
[62] J. Philippe *et al.*, "Application-oriented performance of RF CMOS technologies on flexible substrates," in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 15.7.1-15.7.4.
[63] A. Kiayani, V. Lehtinen, L. Anttila, T. Lahteensuo, and M. Valkama, "Linearity Challenges of LTE-Advanced Mobile Transmitters: Requirements and Potential Solutions," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 170–179, 2017.
[64] C. Maleville, "Engineered substrates for Moore and more than Moore's law: Device scaling: Entering the substrate era," in *2015 IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S)*, Rohnert Park, CA, USA, 2015, pp. 1–5.
[65] "Photonic Integrated Circuit Market | Growth, Trends, and Forecast (2019 - 2024)." [Online]. Available: https://www.mordorintelligence.com/industry-reports/hybrid-photonic-integrated-circuit-market.
[66] M. Huang *et al.*, "Germanium on Silicon Avalanche Photodiode," *IEEE J. Sel. Top. Quantum Electron.*, vol. 24, no. 2, pp. 1–11, Mar. 2018.
[67] L. Pavesi and D. J. Lockwood, "Silicon photonics iii," *Top. Appl. Phys.*, vol. 119, 2016.
[68] A. Rylyakov *et al.*, "A 3.9 ns 8.9 mW 4$\times$ 4 silicon photonic switch hybrid integrated with CMOS driver," in *2011 IEEE International Solid-State Circuits Conference*, 2011, pp. 222–224.
[69] J.-O. Plouchart, "Applications of SOI Technologies to Communication," in *2011 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*, Waikoloa, HI, USA, 2011, pp. 1–4.
[70] B. Olbrechts, "A Novel Approach for Active Pressure Sensors in Thin Film SOI Technology," *Procedia Eng.*, p. 4, 2011.
[71] J.-P. Raskin, "SOI technologies from microelectronics to microsystems-meeting the More than Moore roadmap requirements," *ECS Trans.*, vol. 49, no. 1, pp. 15–23, 2012.
[72] F. Iker, N. Andre, T. Pardoen, and J. Raskin, "Three-Dimensional Self-Assembled Sensors in Thin-Film SOI Technology," *J. Microelectromechanical Syst.*, vol. 15, no. 6, pp. 1687–1697, Dec. 2006.

Chapter 2: Towards high-end RF electronics and laser material processing: State of the art

2.0 Introduction

This chapter presents the state of the art with respect to two different aspects related to the thesis work: 1) Laser processing of silicon 2) Selected circuits of frontend modules studied in our work, which are fabricated on RF-SOI technology. Additionally, examples of RF circuits are also presented which highlight the performance gains obtained as a result of improved substrate parameters.

In section 2.1, methods of etching silicon using existing technologies are compared. The different capabilities of femtosecond laser processing for micro/nanostructuration of silicon are discussed. The viability of femtosecond laser for high speed and large area etching of silicon is also reviewed. Section 2.2 presents the design of integrated inductors and how SOI substrates have an added advantage. Furthermore, several studies are reported for suspended inductors where part or whole of the substrate is etched away using micromachining methods. Section 2.3 reviews the different technologies for RF switches with a comparison of performance between bulk and SOI technologies. Section 2.4 also makes a comparison of Figure of Merit (FoM) for LNA on bulk and SOI technologies. Furthermore, the performance gains obtained by improvements in inductor Q-factor are highlighted.

Contents

2.1 Micromachining of silicon 45
2.1.1 Overview of methods for silicon etching 45
2.1.2 Femtosecond Laser processing: An unique tool for micro/nanostructuring of silicon 47
2.1.3 Laser micromachining: A potential candidate for high speed and large area silicon etching 50
2.2 Inductors 52
2.2.1 Factors affecting inductor performance 52
2.2.2 Inductors on SOI 55
2.2.2.1 Q-factor using different substrate resistivities and BOX thickness 55
2.2.2.2 Inductor on commercial SOI and impact on device performance 57
2.2.3 Suspended inductors 58
2.3 RF Switch 59
2.3.1 Overview of different implementations and technologies 61
2.3.2 Performance comparison: Bulk vs. SOI 62
2.4 Low noise amplifier (LNA) 65
2.4.1 LNA FoM: Bulk vs. SOI 65
2.4.2 LNA performance comparison by optimization of passives 66
Concluding remarks 69
References 71

2.1 Micromachining of silicon

Controlled etching of silicon is an important step to be realized in the experimental part of the thesis. It is thus important to summarize and compare the list of available techniques for silicon etching. After that, different processing capabilities of femtosecond lasers are reviewed. Finally, an outlook is presented on the usage of ultrafast laser for high throughput applications.

2.1.1 Overview of methods for silicon etching

In this section, methods are reviewed for removal of silicon on a macro scale. The etching methods can be broadly divided into wet etching and dry etching methods. Table 2.1 summarizes the different etching methods for silicon.

Table 2.1: Summary of standard silicon etching processes. Adapted from [1] and [2]

Type	Process Name/ Etchants	Possible masking layers	Etch rate (µm/min)	Ref
Dry	Bosch	Photoresist, SiO_2	2 - 30	[3]
	Cryogenic	SiO_2, Metal	7 - 15	[3], [4]
	XeF_2	SiO_2, Photoresist, Metals	3	[4]
Wet	HNA	Si_3N_4	4 - 90	[5], [6]
	KOH	Si_3N_4, SiO_2, SiC	0.003 - 10.5	[2]
	EDP	SiO_2, Si_3N_4, Metals	0.83	[2]
	TMAH	SiO_2, Si_3N_4	0.04 - 1.35	[2]

It can be seen that from Table 2.1, there is a breadth of methods to etch silicon with a wide variation of etch rates. This variation comes from the choice of multiple processing parameters like plasma power, gas flow, pressure etc. in case of dry etching and concentration of etchants, temperature etc. in case of wet etching. It is to be noted that many of the wet etching techniques like KOH etch are sensitive to the crystallographic orientation of silicon as they preferentially etch (100) planes.

The fastest way to etch silicon is by making use of HNA solution. HNA is an acidic solution comprising of three different acids namely hydrofluoric acid (HF), nitric acid (HNO_3), and acetic acid (CH_3COOH). The etching process is a redox reaction in which HNO_3 acts at the oxidizing agent for silicon. The oxidized silicon is soluble in HF which causes the oxidized layer to be dissolved exposing

silicon. The exposed silicon is again oxidized by HNO_3 and the cycle repeats. Acetic acid is used to dilute the concentration of the other 2 strong acids. The etch rate is a function of relative concentrations of the acids. For example, a 2:1 ratio of HNO_3 (69%): HF (49%) gives a very high etch rate of 50 µm min^{-1}.

However, this method is not applicable for our application since the etch stop layer is the buried oxide (BOX). Also, such strong reagents are not preferable for pre-fabricated circuits because of their highly corrosive nature. The other wet etching methods listed in Table 2.1 are also not suitable for backside silicon etching because of slow etch rates, presence of metallic ions (for etchants like KOH) and difficulty in maintaining vertical profile at high aspect ratios [7]. The etching profile of TMAH is illustrated in Fig. 2.1a which shows a slanted profile.

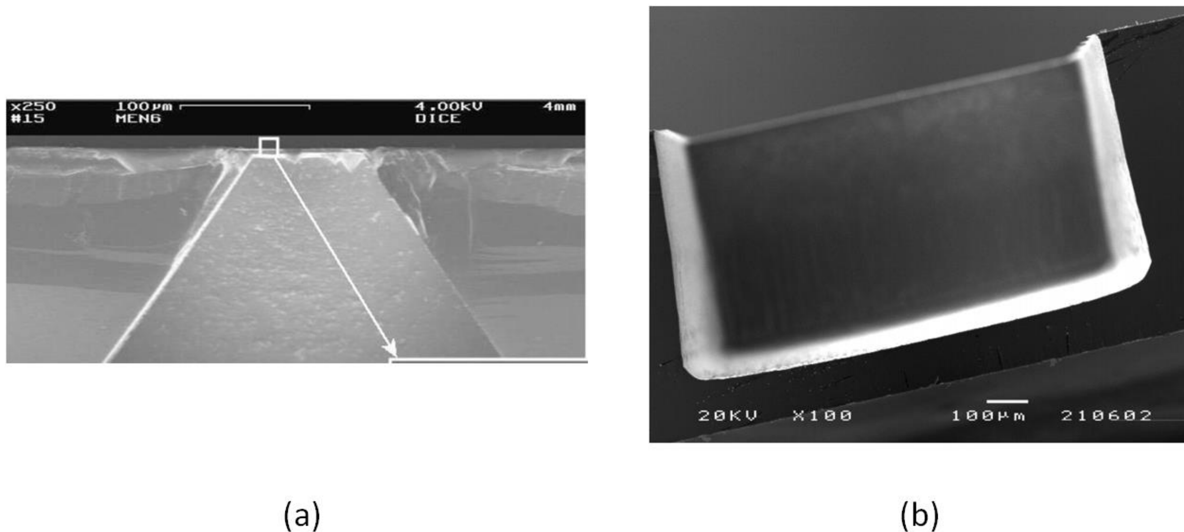


Fig. 2.1: Example of macro-sized cavities etched in silicon using (a) TMAH: Etch Rate – 1.2 µm min^{-1} [8] (b) Advanced silicon etching ,a variant of Bosch process: Etch Rate – 15 µm min^{-1} [9]

The Bosch process is a commonly used method for etching of silicon for high aspect ratio geometries. This can be readily adopted for etching of macro-sized cavities as seen in Fig. 2.1b. The etch profile is highly vertical as seen in Fig. 2.1b and the etch rates can reach up to 30 µm min^{-1}. This is the best known method that can be used for wafer scale etching of silicon with commercially developed technologies which can handle wafer sizes up to 200 mm [9].

2.1.2 Femtosecond Laser processing: A unique tool for micro/nanostructuring of silicon

Ultrafast lasers have unique capabilities for micro/nanostructuring of materials which exploit a variety of surface phenomena. While our material of interest is silicon, techniques used on metals and alloys are generally applicable to silicon as well. Hence, examples of both silicon and other materials are presented here. A few examples of material processing have been already presented in Chapter 1. Additional examples are presented in this section with an attempt to gain an insight into the processing capabilities at different length scales.

Laser-Induced Periodic Surface Structures (LIPSS) is a frequently used method for periodic micro/nano-structuring. The microstructuring of titanium using a femtosecond laser is shown in Fig. 2.2. The period can be tuned by using the appropriate wavelength and pulse width of the laser source [10].

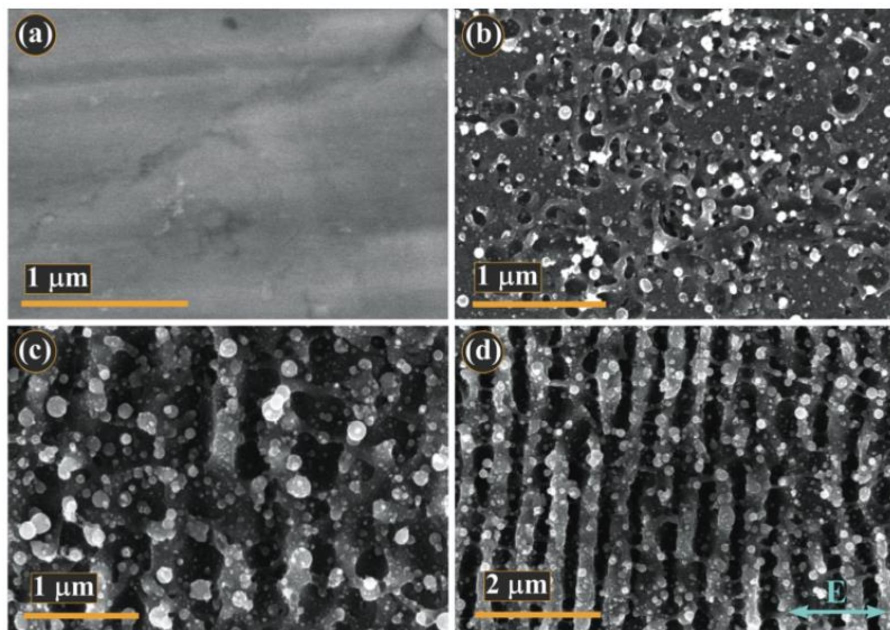


Fig. 2.2: LIPSS on Titanium illustrated with SEM images taken after (a) Initial surface (b) 2 shots (c) 10 shots (d) 40 shots [10]

The microstructuring gives rise to very interesting surface properties. Fig. 2.3 shows the application of microstructuring to colour printing on stainless steel of a well known painting of Vincent Van Gogh. The different colours are obtained by locally controlling the direction of LIPSS by appropriately tilting the polarization. The orientation of LIPSS is perpendicular to the direction of laser polarization.

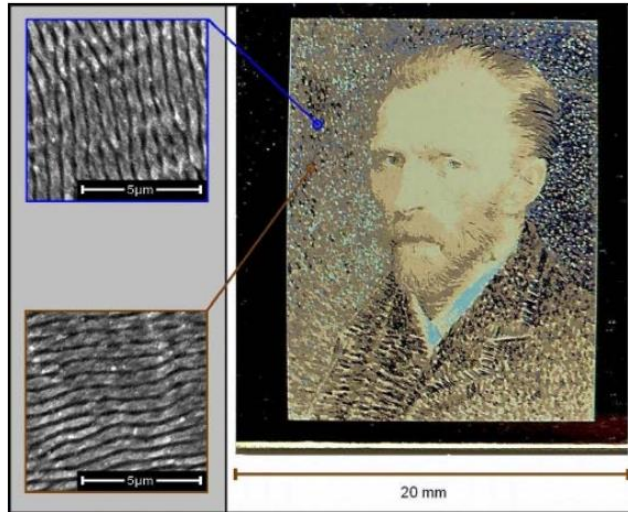


Fig. 2.3: Colour printing on stainless steel using LIPSS [11]

This method applied to silicon has interesting applications for photovoltaics. It can be seen in Fig. 2.4 that by formation of LIPSS, the surface can be made more absorptive. This can help improvement of solar cell efficiency. It was shown in the work of Halbwax et al that by carefully controlling the morphology of the surface structures, absorption can be as high as 94% [12]. Additionally, microstructuring can be used to control the hydrophobic properties of silicon. Superhydrophobic surfaces has been demonstrated on silicon with water contact angle of 160° which have self cleaning properties [13]. Furthermore, by careful control of the laser parameters, a surface can be created with a controlled gradient in hydrophobicity. This can serve several applications in microfluidics, biomedicine and biochemical sensors [10] .

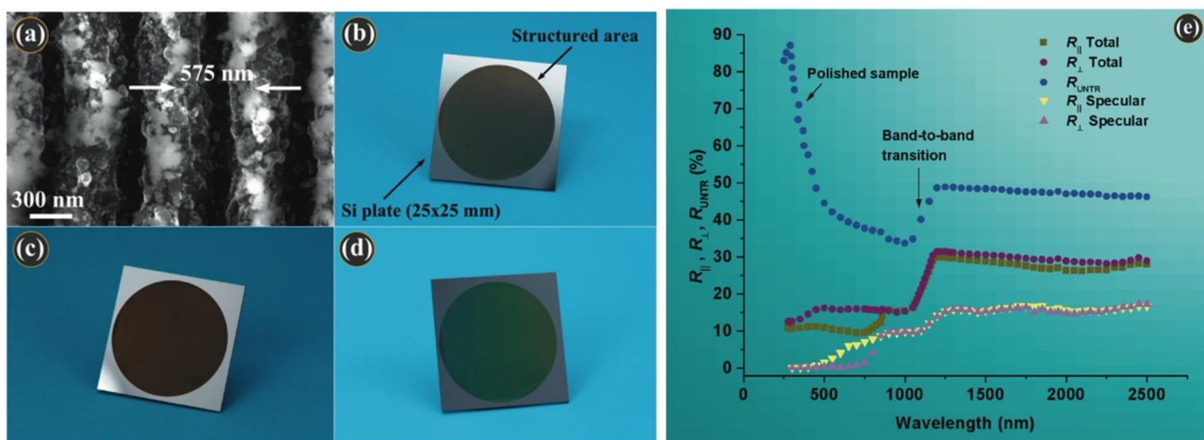


Fig. 2.4: (a) LIPSS on silicon (b,c,d) Structured silicon views from different angles (e) Reflectance measurements of microstructured silicon

Another unique possibility of femtosecond laser is selective removal of thin films on silicon and other substrates. There are several possible mechanisms that can be used for selective ablation. One of the methods is based on delamination of thin film due to absorption of laser energy in silicon. This is applicable to transparent materials. A quasi-metallic absorption mechanism results in the formation of dense electron-hole plasma close to the silicon surface which finally leads to the delamination of thin transparent material. This method has been shown to work for several materials like SiO_2, Al_2O_3 and SiN_x which is of potential interest for solar cell applications [14]. Also, optical gratings can be fabricated without the need for a mask.

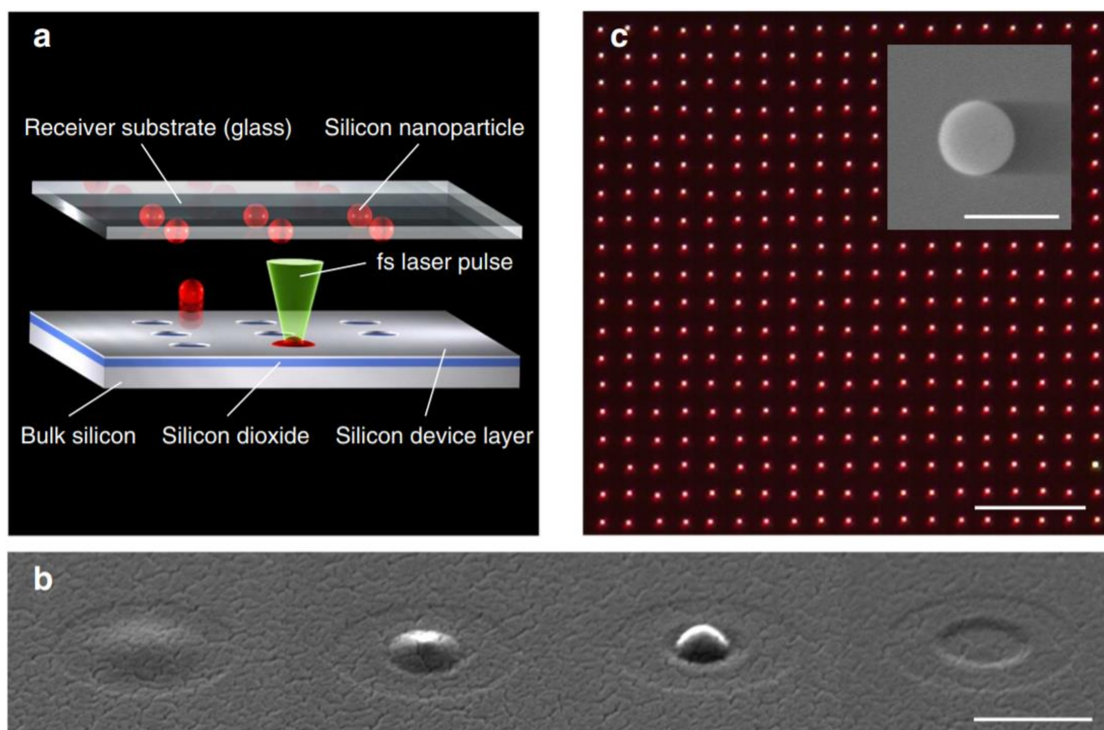


Fig. 2.5: (a) Method of fabrication of nanoparticles array by ejecting nanoparticles from a parent substrate to a host substrate (b) SEM images showing the formation of nanoparticle with the application of higher laser fluences (c) Nanoparticle array fabricated on glass substrate with the inset showing a single nanoparticles [15]

Femtosecond laser can also be used for fabrication of nanoparticle array on glass substrates. The fabrication method and the obtained arrays are shown in Fig. 2.5. Nanoparticles are ejected from parent substrate onto a transparent host substrate due to impinging laser pulse. The deposited nanoparticle is in the amorphous form. A second pulse irradiated on these nanoparticles causes annealing and changes the amorphous phase to crystalline phase. Such arrays have applications in silicon photonics.

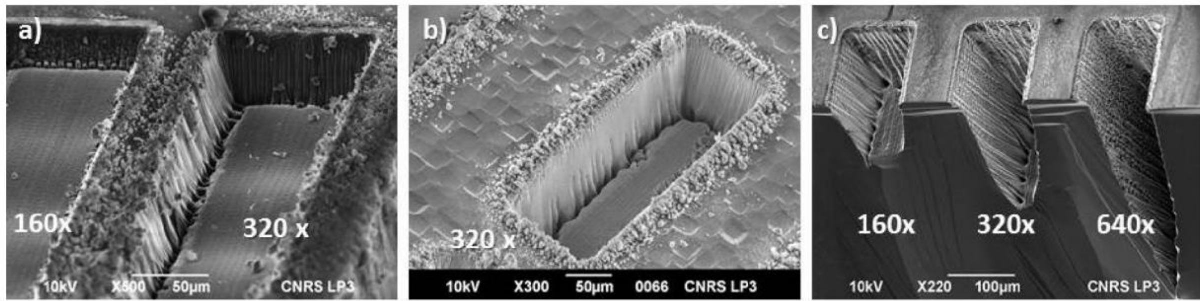


Fig. 2.6: Examples of milled cavities in silicon

So far, structuring has been discussed on the micro and nanoscale. Laser processing is equally effective for macroscale material processing. Fig. 2.6 shows the examples of cavities engraved in silicon using laser milling in the work of Sikora et al [16]. These cavities have been used to speed up the local silicon sample thinning for electron microscopy study of devices. A common method used for this application is Focused Ion Beam (FIB) milling which is very slow. Laser processing can remove a significant portion of the thickness rapidly which can then be further processed by FIB milling. Another example of macroscale processing is the fabrication of microchannel networks in silicon having multiple depths [17]. The fabricated structures served as gas exchangers in artificial lungs. A volume removal rate of 0.5 mm^3 min^{-1} was reported in this work.

Several examples have been discussed to demonstrate the processing capabilities of silicon at different scales. These are by no means exhaustive but serve to demonstrate the flexibility of laser processing. In the next section, the potential of laser for macro scale high throughput processing is presented which is a topic of specific interest for our application.

2.1.3 Laser micromachining: A potential candidate for high speed and large area silicon etching

The different etching methods for silicon have already been reviewed. While Bosch process and its variants are more suitable than wet etching for backside etching of silicon, there is an additional challenge to address. For our targeted study of prefabricated circuits, it is essential to have the proper alignment of the cavity in order to remove silicon on the backside in the targeted area. For that sake, appropriate etch mask needs to be defined on the backside which is well aligned to the front side. This would necessitate a mask set, a good alignment methodology and photolithography steps to remove silicon under RF circuits.

A maskless process would reduce considerably the number of steps for backside etching. Laser micromachining is one such process which has a high etch rate and without requirement of mask. A measure of the quality of etching using laser is the Heat Affected Zone (HAZ), which represents the region below the machined area where material modification takes place due to processes like melting, amorphization etc. The heat affected zone in case of femtosecond laser can be as small as 150 nm in the case of silicon [18].

It is also easier to implement the alignment methodology. Alignment holes and lines can be directly scribed on the front side which are visible on the backside. Using these alignment marks as a reference, laser milling can be performed in the desired regions to remove silicon without the need for masks and photolithography. This makes a good case for laser micromachining to be used in the targeted application of silicon removal.

In order to compare the throughput capabilities of laser micromachining with other processes, the etch rates need to be quantified. However, the calculation of etch rate is not straightforward like for other processes. The most commonly quoted values in literature is the etch depth per scan of the laser for scribing of lines. The time taken for one scan of the laser depends on the area of silicon to be processed. For a larger area, the beam needs to cover a longer trajectory and hence longer time as compared to smaller area in order to have the same depth.

Thus, it is necessary first to define an appropriate area of etch in order to calculate the ablation rate in µm min^{-1} for the chosen area. From a practical point of view, taking into account a wafer scale process, we can choose a 200 mm wafer as a reference. However, since we are interested in local ablation of certain chosen areas within the die, the area of etch can be taken as 10% of the area of the wafer. This gives the reference area of etch as 4000 mm^2. The etch rate in µm min^{-1} calculated for this reference area is termed as ER_{200}.

The state-of-the-art laser micromachining processes use a high laser power source and high scanner speeds in order to achieve high throughput. The studies on high throughput micromachining of silicon are sparse. Lopez et al. reported a volumetric ablation rate of 1.4 mm^3 min^{-1} at an average power of 6 W which gives a ER_{200} of 0.35 µm min^{-1}[19]. This is to our best knowledge the highest etch rate reported for silicon. However, there is a potential to improve this value significantly with existing systems.

Laser micromachining experiments have been reported using high average power for other materials. Schille et al. have investigated high throughput laser

micromachining for different materials [20]. The laser source used in this work has a wavelength of 1064 nm, pulse width of 10ps and average laser power of 187 W. This source is used in conjunction with a polygon scanner capable of scan speed of up to 800 m s^{-1}. The material removal rates are tabulated in Table 2.2. It can be seen that ER_{200} can reach up to 32.3 μm min^{-1}. This demonstrates the scalability of etch rates with the improvement of laser system specifications. The etch rates of silicon scales at approximately the same rate as aluminium with increasing laser power [19]. With this assumption, a rough estimate of ER_{200} value is 5-7 μm min^{-1} is obtained. With increasing average power output capabilities of commercial lasers, the etch rates can approach the Bosch process. For instance, development of 1 kW average power femtosecond laser was reported in collaboration with TRUMPF lasers [21]. Hence, further improvements can be expected with the commercialization of high power lasers.

Table 2.2: High throughput volume ablation rates for different materials [20]

Material	Type	Threshold Fluence (J cm^{-2})	Volume ablation rate (mm^3 min^{-1})	ER_{200} (μm min^{-1})
Aluminium (Al)	Metal	0.16	27.8	7
Copper (Cu)	Metal	0.28	21.4	5.4
Stainless Steel	Alloy	0.09	15.3	3.8
Aluminium Oxide (Al_2O_3)	Ceramic	0.81	129.1	32.3

2.2 Inductors

This section marks the beginning of discussion on RF circuits of frontend modules. Inductors are essential components in integrated front end module applications. For RF circuit design, keeping the inductor losses low is important for many applications. In this section, a brief introduction on inductor design is provided in order to understand some important tradeoffs. This is followed by inductor performance on SOI substrate and freely standing suspended inductors.

2.2.1 Factors affecting inductor performance

The realization of high quality factor integrated inductors is important for several RF design cases like 1) providing input impedance matching of LNA while keeping the noise figure low [22] 2) high quality LC resonators [23] for Voltage Controlled Oscillators (VCO) etc. This section outlines different studies on integrated inductors. The quality factor, values of inductance and resonance frequency are highlighted in these studies. Most of the studies reported on bulk

CMOS processes use a thick insulating layer between the substrate and the inductor trace to minimize the substrate coupling.

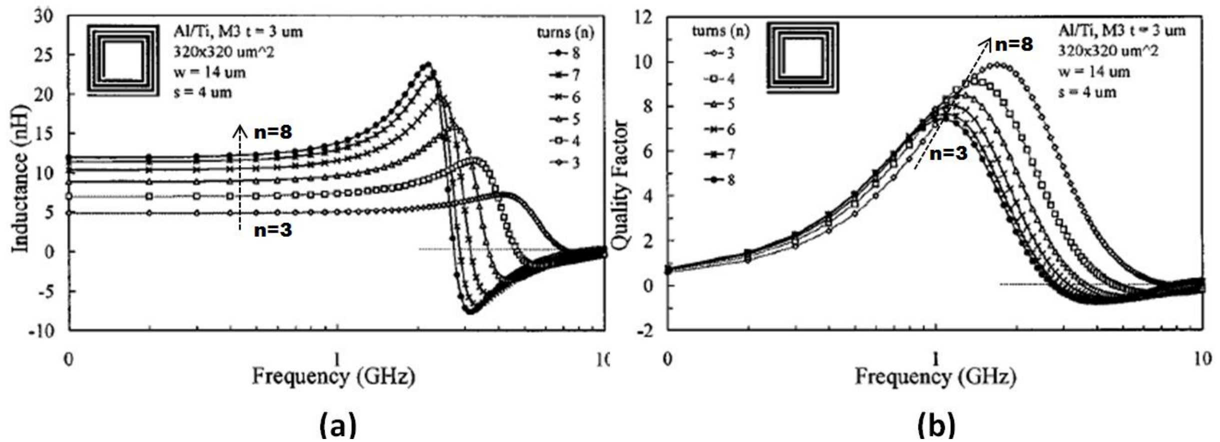
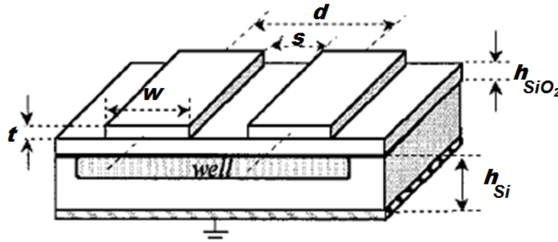


Fig. 2.7 : (a) Inductance for varying number of turns (b) Q-factor for varying number of turns [24]

There are numerous factors that determine the inductor RF behaviour with respect to inductance value, resonance frequency and quality factor. A thorough review of inductors has been provided by Koutsoyannopoulos et al. [24]. An example is taken to illustrate the inductor behaviour for varying number of turns. Fig. 2.7 summarizes the value of inductances obtained for varying number of turns of the inductor. For higher number of turns, the inductance value is higher due to increase in length of the inductor. But the quality factor reduces since the ohmic resistance of the inductor and the turn to turn coupling results in increased inductor losses. The resonance frequency also reduces with the increasing number of turns. This is because of higher capacitance between the adjacent turns of the inductor. In general, with the increase in inductance, the maximum quality factor and the frequency of self-resonance and peak Q-factor reduce. It is thus challenging when an application demands high quality factor and high values of inductance.



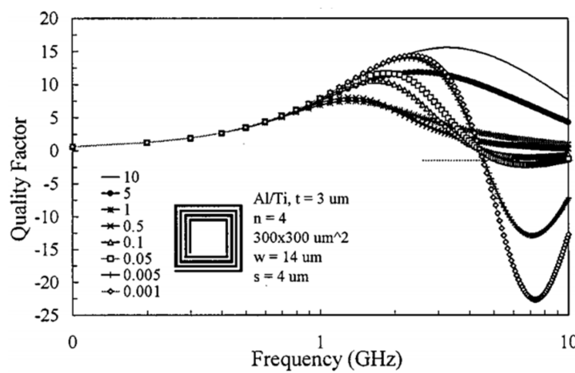
			Q_{max}	L	f_{res}
Conductor thickness (t)	↑	⇒	↗	—	—
Conductor sheet resistance (Rsh)	↑	⇒	↘	—	—
Insulator thickness (h)	↑	⇒	↗	—	↗
Substrate resistivity (ρ)	↑	⇒	↘↗	—	↗
Area	↑	⇒	↘	↗	↘
Number of turns (n)	↑	⇒	↘	↗	↘
Track width (w)	↑	⇒	↗	↘	↘
Multi-layer inductor, extra layer	↑	⇒	↘	↗	↘

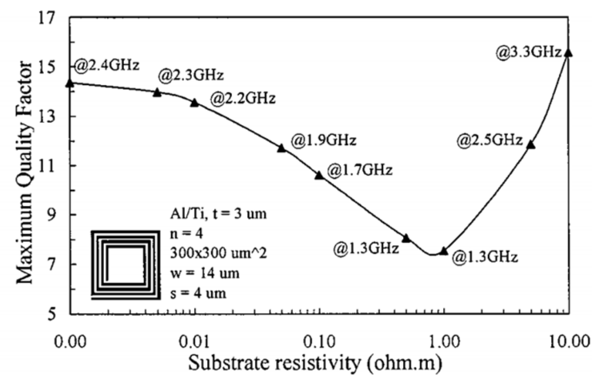
↗: increase, ↘: decrease, —: almost constant, ↘↗: exhibits minimum

(a) (b)

Fig. 2.8: (a) Characteristic dimensions of an inductor showing segments of two neighbouring turns (b) Variation of inductance, maximum quality factor and resonance frequency with variation in different dimensions and design of the inductor [24]

The effect of choice of number of turns has been discussed. There are a number of parameters that can be tuned for an inductor in order to obtain the desired inductance with the required Q-factor and resonance frequency. Fig. 2.8 provides a summary of different parameters that can be varied for an inductor and the consequence on RF performance. It can be seen that a combination of material and geometrical parameters determine the performance of the inductor.





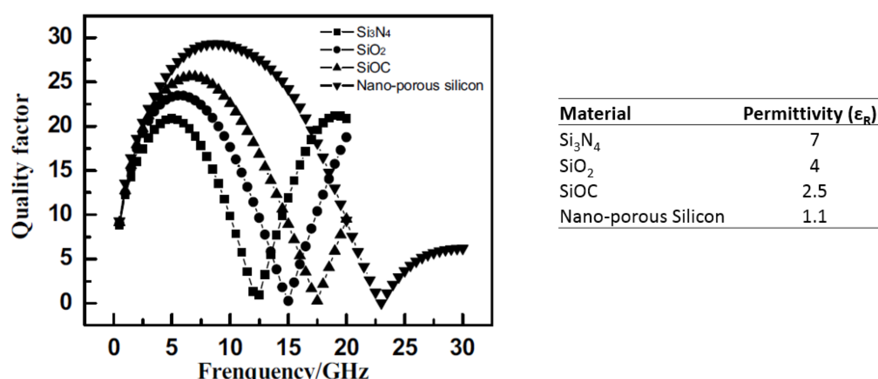
(a) (b)

Fig. 2.9: (a) Inductance with varying substrate resistivity (indicated in legend in units of Ω.m) (b) Q-factor for varying substrate resistivity [24]

It is interesting to notice the effect of substrate resistivity on the maximum Q-factor. With increase in substrate resistivity, it could be expected that the resulting losses in the substrate would be smaller. Correspondingly, the Q-factor would increase with increasing substrate resistivity. However, the maximum Q-factor varies non-monotonically as a function of substrate resistivity as seen in Fig. 2.9. It initially reduces with increasing resistivity and reaches a minimum.

After this point, the Q-factor again starts to augment with higher resistivities. The same effect has been reported by Burghartz et al. with a minimum of Q-factor obtained at a substrate resistivity of 10 Ω.cm [25]. The reduction of Q-factor has been attributed to a specific regime of operation between substrate resistivity values of 1 Ω.cm - 10 Ω.cm called the resonator mode regime. In this regime, resonance was found to occur through the oxide capacitance and substrate resistance.

In addition to all the mentioned factors, the choice of dielectric also plays an important role in the determination of inductor quality factor and resonance frequency. The capacitance from the inductor metal trace to the substrate scales with the dielectric constant and thickness of dielectric. At a given thickness, lower dielectric constant is advantageous because of the reduction in this capacitance. It can be seen in Fig. 2.10 that, by choosing a low dielectric constant material (nanoporous silicon) over a high dielectric constant material (Si_3N_4), the quality factor of the inductor improves by 50%.



Material	Permittivity (ε_R)
Si_3N_4	7
SiO_2	4
SiOC	2.5
Nano-porous Silicon	1.1

Fig. 2.10: Quality factor obtained for the use of different dielectrics [26]

2.2.2 Inductors on SOI

2.2.2.1 Q-factor using different substrate resistivities and BOX thickness

In the previous section, design of inductors was discussed which applies equally to all semiconductor fabrication technology platforms. In this section, some studies dedicated to the inductors fabricated using SOI substrate are outlined.

For an SOI substrate, two important substrate parameters are substrate resistivity and thickness of buried oxide. Eggert et. al studied integrated RF spiral inductors in SOI process for different resistivities [27]. The inductors studied had values 5 – 50 nH with a Q-factor up to 12 and number of turns ranged from 4 - 11. The measured inductor parameters are shown in Fig. 2.11. The thickness of buried oxide (BOX) used in this study is 400 nm. This thickness of oxide is typical in SOI-CMOS processes. It can be seen that by changing the

substrate resistivity from 20 Ω.cm to 10 kΩ.cm, the quality factor nearly doubles. The inductance and the self-resonant frequency do not change appreciably. However, the peak of Q-factor is obtained at different frequencies.

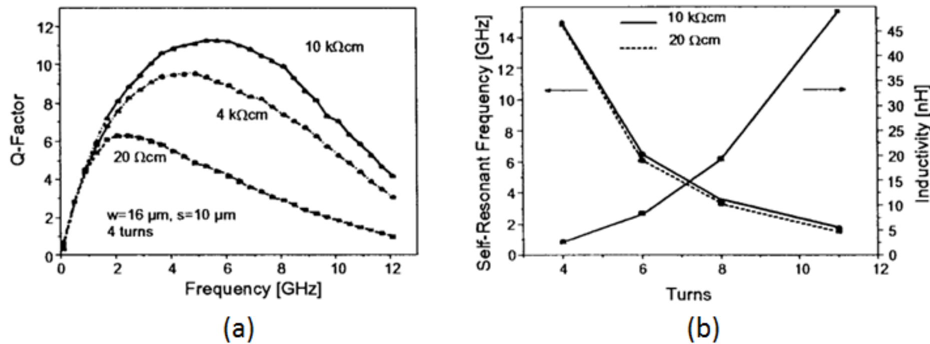


Fig. 2.11: (a) Q-factor for different substrate resistivities (b) Self-resonance and inductance values for different substrate resistivities [27]

When the substrate resistivity is fixed, the inductor design can also be optimized by the choice of thickness of buried oxide. As seen in Fig. 2.12a, the peak Q-factor increases with increasing thickness of oxide. If the thickness of oxide is sufficiently high, then the influence of substrate resistivity on inductor behaviour is minimum. As shown in Fig. 2.12b, the quality factor can be as high as 100 for an inductance value of 2 nH even with a low substrate resistivity of bulk CMOS wafers. While it is possible to have high buried oxide thickness option for SOI wafer, it is usually not preferred. This is because of the poor thermal conductivity of buried oxide.

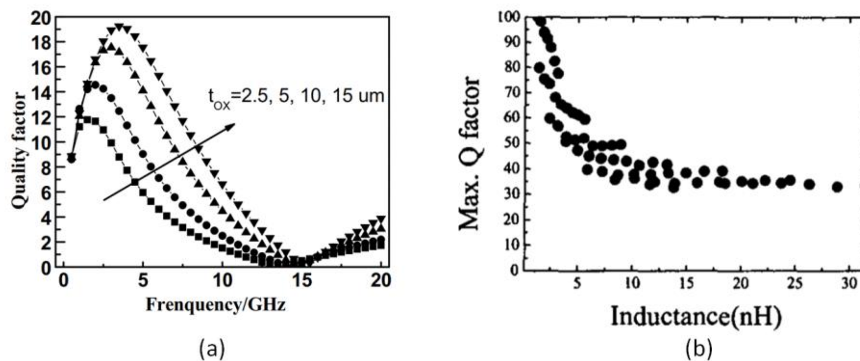


Fig. 2.12: (a) Q-factor obtained for varying thickness of oxide layer (b) Inductance values obtained on bulk Si wafers with a oxide thickness of 25 μm [26], [28]

2.2.2.2 Inductor on commercial SOI and impact on device performance

To have a better idea of the inductance performance on commercial SOI and its implication on application performance, the design of a 0.8nH inductance is reported on IBM 0.12 µm SOI-CMOS technology [29]. It can be seen in Fig. 2.13 that there is nearly a five fold improvement in Q-factor by changing substrate and metal stacking conditions. When using a Regular Resistivity Substrate (RRS) with a single metal line for the inductor, Q-factor is ~12. It increases to ~50 by stacking 4 metal lines in parallel and using a high resistivity substrate.

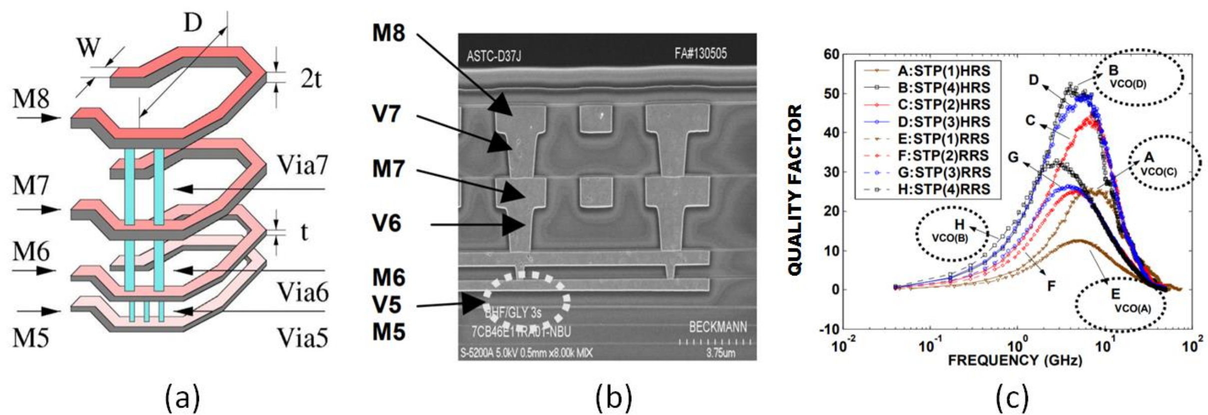


Fig. 2.13: (a)Parallel stacking of metal lines to improve Q factor (b) Cross section of the inductor (c) Q-factor curves for varying number of parallelly stacked metal lines indicated in parantheses for RRS (10-20 Ω.cm) and HRS (>300 Ω.cm) substrate types. [29]

This inductor is used in an LC tank circuit which is part of a Voltage Controlled Oscillator (VCO) as shown in Fig. 2.14. The performance of the VCO is highly dependent on Q-factor of the LC tank circuit consisting of inductor and varactor in parallel. The limiting value of Q-factor of LC tank is the inductor losses. The Q-factor values for inductor are much lower as compared to varactor. Hence inductor Q-factor dominates the device behaviour. As seen in Fig. 2.14b, by using high-Q inductors on HRS substrate, phase noise can be improved by ~15 dBc/Hz.

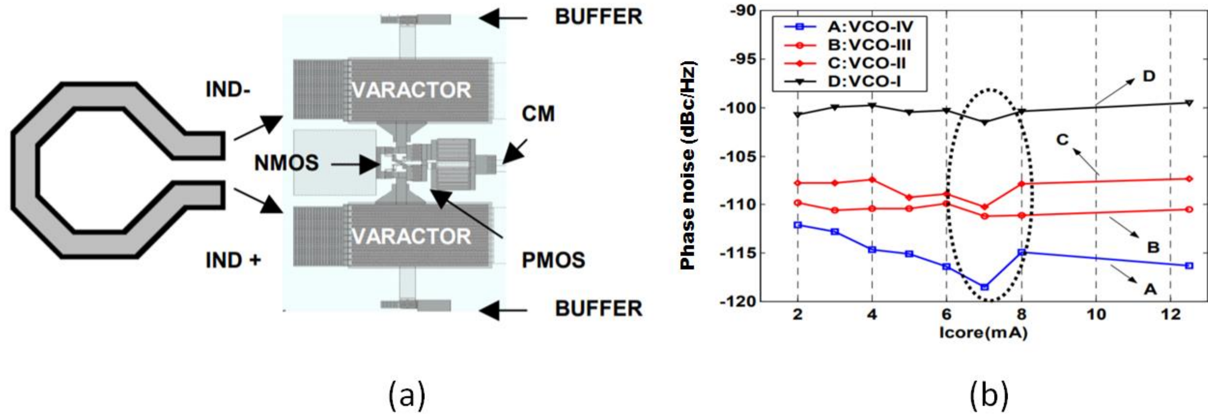


Fig. 2.14: (a) VCO circuit containing parallel LC tank circuit (b) Phase noise measured for different VCOs with varying Q-factor values (A) VCO-IV: Q = 52,4-metal HRS (B) VCO-III: Q = 22, 1-metal HRS (C) VCO-II: Q = 30, 4-metal RRS (D) VCO-I: Q = 11, 1-metal RRS

2.2.3 Suspended inductors

A number of studies have been performed to suspend the inductors in air by making use of bulk micromachining processes. The motivation for suspension of inductors is to reduce the coupling of inductor fields to the substrate. This reduces the losses due to electric and magnetic coupling in the substrate. Fig. 2.15 shows some examples of micromachined inductors with partial/complete substrate removal [30], [31].

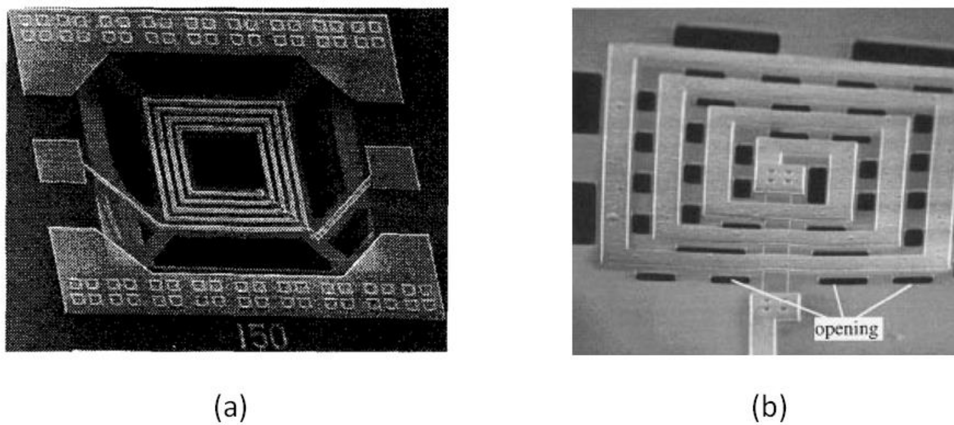


Fig. 2.15: SEM micrographs of micromachined inductors on silicon showing (a) Full removal of substrate (b) Partial removal of substrate [30], [31]

The obtained inductance values and corresponding Q-factors in various works are summarized in Table 2.3. It can be seen that high values of inductance can be obtained with a high Q-factor on silicon substrates. The highest values of quality factor are obtained using insulating substrates. On silicon substrate, relatively high Q-factors were reported in the work of Jiang et al [32]. In this work, a peak Q-factor of 36 was reported for an inductance value of 2.7 nH with a frequency of

peak Q-factor equal to 5.2 GHz. Such high values of Q-factors can help boost performance of RF circuits for applications like GSM and WiFi. It can also be seen in Table 2.3 that relatively large range of inductance values can be obtained by using the method of inductor suspension.

Table 2.3: Performance summary of suspended inductors on different substrates. Adapted from [33]

Inductance Value(s) (nH)	Peak Q Factor	Suspending Method	Substrate	Reference
1-30	50-70	Etch	Alumina	[34]
1-10	70	Double exposed PR mold	Silicon	[35]
20	<10	Etch	Silicon	[36]
7	<15	Etch	SOI	[37]
10-40	50-60	Etch	Glass	[38]
1-10	25-35	Etch	Silicon	[32]
0.3-3	80- >100	PR mold	Ceramic filled fiberglass	[39]
2-10	30	Double exposed PR mold	Silicon	[40]
23	<10	Etch	Silicon	[41]
1.8	30	Flip chip	Silicon	[42]
18	40	Flip chip	Silicon	[43]
4	35-40	PR mold	Glass	[44]

2.3 RF Switch

RF switch is the block immediately adjacent to the antenna in the front end which enables switching between transmit and receive operations in different frequency bands. This facilitates proper routing of RF signals during the operation of the transceiver. There are several ways to implement an RF switch like MEMS [45], PIN diodes [46] and transistors. This section will focus on RF switch designs using transistors. The block diagram of the front end module is given in Fig. 2.16. It can be seen that there are many switches within the module. We are interested in the large throw count antenna switch module. A large number of transistors are stacked in these modules as shown in Fig. 2.17.

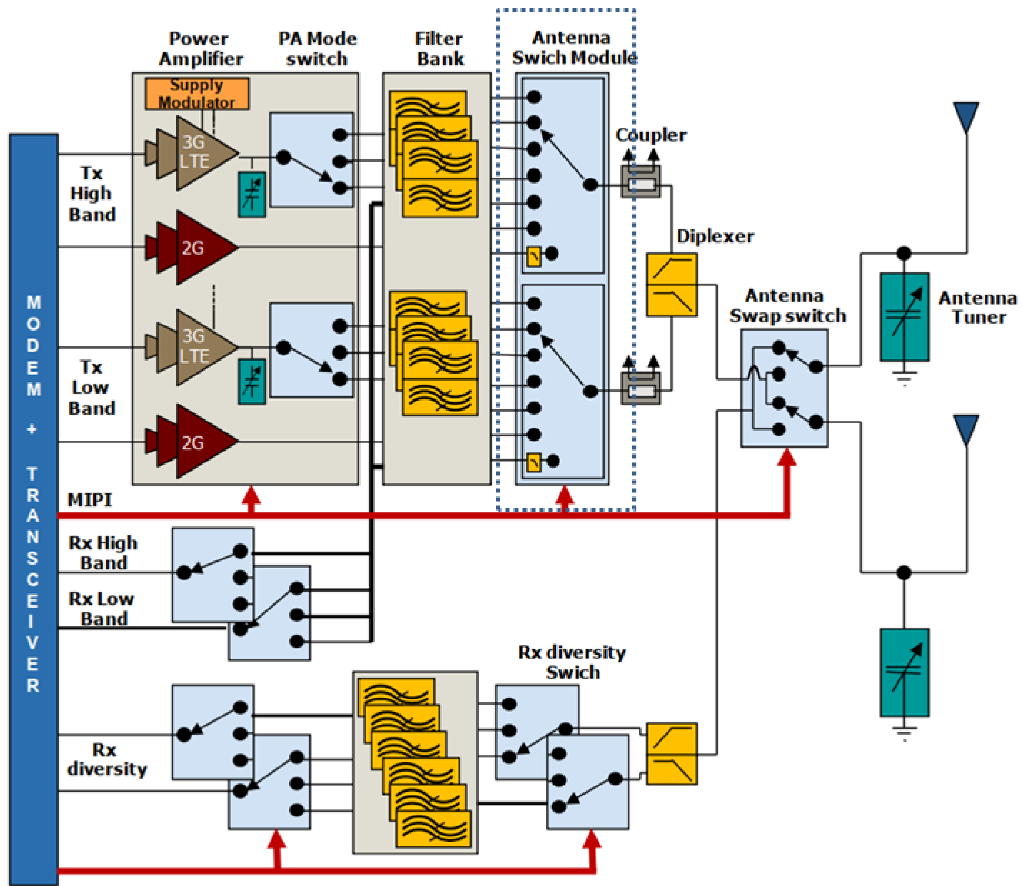


Fig. 2.16: Front end module components for LTE with switch module outlined [74]

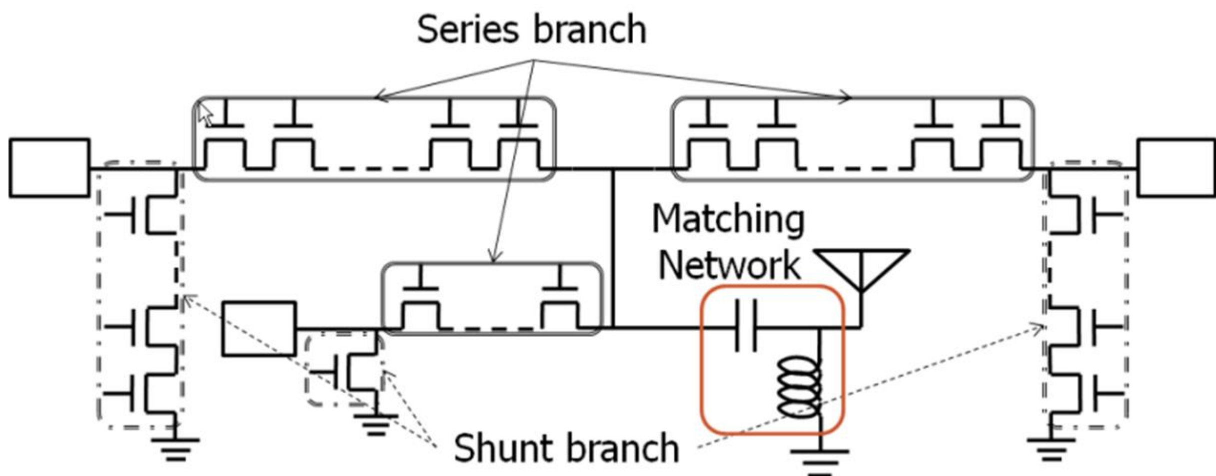


Fig. 2.17: Series shunt implementation of the switch showing transistor stacking in the series and shunt branches [75]

2.3.1 Overview of different implementations and technologies

There are many ways to implement an RF switch. The throw count of the switch depends on the number of frequency bands supported by the front end. The demand of frontends to support multiple bands has resulted in the increase of throw count of the switch of up to 14 – 16 [47]. In order to address such high throw counts, the series-shunt implementation has become the norm which uses stacked transistors to achieve switching operation (Fig. 2.17). The configuration of the switch largely depends on the architecture of the RF frontend. Two different switch configurations are shown as examples in Fig. 2.18. The SP6T configuration is used to support operation on the GSM band which uses a single antenna. The 2 x SP6T configuration is used when separate antennas are used to receive low and high frequencies. This frontend is capable of supporting quad-band GSM/EDGE standard and also 6 UMTS/LTE bands.

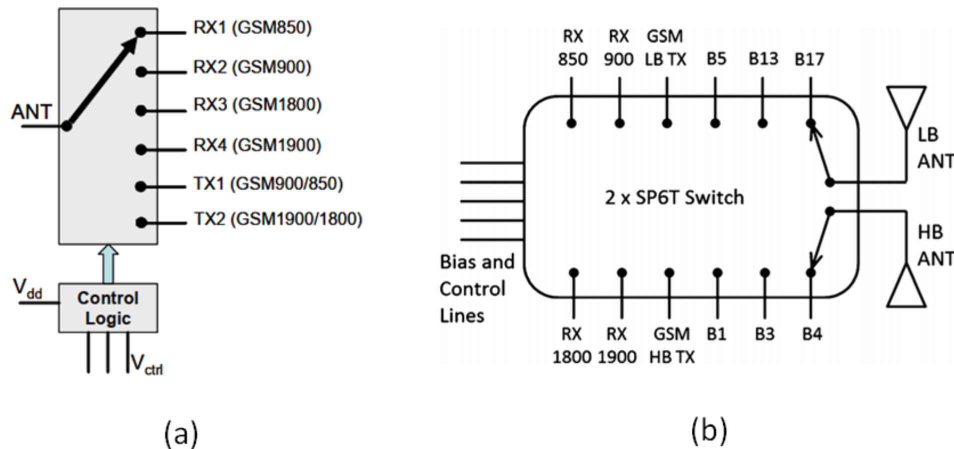


Fig. 2.18: Examples of implementation of switch on SOI technology (a) S96T: GSM only (b) 2xSP6T: GSM and LTE support [47], [48]

For any implementation of the switch, two important parameters can be listed. The first one is the resistance in the on state of the switch referred to as R_{on}. The second one is the capacitance of the switch in the off state referred to as C_{off}. While R_{on} determines the losses in the on-state of the switch, C_{off} determines the isolation of the electrical signal in the off-state. A switch design involves trade-off between these two parameters and they are tuned depending on the loss and isolation requirements. An overall measure of switch performance can be determined by taking the product of R_{on} and C_{off} and lower values of this product means better performance. It can be seen in Table 2.4 that $R_{on}*C_{off}$ value is very competitive for thin film SOI [49].

Table 2.4: Ron*Coff values for different process technologies [49]

Process	Device	Ron [Ω-mm]	Coff [fF/mm]	Ron*Coff [fs]
0.18 µm thick-film SOI	5V NFET Lg=0.6 µm 13.0nm gate ox	1.9	255	485
0.18 µm thin-film SOI	2.5V NFET Lg=0.32 µm 5.2nm gate ox	0.8	310	250
0.5 µm SOS	NFET 10.0nm gate ox	2.8	270	756
0.25 µm SOS	NFET 5.0nm gate ox	1.6	280	448
pHEMT	Single gate	1.4	160	224

2.3.2 Performance comparison: Bulk vs. SOI

Switch designs involve various tradeoffs. The most important is the trade-off between R_{on} and C_{off}. As shown in Fig. 2.19, when the transistor stack number increases, R_{on} increases but simultaneously C_{off} decreases. Hence, insertion loss increases while isolation is improved. However, it is desirable to use a higher stack number for better R_{on}*C_{off}. Another example is the trade-off between linearity and insertion loss. For instance, use of body biasing in SOI technology can significantly lower the stray capacitances of the transistor and can reduce the third harmonics generation and improve isolation [50]. But, because of additional leakage through body contact insertion loss is higher. By using floating body devices, the insertion losses can be reduced but linearity is compromised.

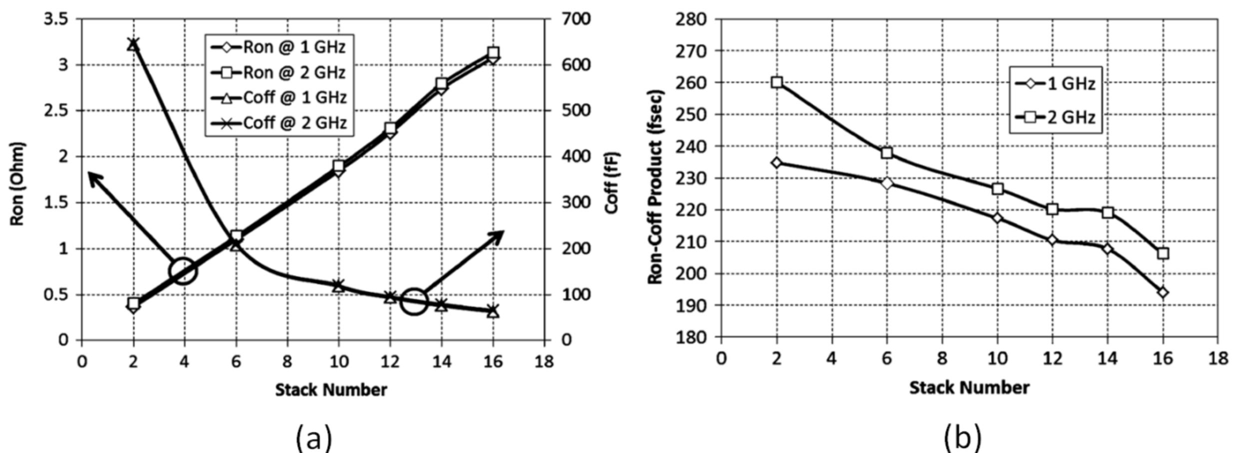


Fig. 2.19: (a) R_{on} and C_{off} as a function of stack number (b) R_{on}*C_{off} product as a function of stack number. The switch is fabricated on 0.18 µm thin-film SOI technology [47]

In addition to desired R_{on} and C_{off} levels, several linearity requirements also have to be met which are specific to the wireless standard. For instance, 2nd and 3rd harmonic linearity and intermodulation distortion are important for WCDMA standard. For high power operation, 1-dB compression point is a limiting factor. A few examples of different types of switches realized versus technologies and frequencies are shown in Table 2.5. The chosen examples mostly depict the switch performance for GSM900 and GSM1800 bands.

Table 2.5: Performance specifications of switches realized on different technologies for GSM frequency bands

Ref	[51]	[52]	[53]		[48]		[50]		[54]		[55]
Throw count	2	2	6		6		8		9		2
Freq (GHz)	0.9	1.9	0.9	1.9	0.9	1.9	2	0.9	1.9	0.9	1.9
Technology	130 nm bulk	180 nm bulk	180 nm SOI		130 nm SOI		180 nm SOI		180 nm SOI		180 nm SOI
IL (dB) (Tx)	0.5	1.6	0.43	0.75	0.7	0.7	0.7	0.47	0.53	0.45	0.65
ISO (dB) (Tx-Rx)	37	20	45	37	40	30	26	36	27.6	42	33
P_{1dB} (input)	31.3	33.5	37.2	35.6	36	34	40	x	38	37.5	x
2nd harmonic (dBc)	x	-46	-86	-85	-73	-74	-81	-75.3	-70.6	-106	-95
3rd harmonic (dBc)	x	-60	-82	-80	-76	-71	-82	-66.5	-60.4	-113	-107

The studies reported in Table 2.5 focus mostly on switches realized on SOI technology. The best switch performance with respect to $2^{nd}/3^{rd}$ harmonic distortion, to our knowledge, has been reported by Tombak et al. with reported levels of harmonic generation <-100 dBc in the GSM900 band [54]. While throw counts up to 9 have been reported here, the advancement of frontends demands even higher throw counts. SOI technology can address such demands with throw counts up to 16 having already been demonstrated [56]. With the increase of throw count of switches, the number of off branches also increases and correspondingly leads to enhanced combined capacitive effects of the off branches. The net effect is the increase in insertion loss through the switch because of parasitic coupling in the OFF branches. The substrate coupling in the OFF branches also degrades the linearity performance of the switch.

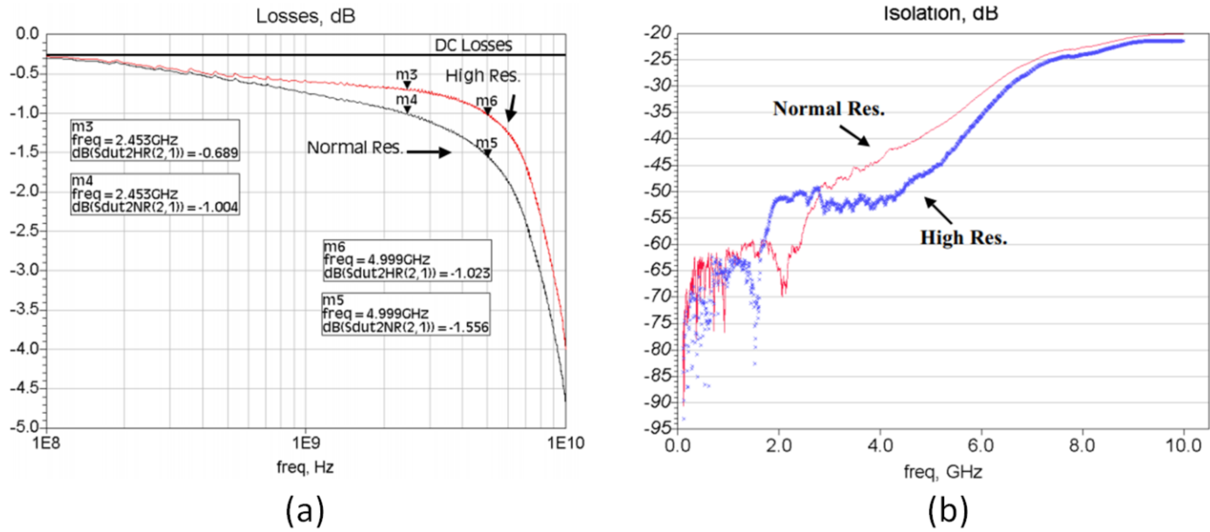


Fig. 2.20: Switch characteristics comparison of bulk vs. SOI SPDT switch on 0.25 µm CMOS process (a) Insertion loss (b) Isolation [57]

In order to have a clearer understanding of the performance of bulk CMOS vs. SOI, same SPDT switch design is studied on the two substrates [57]. As seen in Fig. 2.20, there is an improvement in insertion loss of ~0.3 dB at 2.4 GHz and ~0.55 dB at 5 GHz showing the superior performance of SOI substrate.

Power handling is also an important factor for the switch. The stack number is chosen so as to be able to handle the power requirements. The maximum power handled by the switch assuming an equal division of voltage in the stack is (taken from [50]):

$$P_{max} = \frac{n^2 \min[(V_{bvds}, V_{dsmax})]^2}{2Z_0} \tag{2.1}$$

$$V_{DSmax} = 2\left|V_{goff} - V_{th}\right| \tag{2.2}$$

where

V_{bvds} – *Drain to source breakdown voltage of transistors*

V_{dsmax} – *Maximum voltage swing to prevent turning on of source/drain diodes*

V_{th} – *Threshold voltage of the transistor*

Z_0 – *Characteristic load/source impedance*

The power handling capability is determined by two factors in the OFF branches: Drain to source breakdown and forward bias of drain/source diodes. It is clear from equation 2.1 that the stack number has to be increased for increased power handling. This also increases the physical dimensions of the switch.

2.4 Low noise amplifier (LNA)

Low noise amplifier is the usually first stage in the receiver chain of a radio system. In modern RF frontends supporting multiple frequency bands, it is preceded by a switch which routes the antenna signal to the desired band. The receiver chain may optionally include pre-select filter before the LNA.

2.4.1 LNA FoM: Bulk vs. SOI

The purpose of using a LNA is to reduce the overall noise figure of the receiver. The noise performance of the first stage is crucial in determining the noise performance of the entire receiver. In addition to low noise, the LNA also has to be able to address the linearity constraints like 1-dB compression point and third order intercept point (IP3). The design of LNA requires tradeoffs between various aspects like area, linearity, noise, power etc. Traditionally, GaAs and bipolar technologies were favoured for LNA fabrication. However, with the shrinking of transistor size and advent of silicon based substrate technologies like SOI, the possible performance levels have drastically increased.

Table 2.6: Summary of LNAs realized in CMOS technologies. Adapted from [58]

Ref	Gate length	Freq	Voltage	Power	Gain	NF	$\|S_{11}\|$	$\|S_{22}\|$	P_{1dB}	IIP3	FOM
Unit	**nm**	**GHz**	**V**	**mW**	**dB**	**dB**	**dB**	**dB**	**dBm**	**dBm**	**MHz**
[58]	180	5.8	0.6	0.834	13.92	3.32	12.74	-13.38	-22.2	-11.5	2.1
[59]	90	5.5	1.2	9.72	12.3	2.7	-10.3	-19	-11	-3	1
	90	5.5	0.6	1.0	9.2	3.6	-10	-14	-15.8	-7.25	2.31
[23]	90	5.5	0.8	5.4	14.4	2.9	-13.4	-10.7	-18.4	-6.2	1.3
[60]	180	5.0	0.6	0.9	9.2	4.5	-12	-21	-27	-16.0	0.2
	180	5.2	0.6	1.08	10	3.37	-13.4	-10.6	-18	-8.6	1.8
[61]	180	5.0	0.6	1.68	14.1	3.65	-12.7	-14	-25	-17.1	0.2
[62]	180	5.2	1.8	12.4	16.5	1.1	<-20	-13	-19.9	-11.5	0.7
[63]	180	5.8	1.8	3.42	9.4	2.5	-13.5	-14.8		-7.6	1.1
[64]	180	5.4	1.8	2.7	21	2.8	<-10	x	x	-23.0	0.1
[65]	130	5.65	1.2	6.4	14.9	4.8	-32.4	x	x	-4.2	0.9
[66]	65	10	1.0	13.7	10.5	3.3	<-12	<-16	x	-3.5	1.0
[67]	130	6	1.2	2.79	18.9	3.8	-21	x	-16.3	-5.6	3.7

The performance of LNA can be summarized by defining a figure of merit (FoM) taking into account several parameters. The FoM considering linearity, power dissipation and noise performance has been defined by Brederlow et al. [68]. It is given by the following equation:

$$FOM = \frac{Gain*IIP_3*f_c}{(NF-1)*Power} \quad (2.3)$$

Table 2.6 summarizes results obtained on different LNA designs on CMOS technologies. It can be seen that the FOM ranges from 0.1 – 3.7. The SOI substrate enables integration of passives and amplifier circuit on the same die. The insulation provided by buried oxide mitigates substrate coupling effects and hence yields better performance. Some designs realized on SOI technology are summarized in Table 2.7. The superiority of SOI technology over CMOS is visible when comparing the FOMs. There is 1-2 orders of magnitude improvement in FOM by the use of SOI technology. The use of advanced SOI substrates like TR-SOI provides further improvement in FOM partially due to passives Q factor increase.

Table 2.7: LNA designs on SOI technology. Taken from [69]

Ref	[69] TR-SOI	[69] HR-SOI	[70]	[71]	[72]
Tech [µm]	0.13 SOI	0.13 SOI	0.18 SOI	0.18 SOI	0.13 SOI
Freq [GHz]	5.9		1	5	1.8
NF [dB]	1.34	1.52	1.3	0.95	0.8
V_{DD} [V]	1.2	1.2	2.5	1.5	1.5
P_{DC} [mW]	9.6	9.6	50	12	6.6
Gain [dB]	9	8.8	10.2	11	11
$\lvert S_{11} \rvert$ [dB]	-14.2	-13.1	-10	-33	-12
$\lvert S_{22} \rvert$ [dB]	-29	-27	/	-28	-10
P_{in-1dB}	-2	-2	3	-7	-3.5
IIP3 [dBm]	12.7	12.9	22	5	4.6
Size [mm²]	0.85	0.85	0.93	/	0.8
Lin on chip?	yes	yes	no	partially	no
FOM [MHz]	89.2	78.7	29.4	19.11	13.8

2.4.2 LNA performance comparison by optimization of passives

An important element in the design of LNA is the use of passives. While SOI technology offers considerable improvement in reduction of losses in passives, thin buried oxide thickness (10s of nm – 1 µm) between the active layer and the substrate poses a problem. Electric and magnetic coupling to substrate cannot be fully avoided. The use of high resistivity substrate mitigates the effect of coupling. Despite this, substrate effects are non-negligible and there is still scope for improvement. Superior quality factor of suspended inductors has already been shown in the previous section.

In our work, we target the substrate removal of inductors in LNA to improve the quality factor. Here, some results of impact of quality factor of inductors on LNA

noise performance are discussed. The first example is based on the simulation work of Shashank et al [73].

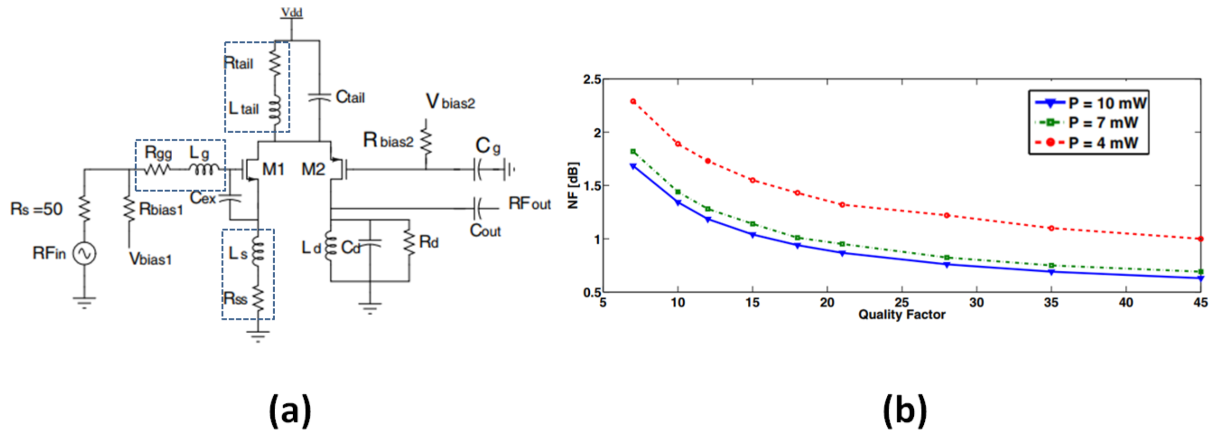


Fig. 2.21: (a) Folded cascode architecture with outlined circuit components representing the different inductors tuned in the design (b) Noise figure as a function of quality factor [73]

Fig. 2.21a shows the folded cascode architecture used for the LNA comprising of four inductors. The inductor on the input side L_g (gate inductor) is very important in determining the noise behaviour of the circuit. As shown in Fig. 2.21b, for different powers of operation, the noise figure can be reduced by ~60% by changing the quality factor of the inductor from 7 to 45. In this simulation, the self resistance of the other inductors is set to zero which doesn't represent a practical situation. Hence, the actual improvement in noise figure may vary. However, it is possible to distinguish the relative importance of different inductors in LNA design.

As shown in Fig. 2.22, when the quality factor of the source degeneration inductor L_s is varied keeping the other inductors ideal with zero resistance, the improvement in noise figure is 0.1 – 0.2 dB. The same observation can be made for variation of quality factor of the tail inductor L_{tail}. These two inductors have relatively less impact on the noise figure of the LNA as compared to the input inductor.

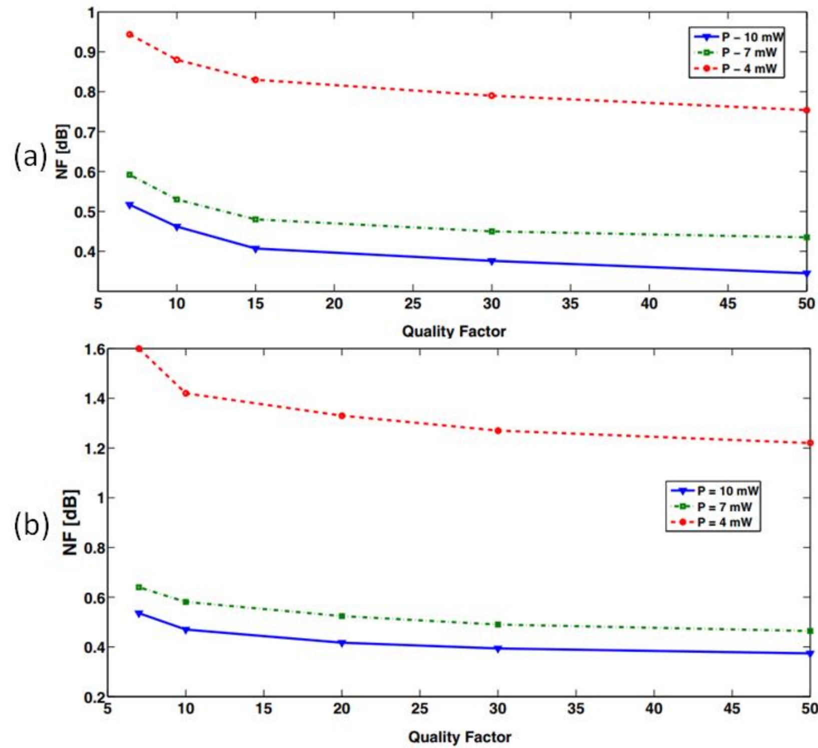


Fig. 2.22: Effect of (a) source degeneration and (b) tail inductor on noise figure of LNA circuit [73]

In another study, Belostotski et al. have described in detail the design of narrowband LNA using integrated gate inductor. Several design methodologies have been proposed and the noise factor has been studied as a function of quality factor of the inductance. As shown in Fig. 2.23, the noise figure reduces with the use of improved quality factor of gate inductance for constrained gain and power designs. The cascode topology of LNA was used in this study. It is also to be noted that with changing inductor Q-factor, the Q_s of the LNA circuit varies. Q_s is not a quality factor of the input network but it represents the relationship between the capacitive and resistive elements of the input network. Hence, the design needs to be appropriately modified to accommodate the optimum Qs requirement in order to meet the design constraints. For example, adjustment of gate source capacitance by using an external capacitance.

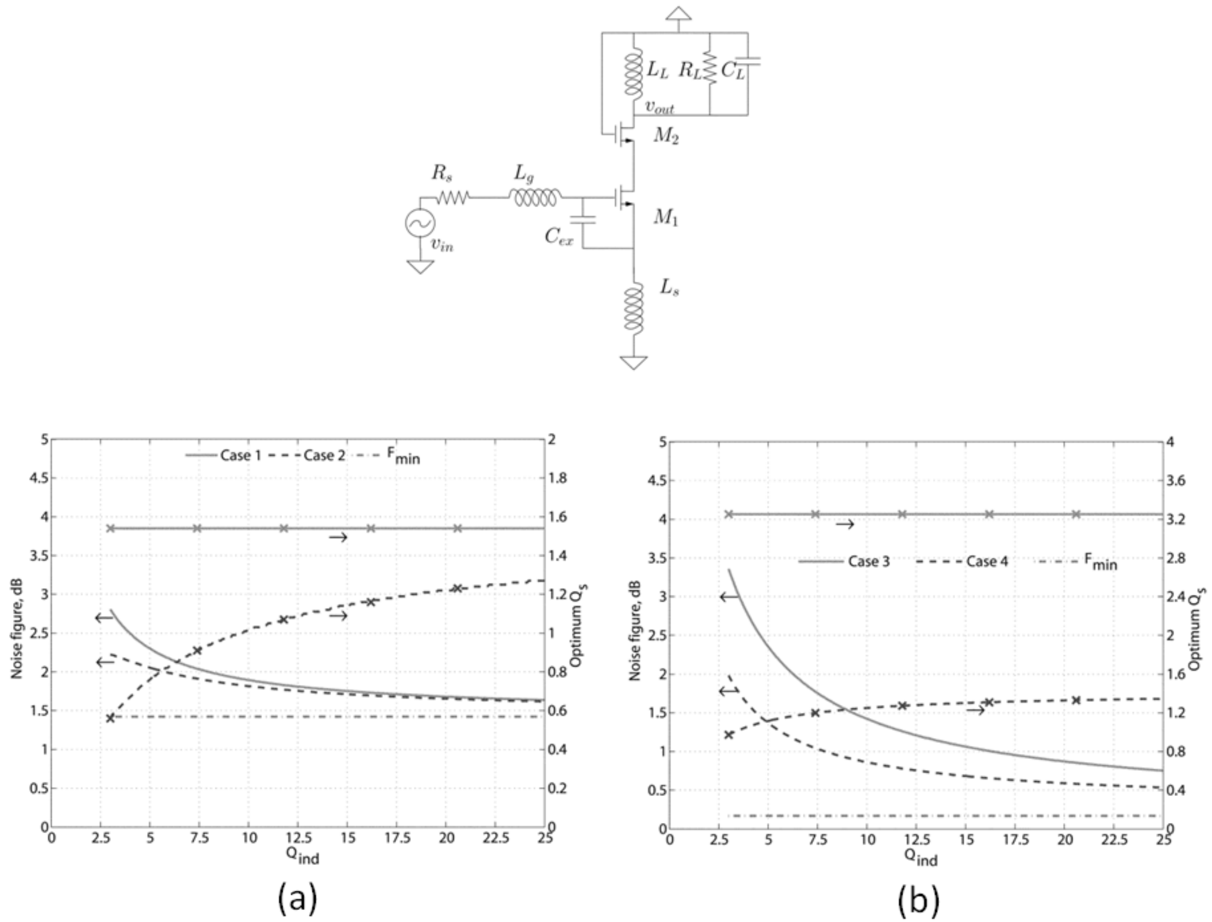


Fig. 2.23: Effect of gate inductor quality factor on noise performance for (a) Constrained gain (b) Constrained power optimization [22]

Concluding remarks

In this chapter, the different processing capabilities offered by femtosecond laser on different scales have been reviewed. It is seen that laser processing is a very flexible tool for both large scale and small scale processing with a high throughput capability. The figures of merit of RF components (inductor, switch and LNA) have been studied. The overall superiority of SOI with respect to CMOS is evident. Additionally, improving the substrate has been shown to improve device performance for different design cases. The case for substrate removal using laser processing for enhanced RF performance has been established. The specific RF functions chosen for substrate removal study are elaborated in Chapter 3. The results pertaining to characterization of functions after substrate removal is presented in Chapter 5.

References

[1] C. Iliescu, H. Taylor, M. Avram, J. Miao, and S. Franssila, "A practical guide for the fabrication of microfluidic devices using glass and silicon," *Biomicrofluidics*, vol. 6, no. 1, p. 016505, Mar. 2012.

[2] X. G. Zhang, *Electrochemistry of Silicon and Its Oxide*. Springer Science & Business Media, 2007.

[3] H. V. Jansen, M. J. de Boer, S. Unnikrishnan, M. C. Louwerse, and M. C. Elwenspoek, "Black silicon method: X. A review on high speed and selective plasma etching of silicon with profile control: an in-depth comparison between Bosch and cryostat DRIE processes as a roadmap to next generation equipment," *J. Micromechanics Microengineering*, vol. 19, no. 3, p. 033001, Mar. 2009.

[4] M. Puech, N. Launay, N. Arnal, P. Godinat, and J. Gruffat, "A Novel Plasma Release Process and a Super High Aspect Rat...," p. 6, 2003.

[5] W. Kern, *Thin Film Processes II*. Elsevier, 2012.

[6] B. Schwartz and H. Robbins, "Chemical Etching of Silicon III . A Temperature Study in the Acid System," *J. Electrochem. Soc.*, vol. 108, no. 4, pp. 365–372, Apr. 1961.

[7] P. K. Guha *et al.*, "Novel design and characterisation of SOI CMOS micro-hotplates for high temperature gas sensors," *Sens. Actuators B Chem.*, vol. 127, no. 1, pp. 260–266, Oct. 2007.

[8] J. Laconte, C. Dupont, D. Flandre, and J.-P. Raskin, "SOI CMOS Compatible Low-Power Microheater Optimization for the Fabrication of Smart Gas Sensors," *IEEE Sens. J.*, vol. 4, no. 5, pp. 670–680, Oct. 2004.

[9] M. E. McNie *et al.*, "Performance enhancement and evaluation of deep dry etching on a production cluster platform," presented at the Micromachining and Microfabrication, San Jose, CA, 2003, p. 34.

[10] A. Y. Vorobyev and C. Guo, "Direct femtosecond laser surface nano/microstructuring and its applications: Direct femtosecond laser surface nano/microstructuring and its applications," *Laser Photonics Rev.*, vol. 7, no. 3, pp. 385–407, May 2013.

[11] B. Dusser *et al.*, "Controlled nanostructrures formation by ultra fast laser pulses for color marking," *Opt. Express*, vol. 18, no. 3, pp. 2913–2924, 2010.

[12] M. Halbwax *et al.*, "Micro and nano-structuration of silicon by femtosecond laser: Application to silicon photovoltaic cells fabrication," *Thin Solid Films*, vol. 516, no. 20, pp. 6791–6795, Aug. 2008.

[13] V. Zorba *et al.*, "Making silicon hydrophobic: wettability control by two-lengthscale simultaneous patterning with femtosecond laser irradiation," *Nanotechnology*, vol. 17, no. 13, pp. 3234–3238, Jul. 2006.

[14] T. Rublack and G. Seifert, "Femtosecond laser delamination of thin transparent layers from semiconducting substrates [Invited]," *Opt. Mater. Express*, vol. 1, no. 4, p. 543, Aug. 2011.

[15] U. Zywietz, A. B. Evlyukhin, C. Reinhardt, and B. N. Chichkov, "Laser printing of silicon nanoparticles with resonant optical electric and magnetic responses," *Nat. Commun.*, vol. 5, no. 1, p. 3402, May 2014.

[16] A. Sikora *et al.*, "Picosecond laser micromachining prior to FIB milling for electronic microscopy sample preparation," *Appl. Surf. Sci.*, vol. 418, pp. 607–615, Oct. 2017.

[17] D. H. Kam, L. Shah, and J. Mazumder, "Femtosecond laser machining of multi-depth microchannel networks onto silicon," *J. Micromechanics Microengineering*, vol. 21, no. 4, p. 045027, Apr. 2011.

[18] E. Coyne, J. P. Magee, P. Mannion, G. M. O'Connor, and T. J. Glynn, "STEM (scanning transmission electron microscopy) analysis of femtosecond laser pulse

induced damage to bulk silicon," *Appl. Phys. A*, vol. 81, no. 2, pp. 371–378, Jul. 2005.
[19] J. Lopez, E. Mottay, C. Loumena, Y. Zaouter, and M. Faucon, "Micromachining of metal and silicon using high average power ultrafast fiber lasers," in *Pacific International Conference on Applications of Lasers and Optics*, Wuhan, People's Republic of China, 2010, p. M503.
[20] J. Schille, L. Schneider, A. Streek, S. Kloetzer, and U. Loeschner, "High-throughput machining using a high-average power ultrashort pulse laser and high-speed polygon scanner," *Opt. Eng.*, vol. 55, no. 9, p. 096109, Sep. 2016.
[21] T. Nubbemeyer *et al.*, "1 kW, 200 mJ picosecond thin-disk laser system," *Opt. Lett.*, vol. 42, no. 7, p. 1381, Apr. 2017.
[22] L. Belostotski and J. W. Haslett, "Noise figure optimization of inductively degenerated CMOS LNAs with integrated gate inductors," *IEEE Trans. Circuits Syst. Regul. Pap.*, vol. 53, no. 7, pp. 1409–1422, 2006.
[23] D. Linten *et al.*, "Low-power 5 GHz LNA and VCO in 90 nm RF CMOS," in *2004 Symposium on VLSI Circuits. Digest of Technical Papers (IEEE Cat. No.04CH37525)*, 2004, pp. 372–375.
[24] Y. K. Koutsoyannopoulos and Y. Papananos, "Systematic analysis and modeling of integrated inductors and transformers in RF IC design," *IEEE Trans. Circuits Syst. II Analog Digit. Signal Process.*, vol. 47, no. 8, pp. 699–713, Aug. 2000.
[25] J. N. Burghartz and B. Rejaei, "On the design of RF spiral inductors on silicon," *IEEE Trans. Electron Devices*, vol. 50, no. 3, pp. 718–729, Mar. 2003.
[26] L. Huang *et al.*, "Analysis and optimum design of RF spiral inductors on silicon substrate," in *2009 3rd IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, Beijing, China, 2009, pp. 990–993.
[27] D. Eggert, P. Huebler, A. Huerrich, H. Kueck, W. Budde, and M. Vorwerk, "A SOI-RF-CMOS technology on high resistivity SIMOX substrates for microwave applications to 5 GHz," *IEEE Trans. Electron Devices*, vol. 44, no. 11, pp. 1981–1989, Nov. 1997.
[28] Dong-Wook Kim *et al.*, "High performance RF passive integration on Si smart substrate," in *2002 IEEE MTT-S International Microwave Symposium Digest (Cat. No.02CH37278)*, Seattle, WA, USA, 2002, vol. 3, pp. 1561–1564.
[29] J. Kim *et al.*, "A Power-Optimized Widely-Tunable 5-GHz Monolithic VCO in a Digital SOI CMOS Technology on High Resistivity Substrate," p. 6.
[30] Y. Sun, H. Van Zejl, J. L. Tauritz, and R. G. F. Baets, "Suspended membrane inductors and capacitors for application in silicon MMIC's," in *IEEE 1996 Microwave and Millimeter-Wave Monolithic Circuits Symposium. Digest of Papers*, San Francisco, CA, USA, 1996, pp. 99–102.
[31] D. Hisamoto, S. Tanaka, T. Tanimoto, and S. Kimura, "Suspended SOI structure for advanced 0.1-μm CMOS RF devices," *IEEE Trans. Electron Devices*, vol. 45, no. 5, pp. 1039–1046, May 1998.
[32] H. Jiang, Y. Wang, J.-L. Yeh, and N. C. Tien, "On-chip spiral inductors suspended over deep copper-lined cavities," *IEEE Trans. Microw. Theory Tech.*, vol. 48, no. 12, pp. 2415–2423, 2000.
[33] L. Woodward, "Fabrication of Novel Suspended Inductors," Master's Thesis, University of Waterloo, 2004.
[34] L. Woodward, P. Woo, M. Capanu, I. Koutsaroff, C. R. Selvakumar, and A. Cervin-Lawry, "Novel High-Q Suspended Inductors on Alumina Ceramic Substrates," *MRS Online Proc. Libr. Arch.*, vol. 833, ed 2004.

[35] J.-B. Yoon, Y.-S. Choi, B.-I. Kim, Y. Eo, and E. Yoon, "CMOS-compatible surface-micromachined suspended-spiral inductors for multi-GHz silicon RF ICs," *IEEE Electron Device Lett.*, vol. 23, no. 10, pp. 591–593, 2002.
[36] C.-H. Chen, Y.-K. Fang, C.-W. Yang, and C. S. Tang, "A deep submicron CMOS process compatible suspending high-Q inductor," *IEEE Electron Device Lett.*, vol. 22, no. 11, pp. 522–523, 2001.
[37] D. Hisamoto, S. Tanaka, T. Tanimoto, Y. Nakamura, and S. Kimura, "Silicon RF devices fabricated by ULSI processes featuring 0.1-/spl mu/m SOI-CMOS and suspended inductors," in *1996 Symposium on VLSI Technology. Digest of Technical Papers*, 1996, pp. 104–105.
[38] J. Y. Park and M. G. Allen, "Packaging-compatible high Q microinductors and microfilters for wireless applications," *IEEE Trans. Adv. Packag.*, vol. 22, no. 2, pp. 207–213, 1999.
[39] S. Pinel, F. Cros, S. Nuttinck, S.-W. Yoon, M. G. Allen, and J. Laskar, "Very high-Q inductors using RF-MEMS technology for System-On-Package wireless communication integrated module," in *IEEE MTT-S International Microwave Symposium Digest, 2003*, 2003, vol. 3, pp. 1497–1500.
[40] Y.-S. Choi, E. Yoon, and J.-B. Yoon, "Encapsulation of the micromachined air-suspended inductors," in *IEEE MTT-S International Microwave Symposium Digest, 2003*, 2003, vol. 3, pp. 1637–1640.
[41] M.-C. Hsieh, Y.-K. Fang, C.-H. Chen, S.-M. Chen, and W.-K. Yeh, "Design and fabrication of deep submicron CMOS technology compatible suspended high-Q spiral inductors," *IEEE Trans. Electron Devices*, vol. 51, no. 3, pp. 324–331, 2004.
[42] J. Zeng, A. J. Pang, C. H. Wang, and A. J. Sangster, "Flip chip assembled MEMS inductors," *Electron. Lett.*, vol. 41, no. 8, pp. 480–481, 2005.
[43] P. J. Bell, V. M. Bright, and Z. Popovic, "Micro-bias-tees using micromachined flip-chip inductors," in *IEEE MTT-S International Microwave Symposium Digest, 2003*, 2003, vol. 1, pp. 491–494.
[44] X.-N. Wang, X.-L. Zhao, Y. Zhou, X.-H. Dai, and B.-C. Cai, "Fabrication and performance of a novel suspended RF spiral inductor," *IEEE Trans. Electron Devices*, vol. 51, no. 5, pp. 814–816, 2004.
[45] G. M. Rebeiz and J. B. Muldavin, "RF MEMS switches and switch circuits," *IEEE Microw. Mag.*, vol. 2, no. 4, pp. 59–71, Dec. 2001.
[46] R. V. Garver, "Microwave diode control devices," *Dedham Mass Artech House Inc 1976 380 P*, 1976.
[47] A. Tombak, M. S. Carroll, D. C. Kerr, J.-B. Pierres, and E. Spears, "Design of high-order switches for multimode applications on a silicon-on-insulator technology," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 10, pp. 3639–3649, 2013.
[48] C. Tinella *et al.*, "0.13/spl mu/m CMOS SOI SP6T antenna switch for multi-standard handsets," in *Digest of Papers. 2006 Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 2006, pp. 4–pp.
[49] P. Hindle, "The State of RF and Microwave Switches." [Online]. Available: https://www.microwavejournal.com/articles/10269-the-state-of-rf-and-microwave-switches?v=preview. [Accessed: 01-Apr-2019].
[50] K. Yu, G. Zhang, L. Huang, J. Lin, and Z. Zhang, "Effects and contrasts of silicon-on-insulator floating-body and body-contacted field-effect transistors to the design of high-performance antenna switches," *IET Microw. Antennas Propag.*, vol. 10, no. 5, pp. 507–516, Apr. 2016.
[51] H. Xu, "A 31.3-dBm bulk CMOS T/R switch using stacked transistors with sub-design-rule channel length in floated p-wells," *IEEE J. Solid-State Circuits*, vol. 42, no. 11, pp. 2528–2534, 2007.

[52] Minsik Ahn, Chang-Ho Lee, Byung Sung Kim, and J. Laskar, "A High-Power CMOS Switch Using A Novel Adaptive Voltage Swing Distribution Method in Multistack FETs," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 4, pp. 849–858, Apr. 2008.
[53] X. S. Wang and C. P. Yue, "A Dual-Band SP6T T/R Switch in SOI CMOS With 37-dBm P(-1dB) for GSM/W-CDMA Handsets," *IEEE Trans. Microw. Theory Tech.*, vol. 62, no. 4, pp. 861–870, 2014.
[54] A. Tombak *et al.*, "Cellular antenna switches for multimode applications based on a Silicon-on-Insulator technology," in *2010 IEEE Radio Frequency Integrated Circuits Symposium*, 2010, pp. 271–274.
[55] D. Im and K. Lee, "Characterization and optimization of partially depleted SOI MOSFETs for high power RF switch applications," *Solid-State Electron.*, vol. 90, pp. 94–98, 2013.
[56] J. Cui, L. Chen, and Y. Liu, "Monolithic Single-Pole Sixteen-Throw T/R Switch for Next-Generation Front-End Module," *IEEE Microw. Wirel. Compon. Lett.*, vol. 24, no. 5, pp. 345–347, May 2014.
[57] C. Tinella, J. M. Fournier, D. Belot, and V. Knopik, "A 0.7dB Insertion Loss CMOS – SOI Antenna Switch with more than 50dB Isolation over the 2.5 to 5GHz Band," p. 5.
[58] M.-T. Lai and H.-W. Tsao, "Ultra-low-power cascaded CMOS LNA with positive feedback and bias optimization," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 5, pp. 1934–1945, 2013.
[59] D. Linten *et al.*, "A 5-GHz fully integrated ESD-protected low-noise amplifier in 90-nm RF CMOS," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1434–1442, 2005.
[60] H. Hsieh and L. Lu, "Design of Ultra-Low-Voltage RF Frontends With Complementary Current-Reused Architectures," *IEEE Trans. Microw. Theory Tech.*, vol. 55, no. 7, pp. 1445–1458, Jul. 2007.
[61] H. Hsieh, J. Wang, and L. Lu, "Gain-Enhancement Techniques for CMOS Folded Cascode LNAs at Low-Voltage Operations," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 8, pp. 1807–1816, Aug. 2008.
[62] and J. Gil, and, and and and, "Complete high-frequency thermal noise modeling of short-channel MOSFETs and design of 5.2-GHz low noise amplifier," *IEEE J. Solid-State Circuits*, vol. 40, no. 3, pp. 726–735, Mar. 2005.
[63] and S. Shekhar and D. J. Allstot, "G/sub m/-boosted common-gate LNA and differential colpitts VCO/QVCO in 0.18-/spl mu/m CMOS," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2609–2619, Dec. 2005.
[64] J. S. Walling, S. Shekhar, and D. J. Allstot, "A g/sub m/-Boosted Current-Reuse LNA in 0.18/spl mu/m CMOS," in *2007 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, 2007, pp. 613–616.
[65] X. Yu and N. M. Neihart, "Analysis and design of a reconfigurable multimode low-noise amplifier utilizing a multitap transformer," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 3, pp. 1236–1246, 2013.
[66] K.-H. Chen and S.-I. Liu, "Inductorless wideband CMOS low-noise amplifiers using noise-canceling technique," *IEEE Trans. Circuits Syst. Regul. Pap.*, vol. 59, no. 2, pp. 305–314, 2012.
[67] N. M. Neihart, J. Brown, and X. Yu, "A dual-band 2.45/6 GHz CMOS LNA utilizing a dual-resonant transformer-based matching network," *IEEE Trans. Circuits Syst. Regul. Pap.*, vol. 59, no. 8, pp. 1743–1751, 2012.
[68] R. Brederlow, W. Weber, J. Sauerer, S. Donnay, P. Wambacq, and M. Vertregt, "A mixed-signal design roadmap," *IEEE Des. Test Comput.*, vol. 18, no. 6, pp. 34–46, 2001.
[69] R. Paulin, P. Cathelin, G. Bertrand, A. Monroy, J. More-He, and T. Schwartzmann, "A 12.7dBm IIP3, 1.34dB NF, 4.9GHz–5.9GHz 802.11a/n LNA in 0.13 μm PD-SOI

CMOS with Body-Contacted transistor," in *2016 IEEE MTT-S International Microwave Symposium (IMS)*, San Francisco, CA, 2016, pp. 1–3.

[70] B.-K. Kim, D. Im, J. Choi, and K. Lee, "A highly linear 1 GHz 1.3 dB NF CMOS low-noise amplifier with complementary transconductance linearization," *IEEE J. Solid-State Circuits*, vol. 49, no. 6, pp. 1286–1302, 2014.

[71] A. Madan, M. J. McPartlin, C. Masse, W. Vaillancourt, and J. D. Cressler, "A 5 GHz 0.95 dB NF Highly Linear Cascode Floating-Body LNA in 180 nm SOI CMOS Technology," *IEEE Microw. Wirel. Compon. Lett.*, vol. 22, no. 4, pp. 200–202, Apr. 2012.

[72] H. Noori, M. Sanner, and N. Yanduru, "A 0.8 dB NF, 4.6 dBm IIP3, 1.8–2.2 GHz, low-power LNA in 130 nm RF SOI CMOS technology," in *2015 Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS)*, 2015, pp. 1–4.

[73] S. Tiwari, V. N. R. Vanukuru, and J. Mukherjee, "Noise figure analysis of 2.5 GHz folded cascode LNA using high-Q layout optimized inductors," in *2015 IEEE Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics (PrimeAsia)*, 2015, pp. 94–97.

[74] White Paper: Innovative RF-SOI Wafers for Wireless Applications, SOITEC, 2013.

[75] T.-Y. Lee and S. Lee, "Modeling of SOI FET for RF switch applications," in 2010 IEEE Radio Frequency Integrated Circuits Symposium, 2010, pp. 479–482.

Chapter 3: Laser processing applied to RF functions: Theoretical background and experimental methodology

3.0 Introduction

This chapter covers the essential theoretical concepts and background literature relevant the thesis. In section 3.1, the basic definitions concerning pulsed laser processing are covered. Following this, a brief description of the different phenomena occurring as a result of laser matter interaction is presented in section 3.2. Here, the special advantage of using femtosecond laser is highlighted as well. In section 3.3, the effect of different experimental conditions viz. ambient conditions, wavelength, pulse width and polarization on laser processing is highlighted. An additional aspect of effect of ablation plume on laser processing is discussed in section 3.4. Following this, a detailed description of the laser processing system used in this work is presented highlighting the processing options in section 3.5. An experimental plan is outlined in section 3.6 to systematically study the laser processing parameters. Finally in section 3.7, the description of different circuits that are studied is presented.

Contents

3.1 Laser processing: Essential concepts....79
3.1.1 Gaussian beam parameters....79
3.1.2 Pulsed laser processing parameters....81
3.2 Laser processing: Description of physical processes and timescales....85
3.3 Effect of different parameters on laser processing of silicon....89
3.3.1 Ambient conditions....90
3.3.1.1 Medium of ablation....90
3.3.1.1 Substrate temperature....91
3.3.2 Laser wavelength....92
3.3.3 Laser pulse width....94
3.3.3.1 Beam distortion effect....94
3.3.3.2 Modification of ablation threshold....95
3.3.4 Polarization....96
3.4 Effect of ablation plume on laser processing....97
3.5 Laser processing: Description of system....100
3.5.1 Laser source and optics....100
3.5.2 Optical attenuator....101
3.5.3 Trepan head....102
3.5.4 Galvanometric scanner....102
3.6 Laser processing: Experimental plan....102
3.7 RF circuits: Description and theory....104
3.7.1 Isolation structures....105
3.7.2 RF Switch....106
3.7.3 Inductors....108
3.7.4 Low Noise Amplifier (LNA)....109
Conclusion....112
References....113

3.1 Laser processing: Essential concepts

3.1.1 Gaussian beam parameters

The laser beam that is most commonly used in laser processing is a Gaussian beam. When the intensity distribution of the beam is observed, by taking its cross section, it features a Gaussian profile. The expression for intensity of a Gaussian beam propagating in the z-direction is given by:

$$I(z,r) = I_0 \left(\frac{w_0}{w(z)}\right)^2 e^{-\frac{2r^2}{(w(z))^2}} \qquad (3.1)$$

$I_0 -$ *peak intensity of the beam at* $z = 0, r = 0$ $(W\ m^{-2})$

$w_0 -$ *beam waist (also referred to as spot* size*/radius)* (m)

The value of the beam waist depends on the lens used in focussing the collimated laser beam, the beam quality factor and the light wavelength. The beam propagation is illustrated in Fig. 3.1. The expression for w_0 is given by [1]:

$$w_0 = \frac{M^2 \lambda f_l}{\pi r_b} \qquad (3.2)$$

$M^2 -$ *Beam quality factor (typically* $1.1 - 1.7$ *for diode laser beams)*

$\lambda -$ *Laser beam wavelength* (m)

$f_l -$ *Focal length of the lens* (m)

$r_b -$ *Beam radius at the input of the lens* (m)

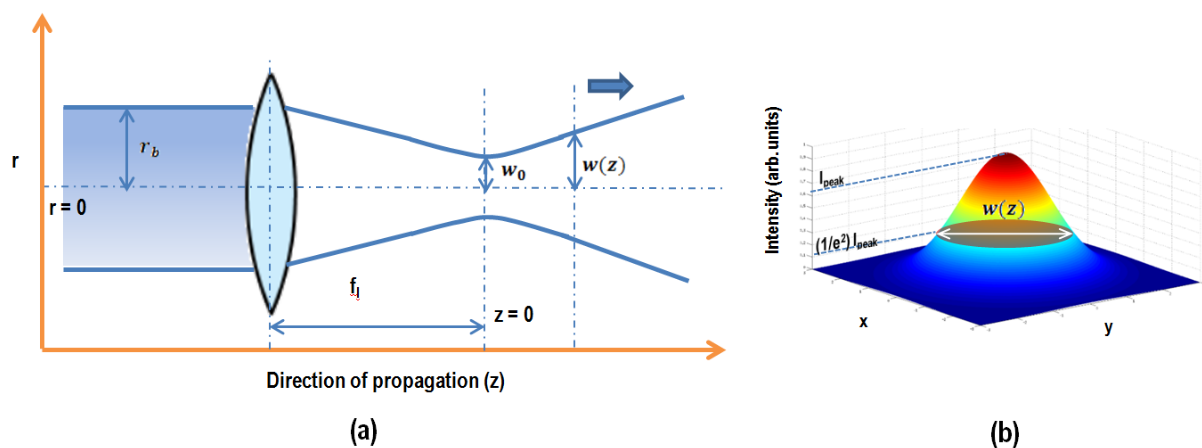


Fig. 3.1: Gaussian beam description showing (a) the variation of beam width along the propagation direction (b) The radial distribution of intensity at an arbitrary z-position

The beam width increases on either side of the beam waist along the direction of propagation. The beam width can be calculated as:

$$w(z) = w_0\sqrt{1 + \left(\frac{z}{z_R}\right)^2} \tag{3.3}$$

$$z_R = \frac{\pi w_0^2}{M^2 \lambda} \tag{3.4}$$

Here z_R is defined as the Rayleigh length which is used to distinguish the far-field and near-field regions of the Gaussain beam. In the near field region ($z \ll z_R$), the contribution from the squared term in (3.3) can be neglected which gives $w(z) \approx w_0$. In the far field region $z \gg z_R$, the beam radius can be approximated as $w(z) = \frac{w_0 z}{z_R}$. In this region, the beam width is a linear function of the distance.

During laser ablation, when material is to be removed over a certain area using multiple passes, another important parameter to be defined is the depth of focus. During every pass, a certain thickness of material is removed and hence the workpiece will no longer be at focus. To check if the surface of the workpiece is within reasonable limits from the focal plane (waist) of the beam, expression (3.3) can be used. A reasonable limit that can be defined for focus is that the beam width at the surface of the workpiece does not exceed $1.05w_0$. Based on this condition, depth of focus (DOF) can be defined as:

$$DOF_{5\%} = \frac{0.32\pi w_0^2}{\lambda} \tag{3.5}$$

If the depth of material to be ablated is less than the depth of focus, then the variation in laser fluence between different passes remains negligible. However, if the depth of the ablated material is significantly higher than the depth of focus, then the lens needs to be lowered between the different passes in order to maintain the focus condition on the workpiece. If the lens position remains unchanged, the fluence and pulse overlap conditions change for each pass which impacts the laser processing quality and ablation rates.

Both Rayleigh range and DOF can be used to determine if the beam is out of focus at the surface. We use the Rayleigh Range to determine if the beam is reasonably focused on the surface.

3.1.2 Pulsed laser processing parameters

The Gaussian beam parameters define the propagation of the beam when the laser output is on. In a pulsed laser system, the laser energy is delivered in pulses at a certain repetition rate. The average power delivered by the laser can be calculated as:

$$P_{avg} = f_{rep}E_p \tag{3.6}$$

$f_{rep} -$ *Pulse repetition rate* (Hz)

$E_p -$ *Energy per laser pulse* (J)

It can be noticed that to deliver a certain average power, the pulse repetition rate and the pulse energy can be varied independently of each other. However, laser processing at low pulse energy and high repetition rate can be significantly different as compared with conditions of high pulse energy and low repetition rate. This is because of varying conditions of fluence and temporal spacing between two pulses.

It is instructive to look at how the energy is distributed during each pulse at the surface of the workpiece by defining the laser fluence. The expression (3.1) can be modified to represent a time averaged intensity or in other words energy distribution over the surface area. This gives the expression for laser fluence as:

$$\phi(r) = \phi_0 e^{-\left(\frac{2r^2}{w_0^2}\right)} \tag{3.7}$$

$\phi_0 -$ *Peak laser fluence at the axis of the beam* $(J\ cm^{-2})$

$$\phi_0 = \frac{2E_P}{\pi w_0^2} \tag{3.8}$$

The expression for peak fluence is obtained by integrating equation (3.7) over the entire surface and equating it to the pulse energy. Sometimes the word peak is omitted from the term and the word fluence is used. Both refer to the same quantity.

Laser fluence influences the regime of ablation and the ablation rate. The material modification begins at a certain threshold value of incident fluence (ϕ_{thm}). At this fluence, different material modifications like amorphization, oxidation etc. takes place, but there is no material ablation. When the laser

fluence is further increased to the ablation threshold (ϕ_{tha}), material ablation occurs. In literature, the most commonly quoted threshold is the ablation threshold. In literature, sometimes averaged threshold fluence is used instead of the peak threshold fluence value which is given by:

$$\phi_{av} = \frac{E_P}{\pi w_0^2} \tag{3.9}$$

This value is obtained by making a simplifying assumption of even distribution of pulse energy E_P over the surface area which is a circle with radius w_0. The use of peak fluence is more rigorous and we use this value in fluence calculations hereinafter.

Thus far, the energy distribution over the surface area during a single pulse has been discussed. During laser micromachining, the workpiece is set in motion relative to the laser beam either by displacement of the stage or the mirrors in the galvanometric scanner. Thus, successive laser pulses are centred at different spatial locations with a certain overlap from one pulse to the next. It is thus necessary to take into account the cumulative effect of multiple pulses on a single spatial location resulting from the convolution of the shooting path and spatial beam shape. This is illustrated in Fig. 3.2.

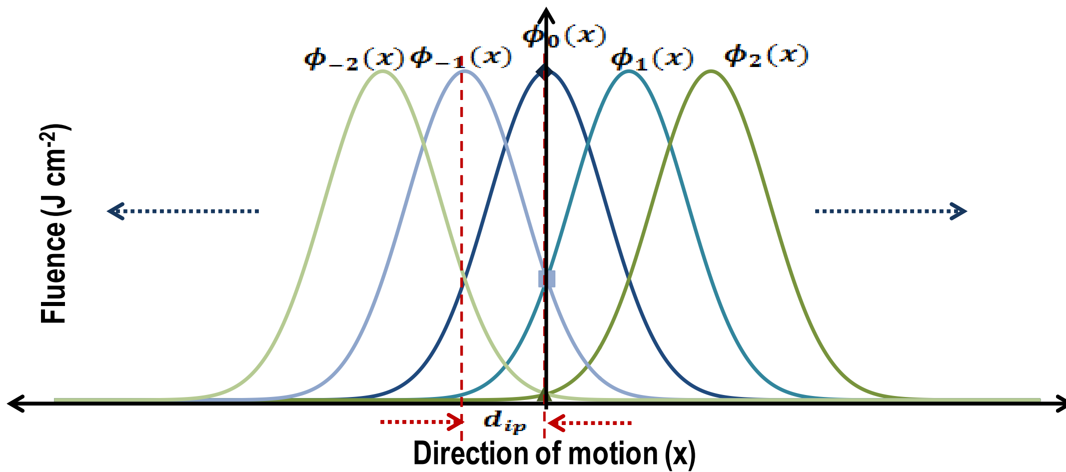


Fig. 3.2: Overlap of pulse energy between successive pulses shown for displacement along a single direction of translation. The markers at $x = 0$ shows the fluence contributions from previous pulses and future pulses

Consider a Gaussian pulse which is centred at $x = 0$ which occurs at $t = t_0$. If the time between pulses is $t_{ip} = \frac{1}{f_{rep}}$, and the velocity of beam is v_b, the distance between successive pulses would be $d_{ip} = t_{ip}v_b$. The total accumulated fluence at $x = 0$ is given by:

$$\phi_{acc} = \ldots\phi_{-2}(0) + \phi_{-1}(0) + \phi_0(0) + \phi_1(0) + \phi_2(0)\ldots$$

$$\phi_N(x) = \phi(x - Nd_{ip})\ occurs\ at\ t = t_0 + Nt_{ip}\ |\ \phi(x) = \phi_0 e^{-\left(\frac{2x^2}{w_0^2}\right)}$$

$$\phi_{acc} = \cdots \phi_0 e^{-\left(\frac{2(2d_{ip})^2}{w_0^2}\right)} + \phi_0 e^{-\left(\frac{2d_{ip}{}^2}{w_0^2}\right)} + \phi_0 + \phi_0 e^{-\left(\frac{2d_{ip}{}^2}{w_0^2}\right)} + \phi_0 e^{-\left(\frac{2(2d_{ip})^2}{w_0^2}\right)} \ldots$$

$$\phi_{acc} = \phi_0\left(1 + 2\left(e^{-\left(\frac{2d_{ip}{}^2}{w_0^2}\right)} + e^{-\left(\frac{2(2d_{ip})^2}{w_0^2}\right)} + \cdots\right)\right)$$

$$\phi_{acc} = \phi_0\left(\sum_{k=-\infty}^{\infty} e^{-\mathrm{c}\mathrm{k}^2}\right)\quad where\ c = \frac{2d_{ip}^2}{w_0^2} \tag{3.10}$$

For conditions of overlap much smaller than the beam waist ($d_{ip} \ll w_0$), a very good approximation of this expression has been provided by Crawford et al. [2].

$$\phi_{acc} = \left(\sqrt{\frac{\pi}{2}}\frac{w_0}{d_{ip}}\right)\phi_0 \tag{3.11}$$

The accumulated fluence is a function of the ratio of the interpulse spacing to the beam waist $\frac{d_{ip}}{w_0}$. As seen in Fig. 3.3, at $\frac{d_{ip}}{w_0} < 0.25$, the accumulated fluence is very sensitive to the interpulse distance, while at higher values of the ratio, the sensitivity is much lower.

Thus far, the discussion is made with respect to calculation of fluence assuming an unchanging plane surface. However, during each pulse, there is ablation and removal of material. As a result of continually changing surface area due to ablation, the conditions of fluence do not remain the same for different pulses. However, expression (3.11) gives an approximate method to calculate a single value for the fluence that is in effect during the micromachining process.

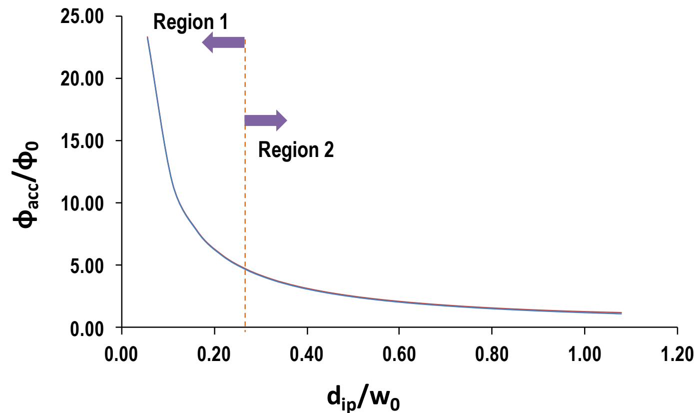


Fig. 3.3: Dependence of accumulated fluence on the ratio of interpulse distance and beam waist showing regions of high sensitivity (region 1) and low sensitivity (region 2)

During laser micromachining, one more important parameter to consider is the incubation effect. Incubation effect refers to the dependence of threshold fluence on the number of laser pulses. If multiple pulses of laser are used on the same spot, the threshold fluence required for ablation of material reduces. This effect is observed on many materials and an example is shown in Fig. 3.4. The dependence of threshold fluence (ϕ_{th}) on the number of laser pulses (N) is given by the following equation:

$$\phi_{th}(N) = \phi_{th}(1)N^{\zeta-1} \tag{3.12}$$

$\zeta - Incubation\ coefficient$

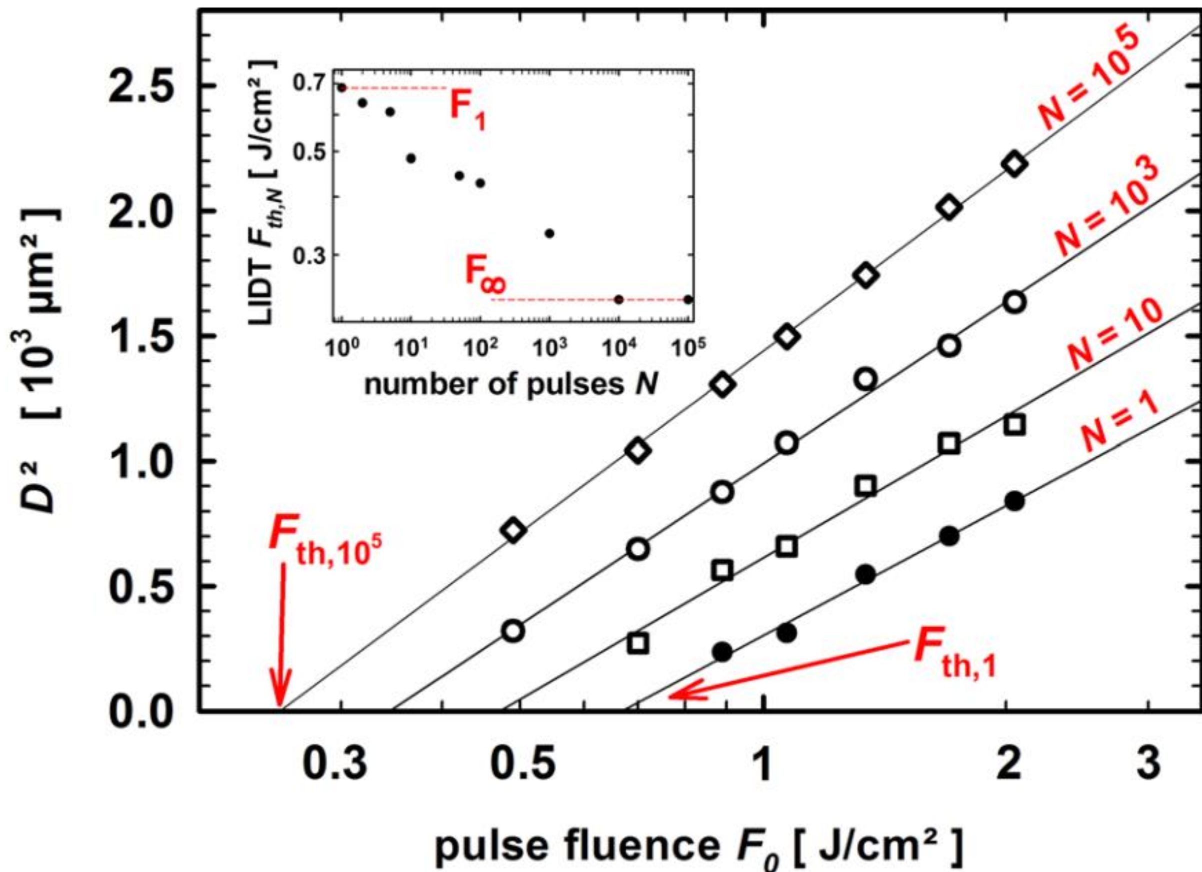


Fig. 3.4: Squared diameter of crater plotted as a function of pulse fluence to estimate threshold fluence for different number of pulses on TiO_2 at repetition rate of 1 kHz. The inset shows the estimated threshold values as a function of number of pulses. [3]

Based on the value of ζ, three types of material responses to laser radiation can be distinguished [4] :

1) $\zeta < 1$ –Incubation effects are present
2) $\zeta = 1$ –No dependence of threshold fluence on number of pulses
3) $\zeta > 1$ –Material hardening with increasing number of pulses

For almost all materials ζ value is < 1 and hence incubation effect needs to be considered. Because of the incubation effect, the ablation process is accelerated due to the lowering of the ablation threshold with successive pulses. Expression (3.12) holds good for stationary multi-pulse ablation. This expression needs to be adapted for dynamic micromachining case. The number of impulsions can be deduced as $\frac{\phi_{acc}}{\phi_0}$. From the experimental graph of ablation diameter as a function of laser fluence, ζ can be determined by curve fitting. While $\phi_{th}(N)$ can be experimentally determined, the theoretical description taking into account incubation effect is involved. Hence, we make use of the concepts of accumulated fluence and number of pulses and in our experiments.

3.2 Laser processing: Description of physical processes and timescales

Laser processing systems have different timescales of operation and one of the most important parameter to consider is the pulse width of the laser source. In our study, we are interested in the relevant processes for ultrashort pulse duration. While there is not a strict definition of the term "ultrashort", it is common to refer to laser beam of pulse width < 1 ps as ultrashort. For such short pulse widths, the dynamics of electron and lattice subsystems can be decoupled. For ultrashort duration, the commonly used model to describe the laser-material interaction is the Two Temperature Model (TTM) [5]. In this model, the material is depicted by two distinct electronic and lattice systems as shown in Fig.3.5 (a, b). Each system can be described by its individual temperature: electrons with temperature T_e and phonons with temperature T_l. These temperatures are thermodynamic quantities which describe the thermal distribution of electrons and phonons. They undergo changes upon interaction with incident laser energy.

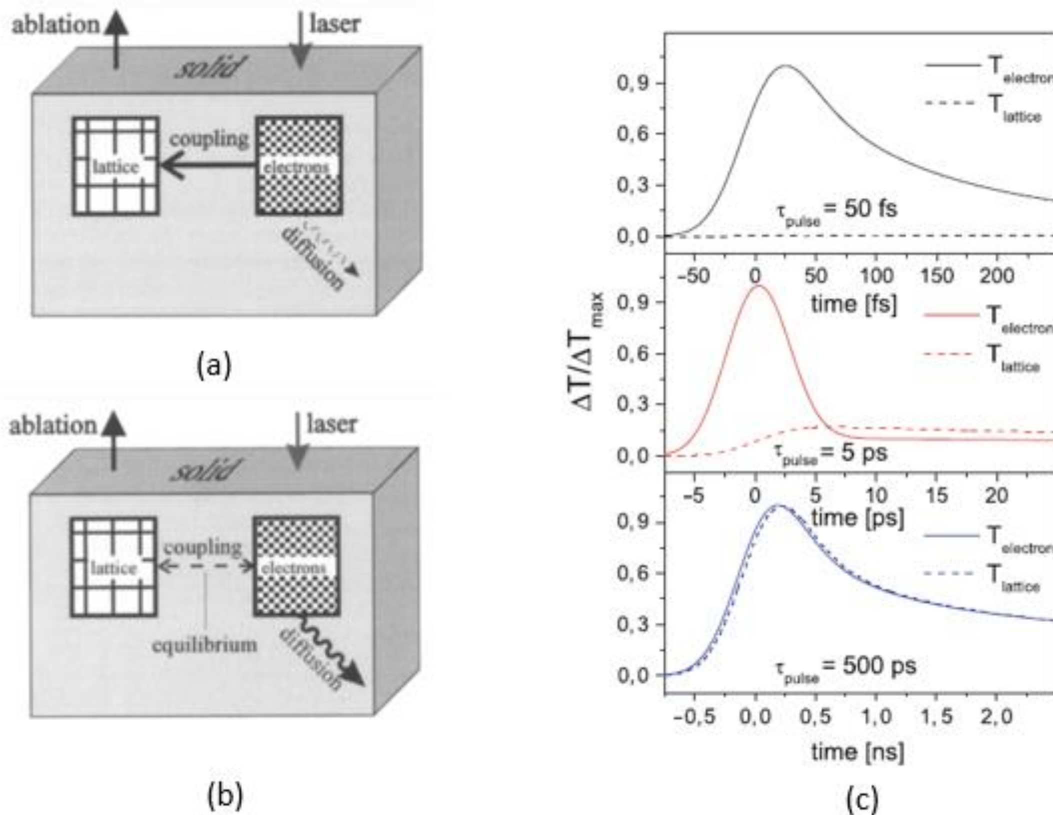


Fig. 3.5: Schematic representation of interaction between the electronic and lattice systems for (a) Femtosecond laser (b) Nanosecond laser (c) Normalized temperature profiles obtained for copper upon laser irradiation for different pulse durations using the TTM [6], [7]

The important feature of ultrashort pulse duration is the decoupling of electronic and lattice systems. As seen in Fig. 3.5c, for pulse duration of 50 fs, the temperature of the electronic system responds to the applied pulse while the lattice temperature remains unperturbed until 300 fs. After several ps, a part of this energy is coupled to the lattice but the temperature rise is not significant. The resulting ablation is referred to as *cold ablation* where there is minimal heating of material due to laser irradiation. In case of pulse durations of the order tenths of nanosecond and higher, the lattice and electronic temperatures are in equilibrium with each other which results in a different physical pathway for ablation called *melt expulsion* which involves melting of material [8].

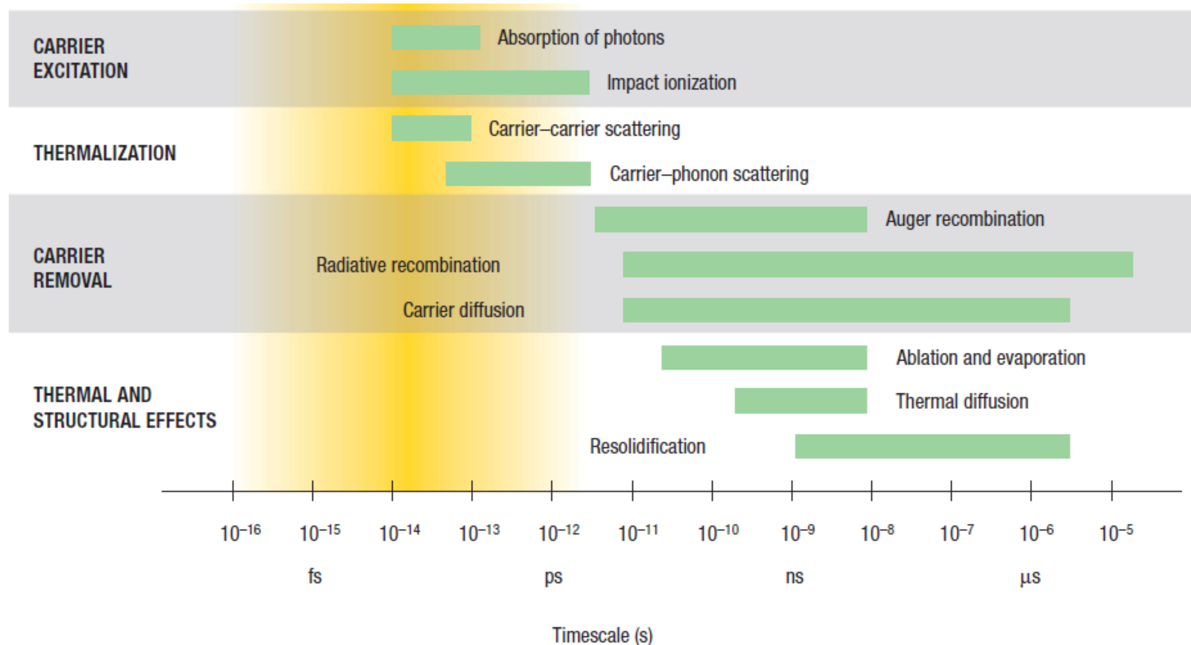


Fig. 3.6: The summary of physical processes upon ultrashort laser irradiation and the typical timescales of these processes. The yellow region of the graph indicates the pulse duration [9]

The processes occurring upon irradiation of a material with femtosecond laser source are summarized in Fig. 3.6. For a semiconductor, all these processes are schematically illustrated with the help of a band diagram in Fig. 3.7. There are 4 classes of processes occurring upon laser beam irradiation: carrier excitation, thermalization, carrier removal and thermal and structural effects. The first process is carrier excitation where the charge carriers of the material are promoted to higher energy states by different mechanisms of absorption. After that, the electron system thermalizes first by means of carrier-carrier scattering and approach a Fermi-Dirac distribution in the timescale of tens to hundreds of femtoseconds. The carrier-phonon scattering process results in the emission of phonons. The emitted phonons after undergoing several scattering processes reach thermal equilibrium with the condensed solid on a timescale of several picoseconds. During the carrier removal phase, the number of free carriers is reduced due to different recombination mechanisms and diffusion of carriers. The phonons also diffuse in the bulk of the material carrying heat deep into the material and cooling the initial laser affected zone. Finally, the material undergoes physical ablation which results in removal of material. After the ablation is complete, condensation and resolidification of the material takes place and the thermal equilibrium with the surroundings is re-established.

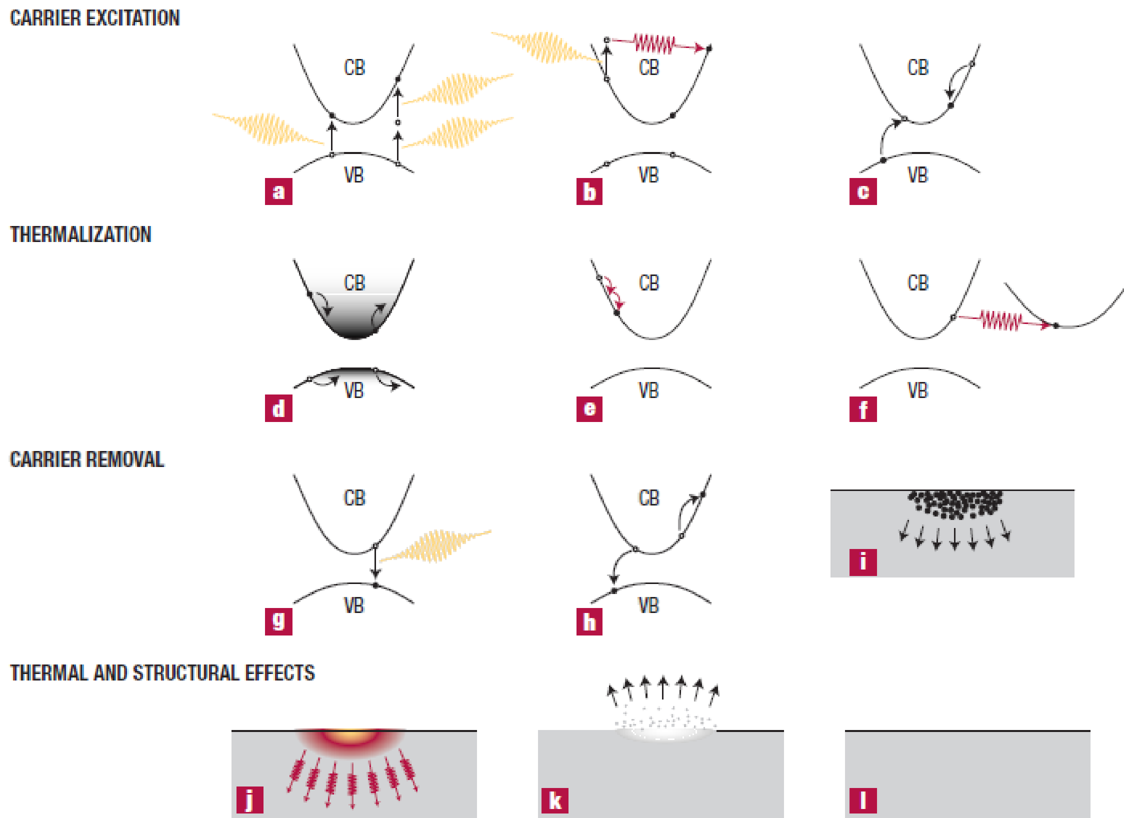


Fig. 3.7: Graphical illustration of sequence of different physical processes due to laser irradiation on a semiconductor (a) Absorption (single and multiphoton) (b) Free carrier absorption/ Inverse bremsstrahlung (c) Impact/Avalanche ionization (d) Non equilibrium carrier distribution after excitation (e) Electron –electron scattering (f) Emission of phonons (g) Radiative recombination (h) Auger recombination (i) Carrier diffusion (j) Thermal diffusion (k) Material ablation (l) Condensation/resolidification [9]

Based on the laser parameters, some processes will be dominant over the others. For instance, absorption of photons due to multi-photon absorption is strongly dependent on the intensity of the laser beam. Hence it is an important absorption mechanism for ultrashort pulses, for which the intensity of the beam is the highest [10]. Also, the timescales of different processes depends on different characteristic times which can vary a lot from one material to another. For example, the thermalization process can be characterized by time t_e for relaxation of electron energy and t_{ph} for relaxation of phonon energy. Using the hyperbolic two temperature model, Kostrykin et al. have calculated these quantities for different metals [11]. Between Au ($t_e = 579\ fs, t_p = 70.6\ ps$) and Al ($t_e = 67\ fs, t_p = 4.6\ ps$), t_e and t_{ph} can vary almost one order of magnitude. While thermalization is just one example, each of the preceding and subsequent processes depends both on the material and laser parameters.

While some of the mechanisms of laser-matter interaction have been introduced in this section, they are not exhaustive. The ablation of material is a complex process to model and will be excluded from the discussion. One important aspect of femtosecond processing is the generation of a high density of charge carriers of the order $10^{22} - 10^{23}$ cm^{-3} by absorption of laser energy. Such high density of charge carriers can cause abrupt changes in the lattice structure on a sub-picosecond timescale which finally leads to ablation on a larger timescale [12]. A detailed description of the ablation mechanisms has been provided by Cheng et al. [13]. The different mechanisms for ablation are broadly listed as spallation [14], phase explosion [15] and critical point phase separation [16] and the dominant mechanism of ablation depends strongly on the fluence.

3.3 Effect of different parameters on laser processing of silicon

This section attempts to outline the necessary background required to understand experimental parameters for laser processing of silicon. There are a large number of parameters that can vary between different laser processing systems and consequently it is necessary to understand how each parameter affects the laser process. This understanding serves two purposes (i) Tuning the parameters of the existing laser system used in this work as per the processing requirements (ii) For parameters which are fixed and not readily tuneable, understanding the consequence of the choice of fixed parameters.

Laser processing involves a complex interplay of several distinct physical processes. These processes that take place during laser irradiation are a direct consequence of the choice of laser processing parameters. A detailed description of the physical processes requires a rigorous treatment and is beyond the scope of this section. The effect of the following parameters on laser processing shall be discussed: ambient conditions, laser pulse width, laser wavelength, polarization. An emphasis is placed on the analysis of threshold fluence as a function of these parameters because this is one of the most important parameters in laser ablation. Wherever possible, the results obtained in literature for silicon are quoted. If they are unavailable, results are presented for metals or other semiconductors as some features are common across these materials. Finally, a brief description of the ablated plume is provided.

3.3.1 Ambient conditions

The surrounding media and the initial substrate temperature constitute the ambient conditions for laser processing. In this subsection, the effect of these two parameters on laser processing is discussed.

3.3.1.1 Medium of ablation

Laser systems have multiple possibilities for medium of ablation and majority of the systems use ambient air atmosphere as the medium. Vacuum is the preferred condition for laser processing when the material degradation in the form of oxidation needs to be avoided. Additionally, vacuum conditions can help better processing for high-aspect ratio etching [17].

Besner et al. studied the diameter of ablation of silicon in vacuum, air and water [18]. This can be used to extrapolate the ablation threshold. The medium does not play a significant role in the low fluence regime (~1 J cm^{-2}). The squared diameter (D^2) follows a linear dependence on the fluence when plotted on a semi-logarithmic scale. However, at higher fluence (~20 J cm^{-2}), the dependence on of D^2 on fluence is no longer linear and varies for all 3 media.

Liu et al. compared the use of water and alcohol as the medium against air [19]. The thermal effects are considerably reduced and the threshold fluence is lowered for the use of alcohol and water as seen in Fig. 3.8. The Gaussian profile for etched craters become clearly visible for water and alcohol at N=500 and N=5000. The ablation depth for same fluence increases for liquids as compared to air in the following order: $L_{water} > L_{alcohol} > L_{air}$.

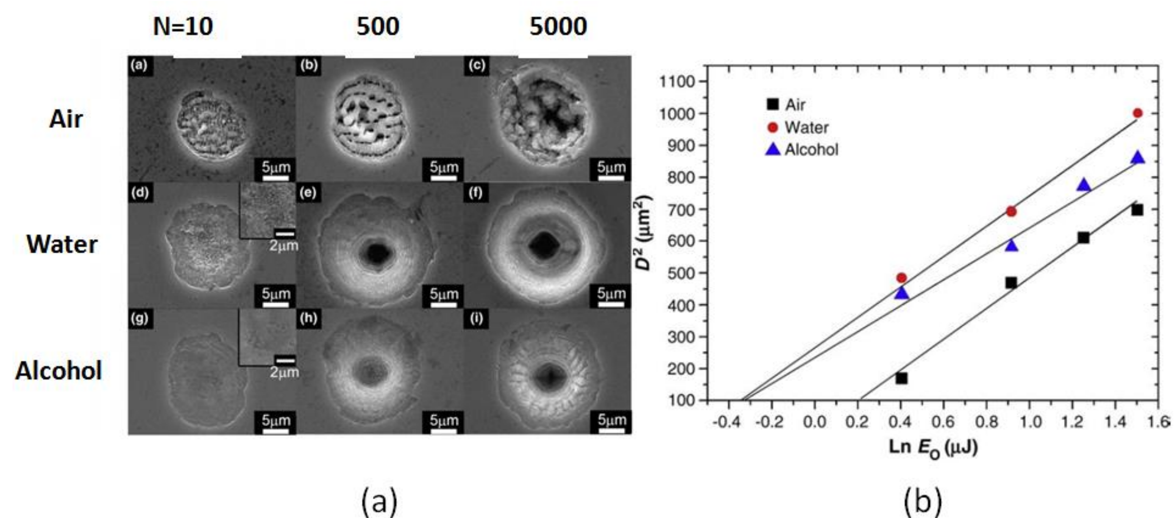


Fig. 3.8: Ablation of silicon in air, alcohol and water (a) SEM micrograph of the obtained crater for different number of impulsions (b) Threshold fluence for Np=5000 for the 3 media by squared diameter regression method [19]

3.3.1.1 Substrate temperature

The effect of substrate temperature on ablation of silicon has been studied in detail by Thorstensen et al. [20]. The pulse width used in this study is 3 ps. In this work, the effect of substrate temperature (25 – 300 °C) on threshold fluence is investigated for 3 different wavelengths: 343 nm, 515 nm and 1030 nm. There is a reduction in the ablation threshold for 1030 nm and 515 nm by 43% and 35% when the temperature is increased from room temperature to 300 °C as shown in Fig. 3.9. It does not change for 343 nm.

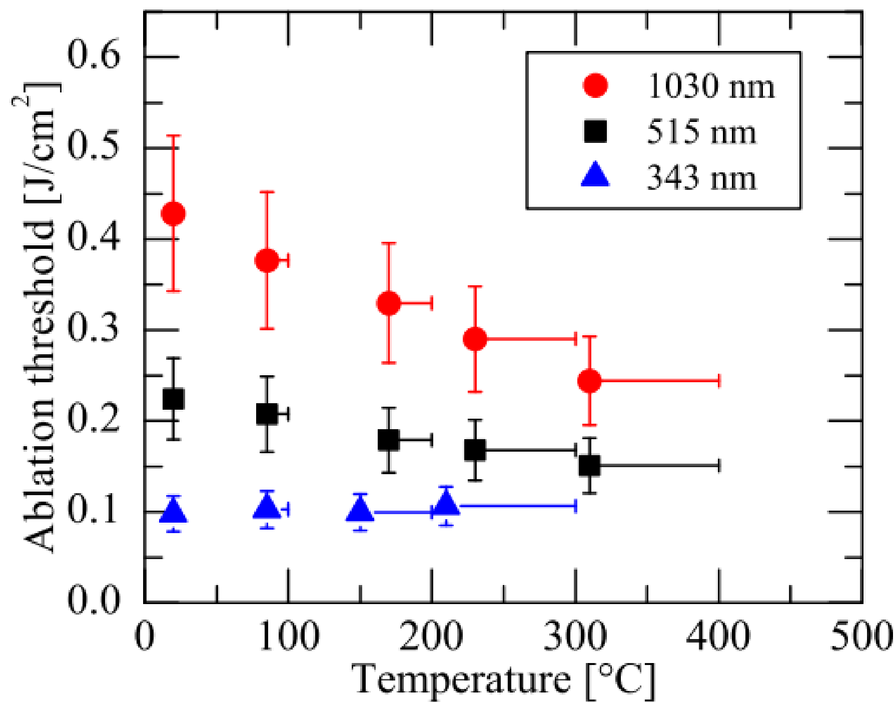


Fig. 3.9: Ablation threshold dependence on the substrate temperature for 3 wavelengths [20]

Simulation of the ablation process also revealed the mechanisms at play at different temperatures for the 3 wavelengths. A two temperature model was used for description of ablation. At room temperature, the dominant physical process is linear absorption for wavelengths of 343 nm and 515 nm and two-photon absorption for 1030 nm. Additionally, at 1030 nm, the free carrier absorption is high which eventually leads to increased energy coupling to the lattice.

The reason for decreased ablation threshold is attributed to the increase in linear absorption coefficient at wavelengths of 515 nm and 1030 nm. Higher absorption coefficient enables reaching of critical electron density required for ablation at smaller fluence thereby lowering the threshold. This temperature dependence is not seen at 343 nm and hence the ablation threshold does not change.

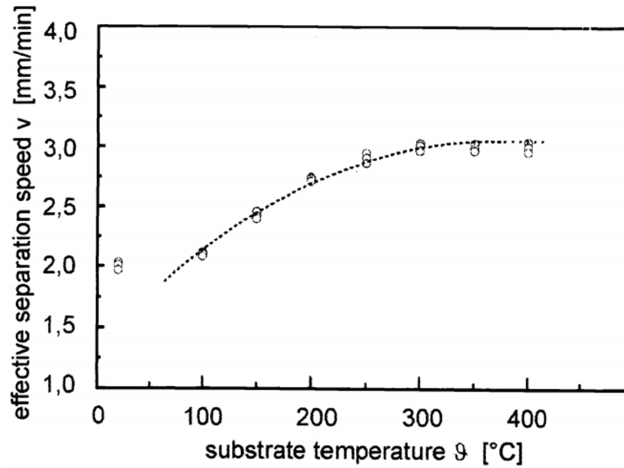


Fig. 3.10: Influence of substrate temperature on effective cutting speed of wafer grade silicon [21]

In the study of Toenshoff et al., different substrate temperatures were utilized for the cutting of silicon. As expected from previous discussion, the effective speed of separation of silicon die increases with increasing substrate temperature while all other parameters remain unchanged (Fig. 3.10). There is nearly a 50% improvement in the effective separation speed by heating the substrate to 300 °C during laser processing.

3.3.2 Laser wavelength

While discussing the effect of substrate temperature, some aspects of the effect of wavelength on the laser source have already been outlined. The reduction of ablation threshold is also seen for femtosecond pulses in the study of InP ablation by [22]. InP is a direct band semiconductor with band gap close to that of silicon and hence similar trends can be expected in the case of silicon.

Table 3.1: Laser ablation crater study depicting ablation threshold, height and volume variation with change in laser wavelength [22]

Wavelength λ (nm)	Pulse width τ_p (fs)	Spot size w_0 (μm)	Ablation threshold φ_{th} (mJ cm^{-2})	Height fitting parameter h_0 (nm)	Volume fitting parameter V_0 (μm^3)
400	175	3.7	73	39	0.28
660	105	4.3	110	44	0.5
800	145	5.4	170	58	0.71
1330	65	6.3	210	61	1.4
2050	100	8.5	260	56	2.8

It is observed in Table 3.1 that ablation threshold increases with increasing wavelength. The ablation threshold affects the width, depth and volume of ablation given by the following expressions:

$$W_{abl} = \sqrt{2w_0^2 \ln\left(\frac{\phi_0}{\phi_{th}}\right)} \tag{3.13}$$

$$h_{abl} = h_0 \ln\left(\frac{\phi_0}{\phi_{th}}\right) \tag{3.14}$$

$$V_{abl} = V_0 \left(\ln\left(\frac{\phi_0}{\phi_{th}}\right)\right)^2 \; where \; V_0 = \frac{\pi\omega_0^2 h_0}{4} \tag{3.15}$$

With increasing wavelength, the ablation threshold increases, and hence for the same peak fluence, the logarithmic term has larger values. Hence, it could be expected that reducing wavelength increases the depth and volume of ablation. However, this is not the case. The fitting parameters for height (h_0) and volume (V_0) are also functions of wavelength. As seen in Table 3.1, both these values increase with increasing wavelength. The overall effect is that increasing the wavelength actually increases the ablation depth and volume at the same fluence. Hence the ablation efficiency which is the volume of matter removed per unit incident energy also increases.

The same observation has been reported for ablation of silicon on a 50 ps-laser source by Sikora et al. [23]. The threshold fluences at 343 nm, 515 nm and 1030 nm are 0.01 J cm^{-2}, 0.15 J cm^{-2}, and 0.83 J cm^{-2} respectively. Despite large differences in ablation threshold, the ablated volume at the same normalized fluence increases for increasing wavelength as observed in Fig. 3.11. And also the ablation efficiency expressed as volume ablated per unit of incident laser energy is also much higher at 1030 nm source as compared to the other 2 wavelengths.

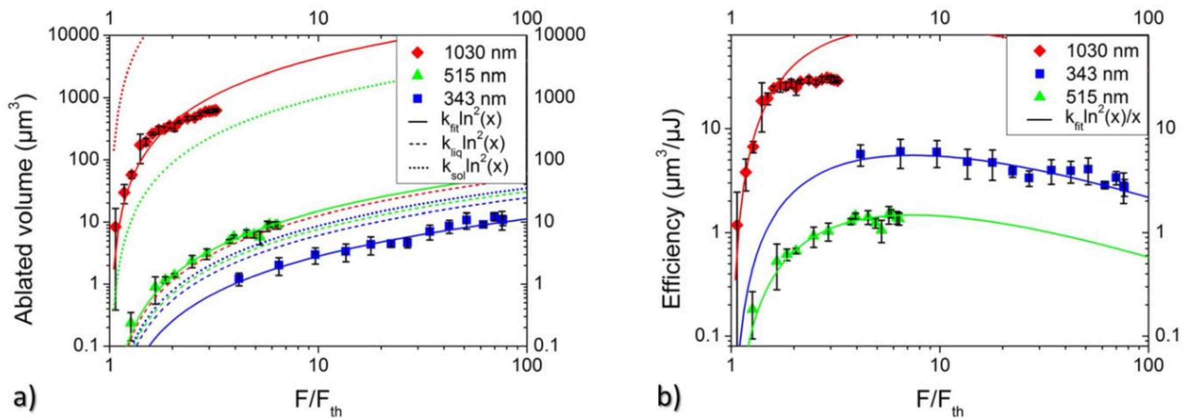


Fig. 3.11:Ablated volume and efficiency as a function of peak fluence normalized to the threshold fluence for three different wavelengths [23].

3.3.3 Laser pulse width

There are two important effects of choice of laser pulse width which are discussed separately (i) Beam distortion effect (ii) Modification of ablation threshold

3.3.3.1 Beam distortion effect

The beam distortion effect refers to the redistribution of incident laser energy because of the high peak intensities of focused laser beam when the laser pulse width is in the femtosecond regime. For the same energy applied at two different pulse widths, the pulse with shorter width has a larger intensity as compared to that with longer width. For instance, going from ns to fs timescale results increase in peak intensity by as much as 6 orders of magnitude.

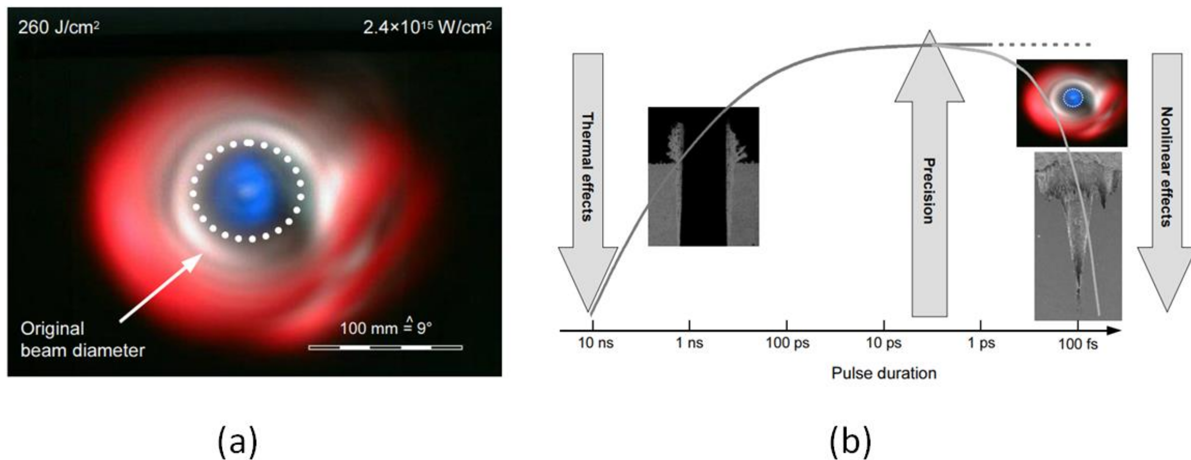


Fig. 3.12: (a) Beam energy distribution taken after focusing of laser beam in air and measuring far-field intensity of the beam. The encircled region represents the outline of the original beam. (b) Different effects occurring at long and short pulse widths. At long pulse widths, melting is severe which causes redeposited slag and at short pulse widths non-linear effects results in non-Gaussian energy deposition profiles and consequently enlarged ablation profiles. [17]

In the regions close to the focal spot, the intensity can result in non-linear interaction of the beam with the air. When the power density is sufficiently high, the beam interacts non-linearly with the atmospheric gas and results in conical emission. This causes emission of broadband radiation within a cone which has a divergence larger than the original beam. The net effect is the deposition of laser energy outside the ideal focal spot for the given beam focussing conditions. The distorted beam profile and its consequence for laser micromachining are illustrated in Fig. 3.12.

3.3.3.2 Modification of ablation threshold

It has already been seen so far that wavelength and ambient conditions can modify the ablation threshold. An additional factor that can modify the ablation threshold is the choice of pulse width of laser. The threshold fluence as function of pulse width (τ_p) can be given by:

$$\phi_{th} \propto \tau_p^m$$

The factor m depends on the pulse duration. For pulse widths longer than few tens of picoseconds, the generally accepted value of m is 0.5 which gives the square root dependence [24]. As the pulse width reduces, the value of m also reduces. The values of m of ~0.2 have been reported for two wavelengths 620 nm and 800 nm as plotted in Fig. 3.13a. It has been suggested by Korfiatis et al. that m is 0.5 when it is thermal diffusion dominated which is the case for longer pulse widths [25]. For very short pulse widths, different processes like non-thermal melting can occur and the value of m for such processes is found to be dependent on the pulse width. The square root dependence for longer pulse durations have been illustrated in the work of Sikora et al. [23]. As seen in Fig. 3.13b, for picosecond pulse duration square root dependence predicts well the ablation threshold.

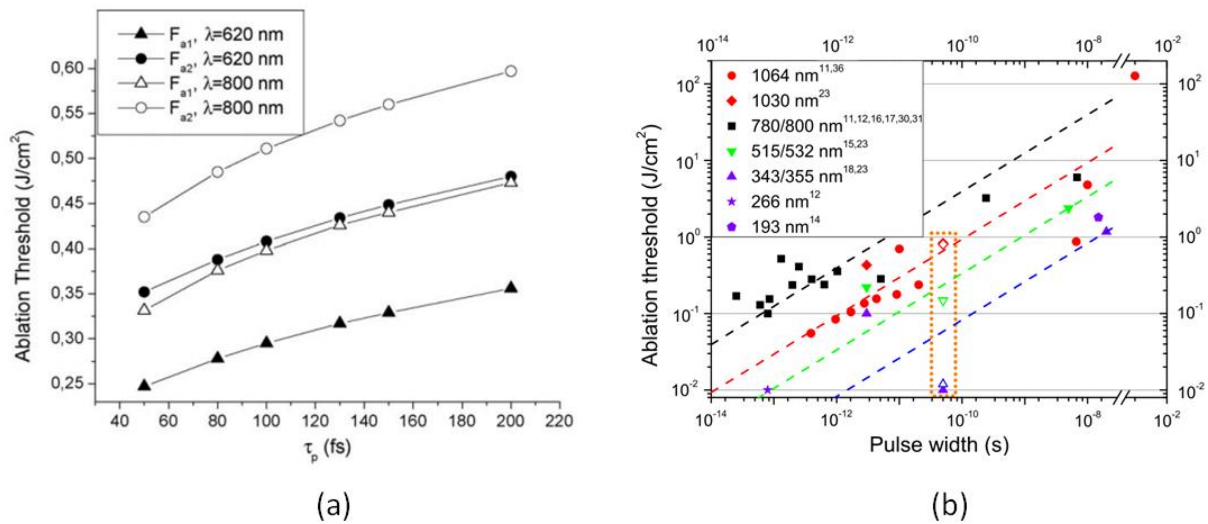


Fig. 3.13: Ablation threshold of silicon in the (a) femtosecond (b) picoseconds regime [23], [25]. In (a), two ablation thresholds are mentioned. Here F_{a1} refers to threshold for onset of ablation and F_{a2} refers to threshold for complete surface ablation

With reduction in threshold fluence, it is expected that lower pulse widths are preferable for efficient micromachining of silicon. Lopez et al. have reported the use of different pulse duration in the engraving of surface grooves on silicon and

other materials [26]. As seen in Fig. 3.14, for all materials, the ablation efficiency increases as the pulse duration reduces. For silicon, the improvement in efficiency is 4 times when the pulse duration is reduced from 34 ps to 900 fs. For some other materials, the improvement in efficiency is even higher.

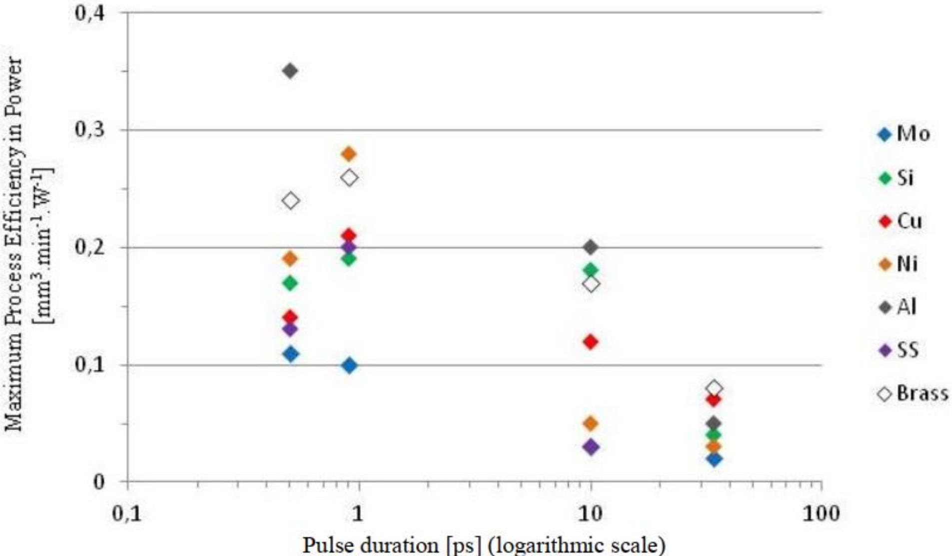


Fig. 3.14: Ablation efficiency of different materials studied by making grooves using a Yb: YAG thin disc laser wavelength of 1030 nm and different pulse widths [26]

3.3.4 Polarization

Polarization is another important aspect of laser processing. The polarization not only influences the absorption of laser radiation by the substrate but also affects the quality of the ablated area.

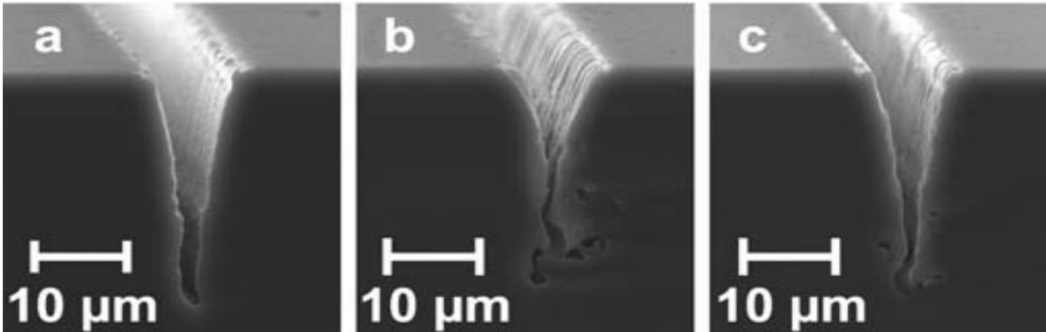


Fig. 3.15: Effect of polarization on the morphology of trenches micromachined in silicon (a) Polarization perpendicular to axis of the trench (b) Polarization parallel to axis of the trench (c) Circular polarization [2]

The scribing of trenches using three different polarizations have been studied by Crawford et al. [2]. The cross sectional profiles of the trenches are shown in Fig. 3.15. The translation of the sample perpendicular to the polarization gives the

best cross-sectional profile. For translation parallel to the direction of polarization, the cross-section is asymmetric and formation of branches is seen. For circularly polarized light, some branching is seen but is overall comparable to perpendicular translation. Similar results have been obtained in other works comprising the study of trenches in silicon [27], [28].

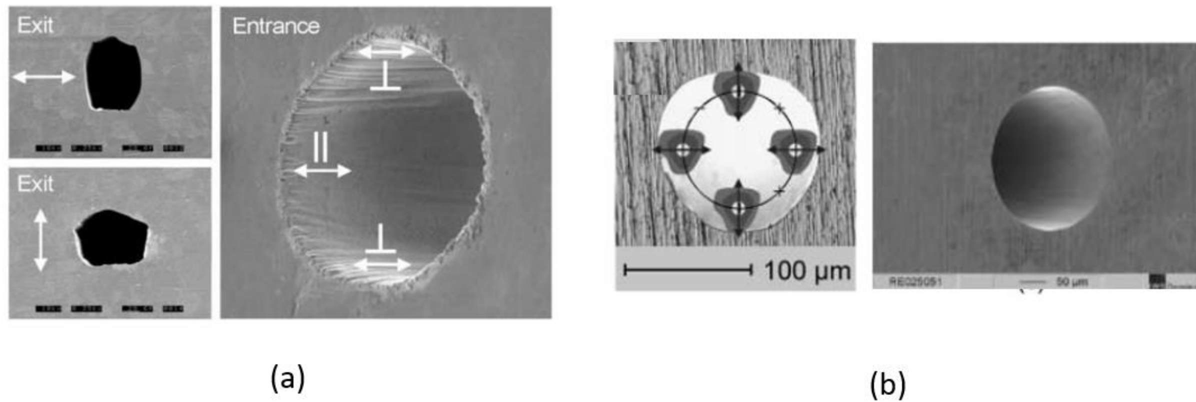


Fig. 3.16: Quality of exit holes during drilling (a) Using fixed direction of polarization showing elongation in direction perpendicular to direction of polarization (b) Dynamic polarization control with local polarization perpendicular to the tangent of the circular hole. [29]

The best polarization depends on the type of ablation geometry needed to achieve. For instance for the drilling of holes, in the work of Föhl et al., quality of exit hole was poor using single direction of polarization [29]. They implemented a system where beam polarization changes synchronously with circular movement of the laser beam as shown in Fig. 3.16. The quality of the exit hole is largely improved. While polarization is often an overlooked parameter in laser processing, its significance in improving quality is high.

3.4 Effect of ablation plume on laser processing

Ablation plume is a collection of atoms, clusters, ions and electrons that is emitted because of absorption of laser energy and resulting ablation. The effect of the ablation plume can last on a much longer timescale than the pulse width of the laser. Ablation plume poses a problem for maintaining the laser processing quality as it can absorb and scatter laser radiation. It is thus essential to understand the ablation plume and design the process in such a way that the impact of the ablation plume can be minimized.

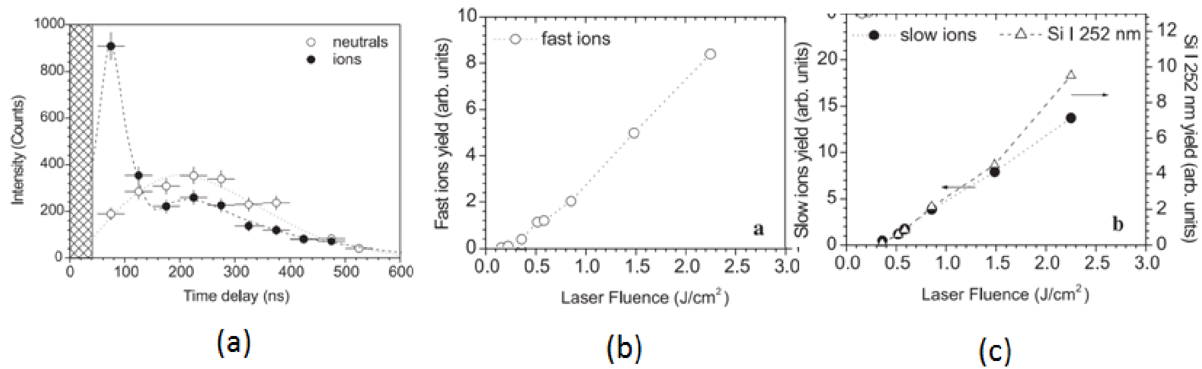


Fig. 3.17: (a) Time evolution of different species of the ablation plume of silicon at a distance of 7 mm from the target surface and peak fluence of 0.5 J cm^{-2} (b) Yield of fast ions as a function of laser fluence (c) Yield of slow ions as a function of laser fluence [30]

The ablated plume of silicon consists of different species which have different characteristic velocities. Fig. 3.17 shows the time evolution of neutrals and ions as a function of time detected using a Langmuir probe. This study by Amoruso et al. revealed that the plume contains neutrals and ions which have very high characteristic velocities of $10^6 - 10^7$ cm s^{-1} [30]. It can also be seen in Fig. 3.17 that the yield of ions increases with increasing fluence because of the total increased quantity of ablated material. Hence, the plume density is also a function of the laser fluence.

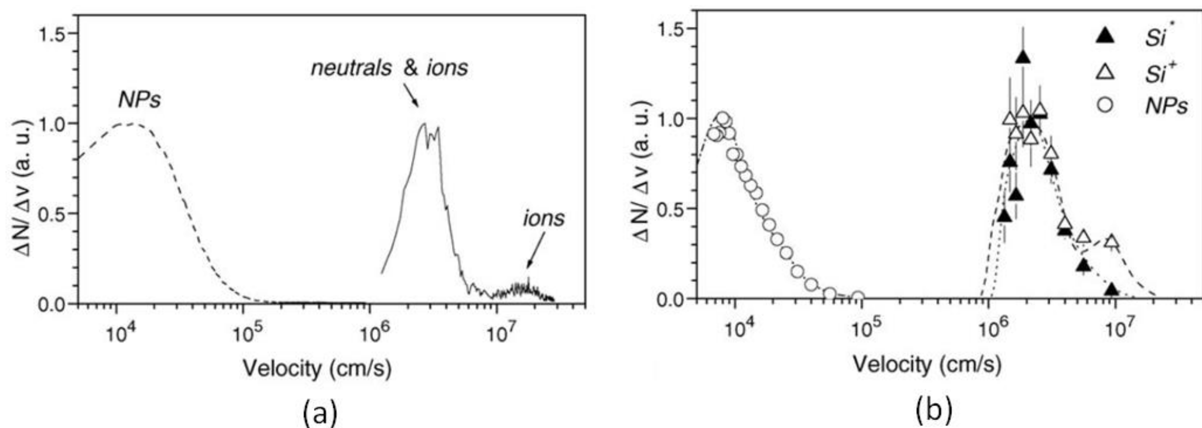


Fig. 3.18: Velocity distributions during ablation of silicon using different laser parameters (a) Pulse width: 0.9 ps, Wavelength: 1055 nm, Fluence = 0.75 J cm^{-2}, Intensity: 8.3 x 10^{11} W cm^{-2} (b) Pulse width : 0.1 ps, Wavelength: 780 nm, Fluence = 0.5 J cm^{-2}, Intensity: 5.0 x 10^{12} W cm^{-2} [31]

In another work by the same authors, a much slower component of plume was also detected consisting of nanoparticles of silicon at two different processing conditions [31]. As Fig. 3.18 shows, the distribution of velocities for nanoparticles has a peak at 10^4 cm s^{-1}.

The slow component of the ablation plume can potentially be a limiting factor during laser processing. The trend in new generation of laser systems is towards the used of repetition rates in the MHz range. For a pulse repetition rate of 1 MHz (pulse separation of 1 µs), the nanoparticle diffusion is 100 µm. Hence, the nanoparticles would still be in the vicinity of the surface. During laser micromachining, the beam is dynamic and each successive pulse has its own ablation front. These different ablation fronts can interact with each other which can result in confinement of the ablation plume in the zone of ablation in the vicinity of the surface.

In addition to the components already mentioned above, a vapour phase of the plume can also exist. The vapour phase is created when the surface is molten and the temperature exceeds the boiling point of silicon. This component is higher at higher laser fluences where amount of laser energy transferred to the lattice is significant and melting of material occurs.

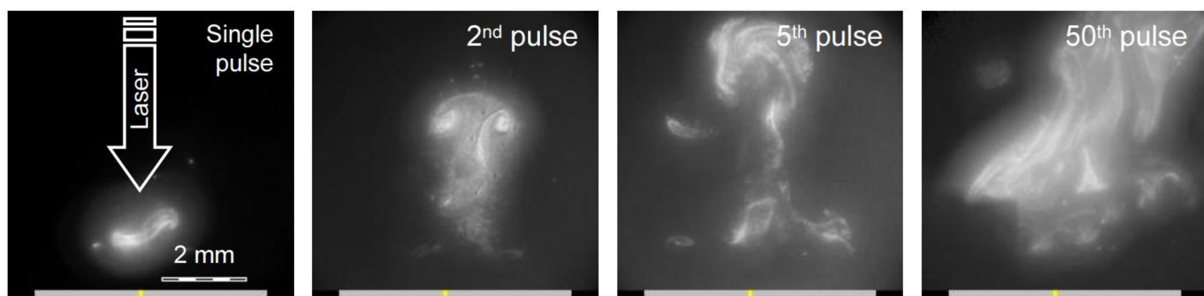


Fig. 3.19: Evolution of laser vapour plume of aluminium for a burst with varying number of pulses in the burst showing accumulation of plume with laser conditions as follows: (i) $\lambda = 800$ nm (ii) Pulse width: 110 fs (iii) Burst frequency (fp): 1 kHz (iv) Peak fluence: 260 J cm^{-2}. Pictures take approximately 500 µs after the last pulse is applied [32]

Femtosecond ablation is often referred to as melt-free ablation. But, a certain finite depth of melt is observed to exist in aluminium at 1 ps pulse width even at low fluence of 0.1 J cm^{-2} [32]. The melt depth is found to increase with fluence. For laser processing of different geometries, to have high removal rates, the laser fluence is usually much larger than the threshold fluence. Hence, melting of material is also to be considered even when using femtosecond pulses [17]. Consequently, a vapour phase of the material also exists in the ablation plume.

An illustration of this is given in Fig. 3.19 where the vapours resulting from ablation of aluminium at high fluence is shown [32]. When a burst of pulses is used with varying number of pulses per burst, for higher number of pulses the vapour aggregates to form a larger volume. This is despite using relatively small pulse frequency in the burst equal to 1 kHz. Hence, the effect of vapour phase is

also non-negligible during micromachining which is essentially a multi pulse process.

3.5 Laser processing: Description of system

3.5.1 Laser source and optics

The laser source that is used in this work is Tangerine from Amplitude Microsystèmes, which is an Yb doped fibre laser. The full schematic of the laser system is shown in Fig. 3.20. The fundamental output wavelength of the laser head is 1030 nm and the pulse width is 350 fs. The laser is also equipped with the SHG/THG module which provides the capability of using two additional wavelengths of 515 nm and 343 nm. The SHG and THG modules have different conversion efficiencies which result in a reduced average power output as compared with the fundamental output. The laser beam parameters are summarized in Table 3.2. The laser beam passes through the various optical components as shown in Fig. 3.20, before it is focused onto the sample.

The laser beam output from the source is linearly polarized with a beam diameter of ~2 mm. This diameter is further increased by beam expanders to reduce the effect of energy density on the coating of optical elements. Optical attenuators and trepan head are placed further along the optical path. The linearly polarized light is converted to circularly polarized light by making use of a quarter wave plate. The beam finally passes through the galvanometric scanner comprising of an x-y mirror system and is focused by means of a telecentric lens. The laser pulse repetition rate can be modified from single shot to 2 MHz. The modulator frequency for peak energy output of the laser for all source wavelengths is 200 kHz (Fig. 3.21). After 200 kHz, the power output remains constant which means that pulse energy reduces with repetition rate.

Table 3.2: Summary of laser beam parameters for the 3 laser sources

Source Wavelength	343 nm	515 nm	1030 nm
Beam quality (M^2)	1.24	1.15	1.18
Beam waist (um)	5.8**	3.9***	8.2**
Rayleigh length (um)	244.2	85	173.8
Numerical aperture	0.05	0.05	0.05
Max power (W) (@ 200 kHz)*	1	3.8	13

* - Power available at the workpiece measured using a power meter

** - Experimentally determined

*** - Calculated from optical system specifications

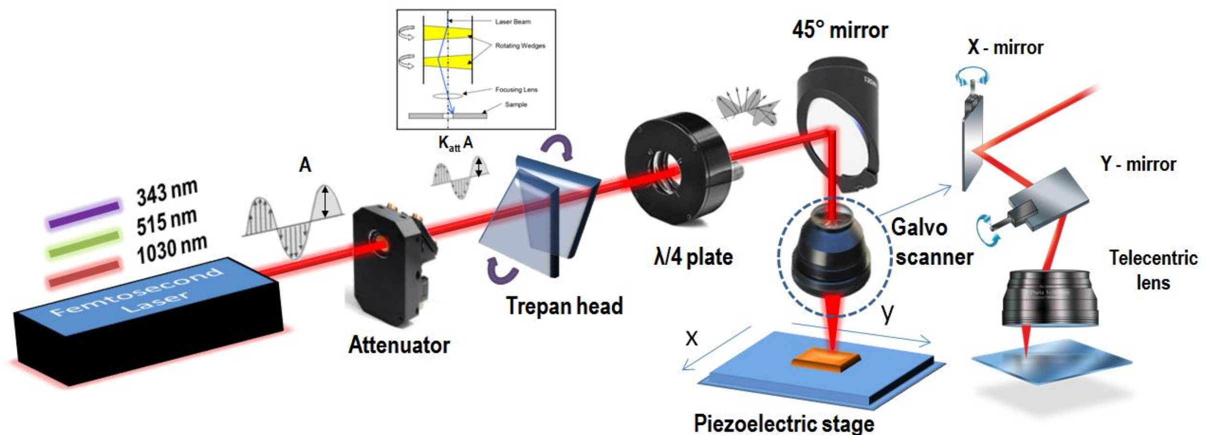


Fig. 3.20: Schematic description of the laser system depicting the important optical components

3.5.2 Optical attenuator

The optical attenuator module is meant for controlling the laser power delivered to the source. The attenuation is done by using a half wave plate coupled to a polarizer and a light trap. The axis of polarization can be changed by rotating the half-wave plate using a servo motor. The degree of attenuation depends upon the relative angle between the axes of the half wave plate and the polarizer.

The attenuation of the beam is measured relative to the average power of the incoming beam. A software calibration of required angle of rotation for desired percentage of attenuation allows having a fine control of the output power.

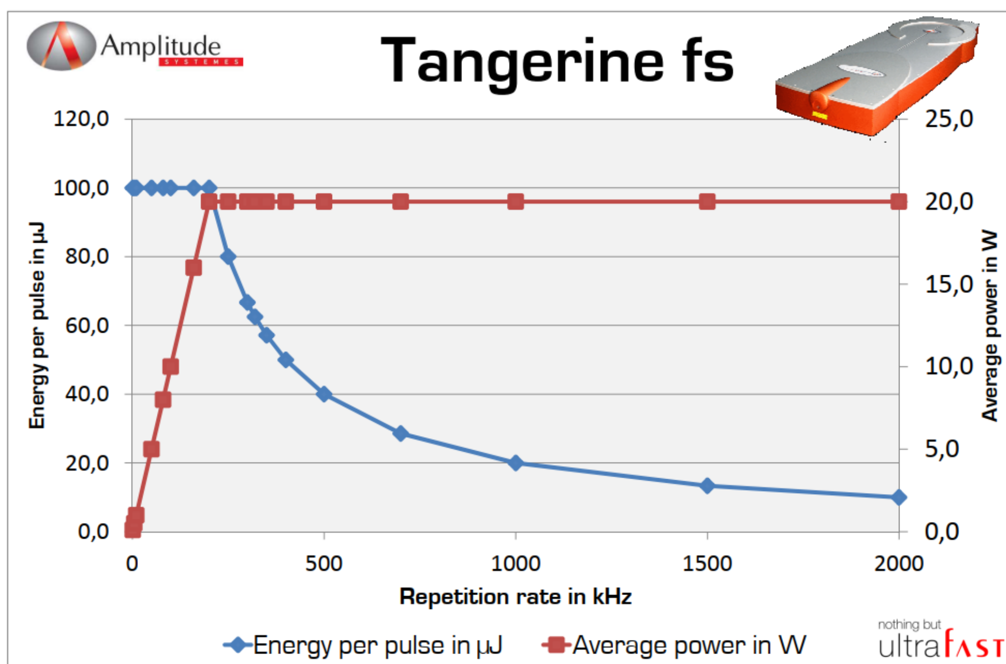


Fig. 3.21: Variation of pulse energy and average power output as a function of repetition rate

3.5.3 Trepan head

The trepan head is meant for deflection of the laser beam mainly in circular and spiral paths. It is based on the concept of Risley prisms.[33] It is typically used for microhole drilling but can be adopted for circular milling operations as well by carefully defining the beam trajectory.

3.5.4 Galvanometric scanner

The galvanometric scanner comprises of a xy mirror system which can deflect the beam within a xy field of 35 mm x 35 mm. The mirror is coupled to the telecentric lens which focuses the beam onto the same plane at the workpiece within the galvo field.

3.6 Laser processing: Experimental plan

The final desired outcome of the laser micromachining process that is going to be developed is to obtain cavities in silicon of varying shapes and dimensions. The emphasis shall be on developing the process for rectangular geometries with side length ranging from 100 µm up to 2.5 mm. The depths of the cavities that are going to be studied will be 20 µm – 650 µm. The cavities shall be characterized using optical profilometry. Different measures of quality shall be determined for the cavity from the obtained profilometer scans. Based on the obtained results, the recipes for milling cavities are adapted to removal of silicon under RF circuits. The following paragraphs explain the different processing options that would be explored to achieve the final objective.

The three available source wavelengths for micromachining are 343 nm, 515 nm and 1030 nm. 343 nm has the highest resolution (smallest spot size) and the lowest output power while 1030 nm has the smallest resolution (largest spot size) and the highest output power. While source power is one factor to consider, the absorption of laser power by silicon is another factor to consider. The absorption coefficient for silicon decreases for increasing wavelengths. The absorption coefficient of silicon can increase by up to 5 orders of magnitude when changing the wavelength from 1030 nm to 343 nm as shown in Fig. 3.22 [34]. While this holds true for low intensity radiation, the same does not hold true for high intensity laser radiation where classical theories cannot be applied. Additionally, non-linear absorption also plays a big role because of high intensity of femtosecond laser pulses. It has been demonstrated through experimental data that the penetration depth (reciprocal of absorption coefficient) varies only by a

few nm between the 343 nm and 1030 nm source [35]. This difference is miniscule as compared to a 5 orders of magnitude difference for low intensity radiation. Hence, 1030 nm source is most suitable because resolution requirement is not stringent and the fluence can be varied over a large range of values. This can help increase the ablation rate as per the specified requirements of the cavity.

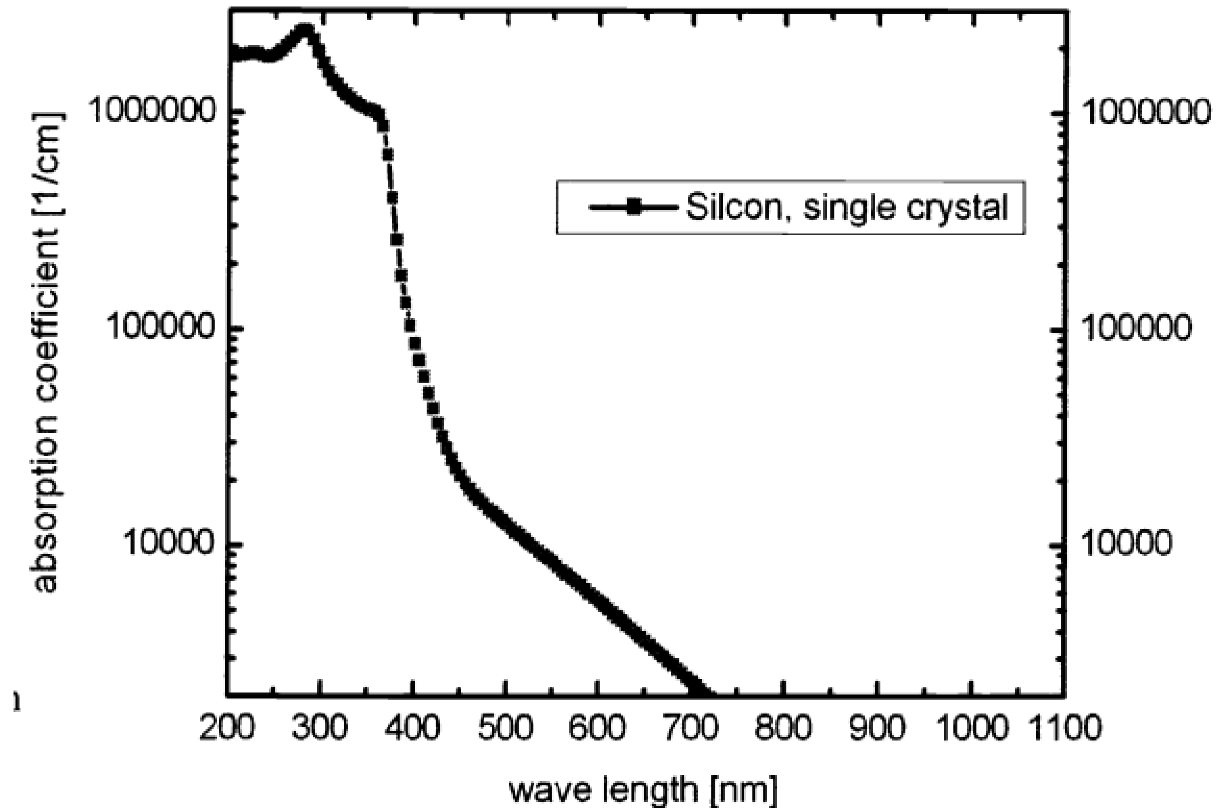


Fig. 3.22: Linear absorption coefficient of silicon as a function of wavelength [34]

Stability of beam movement is essential for good micromachining quality. The beam stability is a function of speed of scanning because the galvo mirrors need to deflect more rapidly in order to deflect the beam in the desired path. A zigzag geometry is scribed using the different sources to trace the beam path in the presence of sharp turns. This will provide the information of maximum usable galvo speed for micromachining applications.

The first set of experiments includes the simplest geometry that can be ablated using laser micromachining which is scribing of lines. The ablation rate can be quantified for different process parameters. The process parameters that are varied are (i) Galvo speed (ii) Fluence. The study of scribing of lines serves two purposes 1) Measure of relative ablation rate between different process parameters 2) Study the relative degree of thermal effects.

Finally, the laser milling is performed over large area using the 343 nm and 1030 nm sources. Laser milling process involves the raster scanning of lines having a certain pitch between the lines. The pitch between two lines is an additional process parameter that will be investigated. Milling involves a more complex 2D beam trajectory and ablation plume from large amounts of removed material which can interact with the beam. Thereby, the final obtained depth would not be the same as determined from scribing lines. The ease of removal of material would depend on the required aspect ratio of the cavity. This is because diffusion of laser plume from the ablated areas is easier for cavities with larger openings (milled area) as compared to smaller openings. This requires that for each choice of processing parameters and geometry, a calibration of depth be performed with respect to number of scans of the laser over the milled area. All of these aspects are subjects of discussion in Chapter 4.

3.7 RF circuits: Description and theory

This section provides the description of RF circuits that are going to be studied after removal of substrate. Circuits are fabricated on the 130nm technology node of ST Microelectronics namely H9SOIFEM which is specialized for RF Front End Modules. Small signal S-parameters characterization will be performed on all circuits. Depending on the circuit application, additional RF analyses such as linearity and noise figure measurements are performed. These studies serve to demonstrate the impact of substrate removal on different RF performance metrics. The circuits that will be characterized are test structures that closely resemble those that are used in real world applications. This allows the direct assessment of practical significance that can be obtained by substrate removal.

When the substrate removal is performed, silicon is removed under the functional area of the circuit. All circuits use aluminium pads for incoming and outgoing RF signals as well as DC biasing. For effective probing, the probe tips need to pierce tightly contact and scratch the bond pads. Hence, silicon is retained under the bond pads as the membrane cannot withstand the strong forces applied by the probe tips during RF probing. This means that, for some circuits, the pads need to be deembedded to obtain the device relevant data.

For most of the experiments, both HR and TR substrate types are characterized. The thickness of silicon and BOX for all characterized substrates is 160 nm and 400 nm respectively.

3.7.1 Isolation structures

Isolation structures are the simplest structures that are studied in this work. They are intended to bring out the effect of substrate coupling in RF circuits. They are long interdigitated fingers separated by varying spacing. The fingers start from metal 1 and they are contacted to the Silicon film through p+ doping. The active areas of each finger are separated by shallow trench isolation. The interdigitated structures are connected to RF pads on either side which allows 2-port RF measurements. When RF signal is sent from one side, coupling takes place both through the layers on top of the BOX as well as through the substrate. A schematic illustration of one such structure is shown in Fig. 3.23.

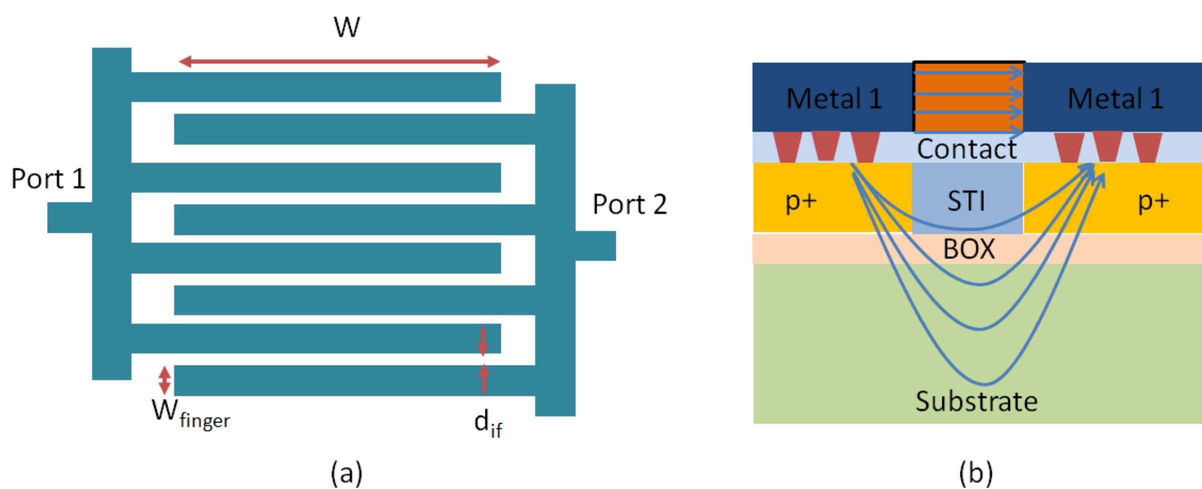


Fig. 3.23: Schematic representation of an isolation structure with its characteristic dimensions shown in (a) Top view (b) Cross section showing the different SOI layers and path for signal coupling between two fingers

While this structure does not have any function, it can reveal information about the impact of substrate on real world circuits. For instance, it can be useful in the analysis of behaviour of multi-turn inductors. The turn to turn coupling plays an important role in determining the inductor properties. A study of isolation structures reveals the behaviour of substrate coupling as a function of frequency. They can also be used for substrate electrical modelling in term of RC network in devices such as RF MOSFET, RF passives, Circuits and EM simulators. In addition, these structures allow studying the biasing dependency of the BOX capacitance.

The parameter of interest for isolation structures is 2-port S-parameters. $|S_{21}|$ is plotted as a function of chuck bias at different frequencies for each isolation structure shown in Table 3.3. In this table, N_{finger} represents the number of

fingers per port. The chuck bias reveals the non-linear behaviour of substrate capacitance at different frequencies which is pronounced in the case of HR substrate. ΔS_{21} allows comparison of substrate coupling before and after substrate removal as a function of frequency. This would highlight the importance of substrate for device performance at different frequencies.

Table 3.3: List of characterized isolation structures and their dimensions

Isolation structure	W (μm)	W_{finger} (μm)	d_{if} (μm)	N_{finger}
ISO-A	269.7	10.7	18.26	6
ISO-B	269.7	10.7	11.26	8
ISO-C	269.7	10.7	5	10
ISO-D	269.7	10.7	3	12

3.7.2 RF Switch

After having characterized the substrate coupling using isolation structures, the first practical case of substrate processing that is going to be studied is the RF switch. It is an important component of an RF frontend and is the first component which is present in between the antenna and rest of the RF components of the frontend. The switch allows the signal to be routed from the antenna to the required circuitry either in Tx or Rx mode within the frontend. The throw count depends on the complexity of the front end.

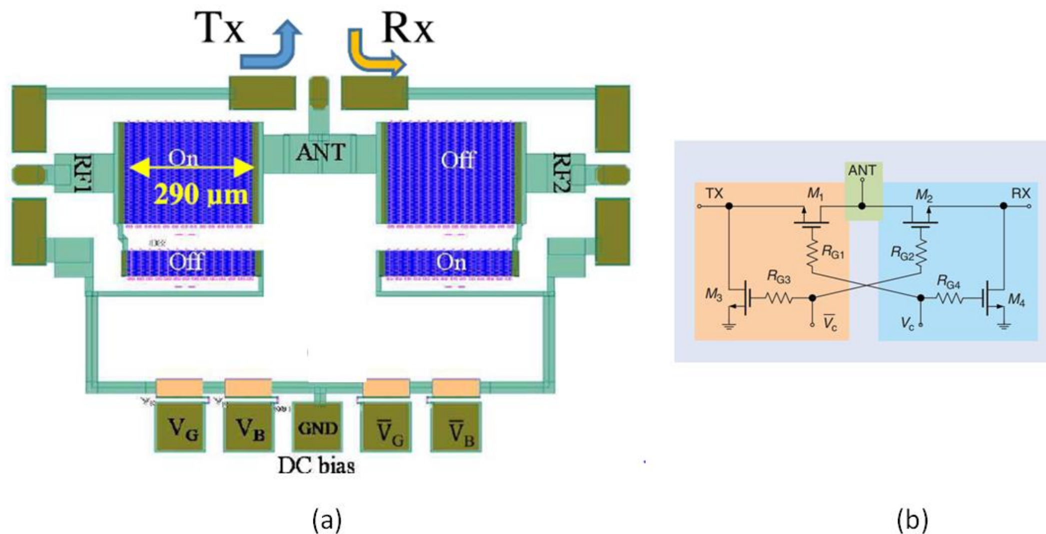
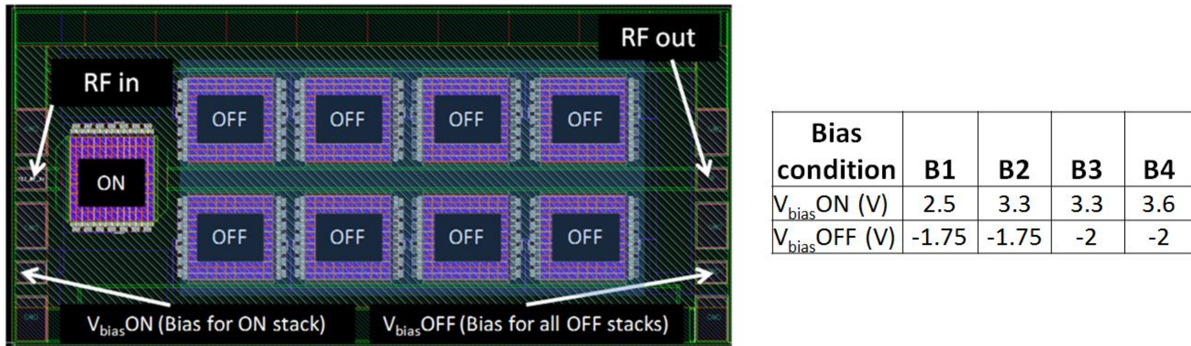


Fig. 3.24: Schematic representation of an SP2T switch showing (a) Transistor stacking and relative sizes of series and shunt branches (b) Equivalent circuit of the switch [36], [37]

There are several possible implementations of an RF switch and the commonly used one on SOI technology is the series-shunt architecture. This architecture is illustrated with an example in Fig. 3.24. The two big transistor stacks are the

series branches and the two smaller branches are the shunt branches. In Tx mode, the series branch of the Tx side is on while the series branch in Rx branch remains off. The opposite is true for Rx mode. The on/off condition of the series and shunt branches on the same side are complementary which means that when series branch is on shunt is off and vice versa. The shunt branch is smaller in size than the series branch and provides isolation for the branch which is turned off where signal can couple through series branch in the off state.



Bias condition	B1	B2	B3	B4
V_{bias}ON (V)	2.5	3.3	3.3	3.6
V_{bias}OFF (V)	-1.75	-1.75	-2	-2

Fig. 3.25: Layout view of the SP9T switch test structure used in the study having GSGSG access. The gate and body bias is tied for all transistor stacks. One DC bias is provided for the ON stack and the second one for other 8 stacks. The different tested bias conditions are shown on the right.

The same concept can be extended for multiple switch branches. In this work, substrate removal is performed on a SP9T test structure shown in Fig. 3.25. The test structure is designed to characterize the linearity with respect to 2nd and 3rd harmonic (H2/H3) measures of the series branch. The shunt branches are not present in the test structure as they are provided only for isolation and do not affect the linearity behaviour. Small signal S-parameter measurements provide a measure of change in impedance of the switch because of substrate removal and also the insertion loss of the switch. Another important parameter in switch design is the isolation when a series branch is in the off condition. This parameter cannot be measured in the present test structure because there are no RF pads in the OFF path.

3.7.3 Inductors

Inductors are important passive elements used ubiquitously in RF design for circuits like LNA, PA, filters etc. Based on the desired application, inductors are designed by carefully adjusting the tradeoffs between power, losses, linearity etc.

The inductors that are used in the study are taken from the 130nm design kit with 4 layers of copper metallization which are stacked in parallel to achieve low series resistance. The design kit provides two parametric cells namely high I (current) and high Q (quality factor) pcells. The parameter of interest being effect of substrate, the high Q pcell is used in the study.

The 6 inductors that are studied are listed in Table 3.4. It can be seen that it comprises of 3 one turn and 3 two turn inductors. They have an octagonal geometry and the dimension is characterized by the inner diameter which is the distance between any two parallel segments of the innermost turn of inductor which has an octagonal geometry. The effective length is varied in each case by increasing the diameter of the inductor for both one turn and two turn cases. The self-inductance values range from 0.2 – 3.25 nH for the inductors. The spacing between adjacent turns for two turn inductors is 10 μm.

Table 3.4: Description of inductors used in the study

Inductor	No. of turns	Width (μm)	Finger Width (μm)	Finger spacing (μm)	Inner Diameter (μm)	No. of fingers	Ls @100 MHz (nH)
1T-S	1	47.22	10.23	2.10	80.00	4	0.20
1T-M	1	47.22	10.23	2.10	240.00	4	0.50
1T-L	1	47.22	10.23	2.10	400.00	4	0.89
2T-S	2	33.13	9.64	2.10	80.00	3	0.61
2T-M	2	33.13	9.64	2.10	240.00	3	1.81
2T-L	2	33.13	9.64	2.10	400.00	3	3.25

The effect of substrate in case of an inductor has been modelled by Liu et al. [38]. The schematic representation of an inductor and the equivalent lumped circuit are showed in Fig. 3.26. In case of single turn inductors, the C_s value is negligible and the substrate coupling occurs along the length of the inductor. In case of two turn inductors, the inter-turn coupling provides an additional path for coupling through the substrate. The coupling occurs both along the length and between the turns of the inductor.

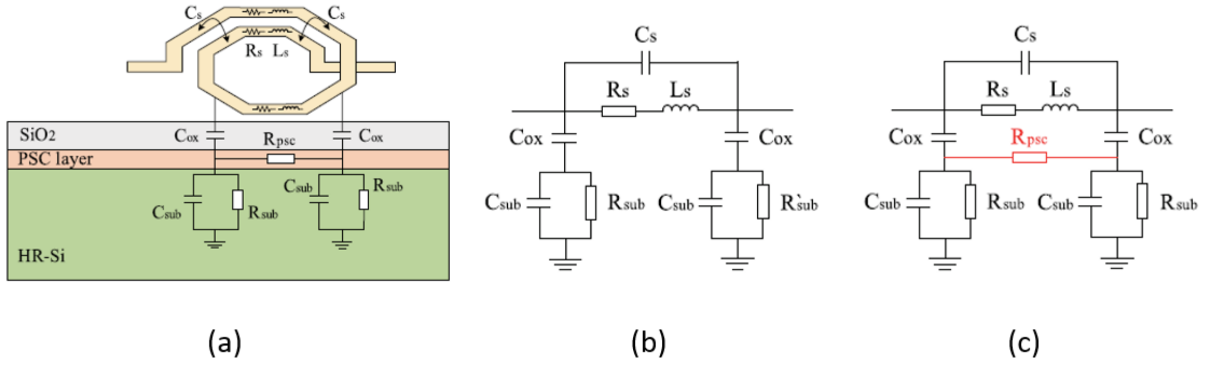


Fig. 3.26: (a) Cross sectional schematic of an inductor fabricated on SOI wafer showing the origin of different capacitances and resistances (b) Lumped circuit equivalent for TR-SOI (c) Lumped circuit equivalent for HR-SOI. Taken from [38]

The important distinguishing factor between the HR-SOI and TR-SOI cases is the presence of Parasitic Surface Conduction (PSC) layer. A low resistance path is created at the interface between silicon and the BOX because of charge on the oxide. This allows an additional parasitic conduction path in HR-SOI case. This path is eliminated in the TR-SOI substrates by introducing a poly-silicon layer which traps the free charges that are generated and hence substrate losses are ameliorated.

3.7.4 Low Noise Amplifier (LNA)

The low noise amplifier is an important component of RF receiver systems and is commonly the first component after the antenna in the receiver chain. As the name suggests, LNA is an amplifier specifically designed to have low noise and hence it plays an important role in the overall sensitivity of the RF receiver. The noise figure of the receiver system is a measure of its sensitivity and is given by:

$$NF_{sys} = 10\log\left(F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \cdots \frac{F_n - 1}{G_1 G_2 G_3 .. G_n}\right) \tag{3.16}$$

Here G_n refers to the gain of the nth stage of the receiver and F_n is the noise factor corresponding to this stage. It can be clearly seen F_1 is very critical for determining the noise figure of the system. The contribution of noise further down the receiver chain $(F_{2,} F_3, F_4 \ldots)$ is divided by the product of gains upto the point under consideration. Hence these terms do not dominate the noise figure. Hence, LNA is chosen as the first stage of the receiver chain to have the lowest noise figure for the receiver. A typical example of receiver architecture is showed in Fig. 3.27 to illustrate this point [39].

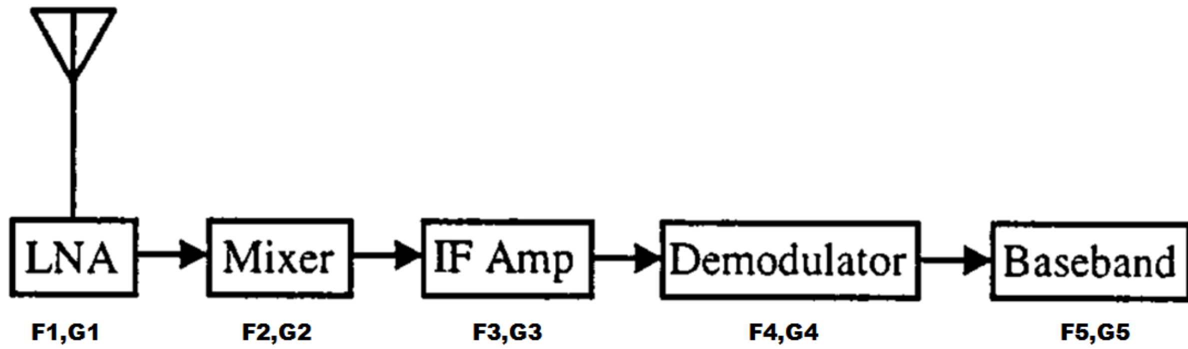


Fig. 3.27: Typical receiver architecture showing the different components and associated gains and noise factors [39]

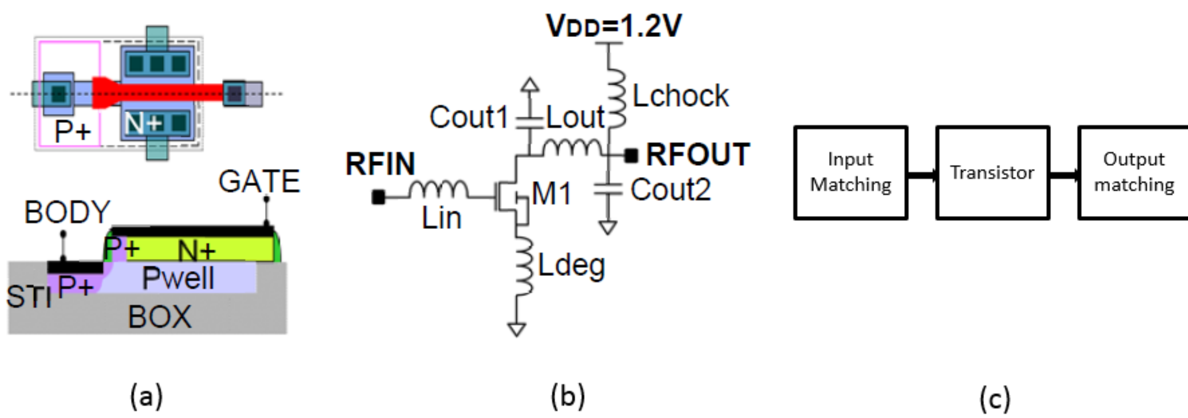


Fig. 3.28: (a) Transistor layout showing the implementation of the body contact (b) Circuit diagram for the common source LNA (c) Input block diagram of the LNA. [40]

While noise is a key factor to consider during the LNA design, there are other important parameters like linearity, gain, bandwidth, stability, reverse isolation etc. that need to be optimized. The LNA that is studied in this work has a common source configuration with inductive source degeneration as shown in Fig. 3.28b. The common source LNA gives a good noise performance compared to other implementations. However, the chosen LNA is designed to provide maximum linearity and not the best noise performance as described in [40].

As seen in Fig. 3.28c, the LNA block diagram consists of three parts namely the input impedance matching network, transistor amplifier and the output impedance matching network. The very first module that a RF signal sees in the receiver chain after the antenna is the input matching network of the LNA. Hence, the design of input matching network for the LNA will have an important contribution to noise figure of the system. A high-Q inductor is commonly used at the input of the LNA to have the required matching conditions at the input.

By improving the quality factor of the input inductor, the noise performance of the LNA can further be improved. This fact is established in the comprehensive noise analysis for common source LNA circuits by Shaeffer et al. [41]. The input stage of common source LNA analyzed in this work is shown in Fig. 3.29. The corresponding equivalent circuit for noise calculations is depicted in Fig. It was postulated that if the gain of the input stage is sufficiently large, the noise factor of the subsequent stages can be neglected.

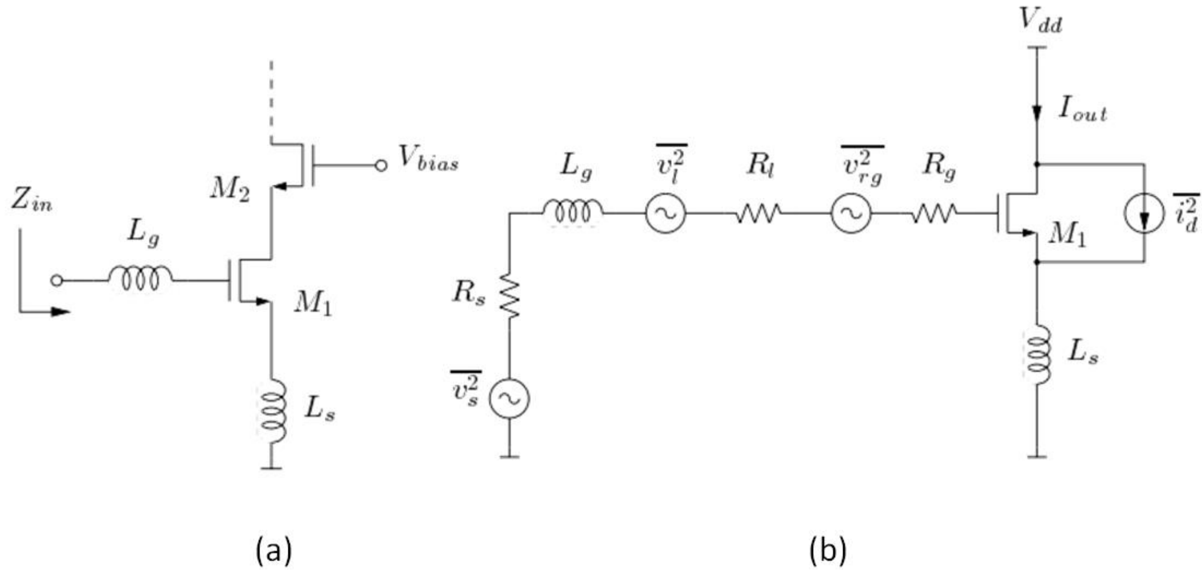


Fig. 3.29: (a) Input stage of a common source LNA (b) Equivalent circuit for calculating the noise at the input stage of LNA [41]

The noise power output density for a bandwidth of 1 Hz was derived as:

$$F = 1 + \frac{R_l}{R_s} + \frac{R_g}{R_s} + \gamma g_{d0} R_s \left(\frac{\omega_0}{\omega_T}\right)^2 \tag{3.17}$$

$R_l/R_s - Series\ resistance\ for\ the\ inductors\ L_g/L_s$ $\qquad R_g - Gate\ resistance$

$g_{d0} - Zero\ bias\ drain\ conductance\ of\ the\ MOSFET$ $\qquad \gamma - Bias\ dependent\ factor$

$\omega_T - Resonant\ frequency\ of\ input\ stage$ $\qquad \omega_0 - Frequency\ of\ the\ input\ signal$

From equation 3.17, the direct link of series resistance of inductor at the input can be established. It can also be seen that the noise contribution from transistor can be optimized independently of the input circuit by improving transistor parameters γ and g_{d0}.

After having discussed the different RF circuits, the different RF characterization steps that are going to be performed are summarized in Table 3.5.

Table 3.5: Summary of RF characterization performed on different RF circuits

Component	Measure 1	Measure 2	Measure 3
Isolation structures	*2-port S-parameters* 10 MHz – 26 GHz V_{chuck}: -2.5 V to 2.5 V	-	-
RF switch	*2-port S-parameters* 100 MHz - 6 GHz	*DC characteristics (Id - Vds)* Vg = 0.1 - 1 V	*Harmonic Distortion* Fundamental: 1.22 GHz 2nd Harmonic: 2.44 GHz 3rd Harmonic: 3.66 GHz
Inductors	*1-port S-parameter* 100 MHz - 110 GHz	-	-
LNA	*2-port S-parameters + Noise Figure* 4 GHz - 7 GHz	*Third order intercept point (IP3)* At 3 frequencies: 4.9 GHz, 5.4 GHz, 5.9 GHz	*1 dB compression point (P1dB)* At 3 frequencies: 4.9 GHz, 5.4 GHz, 5.9 GHz

Conclusion

The essential theoretical definitions and concepts for pulsed laser processing have been introduced. A brief introduction to laser material interaction has been presented to outline different processes and their timescales. The advantages of femtosecond pulse duration also have been discussed. The effect of different parameters like ambient conditions, pulse width, wavelength and polarization on laser processing of silicon has been highlighted using numerous examples in literature. Additionally, the presence of ablation plume has been discussed which shows presence of different components which can potentially have adverse impacts on the laser process. The laser processing system and its features used in this work have been discussed. A plan of experiments has been presented for development of the 2D milling process. The results of these experiments will be presented in Chapter 4.

The description of different circuits of the FEM has been given along with the studies that are planned. Isolation structures serve to provide an understanding the role of substrate in SOI RF circuits. Following this, substrate removal studies are performed on both passive and active circuits. For different circuits, the targeted improvements in performance using substrate removal are as follows: Linearty and insertion loss for the switch, Q-factor for inductors and Noise Figure for LNA. The results of RF characterization of membranes are presented in Chapter 5.

References

[1] "Gaussian Beam Optics." [Online]. Available: http://experimentationlab.berkeley.edu/sites/default/files/MOT/Gaussian-Beam-Optics.pdf.

[2] T. H. R. Crawford, A. Borowiec, and H. K. Haugen, "Femtosecond laser micromachining of grooves in silicon with 800 nm pulses," *Appl. Phys. A*, vol. 80, no. 8, pp. 1717–1724, May 2005.

[3] Z. Sun, M. Lenzner, and W. Rudolph, "Generic incubation law for laser damage and ablation thresholds," *J. Appl. Phys.*, vol. 117, no. 7, p. 073102, Feb. 2015.

[4] S. H. Kim, I.-B. Sohn, and S. Jeong, "Ablation characteristics of aluminum oxide and nitride ceramics during femtosecond laser micromachining," *Appl. Surf. Sci.*, vol. 255, no. 24, pp. 9717–9720, Sep. 2009.

[5] S. I. Anisimov, B. L. Kapeliovich, and T. L. Perelman, "Electron emission from metal surfaces exposed to ultrashort laser pulses," Zh Eksp Teor Fiz, vol. 66, no. 2, pp. 375–377, 1974.

[6] B. N. Chicbkov, C. Momma, and S. Nolte, "Femtosecond, picosecond and nanosecond laser ablation of solids," p. 7.

[7] P. Schaaf, Ed., *Laser Processing of Materials: Fundamentals, Applications and Developments*. Berlin Heidelberg: Springer-Verlag, 2010.

[8] W. Schulz, U. Eppelt, and R. Poprawe, "Review on laser drilling I. Fundamentals, modeling, and simulation," *J. Laser Appl.*, vol. 25, no. 1, p. 012006, Feb. 2013.

[9] S. K. Sundaram and E. Mazur, "Inducing and probing non-thermal transitions in semiconductors using femtosecond laser pulses," *Nat. Mater.*, vol. 1, no. 4, pp. 217–224, Dec. 2002.

[10] X. Liu, D. Du, and G. Mourou, "Laser ablation and micromachining with ultrashort laser pulses," *IEEE J. Quantum Electron.*, vol. 33, no. 10, pp. 1706–1716, Oct. 1997.

[11] V. Kostrykin, M. Niessen, J. Jandeleit, W. Schulz, E.-W. Kreutz, and R. Poprawe, "Picosecond-laser-pulse-induced heat and mass transfer," presented at the High-Power Laser Ablation, Santa Fe, NM, 1998, pp. 971–982.

[12] P. Lorazo, L. J. Lewis, and M. Meunier, "Thermodynamic pathways to melting, ablation, and solidification in absorbing solids under pulsed laser irradiation," *Phys. Rev. B*, vol. 73, no. 13, Apr. 2006.

[13] J. Cheng *et al.*, "A review of ultrafast laser materials micromachining," *Opt. Laser Technol.*, vol. 46, pp. 88–102, Mar. 2013.

[14] D. Perez and L. J. Lewis, "Molecular-dynamics study of ablation of solids under femtosecond laser pulses," *Phys. Rev. B*, vol. 67, no. 18, p. 184102, May 2003.

[15] C. Cheng and X. Xu, "Mechanisms of decomposition of metal during femtosecond laser ablation," *Phys. Rev. B*, vol. 72, no. 16, p. 165415, Oct. 2005.

[16] K. Sokolowski-Tinten *et al.*, "Transient States of Matter during Short Pulse Laser Ablation," *Phys. Rev. Lett.*, vol. 81, no. 1, pp. 224–227, Jul. 1998.

[17] F. Dausinger, H. Hugel, and V. I. Konov, "Micromachining with ultrashort laser pulses: from basic understanding to technical applications," presented at the SPIE Proceedings, 2003, pp. 106–115.

[18] S. Besner, J.-Y. Degorce, A. V. Kabashin, and M. Meunier, "Influence of ambient medium on femtosecond laser processing of silicon," *Appl. Surf. Sci.*, vol. 247, no. 1–4, pp. 163–168, Jul. 2005.

[19] H. Liu *et al.*, "Influence of liquid environments on femtosecond laser ablation of silicon," *Thin Solid Films*, vol. 518, no. 18, pp. 5188–5194, Jul. 2010.

[20] J. Thorstensen and S. Erik Foss, "Temperature dependent ablation threshold in silicon using ultrashort laser pulses," *J. Appl. Phys.*, vol. 112, no. 10, p. 103514, Nov. 2012.

[21] H. K. Toenshoff, A. Ostendorf, K. Koerber, and N. Baersch, "Speed rate improvement for microcutting of thin silicon with femtosecond laser pulses," presented at the LAMP 2002: International Congress on Laser Advanced Materials Processing, Osaka, Japan, 2003, p. 531.

[22] A. Borowiec, H. F. Tiedje, and H. K. Haugen, "Wavelength dependence of the single pulse femtosecond laser ablation threshold of indium phosphide in the 400–2050nm range," *Appl. Surf. Sci.*, vol. 243, no. 1–4, pp. 129–137, Apr. 2005.

[23] A. Sikora, D. Grojo, and M. Sentis, “Wavelength scaling of silicon laser ablation in picosecond regime,” *J. Appl. Phys.*, vol. 122, no. 4, p. 045702, Jul. 2017.
[24] M. D. Perry, B. C. Stuart, P. S. Banks, M. D. Feit, V. Yanovsky, and A. M. Rubenchik, “Ultrashort-pulse laser machining of dielectric materials,” *J. Appl. Phys.*, vol. 85, no. 9, pp. 6803–6810, Apr. 1999.
[25] D. P. Korfiatis, K.-A. Th. Thoma, and J. C. Vardaxoglou, “Numerical modeling of ultrashort-pulse laser ablation of silicon,” *Appl. Surf. Sci.*, vol. 255, no. 17, pp. 7605–7609, Jun. 2009.
[26] J. Lopez *et al.*, “Comparison of picosecond and femtosecond laser ablation for surface engraving of metals and semiconductors,” presented at the SPIE LASE, San Francisco, California, USA, 2012, p. 82430O.
[27] C. Y. Chien and M. C. Gupta, “Pulse width effect in ultrafast laser processing of materials,” *Appl. Phys. A*, vol. 81, no. 6, pp. 1257–1263, Nov. 2005.
[28] M. Halbwax *et al.*, “Micro and nano-structuration of silicon by femtosecond laser: Application to silicon photovoltaic cells fabrication,” *Thin Solid Films*, vol. 516, no. 20, pp. 6791–6795, Aug. 2008.
[29] C. Föhl, D. Breitling, and F. Dausinger, “Influences on hole quality in high precision drilling of steel with ultra-short pulsed laser systems,” *Int. Congr. Appl. Lasers Electro-Opt.*, vol. 2002, no. 1, p. 157512, Oct. 2002.
[30] S. Amoruso *et al.*, “Study of the plasma plume generated during near IR femtosecond laser irradiation of silicon targets,” *Appl. Phys. A*, vol. 79, no. 4–6, pp. 1377–1380, Sep. 2004.
[31] S. Amoruso *et al.*, “Characterization of laser ablation of solid targets with near-infrared laser pulses of 100fs and 1ps duration,” *Appl. Surf. Sci.*, vol. 252, no. 13, pp. 4863–4870, Apr. 2006.
[32] D. Breitling, A. Ruf, and F. Dausinger, “Fundamental aspects in machining of metals with short and ultrashort laser pulses,” presented at the Lasers and Applications in Science and Engineering, San Jose, Ca, 2004, p. 49.
[33] B. A. Dajnowski, “Laser ablation and processing methods and systems,” Mar-2016.
[34] E. D. Palik, *Handbook of optical constants of solids*, vol. 3. Academic press, 1998.
[35] A. Sikora *et al.*, “Picosecond laser micromachining prior to FIB milling for electronic microscopy sample preparation,” *Appl. Surf. Sci.*, vol. 418, pp. 607–615, Oct. 2017.
[36] B. Kazemi Esfeh, M. Rack, S. Makovejev, F. Allibert, and J.-P. Raskin, “A SPDT RF Switch Small- and Large-Signal Characteristics on TR-HR SOI Substrates,” *IEEE J. Electron Devices Soc.*, vol. 6, pp. 543–550, 2018.
[37] X. Li and Y. Zhang, “Flipping the CMOS Switch,” *IEEE Microw. Mag.*, vol. 11, no. 1, pp. 86–96, Feb. 2010.
[38] S. Liu, L. Zhu, F. Allibert, I. Radu, X. Zhu, and Y. Lu, “Physical Models of Planar Spiral Inductor Integrated on the High-Resistivity and Trap-Rich Silicon-on-Insulator Substrates,” *IEEE Trans. Electron Devices*, vol. 64, no. 7, pp. 2775–2781, Jul. 2017.
[39] Wei Jin, Weidong Liu, Chaohe Hai, P. C. H. Chan, and Chenming Hu, “Noise modeling and characterization for 1.5-V 1.8-GHz SOI low-noise amplifier,” *IEEE Trans. Electron Devices*, vol. 48, no. 4, pp. 803–809, Apr. 2001.
[40] R. Paulin, P. Cathelin, G. Bertrand, A. Monroy, J. More-He, and T. Schwartzmann, “A 12.7dBm IIP3, 1.34dB NF, 4.9GHz–5.9GHz 802.11a/n LNA in 0.13 μm PD-SOI CMOS with Body-Contacted transistor,” in *2016 IEEE MTT-S International Microwave Symposium (IMS)*, San Francisco, CA, 2016, pp. 1–3.
[41] D. K. Shaeffer and T. H. Lee, “A 1.5-V, 1.5-GHz CMOS low noise amplifier,” *IEEE J. Solid-State Circuits*, vol. 32, no. 5, pp. 745–759, 1997.

Chapter 4: Femtosecond Laser Assisted Micromachining and Etch (FLAME) for SOI-CMOS technology

4.0 Introduction

This chapter presents the results of first part of the experimental work which is concerned with the formation of reliable membranes on SOI circuits. In section 4.1, the measurement of beam waist is performed for the two laser sources used in the work. In section 4.2, scribing of trenches on silicon is studied in order to highlight the impact of different experimental parameters. Various trench measures are studied in order to understand the different ablation regimes of operation. After having determined the optimal parameters for processing, laser milling is presented in section 4.3. The additional complexities to be taken into consideration for formation of cavities with controlled depth distributions are highlighted here. Finally, the FLAME process is described in section 4.4 which enables the formation of local membranes of circuits over desired areas.

Content

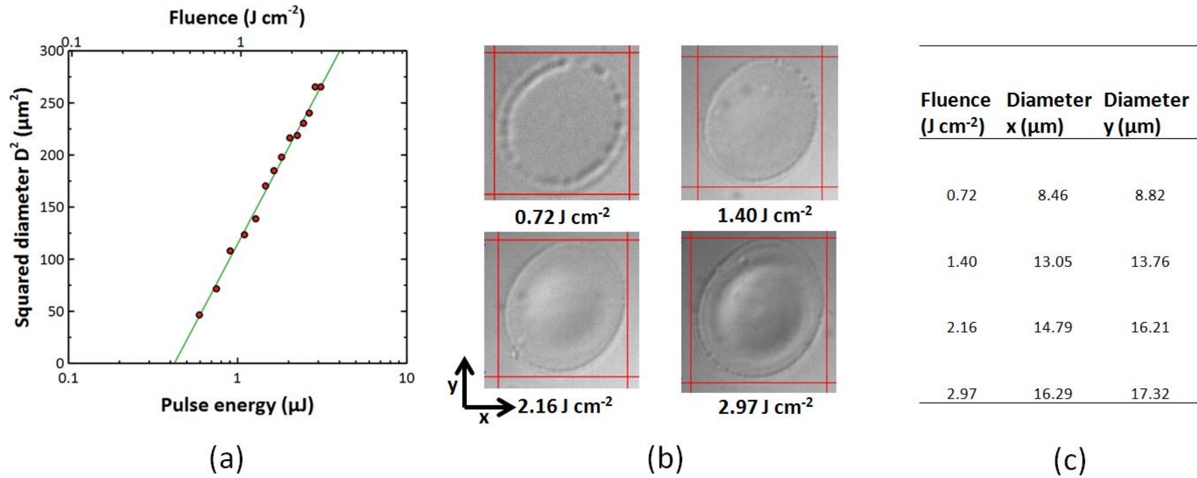
4.1 Determination of laser beam waist and threshold fluence 117
4.2 Scribing of trenches in silicon 119
4.2.1 Galvo stability: 343 nm and 1030 nm line 120
4.2.2 Study of trenches: 343 nm source 121
4.2.2.1 Effect of speed 121
4.2.2.2 Effect of fluence 122
4.2.3 Study of trenches: 1030 nm source 123
4.2.3.1 Effect of speed 124
4.2.3.2 Effect of fluence 124
4.2.3 Detailed study of trench parameters for optimization of milling 126
4.2.3.1 Average trench depth (d_t) 127
4.2.3.2 Recast layer volume (V_r) 131
4.2.3.3 Recast layer average height (h_{avg}) 132
4.2.3.4 Average roughness of trench profile (R_a) 133
4.2.3.5 Trench width (W_t) 134
4.3 Laser milling 136
4.3.1 Choice of milling pitch 136
4.3.2 Determination of depth as a function of number of scans 138
4.3.3 Effects of ablation plume and redeposition on laser milling 139
4.3.3.1 Generation of particles 140
4.3.3.2 Screening of laser radiation 142
4.3.3 Analysis of milling quality 146
4.4 Femtosecond Laser Assisted Micromachining and Etch (FLAME) process 152
Concluding Remarks 154
References 155

4.1 Determination of laser beam waist and threshold fluence

The experiments are performed on n-type wafer grade silicon of resistivity 10 Ω.cm. The surface roughness of the wafer is of the order of few nm. For the application on SOI substrate, the ablation is performed on silicon of resistivity ~10 kΩ.cm. Despite three orders of magnitude difference in resistivity, we found in our experiments that ablation depth did not change appreciably for the two resistivities. Hence, the interpretations made during different studies hold good for both values of resistivities.

The first step in our experiments is the determination of beam waist. This is an important parameter in femtosecond laser ablation. It allows conversion of pulse energy to fluence. Most of the laser ablation studies report fluence instead of pulse energy because fluence is a normalized energy quantity which describes the surface energy distribution for each laser pulse. The beam waist is determined by plotting the squared diameter (D^2) of the ablated region as a function of applied laser energy (E_p). This is a common method of determining beam waist as first described by Liu et al. [1]. The expression of the squared diameter is given as:

$$D^2 = 2w_0^2 \ln(\frac{E_p}{E_{th}}) \tag{4.1}$$



Fluence (J cm^{-2})	Diameter x (µm)	Diameter y (µm)
0.72	8.46	8.82
1.40	13.05	13.76
2.16	14.79	16.21
2.97	16.29	17.32

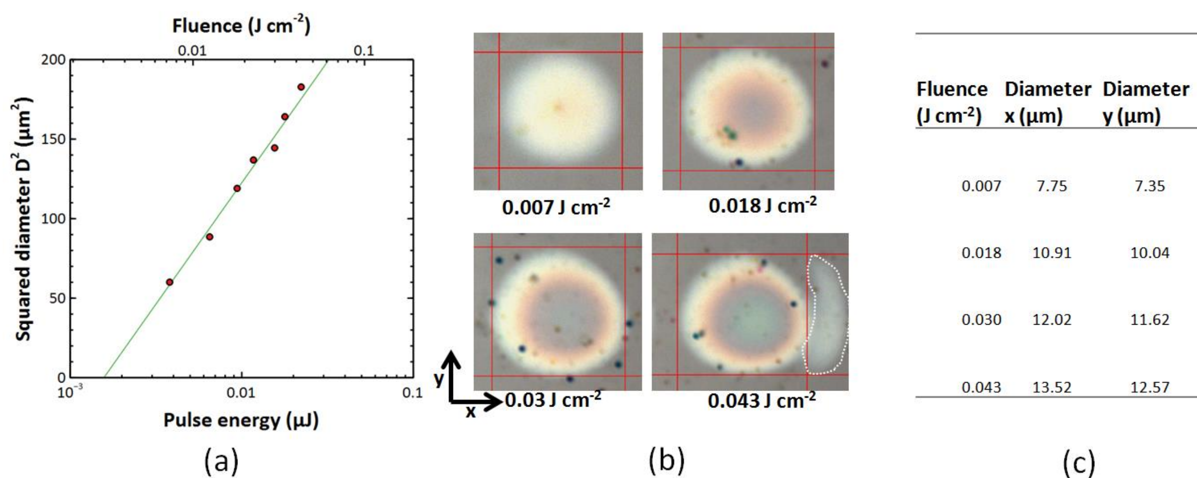
Fig. 4.1: (a) Squared diameter (x) plotted as a function of pulse energy for 1030 nm source. Green line is the least squares fit of the experimental data (b) Optical micrographs of ablated area shown for 4 different fluences (c) Measured diameter in x and y direction for different fluences

By curve fitting of experimental data points using the above equation, the values of two unknowns E_{th} and w_0 are determined. From these values, threshold fluence can be calculated. This calculated value represents the threshold fluence

for modification as the outermost diameter of the laser affected region is measured.

The graph of D^2 as a function of threshold fluence is plotted on a semi-logarithmic scale for the 1030 nm source in Fig. 4.1. From this, the beam waist is determined to be 8.2 μm and the single pulse threshold fluence is 0.42 J cm^{-2}.The optical micrographs of the laser affected areas are shown in Fig. 4.1b. Since there is not a big difference between the measured x and y diameters, a single value of the beam waist is used.

The same experiment is performed on the 343 nm source and plotted in Fig. 4.2. The spot size and threshold fluence are determined to be 5.75 μm and 3 mJ cm^{-2} respectively. The modification threshold for 343 nm is 3 orders of magnitude lower than the 1030 nm source. This is because the threshold fluence reduces with decreasing wavelength and for femtosecond pulses the effect is more pronounced [2] . Fewer points are used for determination of ablation threshold and spot size because of beam distortion effect starting from a fluence of 42.7 mJ cm^{-2}. For higher fluences, the beam distortion effects are clearly observable as depicted in Fig. 4.3. Beam distortion results is asymmetric energy deposition profiles and are not desirable for micromachining applications [3]. An additional possible reason of beam distortion is imperfections in the optical path of the laser beam. A central dark spot is also seen appearing at a fluence of 0.082 J cm^{-2}. This is a well-known phenomenon and corresponds to the region of silicon ablation as described by Bonse et al [4]. The single pulse ablation threshold is between 0.067 J cm^{-2} and 0.082 J cm^{-2}.



Fluence (J cm^{-2})	Diameter x (μm)	Diameter y (μm)
0.007	7.75	7.35
0.018	10.91	10.04
0.030	12.02	11.62
0.043	13.52	12.57

Fig. 4.2: (a) Squared diameter (x) plotted as a function of pulse energy for 343 nm source. Green line is the least squares fit of the experimental data (b) Optical micrographs of ablated area shown for 4 different fluences. Outlined region for 42.7 mJ cm^{-2} shows the distortion of beam and redistribution of incident energy (c) Measured diameter in x and y direction for different fluences

While the single pulse modification threshold has been determined in this experiment, this value of fluence will not be used in calculation of depth as the micromachining is a multi-pulse ablation process with a certain spatial offset between the two pulses which reduces the threshold for ablation. The treatment of multi-pulse ablation is usually done with the determination of incubation co-efficient. This parameter relates the single pulse and multi-pulse ablation threshold (given by equation 3.12) [5]. However, accurate modelling of incubation phenomenon requires a more involved mathematical treatment as described by Žemaitis et al. [6].

A simplified treatment of laser processing of silicon is presented based on the concept of total accumulated fluence and effective number of pulses as described in work of Crawford et al. [7]. This will be discussed in more detail for scribing of trenches in silicon in the next section.

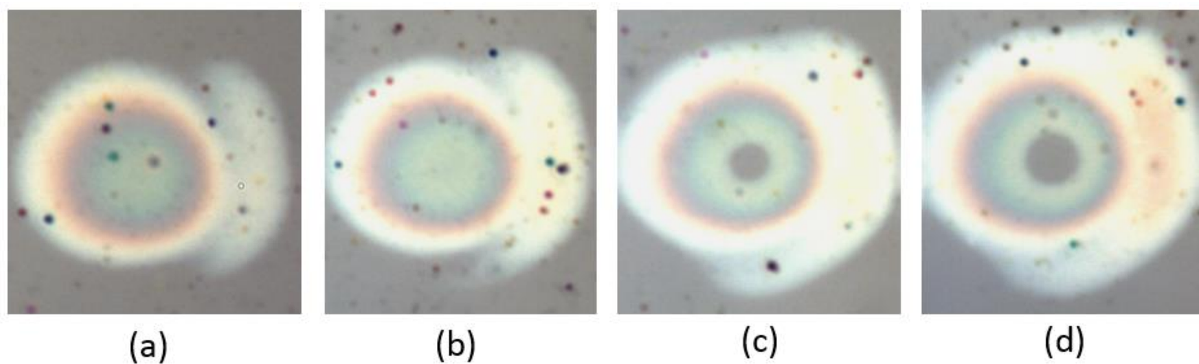


Fig. 4.3: Optical micrographs depicting evolution of laser spot with increasing laser fluence (a) 0.055 J cm^{-2} (b) 0.067 mJ cm^{-2} (c) 0.082 mJ cm^{-2} (d) 0.097 mJ cm^{-2}

4.2 Scribing of trenches in silicon

The scribing of trenches/lines in silicon is analyzed for high resolution 343 nm source and the high power 1030 nm source with the focal plane of the laser beam corresponding to the sample surface. The stability of galvo system is studied first to determine the limits for maximum usable galvo speed. Scribes are made in a serpentine path to determine the stability of the galvo mirror system. Then trenches are studied for varying fluence and scan speeds.

Several analyses on trenches are performed to determine how the laser parameters affect micromachining of silicon. This serves as the starting point for the development of the laser milling process.

4.2.1 Galvo stability: 343 nm and 1030 nm line

Galvo stability is an important factor during 2D movement of the laser beam. The galvo system consists of two separate mirror sources for x and y translation of the beam. During unidimensional movement, only one of the two mirrors needs to move. The stability of the beam is better in this case. However, when 2D geometries have to be traversed, the two mirrors need to work in coordination. This can result in instabilities of the resulting beam trajectory. In our laser system, two sets of mirrors and lenses are used separately for the 343 nm and 1030 nm sources.

In order to study the stability of movement, a serpentine trajectory is used with length 200 µm and separation of 80 µm as shown in Fig. 4.4. For the 343 nm galvo system, the movement is very stable up to 40 mm s^{-1}. At higher speeds, the beam movement is slightly distorted at the turns of the serpentine. But the overall trajectory looks satisfactory for 2D micromachining even at 100 mm s^{-1}. The galvo stability can be improved further by efficient implementation of the PID servo control loop of the scanner.

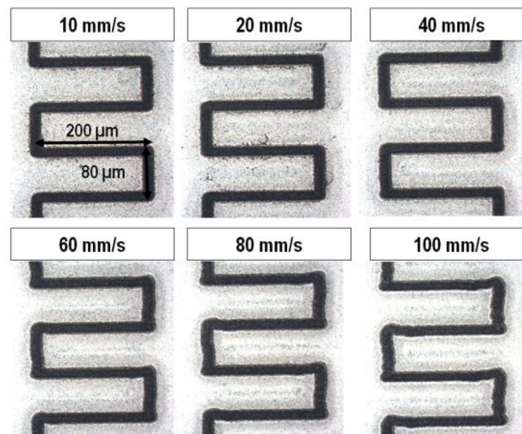


Fig. 4.4: 2D laser scribing at different galvo speeds for 343 nm source

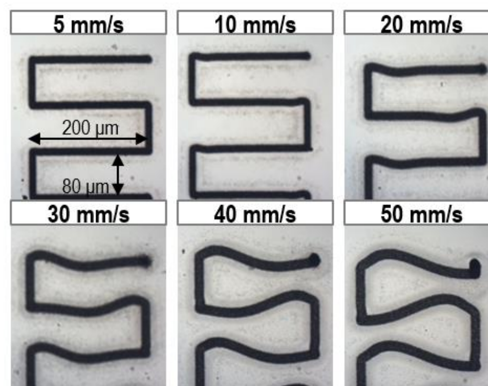


Fig. 4.5: 2D laser scribing at different galvo speeds for 1030 nm source

In the case of 1030 nm source, as seen in Fig. 4.5, the galvo is less stable. The beam stability is satisfactory up to 20 mm s^{-1}. Starting from 30 mm s^{-1}, the beam stability is poor and at 50 mm s^{-1}, the intended trajectory is largely distorted. This instability prevents the use of higher speeds which are preferable for milling applications. It is to be noted that for both galvo systems, unidimensional scribing is stable for all the reported speeds (up to 100 mm s^{-1}). Hence, study of unidimensional trench scribing is possible for all speeds.

4.2.2 Study of trenches: 343 nm source

The profile of the trench is studied along the length of the trench by taking a profile at its central axis by using optical profilometer scans of the trench. The profile is studied for varying scan speeds and laser fluence.

4.2.2.1 Effect of speed

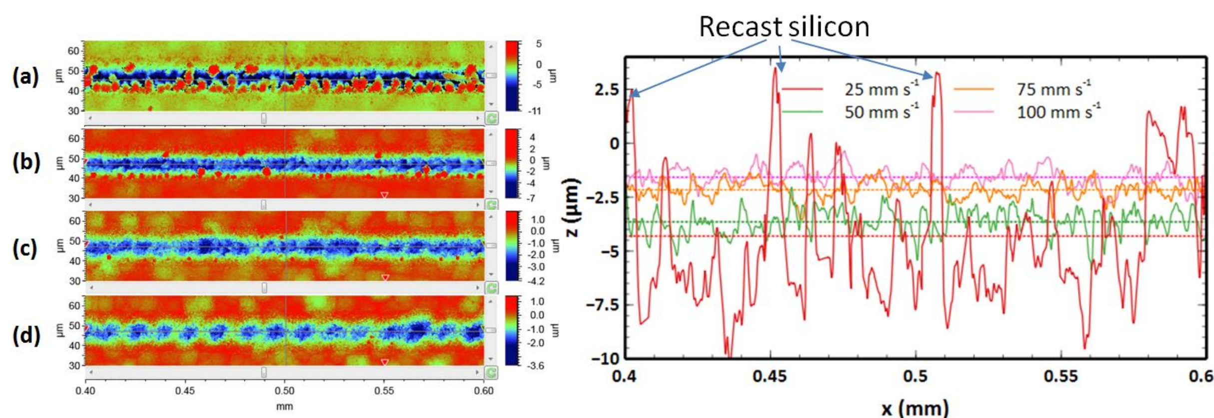


Fig. 4.6: Profile of trenches taken along the centre of the trench for 343 nm source at fluence of 4.2 J cm^{-2} and repetition rate 200 kHz and scan speeds of (a) 25 mm s^{-1} (b) 50 mm s^{-1} (c) 75 mm s^{-1} (d) 100 mm s^{-1}

In order to study the optimal value of scan speed, the profiles are studied at a value of laser power 0.44 W which corresponds to a peak fluence of 4.2 J cm^{-2}. The profilometer scans along with the depth profiles taken at the centre are shown in Fig. 4.6. The profiles reveal that smaller scan speeds result in higher deposition of recast ablated material. The recast layer can be visualized in the profilimeter scans as red spots surrounding the ablated trench. At 25 mm s^{-1}, accumulation of expelled material takes place both within and around the ablated trench. With increasing speed, the quantity of recast ablated material is reduced (as observed by red dots around the ablated area in the profilometer scan). A similar effect has been observed in the study of ablation of InP by Boroweic et al [8].

4.2.2.2 Effect of fluence

It has been determined that higher scan speed is better for scribing of lines. To scale up the ablation rate, higher fluence needs to be used. Fig. 4.7 depicts the results obtained for scribing of trenches at varying laser fluence. The depth variation follows a periodic pattern resembling a sinusoid. The amplitude of the sinusoid increases for increasing laser fluence. At 9 J cm^{-2}, the local depths can reach below the profilometer measurement limit. Thus, high fluence milling at high speeds results in enhanced local depths in the form of micro-holes in the trench which can far exceed the average depth. The periodicity of the profile is found to depend on the scan speed. This is discussed in section 4.2.3.4.

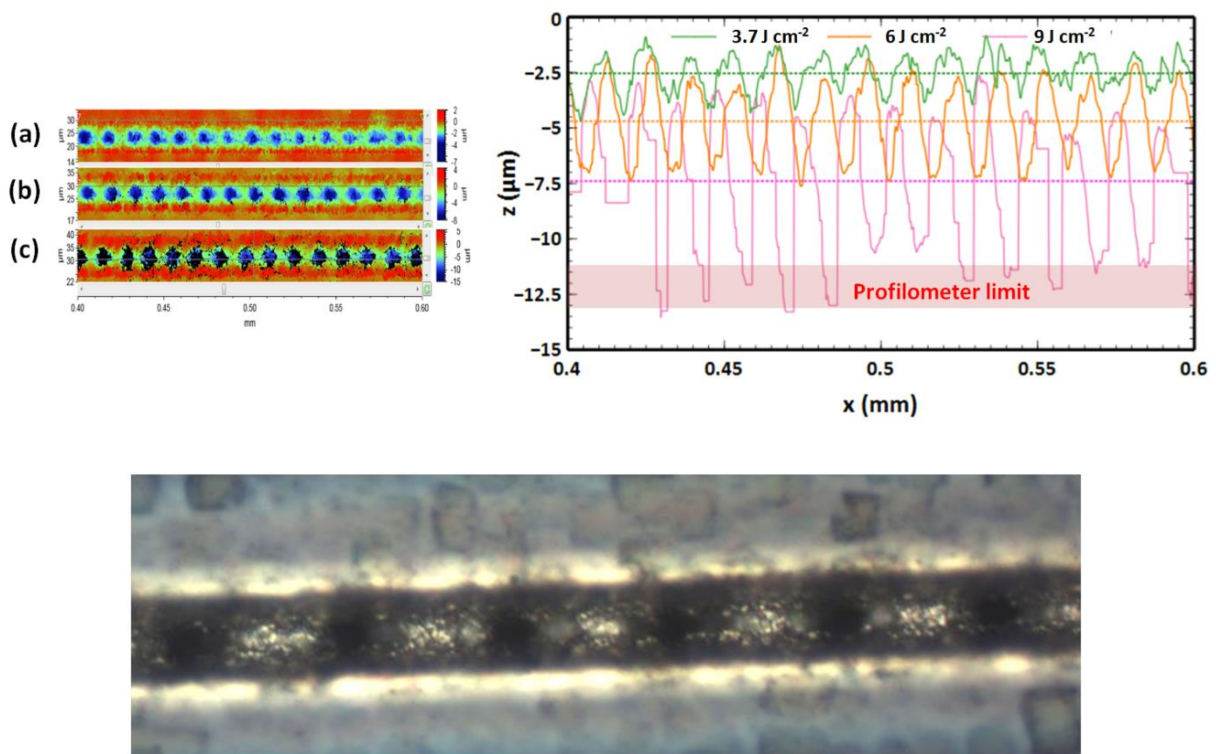


Fig. 4.7: Profile of trenches along the centre of the trench for 343 nm source at scan speed of 100 mm s^{-1} and pulse repetition rate 200 kHz and fluence of (a) 3.7 J cm^{-2} (b) 6 J cm^{-2} (c) 9 J cm^{-2}. The microscope image shows the micro-holes at the bottom of the trench

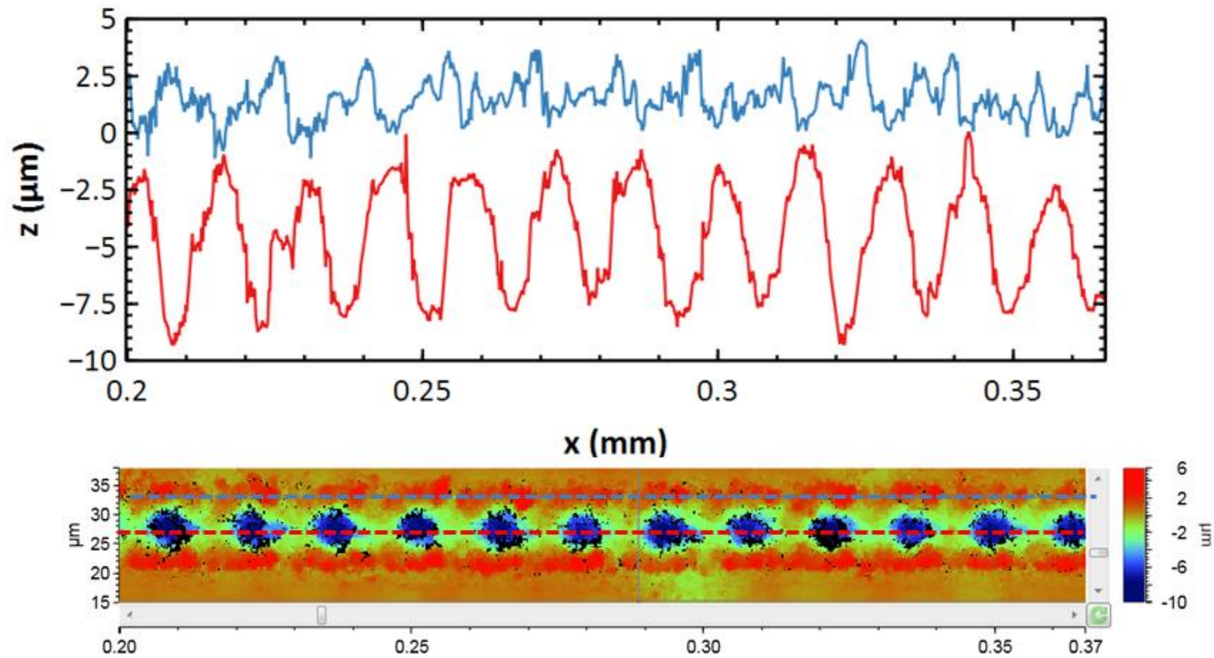


Fig. 4.8: Profile of trench compared with the profile of recast layer for trench scribed at 6 J cm^{-2}. The red line is the profile at trench centre while blue line represents the recast layer profile

The periodic nature of the trench profile seems to come from periodic modulation of absorbed energy in the substrate. This can be confirmed by taking the profile of the trench scribed at 6 J cm^{-2} and comparing it with the recast layer profile as shown in Fig. 4.8. It can be observed that at local regions where depth is higher, the height of the recast layer is also larger. While the exact phenomenon of periodic modulation of depth is not known, such an effect can be confirmed to exist by direct experimental evidence. This has also been reported in literature [7], [9]. The presence of microholes can be detrimental to micromachining quality.

4.2.3 Study of trenches: 1030 nm source

The 1030 nm source has a greater range of pulse energy because it corresponds to fundamental emission of the line. With the use of high pulse energy, larger average power can be obtained at a much lower repetition rate. This is not possible in the 343 nm source as lowering the pulse repetition rate would result in small average power outputs. For instance, reducing the repetition rate from 200 kHz to 20 kHz, maximum possible pulse energy would be 0.1 W which is very low for milling applications.

An average power of 1 W or above is necessary to keep the ablation rate sufficiently high for our application. In case of 1030 nm source, two choices are available (i) High pulse energy and low pulse repetition rate (ii) Low pulse energy

and high pulse repetition rate. With the use of choice (ii), under similar conditions as UV source, microhole features have been observed in 1030 nm source.

Hence, it is interesting to explore choice (i) for micromachining. The pulse repetition rate is reduced so as to obtain a maximum laser power of ~1.5 W. This yields a repetition rate of 30 kHz.

4.2.3.1 Effect of speed

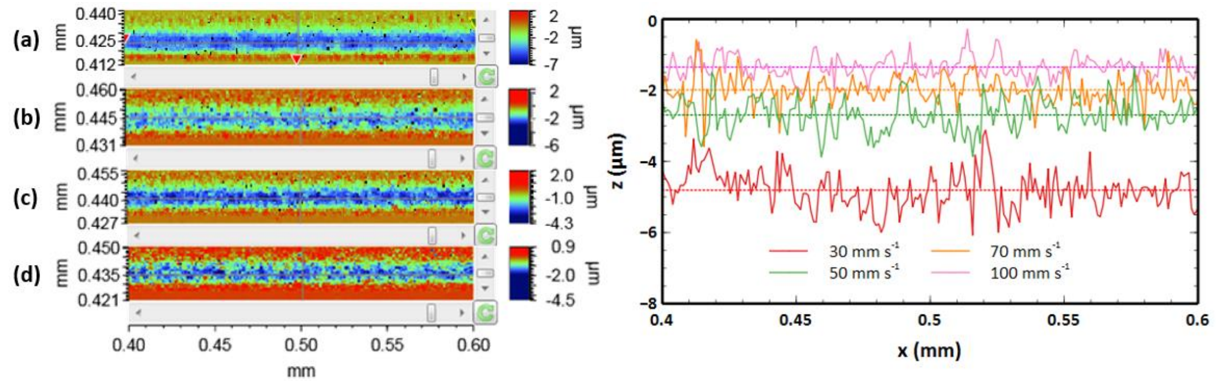


Fig. 4.9: Profile of trenches along the centre of the trench for 1030 nm source at fluence of 13.6 J cm^{-2} and repetition rate 30 kHz and scan speeds of (a) 25 mm s^{-1} (b) 50 mm s^{-1} (c) 75 mm s^{-1} (d) 100 mm s^{-1}

As seen in Fig. 4.9, for the same galvo speed and power, the profiles are improved as compared to 343 nm source. There is no appearance of recast layer close to the trench centre at low scan speed. This allows a larger range of scan speed values to be used for micromachining. The average depths for 1030 nm source are lower are compared to 343 nm source. An important reason for this is that efficiency of ablation is reduced at high fluences [10]. Also, the physical mechanisms of ablation are different for the two wavelengths. The recast layer reduces at higher speeds as in the case of 343 nm source.

4.2.3.2 Effect of fluence

Scribing of lines is studied at higher fluence at a speed of 100 mm s^{-1} where the material removal rates are improved. There is no appearance of periodic microholes like in the case of 343 nm source as seen in Fig. 4.10. However, the obtained depths are smaller because of less efficiency and lesser effective number of pulses which reduces incubation. Additionally, it is seen in the galvo stability experiments that only scan speeds up to 30 mm s^{-1} are usable for 2D micromachining.

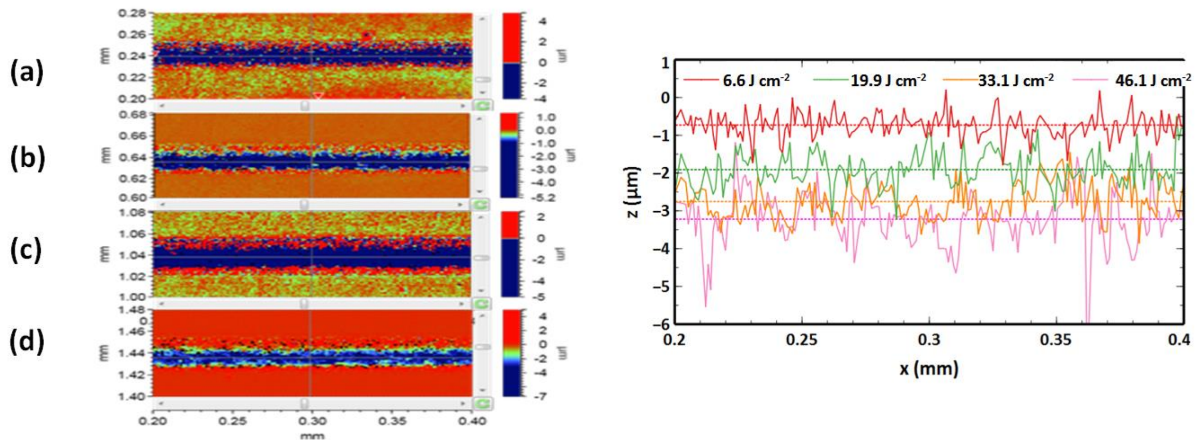


Fig. 4.10: Profile of trenches along the centre of the trench for 1030 nm source at scan speed of 100 mm s^{-1} repetition rate 30 kHz and fluence of (a) 6.6 J cm^{-2} (b) 19.9 J cm^{-2} (c) 33.1 J cm^{-2} (d) 46.1 J cm^{-2}

Hence, the same study is performed for scan speed of 30 mm s^{-1} as shown in Fig. 4.11. In this case, larger depths can be attained without formation of microholes. Therefore, these parameters are found to be suitable for laser milling.

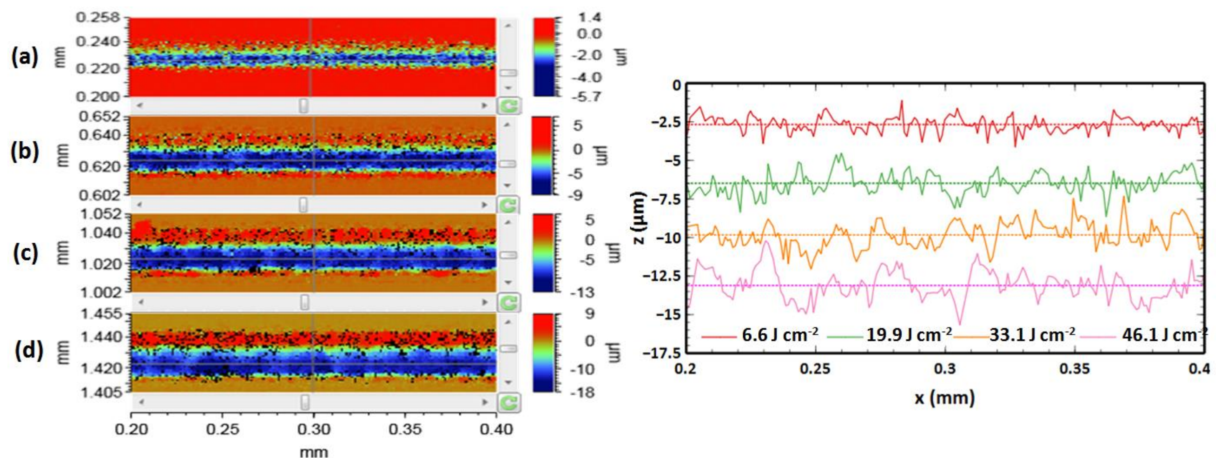


Fig. 4.11: Profile of trenches along the centre of the trench for 1030 nm source at scan speed of 30 mm s^{-1} repetition rate 30 kHz and fluence of (a) 6.6 J cm^{-2} (b) 19.9 J cm^{-2} (c) 33.1 J cm^{-2} (d) 46.1 J cm^{-2}

A basic study of the trenches for a few set of parameters has been seen so far highlighting some aspects of scribing of trenches. In order to further understand the impact of laser process parameters on trench scribing in more detail, a comprehensive study of trenches is described in the following section.

4.2.3 Detailed study of trench parameters for optimization of milling

The axial profiles of trenches have been analyzed so far only for few cases of fluence and speed. It is also seen that the recast layer becomes prominent with increasing depth of the trench. Our final objective is performing 2D milling which is essentially an extension of trench scribing. The trench is the basic unit of 2D milling and hence optimization of 2D milling requires optimization of the trench. Thus, it is necessary to characterize the trenches and study different measures which represent the quality of the trench. In this section, a comprehensive analysis of the trenches is presented for three different cases that have already been mentioned: 1) 343 nm source: High repetition rate (200 kHz) 2) 1030 nm source: High repetition rate (200 kHz) 3) 1030 nm source: Low repetition rate (30 kHz). In each case, the pulse energy is varied such that the range of average power is ~0 – 2 W. This means that for case 3, relatively high values of fluence are reached for the laser pulse. Since we perform different analyses on the recast layer, we take the profilometer scans immediately after laser scribing without performing cleaning of substrate in ultrasonic bath. We define the recast layer as the aggregate of debries/redeposited material in the vicinity of the ablated trench.

The following trench analyses are performed using profilometer scans:

1. Average trench depth (d_t): This value is determined by summing the different height values comprising the axial trench profile and dividing by the number of points.

$$d_t = \frac{(z_1 + z_2 + z_3 + \cdots z_N)}{N} \tag{4.2}$$

2. Recast layer volume (V_r): The points representing the recast layer form a projected 2D surface extending above the substrate. The recast layer volume is calculated by computing the volume enclosed by this surface and the reference level of the substrate (zero level).

3. Recast layer average height (h_{avg}): This is the average height of the points representing the recast layer in the profilometer scan. The recast layer is obtained by masking the points which are at and above the level of the substrate (zero level). After masking, only the points above the level of the substrate are visible in the scan of the trench. By taking an average of all these points, the average height of the recast layer can be calculated.

4. Average roughness of the trench profile (R_a): This value represents the roughness of the cut. This is obtained by calculating the absolute deviation of each point of the profile from the average depth, summing up the deviations and dividing the sum by number of points.

$$R_a = \frac{|z_1 - d_t| + |z_2 - d_t| + |z_3 - d_t| + \cdots |z_N - d_t|}{N} \tag{4.3}$$

5. Trench Inner/Outer width (W_{ti}/W_{to}): This value represents the effective width of cut for the trenches. When a trench is scribed in Silicon, a recast layer is deposited on the sides of the trench. In the cross-sectional profile of the trench, the trench bottom is seen surrounded by the recast layer on either sides. The inner and outer trench widths are defined as shown in Fig. 4.12.

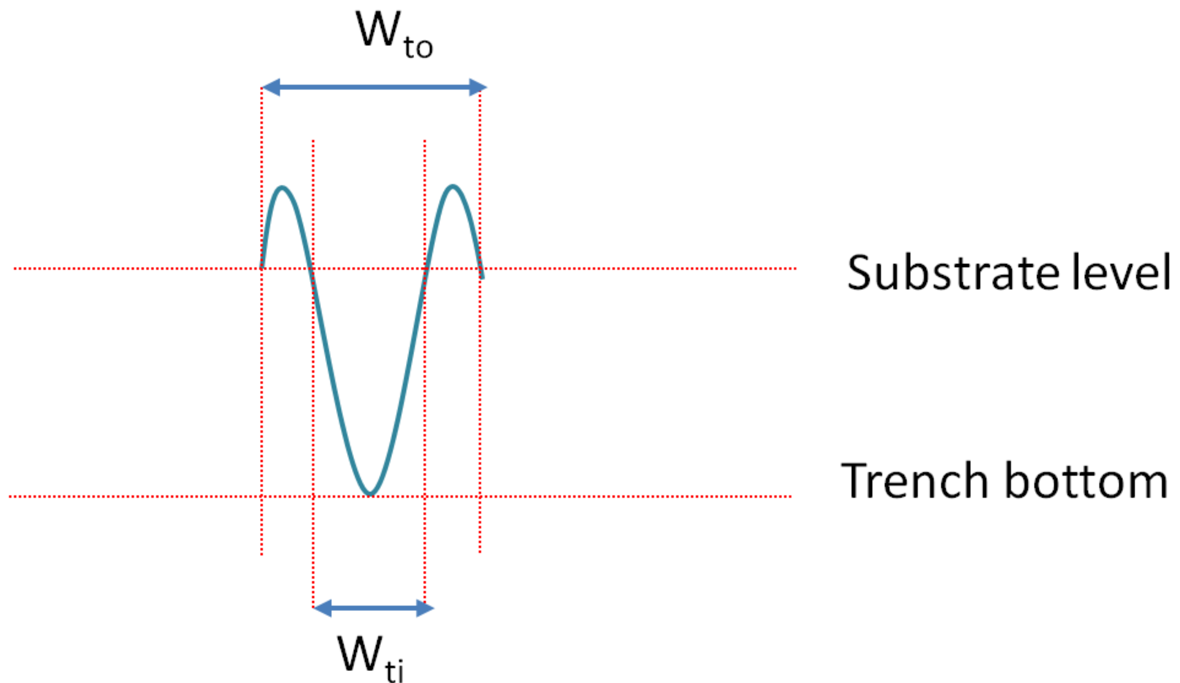


Fig. 4.12: Cross sectional profile of a trench

4.2.3.1 Average trench depth (d_t)

For the 3 cases, the average trench depth is shown in Fig. 4.13. The depths plotted on a semi-logarithmic scale shows two ablation regimes with different ablation thresholds. This is a well known phenomenon which has been reported in metals [10][11], semiconductors [6][7] and dielectrics [13]. The depths in the two regimes can be given by:

$$d_t^{g,s} = l_{g,s} \log\left(\frac{\phi_0}{\phi_{th}^{g,s}}\right)$$

$d_t^{g,s} - \mathit{Average\ trench\ depth\ in\ gentle/strong\ ablation\ regimes}$

$\phi_{th}^{g,s}$ – *Threshold fluence in gentle/strong ablation regimes*

$l_{g,s}$ – *Characteristic ablation depth in gentle/strong ablation regimes*

The transition fluence between these two regimes is determined by the scan speed when all other laser process parameters are fixed. Higher scan speeds allow utilization of higher fluence (average power) in the gentle ablation regime. Thus, higher removal rates can be obtained by using higher scan speeds while still operating in the soft ablation regime.

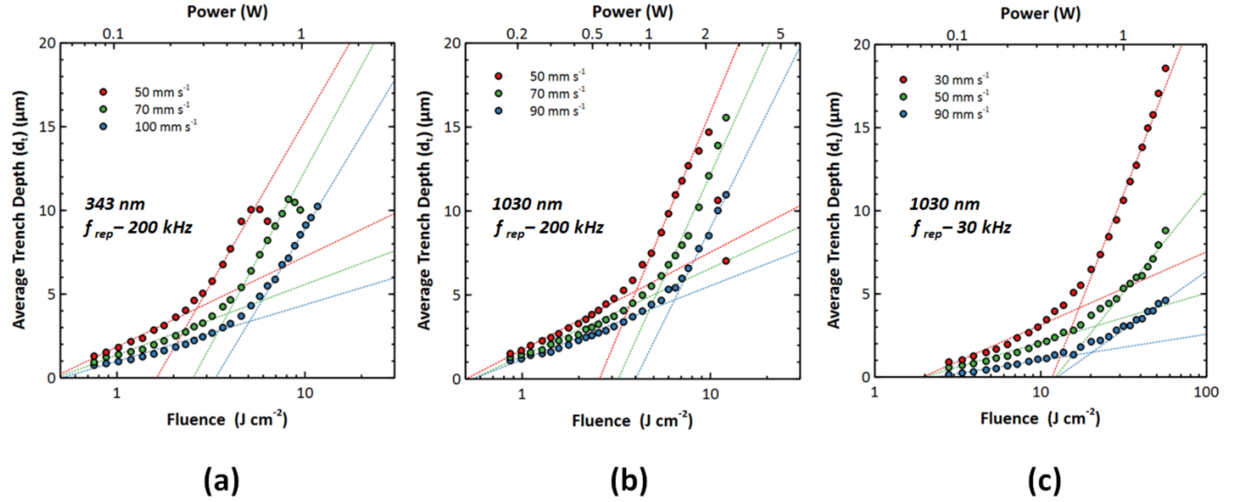


Fig. 4.13: Average trench depth plotted as a function of fluence for 3 cases at different scan speeds

Table 4.1: Parameters extracted from average trench depth using curve fitting

Parameter	Speed (mm s^{-1})	l_g (μm)	ϕ_{th}^{g} ($J\ cm^{-2}$)	l_s (μm)	ϕ_{th}^{s} ($J\ cm^{-2}$)
343 nm f – 200 kHz	50	2.34	0.45	8.45	1.63
	70	1.83	0.48	9.03	2.55
	100	1.47	0.51	8.11	3.35
1030 nm f – 200 kHz	50	2.53	0.5	11.71	2.58
	70	2.26	0.55	10.81	3.24
	90	1.91	0.55	9.83	4
1030 nm f – 30 kHz	30	1.91	1.94	11.12	11.73
	50	1.29	2	5.38	12.25
	90	0.68	2.34	3.07	12.6

From the plots of average depth as a function of fluence, curve fitting is done to fit the data points approximately into two regimes. The extracted values of parameters for the 3 cases at different speeds are tabulated in Table 4.1. On comparing the different parameters at repetition rate of 200 kHz, the observations agree with those in Table 3.1. With increase in wavelength, both the characteristic depths ($l_{g,s}$) and threshold fluences ($\phi_{th}^{g,s}$) increase. In the soft

ablation regime, at repetition rate of 200 kHz, the difference in ablation thresholds between 343 nm and 1030 nm source is small. For 1030 nm source, at repetition rate of 30 kHz, the threshold fluence for gentle ablation is much larger as compared to other two cases. Also, the characteristic depths in the gentle ablation regime are smaller at same speed for this case. Hence, low repetition rate condition is less efficient in terms of ablation efficiency.

The threshold fluence extracted for the gentle regime exceeds the ablation threshold value for single pulse ablation. This is contradictory because silicon is known to exhibit incubation effect and threshold fluence decreases with increasing number of pulses [4]. Scribing of trenches is essentially a multi-pulse ablation process and hence this contradiction should not exist. These threshold values are explained by looking at multi pulse ablation regimes of silicon. In our work, two regimes of ablation are observed. However, in other studies, it has been shown that silicon exhibits 3 regimes for multi-pulse ablation [14], [15]. This additional regime occurs at very low fluences. In our experiments, the lowest fluence that is used is ~0.7 J cm^{-2}. In literature, a fluence value of 0.45 J cm^{-2} for transition from first to second regime has been reported [14]. In this work, an 800 nm laser source was used with a pulse width of 100 fs. Hence, it is possible that we do not observe the first regime of ablation in our experiments for which threshold fluence value could be lower as compared to single pulse ablation threshold. Further investigation is needed to confirm this hypothesis.

However, there are several advantages of using low repetition rate. For the high repetition rate cases, average depth measured from axial trench profile starts to decrease at high fluence. This is observable at speeds of 50 and 70 mm s^{-1} for 343 nm source and at 50 mm s^{-1} for 1030 nm source. This is mostly likely due the formation of debries at the bottom of the trench as observed in [7] (Fig. 8b). The profile taken for 200 kHz repetition rate and 1030 nm source at a fluence of 10.1 J cm^{-2} is shown in Fig. 4.14 also seems to suggest presence of debris. There are locally deep regions of depth ~25 μm. This is likely the depth in which there is penetration of laser energy and ablation occurs. However, presence of debris along the axis of the trench causes the measured average depth to be much smaller (~5 μm). It can be also seen that the trench opening is really narrow because of a prominent recast layer. This reduction of average trench depth and narrowing is not seen for 1030 nm source and repetition rate of 30 kHz for the 3 reported speeds.

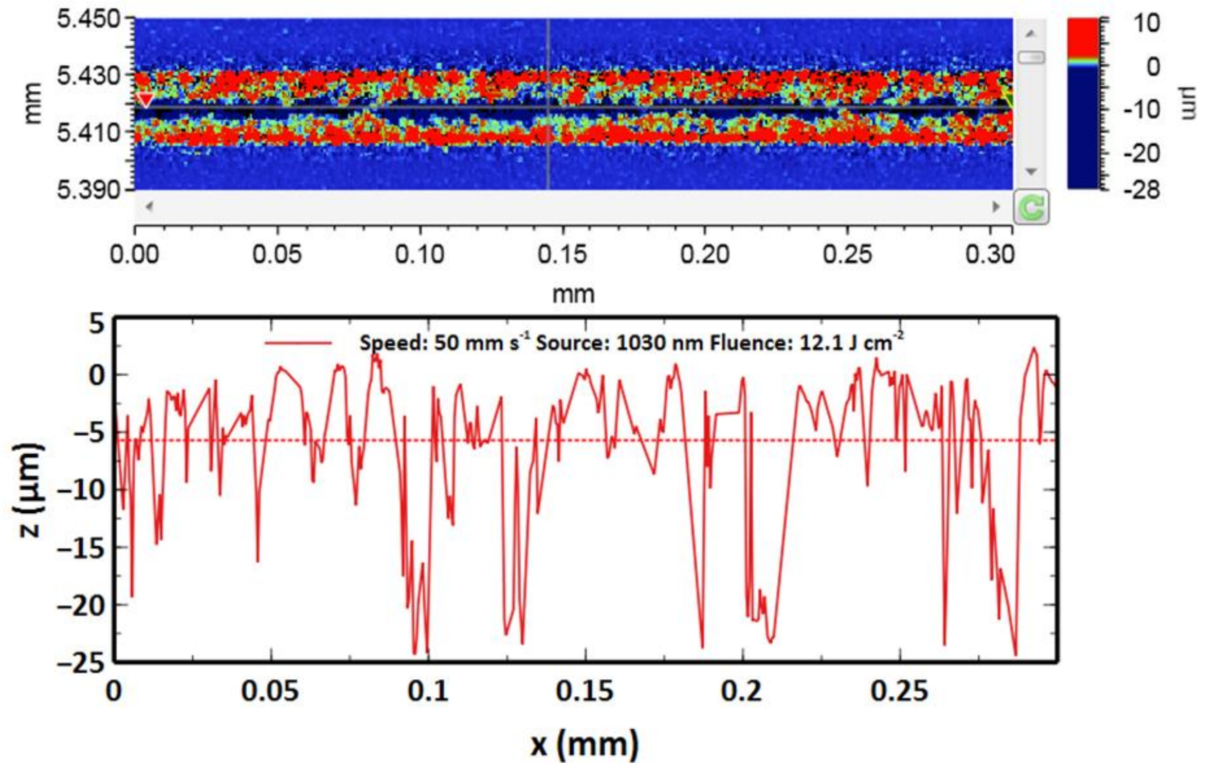


Fig. 4.14: Axial trench profile showing narrowing of trench and possible presence of debris

As for the interpretation of the two regimes, it has been studied in different works where different explanations are proposed. For metals, Nolte et al have used the two temperature model to explain the two ablation regimes [11]. At lower fluences, the optical penetration depth determines the trench depth which is the gentle ablation regime. At higher fluences, the effective heat penetration depth determines the trench depth which is the strong ablation regime. It is to be noted that this analysis holds good for metals for stationary multi-pulse ablation. The presence of similar ablation regimes was confirmed to exist in gold, silicon and silica in a comparative study made by Shaheen et al. [16]. In this study, for silicon, the threshold fluences for the two regimes are found to be 0.55 J cm^{-2} and 2.4 J cm^{-2}, respectively. A laser source of 785 nm wavelength with a pulse width of 130 fs is used in this work. This is very similar to the threshold fluences calculated in our system for 1030 nm source at repetition rate of 200 kHz and speed of 50 mm s^{-1}. At this speed, pulse to pulse distance of 0.25 µm is very small and similar ablation behaviour as static multi-pulse ablation could be expected.

Additional interpretations have been provided in other studies. In the detailed study of grooves (trenches) in silicon by Crawford et al., it is argued that scribing of trenches may not be equivalent to stationary multi-pulse ablation [7]. For scribing of trenches, there is a spatial displacement of laser beam with respect to the sample from one pulse to the next. Each pulse causes ablation and

modification of the surface. Thus, for subsequent pulses, the laser energy is incident on an uneven and changing surface which gives rise to local variation of fluence. Additionally, they found that at low translation speeds, a nearly linear dependence of trench depths as a function of pulse energy. At the same time, on a logarithmic scale, they could fit the data points into 2 regimes. They finally used the 2 ablation regimes model because despite differences, some commonalities exist between stationary and ablation involving translation. However, they expressed the view that linear dependence of ablation depths cannot be ruled out and that two different regimes may not be present.

In our study, we had the same question if a linear model or two-regime model better explains the data. In order to answer this question, several analyses of the recast layer are performed. It will become clear that two ablation regimes are indeed present for all 3 cases. This is discussed in the next two subsections.

Further on, presence of 2 ablation regimes for 1030 nm source at repetition rate of 30 kHz is explained. The spatial offset at a given speed is nearly one order of magnitude larger as compared to repetition rate of 200 kHz. This means that the effective number of pulses calculated from equation 3.11 is much smaller for this case. In the multi-pulse ablation study of sapphire by Ashkenasi et al. , gentle and strong ablation regimes are found to be dependent on two factors (i) Pulse fluence (ii) Number of shots [13]. It is found that at smaller fluence and number of pulses, gentle ablation is seen while at higher fluence and number of pulses, strong ablation occurs. In our experiments with source of 1030 nm and repetition rate of 30 kHz, we increase the fluence but reduce the number of effective pulses. Hence, it could well be expected that two regimes can exist even in this condition.

4.2.3.2 Recast layer volume (V_r)

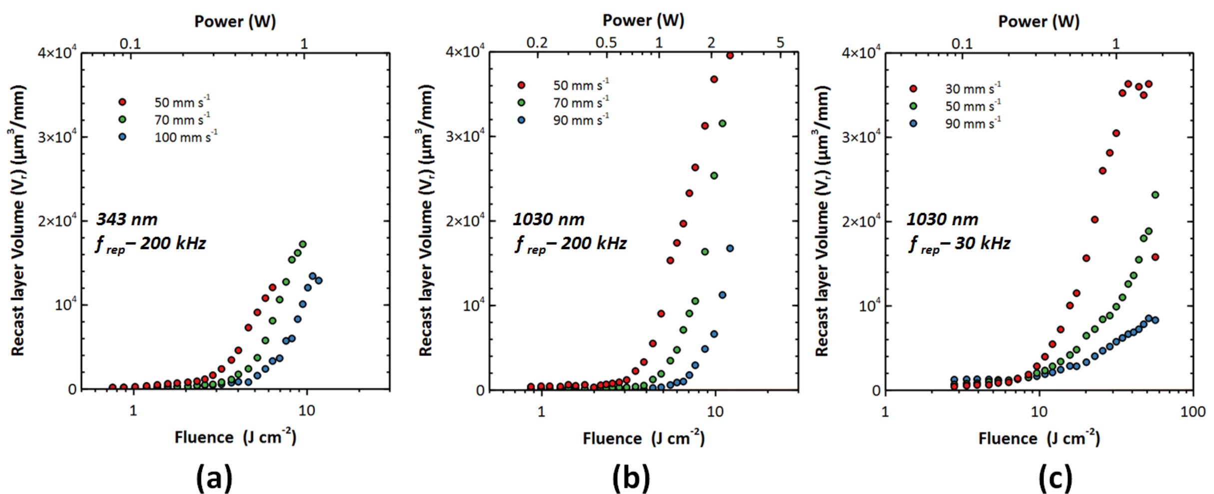


Fig. 4.15: Recast layer volume plotted as a function of fluence for 3 cases at different scan speeds

The recast layer volume obtained for the three cases is shown in Fig. 4.15. The unit μm^3 mm^{-1} refers to the volume of recast layer in μm^3 when scribing a trench of length 1 mm. This measure is a key indicator of the presence of two ablation regimes. For all 3 cases, the volume of recast layer is very low in the gentle ablation regime followed by abrupt increase which shows the onset of strong ablation. At a given speed, there is only a very small increase in ablation volume when going from lowest to highest fluence within the gentle ablation regime. Higher fluences out of this regime result in strong ablation and this is accompanied by an increase of volume of recast layer. In the strong ablation regime, the volume of recast layer is much more sensitive to the fluence. A large difference in volume is observed between highest and lowest fluence conditions within the strong ablation regime at a given speed.

By comparing Fig. 4.13 and Fig. 4.15, the threshold fluences for transition from one regime to the other as seen in the recast layer volume data are well correlated with those observed in average trench depth data. For the 1030 nm source and repetition rate of 30 kHz, at speed of 30 mm s^{-1}, volume of recast layer starts to decrease at very high fluences. When the scan is observed, it is seen that some parts of the recast layer is missing along the sides of the trench. This can be attributed to strong expulsion of ablated material at such fluences which possibly avoids formation of recast layer at certain places along the sides of the trench.

4.2.3.3 Recast layer average height (h_{avg})

The recast layer average height is plotted in Fig. 4.16. All the observations that are made for the recast layer volume hold true also for this measure.

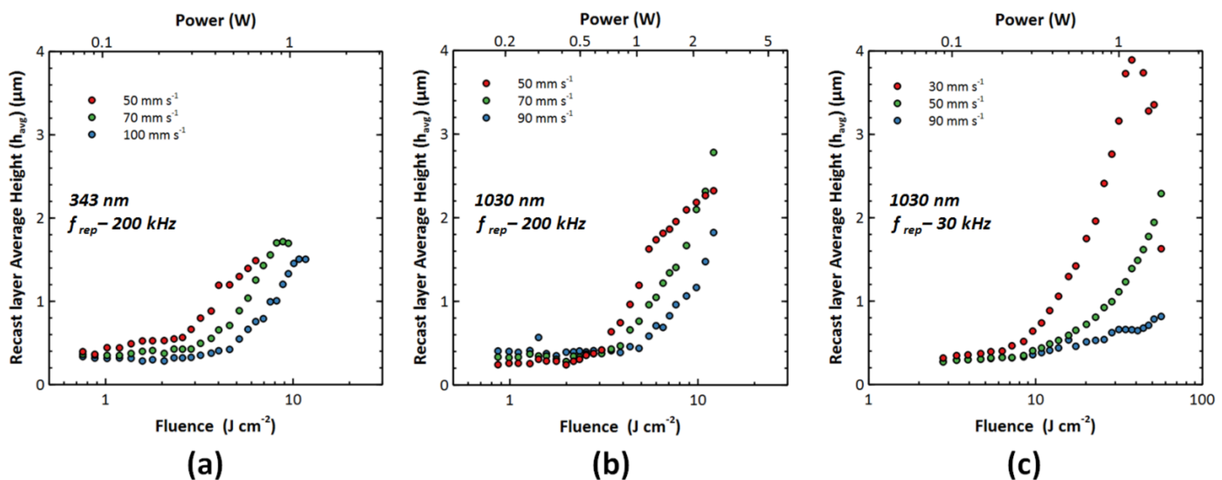


Fig. 4.16: Recast layer average height plotted as a function of fluence for 3 cases at different scan speeds

The recast layer average height for a given speed remains nearly constant during gentle ablation regime. After the onset of strong ablation regime, the height starts increasing. There are small inconsistencies in the observations. For instance, at repetition rate of 200 kHz in the soft ablation regime, the average height value is larger for 50 mm s^{-1} than for higher speeds for 343 nm source. While for 1030 nm source, the average height for 50 mm s^{-1} is lower than other speeds in the same regime. It would be reasonable to expect the average height to be larger for 50 mm s^{-1} than for higher speeds. This is because incubation is strongest at this speed. This inconsistency can be explained by the errors in setting zero reference. The height is calculated with respect to a zero reference position. We observed an error of ~100 – 200 nm in fixing the reference zero level. This error comes from the z-resolution for masking points in the profilometer software. By further optimization of zero reference, the correct trend can be known in the soft ablation regime. However, zero reference errors do not affect the ability to distinguish soft ablation regime from strong ablation.

4.2.3.4 Average roughness of trench profile (R_a)

The average roughness measured from the axial trench profile using equation 4.3 is plotted in Fig. 4.17. It can be seen that for the two 200 kHz repetition rate cases, very high roughness values of 2 – 3 µm can be reached. The profile is highly corrugated at high fluences as it had been already seen in Fig. 4.7. For the two cases, the increase in trench profile roughness shows the same behaviour as for the other measures already discussed. A nearly constant roughness is seen in the gentle ablation regime followed by an increase in roughness with respect to fluences in the strong ablation regime.

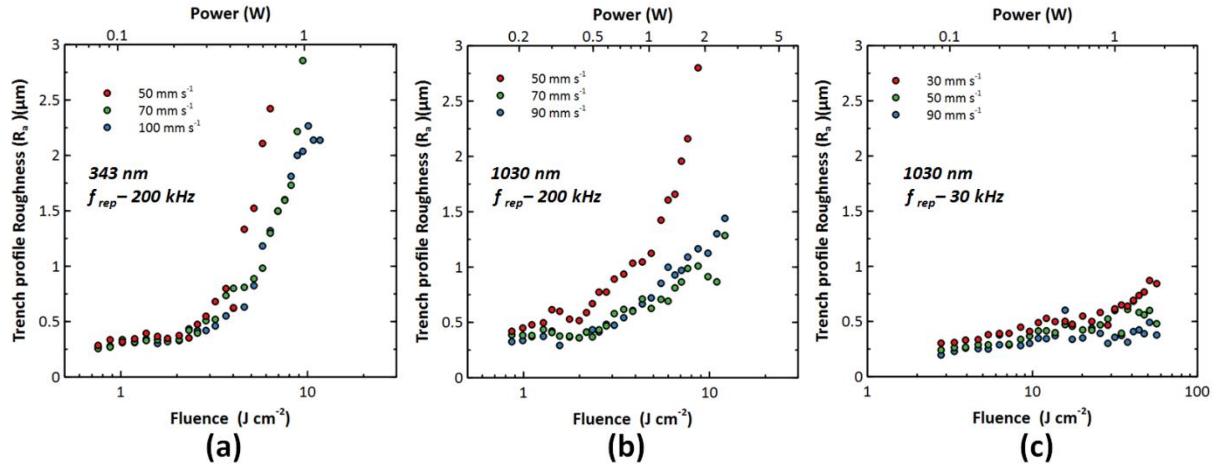


Fig. 4.17: Average roughness of trench plotted as a function of fluence for 3 cases at different scan speeds

Surprisingly, for the 1030 source at repetition rate of 30 kHz, the surface roughness shows a different behaviour. It is nearly constant over both the regimes with only a small increase at high fluences and the maximum value is < 1 μm. In laser milling, the aim for our application is to minimize surface roughness and low repetition rate with high fluence can potentially help minimize the surface roughness of milled area.

A nearly sinusoid-like profile is observed for the two cases with repetition rate of 200 kHz as it was shown in Fig. 4.7. It is found that the period of the sinusoid is proportional to the scan speed. For the 343 nm source, a Fourier analysis of profiles revealed periods of 6.9, 9.6, and 13.8 μm for speeds of 50, 70 and 100 mm s^{-1} respectively. The periodicity was not found to depend on the fluence.

While average roughness is one measure of roughness that we have showed, Root Mean Squared (RMS) roughness may also be used. Same trends can be observed by plotting the RMS roughness as a function of fluence. We find that both values are correlated with RMS roughness slightly higher than average roughness.

4.2.3.5 Trench width (W_t)

For measurement of trench width, the 1D cross-sectional profiles are taken perpendicular to the axis of the trench. The inner and outer trench widths are described as shown in Fig. 4.12. The width of the recast layer can vary considerably along the length of the axis. Hence, in order to obtain an average width, 10 cross-sectional profiles are taken at different points along the length of the axis. The average widths calculated from these profiles are plotted in Fig. 4.18. Only the inner trench widths are shown. When the outer trench widths were plotted, a large scatter is observed and no clear trend can be seen. Hence, this data is not presented here.

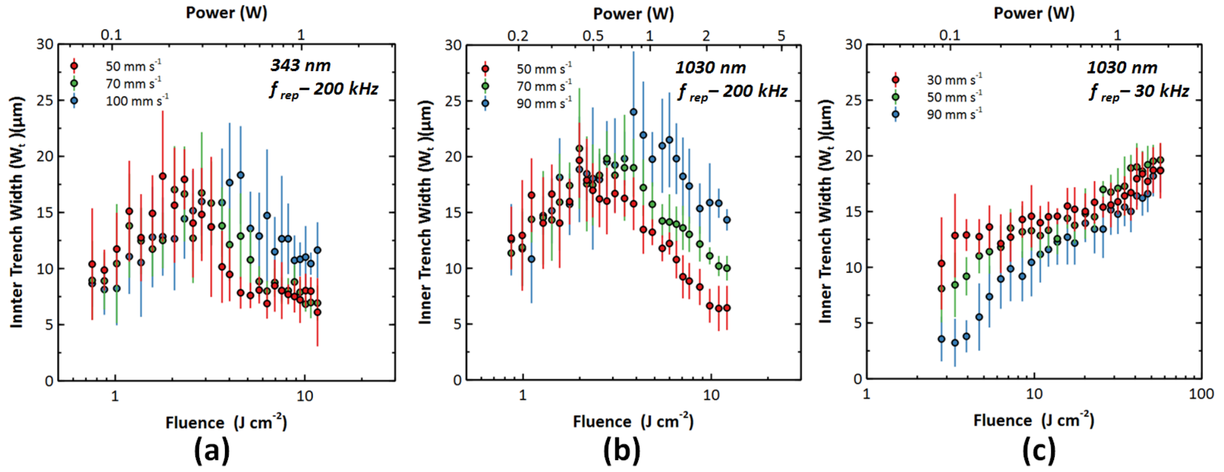


Fig. 4.18: Inner trench width plotted as a function of fluence for 3 cases at different scan speeds. The error bars represent symmetric error of one standard deviation

It should be noted that 343 nm source has a smaller beam waist of ~5.8 μm as compared to 1030 nm with beam waist of ~8.2 μm. The width of cut scales directly with the beam waist and hence a bigger width of cut can be expected for 1030 nm source (Equation 3.13). However, it is also determined additionally by the threshold fluence which is smaller for 343 nm source. On comparing the two high repetition rate cases, an overall observation indicates higher width of cut is indeed seen for 1030 nm source. However, the scatter of data is not negligible for these data points, which makes it difficult to compare the two cases. Nevertheless, it is seen for both cases that in the soft ablation regime, the width of trench increases.

An additional important observation is made for the two cases. With the onset of strong ablation regime, the trench width starts to decrease. This is most likely due to accumulation of recast material on the sidewalls of the trench as seen in [7] (Fig. 8b). The reduction in trench width makes the milling process less efficient. This is because for each subsequent line traversed in the milling trajectory, a considerable amount of energy is deposited on the recast layer which does not result in any net ablation from the substrate.

For the last case of 1030 source with repetition rate of 30 kHz, the trench widths in the gentle ablation regime is more sensitive to the fluence especially at high galvo speeds. A very low width of cut of ~3 μm can be obtained at galvo speed of 90 mm s^{-1}. The scatter of data is smaller as compared to the other two cases indicating the presence of a more uniform recast layer. The narrowing of trench width in the strong ablation regime is not seen for this case. This would be useful in preventing loss of laser energy to recast layer and improving milling efficiency.

Thus, trenches have been studied for 3 different cases and following important observations have been made. At same conditions of repetition rate and speed, 1030 nm source gives a higher material removal rate as compared to 343 nm. However, several disadvantages of using high repetition rate for the two sources have been highlighted. The use of 1030 nm source at low repetition rate gives a lower material removal rate. However, other advantages such as reduction of debries, low profile roughness at high depth, absence of trench narrowing effect make this case suitable for usage in laser milling. However, additional factors come into play during laser milling. This is discussed in the next section.

4.3 Laser milling

The study of trenches provides reasonable understanding of optimal processing conditions for laser milling. However, additional factors have to be considered for laser milling. This section covers the topics of laser milling pitch, determination of depth, analysis of quality of cavity, effect of ablation plume. The value of pitch (line to line spacing) that is going to be used for milling is discussed first.

4.3.1 Choice of milling pitch

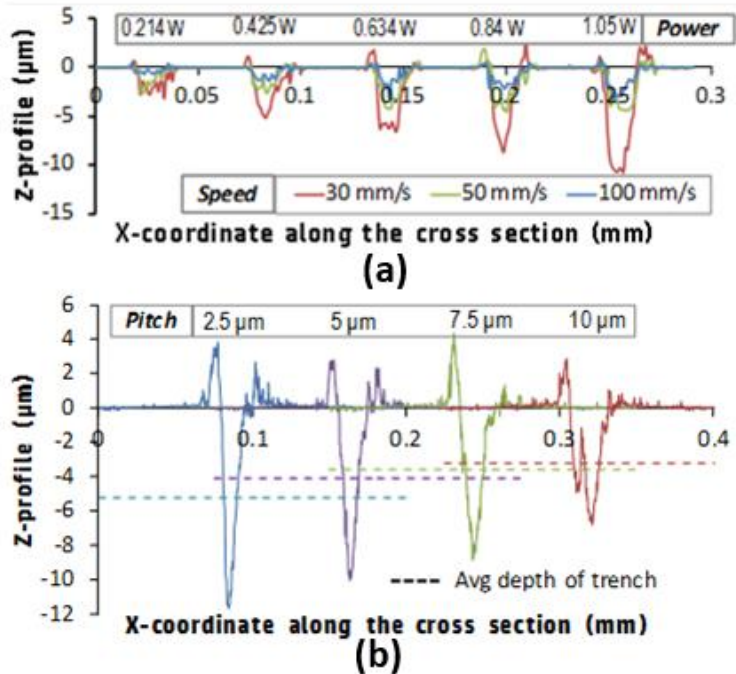


Fig. 4.19: (a) Cross-sectional profiles of single trench as a function of varying laser power at different scan speeds on 1030 nm source (b) Cross-sectional profiles for two lines scribed parallel to each other with varying pitch. Laser parameters: (i) Power: 0.26 W (ii) Scan speed: 20 mm s^{-1} (iii) Fluence: 6.4 J cm^{-2}

Pitch is an important factor to consider during laser milling. The primary consequence of changing the pitch is the change of the depth per scan of laser of a milled area. For scribing a single trench, both the width and depth of trench increases with increasing laser power as shown in Fig. 4.19a. In milling, lines are scanned with a certain pitch between one line to the next. The milling process can be viewed as starting from a single trench and widening the trench by moving the laser beam line by line parallel to the first trench which effectively widens and deepens it.

The effect of pitch can be explained by studying the scribing of two parallel lines. For the first line, by beam traversal over a certain length, a trench is scribed over this length. For the second parallel line, the beam traverses the same length of the first trench but with an offset in the transverse direction which is equal to the pitch. The pitch is set to a value smaller than the width of the trench. Hence, during scribing of the second parallel line, a part of the energy is deposited in the first trench and a part outside the first trench. At the end of the second traversal, the resulting geometry also resembles a trench but with an increased width and depth. This is the working principle of laser milling. The same explanation can be extended for laser milling where laser beam traverses multiple lines in a sequential manner and finally results in removal of a 3D volume with the x-y dimensions set by the boundaries of the beam trajectory.

At smaller pitch, a smaller offset results in a larger portion of the beam energy to be deposited in the region of the first ablated trench. This causes further increase of trench depth while only a small part of energy is deposited outside and the width of trench is extended a little. At higher pitches, smaller portion of the energy is deposited in the region of first trench and a larger portion is used to ablate material outside the trench. The net effect is smaller depth and increased width of the trench. Hence, the pitch sets the depth per scan for milling. These observations can be made by looking at profiles of scribing two parallel lines as shown in Fig. 4.19b. At high pitch, it is also seen that two minima appear instead of one. This will increase the overall surface roughness. A detailed study of impact of pitch on roughness of stainless steel is presented in the work of Audouard et al. [17]. The optimal pitch to minimize corrugation is dependent on the fluence used. While this is useful for very sensitive optimization of roughness, we find that roughness is dominated by factors other than the pitch when the removal rate is high. These factors are discussed in section 4.3.3. We keep the pitch fixed at 10 μm for our milling experiments.

4.3.2 Determination of depth as a function of number of scans

During laser milling, a certain depth needs to be removed in a chosen area. In order to remove the required depth, multiple scans of the laser beam over the milling area are needed. It is preferable to remove only a small depth per scan. Higher scan speeds allow the use of higher fluence while keeping the depth per scan low and also reduces process time. Determination of depth as a function of number of scans is necessary to obtain the removal rate. In principle, this can be known by estimating the depth of cavity using the depth of trench for chosen laser parameters and the pitch. However, additional factors come into play during milling process.

Fig. 4.20a shows the cavity depth obtained for same set of laser parameters but different area of milling. It can be seen that for small areas, the milling depth is reduced. It is to be noted that in this case, very low repetition rate is used which yields low removal rates. Hence, even under the best conditions of milling, differences are to be expected. The differences are due to ablation plume and accompanying redeposition which is discussed later. Hence, plotting the cavity depth as a function of number of scans provides the depth per scan for the chosen set of milling parameters.

The depth per scan is necessary for another reason. For every scan of the laser, a certain depth of material is ablated. The beam focus is initially at the surface and after one scan, material is ablated and the beam is exposed onto newly open surface for the following scan. If the beam position is unaltered in the z-direction, the spot size increases for each scan. This can pose problem during laser milling because it changes the conditions of fluence and pulse overlap for each subsequent pass.

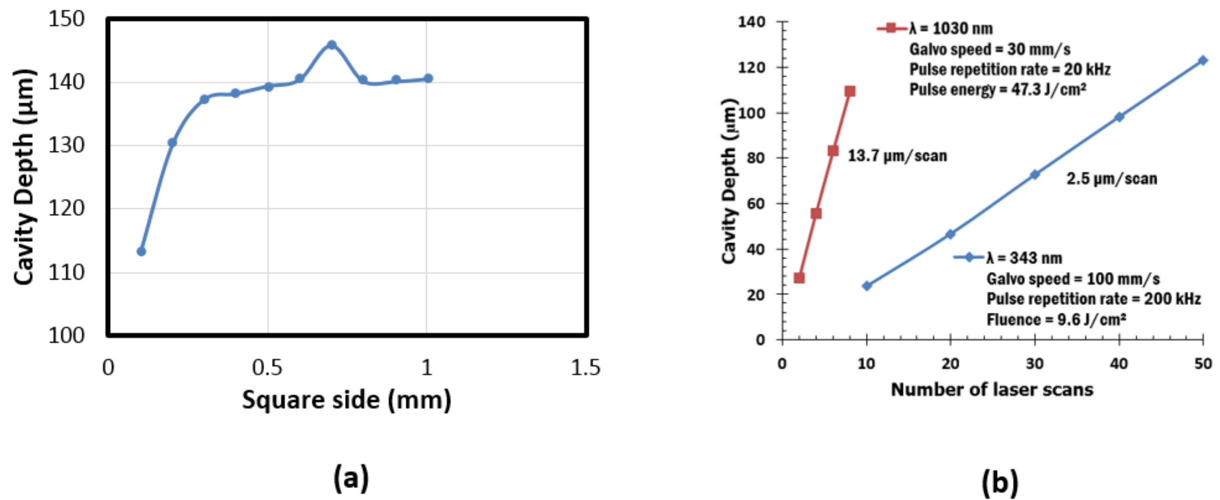


Fig. 4.20: (a) Cavity depth obtained for different squares plotted as a function of square side. Milling parameters (i) Source: 1030 nm (ii) Scan speed : 5 mm s^{-1} (iii) Repetition rate: 2 kHz (iv) Fluence: 56.8 J cm^{-2}(v) Number of passes: 15 (b) Plot of depth as a function of different milling parameters for the two laser sources having average power of 1 W for square of side 1 mm.

However, in Fig 4.20b, we find that despite changing conditions of spot size during each pass, the depth is still linear with respect to number of scans. This is because the spot size does not change appreciably if the distance from the beam waist to the exposed surface is less than the Rayleigh Range (RR) which is determined by the laser optics and wavelength. In our system, the RR values for 343 nm source and 1030 nm source are 244.2 µm and 173.8 µm respectively. If milling depths needs to be higher than RR, then it is preferable to lower the lens each scan by a value equal to depth per scan. This way, same conditions of fluence and pulse overlap can be maintained during each scan.

In similar milling experiments by Zhang et al. , even under tight focusing conditions (numerical aperture of 0.1), there was very little impact of defocusing distance (up to 200 µm) on removal rate [18]. Hence, for deep cavity milling, precise control of focus would not be necessary and lens can be lowered approximately to keep the focus point approximately close to the surface.

4.3.3 Effects of ablation plume and redeposition on laser milling

In laser milling experiments, it is observed that ablation plume and redeposition poses a challenge for quality of laser milling. Several experiments are reported to demonstrate the effect of ablation plume on laser milling using both 343 nm and 1030 nm source. Two effects that are observed because of laser plume are: (i) Generation of particles (ii) Screening of laser radiation

4.3.3.1 Generation of particles

This is an important factor to consider during laser milling and limits the line to line pitch and range of fluences that are usable for milling. As shown in Fig. 4.21, the quantity of generated particles depends on both the pitch and peak fluence. The total energy incident for each cavity is nearly kept the same by balancing the pulse repetition rate and the pulse energy for varying fluence. However, because of difficulties in balancing the two, there are noticeable differences in average power for varying fluences. However, this is not a significant factor in this experiment as it is only a qualitative analysis. In addition, for varying pitch, the number of passes is varied to have a nearly constant total energy condition.

At lower fluences and smaller pitch, the formation of particle aggregates is higher. This is possibly because of higher repetition rate used at lower fluences which promotes build-up of ablation plume. Most of these particles are loosely bound to the substrate and can be removed in a strong air or nitrogen flow. However, it is not desirable to have them as the accumulation of particles interferes with the laser process. There is circulation of generated particles during milling (as observed in the on-axis laser camera) which interferes with the incoming laser radiation. While this does not completely hinder the laser process, the ablation depth is reduced and the uniformity of the milled surface is degraded.

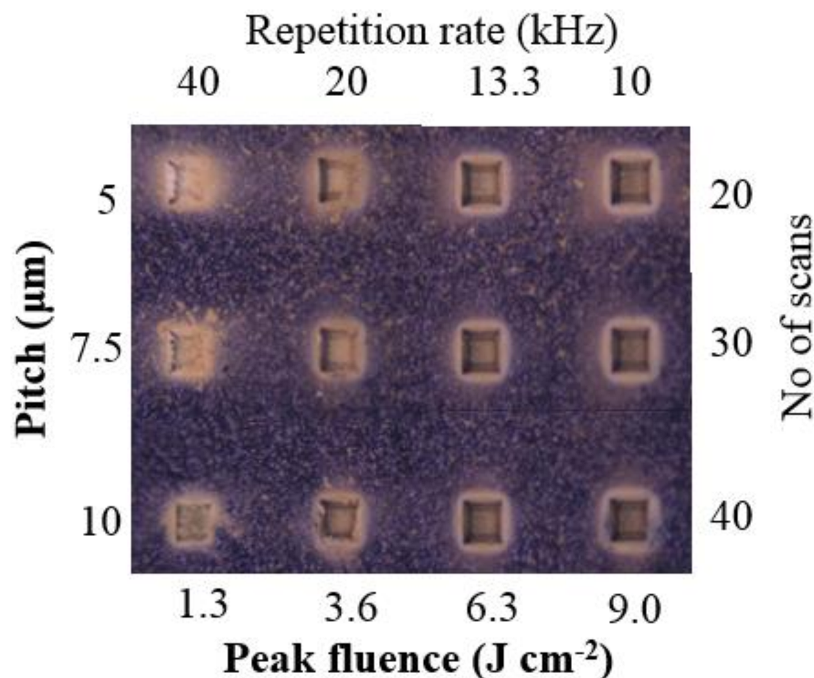


Fig. 4.21: Particle generation and redeposition as a function of laser fluence and pitch. Milling parameters (i) Source: 1030 nm (i) Area: 0.25 x 0.25 mm^2 (ii) Power: Column 1: 0.054 W, Column 2: 0.072 W, Column 3: 0.085 W, Column 4: 0.09 W) (iii) Scan speed: 10 mm s^{-1}

The generation of particles depends also on the aspect ratio of the cavity. Higher aspect ratios promote confinement of ablation plume in the vicinity of the ablated area. This can promote creation and redeposition of particles. This is illustrated in Fig. 4.22. For all cavities, the xy dimensions of the inner 4 squares are the same with a side of 250 µm. In the 3 cases, the boundaries surrounding the 4 smaller squares are varied.

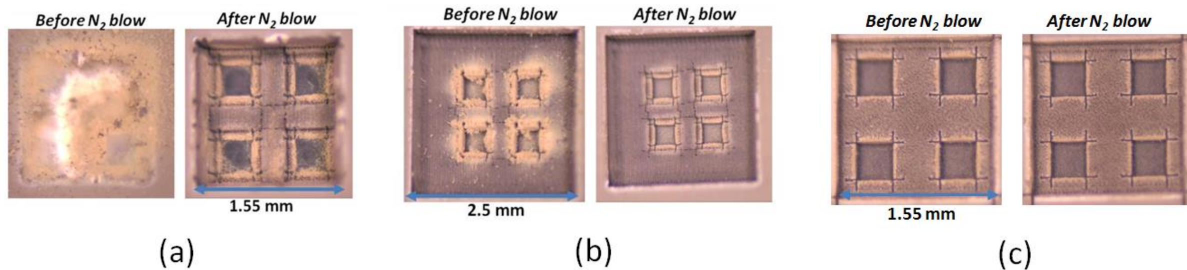


Fig. 4.22: Generation of particles during milling on 1030 nm source: (a) Case 1: Narrow opening, deep cavity (b) Case 2: Wide opening, deep cavity (c) Case 3: Narrow opening, shallow cavity. Large cavity milling parameters: Scan Speed - 50 mm s^{-1}, Fluence – 61.5 J cm^{-2}, No. of scans- 12. Small cavity milling parameters: Scan speed – 10 mm s^{-1}, Fluence – 7.1 J cm^{-2}, No. of scans – 80

In case 1, the 4 squares are milled inside a cavity of depth 450 µm and a narrow square opening of side 1.55 mm. The particle aggregates generated from the ablation plume completely covers the region of milling at the end of process. After process is complete, by placing the milled cavity in a strong flow of N_2 gas, the particles aggregates are removed. It is found that the 4 squares are formed despite the presence of particle aggregates. The average depth of the 4 cavities is found to be 42.3 µm. In case 2, where the depth of cavity is same as case 1, by increasing the size of the opening to 2.5 mm, it can be seen that the formation of particle aggregates can be greatly reduced. The average depth of the four square cavities is 73.1 µm. Thus, by allowing ablation plume to better diffuse away from the region of milling, its negative effect can be reduced.

Finally for our intended application of creating membranes of SOI dies, the thickness of the die can be reduced by grinding. If the thickness of the SOI wafer is reduced from 750 µm to 250 µm, the thickness of material to be removed using laser processing is reduced by 2/3rd of the original thickness. This reduced thickness allows better quality of laser processing as seen in case 3 of Fig. 4.22. The same 4 squares are milled inside a cavity of depth 100 µm with the same square opening of 1.55 mm as in case 1. The average depth of the 4 cavities in this case is 86.1 µm and the redeposition of particle aggregates is negligible in this case. All these experiments reveal the importance of removing the ablation plume from the milling region in order to achieve better milling.

4.3.3.2 Screening of laser radiation

To remove a certain depth of material using laser milling, multiple scans are required over the milling area. Each scan results in ablation and removal of material in the form of a dense plume which can contain different phases of the ablated material. The plume can undergo phase transformations and recondensation. It has already been shown that ejected matter can be redeposited in the form of loosely bound particles. The ablation plume can additionally shield the laser beam intensity in the milling area. In some applications, the ablation plume has been reported to last even up to 1 s [19].

We confirm the presence of ablation plasma in the milling area with the help of air flow. Laser milling is performed in the presence of a nitrogen gas flow as shown in Fig. 4.23. It can be seen that without air flow, the milling process is clean and no particle aggregates are seen around the milling area. When a nitrogen gas flow is introduced, there is accumulation of particles around the bottom part of the milling area. These particles are possibly formed as the ablation plume is carried away from the milling region and cooled down in the presence of nitrogen flow. A depth profile of these particle aggregates reveals heights up to 150 µm. This suggests that the quantity of ablation plume in the milling area is not negligible. Also, the formation of particles in presence of air flow suggests that particle formation is due to cooling down of the ablation plume. Despite the presence of an air flow in the second case, the average depth remains the same for the two cases. The particles are weakly bound to the substrate and can be easily removed in a strong flow of air or ultrasonic cleaning.

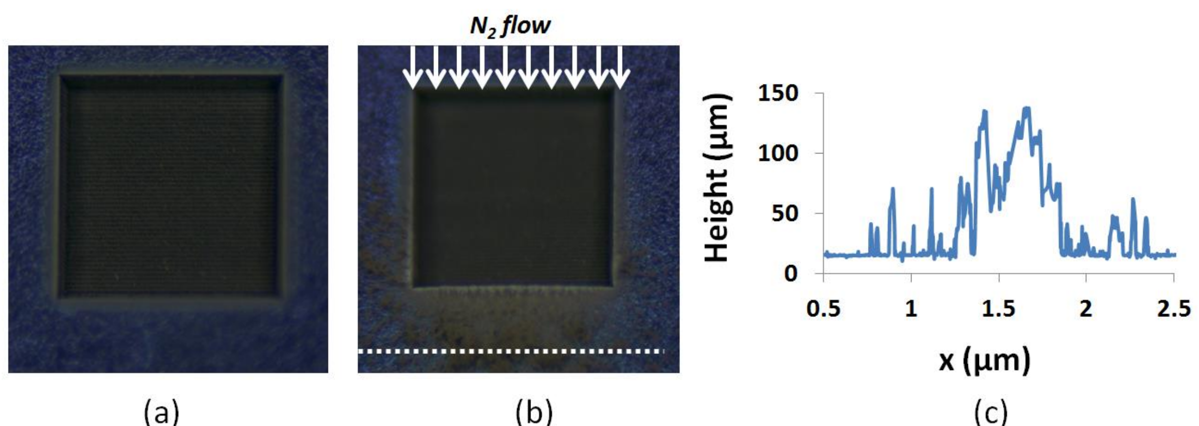


Fig. 4.23: Demonstration of presence of ablation plume in the region of ablation (a) Milling process without air flow (b) Milling process with air flow (c) Depth profiling of particle aggregates along the white dotted line. Milling parameters (i) Scan speed: 50 mm s^{-1} (ii) Fluence: 9.8 J cm^{-2} (iii) Pulse repetition rate: 100 kHz (iv) No. of scans: 50

In our experiments, we visually observe the presence of a thick plume during laser processing which diffuses away from the sample. In order to understand the timescale of diffusion of ablation plume, one scan of laser milling is performed over a square of area 1 mm^2. Once the scan is complete, it generates an ablation plume in the vicinity of the milling area. The study of dynamics of plume species by a single pulse ablation of silicon shows very high particle velocities of ions and neutrals order $10^6 – 10^7$ cm s^{-1} [20]. However, it was also shown in the study that there is a slow component containing nanoparticles which have much slower velocities of the order 10^3 cm s^{-1}. If the plume is still present over the milling area at the beginning of the next scan, shielding effect results in a decrease of laser energy deposited in the milling area and consequently a lower depth of ablation.

When a delay is introduced between milling scans, this allows for generated plume to diffuse away from the vicinity of the milling area. If the delay is long enough, the ablation plume would have diffused completely away from ablation area and no difference is milling would be noticed for longer delays. Thus, the study of milling process by introducing a variable delay between two consecutive scans reveals the timescale on which the ablation plume affects the laser process.

Table 4.2: Measured quantities at different values of delay for 1-way milling

Delay (s)	Average Depth (µm)	Depth - Top half (µm)	Depth - Top 10% (µm)	Depth - Bottom half (µm)	Depth - Bottom 10% (µm)
0	145.4	144.9	141.2	146.2	146.3
1	159.3	154.5	148.2	164.5	173.1
2	158.4	153.6	148.7	163.3	169.9
3	156.8	152.8	146.8	161.2	165.6

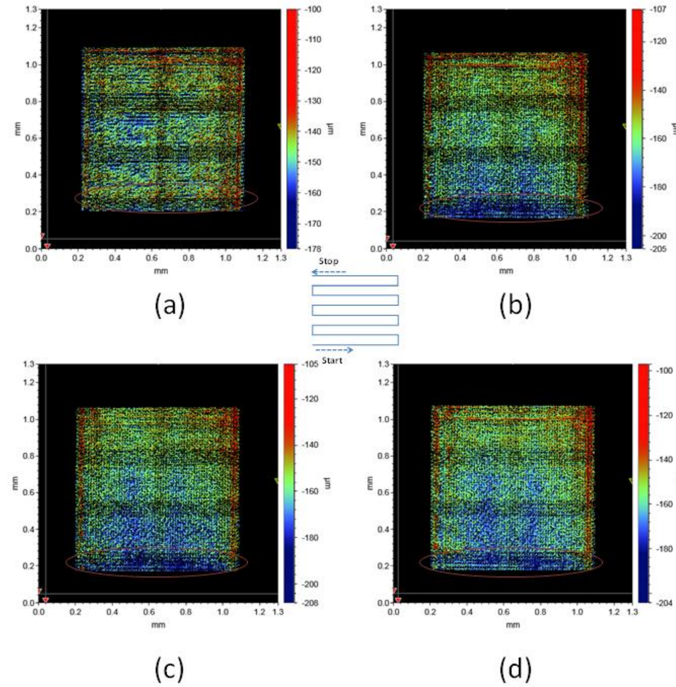


Fig. 4.24: Profilometer scans of milling with beam trajectory for each scan shown in the middle. Milling parameters (i) Scan speed: 100 mm s^{-1} (ii) Fluence: 9.9 J cm^{-2} (iii) Pulse repetition rate: 200 kHz (iv) No. of scans: 50 (v) Delay between successive scans: (a) 0 s (b) 1 s (c) 2 s (d) 3 s

The depth profiles obtained for different scan to scan delays are shown in Fig. 4.24. Different values of depths are deduced from these scans and are tabulated in Table 4.2. An inherent laser system delay of 45 ms is to be added to each of the delay values. The average depth is taken over a reduced square window of side 0.9 mm centred on the cavity of side 1 mm. The depths of top half and top 10% are taken from windowing rectangles of dimensions 0.45 mm x 0.9 mm and 0.09 mm x 0.9 mm with the longer edge corresponding to the top side of the reduced square window. The same holds true for bottom side quantities where the longer edge corresponds to bottom side of the reduced square window.

The average depth of cavity obtained for no delay between scans is lower by 11 – 14 µm as compared with cavities with delay. The differences between cavities

having some scan to scan delay are negligible. The scan trajectory for each scan is shown in the centre of Fig. 4.24. As the laser beam moves in a serpentine manner from bottom to the top, the laser plume builds up slowly and by the time the top end of the cavity is reached, the increased laser plume results in reduced energy absorption at the top end of the cavity. Hence, the top side depths are smaller than bottom side depths. If there is no delay between scans, by the time the next scan starts, the plume is still not fully cleared from the vicinity of the milling area. Hence, at the beginning of the following scan, the plume screening effects are still seen.

When a delay is introduced, it allows plume to diffuse away from the milling area. When the next scan is started in a plume free environment, the obtained depths will be higher. The interesting region is the bottom 10% which represents the starting part of each laser scan. Here, the effects of plume are least seen and hence the ablation depth is noticeably higher in this region as compared to the average depth of the cavity. When there is no delay, because of plume screening, depth at bottom 10% and average depth remains about the same.

To confirm this hypothesis further, the delay experiment is extended but milling trajectory is two-way instead of one way as illustrated in Fig. 4.25. Here, only two cases are discussed namely zero delay and delay of 1 s. The same laser system delay of 45 ms is to be considered here as well.

Table 4.3: Measured quantities at different values of delay for 2-way milling

Delay (s)	Average Depth (μm)	Depth - Top half (μm)	Depth - Top 10% (μm)	Depth - Bottom half (μm)	Depth - Bottom 10% (μm)
0	147.2	146.2	142.9	148.1	146.9
1	167.2	166.2	168.1	168.2	170.4

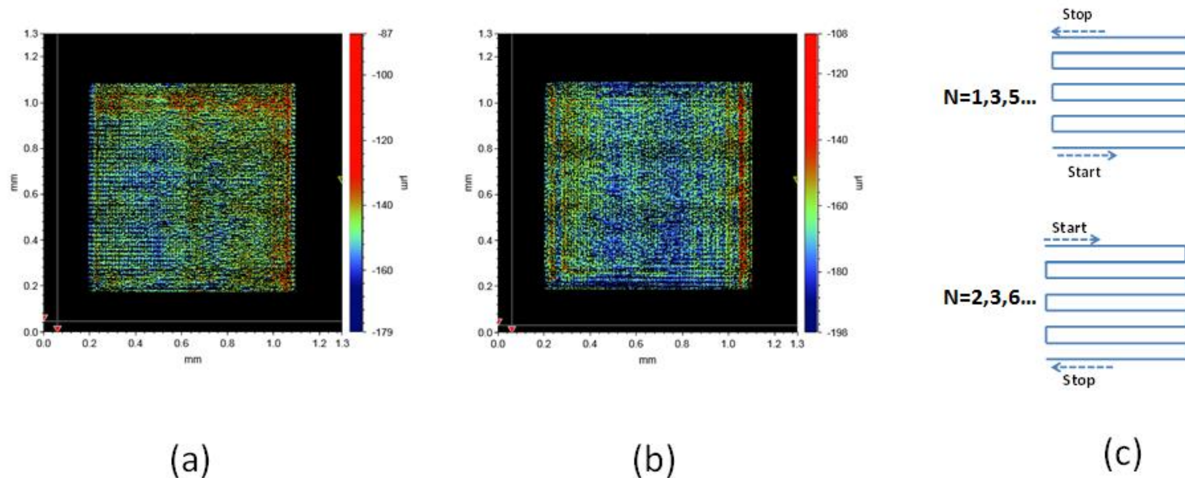


Fig. 4.25: Profilometer scans of 2-way milling (i) Scan speed: 100 mm s^{-1} (ii) Fluence: 9.9 J cm^{-2} (iii) Pulse repetition rate: 200 kHz (iv) No of scans: 50 (v) Delay between successive scans: (a) 0 s (b) 1 s (c) Two-way milling trajectory

Because of the symmetry of the trajectory, it can be seen that the top and bottom depth quantities are comparable with small differences. Without delay, the depth quantities for two-way case are only marginally better as compared to one way case as seen in Table 4.3. With a scan to scan delay of 1 s, both the top 10% and bottom 10% depths are similar and are comparable to bottom 10% depth of one-way case. Additionally, in two-way case the average depth is increased by a few microns and variation between different depth quantities is minimized. Thus, by using 2-way or possibly more directions of milling and appropriate use of delays, the milling quality can be improved by alleviation of ablation plume related effects.

4.3.3 Analysis of milling quality

When the profilometer scan is observed for any milled cavity, the surface is not perfectly flat and there are local variations in depths at different regions in the cavity. One way to characterize this is the calculation of average and RMS roughness of the surface. But, these metrics will not provide any information about the numerical values of the maximum (peak) and minimum (valley) values of the depths. These values will be necessary for the subsequent development of substrate removal process that is used in this work. While this value can be known directly from the profilometer scan, it is highly susceptible to noise in the profilometer scan. There is an indirect method to determine the peak to valley distance. It can be approximately known by multiplying the RMS roughness by 6. This is because it is empirically observed that the distribution of depths can be

well approximated by a Gaussian distribution and the standard deviation of this distribution is equal to the RMS roughness value.

In general, the milled cavities can have very different profiles showing different variation of local depths depending on the processing parameters as seen in Fig. 4.26. By plotting the distribution of depths, the global variation of depths can be understood. As seen in Fig. 4.26c, the depth distribution is plotted for the two cavities which can be well approximated by Gaussian.

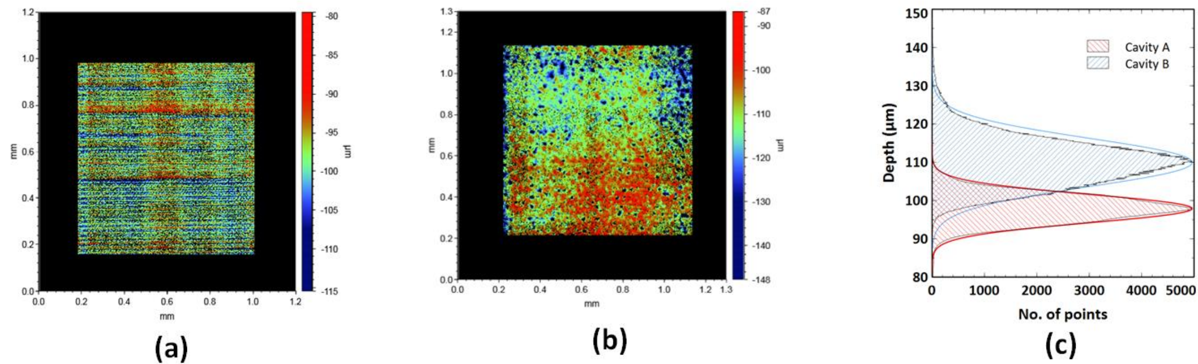


Fig. 4.26: Milling profiles taken for two different square cavities of side 1 mm milled on two different sources and different laser parameters (a) (i) Source: 343 nm (ii) Scan speed : 100 mm s^{-1} (iii) Repetition rate: 200 kHz (iv) Fluence: 9.6 J cm^{-2} (v) Number of scans: 40 (b) (i) Source: 1030 nm (ii) Scan speed : 30 mm s^{-1} (iii) Repetition rate: 20 kHz (iv) Fluence: 51.1 J cm^{-2} (v) Number of scans: 8 (c) Plot of depth distributions along with Gaussian fit

For nearly the same average depth of the two cavities in Fig. 4.26a and Fig. 4.26b (98.2 µm and 109.2 µm respectively), the roughness values vary considerably. The average and RMS roughness of 2.7 and 3.3 µm is obtained for 343 nm source. For the 1030 nm source, the average roughness is 4.6 and RMS roughness is 5.9 µm. It is to be noted that for 343 nm source, low removal rate is used and for 1030 nm source, high removal rate is used. Hence, the roughness is dependent also on the removal rate.

For 1030 nm case, at low repetition rate, the value of average roughness for the cavity is 4.6 µm. This is much larger than the value of roughness expected from average roughness measured from trench axial profile (< 1 µm). In order to study systematically the effect of removal rate on roughness, milling of cavities is performed at repetition rate of 30 kHz using 1030 nm source and scan speed of 20 mm s^{-1}. The fluence is varied in these experiments to vary the material removal rate. The SEM images of the bottom of the cavities for one scan of the laser beam are shown in Fig. 4.27. It can be seen that overall morphology looks good with presence of few microholes notably for fluence of 4.7 J cm^{-2}.

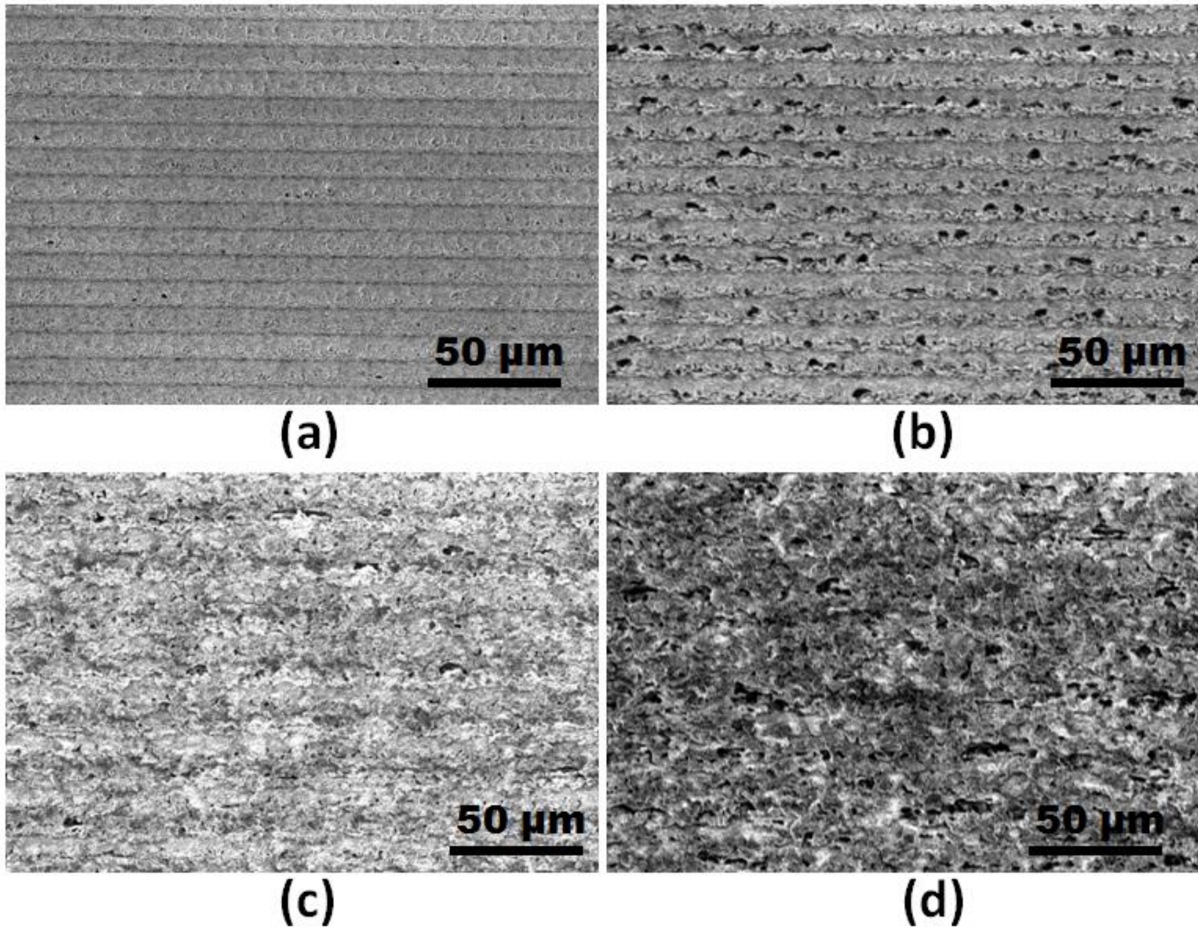


Fig. 4.27: Morphology of the bottom of the trench after cleaning in ultrasonic bath. Milling using 1030 nm source with parameters: Single scan; Scan speed - 20 mm s^{-1}; Repetition rate - 30 kHz; Pitch - 10 µm; Fluence (a) 2.8 J cm^{-2} (b) 4.7 J cm^{-2} (c) 12.7 J cm^{-2} (d) 17.4 J cm^{-2}

Table 4.4: Calculated quantities for milled cavities (single scan) on 1030 nm source at repetition rate of 30 kHz and scan speed of 20 mm s^{-1}

	Fluence (J cm^{-2})			
Parameter	**2.8**	**4.7**	**12.1**	**17.4**
Before ultrasonic cleaning				
Average depth of cavity (µm)	1.3	1.5	2.3	4.1
After ultrasonic cleaning (in IPA for 5 minutes)				
Average depth of cavity (µm)	1.9	2.9	6	8.9
Average depth of trench profile (µm)	1.9	4.2	6.9	10.4
Average roughness of trench profile (µm)	0.2	0.4	0.4	0.6
Average roughness of cavity (µm)	0.2	0.5	0.8	1.2
RMS roughness of cavity (µm)	0.2	0.7	1	1.5
RMS roughness of trench profile (µm)	0.2	0.5	0.6	0.7

Several parameters are computed for the cavity which is summarized in Table 4.4. It can be seen that depth of cavity after ultrasonic cleaning increases in all

cases. With increase in fluence, the difference is larger. This means that a significant quantity of material at the bottom of the cavity is loosely bound suggesting a redeposition of ablated material. At lowest fluence of 2.8 J cm^{-2}, the parameters for the cavity and trench are excellently correlated. With increasing fluence, the difference in roughness parameters between cavity and trench profile becomes larger. This increase in roughness can be attributed to presence of ablation plume and redeposited material in the vicinity of the ablation area which can interact with the laser beam.

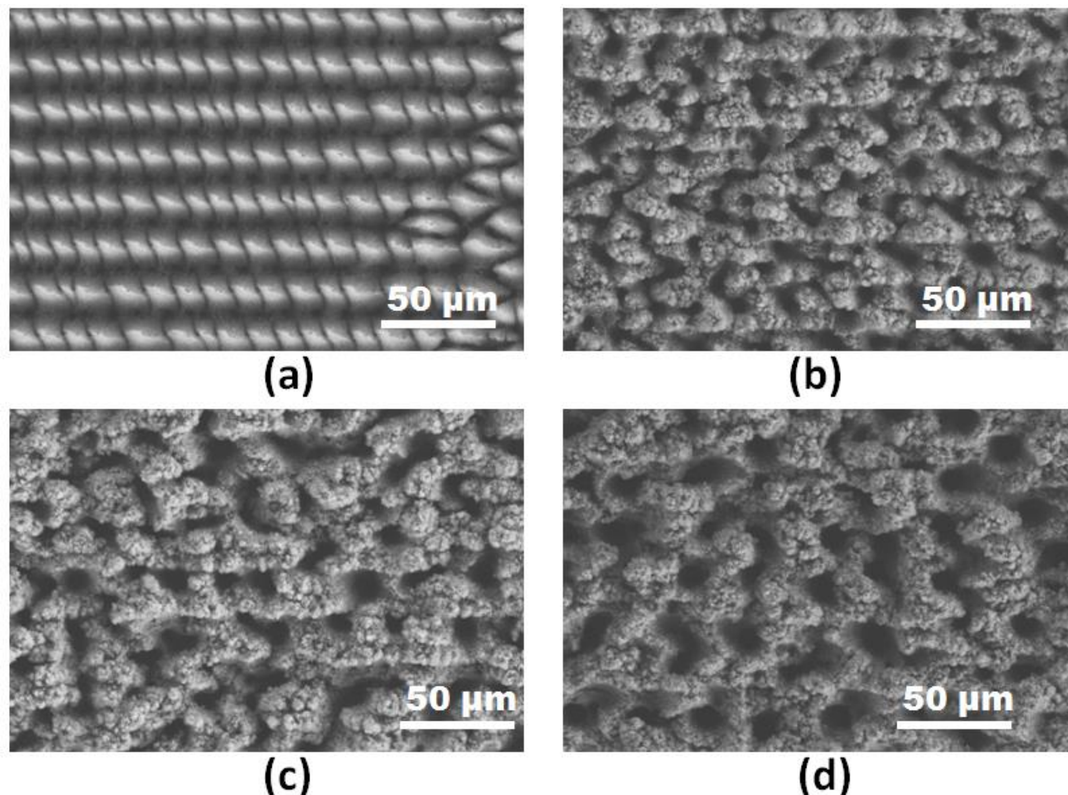


Fig. 4.28: Morphology of the bottom of the trench after cleaning in ultrasonic bath. Multipass milling (Depth ~ 100 µm) using 1030 nm source with parameters: Scan speed - 20 mm s^{-1}; Repetition rate - 30 kHz; Pitch - 10 µm Fluence/ Number of scans (a) 2.8 J cm^{-2}/65 (b) 4.7 J cm^{-2}/28 (c) 12.7 J cm^{-2}/17 (d) 17.4 J cm^{-2}/10

Having studied the milling for a single pass, the number of passes is increased to mill a cavity with a depth ~100 µm. In this experiment, the area of cavity is also varied from square of side length 100 µm to 1 mm. The morphology of the bottom of the cavity is shown in Fig. 4.28. The morphology of the cavity is drastically different from that observed for single pass. A large number of microholes are seen at the bottom of the cavity. The number of holes decreases while the size of the holes increases at higher fluence. At the lowest fluence of 2.8 J cm^{-2}, a regular array of microstructures approximately in the form of pyramids is seen. At higher

fluences, such regularly ordered microstructures are not seen. Additionally, it is observed that for the low fluence of 2.8 J cm^{-2}, a lot of debris are formed and the smaller cavities were almost fully covered in particle debris as it was seen in Fig. 4.21. For higher fluences, quantity of debris was much lower.

The average depth is plotted in Fig. 4.29a for all cavities before and after ultrasonic cleaning to quantify redeposited material. It can be seen that in all cases redeposited matter exists at the bottom of the cavity for all cases. The average difference in depths before and after ultrasonic cleaning are 3, 1.7 and 1.5 µm for fluences of 4.7, 12.7, and 17.4 J cm^{-2} respectively. For fluence of 2.8 J cm^{-2}, the data before ultrasonic cleaning is not taken because many cavities are covered in particle debris. In Fig. 4.29b, the average depths are found in a similar way but the number of passes is doubled to have nearly twice the depth. The average difference in depths before and after ultrasonic cleaning are 5.2 and 3.7 µm for fluences of 12.7 and 17.4 J cm^{-2} respectively. Thus, when the cavity becomes deeper, the quantity of redeposited material increases. An efficiency factor is determined using the following formula:

$$\eta_f = \frac{d_{cav}}{N_p d_t} \tag{4.4}$$

$d_{cav} - Average\ depth\ of\ cavity$

$N_p - Number\ of\ scans\ of\ laser$

$d_t - Average\ depth\ obtained\ from\ axial\ trench\ profile$

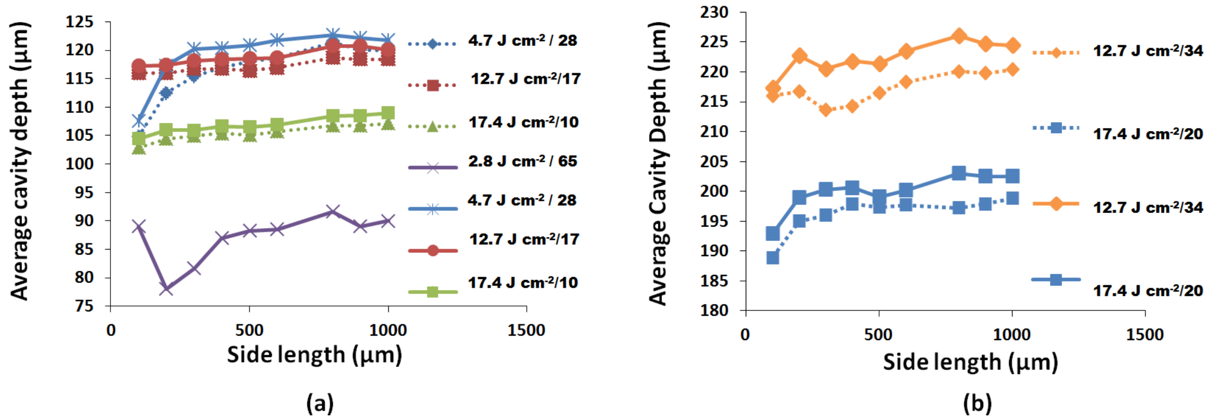


Fig. 4.29: Average depth of cavities milled on the 1030 nm source at repetition rate of 30 kHz for different fluences/ number of passes indicated in the legend. Dotted and solid lines represent data before and after ultrasonic cleaning respectively

The different parameters calculated are summarized in Table 4.5. The efficiency factors for 1 mm^2 cavities milled at 12.7 and 17.4 J cm^{-2} respectively are found to

be 0.906 and 0.917 respectively when cavity depth is ~100 µm. On doubling the number of passes, the efficiency factor reduces to 0.846 and 0.852 respectively. Thus, it can be said that with increasing depth of the cavities, the ablation plume and redeposited material plays an important role in determining depth and roughness parameters. With the reduction in efficiency factor, the surface roughness also increases. This is clearly noticed for the fluence of 2.8 J cm^{-2}, where the efficiency factor is much smaller than other cases. When the particle debris are small the quality of side walls is good as seen in Fig. 4.30. In our application, the sidewall quality is equally important as the roughness parameters.

Table 4.5: Calculated quantities for milled cavities (multiple scans) on 1030 nm source at repetition rate of 30 kHz and scan speed of 20 mm s^{-1}

Fluence (J cm^{-2})	No of scans	Depth (µm)	Average roughness (µm)	RMS roughness (µm)	Efficiency factor
2.8	65	90	9.9	11.3	0.682
4.7	28	121.8	3.9	5.1	0.939
12.7	17	120.1	4.3	5.5	0.906
	34	224.5	5.1	6.4	0.846
17.4	10	108.9	4.8	6.1	0.917
	20	202.5	5.9	7.3	0.852

Finally, we report a way to reduce the number of microholes in the cavity. It is found that by maintaining the focus of the beam slightly inside the material, the number of holes is reduced as seen in Fig. 4.31. The other parameters such as average depth and roughness do not change appreciably.

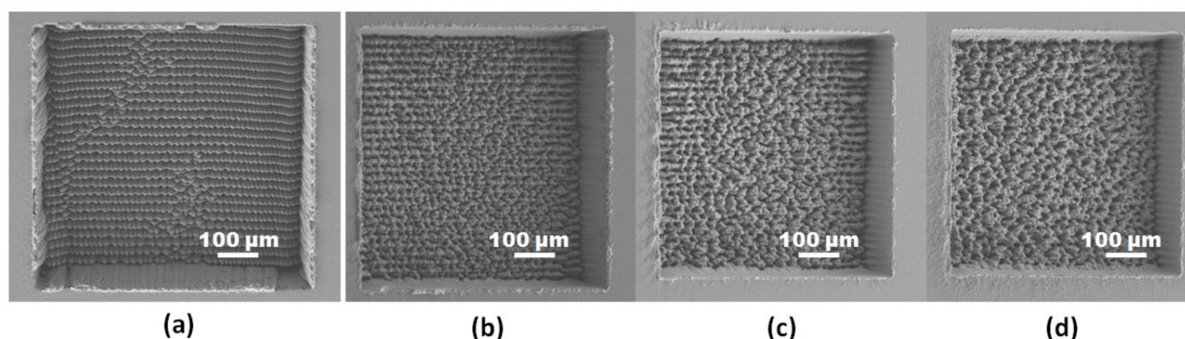


Fig. 4.30: Cavity side walls after cleaning in ultrasonic bath. Milling parameters used are scan speed of 20 mm s^{-1} and repetition rate of 30 kHz with fluence/ number of scans (a) 2.8 J cm^{-2}/65 (b) 4.7 J cm^{-2}/28 (c) 12.7 J cm^{-2}/17 (d) 17.4 J cm^{-2}/10

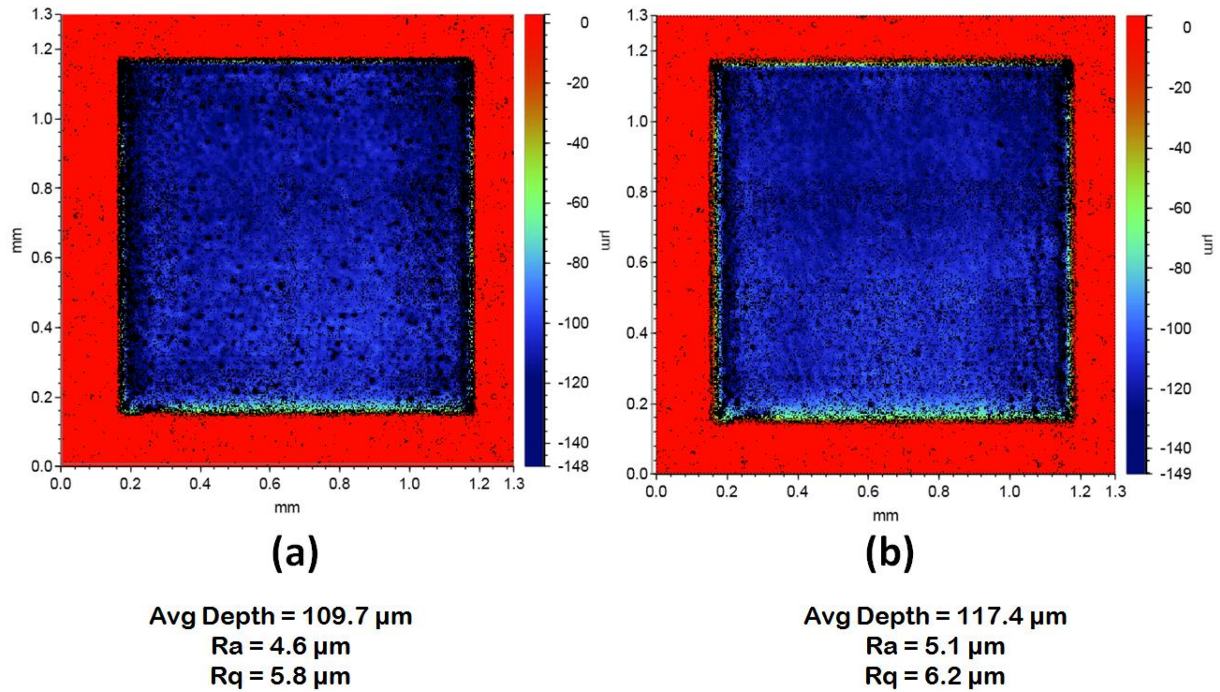


Fig. 4.31: Reduction of holes by changing the focus for each pass. Milling parameters: (i) Speed: 30 mm s^{-1} (ii) Repetition rate: 20 kHz (iii) Fluence: 51.1 J cm^{-2} (iv) No. of scans: 8 (iv) Focus change after each scan (a) No focus change (b) 45 µm into the sample

4.4 Femtosecond Laser Assisted Micromachining and Etch (FLAME) process

Laser milling has been described in detail in previous section. This section describes the application of laser milling for creating suspended membranes in SOI. The Femtosecond Laser Assisted Micromachining Etch (FLAME) process is a simple 4 step process for removal of handler silicon on SOI substrate as shown in Fig. 4.32. The first step is the lamination of etch protect layer which is Ajinomoto GX-T31 dry film. Following this step, laser milling is performed in the desired area until a thin layer of silicon remains. This small thickness of silicon is etched in XeF_2 gas which etches Silicon selectively over SiO_2. Finally, the etch protect layer is stripped by dissolving the dry film in acetone.

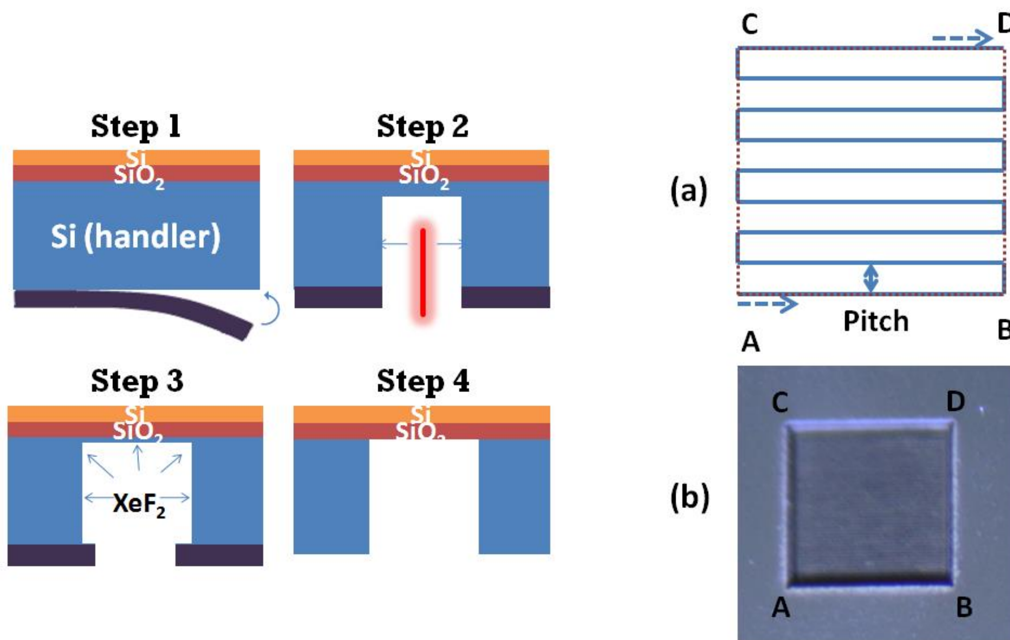


Fig. 4.32:FLAME process steps (1) Lamination of Etch protect layer (2) Laser milling under SOI circuit (3) XeF2 dry etching (4) Stripping of etch protect layer. (a) Hatching of area ABCD to define trajectory of laser beam (b) View of cavity obtained in area ABCD

An important additional step not described in Fig 4.32 is the scribing of alignment lines on the front side which are visible on the back side. Lines are scribed over rectangular features close to the milling area as shown in Fig 4.33. These serve as reference points on the backside. Once lines are scribed on the front side, the offset between the desired and actual position of the x and y axis is measured using a microscope as seen in Fig. 4.33. This can be used to further correct minor random offsets in the laser system to further improve alignment. The line is scribed with surface out of focus to increase the spot size and hence the width of line so that the visibility is better on the backside.

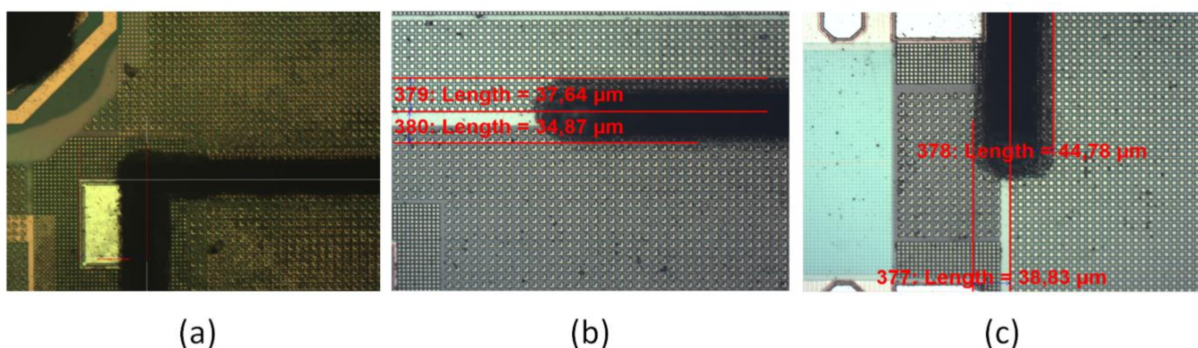


Fig. 4.33: Alignment methodology for creating reference axes on backside (a) Origin (b) x-axis (c) y-axis

Finally, to remove the remaining silicon after laser milling, XeF_2 dry etching is used which etches only silicon. The etching is performed in pulses where the gas

is introduced into a chamber and pumped out after a defined etching time per cycle. The process is repeated over several cycles to remove the desired thickness. For our application, anisotropic smooth etching is preferred. But, XeF_2 is an isotropic rough etching method as seen in Fig 4.34.

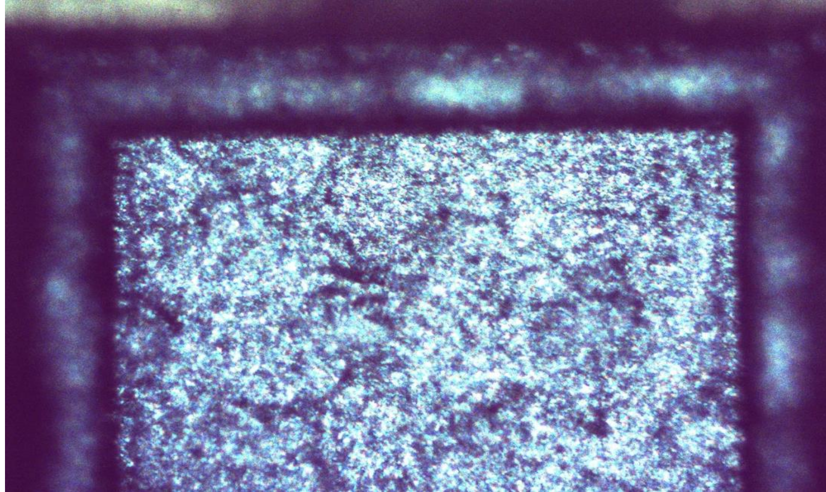


Fig. 4.34: Morphology of cavity after 25 cycles of etching in XeF_2

It is nevertheless chosen over Bosch process because of excellent selectivity of etch of $Si:SiO_2$ of 1000:1. The etch rate mainly depends on the pressure of XeF_2 gas used during etching. For most of our experiments, we use the lowest possible stable pressure in the chamber which is 2 Torr and cycle time of 10 s. For further understanding of XeF_2 etching process and its application to SOI technology, the reader is referred to [21]–[25].

Concluding Remarks

This chapter has described the set of experiments necessary for the development of the FLAME process. The different process parameters have been carefully studied by scribing of simple trenches on silicon. The presence of two ablation regimes has been established. Different measures of quality have been discussed highlighting the parameters which are suitable for milling. Laser milling has also been described and additional factors that impact laser milling have been highlighted. The different effects of ablation plume and redeposition during laser processing also have been highlighted. The FLAME process steps have been elaborated which enables creation of membranes of SOI circuits. This will be the subject matter of the next chapter where the developed process will be applied on different circuits already described in Chapter 3.

References

[1] J. M. Liu, “Simple technique for measurements of pulsed Gaussian-beam spot sizes,” *Opt. Lett.*, vol. 7, no. 5, p. 196, May 1982.

[2] A. Sikora, D. Grojo, and M. Sentis, “Wavelength scaling of silicon laser ablation in picosecond regime,” *J. Appl. Phys.*, vol. 122, no. 4, p. 045702, Jul. 2017.

[3] D. Breitling, A. Ruf, and F. Dausinger, “Fundamental aspects in machining of metals with short and ultrashort laser pulses,” presented at the Lasers and Applications in Science and Engineering, San Jose, Ca, 2004, p. 49.

[4] J. Bonse, S. Baudach, J. Krüger, W. Kautek, and M. Lenzner, “Femtosecond laser ablation of silicon–modification thresholds and morphology,” *Appl. Phys. A*, vol. 74, no. 1, pp. 19–25, Jan. 2002.

[5] Z. Sun, M. Lenzner, and W. Rudolph, “Generic incubation law for laser damage and ablation thresholds,” *J. Appl. Phys.*, vol. 117, no. 7, p. 073102, Feb. 2015.

[6] A. Žemaitis, M. Gaidys, M. Brikas, P. Gečys, G. Račiukaitis, and M. Gedvilas, “Advanced laser scanning for highly-efficient ablation and ultrafast surface structuring: experiment and model,” *Sci. Rep.*, vol. 8, no. 1, Dec. 2018.

[7] T. H. R. Crawford, A. Borowiec, and H. K. Haugen, “Femtosecond laser micromachining of grooves in silicon with 800 nm pulses,” *Appl. Phys. A*, vol. 80, no. 8, pp. 1717–1724, May 2005.

[8] A. Borowiec and H. K. Haugen, “Femtosecond laser micromachining of grooves in indium phosphide,” *Appl. Phys. A*, vol. 79, no. 3, pp. 521–529, Aug. 2004.

[9] M. Domke, B. Egle, G. Piredda, G. Fasching, M. Bodea, and E. Schwarz, “Controlling depth and distance of the hole formations at the bottom of laser-scribed trenches in silicon using fs-pulses,” presented at the SPIE LASE, San Francisco, California, United States, 2015, p. 93501J.

[10] B. Jaeggi, B. Neuenschwander, S. Remund, and T. Kramer, “Influence of the pulse duration and the experimental approach onto the specific removal rate for ultra-short pulses,” presented at the SPIE LASE, San Francisco, California, United States, 2017, p. 100910J.

[11] S. Nolte *et al.*, “Ablation of metals by ultrashort laser pulses,” *JOSA B*, vol. 14, no. 10, pp. 2716–2722, Oct. 1997.

[12] K. Furusawa, K. Takahashi, H. Kumagai, K. Midorikawa, and M. Obara, “Ablation characteristics of Au, Ag, and Cu metals using a femtosecond Ti: sapphire laser,” *Appl. Phys. A*, vol. 69, no. 1, pp. S359–S366, 1999.

[13] D. Ashkenasi, A. Rosenfeld, H. Varel, M. Wähmer, and E. E. B. Campbell, “Laser processing of sapphire with picosecond and sub-picosecond pulses,” *Appl. Surf. Sci.*, vol. 120, no. 1–2, pp. 65–80, 1997.

[14] Y. Izawa, Y. Setuhara, M. Hashida, M. Fujita, and Y. Izawa, “Ablation and Amorphization of Crystalline Si by Femtosecond and Picosecond Laser Irradiation,” *Jpn. J. Appl. Phys.*, vol. 45, no. 7, pp. 5791–5794, Jul. 2006.

[15] M. Hashida, Y. Miyasaka, Y. Ikuta, S. Tokita, and S. Sakabe, “Crystal structures on a copper thin film with a surface of periodic self-organized nanostructures induced by femtosecond laser pulses,” *Phys. Rev. B*, vol. 83, no. 23, p. 235413, 2011.

[16] M. E. Shaheen, J. E. Gagnon, and B. J. Fryer, “Femtosecond laser ablation behavior of gold, crystalline silicon, and fused silica: a comparative study,” *Laser Phys.*, vol. 24, no. 10, p. 106102, Oct. 2014.

[17] E. Audouard, J. Lopez, B. Ancelot, K. Gaudfrin, R. Kling, and E. Mottay, “Optimization of surface engraving quality with ultrafast lasers,” *J. Laser Appl.*, vol. 29, no. 2, p. 022210, May 2017.

[18] P. Zhang, L. Chen, J. Chen, and Y. Tu, "Material removal effect of microchannel processing by femtosecond laser," *Opt. Lasers Eng.*, vol. 98, pp. 69–75, Nov. 2017.
[19] C. Doerbecker, H. Lubatschowski, S. Lohmann, C. Ruff, O. Kermani, and W. Ertmer, "Influence of the ablation plume on the removal process during ArF-excimer laser photoablation," presented at the BiOS Europe '95, Barcelona, Spain, 1996, pp. 2–9.
[20] S. Amoruso *et al.*, "Characterization of laser ablation of solid targets with near-infrared laser pulses of 100fs and 1ps duration," *Appl. Surf. Sci.*, vol. 252, no. 13, pp. 4863–4870, Apr. 2006.
[21] P. B. Chu *et al.*, "Controlled pulse-etching with xenon difluoride," in *Proceedings of international solid state sensors and actuators conference (Transducers' 97)*, 1997, vol. 1, pp. 665–668.
[22] M. Haras *et al.*, "Fabrication of thin-film silicon membranes with phononic crystals for thermal conductivity measurements," *IEEE Electron Device Lett.*, vol. 37, no. 10, pp. 1358–1361, 2016.
[23] A. Lecavelier des Etangs-Levallois *et al.*, "A converging route towards very high frequency, mechanically flexible, and performance stable integrated electronics," *J. Appl. Phys.*, vol. 113, no. 15, p. 153701, 2013.
[24] J. Philippe *et al.*, "Application-oriented performance of RF CMOS technologies on flexible substrates," in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 15.7.1-15.7.4.
[25] A. L. des Etangs-Levallois *et al.*, "150-GHz RF SOI-CMOS technology in ultrathin regime on organic substrate," *IEEE Electron Device Lett.*, vol. 32, no. 11, pp. 1510–1512, 2011.

Chapter 5: FLAME membranes for high-performance RF frontends

5.0 Introduction

In this chapter, the experimental studies of different SOI circuits listed in Chapter 3 are presented. The FLAME process has already been described in detail in Chapter 4 which deals with the suspension of membranes on SOI circuits. Applying these methods, membranes are created and characterized which is presented here. In section 5.1, isolation structures are studied to highlight the impact of substrate on coupling between 2 RF accesses through interdigitated structures by two port S-parameter characterization. After this, the RF SP9T switch is studied in section 5.2. The measured parameters are switch DC characteristics, small signal losses, large signal linearity, and power handling. This is the circuit with largest physical dimensions studied in this work. Next, integrated inductors are presented in section 5.3 comprising of both single and two turn inductors. The impact of substrate on Q-factor is assessed here by making one port S-parameter measurements. Finally, the impact of inductors on LNA performances (Gain, IIP3, P1dB and Noise Figure) is examined in section 5.4. Different substrate removed configurations are studied with the object of understanding the relative importance of each inductor. For some of the measurements, a brief description of the measurement bench has also been provided.

Contents

5.1 Isolation structures 159
5.1.1 FLAME process parameters 159
5.1.2 S-parameter characterization 159
5.2 RF SP9T switch 163
5.2.1 FLAME process parameters 164
5.2.2 DC characterization 166
5.2.3 S-parameter characterization 167
5.2.4 Large signal characterization 170
5.2.4.1 Measurement setup 170
5.2.4.2 Results 172
5.3 Inductors 176
5.3.1 FLAME process parameters 176
5.3.2 S-parameter characterization 177
5.3.3 Modelling of inductors on membranes 185
5.4 Low noise amplifier (LNA) 189
5.4.1 FLAME process parameters 189
5.4.2 Noise figure measurement 190
5.4.3 Standalone input inductor measurement 191
5.4.4 Linearity measurements 192
Concluding Remarks 194
References 195

5.1 Isolation structures

Isolation structures are simple interdigitated structures designed to study substrate coupling. Two isolation structures are studied with different number of fingers and spacing as described in section 3.7.1. Small signal S_{21} is the parameter of interest for this study which quantifies the coupling through substrate. The study of isolation structures serves to distinguish between HR and TR substrates. Also, after post laser processing, since the substrate is completely removed, the coupling occurs only through the intermetallic capacitance. There shouldn't be any difference observed in coupling values between HR and TR substrate.

5.1.1 FLAME process parameters

750 µm dies are grinded to have a starting thickness of ~250 µm before FLAME process. 1 device is processed for each substrate type. All the isolation structures have the same area of membrane to be created. A two-step milling process is used with the bigger cavity being a square of dimension 1.5 mm. Two smaller cavities are created for ISO-A and ISO-B within this bigger cavity. The parameters used for milling are summarized in Table 5.1. The drawn layout area used for laser milling of smaller cavity is 0.24 x 0.26 mm^2. This area is suitable for the isolation structures for which the targeted membrane area is 0.34 x 0.36 mm^2. XeF_2 etching increases the milling area defined in the laser trajectory (drawn layout area) by 100 µm in either direction because of lateral etching. For XeF_2 etching, a pressure of 2 Torr is used with a cycle period of 10 s and it takes ~50 cycles to fully suspend the circuit.

Table 5.1: FLAME process parameters for isolation structures

Parameter	Fast step	Slow step
Scanner speed (mm s^{-1})	20	5
Pulse repetition freq (kHz)	30	2
Laser power (W)	1.006	0.0134
Number of passes	4	180
Fluence (J cm^{-2})	32.1	6.4
Depth Window (µm)	49.9	36.8
Removal rate (x 10^6 µm^3 s^{-1})	4.75	0.02

5.1.2 S-parameter characterization

2-port S-parameter characterization is done over a frequency range of 20 MHz to 26 GHz using setup shown in Fig. 5.1. The measurements are performed using

R&S ZVA 67 vector network analyzer. Infinity probes (GSG type) from Cascade Microtec are used to make contact with the RF pads on the die. The substrate is also biased by biasing the chuck. S-parameter measurement is performed for different substrate bias conditions with the bias values swept from -2.5 V to +2.5 V with a step of 0.1 V. The isolation between the two ports is characterized by S_{21} (dB) parameter. Lower value of S_{21} corresponds to better isolation between the two ports. Isolation values before laser processing for different frequencies plotted as a function of chuck bias are shown in Fig 5.2. HR substrate shows a strong dependence of S_{21} on chuck bias for MHz frequencies. This is because of the non-linear MOS like behaviour of the BOX capacitance of HR substrate. However, for TR substrate, the S_{21} response at a given frequency is flat as a function of chuck bias. This is because of the elimination of parasitic surface conduction layer by adding a trap-rich layer at the interface. At frequencies ≥5 GHz, the S_{21} curves are flat for both substrates. Additionally, S_{21} behaviour of HR and TR substrate is the same at these frequencies. Thus, at high frequencies HR and TR substrate behave quite similarly with respect to coupling. By removing the substrate under these isolation structures, it would be possible to determine if the substrate plays an important role at these frequencies or if the BEOL parasitics dominate the S_{21} behaviour.

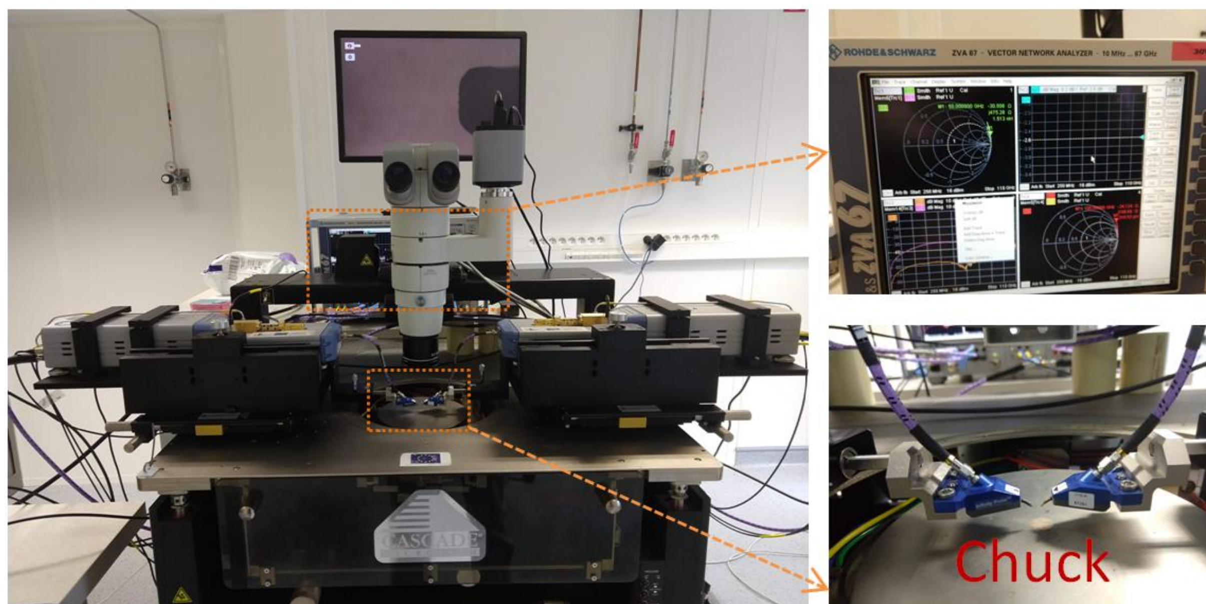


Fig. 5.1: 2-port S-parameters measurement bench

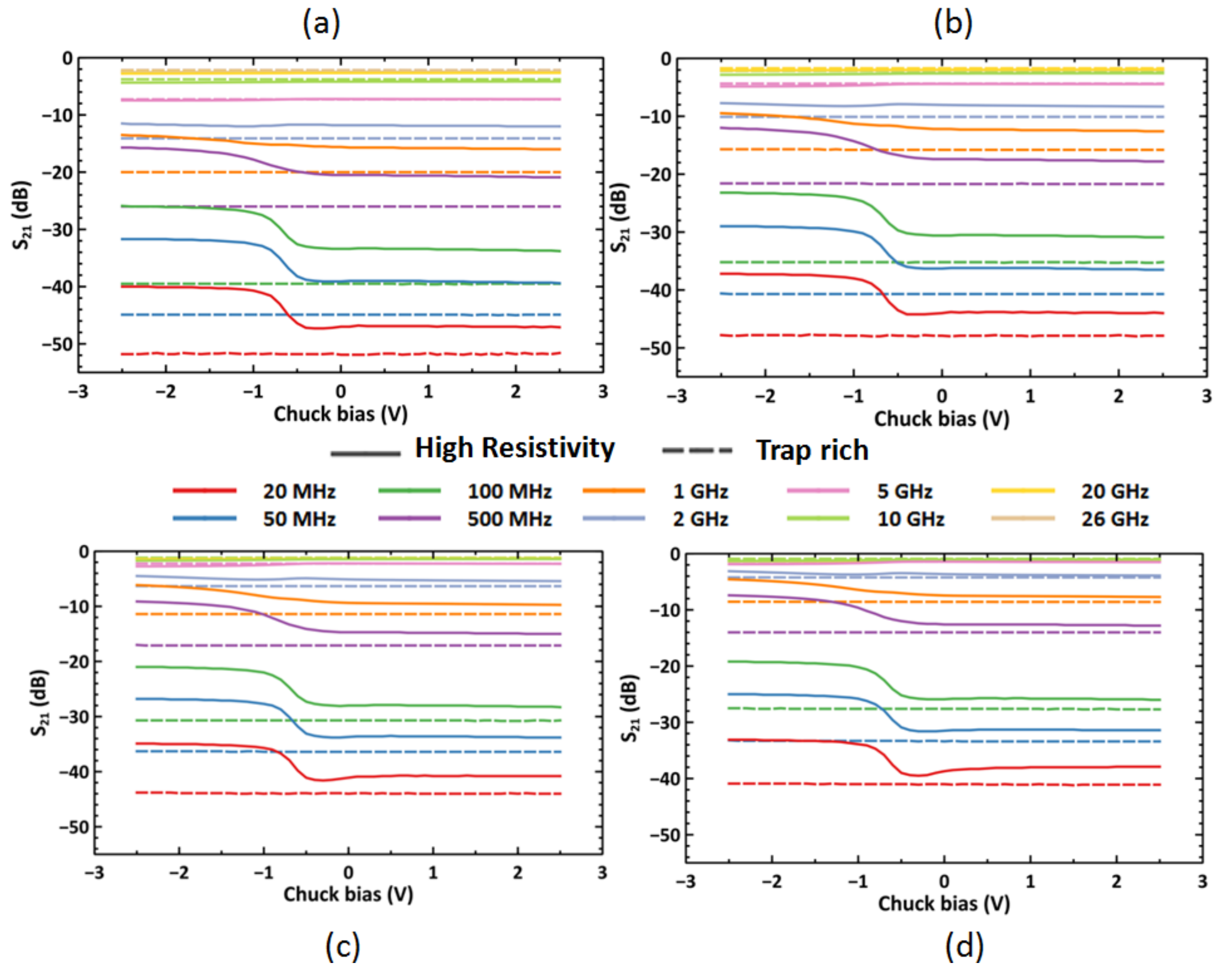


Fig. 5.2: Isolation values before FLAME process (a) ISO-A (b) ISO-B (c) ISO-C (d) ISO-D

Substrate removal is performed on ISO-A and ISO-B for both substrate types. S_{21} values are plotted after substrate removal in Fig. 5.3. S_{21} differences between the two substrate types are not seen after FLAME process. At all frequencies, the isolation values for the 2 substrates are equal. These results indicate that the FLAME process for suspending RF circuits works well and the physical integrity of the circuit stack is maintained after substrate removal. In ISO-A, there is a very small but noticeable difference in isolation between the two substrates. These minor differences occur because of small differences in targeted and actual area of membrane for the two cases.

The etched area of the membrane is visualized using dual-light microscopy (DLM) where both the front side and back side are illuminated with varying intensities. The high intensity backlight through the cavity appears brighter in the microscope image outlining the area of membrane. The DLM images for ISO-A and ISO-B are shown in Fig. 5.4. It can be seen that bulb like features appear at some places around the outline of the membrane. These bulb-like features are

a result of uneven local etching characteristic of XeF_2. Some places have an enhanced etch rate in the lateral direction as compared to the others. These features can be avoided by fine tuning the process in order to have fewer cycles of XeF_2 etching.

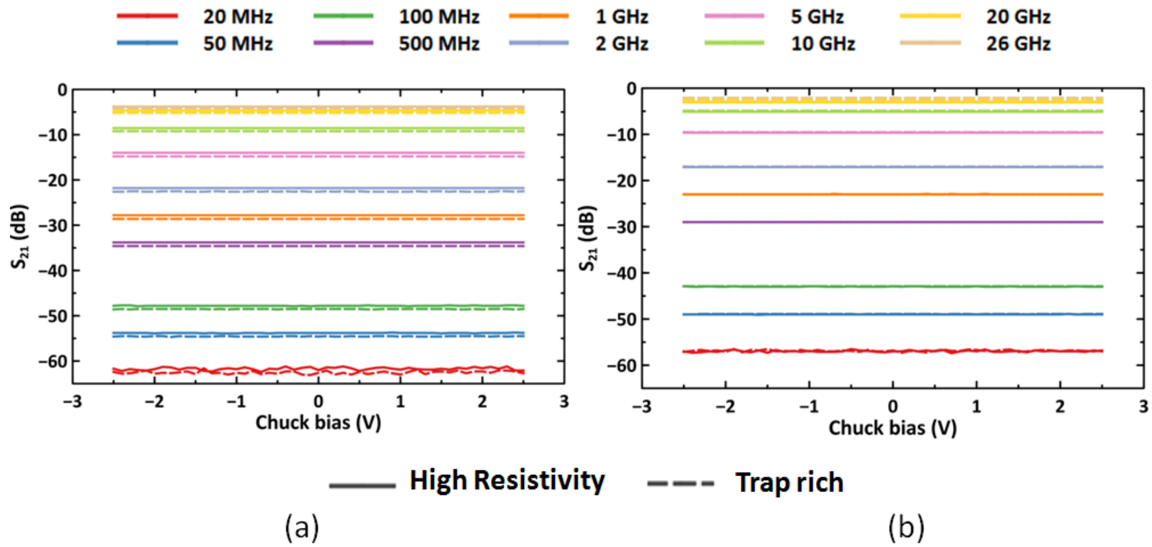


Fig. 5.3: Isolation values after FLAME process (a) ISO-A (b) ISO-B

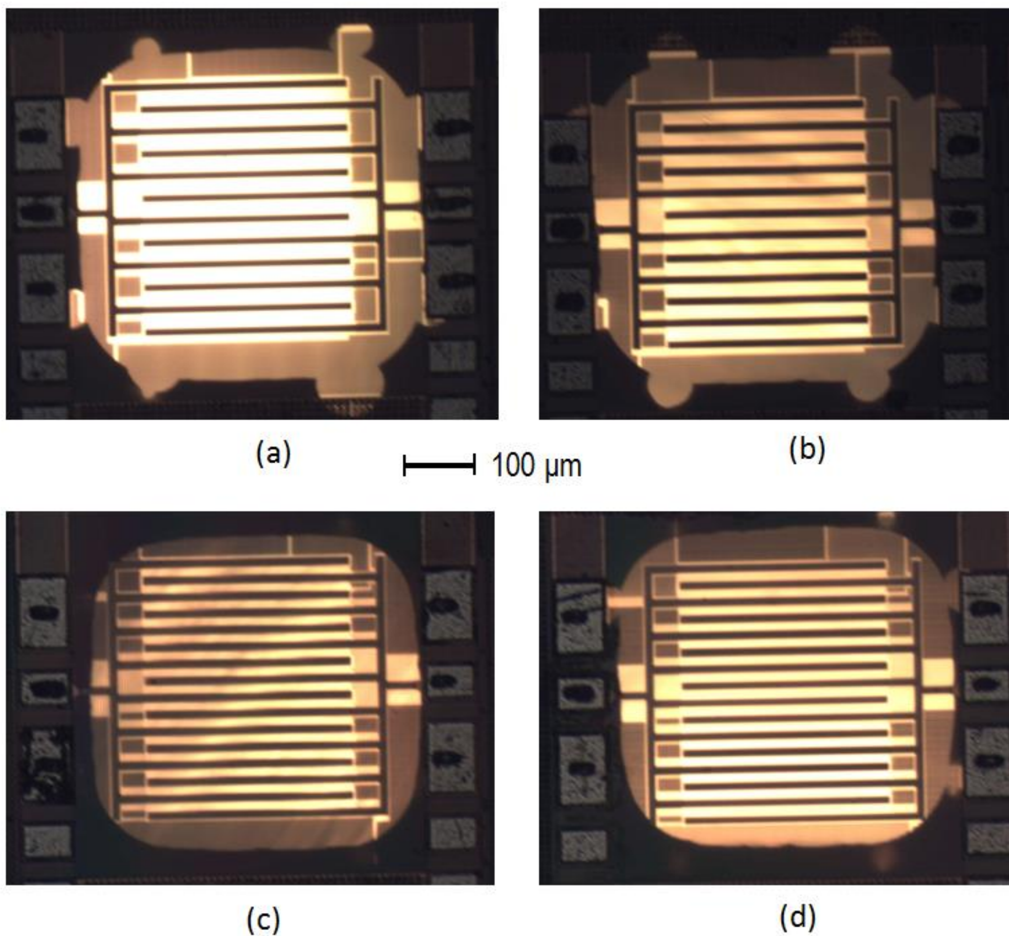


Fig. 5.4: DLM images after FLAME process (a) ISO-A HR (b) ISO-A TR (c) ISO-B HR (d) ISO-B HR

In order to compare the isolation performances before and post FLAME process, the difference in isolation ($\Delta S_{21} = S_{21\,(HR/TR)} - S_{21\,(SR)}$) where SR refers to Substrate Removed is plotted for ISO-A and ISO-B in Fig. 5.5. For MHz frequencies up to 100 MHz, the difference in isolation can exceed 20 dB for HR substrate for both isolation structures. The difference in isolation is smaller for TR substrate with values of 10.8 dB and 9 dB for ISO-A and ISO-B respectively at 20 MHz. With increasing frequency, the difference becomes smaller and approaches 0 dB. This indicates that at higher frequencies coupling through the intermetallic capacitance dominates the behaviour of the circuit. However, at frequencies of operation of practical interest like 5 GHz, there is still a significant difference in isolation. The ΔS_{21} values (HR/TR) are 6.8/7.5 dB for ISO-A and 5.2 dB for ISO-B at 5 GHz and chuck bias of 0 V. Thus, coupling due to substrate parasitics can be considerably reduced by suspending the substrate.

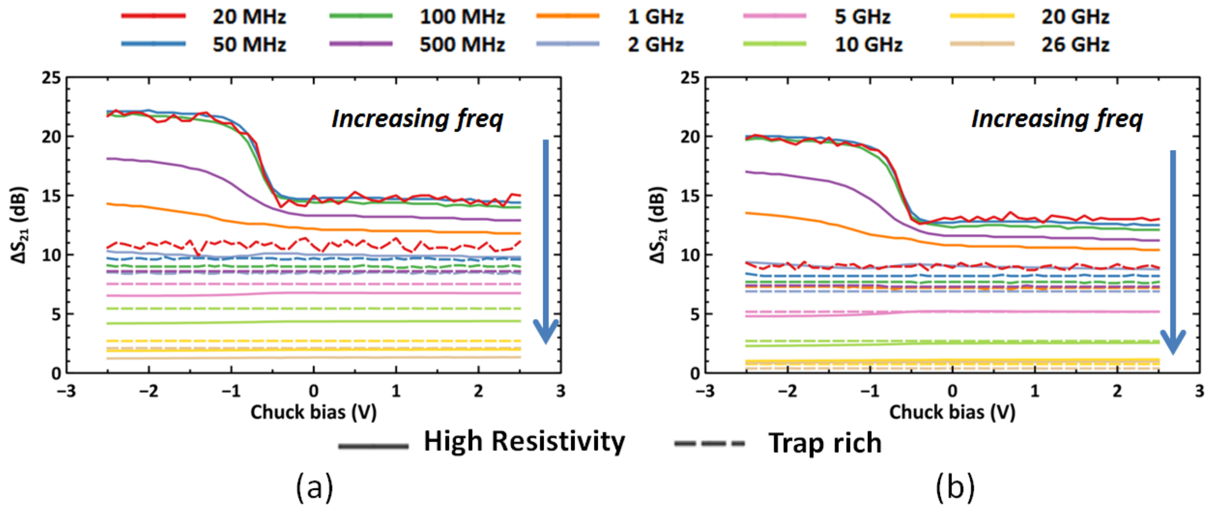


Fig. 5.5: Difference in isolation values ($\Delta S_{21} = S_{21(Before\ FLAME)} - S_{21(After\ FLAME)}$)(a) ISO-A (b) ISO-B

5.2 RF SP9T switch

RF SP9T switch chosen in this study is a test structure designed to analyze the performance parameters close to real world operating conditions of a T/R switch. The design of the switch has been described in section 3.7.2. In this section, the laser processing is first outlined for SP9T switch. Following the description of process, the DC and RF performance parameters are analyzed before and after substrate removal for 2 different channel lengths and different bias conditions for both substrates.

5.2.1 FLAME process parameters

The starting thickness of substrate is 750 µm for SP9T switch. Two-step milling is used for suspending membranes of SP9T switch circuits. In the first step, an area of 1.66 x 2.12 mm^2 is milled to get a larger cavity of 450 µm depth. Within this cavity, the 2nd cavity is milled until remaining thickness of silicon to be etched is ~100 µm. The targeted area for 2nd step of milling of the switch membrane is 1.26 x 1.72 mm^2. A margin of 200 µm is used between outer and inner milling areas. The milling parameters are listed in Table 5.2. For step 1, fast removal parameters are used and step 2 slow removal parameters are used. The obtained cavities after step 1 and 2 are shown in Fig. 5.6. The milling parameters for slow step are chosen such that a removal rate is one order of magnitude lesser than the fast step. This is to avoid accumulation of ablated plume close to the milling area. Also, lesser fluence reduces the melting effects and heat affected zone.

Table 5.2: Milling parameters of SP9T switch

Parameter	Fast step	Slow step
Scanner speed (mm s^{-1})	30	20
Pulse repetition freq (kHz)	30	40
Laser power (W)	1.45	0.26
Number of passes	10	30
Fluence (J cm^{-2})	34.7	8.3
Depth Window (µm)	116.1	121.2
Removal rate (x 10^6 $µm^3$ s^{-1})	6.87	0.35

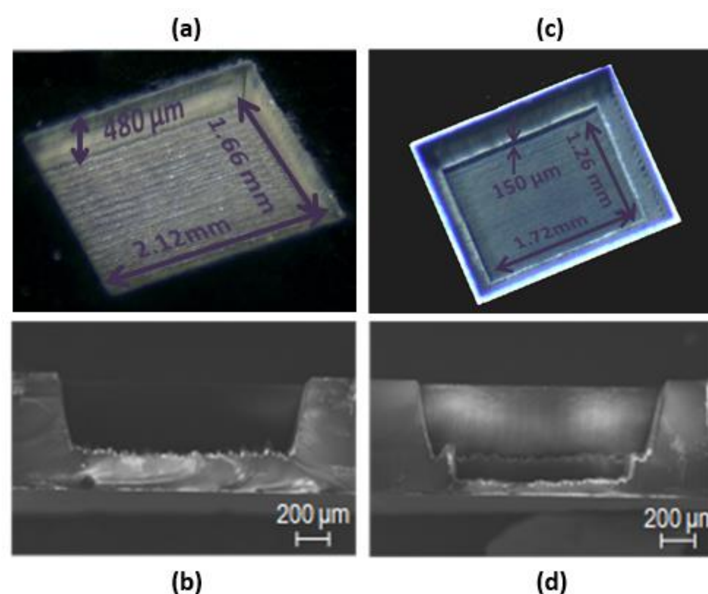


Fig. 5.6: (a, c) Optical microscope images of cavities obtained for 2-step milling process of SP9T switch. (b,d) Cross sectional images of the fast and slow milling steps respectively.

At the end of two step milling, optical profilometer scan is taken for depth profiling. Finally, XeF_2 etching is performed on the cavity to suspend the membranes. An etch pressure of 3 Torr and 100 cycles are used to complete the etching. The progression of etching is shown in Fig. 5.7. Because of a large depth window, in some regions etching is already complete at the end of 25 cycles. In order to fully suspend the structures, additional 75 cycles are necessary. At the end of etching process, the circuit is fully visible from the backside and closer inspection reveals the absence of any residues of silicon at the end of the etching. The DLM image is taken to know the outline of the membrane at the end of the FLAME process.

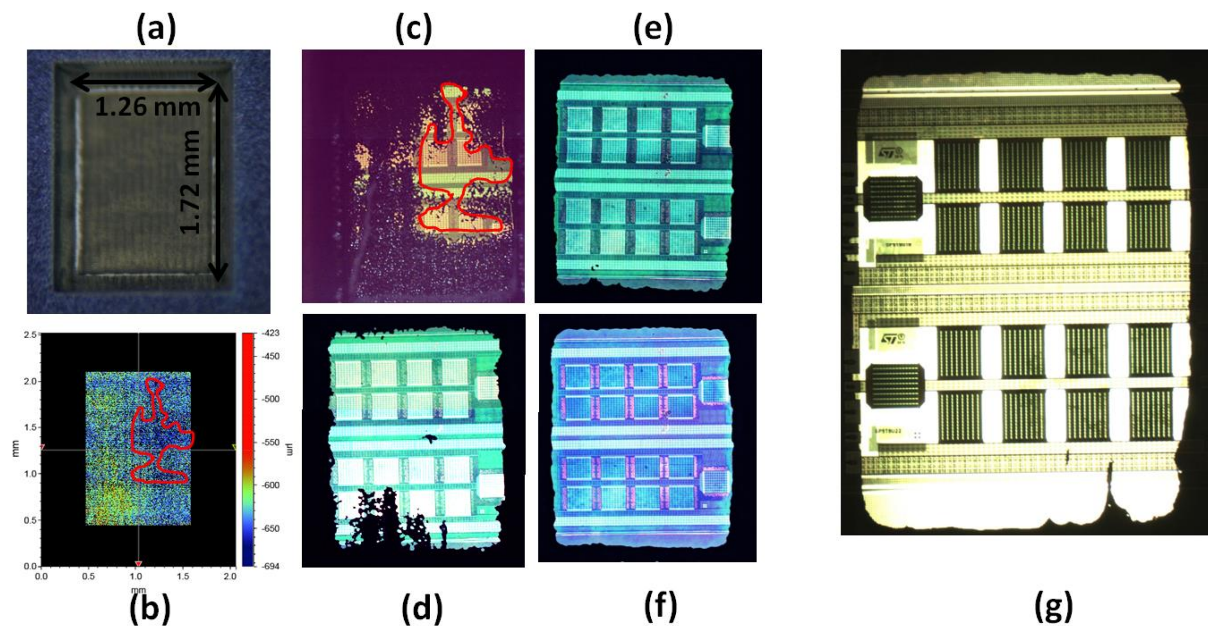


Fig. 5.7: (a) Optical micrograph of switch cavity after laser ablation (b) Profilometer scan of the ablated area with outlined region showing locally deep regions (c) After 25 cycles of XeF_2 etch. Outlined area etched first suggesting a uniform vertical etch rate of cavity (d) After 50 cycles of XeF_2 etch (e) After 75 cycles of XeF_2 etch (f) After 100 cycles of XeF_2 etch (g) View of the switch front side using backlight illumination from the cavity side.

The SP9T switch was the first studied circuit using the FLAME process. Because of unoptimized alignment procedure, small changes in laser power and poor repeatability of XeF_2 etching step, the membrane offset is typically $\pm 20 - 40$ µm in the x-direction from the intended area of etch. The FLAME process has been used for other studied circuits and xy-offset is substantially lesser < 10 µm. The y-direction offset is not considered because additional etch area is chosen so that part of etched area is outside the functional area of the switch. Any differences in y-offset would still completely etch the full switch area in the y-direction.

5.2.2 DC characterization

The removal of substrate changes the RF behaviour of the circuit while DC performance is not affected. Hence, DC characterization is performed on the RF switch to determine the quality of the substrate removal process. I_D-V_{DS} characterization is performed on one device on TR substrate. As seen in Fig. 5.8, for both channel lengths, there is very good agreement between the measurements before and after substrate removal.

For the same applied gate voltage, the drain current is smaller for 220 nm channel length as compared to 180 nm channel length. Hence, a smaller on state resistance and insertion loss is expected for 180 nm switch. Also, increasing gate voltage reduces the resistance as observed by the slope of the IV curve in the linear region. The transistor operates in the linear region during the ON state. Hence, use of smaller gate length and higher gate drive is favourable for lower ON state resistance. This point is further elaborated when discussing S-parameter measurements.

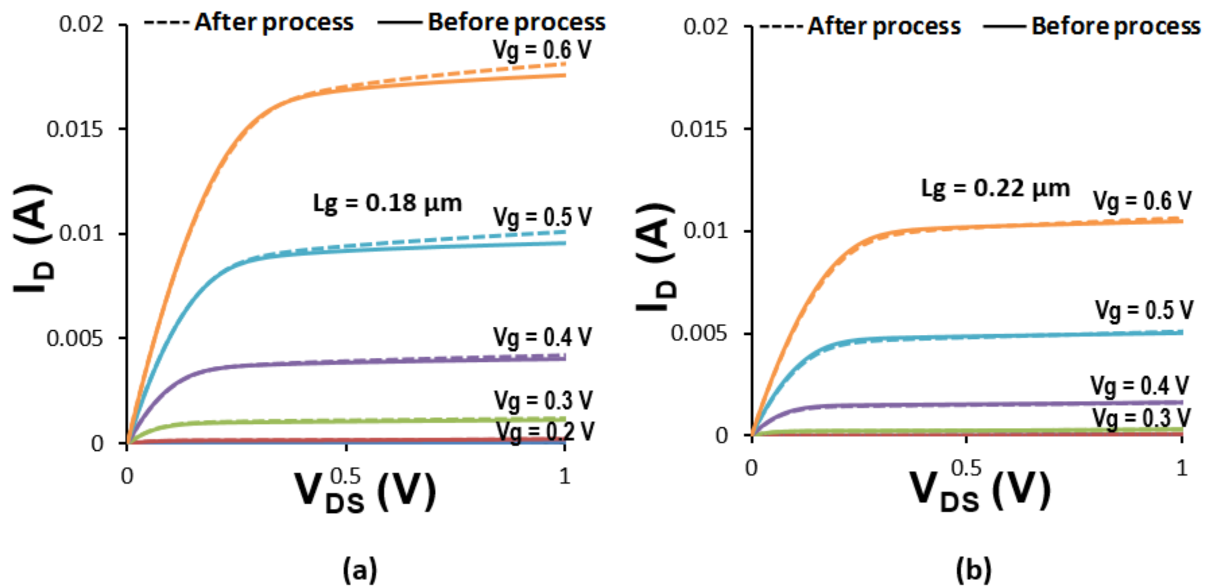


Fig. 5.8: DC characteristics of the RF switch on TR substrate for 2 channel lengths before and after substrate removal (a) 180 nm (b) 220 nm

With the increase in gate voltage beyond 0.6 V, self-heating of the transistor stack becomes visible in the IV curves as depicted in Fig. 5.9. In the saturation region of the transistor, the drain current drops with increasing gate drain voltage. This is because of heating induced by the DC power dissipated in the ON stack. The heating effect starts approximately at gate drive > 0.7 V. This corresponds to current density values of 12.4 µA/µm and 8.1 µA/µm for gate

lengths 180 nm and 220 nm respectively. Taking maximum V_{DS} of 2 V at this current density, self-heating can be seen if the static power dissipation is more than 56 mW and 37 mW for channel lengths of 180 nm and 220 nm respectively. With higher gate drives the self-heating gets worse with a large deviation from IV characteristics before laser processing. This will limit the maximum power of operation of the switch.

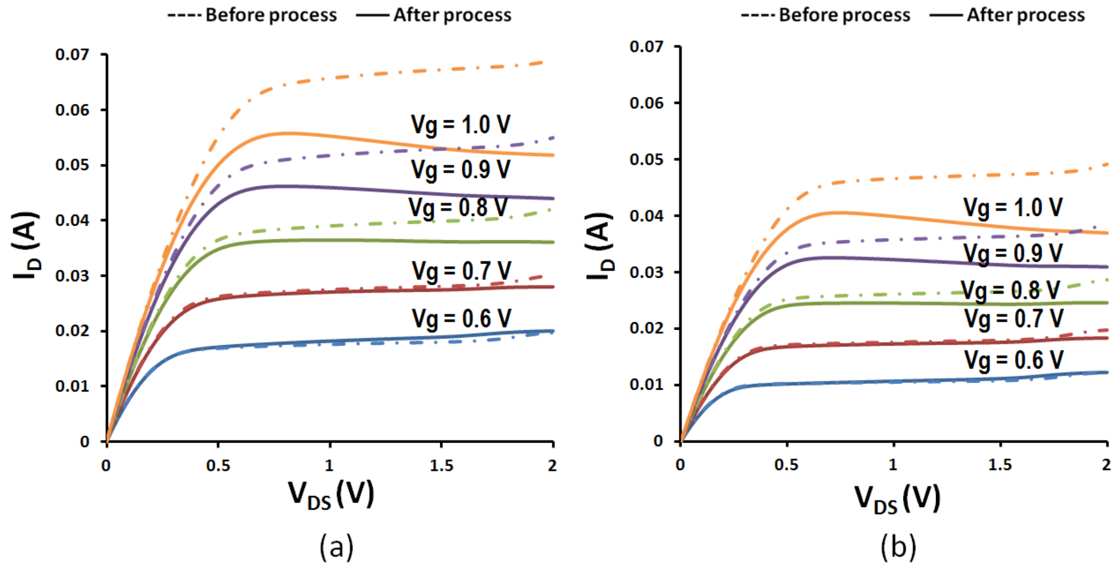


Fig. 5.9: Self heating of RF switch seen in DC characterization on TR-SOI (a) 180 nm (b) 220 nm

5.2.3 S-parameter characterization

S-parameter measurements are performed for the SP9T switch to determine the insertion loss and matching of the switch. Two samples are characterized before and after substrate removal for each substrate type. The value of insertion loss can be taken approximately as $-S_{21}$ (dB) because the magnitude of the reflected signal on both ports are negligible compared to transmitted signal. The results are shown in Fig. 5.10 for both HR and TR substrate. It can be seen that the insertion loss is reduced after substrate removal. Of the 4 bias conditions only two bias conditions are reported because the differences between B2, B3 and B4 are negligible under small signal conditions. Before substrate removal, bias condition B2 with a gate overdrive of 3.3 V is favourable for smaller losses as compared to bias B1 of 2.5 V, which is the process technology specification. Also, insertion loss is lesser for 180 nm channel length as compared to 220 nm. Both of these improvements can be attributed to reduction of on-state resistance of the transistor stack which dominates the ON-behaviour of the switch. After substrate removal, the same trend is retained with respect to channel length and gate voltage. The insertion losses remain about the same at low frequency and reduce more notably at high frequencies. This is because at MHz frequencies switch

insertion losses are dominated by Ron which does not change appreciably after FLAME. The small differences observed at low frequencies can be attributed to varying contact resistance which appears in series with Ron and contributes to insertion losses. At higher frequencies, the insertion losses are dominated by the OFF state branches. These branches result in losses due to substrate coupling due to their relatively large size. At 6 GHz, for HR substrate the improvement in S_{21} is 0.4 dB. For TR substrate, the improvement in S_{21} is 0.28 dB which is smaller as compared to HR substrate because presence of trap-rich layer which alleviates substrate losses. This number is nevertheless significant for the performance of the switch. Converting dB to absolute values, at small signal, the P_{out}/P_{in} ratio is calculated as 0.807 and 0.861 for TR substrate and post substrate removal cases respectively. The improvement in P_{out}/P_{in} ratio is nearly 6.8%. The improvement in power dissipation is potentially higher under large signal conditions.

Ideally, the switch acts as a short circuit between the source and load and there are no reflections at the switch input port. But the switch has finite impedance and hence reflections occur at the input port. The reflected signal can be quantified by S_{11} (dB) which is plotted for HR and TR substrate in Fig. 5.11. The signal reflection is a function of both the on state resistance of the ON branch and the combined capacitance of the OFF branches. It can be seen that reflection is higher for gate length of 180 nm as compared to 220 nm. The higher reflection can be attributed to larger combined capacitance of OFF transistor stacks. Smaller channel length results in increased effective source to drain coupling and hence a larger OFF state capacitance. Thus, larger channel length is better to reduce mismatch losses.

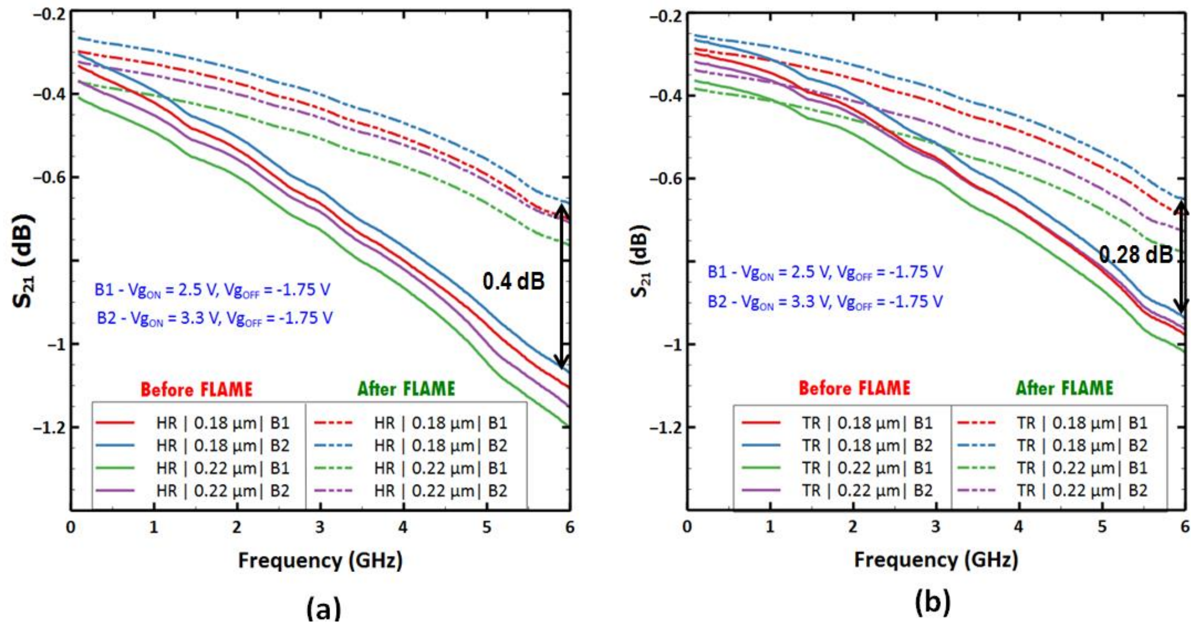


Fig. 5.10: S_{21} plot for two different channel lengths and bias conditions before and after FLAME (a) HR-SOI substrate (b) TR-SOI substrate

By removal of substrate, the parasitics due to the substrate are reduced which helps improvement of matching. The improvement of both matching and insertion loss is prominent at higher frequencies. There is ~4 dB improvement in return loss at 6 GHz for both substrates.

For both HR and TR substrate the reported improvements are conservative as observed from contact resistance. The total series resistance of the switch including the pads is extracted using the S_{11} measurement at 100 MHz. It is seen that for measurements after substrate removal, the extracted resistance is ~0.5 Ω higher than for measurements before substrate removal. This is because pads are probed multiple times before substrate removal. For aluminium pads, we see that contact degrades after multiple instances of probing. The consequence is the increased series resistance of the pad. This contributes to higher measured values of both insertion losses and matching parameter.

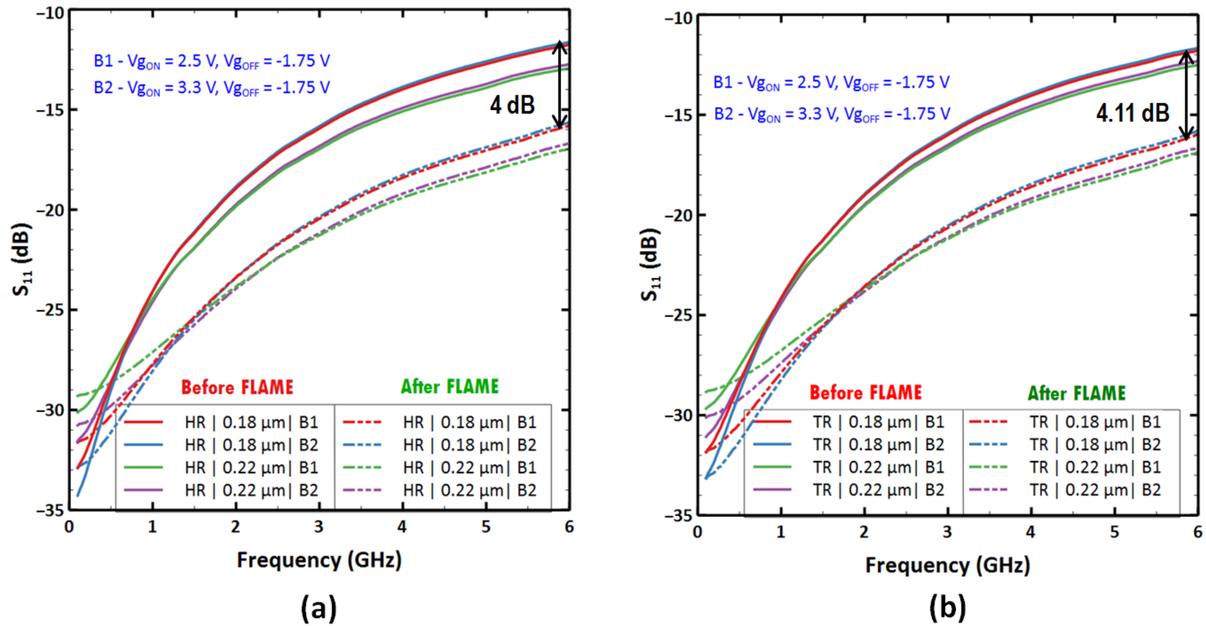


Fig. 5.11: S_{11} plot for two different channel length and bias conditions before and after FLAME (a) HR-SOI substrate (b) TR-SOI substrate

5.2.4 Large signal characterization

For the SP9T switch, large signal measurement of interest is the harmonic distortion. This is because the switch is designed to operate in the GSM-900 and GSM-1800 bands. For linearity measurement, a low noise floor bench has been realized which is described in the next section.

5.2.4.1 Measurement setup

The setup for linearity measurement of the switch is depicted in Fig. 5.12. Each component of the bench produces its own harmonics and it becomes necessary to eliminate the non-linearity of the bench components as much as possible. At the input stage, the fundamental signal from the VNA is selected by the low pass filter and 2nd and 3rd harmonics are rejected. The power amplifier provides the necessary power to the device and the harmonics of the PA is again filtered by using another low pass filter. The bidirectional coupler allows the measurement of incident and reflected signals on the input side. DC biasing for the switch transistor stacks are provided by using bias-T for RF isolation.

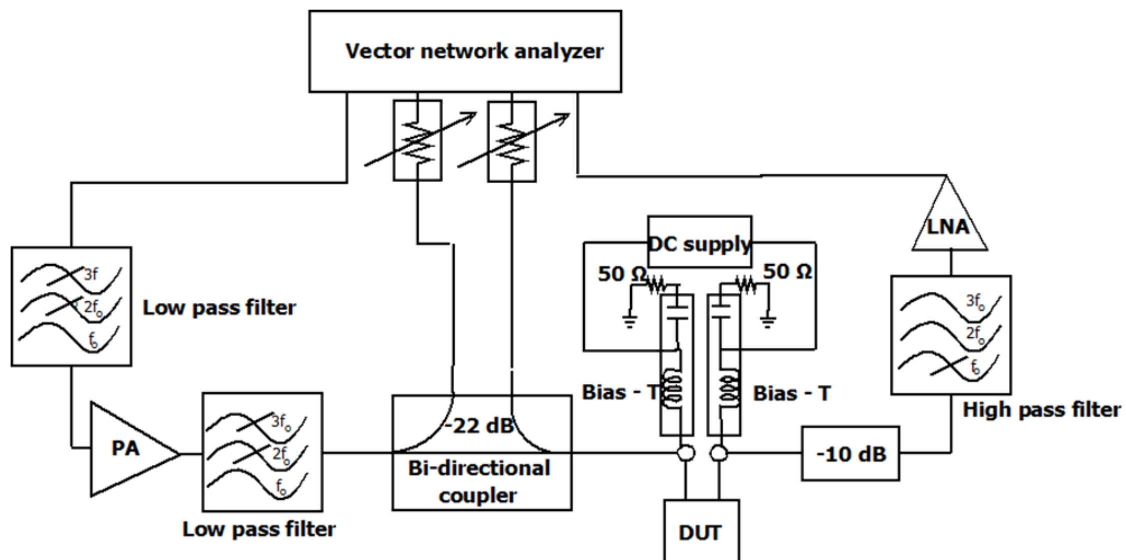


Fig. 5.12: Schematic of 2nd and 3rd harmonic distortion measurement bench

On the output side, a 10 dB attenuator provides 50 Ω matching at the switch output port. Additionally, it serves to attenuate the fundamental signal by 20 dB noting that the fundamental signal is reflected at the high pass filter. While a higher value of attenuator provides better isolation of reflected signal, the noise floor is compromised. An isolator at the output port of the switch is the ideal solution to isolate the reflected fundamental frequency. But, because of poor linearity of the isolator, a simple attenuator is preferred. Hence, a small part of the signal at fundamental frequency reflected from the high pass filter is present at the switch output port. The LNA amplifies the harmonics generated by the switch which is measured by the vector network analyzer. The measurement setup is calibrated using Labview and a thru structure is used on the Calkit (Cascade Microtec: 101-190) to verify the calibration. The final scalar power measurement of 2nd and 3rd harmonics is recorded on Labview. Fig. 5.13 shows the measured harmonics on the Calkit Thru. It can be seen that the noise floor of the setup is -140 dBm. The excellent dynamic range offered by the setup allows for measurement of switch harmonics starting from lower values of injected power.

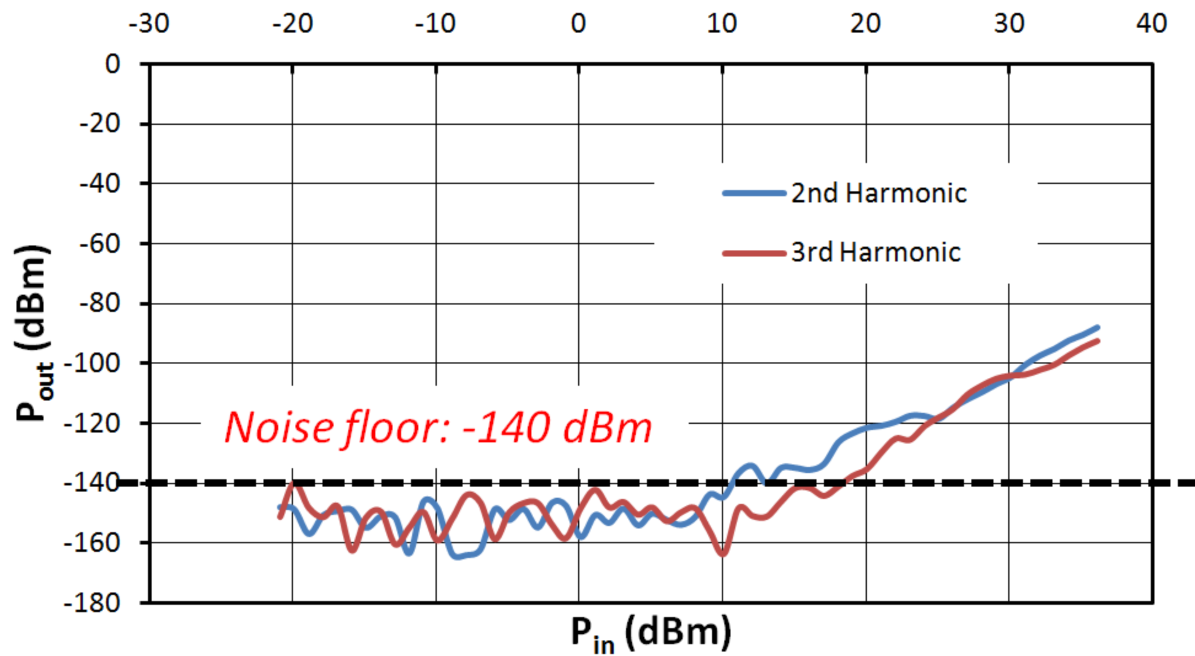


Fig. 5.13: Linearity measurement of the bench using a thru line on Calkit

5.2.4.2 Results

The number of processed samples for H2/H3 characterization is 4 and 5 for HR and TR substrate respectively. The dispersion of measured data is small before substrate removal and hence the same would be expected after substrate removal if the process is well controlled. However, after processing some samples have inconsistencies in measured data with high dispersion from mean values and these points are not considered. The reasons for such inconsistencies could be multiple such as processing defects, improper probe contact to the pad due to pad degradation after probing several times, non-linearities introduced by the pad, and offset between targeted area and removed area.

An illustration of effect of etched area on linearity is shown in Fig. 5.14, where the handler silicon of same switch is etched to varying degree and the linearity is measured after each etch. The first etch results in an etch area which leaves a part of silicon left under the ON branch. After the device is characterized, it is etched further laterally by exposing the same sample to 10 more cycles of XeF_2 etching. The device is characterized the 2nd time. The same procedure is followed for the 3rd etch. It can be seen that small difference in areas exposed after etching can have a big impact on the linearity measurement. Additionally, the thickness of silicon under the pads also reduces with subsequent etch steps.

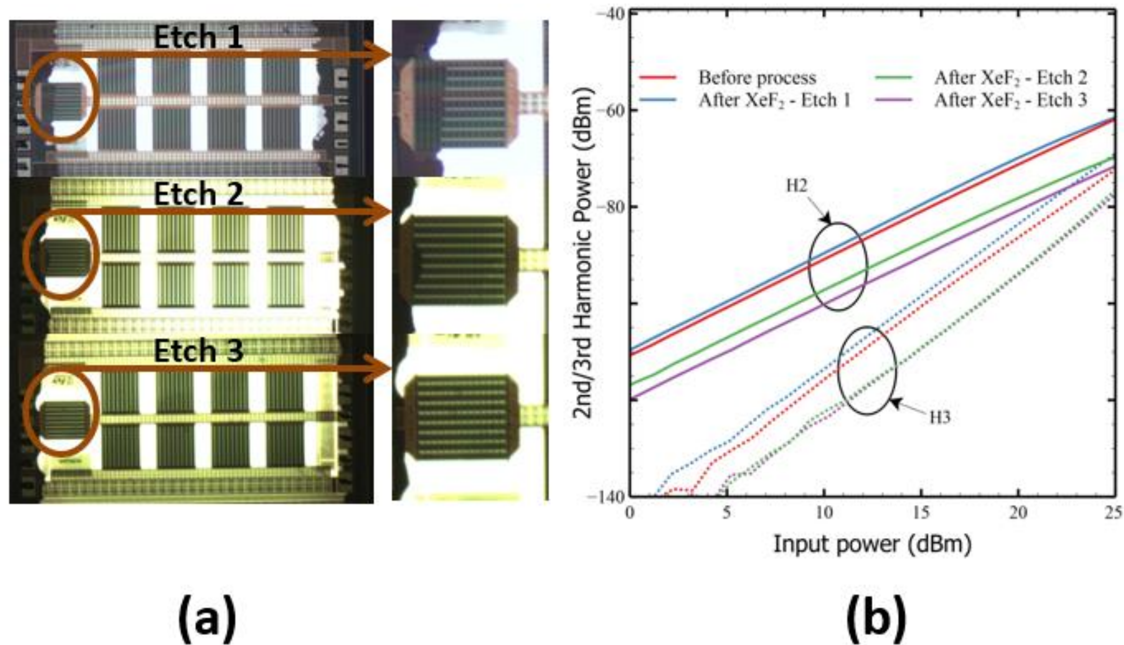


Fig. 5.14: Harmonic distortion dependence on area of silicon etched measured at bias B1 on HR substrate

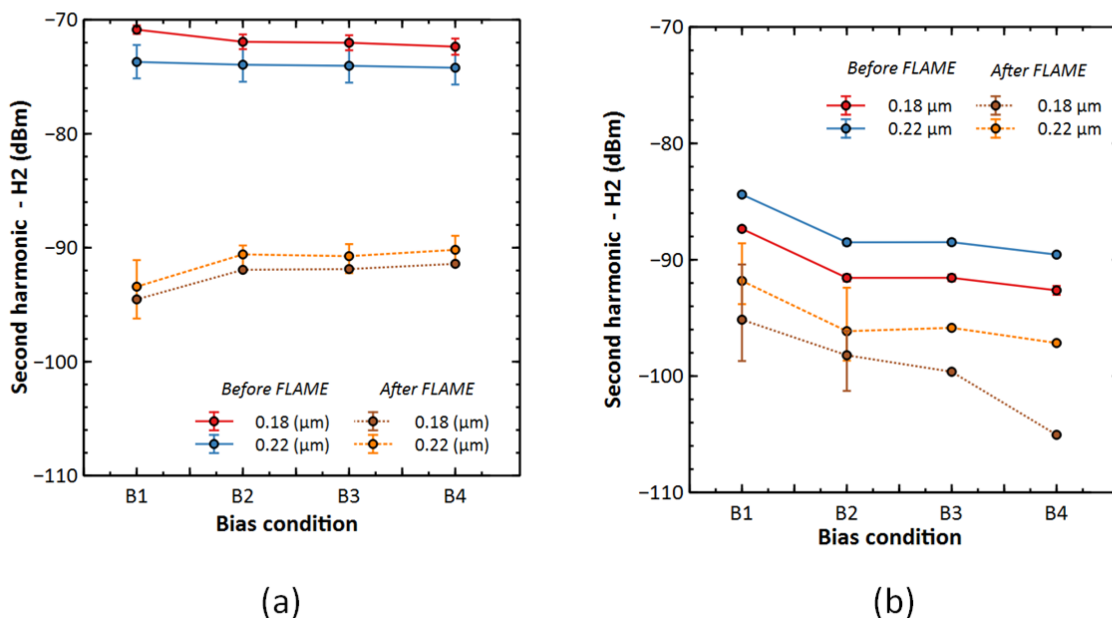


Fig. 5.15: 2nd harmonic distortion before and after FLAME process (a) HR-SOI (b) TR-SOI. Symmetric error bars represent single standard deviation.

The second harmonic (H2) measurements are performed for the switch before and after FLAME process for HR and TR substrate and the results are summarized in Fig. 5.15. It has been indicated by Wei et al. that H2 depends on circuit symmetry and parasitic from gate to ground [1]. Based on this argument, the additional contribution of BOX capacitance and the substrate impedance network

increases the gate parasitics. A removal of substrate would mean elimination of these parasitics and hence lesser harmonic distortion could be expected. Accordingly, there is an improvement in linearity for both TR and HR substrate. The improvement in H2 for HR substrate is the key result that is obtained in the experiment. For HR substrate, H2 is higher for gate length of 0.18 μm as compared to 0.22 μm before substrate removal. This is contrary to the hypothesis that smaller channel length means smaller gate parasitics and hence lesser gate distortion. However, the H2 performance depends not only on the substrate parasitics. In a SP9T switch, there are 8 OFF branches when one branch is ON. The non-linearity of these branches depends on the non-linear behaviour of gate – source and gate – drain capacitances [2]. In order to model the overall linearity of the switch, the combined effects of different parasitics and branches on linearity is to be considered. For HR substrate, the H2 is 3-4 dB higher for channel length of 0.18 μm as compared to 0.22 μm. For TR substrate, the opposite effect is observed where H2 is worse for 0.22 μm channel length. After substrate removal, H2 is larger for 0.22 μm channel length for both HR and TR substrate cases. In all cases, before and after substrate removal, increasing the gate voltage for ON stack improves the H2 slightly.

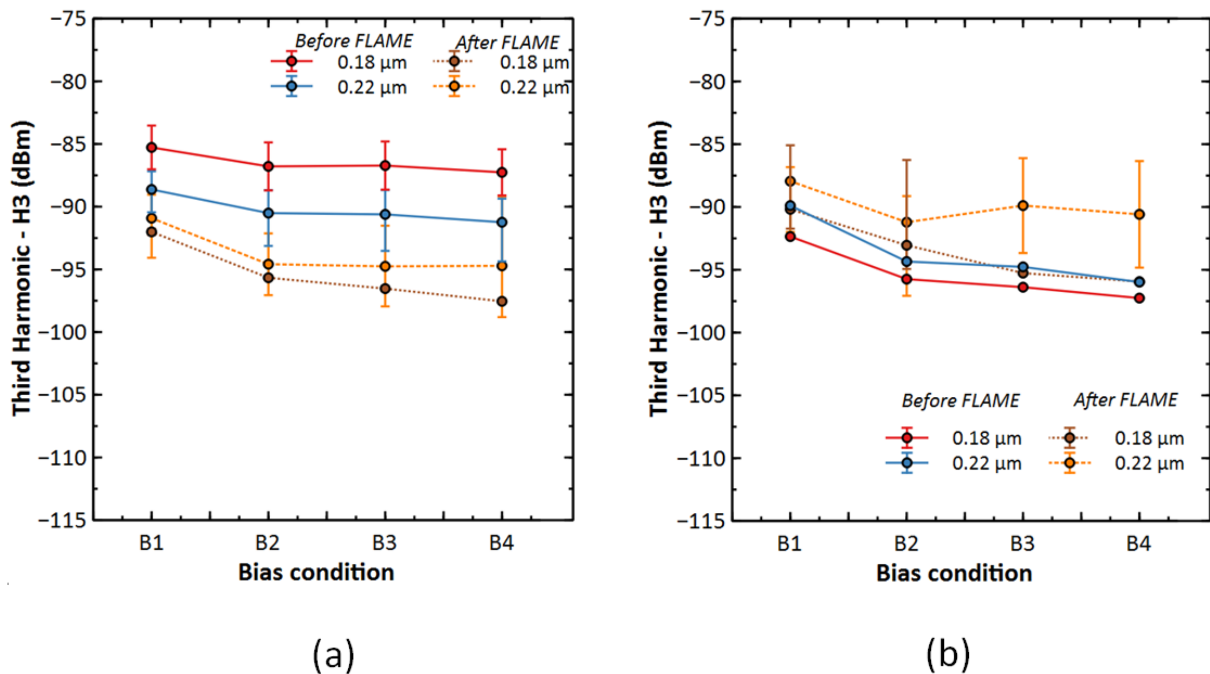


Fig. 5.16: 3rd harmonic distortion before and after FLAME process (a) HR substrate (b) TR substrate. Symmetric error bars represent standard deviation.

The 3rd harmonic (H3) measurements are summarized in Fig. 5.16. The same observations can be made here as in case of 2nd harmonic with respect to channel length and gate voltage dependence before and after substrate removal. For HR

substrate, there is 6-7 dB improvement of H3. In case of TR substrate, H3 increases marginally after substrate removal by 2-3 dB. One reason for this apparent increase in H3 is that the transmitted power in the switch is higher after substrate removal because of improved matching. Since, the transmitted signal cannot be measured in the existing setup; the increase in transmitted power after substrate removal cannot be quantified. In summary, H2 and H3 improvement is seen for HR substrate. For TR substrate, only H2 improvement is seen. The accuracy of the results can be improved by processing more samples.

The results of harmonic distortion are summarized as follows. For HR substrate at bias condition B1 and gate length of 0.18 μm the improvement obtained after substrate removal is ΔH2 = 5.6 dB and ΔH3 = 22.7 dB. For TR substrate at same bias condition and channel length, the improvement obtained after substrate removal is ΔH2 = 7.2 dB and ΔH3 = -2.2 dB. The negative sign of ΔH3 indicates a very small increase in H3 after substrate processing.

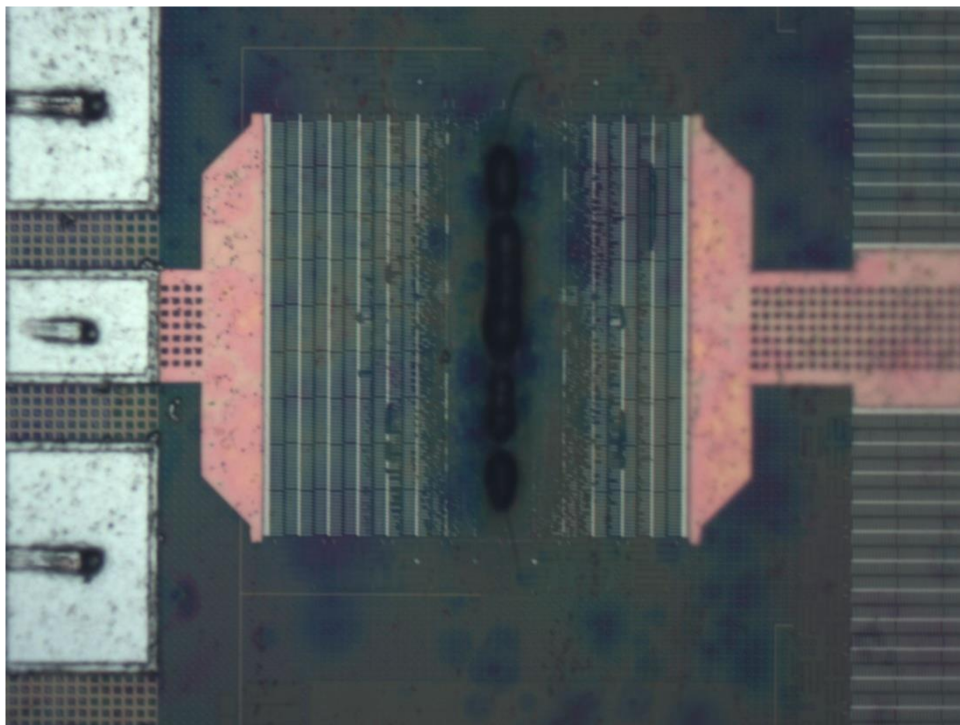


Fig. 5.17: Failure of ON transistor caused by excessive self heating at high input power

DC measurements previously indicated self-heating of the on transistor. It is thus interesting to measure the power handling capacity of the switch. Fig. 5.17 shows the burned membrane because of transistor overheating at an input power of 30 dBm. While the membranes cannot support high power rating, there are significant benefits for low power applications.

5.3 Inductors

Integrated inductors are important components for realization of high-performance RF and microwave circuits in SOI technology. Foundries provide different types of inductor libraries to support the required applications. For instance high-current libraries are intended to support high power circuits while high-Q libraries support better low-power RF design. Here, Q refers to the quality factor of the inductor which is a ratio between the imaginary and real impedance of the inductor. The quality factor is a measure of losses in the inductor and higher quality factor at a given frequency indicates lower losses. In this study, the substrate effects on quality factor are studied for high-Q inductors. The inductor dimensions and corresponding inductance values are indicated in section 3.7.3. The experimentally determined quality factors for the inductors are reported in this section.

5.3.1 FLAME process parameters

The FLAME process parameters used for the inductors are very similar to the ones used for isolation structures. A starting thickness of ~250 μm is used here as well. 1-port measurement is used for RF characterization.

Table 5.3: Milling parameters for inductor study

Parameter	Fast step	Slow step
Scanner speed (mm s^{-1})	20	10
Pulse repetition freq (kHz)	30	4
Laser power (W)	1.006	0.0268
Number of passes	4	160
Fluence (J cm^{-2})	32.1	6.4
Depth Window (μm)	49.9	36.8
Removal rate (x 10^6 μm^3 s^{-1})	4.48	0.051

There are two sets of inductors in the study with the first set corresponding to 1-turn and second set corresponding to 2-turn symmetric inductors. In each set, there is one small, medium and large inductor each. Table 5.3 summarizes the FLAME process parameters used for inductors. A 2-step milling process is used with the 1st step having a square milling area of side 1.05 μm and a depth of ~100 μm. In order to have repeatability of laser processing, the same membrane area is chosen for the second step. The chosen area is a square of side 0.95 μm which can support either the combination small and medium inductor or the large inductor alone as seen in Fig. 5.18. In the layout the drawn side length of square

is 0.85 μm taking into consideration lateral expansion after isotropic XeF_2 etch. The depth of ablation for the second square is 85 μm. The XeF_2 etching parameters are the same as for isolation structures (see section 5.1.1).

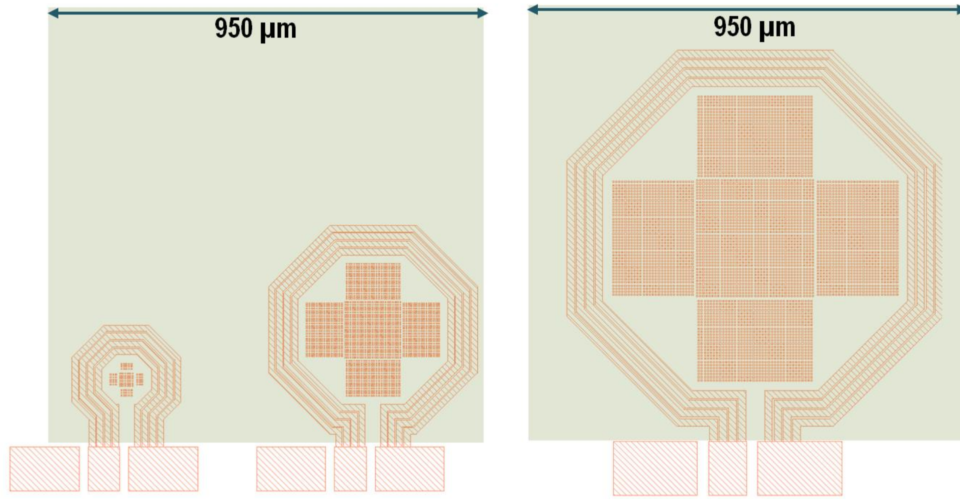


Fig. 5.18: Membrane area showing suitability for (a) Small and medium sized inductors (b) Large inductor

5.3.2 S-parameter characterization

S-parameter characterization is performed over a frequency range of 100 MHz – 110 GHz. A low IF bandwidth of 50 Hz is used to improve S-parameter measurement. The upper limit for the frequency is set by the smallest inductor in the study 1T-S with an inductance of 0.2 nH. By using 110 GHz frequency, the Q-factor response can be viewed fully on either side of the peak Q-factor. Since, a 1-port setup is used, and the parasitics are not negligible because of the chosen inductance values, pad deembedding is necessary. An open pad is used to determine the pad parasitics which is then later deembedded from the inductor measurements. The dembedding structure and corresponding circuit representation is shown in Fig. 5.19. The impedance of the inductor is calculated as:

$$Z_{ind} = \left(Y_{meas} - Y_{pad}\right)^{-1}$$

After de-embedding, the Q-factor is calculated as:

$$Q_{ind} = \frac{imag(Z_{ind})}{real(Z_{ind})}$$

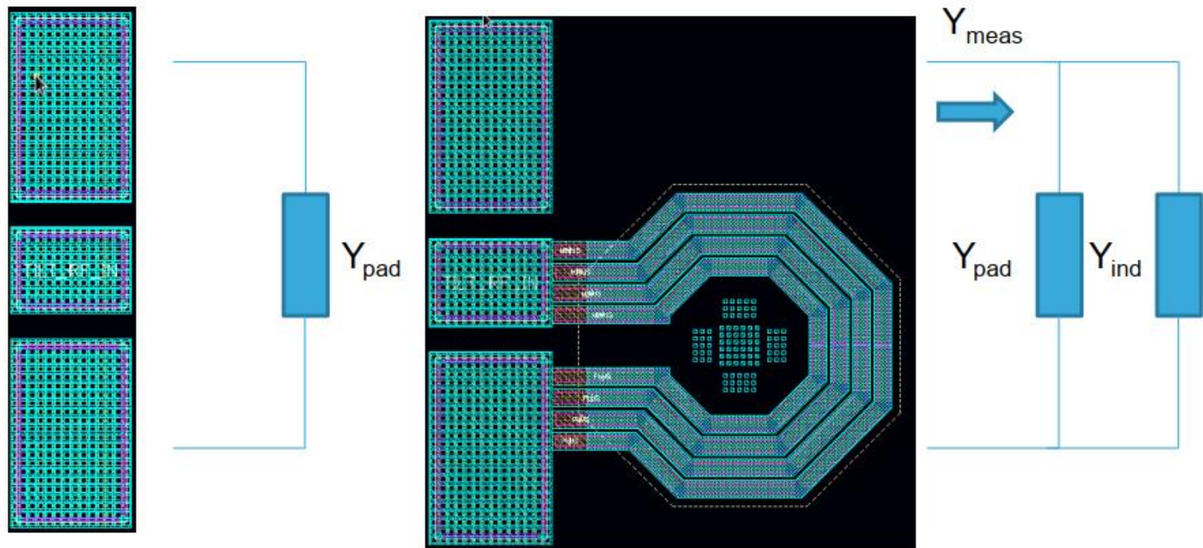


Fig. 5.19: Open pad deembedding methodology

After using the deembedding methodology as described, the Q-factor for the inductors are shown for HR substrate in Fig. 5.20 along with the self-resistance which is the real part of Z_{ind}. Because the self-resistance values of the inductors are of the order few hundreds of mΩ, the contact resistance is not negligible compared to the inductor self-resistance. For small inductors, the contact resistance can be even larger than the self-resistance. This results in a large spread in the Q-factor curves and the extracted values of the peak Q-factor.

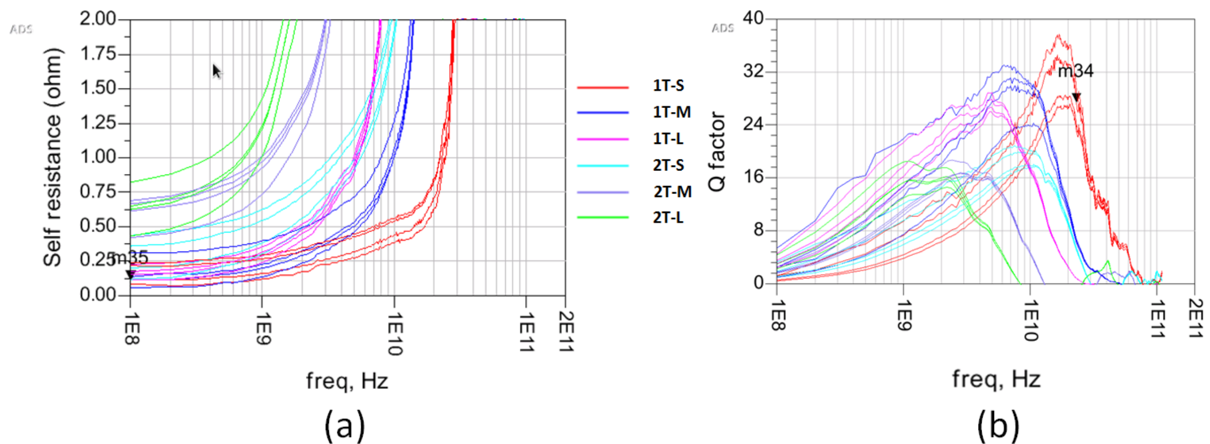


Fig. 5.20: Measured values of self-resistance and Q-factor for the inductor on 4 different dies for HR substrate

In order to remove the contribution of contact resistance from the measurement, EM simulation is performed and the self-resistance of the inductor is extracted at 100 MHz. Any measured resistance which is in excess of this value is counted as the contact resistance and subtracted from the measured value. The self-

resistance and Q-factor curves after performing this correction are shown in Fig. 5.21. The dispersion is greatly reduced after this correction. The same procedure is followed for TR substrate as well.

A small inductance of ~50 pH is also introduced by the contact which also significantly impacts the Q-factor calculation for small inductors. In Fig. 5.21, the contact inductance is not removed from Z_{ind}. For the smallest inductor 1T-S, it results in an overestimation of peak Q-factor by 40%. Hence, the contact inductance is removed from the measurement in the same way as the contact resistance using the inductance value obtained from EM simulation at 1 GHz. The EM simulation is only used to extract parameters for corrections in measurement which does not depend on the substrate type. In order to precisely perform the EM simulation, the substrate needs to be accurately calibrated over the frequency range of simulation. This is a complicated task which is beyond the scope of the thesis and only experimentally measured values are reported.

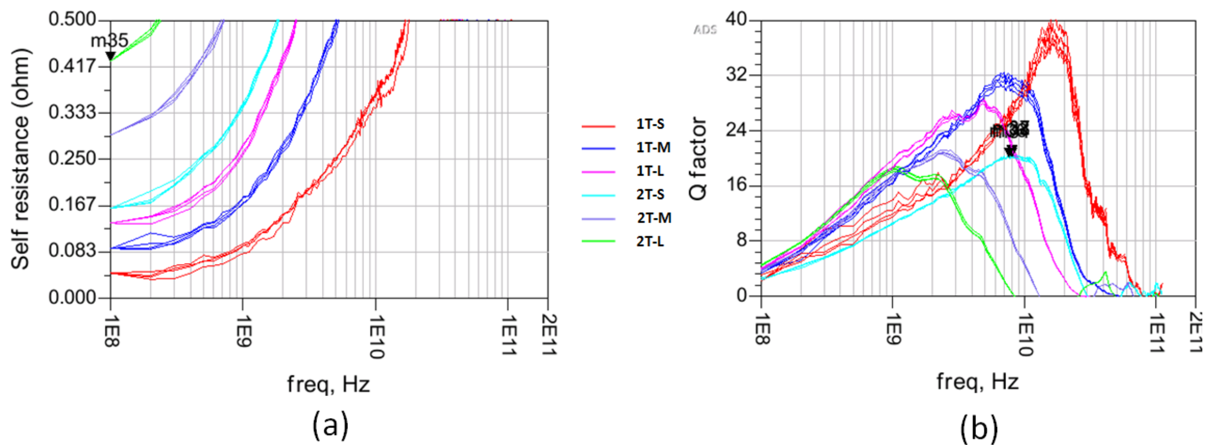


Fig. 5.21: Corrected values of self-resistance and Q-factor for the inductor on 4 different dies for HR substrate. Here, the contact inductance is not removed from Z_{ind}.

The FLAME process is performed to suspend the inductors. One suspended inductor measurement is done for each of the 6 inductors for which 4 dies (1 membrane per die) are processed for each substrate type. The correction of contact resistance and inductance is performed on the measurements after FLAME process. The results for one-turn inductors before and after substrate processing are shown in Fig. 5.22. Before substrate removal, the peak Q-factor value of the 3 inductors is ~30. There is very little difference between HR and TR substrate. For 1T-L, a small difference is visible with TR substrate having a marginally higher Q-factor than HR substrate. After substrate removal, the peak Q-factor of the inductor increases for all 3 inductors. This improvement comes from the reduction of frequency dependent series resistance. This resistance

captures losses due to different mechanisms like skin effect, proximity effect and substrate losses due to electric and magnetic coupling, which are all highly frequency dependent. The substrate related loss mechanisms are eliminated by FLAME process which reduces series resistance and hence results in improvement of quality factor. We can also note a decrease of the oxide capacitance resulting in an increase of the cut off frequency.

For 1T-S and 1T-M cases, the self-resistance does not show a monotonic response and results in unrealistic values of Q-factor. This could be due to secondary resonances in the inductor. A different approach is used to measure approximately the Q-factor of 1T-S and 1T-M cases which is described later.

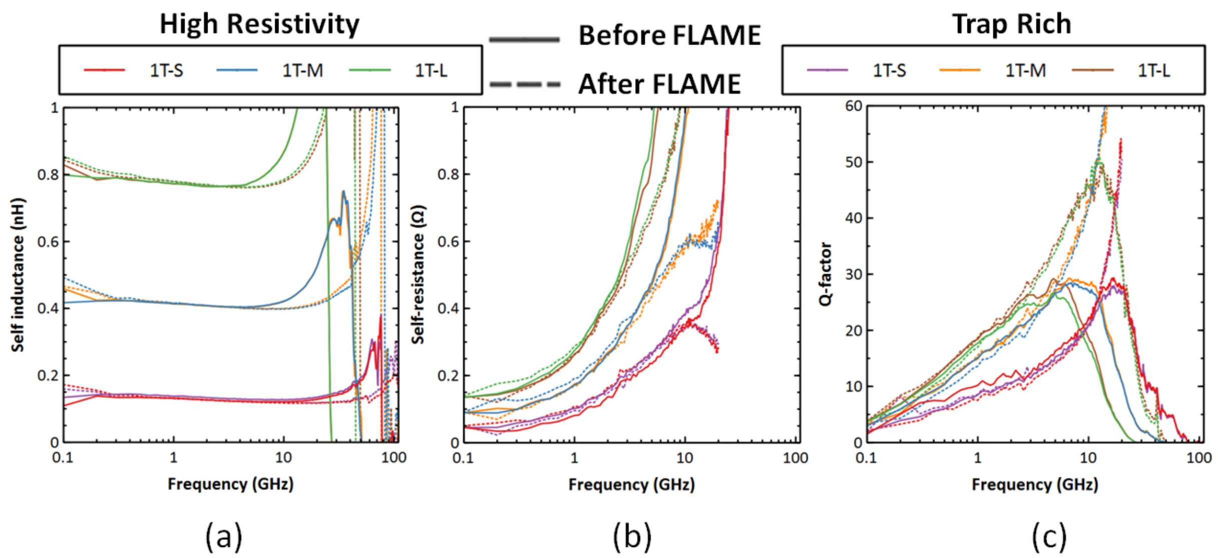


Fig. 5.22: Inductor parameters for single turn inductors before and after FLAME (a) Self-inductance (b) Self resistance (c) Q-factor

The DLM images for the processed inductors are shown in Fig. 5.22. These images do show the etching border close to the bond pad so as to check that silicon etching stops as close to the bond pad as possible and that bond pad remains intact. For all the inductors, there is no etching under the bond pad and the boundary of the membrane is very close to the bond pad. It is also to be noted that because of two step milling, the bond pad is within the area of the bigger cavity. At the end of the FLAME process, the substrate thickness under the bond pad is reduced to ~40 µm. Hence, the pad parasitics reduces and a bond pad of same thickness should be used for deembedding accuracy. However, it is assumed that the bond pad remains in the same configuration after substrate removal and the pad parasitics extracted before substrate removal is used for pad deembedding. This leads to a slight overestimation of Q-factor.

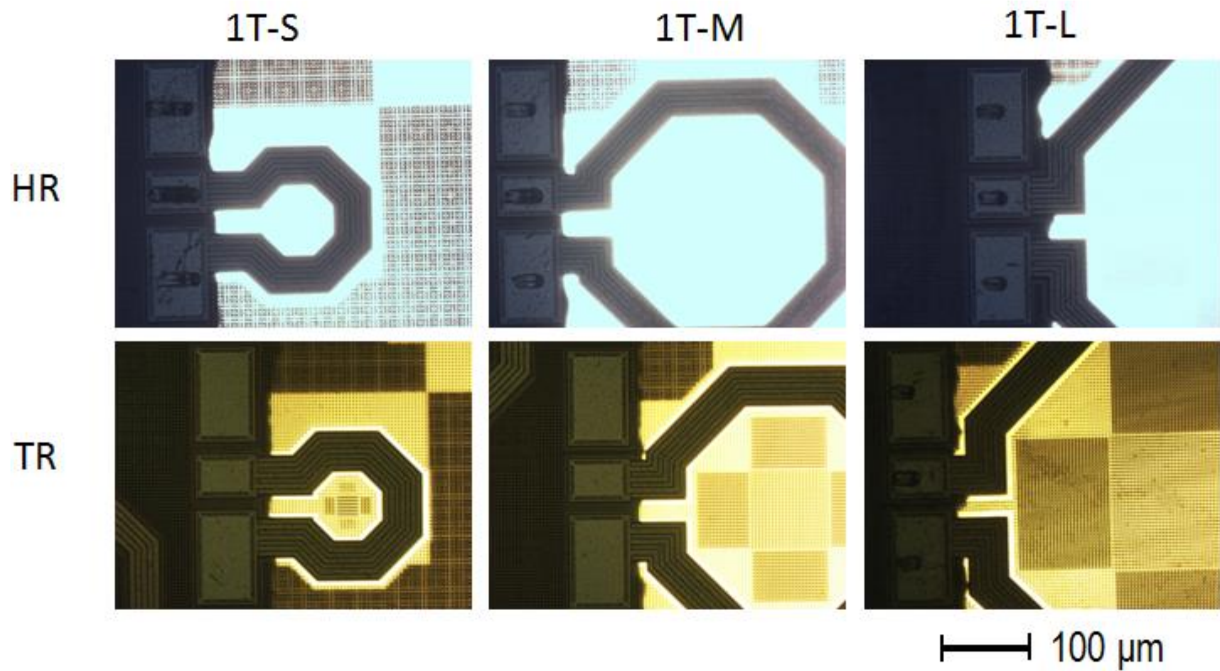


Fig. 5.23: DLM images of single turn inductors taken close to the bond pad

For the two turn inductors, the difference between HR and TR substrate is more visible. The measured parameters for 2 turn inductors are shown in Fig. 5.24. It can be seen that for 2T-M and 2T-L, TR substrate has superior Q-factor as compared to HR substrate. For 2T-S, the difference is negligible. Taking into consideration the results from both single turn and two turn cases, the difference in HR and TR substrate depends on the value of the inductance itself. For smaller inductance values (< 1 nH), the difference in HR and TR substrate is negligible. After FLAME process, the peak Q-factor values increases for all 3 cases. The inductor parameters for HR and TR substrate after FLAME process agree very well like in the case of one turn inductors. At low frequencies, the Q-factor curves after FLAME process follow the curves before inductor removal. At these frequencies, there is no difference between HR/TR both before and after FLAME. The difference in Q-factor is observed only at high frequencies.

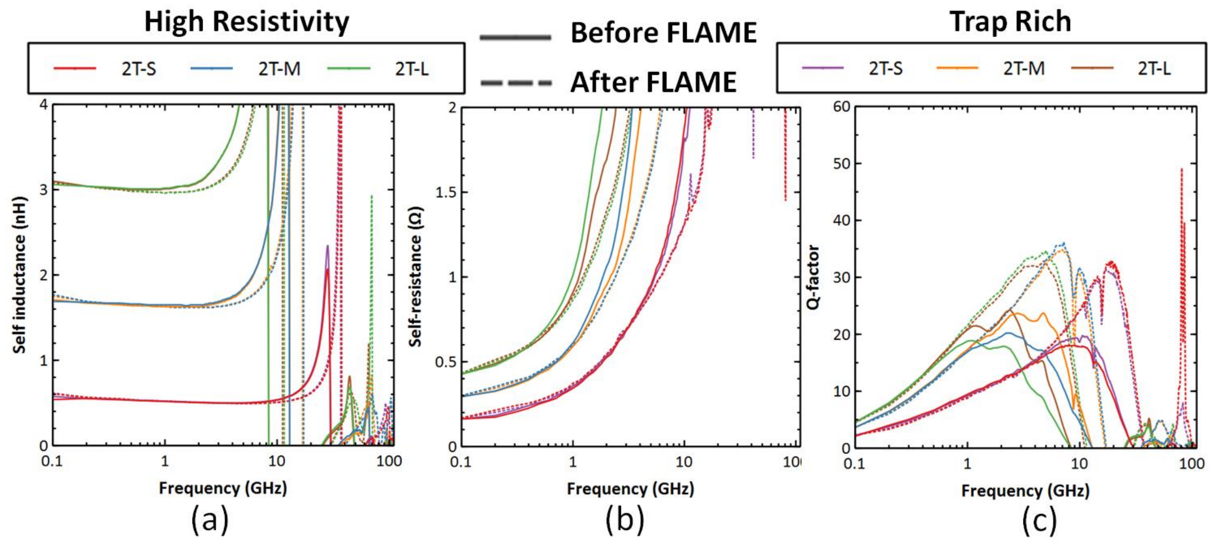


Fig. 5.24: Inductor parameters for two turn inductors before and after FLAME (a) Self-inductance (b) Self resistance (c) Q-factor

The DLM images of the membranes of two turn inductors are shown in Fig. 5.25. The membrane borders for HR and TR substrate are placed very close to the bond pad. The same substrate thickness of ~40 µm is present under the bond pad after FLAME process.

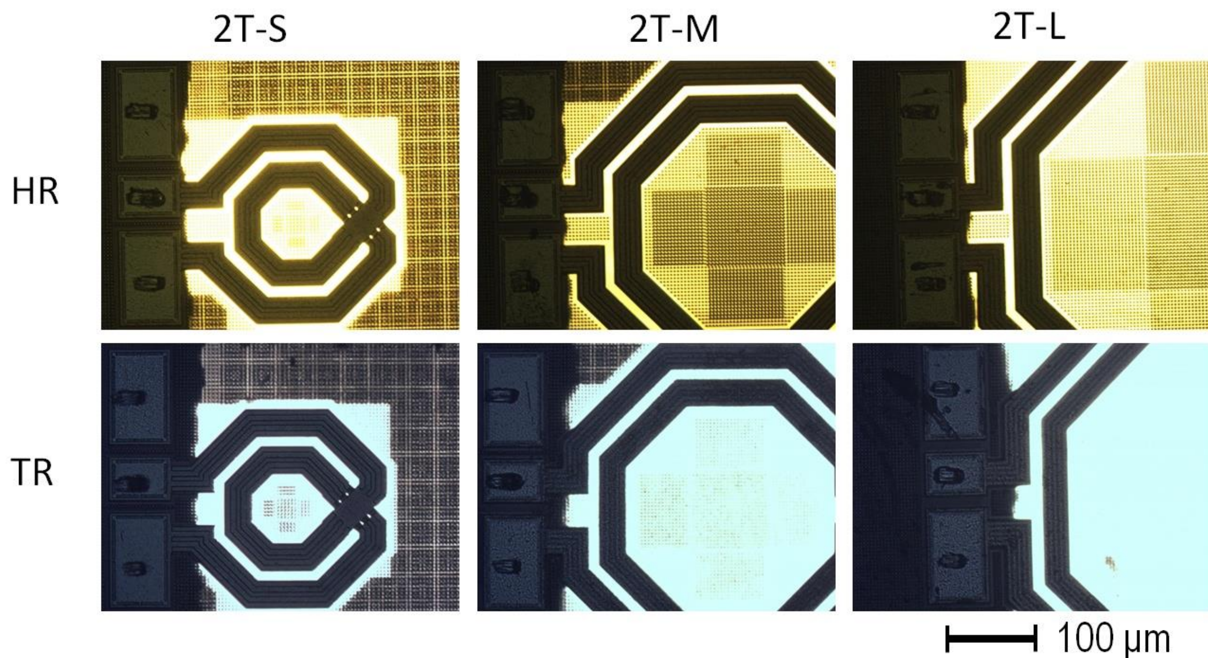


Fig. 5.25: DLM images of two turn inductors taken close to the bond pad

The results of the study on inductors are summarized in Table 5.4. The peak Q factor and the corresponding frequency and self-resistance are reported here. It can be seen that for the single turn case, there is no noticeable difference between

HR and TR substrate before FLAME process. A noticeable difference in Q-factor is seen for inductance values > 1 nH. This suggests that the impact of substrate on the Q-factor is larger for higher values of inductances. The same trend is observed for frequency of peak Q-factor.

After FLAME process, there is no difference in inductor configuration between the two substrate types as substrate is completely etched. Accordingly, the Q-factor values are nearly same for the two substrate types after FLAME process with a marginally high value of Q-factor for HR substrate. This could be as a result of reduced pad parasitics for HR as compared to TR substrate after FLAME process because of reduced substrate thickness (~40 μm) under bond pad. Because the bond pad (with substrate thickness~250 μm) before FLAME process is used for deembedding, the final result could result in a marginal overestimation of Q-factor for HR substrate. It is also to be noted that at peak value of Q-factor, the series resistance is higher than before FLAME process. This is because both inductive reactance and series resistance together determine the Q-factor curve. With increase in frequency, both parameters increase and at a certain frequency the Q-factor reaches the maximum. When the resistive losses are reduced, the frequency value corresponding to peak Q-factor is shifted to higher frequencies. Hence, when operating at near peak Q-factor conditions before and after FLAME process, the series resistance values are different.

The frequency corresponding to peak Q-factor also increases after substrate removal. It is also to be noted in Fig. 5.24 that at low frequencies, there is no difference between HR/TR substrates before/after FLAME. The benefits of removing the substrate are seen only at higher frequencies.

Table 5.4: Summary of inductor measurements

Ind	Before FLAME						After FLAME					
	Qp HR	**Qp TR**	**fp HR (GHz)**	**fp TR (GHz)**	**Rs HR (Ω)**	**Rs TR (Ω)**	**Qp HR**	**Qp TR**	**fp HR (GHz)**	**f TR (GHz)**	**Rs HR (Ω)**	**Rs TR (Ω)**
1T-S	29.4	28.0	17	17	0.46	0.50	x	x	x	x	x	x
1T-M	28.5	29.3	7	6.6	0.63	0.58	x	x	x	x	x	x
1T-L	26.6	29.2	5	4.8	0.92	0.80	51.0	49.7	12.6	13.6	1.25	1.39
2T-S	18.1	19.8	8	10.6	1.43	1.83	33.0	31.3	19.2	17.4	2.18	2.01
2T-M	20.2	23.7	2.4	4.8	1.24	2.29	36.2	34.9	7.2	7.00	2.31	2.34
2T-L	18.9	24.3	1.1	2.4	1.11	1.96	34.7	32.9	5.00	4.80	3.14	3.18

Qp – Peak Q factor, fp – Frequency corresponding to Qp, Rs – Series resistance @ fp

In order to study the Q-factor of the single turn small and medium sized inductors, a different approach is used where pad parasitics can be partially

removed. We observed that the open pad deembedding methodology does not work well for small inductors and as a result the extracted parameters are erroneous. Ideally, if silicon bond pad is completely removed, deembedding will not be necessary. But silicon is required under bond pad for mechanical support of bond pads during probe landing. However, partial etching of silicon is possible without compromising the mechanical integrity of the circuit during probing. The partial etching of silicon close to the bond pads realized on TR substrate using FLAME process is shown in Fig. 5.26. Case BP1 represents the standard case where silicon is present over full area of the bond pad. In cases BP2 and BP3, approximated 25% and 50% of the bond pad area does not have handler silicon.

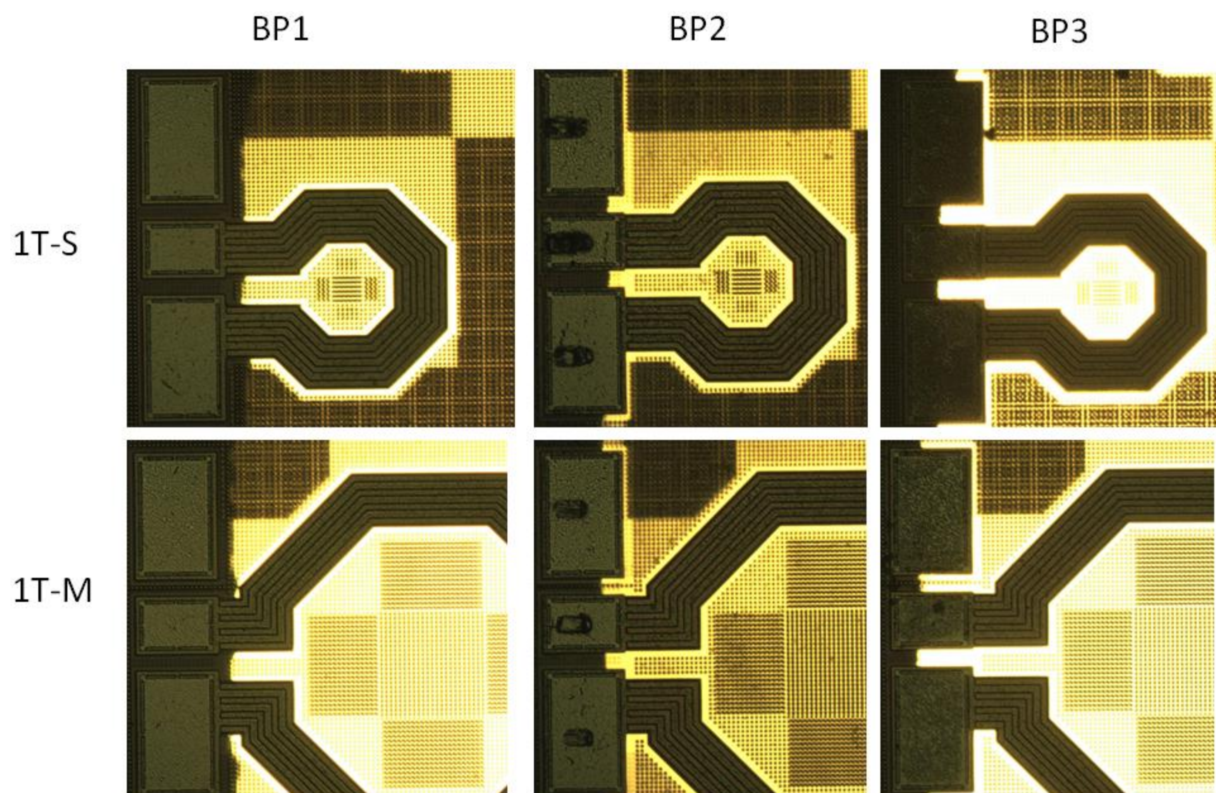


Fig. 5.26: DLM images of single turn inductors 1T-S and 1T-M with different bond pad configurations

The S-parameter characterization of TR substrate after FLAME process with different bond pad configurations is summarized in Fig. 5.27. In these measurements, only contact resistance and inductance are corrected and no deembedding is performed. For the 1T-M inductor, the self resistance increases monotonically as expected and for bond pad configuration BP3 a peak Q-factor of ~70 is observed at ~24 GHz. The actual Q-factor of the suspended 1T-M inductor would be larger than this value as some pad parasitics still exist. However, in the

case of 1T-S inductor, the self-resistance shows the same behaviour as with regular deembedding procedure. This needs further investigation.

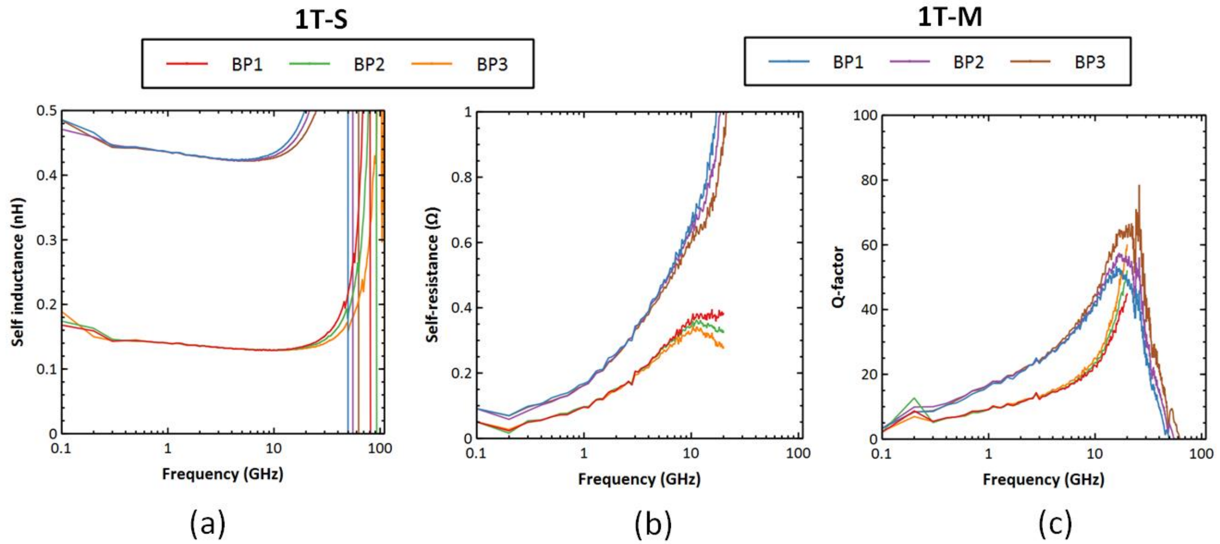


Fig. 5.27: Extracted parameters of inductors on partially removed TR pad without deembedding

In summary, the peak Q-factor of all the inductors and the corresponding frequency increases by substrate removal. The peak Q-factor increases by a factor of 1.3 – 1.7 for TR substrate and 1.8 – 1.9 for HR substrate. For small inductors < 0.5 nH, the Q-factor increases by a factor > 2. However, reliable measurement of small inductors is challenging because the inductor impedance values are comparable to the pad parasitics and the contact introduces additional resistance and inductance.

5.3.3 Modelling of inductors on membranes

Small signal measurements discussed so far have clearly indicated the benefits of removing substrate for improvement of Q-factor even in low loss substrates like HR-SOI and TR-SOI. For circuit design, it would be beneficial to have a SPICE modelling of suspended inductors. For lossy substrates, pi-model shown in Fig. 5.28 is commonly used to describe the substrate loss mechanisms [3]–[5]. Based on specific design cases, in other works, the pi-model has been extended to accommodate skin and current crowding effects [6], [7]. Using pi-model, the behaviour of inductors in lossy substrates has been well understood. In such substrates, the displacement and eddy currents flowing in the substrate limit the Q-factor of inductor.

A comprehensive study of inductors on nearly lossless Silicon-on-Sapphire substrates has been reported by Kuhn et al. [8]. We use this as the basis for

modelling of inductors on SOI membranes. When substrate losses are mitigated, inductor metal losses become the dominant loss source. The inductor can be represented by a simple 3-elements model as shown in Fig. 5.29a. In this model, L_s represents the inductance and C_p represents the capacitive element responsible for self resonance. C_p is the capacitance arising from turn to turn coupling in multi-turn inductors. These two elements are frequency independent. The self-resistance R_s is frequency dependent which captures the different loss mechanisms which are listed below:

1. Current crowding effect: This effect results in change of current density within a specific turn of an inductor caused by magnetic field due to current flowing in a neighbouring turn [9]. The net result is the reduction of effective width of the metal trace. This is the dominant loss mechanism at frequency < 2 GHz.

2. Skin effect: This effect is the increase in current density near the edges of the conductor. This is because the flow of RF current in a conductor which creates an internal magnetic field and affects the current distribution. The net result is the reduction of effective thickness of the metal trace. This is the dominant loss mechanism at frequency > 2 GHz.

   Both of these effects are additive which makes it difficult to distinguish the contribution of each separately [10].

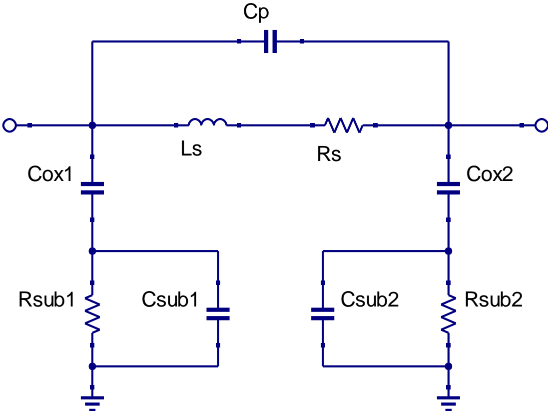


Fig. 5.28: Pi-Model of integrated inductor on silicon technology

In the simplest form, the inductor on membrane can be modelled using three components as shown in Fig. 5.29a. However, an expression of the frequency dependent series resistance is to be calculated in order to model the effects that have already been described. In order to have a frequency independent model, Kuhn et al. proposed the 6-elements model which is shown in Fig. 5.29b. The 6-

element model consists of a mutual inductance with resistors modelling the different frequency dependent metal losses. Using this model as the basis, we have used circuit transformation of the 6-elements model proposed in [11] to obtain a modified 6-elements model which consists of 2 inductors, 1 capacitor and 3 resistors. We take the first Cauer form of the boxed part in Fig. 5.29b as proposed in [11]. We use the 2nd order Cauer form (N = 2) shown in Fig. 5.29c for our modelling purposes.

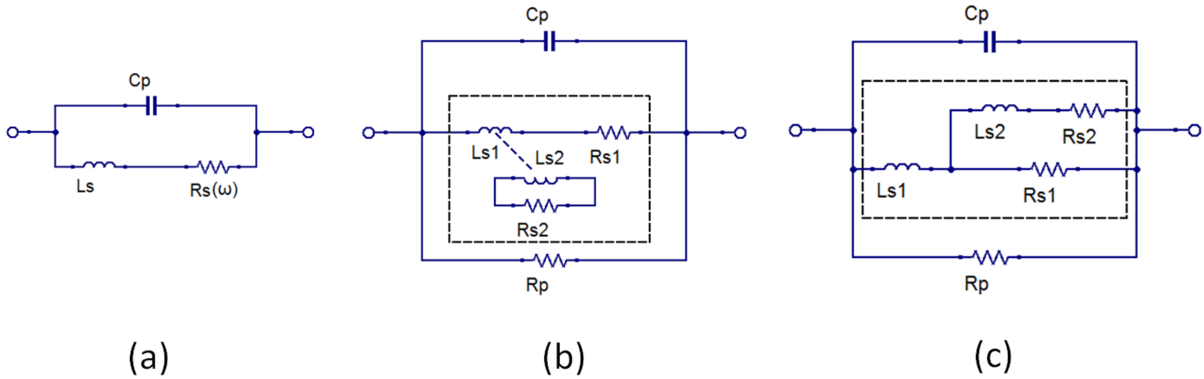


Fig. 5.29: Models of inductors on nearly lossless substrates (a) 3-elements frequency dependent model (b) 6-elements model (c) Modified 6-elements model

The procedure for finding the elements of the model is outlined below:

1. Compute the complex impedance of the inductor from the S-parameter characterization to obtain $Z_{ind}(\omega)$.
2. Calculate the inductance L_{ind} as a function of frequency using $L_{ind} = \frac{imag(Z_{ind}(\omega))}{\omega}$
3. L_{s1} is set to the value of inductance extracted at 1 GHz. $L_{s1} = L_{ind}(1\ GHz)$.
4. With the increase in frequency, the value of L_{ind} drops to zero. This corresponds to the frequency of self-resonance ω_r. Using this value of frequency C_p is calculated as $C_p = \frac{1}{L_{s1}\omega_r^2}$.
5. The value of resistance of inductor R_{ind} is extracted using $R_{ind} = real(Zind(\omega))$. At low frequency, the dominant elements are Rs_1 and Rs_2. Since they appear in parallel, a condition is set in the model: $R_{ind}(100\ MHz) = Rs_1||Rs_2$

6. Finally, the parameters R_p, R_{s1}, L_{s2} are varied to obtain the best fit. Rs_2 is not an independent variable because of the constraint imposed in step 5. Additionally, $L_{s1}$is also adjusted for each value of L_{s2} used during fitting process. This is because L_{s2} which models frequency dependent resistance also adds an inductive component of reactance which increases the total inductance. L_{s1} is changed such that the value of inductance calculated from the model is equal to $L_{ind}(1\ GHz)$.

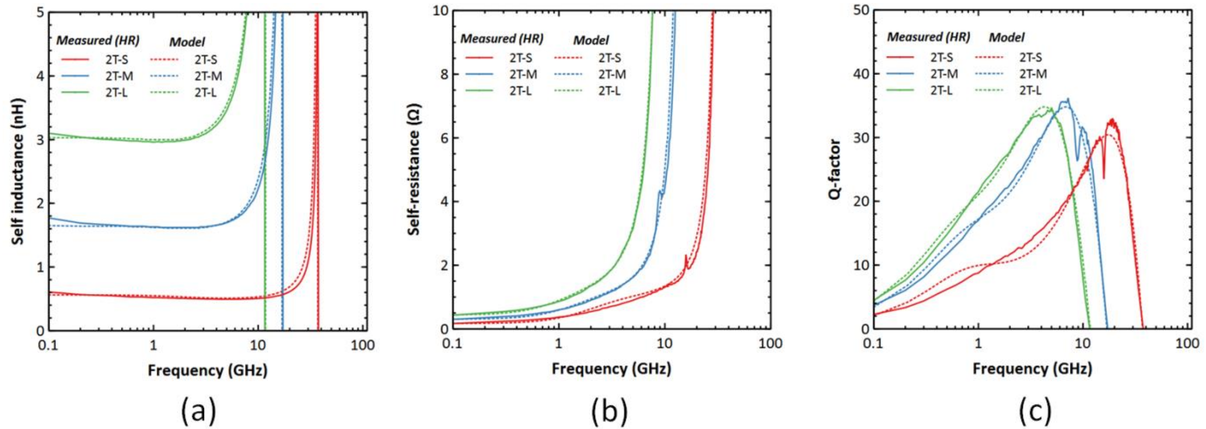


Fig. 5.30: Comparison of measured vs. modelled inductor parameters

The results of modelling are depicted in Fig. 5.30. The model parameters extracted for the modified 6-elements model is tabulated in Table 5.5. It can be seen that the model has good agreement with the measured data. Both the peak value of Q-factor and frequency of the peak can be modelled quite well using the modified 6-elements model. There is a small deviation between measured and modelled Q-factor at frequencies less than the peak Q-factor. This is because at these frequencies, the Q-factor is very sensitive to the value of series resistance. The model gives values of series resistance slightly differently from the measured data. This difference also reflects in the Q-factor curves. This could potentially be improved by using a higher order model with more elements.

Table 5.5: Modified 6-element model parameters for 2-turn inductors

Ind	L_{ind} @ 1 GHz (nH)	R_{ind} @ 100 MHz (Ω)	R_{s1} (Ω)	R_{s2} (Ω)	L_{s1} (nH)	L_{s2} (nH)	R_p (Ω)	C_p (fF)
2T-S	0.61	0.17	1.1	0.20	0.49	0.1	10000	37.5
2T-L	1.81	0.30	1.2	0.40	1.55	0.17	10000	57
2T-M	3.25	0.43	1.4	0.62	2.9	0.28	10000	63.5

Some important observations can be made by looking at the values of the model parameters in Table 5.5. Both R_{s1} and R_{s2} increase with the increase in size (diameter) of the inductor. This is because the resistance depends on the length of the conductor which increases with increased diameter of inductor. L_{s1} and L_{s2} increase similarly due to length dependence of inductance. It is to be noted that higher L_{s2} means that it blocks current flowing through R_{s2} starting from lower frequencies. If the series resistance curve is observed in Fig. 5.30b, it can be seen that resistance increases slowly initially and starts rising rapidly beyond a certain frequency. This frequency is higher for inductors of smaller size. In the model, the frequency of transition is captured by L_{s1}. Higher L_{s1} means that the knee point occurs at smaller frequencies. R_p is used for fine adjustment of peak Q-factor and corresponding frequency. We found that a value of 10 kΩ is suitable for all inductors. Finally, C_p models the frequency of self resonance where the inductor value and the Q-factor drop to zero. Smaller inductors means smaller coupling between turns. Hence, for small size inductors, lower C_p shifts the self resonant frequency to higher values. For the smallest inductor, the self resonant frequency is ~37.5 GHz. Thus the model is able to predict inductor behaviour even at large frequencies.

5.4 Low noise amplifier (LNA)

The LNA design is described in section 3.7.4. The motivation behind processing of LNA is study of impact of spiral inductors on linearity and noise figure of LNA. Spiral inductors are key components in LNA design used for providing input and output matching. The inductor on the input side connected to the gate in a common source configuration is the most interesting because its contribution to the noise figure is the highest of all inductors. The Q-factor for this inductor is separately characterized in order to correlate the LNA operation

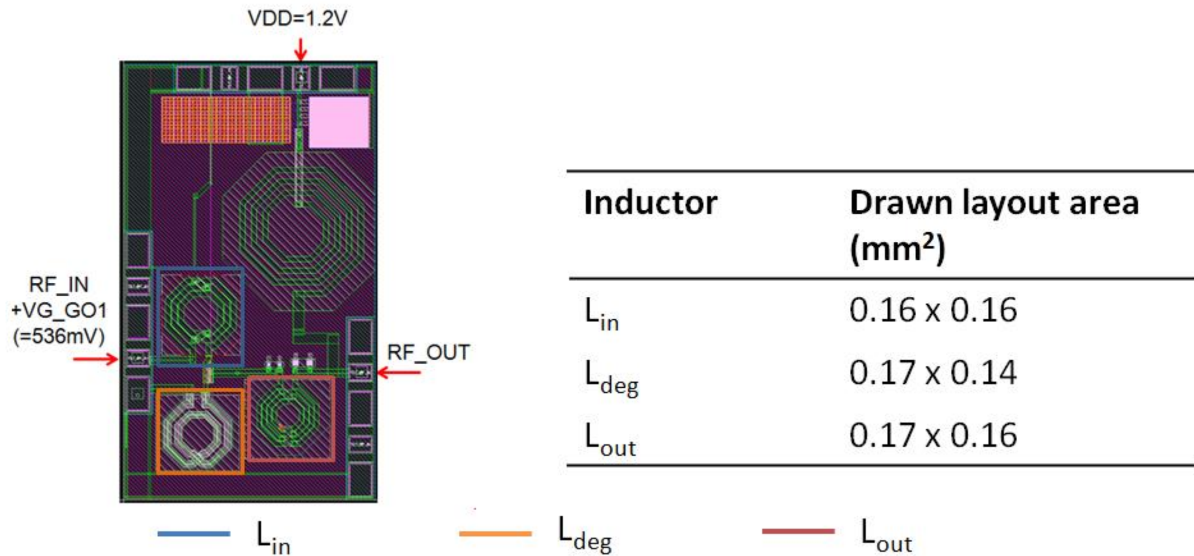
5.4.1 FLAME process parameters

The same parameters used for isolation structures are applicable to LNA processing as well. The LNA layout along with drawn layout area for laser processing is shown in Fig. 5.31. The membrane dimensions of the 3 inductors are nearly same. Three different dies of HR type are processed to obtain LNA with 3 cases of suspended inductors as follows:

Case 1: L_{deg} only

Case 2: L_{deg} and L_{in}

Case 3: L_{deg}, L_{in} and L_{out}



Inductor	Drawn layout area (mm^2)
L_{in}	0.16 x 0.16
L_{deg}	0.17 x 0.14
L_{out}	0.17 x 0.16

Fig. 5.31: LNA layout showing the different inductors and corresponding drawn layout area for laser processing

The DLM images of the suspended inductors for the 3 cases are shown in Fig. 5.32. In case 3, the two cavities for L_{deg} and L_{out} become joined after XeF_2 etching because of their proximity.

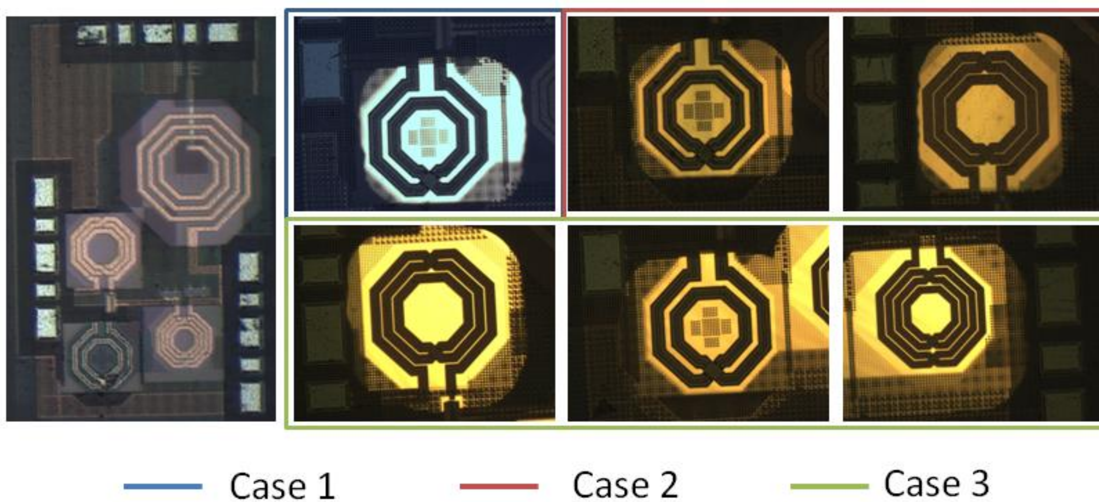


Fig. 5.32: DLM images of suspended inductors for 3 test cases

5.4.2 Noise figure measurement

Small signal S-parameters along with noise figure is measured for the LNA on the same bench at an input source inductance of 50 Ω with bias conditions as shown in Fig. 5.31. Three dies are characterized for HR and TR substrate prior to

FLAME process. The measured noise figure and gain are plotted in Fig. 5.33. The noise measurement shows very little scatter in data across the 3 devices for both HR and TR substrate. The minimum NF is obtained at ~6 GHz with TR substrate showing 0.1 dB lower noise figure as compared to HR substrate. Since the difference in noise figures of HR and TR substrate is very small, care has been taken to make sure that there are no errors due to calibration. A reference device is measured each time after calibration to ensure that all measurements are accurately calibrated.

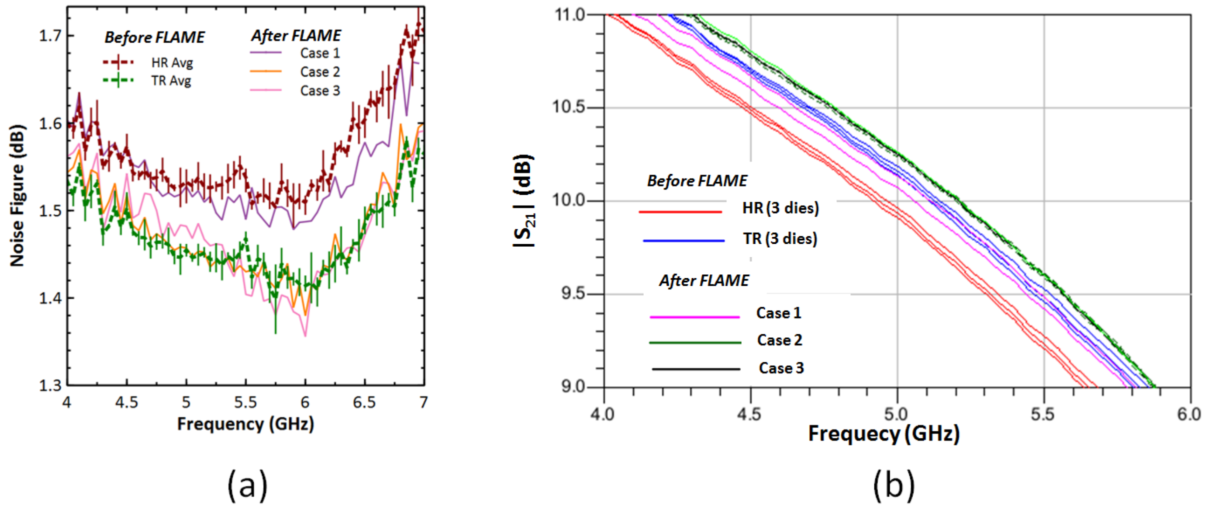


Fig. 5.33: Comparison of (a) Noise figure with error bars representing min-max values (b) Gain before and after FLAME process

Case 1 corresponds to substrate removal of source degeneration inductor (L_{deg}). There is only a small improvement in NF as compared with HR substrate which increases at higher frequencies. This effect is probably related to the frequency dependence of Q-factor with higher improvement in Q-factor at higher frequencies. In case 2, the gate inductor (L_{in}) is additionally suspended along with Ldeg. The NF behaviour of case 2 is well matched with TR substrate with ~0.1 dB improvement in gain. With all 3 inductors suspended, the NF and gain values are identical as case 2. At ~6 GHz, there is a marginal improvement of ~0.03 dB and ~0.05 dB for cases 2 and 3 respectively. These results indicate the importance of input inductor for the noise behaviour of the LNA. This inductor is characterized separately which is presented in the next section.

5.4.3 Standalone input inductor measurement

In order to study the change in input inductor characteristics after substrate removal, a standalone version of this inductor is tested with a one port configuration. The deembedding methodology is the same as explained in section

5.3.2. However, since the EM simulation data for this inductor is not accessible, the contact inductance and resistance correction is not performed as before. However, the spread in measured resistance and inductance at 100 MHz is negligible and hence qualitative comparisons are valid with possible variations in actual and extracted parameters. The results are shown in Fig. 5.34.

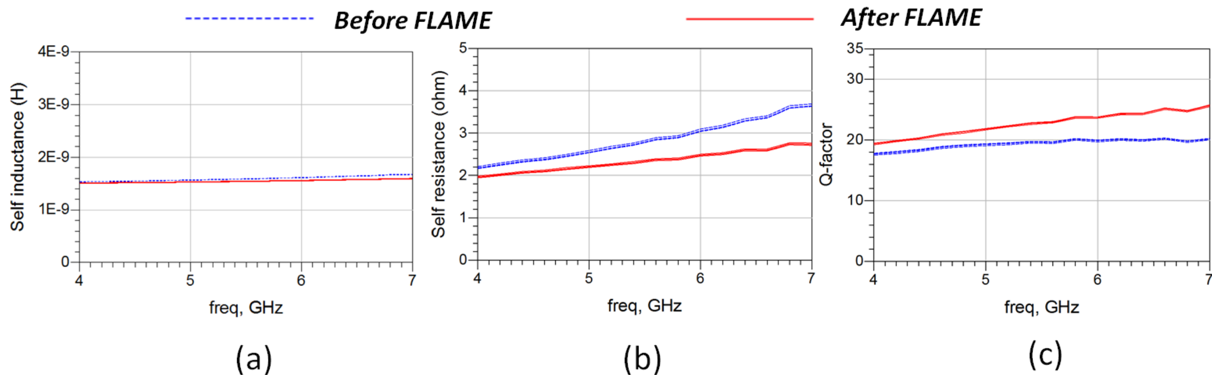


Fig. 5.34: Extracted parameters of input inductor from single port S-parameter characterization for HR substrate (a) Self Inductance (b) Self resistance (c) Q-factor

The measurements show that over the frequency range of operation of the LNA, the quality factor of the inductor increases. The improvement in quality factor is higher at higher frequencies. The $\Delta Q = Q_{after} - Q_{before}$ at frequencies of 4.9 GHz and 5.9 GHz are 2.3 and 3.7 respectively. This small improvement in Q-factor helps reduce the overall resistance on the input side which gives better noise performance. A possible secondary effect of substrate removal is the elimination of coupling of substrate noise to the inductor through the parasitic BOX capacitance [12].

In addition, the presented impedance by the gate inductance to the LNA transistor has changed between the 2 processes: before and after FLAME and this impacts also LNA noise figure. In the previous section, it has been shown that inductances peak Q factor can be significantly improved by FLAME and to better quantify substrate removal effect on LNA noise figure a redesign with inductor optimization once substrate is removed to match the transistor near Q peak should demonstrate higher improvement.

5.4.4 Linearity measurements

The linearity measurement comprises of determination of output compression point (P_{1dB}) and the third order intercept point referred to the input (IIP3). Both

the measurements are realized on the same setup as shown in Fig. 5.35. As in the case of NF measurements, 3 HR and TR devices are measured prior to substrate removal. Linearity measurements are presented for 3 different frequencies of 4.9, 5.4 and 5.9 GHz.

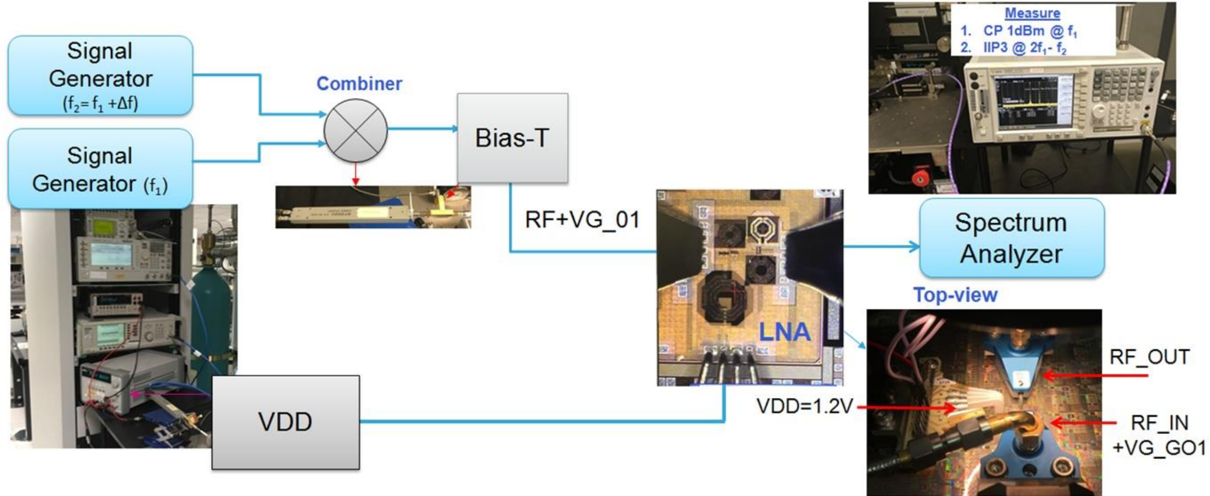


Fig. 5.35: Measurement setup for P_{1dB} and IIP3

The summary of IIP3 measurements is presented in Table 5.6. There is a marginal improvement of IIP3 for cases 2 and 3 when comparing the average values. However, when comparing the maximum measured values before substrate processing, the improvement seems very negligible.

Table 5.6: Summary of LNA IIP3 measurements

	Wafer Type	**Value**		
Frequency (GHz)		4.9	5.4	5.9
Third order intercept point (IIP3) (dBm)				
Simulated		6.7	9.2	11.7
Measured	HR-SOI (avg)	6.84	9.19	11.65
	TR-SOI (avg)	6.74	8.96	11.45
	HR-SOI (max)	6.93	9.35	11.70
	TR-SOI (max)	6.95	9.21	11.82
	Case 2	7.12	9.30	12.01
	Case 3	6.86	9.46	11.51

The summary of P_{1dB} measurements is presented in Table 5.7. There is a marginal reduction of P_{1dB} for cases 2 and 3 when comparing the average values. However, once again when comparing the minimum measured values, the change

in P_{1dB} values are negligible before and after substrate processing. Hence, the substrate removal does not affect significantly the linearity figures of merit.

Table 5.7: Summary of LNA P_{1dB} measurements

	Wafer Type	Value		
Frequency (GHz)		4.9	5.4	5.9
1-dB compression point (P_{1dB}) (dBm)				
Measured	HR-SOI (avg)	-5.21	-5.07	-4.45
	TR-SOI (avg)	-5.21	-5.18	-4.46
	HR-SOI (min)	-5.25	-5.15	-4.57
	TR-SOI (min)	-5.25	-5.25	-4.51
	Case 2	-5.75	-5.34	-5.12
	Case 3	-5.78	-5.41	-5.19

Considering all the measured data in LNA, the noise figure performance after substrate removal of inductors on HR substrate is comparable to TR substrate. The linearity performance is overall retained with possibly a small improvement in IIP3 and small reduction in P_{1dB}.

Concluding Remarks

In this chapter, some of the components of the frontend module have been analyzed. The FLAME process has been applied to create SOI membranes without any handler substrate under the device operating area. RF characterization has been performed to study the improvements seen with improved substrate conditions. Isolation structures have demonstrated that the difference in substrate coupling is significant at lower frequencies and non-negligible at higher frequencies. Thus, device improvements can be expected over a wide frequency range. The study of RF switch was the first demonstration of active device improvement by substrate removal. Both losses at small signal and harmonic distortion at higher signal power reduce with substrate removal. Q-factor measurements on RF inductors also have been studied demonstrating the performance boost of passive devices on SOI technology. Following the study on inductors, the impact of improved Q-factor inductors on LNA performance has been studied. The input inductor has been shown to play an important role in determining LNA performance. Noise figure and gain improvements are seen while linearity remains mostly unchanged. All these results show impact of

substrate on device performance and how it can be enhanced by local substrate removal. As the BOX thickness reduces with coming generations of SOI, substrate coupling challenges are expected to increase. Thus, substrate removal could provide the potential boost for highly demanding applications.

References

[1] C.-J. Wei *et al.*, “Analysis and modeling on linearity for multi-throw TX/RX switches,” in *2009 3rd IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, Beijing, China, 2009, pp. 5–8.

[2] D. Prikhodko *et al.*, “Enhanced linearity technique for multithrow TX/RX switches.,” in *2008 IEEE Radio and Wireless Symposium*, Orlando, FL, USA, 2008, pp. 403–406.

[3] J. N. Burghartz, M. Soyuer, and K. A. Jenkins, “Microwave inductors and capacitors in standard multilevel interconnect silicon technology,” *IEEE Trans. Microw. Theory Tech.*, vol. 44, no. 1, pp. 100–104, 1996.

[4] J. R. Long and M. A. Copeland, “The modeling, characterization, and design of monolithic inductors for silicon RF IC’s,” *IEEE J. Solid-State Circuits*, vol. 32, no. 3, pp. 357–369, 1997.

[5] C. P. Yue and S. S. Wong, “Physical modeling of spiral inductors on silicon,” *IEEE Trans. Electron Devices*, vol. 47, no. 3, pp. 560–568, Mar. 2000.

[6] P. Arcioni, R. Castello, L. Perregrini, E. Sacchi, and F. Svelto, “An improved lumped-element equivalent circuit for on silicon integrated inductors,” in *Proceedings RAWCON 98. 1998 IEEE Radio and Wireless Conference (Cat. No. 98EX194)*, 1998, pp. 301–304.

[7] J. Craninckx and M. S. Steyaert, “A 1.8-GHz low-phase-noise CMOS VCO using optimized hollow spiral inductors,” *IEEE J. Solid-State Circuits*, vol. 32, no. 5, pp. 736–744, 1997.

[8] W. B. Kuhn, Xin He, and M. Mojarradi, “Modeling spiral inductors in SOS processes,” *IEEE Trans. Electron Devices*, vol. 51, no. 5, pp. 677–683, May 2004.

[9] W. B. Kuhn and N. M. Ibrahim, “Analysis of current crowding effects in multiturn spiral inductors,” *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 1, pp. 31–38, 2001.

[10] J. N. Burghartz and B. Rejaei, “On the design of RF spiral inductors on silicon,” *IEEE Trans. Electron Devices*, vol. 50, no. 3, pp. 718–729, Mar. 2003.

[11] A. Nieuwoudt and Y. Massoud, “Analytical wide-band modeling of high frequency resistance in integrated spiral inductors,” *Analog Integr. Circuits Signal Process.*, vol. 50, no. 2, pp. 133–136, Dec. 2006.

[12] A. L. L. Pun, T. Yeung, J. Lau, J. R. Clement, and D. K. Su, “Substrate noise coupling through planar spiral inductor,” *IEEE J. Solid-State Circuits*, vol. 33, no. 6, pp. 877–884, Jun. 1998.

Conclusion and perspectives

The FLAME process for suspending SOI circuits has been successfully implemented. From the available options for laser processing, a systematic study was performed to understand the effect of different process parameters on milling quality. By studying scribing of trenches, it has been shown that a low repetition rate combined with high fluence is more suitable for high removal rate milling. Following this, 2D milling has been described where additional phenomena like generation of particles and screening of laser radiation have been discussed. Shallow cavities (depth < 10 µm) of small roughness and good morphology have been reported. Milling of deep cavities (depth > 100 µm) showed certain limitations in terms of increased roughness and coarse morphology. The reasons for reduced quality of milling have been discussed. It has been shown that morphology can be improved by maintaining the focus slightly inside the sample. The morphology of the obtained sidewalls is smooth which is well suited for our application. Finally, the FLAME process is discussed which describes the selective etching of silicon in XeF_2 after milling to obtain membranes on SOI.

The FLAME process has been used to suspend different functions of the frontend module and RF characterizations have been performed. The RF measurements on isolation structures revealed the importance of substrate coupling. It was shown that substrate coupling is higher at MHz frequencies and remains significant even at frequencies as high as 5 GHz. On SP9T switch, the removal of substrate showed improved performance under both small and large signal conditions. For HR-SOI substrate, the improvements obtained are: Insertion loss (ΔS_{21}): 0.4 dB, Return loss (ΔS_{11}): 4 dB, 2nd harmonic: 22.7 dB, 3rd harmonic: 5.6 dB. For TR-SOI substrate, the improvements obtained were: Insertion loss (ΔS_{21}) (dB): 0.28 dB, Return loss (ΔS_{11}): 4 dB, 2nd harmonic ($\Delta H2$): 7.8 dB, 3rd harmonic ($\Delta H3$): -2.2 dB. Substrate removal has been subsequently used on inductors for the improvement of Q-factor. Both single and two turn inductors have been characterized with inductance values ranging from 0.2 – 3.3 nH. For large inductors ($L > 0.5$ nH), the increase in Q-factor that is observed is 1.8 – 1.9 for HR-SOI and 1.3 – 1.7 for TR-SOI substrate. For smaller inductors ($L < 0.5$ nH), the Q-factor improvement is more than 2. Finally, inductor membranes have been used to study the impact of LNA inductors on Q-factor. The input inductor has been shown to be most important for improvement of noise figure. By substrate removal of input inductor on HR-SOI substrate, noise figure of LNA is reduced by ~0.1 dB. This noise figure is comparable to that obtained on TR-SOI

without any substrate removal. There is also a small improvement in gain and IIP_3 along with a small reduction in P_{1dB}.

These results suggest that there can be significant benefits by use of substrate engineering on SOI wafers. Substrate removal not only enables direct performance enhancement of full circuits like switches but also paves way for improvement of substrate modelling. Substrate modelling of SOI substrates is a challenging task for RF design engineers. By making use of FLAME process, different designs can be characterized both with and without substrate thereby facilitating the development of substrate models. Additionally, the FLAME process provides the flexibility of hierarchical studies of substrate removal. We demonstrated this possibility by studying the impact of substrate removal of LNA inductors. In this study, the impact of substrate removal of different inductors in the LNA design was systematically performed. The relative importance of different inductors for overall LNA performance was demonstrated.

This leads us to the first perspective of the work: Development of physical simulation methodology of RF switch to be able to model harmonic distortion. A strategy has been developed for modelling switch by using mixed mode simulation that couples Poisson and drift-diffusion transport equations to Spice-like circuit analysis. In the mixed mode simulation, it is possible to connect lumped elements (Eg. resistors) with physical structures (Eg. transistors) and obtain solutions using drift-diffusion equations on physical structures. After the simulations are appropriately calibrated to the measured results, more insights into the substrate non-linearity can be obtained. In this direction, some progress has already been made with transistor structures for physical simulations realized using Silvaco Athena process simulator. The physical transistor model calibration is under progress. It is worth noting that the effect of deep levels traps associated to the buried polysilicon layer of HR substrate was properly taken into account for the simulation of harmonics generation in coplanar waveguides [1].

The second perspective is the study of LNA noise figure by using an inductor of much higher quality factor. While it is shown that noise figure improves with improvement in Q-factor of input inductor (L~1.5 nH), after substrate removal, the improvement is only marginal in the range of frequency of operation of the LNA. In the study of two turn inductors, it was shown that ~1.5 nH inductance was realized with a peak Q-factor of ~35 at frequency of ~7 GHz. Thus, it would be interesting to simulate the noise figure of the LNA using Q-factor ~30 and see

how it impacts the noise performance of the LNA. It could result potentially in higher improvements in noise figure.

As it was seen in the case of the switch, the power handling is limited due to excessive heating in the ON stack. To improve this, backside deposition of thermally conductive layers is a viable solution. In order to maintain low losses, a dielectric is preferable. AlN could be the material as it has high thermal conductivity and low dielectric losses [2]. Backside deposition of AlN after FLAME process is a possibility to explore. AlN is compatible with CMOS process and sputter deposition can be done from 100 - 800 °C [3].

As for the improvement of laser process, there are several possibilities to explore. For the improvement of morphology, it was seen that maintaining the focus inside the sample yielded better results. This strategy needs to be more carefully studied. We also saw in Chapter 3 that ablation in medium like water, the ablation quality is improved. Alternately, vacuum conditions are better than ambient environment as it allows free expansion of ablated plume and potentially reduce redeposition and debris. In the work of Matsumura et al., it was seen that formation of debris reduced in the presence of vacuum environment [4]. This can possibly allow reduce negative effects of debris on laser processing and help obtain better quality cavities.

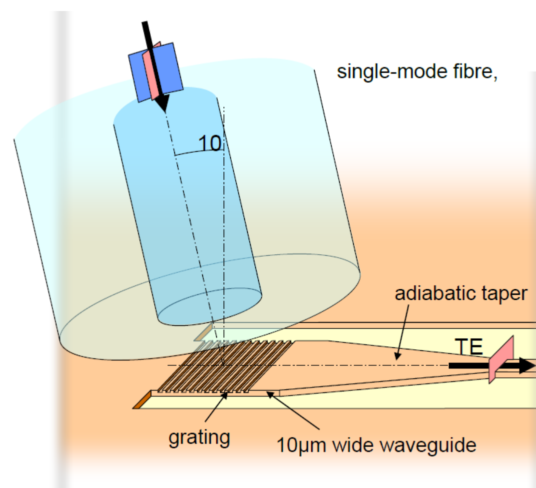


Fig. C.1: Coupling of light from fibre to photonic IC using grating coupler [5]

An additional avenue to explore is the study of backside coupling of light to photonic IC. Fig. C.1 shows light coupling scheme where light is focused on a grating coupler located on the front side of the die. The problem with this approach is the optical losses in different BEOL layers. Using the developed the FLAME process, membranes can be made to couple light from the backside. This

can help reduce optical losses. In general, FLAME process is compatible for any of the applications involving SOI membranes as listed in Chapter 1.

References

[1] J. Philippe, "Intégration hétérogène de systèmes communicants CMOS-SOI en gamme millimétrique sur substrat flexible," PhD Thesis, Lille 1, 2017.

[2] S. Kume, M. Yasuoka, S.-K. Lee, A. Kan, H. Ogawa, and K. Watari, "Dielectric and thermal properties of AlN ceramics," *J. Eur. Ceram. Soc.*, vol. 27, no. 8–9, pp. 2967–2971, Jan. 2007.

[3] S. Trolier-McKinstry and P. Muralt, "Thin Film Piezoelectrics for MEMS," *J. Electroceramics*, vol. 12, no. 1/2, pp. 7–17, Jan. 2004.

[4] T. Matsumura, A. Kazama, and T. Yagi, "Generation of debris in the femtosecond laser machining of a silicon substrate," *Appl. Phys. A*, vol. 81, no. 7, pp. 1393–1398, Nov. 2005.

[5] W. Bogaerts, "Lecture: Coupling light to silicon photonic circuits," *HELIOS Nov*, 2009.

Abstract

In semiconductor industry, the More-than-Moore approach is a key enabler for enhanced system performance, better integration and improved diversity of applications. Within the focus area of RF/microwave systems, it is essential to develop different functionalities which are optimized for various requirements like linearity, losses, sensitivity etc. While Silicon-on-Insulator (SOI) technology offers competitive solutions for RF/microwave market, it has been demonstrated in previous studies that SOI substrate engineering results in further performance gains. In this context, the specific goal of our work is the investigation of substrate processing of SOI RF functions using femtosecond laser ablation. The objective is to remove silicon handler substrate under the active area of the RF functions to obtain SOI membranes which have reduced RF losses and improved linearity. In this work, we have developed the Femtosecond Laser Assisted Micromachining and Etch (FLAME) process to suspend RF functions integrated on a SOI substrate. A high specific ablation rate of 8.5 x 106 µm3 s-1 has been achieved to produce membranes with a surface area ranging from few hundreds of µm2 to several mm2. RF characterization has been performed on different suspended RF functions: switches, inductors and low noise amplifiers (LNA). A comparison with high-resistivity SOI substrates shows superior performance of RF functions integrated in suspended membranes. For the SP9T switch, harmonic distortion measurements showed an improvement of 23 dB and 6 dB of the second and third harmonic, respectively. Small signal measurements of inductors on membranes revealed a near doubling of the quality factor of inductors up to 3.2 nH. Substrate removal of input matching inductor on LNA resulted in reduction of noise figure by ~0.1 dB. These results highlight the great potential for use of substrate processing for improvement of RF performance in CMOS technology. In addition, for short loop analysis needs, the FLAME method allows to quantify the influence of the substrate on losses and linearity very quickly without the need for total substrate removal. Another distinctive advantage of this methodology is the ability to quantify the substrate effect on a full circuit by suspending a specific component while keeping other components unaffected. The developed fabrication methods are equally usable for sensor applications on SOI technology, which provides an overall added value in line with the More-than-Moore paradigm.

Résumé

Dans l'industrie des semi-conducteurs, l'approche More-than-Moore constitue un facteur clé pour améliorer les performances du système, l'intégration et la diversification des applications. Dans le domaine des systèmes RF/hyperfréquences, il est essentiel de développer des fonctionnalités optimisées pour diverses exigences comme la linéarité, les pertes, la sensibilité, etc. Bien que la technologie silicium-sur-isolant (SOI) offre des solutions concurrentielles pour le marché des radiofréquences et des hyperfréquences, il a été démontré dans des études antérieures que l'ingénierie des substrats SOI permet d'améliorer encore les performances. Dans ce contexte, l'objet spécifique de ce travail de thèse a été d'étudier le traitement des substrats porteurs de tranches SOI (Silicium-sur-Isolant). L'objectif a consisté à enlever le substrat de silicium sous la zone active des fonctions RF pour obtenir des membranes SOI menant à des pertes RF réduites et une amélioration de la linéarité. Nous avons donc développé le procédé de micro-usinage et de gravure assistée par laser femtoseconde FLAME (Femtosecond Laser Assisted Micromachining and Etch) pour suspendre en membrane les fonctions RF intégrées sur un substrat SOI. Un taux d'ablation spécifique élevé de 8,5 x 10^6 μm^3 s^{-1} a été obtenu pour produire des membranes dont la surface varie de quelques centaines de μm^2 à plusieurs mm^2. La caractérisation RF a été réalisée sur différentes fonctions RF suspendues : commutateurs, inductances et amplificateurs à faible bruit (LNA). Une comparaison avec des substrats SOI à haute résistivité montre des performances supérieures pour les fonctions RF intégrées en membranes. Pour le commutateur, les mesures de distorsion harmonique ont montré une amélioration de 23 dB et 6 dB des secondes et troisièmes harmoniques, respectivement. Des mesures en régime petit signal d'inductance sur membranes ont révélé un quasi-doublement du facteur de qualité Q jusqu'à 3,2 nH. L'élimination du substrat de l'inductance d'adaptation d'entrée des LNA entraine une réduction du facteur de bruit de ~0,1 dB. Ces résultats mettent en évidence le potentiel important que constitue l'ingénierie des substrats pour l'amélioration des performances RF des technologies CMOS. De plus, pour les besoins d'analyse en boucle courte, la méthode FLAME permet de quantifier très rapidement l'influence du substrat sur les pertes et la linéarité sans avoir recours à des techniques d'élimination complète. Un autre avantage distinctif de cette méthode est la possibilité de quantifier l'effet du substrat sur un circuit complet en suspendant un composant spécifique sans affecter les autres. Les méthodes de fabrication développées sont également applicables aux capteurs en technologie SOI, ce qui apporte une valeur ajoutée globale en ligne avec le paradigme More-than-Moore.

List of publications

A. Bhaskar, J. Philippe, M. Berthomé, E. Okada, J-F. Robillard, D. Gloria, C. Gaquière, E. Dubois. “Femtosecond laser micromachining of crystalline silicon for ablation of deep macrosized cavities for Silicon-On-Insulator Applications”, *In SPIE LASE - Laser based Micro- And Nanoprocessing XIII*, San Francisco, Feb 2019.

A. Bhaskar, J. Philippe, M. Berthomé, E. Okada, J-F. Robillard, D. Gloria, C. Gaquière, E. Dubois. “Large-area femtosecond laser ablation of Silicon to create membrane with high performance CMOS-SOI RF functions.” *In IEEE Electronic System Integration Conference (ESTC)*, Dresden, Sept 2018.

J. Philippe, A. Bhaskar, E.Okada, F.Braud, J-F. Robillard, F. Danneville, C. Raynaud, D. Gloria, E. Dubois. “Thermal Analysis of Ultimately-Thinned-and-Transfer-Bonded CMOS on Mechanically Flexible Foils” *In IEEE Electron Device Letters*

E. Dubois, J. Philippe, A. Bhaskar, F. Braud, J-F. Robillard, D. Gloria. “High performance mechanically flexible CMOS technology for system moore integration*”, In 19th International Symposium on the Physics of Semiconductors and Applications, ISPSA 2018*, Jeju, Korea, July 1-5, 2018, paper MoB2-1

D. Zhou, F. Braud, A. Bhaskar, Q. Theret, E. Dubois, J.F. Robillard. “Fabrication of large-scale free-standing Si membrane using laser ablation” *In 16èmes Journées de la Matière Condensée, JMC 2018,* Grenoble, France, 27-31 août, 2018, paper PMQ8-P9, 517-517