

Integrated photonics for imaging: novel sources, architectures and applications

by

Marc de Cea Falco

M.S., Massachusetts Institute of Technology (2020)

B.S., Universitat Politècnica de Catalunya (2017)

Submitted to the Department of Electrical Engineering and Computer
Science

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

© 2024 Marc de Cea Falco. All rights reserved.

The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Authored by: Marc de Cea Falco
Department of Electrical Engineering and Computer Science
February 6, 2024

Certified by: Rajeev J. Ram
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by: Leslie A. Kolodziejcki
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Integrated photonics for imaging: novel sources, architectures and applications

by

Marc de Cea Falco

Submitted to the Department of Electrical Engineering and Computer Science
on February 6, 2024, in partial fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY

Abstract

Integrated photonics is considered a promising technology to tackle many current technological bottlenecks including data communication, computation, sensing and healthcare. Consequently, significant efforts have been dedicated to the study of integrated optical devices and systems for these applications.

However, despite its disrupting potential, the use of integrated photonics for imaging applications has only recently started to be considered, and its scope has mostly been limited to Light Detection and Ranging (LIDAR) for autonomous driving.

This thesis deals with the development and experimental demonstration of integrated optical systems for imaging applications. In particular, in the first part of this thesis we address the realization of arrays of silicon LEDs (both free space coupled and waveguide coupled) in CMOS photonics processes. We show that such sources can be used to realize compact lensless holographic microscopes, and we leverage CMOS integration to demonstrate a novel reflection holography architecture. We also show how silicon waveguide coupled light sources can be used to realize truly monolithic, highly multiplexed refractive index sensors at low cost and small form factor. In the second part of this thesis we explore the scaling limits of traditional Optical Phased Arrays (OPAs) for beam steering applications and study the use of non-uniformly spaced OPAs to circumvent such limitations.

This thesis lays out novel light sources, architectures and applications that leverage silicon photonics platforms to realize compact, low cost and high performance imaging systems. This serves as a stepping stone toward the transformation of currently expensive and specialized imaging systems into the consumer market.

Thesis Supervisor: Rajeev J. Ram

Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

This thesis is the result of many years of work, and it would not be a reality without the help of a great number of great people. First of all, I would like to thank Prof. Rajeev Ram for his invaluable and continued mentorship and guidance. It is hard to imagine a better understanding and approach to science and engineering than his - from approaching a problem from first principles to the constant questioning of assumptions and motivations to pursue a given project, to the continued pursuit of pushing the boundaries of what's possible, - these are fundamental principles that I will carry forward all along my professional career.

I also want to thank the members of my committee, Prof. George Barbastathis and Prof. Jelena Notaros, for their critical help in making this thesis a reality with thorough comments and on-point advice.

Special thanks to Dr. Amir H. Atabaki, who taught me all about scientific research and experimental characterization when I first arrived in the group. Without his infinite knowledge, talent, patience and kindness this thesis would have never even started.

A big shout-out to all the members of the Physical Optics and Electronics (POE) group at MIT, past and present, with whom I have shared not only an office, but also life - victories, defeats, tears, laughs and drinks. They are all brilliant researchers, but also great friends. Special mention to Nili Persits for our coffee-therapy sessions and to Gavin West for sharing the PhD path. I would also like to acknowledge Dr. Jaehwan Kim, Dr. Dodd Gray, Daniel Kramnik, Dr. Zheng Li and Dr. Iksung Kang.

I also want to acknowledge here everyone who made a difference, who lent a hand to make pursuing my PhD a reality: school and university teachers, research collaborators, colleagues at MIT and elsewhere - many, many thanks.

I did my PhD in Boston, but spent almost a year in San Francisco and I'm from

Barcelona, and I've met great people in all these places. Even if they didn't help directly in this thesis, they made this journey one worth living, and for that I am forever indebted: Barcelona (David, Marc, Alberto, Nacho, Cristina, Maria, Pablo, Albert, Anna), San Francisco (Jaime, David, Ignacio, Ricardo, Nicole, Clara, Edu, Cera, Nathan) and Boston (Armand, Guillem, Alex, Helena, Camille, Peter, Ola, Issac, John, Ondrej, Josep, Julia, Melanie). Juanjo will always have an honors mention for being there and being true no matter the circumstances.

Last, but definitely not least, infinite thanks to Guillermo, Mercedes and Marina - my family - for their unwavering support even when that meant me being away, for always believing in me and for always being proud no matter what.

Contents

1	Introduction and background	19
1.1	Integrated Optics	19
1.1.1	Silicon Photonics	21
1.2	Integrated optics for imaging: state of the art	23
1.3	Lensless microscopy	31
1.3.1	Shadow imaging	32
1.3.2	Lensless holography	35
1.3.3	Lensless imaging: The opportunity for integrated optics	40
1.4	Optical beamforming with OPAs	50
1.4.1	OPA theory and operating principle	51
1.4.2	OPA beamformers for LIDAR and Free Space Optical Communi- cations	57
1.4.3	Realization of integrated OPAs	65
1.4.4	OPA demonstrations: state of the art	74
1.5	Summary	79
1.6	Thesis outline	80
2	Lensless Holography with Near-Point CMOS Light Sources	83

2.1	Compact lensless holography with a single near-point CMOS light source	84
2.1.1	Near-point CMOS surface emitting LEDs	85
2.1.2	Lensless holography setup	88
2.1.3	Computational reconstruction: increasing quality with Deep Image Prior	89
2.2	Exploiting CMOS: realizing sub-wavelength spacing emitter arrays	94
2.3	Subpixel shift lensless holography with sub-wavelength CMOS emitter arrays	100
2.3.1	Reflection lensless holography - Geometrical considerations	102
2.3.2	Demonstration of subpixel shift reflection holography with CMOS LED arrays	111
2.4	Conclusion	131
3	Native Single-mode Waveguide-coupled LEDs in Unmodified Silicon Photonics Fabrication Platforms	133
3.1	Device designs	135
3.2	Operation principle and main physics	138
3.3	Device modeling	140
3.3.1	Electrical to optical efficiency: Sentaurus TCAD model	140
3.3.2	Mode coupling efficiency: Lumerical FDTD model	144
3.4	Experimental demonstration	148
3.4.1	Characterization results	148
3.4.2	On-chip detection of the generated light	154
3.5	Conclusion	161
4	Fully integrated, monolithic refractive index sensors in CMOS photonic fabrication processes	165

4.1	Refractive index sensing: a brief overview	166
4.2	Refractive index sensing with native waveguide coupled LEDs	171
4.2.1	Achievable performance	174
4.3	Experimental demonstration of resonator based refractive index sensing in a CMOS photonic platform	186
4.4	Conclusion	191
5	OPA architectures for advanced optical beamforming applications	193
5.1	Advanced OPA-based beamforming applications	194
5.1.1	Performance requirements	197
5.1.2	A suitable system-level Figure of Merit (FoM) for comparing OPA architectures	198
5.1.3	What now?	202
5.2	Non-uniform OPAs	208
5.2.1	Performance tradeoffs	212
5.2.2	System-level study	227
5.3	Conclusion	243
6	Conclusions and Future Work	247
6.1	Lensless Holography with Near-Point CMOS Light Sources	250
6.1.1	Improving the performance of near-point CMOS LEDs	250
6.1.2	Scaling the CMOS LED array to a large number of elements	251
6.1.3	Further exploiting computational techniques	252
6.2	Native Single-mode Waveguide-coupled LEDs in Unmodified Silicon Pho- tonics Fabrication Platforms	254
6.2.1	Further characterization	254

6.2.2	Device optimization	255
6.3	Fully integrated, monolithic refractive index sensors in CMOS photonic fabrication processes	257
6.3.1	NanoSOI tape-out: Focus on learning lessons in refractometric sensing	258
6.3.2	GlobalFoundries 45CLO tape-out: Experimental demonstration of fully monolithic refractive index sensors	260
6.4	OPA architectures for advanced optical beamforming applications	262
6.4.1	Round OPAs for optical power limited scenarios	262
6.4.2	Realization of OPA-based beamsteering systems for real-world FSOC applications	269
A	A survey of phase shifters for OPA-based beam steering applications	271
A.1	Thermo-optic phase shifters	272
A.2	Plasma dispersion phase shifters	273
A.3	Other phase shifters	276

List of Figures

1-1	Silicon photonic market value forecast by application (2022) [19].	23
1-2	Schematic of a generic imaging system.	24
1-3	The benefits of integrated photonics for imaging using a beam splitter as an example.	26
1-4	Comparison between lens-based microscopy (top) and lens-free microscopy (bottom).	32
1-5	Schematic of a generic lensless imaging system.	33
1-6	Lensless holography resolution limits due to coherence of the illumination source.	38
1-7	(a) Minimum source to sample distance necessary to achieve sufficient spatial coherence to have diffraction limited resolution as a function of emitter size.	44
1-8	Different approaches for the use of integrated optics for lensless imaging.	45
1-9	DBR reflectivity as a function of the number of pairs of high index - low index materials for a silicon platform (blue) and a silicon nitride (orange) platform.	46
1-10	A nominal 2D OPA.	52
1-11	OPA beam steering performance.	55

1-12 OPA 2D beam steering configurations.	57
1-13 OPA beamforming applications.	60
1-14 (a) Angular variance of the beam centroid due to beam wander as a function of link distance for different turbulence strengths and receiver aperture diameters.	62
1-15 The generic architecture of an OPA.	65
1-16 (a) Best axis and (b) worst axis performance of OPAs reported in the literature.	75
2-1 Near point CMOS surface emitting LEDs.	86
2-2 Lensless holography setup with CMOS LED as illumination source.	89
2-3 (a) Generic DIP architecture, which tries to minimize the difference between an experimental measurement g and the estimation of that measurement generated with a neural network $\hat{g} = H\hat{f}$	92
2-4 DIP-based holographic reconstruction.	95
2-5 A 4 element array of CMOS near-point source LEDs.	98
2-6 CMOS LED array spot size and spacing.	101
2-7 Reflection holography.	103
2-8 Geometric considerations for reflection holography.	105
2-9 Depth estimation with reflection holography.	109
2-10 Process to realize subpixel shift holography with illumination arrays.	111
2-11 CMOS LED array reflection holography setup.	113
2-12 Robustness to noise of different image registration algorithms.	116
2-13 Evaluation of subpixel shift in five different samples acquired with our CMOS LED array reflection holography setup.	117

2-14	Evaluation of superresolution image generation algorithms with a synthetic hologram.	122
2-15	High resolution holograms generated using the Structure-Adaptive Normalized (SAN) Convolution algorithm for our experimental samples. . . .	123
2-16	Peak Signal to Noise Ratio (PSNR) of different approaches for noise suppression.	124
2-17	Hologram reconstructions of subhologram 1 for three different samples. .	126
2-18	Hologram reconstructions of subhologram 2 for three different samples. .	129
2-19	Modified DIP architecture for noise robustness, where we include the presence of noise as an extra step to the hologram generation.	130
3-1	Native waveguide coupled silicon LED designs.	137
3-2	Main physical mechanisms affecting the generated optical power and efficiency of our silicon light sources.	139
3-3	Sentaurus TCAD models of (a) the resonant strip waveguide Si LED, and (b) the rib waveguide Si LEDs (both linear and resonant).	141
3-4	Sentaurus TCAD simulation of resonant strip waveguide (a, b) linear rib waveguide (c, d) and resonant rib waveguide (e,f) silicon LEDs.	142
3-5	Lumerical FDTD simulation models for (a) the strip waveguide resonant Si LED, (b) the linear rib waveguide Si LED and (c) the resonant rib waveguide Si LED.	147
3-6	Optical mode coupling efficiency η_c obtained with Lumerical FDTD as a function of wavelength for the resonant strip waveguide LED (a), the linear rib waveguide LED (b) and the resonant rib waveguide LED (c). .	148
3-7	Experimental characterization of waveguide coupled LEDs.	150

3-8	Illustration of the procedure to estimate the power in the waveguide for the resonant LED.	153
3-9	Comparison of the simulated and measured optical power and efficiency for both the resonant and linear Si LEDs.	154
3-10	On-chip detection of the light generated by the resonant strip waveguide Si LED through a Si resonant PD.	156
3-11	Ruling out heating as the source of photocurrent.	158
3-12	On-chip detection of the light generated by the resonant rib waveguide Si LED through a Ge linear PD.	162
4-1	Refractive index sensor with optical ring resonators.	168
4-2	Monolithically integrated refractive index sensor.	171
4-3	Sensing schemes for our fully integrated refractometric sensor, based on a feedback system that maintains alignment between a sensing and a tracking ring.	174
4-4	Model parameters for each of the components in the fully integrated refractometric sensor.	176
4-5	Optimizing the FWHM of the add-drop sensing ring in Sensing Scheme 1.	178
4-6	Sensitivity of Sensing Scheme 1 as a function of the total power generated by the LED in (a) resonance wavelength change, and (b) change in refractive index of the analyte.	179
4-7	Change in readout signal as a function of the Q factor of the tracking ring for different FSRs.	180
4-8	Sensitivity of sensing Scheme 2 as a function of the total power generated by the LED in (a) resonance wavelength change, and (b) change in refractive index of the analyte.	182

4-9	Sensitivity of Sensing Scheme 1 (a) and Sensing Scheme 2 (b) as a function of the group index mismatch between sensing and tracking ring.	185
4-10	Demonstration of refractive index sensing with an integrated poly-silicon ring resonator.	187
5-1	OPA-based reconfigurable FSOC networks.	196
5-2	The radiation pattern for a uniformly spaced (blue) and a non-uniformly spaced (orange) 1D OPA with the same total extent of 2 mm.	208
5-3	Non-uniform OPA emitter placement approaches.	210
5-4	The thinned array curse.	214
5-5	The effects of Side Lobe Suppression Ratio (SLSR).	217
5-6	Side Lobe Suppression Ratio (SLSR) characteristics of non-uniform OPAs.	219
5-7	Increasing SLSR by amplitude apodization.	225
5-8	System-level optical output power considerations for non-uniform OPAs.	230
5-9	OPA optical and electrical power consumption, dissipation and loss breakdown.	233
5-10	System-level electrical power consumption considerations for non-uniform OPAs in Scenario 1 (non-OPA based receiver).	240
5-11	Bump density requirements for OPAs as a function of the number of elements.	244
5-12	OPA configuration decision chart.	245
6-1	Scaling the CMOS LED array to large element number.	253
6-2	Use of blocking junctions to reduce effect of surface recombination in the electrical to optical conversion efficiency of waveguide-coupled LEDs.	255
6-3	Structures to increase coupling efficiency of waveguide-coupled LEDs.	258
6-4	Test structures for refractometric sensing included in the NanoSOI tape-out.	260

6-5	Test structures for the demonstration of the fully integrated refractometric sensor included in the 45CLO tape-out.	261
6-6	Round OPA gains.	264
6-7	OPA configuration decision chart including round OPAs.	266
6-8	Round OPA routing through pruned H-trees.	268
A-1	A comparison of the performance of difference phase shifters used for the realization of OPAs.	274
A-2	High performance phase shifters reported in the literature.	279

List of Tables

1.1	Survey of commercial lens-less and lens-based holographic and shadow microscope systems.	42
1.2	Grating arrays with the smallest grating spacing demonstrated in the literature.	47
1.3	NanoLED arrays reported in the literature.	49
1.4	Summary of OPA performance metrics and formulas.	56
1.5	Requirements for OPA-based LIDAR systems for automotive applications.	59
1.6	Requirements for OPA-based FSOC systems.	64
1.7	Requirements in emitter spacing and array size for OPA-based FSOC and LIDAR systems.	68
1.8	Requirements in phase shifter power and area for OPA-based FSOC and LIDAR systems.	72
1.9	A survey of OPA-based beam steering systems reported in the literature (Part 1).	76
1.10	A survey of OPA-based beam steering systems reported in the literature (Part 2).	77
1.11	A survey of OPA-based beam steering systems reported in the literature (Part 3).	78

2.1	Comparison of our CMOS nano-LED array with other nano-LED arrays reported in the literature.	100
4.1	Performance of optical refractive index sensors with different levels of integration.	170
4.2	Performance of optical refractive index sensors with different levels of integration.	183
5.1	Requirements for viable (left column) and advanced (right column) OPA-based LIDAR systems for automotive applications.	196
5.2	Requirements for simple (left column) and advanced (right column) OPA-based FSOC systems.	197
5.3	Requirements in phase shifter power and area for OPA-based FSOC and LIDAR systems.	199
5.4	A survey of non-uniform OPAs reported in the literature.	207
A.1	A survey of thermo-optic phase shifters reported in the literature.	280
A.2	A survey of efficient silicon phase shifters based on the plasma dispersion effect.	281
A.3	A survey of efficient phase shifters reported in the literature not based in the thermo-optic effect or the plasma dispersion effect.	282

Chapter 1

Introduction and background

1.1 Integrated Optics

Ever since the first laser source was developed around 1960, the use of optical beams to perform a wide variety of operations has increased to the point where it is hard to imagine modern life without it: medicine, biology, communications, imaging, defense, basic science, space exploration, navigation, consumer products, etc. All these and many others are fields where optical beams are routinely used. Most of these applications require not only the generation of an optical beam, but also its manipulation to adapt it to some specific requirements. Traditionally, the components necessary to perform these manipulations (prisms, mirrors, collimators, electro-optic modulators, photodetectors, beam combiners, optical isolators, etc.) are bulky, to the point where a dedicated table several meters long is necessary to implement many useful optical systems.

As a response, Stewart Miller proposed the integration of all these components in a single, planar chip in his seminal paper of 1969 [1] which started the field of integrated optics. By exploiting the large index of refraction contrast between a 'core' and a 'cladding'

material, nano- and micro-meter size structures that could guide and manipulate light were proposed, allowing for the miniaturization of complex optical systems, which came accompanied by a decrease in its cost and, ultimately, allowed access to the consumer market. Since then, an impressive range of integrated optical devices and systems have been proposed, realized and commercialized.

Different material platforms have been used over the years to realize integrated optical systems. Early demonstrations focused on platforms based on III-V semiconductors such as gallium arsenide (GaAs) [2] or indium phosphide (InP) [3], mainly because of their the efficient electro-optic mechanisms and its direct bandgap, which allows for the realization of on-chip lasers. III-V materials have, nevertheless, several drawbacks [4]: (1) Their maximum available wafer size to date is limited to about 200 mm (8") - with the typical size being 4" -, which raises concerns about its scalability and increases manufacturing costs; (2) Its relatively low index contrast makes the devices large compared to other material platforms; and (3) the realization of complex electrical circuits in the same chip is not easy due to a limited transistor yield. This prevents the monolithic integration of electronic and optical devices, which is highly desirable since nearly all modern photonic systems require control electronics.

Consequently, several material platforms have been explored as an alternative, with silicon (Si) [5] and silicon nitride (SiN) [6] being the most widely used both in academia and industry due to their favorable optical properties and compatibility with Complementary Metal Oxide Semiconductor (CMOS) fabrication processes. Over the last decades, additional materials have been used either in standalone platforms or in conjunction with silicon or silicon nitride to enhance device and system performance. Notable examples include Lithium Niobate [7], electro-optic polymers [8], 2D materials [9] and transparent conductive oxides [10]. While highly promising, the use of these materials poses an important challenge in terms of cost, ease of fabrication and scalability given that these are

not readily CMOS compatible.

In this chapter we will start by introducing the field of integrated optics, focusing on silicon photonics in particular (Section 1.1.1). We will then discuss how integrated optics can be applied to imaging applications and review the current state of the art (Section 1.2). Afterwards, we will delve into two specific imaging applications where integrated optics hold a great potential for impact: lensless microscopy (Section 1.3) and optical phased array beamforming (Section 1.4). We will introduce the theory behind these and derive requirements and figures of merit for integrated optical systems targeted at these applications. Finally, we will summarize our findings in Section 1.5 and provide a brief overview of this thesis in Section 1.6.

1.1.1 Silicon Photonics

The potential to use Si as a material for integrated optics has been recognized since the 1980s [11, 12]. The fact that nearly every electronic system manufactured nowadays is made of silicon makes a strong case towards its use in integrated optics: all the knowledge, research, fabrication and manufacturing processes, infrastructure and technologies developed over the last century by the microelectronics industry can be readily exploited to realize silicon photonic systems. Not only this, but using silicon allows for the fabrication of complex, high performance electronic circuits in the same chip ('monolithically'), alongside the photonic components. Therefore, complex control circuitry, as well as functionality requiring both electrical and optical operations can be implemented without the need of co-packaging multiple chips.

Furthermore, silicon is an inexpensive, extremely well understood material with a high quality native oxide. The high index contrast between this oxide and the silicon allows for the realization of high optical confinement structures in standard Silicon on Insulator

(SOI) platforms [13], offering a way to fabricate low loss optical waveguides and very compact photonic devices.

Once these advantages were recognized both by industry and academia, rapid development of a great variety of high performance optical and electro-optical devices and systems in silicon platforms followed [5, 14]. While most of the early development was directed towards the telecom and datacom spaces, a range of optical systems for a variety of applications have been demonstrated in silicon photonics platforms. Examples include sensing [15], nonlinear optics [16], quantum optics [17] and computation [18]. It is projected that the market value of silicon photonics will be around \$1 billion by 2027, with an impressive compound annual growth rate of 36% [19] (Fig. 1-1).

SiN is another popular material platform. Compared to Si, SiN has significantly lower optical propagation losses and is transparent at visible wavelengths, which makes it an attractive platform for applications highly sensitive to loss (such as quantum or nonlinear operations) or that require operation at visible wavelengths. On the other hand, SiN is an insulating dielectric material so the realization of active components (such as modulators or photodetectors) is challenging.

As the silicon photonic industry grows and fabrication processes mature, new processes incorporating more than one photonic material are arising. Photonic fabrication platforms offering monolithic integration of silicon and silicon nitride waveguides alongside high performance transistors are now standard and commercially available [20, 21], offering great design flexibility and allowing to take advantage of "the best of all worlds". Furthermore, these platforms also include germanium (Ge) for the realization of high responsivity photodetection at near-infrared wavelengths.

The most significant limitation of photonic platforms based on Si or SiN is the lack of monolithic light sources, which makes it necessary to have an external chip where the light is generated, and requires a means of coupling the generated light into the Si/SiN

2021-2027 silicon photonic die forecast by application

(Source: Silicon Photonics 2022, Yole Intelligence, July 2022)

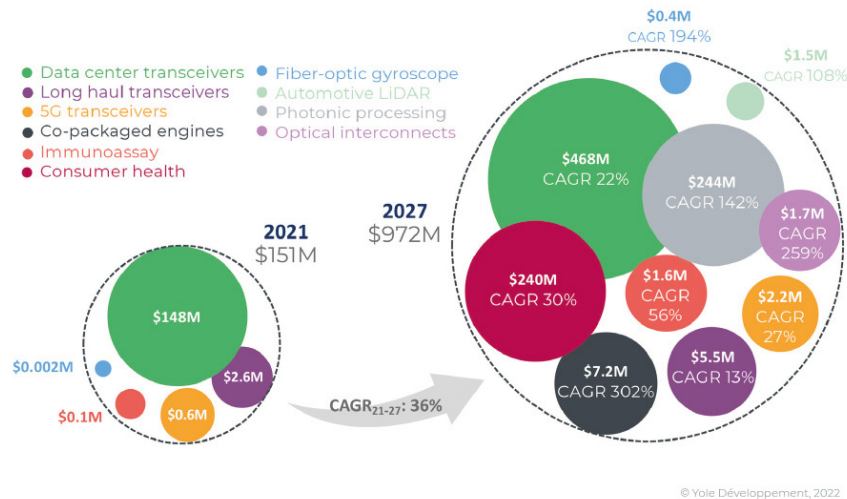


Figure 1-1: Silicon photonic market value forecast by application (2022) [19].

chip. This is usually achieved either through fiber coupling or by hybrid integration of the source into the photonic chip [22].

1.2 Integrated optics for imaging: state of the art

Integrated optics in general, and silicon photonics in particular, is seen as a promising technology with the potential to impact a range of fields including communications, computing and consumer healthcare. These are, in fact, the three biggest growth drivers of the silicon photonics industry [19] (Fig. 1-1).

In comparison, little attention has been directed toward the investigation of integrated photonics for imaging applications. In fact, a single imaging application (automotive LIDAR) appears in the market forecast shown in Fig. 1-1, with only 0.15% of the market

share in 2027 ¹.

Nevertheless, integrated optics has the potential to allow for the realization of compact, high performance imaging systems with advanced capabilities at a lower cost. Miller's original vision of exploiting integrated optics as a means to substitute free space optical components traditionally used in optical systems can easily be applied to imaging systems, which require expensive and bulky components such as lenses, collimators, spatial light modulators, etc.

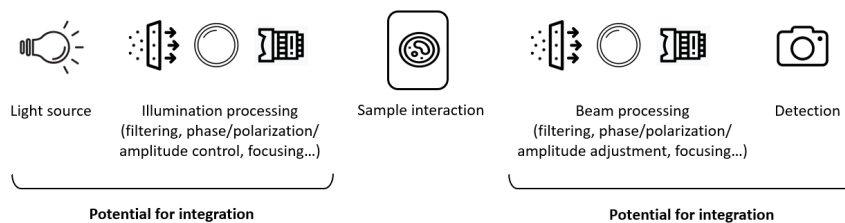


Figure 1-2: Schematic of a generic imaging system. Integrated optics can be leveraged all along the imaging path.

In fact, integrated optics can be leveraged in virtually all the components of a generic imaging system (Fig. 1-2), from generation and manipulation of the illumination to the collection of the beam that has interacted with the sample. Several advantages arise when using integrated optics for imaging (which we illustrate with a beam splitter in Fig. 1-3):

- It allows for the miniaturization of components and therefore allows for the realization of small form factor, compact imaging systems. This results in lower cost devices and can open up the use of these systems in novel spaces such as consumer electronics.

¹The report in [19] is made by Yole Group, which has been publishing market reports on the silicon photonics industry for over 20 years. Most market reports on silicon photonics made by other companies include imaging in the "other applications" category, so its specific market share cannot be obtained.

- Related to the above point, the fact that components can be miniaturized and cheaply fabricated allows us to replicate devices and systems in the same substrate almost *for free*, a feature which is inaccessible in traditional bulk optics imaging systems.

As an example of the above two points, take a beam splitter. A bulk optics beam splitter costs on the order of \$100 and is several centimeters in size. An integrated optical beam splitter (usually referred to as a "3 dB coupler") occupies less than 100 μm and costs less than a penny (at scale, CMOS fabrication processes cost around \$1/ mm^2).

- Since integrated optical systems rely on light propagating along waveguides, alignment of the different components is ensured as long as they are connected by a waveguide. This is of course not true in bulk optics systems, where careful alignment is required. Additionally, routing of the optical beams is significantly easier as compact waveguide bends with arbitrary angles can easily be realized in integrated optics, compared to the use of mirrors in bulk optics.
- Integrated optics offers an avenue towards realizing reconfigurable imaging systems, which is challenging in traditional imagers based on bulk optics. Phase, amplitude, wavelength and even spatial control of light in a photonic integrated circuit can be achieved by exploiting active photonic and electro-optic devices.
- Integrated optics also allows for the realization of small emitting apertures and, moreover, enables us to place multiple apertures close to each other, something which is extremely hard to realize in bulk optics systems. As we will review below, this is of paramount importance for the realization of beamforming systems such as those used for light detection and ranging (LIDAR) or free space optical

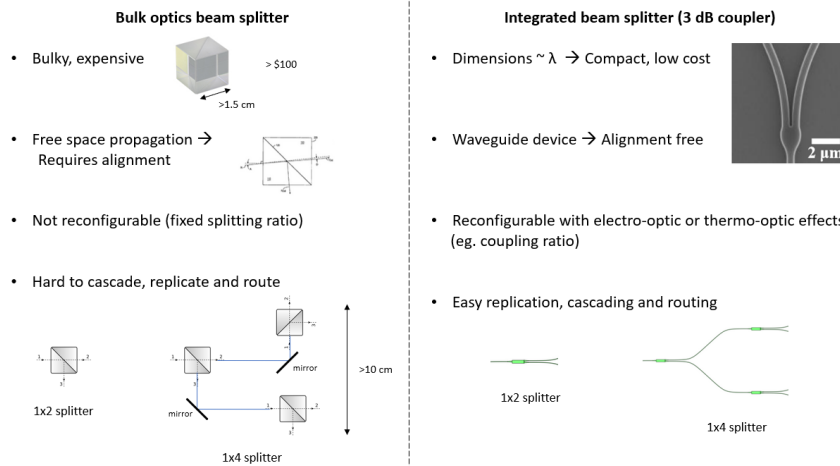


Figure 1-3: The benefits of integrated photonics for imaging using a beam splitter as an example. The micrograph of the integrated beam splitter is from [23].

communications.

While not the main focus of research on the integrated photonics field, these advantages have been recognized and a variety of integrated devices and systems for imaging have been demonstrated and reported in the literature. These can be divided in 6 main categories with distinct functionalities and applications ²:

- Routing and wavelength multiplexing and demultiplexing: As already mentioned, the high confinement of the optical mode in optical waveguides offers the possibility of routing different optical signals in a small area (depending on material platform and operating wavelength, the bend radius of integrated waveguides are typically in the range of 10 - 200 μm), which is highly beneficial in applications requiring a remapping of multiple generated/collected beams, such as in stellar interferometric imaging [24] or spatially multiplexed imaging [25].

²In the interest of brevity, we provide only a brief overview of these functionalities. The reader is pointed to the cited references for more detail.

Similarly, optical waveguides can also be leveraged for wavelength multiplexing and demultiplexing, which has applications in multi-wavelength imaging systems [26] and spectroscopic/diffractive imaging [27, 28].

- Flat optics: Integrated optics allow for the fabrication of structures with dimensions on the order of the operating wavelength, which opens up the avenue for the realization of optical antennas and metasurfaces that allow for wavefront manipulation in flat photonic chips, circumventing the need for bulky and expensive objectives and lenses [29].

While these are not technically integrated optics because the light propagates through the chip and not along the chip, flat optics has been successfully integrated with photonic integrated circuits for imaging [30, 31].

- Sources and devices for arrayed illumination: Imaging schemes where multiple images are acquired with either the illumination source, the sample or the detector at different positions (usually called sub-pixel sampling) are common to enhance resolution [32, 33]. Traditionally, this is achieved by having an accurate mechanically moving stage, which results in increased complexity and cost. The use of arrays of illuminating sources such as a micro-LED array [34] or a photonic integrated circuit with multiple, individually addressable emitting apertures [35] offers an alternative that does not require mechanical parts.
- Structured and non-standard illumination generation: Structured illumination is another approach to achieve superresolution imaging in which multiple images are acquired with different spatially structured illumination beams and subsequently combined [36]. Multiple integrated optics approaches to generate the spatially structured illumination beams have been reported in the literature [37–39], allowing

for a significant reduction in system complexity and cost compared to traditional approaches.

Additionally, integrated optics has also been leveraged to generate non standard illumination. Examples include evanescent field illumination with single- and multi-mode waveguides [40, 41] and generation of orbital angular momentum (OAM) beams [42].

- Sources and devices for lensless imaging: As the name implies, lensless imaging consists of the acquisition of images without the use of focusing lenses [43, 44], and usually requires computational post-processing or the acquisition of multiple images to recover the desired scene (lensless imaging will be covered in detail in Section 1.3). Several integrated optics approaches for lensless imaging have been reported in the literature, either exploiting small emitting apertures to achieve large magnification [45] or using angular or spatially selective emitting/collecting apertures [46, 47].
- Optical Phased Arrays (OPAs): In analogy with radio-frequency (RF) phased arrays, optical phased arrays allow for the accurate control of the phase difference between closely spaced emitting/receiving optical antennas, which allows for the control of the emission properties of the array in which is traditionally known as beamforming [48] (OPAs and beamforming will be covered in detail in Section 1.4). Integrated OPAs offer significant advantages in terms of compactness, cost, scalability and complexity compared to non-integrated solutions, which usually use slow, bulky deformable mirrors or liquid crystals. Consequently, extensive efforts have been made toward the realization of integrated OPAs, and it arguably constitutes the most heavily researched integrated optical device that can be applied to imaging (see [48] for a review of existing demonstrations). This is due to the high impact of

the applications enabled by OPAs, which include LIDAR [49] and free space optical communications [50].

It is clear from our discussion up to this point that integrated optics has a lot to offer for improving and simplifying imaging systems. Why, then, is integrated optics still not widely used in imaging systems? This can be explained by two main reasons:

1. The need for in- and out-coupling from the photonic integrated circuit to free space. Imaging requires the light to interact with the sample/scene we want to observe, which means that the light generated and/or processed in the emitter chip needs to be out-coupled to free space and in-coupled to the receiver chip after interacting with the scene.

Simple, efficient and controllable light coupling is one of the outstanding challenges in integrated photonics, even for non-imaging applications such as communication and computing. Nevertheless, the challenge is even bigger for imaging systems, as they ideally require individual control (in amplitude and phase) of small, tightly packed efficient radiating apertures (see Sections 1.3 and 1.4 for a detailed discussion).

2. The lack of a scalable photonic fabrication platform capable of implementing all required functionalities. As discussed previously, silicon and silicon nitride are arguably the only platforms available at the moment that can operate at large fabrication volumes. Nevertheless, they both lack monolithic sources, silicon waveguides cannot operate at visible wavelengths (which are relevant for biological imaging) and silicon nitride lacks active components. If multiple chips and/or material platforms are required for the realization of useful imaging systems, then the appeal of integrated optics is lost.

It is worth noting that this is changing thanks to the emergence of platforms that monolithically integrate silicon and silicon nitride, as well as advancements on heterogeneous integration of light sources in photonic fabrication platforms.

This thesis proposes solutions to both of these challenges for specific applications:

1. For lensless imaging systems, we discuss the development and use of near point source surface emitting LED arrays made in silicon for holographic imaging, allowing for compact, high magnification imaging with simple, low cost setups. This is discussed in Chapter 2.
2. We propose and demonstrate the realization of waveguide coupled Light Emitting Diodes (LEDs) in commercial silicon photonics platforms without any process modification. The availability of native light sources in silicon can eliminate the need for heterogeneous integration in certain applications, significantly reducing the barrier to the realization of truly integrated imaging systems. This is discussed in Chapter 3.
3. To demonstrate the potential of our silicon waveguide coupled LEDs, we analyze in detail the feasibility of realizing fully monolithic refractive index sensors in silicon photonics platforms. We show that performance comparable to state of the art systems can be achieved at a fraction of the cost and in a much more compact form factor. This is presented in Chapter 4.
4. In Chapter 5 we investigate the use of non-uniform OPA architectures to provide a scalable path toward the realization of systems for advanced optical beamforming applications. These architectures can ease the requirements in terms of area and power consumption, which is critical since, as we will see in Section 1.4.2, no

currently demonstrated phase shifter (one of the most critical components in an OPA) can deliver the required performance for traditional OPA architectures.

Since we focus on lensless microscopy and beamforming, in the next two sections we analyze these in detail and derive figures of merit and performance requirements for integrated optical devices and systems tailored to these applications.

1.3 Lensless microscopy

While tremendously successful and useful, traditional high performance optical microscopes are complex, bulky and expensive due to the need of high numerical aperture and high magnification objectives.

Lensless imaging offers an alternative to traditional microscopy (and imaging in general) by acquiring images without being *processed* by a lens [43, 44]³. Over the last decade, lensless imaging approaches with sub-micrometer resolution have been reported (e.g. [51, 52]), demonstrating that these techniques can compete with traditional microscopy without compromising resolution.

Compared to lens-based microscopy, lensless imaging systems are significantly simpler (Fig. 1-4) and therefore cheaper, more compact and portable, have a larger field of view for the same magnification when compared to a traditional microscope and can easily achieve depth resolved (3D) imaging. However, lensless microscopy requires computational processing of the acquired image and has a light collection capability limited by the imaging array itself (in traditional microscopy increased light collection can easily be achieved using lenses in front of the imager).

³These are two excellent references for the reader interested in a more detailed analysis of lensless imaging than the one presented here.

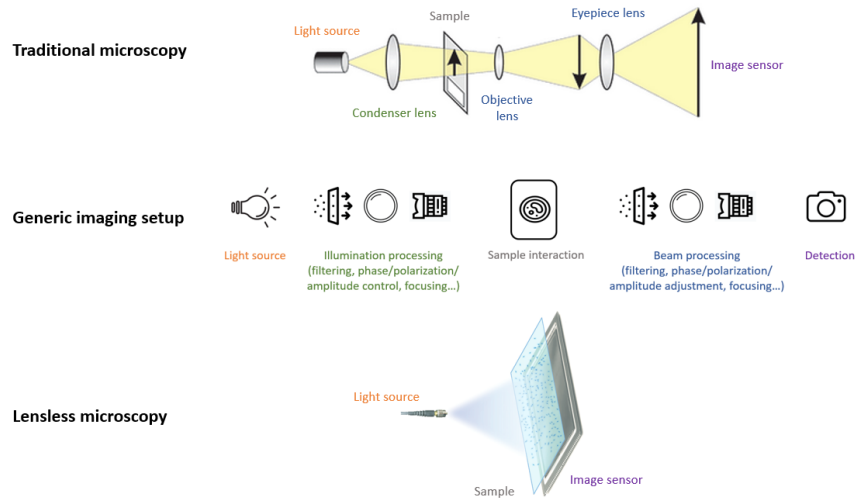


Figure 1-4: Comparison between lens-based microscopy (top) and lens-free microscopy (bottom). The middle shows the schematic of a generic imaging setup. Lens-free microscopy results in a much simpler system. The top schematic is adapted from [53], and the bottom from [44].

There are two main techniques for lensless imaging, namely shadow imaging and holographic imaging. While similar, they differ in the requirements for the illumination source and have different resolution and performance limitations. We discuss them separately in the following sections.

1.3.1 Shadow imaging

In shadow imaging, a light source is used to illuminate the sample of interest and the resulting shadow is recorded with an imaging array. A generic setup for shadow imaging is shown in Fig. 1-5 along with a sample image (bottom left of Fig. 1-5). Shadow imaging has been applied to the monitoring of cell division and motility, as well as cytometry of a wide range of cells including red blood cells, leukocytes, sperm cells and cardiomyocytes amongst others [44].

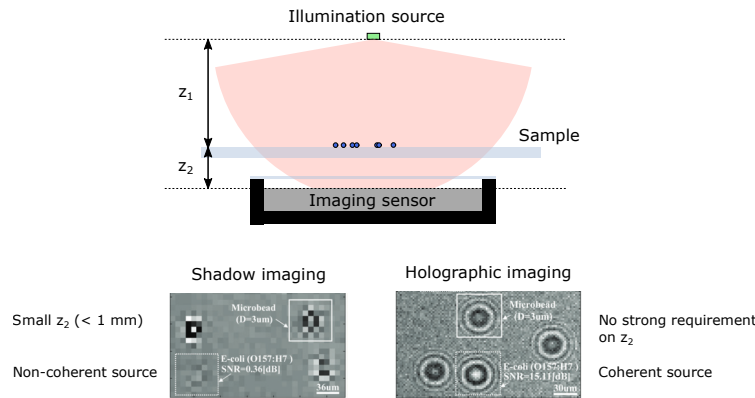


Figure 1-5: Schematic of a generic lensless imaging system. Representative images in shadow imaging (left) and holographic imaging (right) approaches are shown. The shadow and holographic images are reproduced from [54].

Source requirements for shadow imaging

Standard shadow imaging poses almost no constraints on the illumination source. The most important requirement is power, so that a sufficiently high signal to noise ratio (SNR) can be achieved.

Since it does not rely on coherent interactions (only on the formation of a shadow), there are no requirements for either temporal or spatial coherence. In fact, shadow imaging has been demonstrated using sunlight as the illumination source [55]. While not required, using a source with high spatial coherence results in shadows with decreased edge blurriness, which can be important for improving imaging quality. Similarly, using temporally and spatially coherent beams for shadow imaging so that an interference pattern instead of a shadow is obtained can result in improved performance in some cases [54].

Geometric considerations for shadow imaging

In any generic lensless imaging setup (either shadow or holographic) there are two geometric distances that have an important impact in the image quality: the distance

between the source and the sample z_1 and the distance between the sample and the imaging sensor z_2 (these are depicted in Fig. 1-5).

In shadow imaging it is of interest to minimize the sample - sensor distance z_2 so as to minimize diffraction of the light scattered from the sample, as increasing diffraction results in increasingly distorted shadow that has less resemblance to the original object. In practical setups, the minimum achievable z_2 is usually limited by the thickness of the sample holder (usually a quartz coverslip) and by the thin protective glass layer present in imaging sensors for protection purposes.

Resolution limits for shadow imaging and super-resolution approaches

There are two main resolution limits for shadow imaging:

- The diffraction of the generated shadow in the propagation path from the sample to the sensor.
- The pixel size of the image sensor. Modern CMOS imaging sensors have pixel sizes ranging from 1-10 μm , so practical resolution limits of shadow imaging are on that order.

The resolution limitation imposed by the pixel size can be overcome by exploiting sub-pixel sampling, where multiple images of the same sample spatially shifted a distance smaller than the pixel spacing are acquired. By computationally merging the multiple acquired low resolution images, a composite high resolution image is obtained.

For shadow imaging, sub-pixel sampling has been demonstrated with the use of illumination at different angles [55] or by physically displacing the sample by small amounts using microfluidics [56]. Another possible approach is the use of an array of sub-micron apertures where the light transmitted through each aperture is associated to a single region of the imager array. By scanning the sample along the aperture array, different

portions of the sample can be imaged at a time [57]. In these cases, the resolution is limited by the size of the aperture and the achievable resolution of the sub-pixel sampling movement.

1.3.2 Lensless holography

Holographic imaging employs the same setup as shadow imaging (Fig. 1-5) with the difference that now the illumination source is partially coherent. As a result, a coherent diffraction pattern instead of a shadow is obtained at the imaging sensor (bottom right in Fig. 1-5). Such a coherent diffraction pattern contains complex field information (both amplitude and phase), which results in (1) generally better reconstruction quality than shadow imaging and (2) the possibility to reconstruct the image at different propagation distances (i.e, it allows for reconstructing a 3D image).

The recorded diffraction pattern is an in-line hologram resulting from the interference between the light scattered by the sample and reference light going through the sample undisturbed:

$$I_{sensor}(x, y) = |E_{ref}(x, y) + E_{scat}(x, y)|^2 = I_{ref}(x, y) + I_{scat}(x, y) + E_{ref}^*(x, y)E_{scat}(x, y) + E_{ref}(x, y)E_{scat}^*(x, y) \quad (1.1)$$

In the equation above, I_{sensor} is the image acquired at the sensor plane, E_{ref} (I_{ref}) the electric field (intensity) of the unscattered reference wave, E_{scat} (I_{scat}) the electric field (intensity) of the light scattered by the sample.

With knowledge of the reference electric field E_{ref} (usually a plane or spherical wave) we can retrieve E_{scat} and therefore reconstruct the sample, a computational process usu-

ally referred to as backpropagation. A great range of backpropagation algorithms, each with a different set of advantages and limitations and with different levels of complexity and assumptions have been employed and reported in the literature. Since this is not the focus of this thesis, the interested reader is pointed to references [44, 58, 59] as good starting points for discussion of backpropagation algorithms.

A variety of lensless holographic imaging systems have been demonstrated, with applications ranging from tissue and cell imaging to air quality control and particle tracking [44, 58].

Source requirements for lensless holography

Since holographic imaging requires coherent interaction, suitable illumination sources that are at least partially spatially and temporally coherent are required. The coherence properties of the light determine the maximum angle θ_{max} at which there will be a coherent interference, which essentially sets the equivalent numerical aperture (NA) of the lensless holographic system and thus its achievable resolution [58]:

$$\Delta x = \frac{\lambda}{2NA} = \frac{\lambda}{2n \sin(\theta_{max})} \quad (1.2)$$

Temporal coherence describes the correlation of a light wave with itself at a different instance in time [60], and is related to the spectral bandwidth of the source $\Delta\lambda$ (Fig. 1-6, left). The temporal coherence length of a source is defined as:

$$L_{coh,temp} = \frac{2 \ln(2)}{\pi} \frac{\lambda^2}{\Delta\lambda} \quad (1.3)$$

Coherent interaction of the rays generated by the same source will occur as long as the difference in propagation length between them is below $L_{coh,temp}$, which results in:

$$\theta_{max} \leq \arccos\left(\frac{z_2}{z_2 + L_{coh,temp}}\right) \quad (1.4)$$

Where z_2 is the sample - camera distance as depicted in Fig. 1-5. Temporal coherence length and resolution limits for different wavelengths and z_2 are shown in the left column of Fig. 1-6. We can see how a good temporal coherence resolution limit is achieved as long as we ensure $L_{coh,temp} > 5z_2$.

Complementarily, spatial coherence describes the correlation of spatially separated portion of the same wavefront [60] and is related to the spatial extent of the light source D (Fig. 1-6, right). Similar to the case of temporal coherence, coherent interaction of the light will take place for rays whose difference in propagation length is below the spatial coherence length of the light at the sample plane, which according to the Van Cittert - Zernike theorem is proportional to $\lambda z_1/D$ [61]. This translates to a maximum coherent interaction angle of [58]:

$$\theta_{max} \leq \arctan\left(\frac{0.61\lambda z_1/D}{z_2}\right) \quad (1.5)$$

Notice how spatial coherence length is not only a property of the light source dimension D but also depends on the distance to the sample z_1 . This is why stellar objects (with large D) can be used as a spatially coherent source (z_1 is also large). Spatial coherence length and resolution limits for different source sizes (D) and imaging distances (z_1, z_2) are shown in the right column of Fig. 1-6 for a wavelength of $1 \mu\text{m}$. In general, good spatial coherence resolution limit is achieved when $\lambda z_1/Dz_2 > 10$.

On top of spatial coherence considerations, the spatial extent of the source also affects the smearing of the generated hologram at the imaging plane. Essentially, each infinitesimal point at the illumination source generates a slightly displaced hologram at the imaging plane. Mathematically [58]:

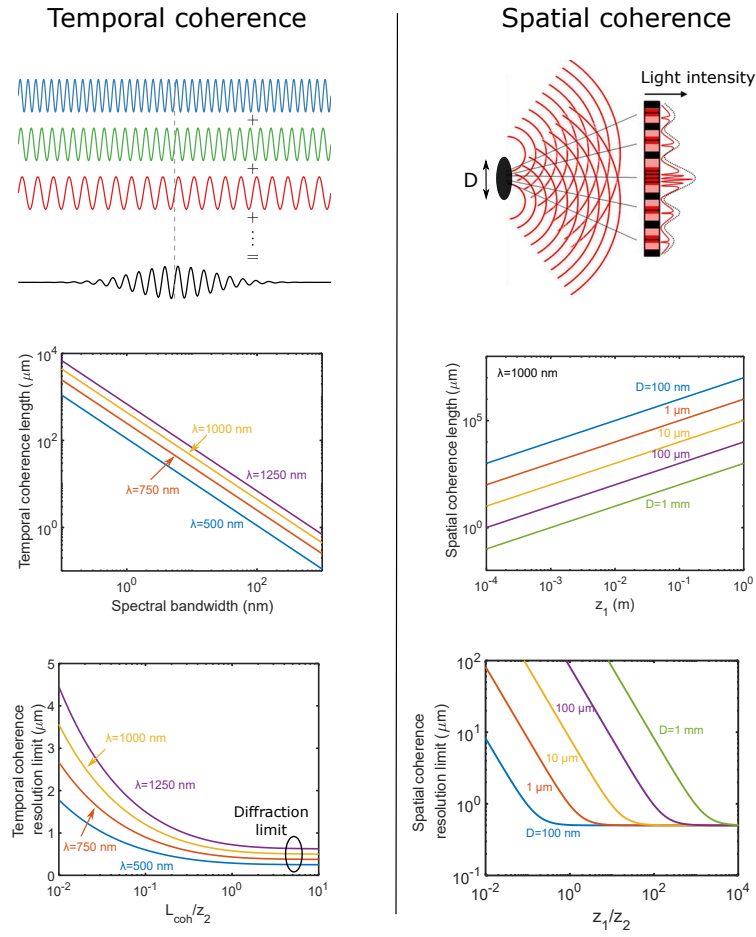


Figure 1-6: Lensless holography resolution limits due to coherence of the illumination source. The limits associated to temporal coherence are shown in the left, while the limits associated to spatial coherence are shown in the right. The limits associated to temporal coherence are shown in the left, while the limits associated to spatial coherence are shown in the right. The schematic of temporal coherence is adapted from [62], and that of spatial coherence from [63].

$$I_{meas} = I_{point\ source} \otimes S\left(-\frac{z_1}{z_2}x, -\frac{z_1}{z_2}y\right) \quad (1.6)$$

Above, $I_{point\ source}$ is the hologram that would be generated if the source was an ideal point source, \otimes denotes convolution and $S(x, y)$ describes the illumination source.

Notice how the aperture gets demagnified by a factor z_1/z_2 ⁴. To eliminate this effect, we need to ensure that the extent of the smearing is below the pixel size of the imager: $Dz_2/z_1 < T$, where T is the pixel size.

Besides coherence properties, the light source employed in lensless holographic imaging needs to have a high enough power to achieve a good SNR.

Geometric considerations for lensless holography

From our discussion on source requirements above, it is clear that the geometry of the lensless holographic setup plays an important role in achievable resolution. Limitations stemming from partial coherence of the illumination source can be overcome by adjusting z_1 and z_2 accordingly. In particular, we have seen how temporal coherence can be ensured as long as $L_{coh,temp} > 5z_2$, which means that if we use a low coherence source we need to decrease z_2 . We have also seen how spatial coherence is achievable if $\lambda z_1/Dz_2 > 10$, so for spatially extended sources we need to increase z_1 .

Besides these considerations, it is relevant to notice that when using a high NA illumination source it is possible to achieve a magnification of the hologram by a factor $M = (z_1 + z_2)/z_1$. Magnification can result in a resolution improvement at the expense of a smaller field of view.

Resolution limits for lensless holography and super-resolution approaches

In addition to the resolution limits mentioned above, and similarly to the case of shadow imaging, the pixel size of the imaging array imposes limitations on the achievable resolution. In particular, the pixel pitch of the imager needs to be fine enough to capture the interference fringes generated by the coherent interaction of the scattered and

⁴This is easy to see from geometrical optics considerations by drawing rays from two points separated a given distance D at the source plane passing through the same spatial point in the sample plane and propagating to the image plane. If z_1 is the source to sample distance and z_2 the sample to imager distance, then the distance between the two rays at the sample plane is inverted with respect to the source plane and its magnitude given by Dz_1/z_2 .

reference beams. In commercial imagers, this limits the achievable resolution to about $1 \mu\text{m}$.

Nevertheless, the pixel pitch limitation can be overcome by using sub-pixel sampling in the same fashion as we discussed for shadow imaging. A wide range of sub-pixel sampling approaches have been reported in the literature, including moving the sample [64], moving the illumination source [65] or using an array of emitters [51]. Alternative approaches that do not rely on in-plane movement of the sample/illumination have also been reported, including acquiring holograms at slightly different wavelengths [66] or varying the propagation distance z_2 [67]. Resolution as low as 225 nm has been demonstrated using these techniques coupled with advanced image reconstruction algorithms [68].

1.3.3 Lensless imaging: The opportunity for integrated optics

It is fair to say that the lensless imaging community has focused on the development of new computational frameworks and algorithms to improve resolution and imaging quality. While this is of course of paramount importance, it has put the development of new hardware in general, and the exploitation of integrated optics in particular, on a second priority. There exists therefore a mostly unexplored opportunity space for integrated optics in lensless imaging.

It is important to realize that, in contrast to conventional microscopy, the main resolution limitations in lensless imaging come from the properties of the illumination source (mainly its coherence) and the imaging sensor characteristics (mainly pixel pitch, which in modern CMOS imagers is limited to around $1 \mu\text{m}$).

Good temporal coherence sources are commercially available, either through lasers or narrowband light emitting diodes. Good spatial coherence can be achieved by ensuring a small emitting aperture D combined with large enough source to sample distance z_1 (Eq.

1.5). In typical lensless imaging setups, spatial coherence is achieved by using either a pinhole or an optical fiber as the imaging aperture (with typical diameters ranging from $1 \mu\text{m}$ to $10 \mu\text{m}$) and a distance z_1 on the order of 5-10 cm. While effective, this results in relatively bulky imaging systems with dimensions on the order of tens of cm, and in the case of pinholes requires careful alignment to the source.

We can thus identify a first promising avenue for integrated optics, which is **the realization of small aperture illumination sources (ideally with sub-micrometer size)**. As discussed, non-integrated approaches limit the aperture size to above $1 \mu\text{m}$ and are typically closer to 10 - 20 μm , which makes it necessary to have z_1 on the order of 5 cm (Fig. 1-7(a)). Due to the very low minimum feature size in integrated optical platforms (around 100 nm), integrated optics has the potential to realize source apertures below $1 \mu\text{m}$, which would allow for a decrease in z_1 and therefore for the realization of more compact lensless imaging setups. Furthermore, realizing such small apertures in an integrated chip would bypass the need for either a pinhole or an optical fiber, further simplifying the system and thus reducing its cost and increasing robustness.

As we have already discussed, once we have realized an individual source aperture with the desired spatial and temporal coherence properties in an integrated optical platform, realizing arbitrarily large arrays of such sources comes at almost no additional cost. This constitutes a second highly promising avenue for integrated optics in lensless imaging systems: **the realization of arrays of sub-micrometer light sources with sub-pixel spacing**. This would allow the use of sub-pixel sampling techniques to overcome the resolution limitation set by the imaging sensor pixel pitch.

Previous sub-pixel sampling techniques exploit either precise mechanical movement of the sample or the imaging source [32, 52] or source arrays with large spacing where movement demagnification results in sub-pixel shifts [65] (which require a large z_1/z_2 ratio and therefore result in bulky systems). Commercially available lensless holographic

Company	Brief description	λ	FoV	Lateral resolution	Size & weight	Frame rate
Lyncee tec [69]	Lensless holography	NR	6 x 6 mm	10 μm	NR	NR
Lyncee tec [70, 71]	Lens holography	666 nm	1.32 x 1.32 mm 330 x 330 μm 165 x 165 μm	2.85 μm 0.67 μm 0.44 μm	600x350x500 mm 30 kg	25 fps
Ovizio [72]	Lens holography	630 nm	320 x 320 μm	1.5 μm	322x212x400 mm 12 kg	15 fps
MetroLaser [73]	Lens holography	650 nm	NR	0.8 μm	550x100x150 mm 5 kg	12 fps
Phi [74]	Lens holography	635 nm	500 x 500 μm	1 μm	290x170x185 mm 3.75 kg	1 fps
Holmarc [75]	Lens holography	650 nm	180 x 120 μm	1 μm	NR	NR
Dantec dynamics [76]	Lens shadow imaging for sizing particles	532 nm	2.1 x 1.3 mm 11.6 x 7.3 mm 13.5 x 8.4 mm	2.7 μm 11.5 μm 13.3 μm	NR	16 fps 800 fps

Table 1.1: Survey of commercial lens-less and lens-based holographic and shadow microscope systems.

microscopes, while offering a large FoV, have a limited resolution of 10 μm [69] (Table 1.1). Higher resolution can be achieved using lens-based holographic microscopes, but these are bulky, expensive and have a small FoV.

In contrast, the availability of **an array of individually addressable, sub-micrometer illuminating apertures spaced by sub-pixel distances** would allow for large FoV superresolution lensless imaging in a highly compact form factor, limited only by physical limitations in how close can the source, sample and imaging sensor be placed.

With this in mind we can derive a set of desirable characteristics for such an array:

1. Individual aperture size: From Fig. 1-7(a) we can see how to achieve z_1 below 5 cm we need imaging apertures below 1 μm and ideally closer to 250 nm.

2. Aperture spacing: Subpixel sampling techniques are capable of overcoming the limitation set by the pixel size, but can't overcome the diffraction limit without the use of additional techniques such as structured illumination or relying on a nonlinear response of the sample being imaged [77]. There is therefore no gain in achieving an aperture spacing below the diffraction limit, which makes our desirable aperture spacing be at most $\lambda/2$.
3. Optical spectrum: We have discussed how a finite optical spectrum limits the temporal coherence of the source and how this affects achievable resolution (Eq. 1.4) in lensless holography. Figure 1-7(b) shows the maximum z_2 distance acceptable to reach diffraction limited resolution for a given source optical bandwidth at $1 \mu\text{m}$ wavelength. Even for highly temporally coherent beams with sub-nm optical bandwidth (achievable only with a laser source) required z_2 are below 1 mm. Since the minimum sample - camera distance is limited to about $100 \mu\text{m}$ due to the camera protecting layer, an optical bandwidth below 5 nm is desirable. As we will see, this is challenging to achieve in an integrated source.

Nevertheless, there exist computational techniques to overcome the limits set by the source temporal coherence [78], which arguably make the requirements on optical spectrum the lowest priority.

While not directly related to the imaging array characteristics, it is worth noting how the realization of a spatially coherent source that allows for small z_1 can enable lensless setups with high magnification in a small form factor. For instance, if we can achieve a z_1 of 1 mm (requiring ≈ 100 nm diameter source) a 10x magnification system can be achieved with a z_2 of only 10 mm.

One possible concern with the realization of very small imaging apertures placed close to the sample is the resulting small field of view, which is basically related to the

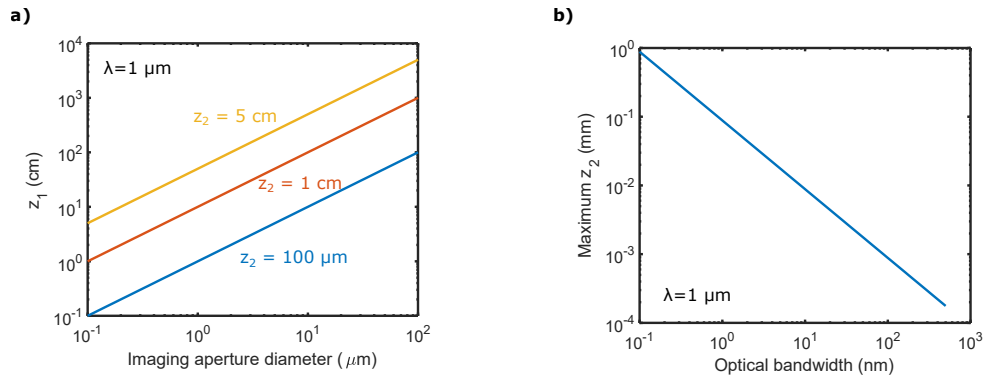


Figure 1-7: (a) Minimum source to sample distance necessary to achieve sufficient spatial coherence to have diffraction limited resolution as a function of emitter size. (b) Maximum sample to camera distance necessary to achieve a high enough temporal coherence to have diffraction limited resolution as a function of emitter optical bandwidth.

size of the illuminating spot at the sample plane. While this is a fair consideration for setups employing a single source aperture, the use of an array of individually addressable illuminating apertures overcomes such a limitation. In this case, the achievable field of view is set by the array size or the imaging sensor size (whichever is smaller).

We also notice that such an array would effectively allow for the realization of reconfigurable grayscale amplitude modulation masks by individually controlling the power emitted by each aperture, offering an easy route towards the realization of coded apertures. This has been shown to allow for high resolution imaging with a larger SNR compared to pinhole-based imaging [43, 79].

A desirable extension of the illuminating array described above would additionally allow for the control of the relative phase between different illumination apertures, which would allow for the realization of phase masks. Phase masks are desirable over amplitude masks because they do not discard any of the available power, resulting in larger SNR and therefore better imaging quality.

Summarizing our discussion above, we have shown how an **integrated array of**

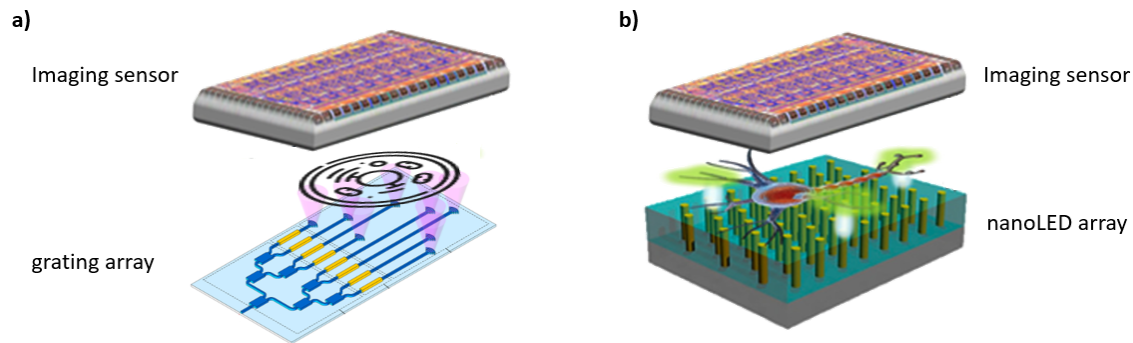


Figure 1-8: Different approaches for the use of integrated optics for lensless imaging. (a) Using a grating array. (b) Using an emitter array. Figures adapted from [80] and [81].

individually addressable sub-micrometer sized emitting apertures spaced by half the emission wavelength and with a narrow emission spectrum (ideally below 5 nm) could enable the realization of lensless imaging setups with diffraction limited resolution and a large field of view in a compact and low cost manner. This is in contrast with current solutions that have either large FoV and low resolution or high resolution but small FoV (Table 1.1).

The realization of such an array is nevertheless quite challenging. We envision the realization of such an array is possible with two distinct approaches (Fig. 1-8):

1. By employing an array of apertures where the light is externally delivered (either actively or passively) to the desired aperture (Fig. 1-8(a)). We will term this approach "grating array" because the illuminating apertures are made of gratings.
2. By employing an array of individually addressable light emitters where the aperture is the emitter itself, i.e., a nano- or micro- LED or laser array (Fig. 1-8(b)).

Below we analyze the advantages and challenges for each approach.

Grating arrays

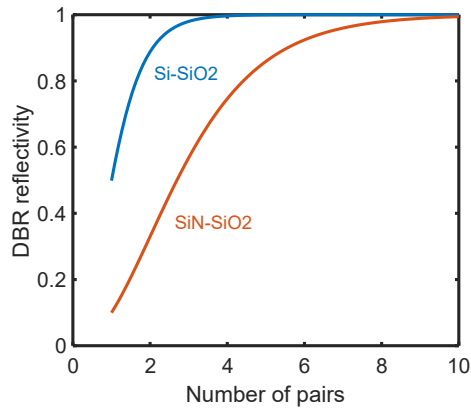


Figure 1-9: DBR reflectivity as a function of the number of pairs of high index - low index materials for a silicon platform (blue) and a silicon nitride (orange) platform.

The use of grating arrays is attractive because it allows for "processing" of the light before being launched. This includes, for example, spectral filtering (which could enable lower spectral bandwidths and thus higher temporal coherence sources) or phase control (which would allow for the realization of phase masks). This is not possible when using an emitter array unless additional components are added to the optical path.

On the flip side, grating arrays require external generation and routing of the light to be emitted. In integrated optical platforms, the emitting apertures are typically realized through grating couplers (GC) which are structures employing slabs of alternating materials to change the in-plane wavevector direction of the mode propagating in a waveguide to an out-of plane direction [88].

The greatest challenge faced by grating arrays to realize small emitting apertures is the existing tradeoff between grating length and scattering strength: to achieve maximum outcoupling of the light propagating in-plane to free space, relatively long structures are required. This results in large emitter elements which limit the minimum achievable aperture spacing and therefore resolution. We can get a feel for this tradeoff by looking

Ref (year)	Array Size	Grating pitch (μm)	Single grating size (μm)	Grating efficiency
[82] (2009)	1 x 16 (1D)	2	1.5 x 31.5	35%
[83] (2018)	1 x 512 (1D)	1.3	0.45 x 1,000	NR
[84] (2019)	1 x 16 (1D)	0.8	0.22 x 0.5 (end facet of waveguide)	$\approx 100\%$
[85] (2019)	1 x 24 (1D)	1.3	0.5 x 1,000	95% (w DBR) 46% without
[86] (2013)	8 x 8 (2D)	9	3 x 2.8	51 %
[87] (2018)	8 x 8 (2D)	11.2	5 x 2	NR

Table 1.2: Grating arrays with the smallest grating spacing demonstrated in the literature. NR = Not Reported.

at the reflectivity of a first order Distributed Bragg Reflector (DBR) [89]⁵:

$$R = \left(\frac{n_2^{2N} - n_1^{2N}}{n_2^{2N} + n_1^{2N}} \right)^2 \quad (1.7)$$

Above, n_1 and n_2 are the refractive indices of the two alternating materials and N is the number of repeated pairs of alternating materials. Figure 1-9 shows the DBR reflectivity as a function of the number of repeated pairs for a silicon (blue) and silicon nitride (orange) based DBR. We can see how higher refractive index contrast ($|n_2 - n_1|$) results in a smaller number of periods required. Assuming that a period length is about $\lambda/2$, we can see how to achieve high efficiency out of plane emission at 1.55 μm wavelength we need at least 3.5 μm ($=4\lambda/2$) for a silicon platform and 6 μm ($=10\lambda/2$) for a silicon nitride platform.

Additionally, the greatest limitation for achieving closely packed emitting apertures is

⁵DBRs are not exactly the same as grating couplers, but they are based on the same principle and we can therefore use them to get an order of magnitude approximation of the length scales necessary to achieve a given out-coupling efficiency.

the need for routing the light to the apertures themselves. Such routing requires the use of optical waveguides, which need to be spaced by several wavelengths in order to limit crosstalk effects. As a consequence, even for 1D emitting arrays with a small number of components the minimum spacing demonstrated in the literature has been limited to 800 nm, with typical values on the order of 2 - 5 μm (Table 1.2). The use of waveguide superlattices engineered for low crosstalk can offer a solution to this problem, but long propagation lengths are required to achieve low enough crosstalk [90–92].

Emitting arrays

Emitting arrays are another approach to the realization of closely spaced emitting apertures. Their main advantage is that the array element both generates and emits the light, so there are no requirements for external light coupling and routing, making the system packaging much simpler compared to grating arrays.

To achieve an array of closely packed emitters it is of course necessary to realize light sources (be it LEDs or lasers) with dimensions smaller than the desired spacing. Therefore, for our needs this means realizing an array of individually addressed sub-micrometer sized emitters with sub-micrometer spacing.

Two main challenges arise for the realization of such an array:

1. Efficiency scaling with size: As the size of a light emitter decreases the ratio of surface area to volume increases, which makes non-radiative surface recombination effects become increasingly important and result in reduced quantum efficiency [93–95]. For instance, reducing the diameter of a blue InGaN LED from 30 μm to 1 μm results in a 50% reduction in quantum efficiency [96]. It is therefore challenging to realize efficient sub-micrometer emitters with high output powers.
2. Electrical contact and routing: Obtaining reliable electrical contact to sub-micrometer sized emitters is technically challenging. Such difficulty is increased when trying

Ref (year)	Material	Array Size	Emitter size (nm)	Emitter pitch	Single emitter power	Emission wavelength	λ /Pitch
[97] (2020)	GaN	6 x 6	400 x 400	0.8 μm	$\approx 1 \mu\text{W}^\dagger$	500 nm	1.6
[97] (2020)	GaN	2 x 32	200 x 200	0.4 μm	$\approx 100 \text{ nW}^\dagger$	500 nm	0.8

Table 1.3: NanoLED arrays reported in the literature. † Estimated from internal quantum efficiencies mentioned in [97] assuming 100% extraction efficiency.

to fabricate individual electrical lines each addressing a sub-micrometer spaced, sub-micrometer size emitter without occluding nearby emitters, without causing crosstalk (where addressing one emitter causes another emitter to also partially turn on) and without shortcircuiting other contact lines [97]. This is particularly challenging in III-V material platforms, which have a limited number of interconnect layers and require integration with CMOS driving electronics. This limits practical emitter pitches to 5 μm and above for large scale systems [98, 99]. Alternatively, realizing these emitter arrays monolithically in a CMOS process would allow to take advantage of the large number of interconnect layers (> 7) available as well as realize the driving electronics on-chip, offering a viable way of achieving large scale, addressable sub- μm apertures.

As a result of these technical challenges, to the best of the author’s knowledge there is only one demonstration of a nanoemitter array with individually, electrically addressable emitters reported in the literature, made of gallium-nitride (GaN) (Table 1.3) [97]. Significant efforts have been devoted to the development of nanoLEDs and nanolasers, but most demonstrations have been limited either to single emitters or rely on optical pumping which makes individual addressing very challenging [100, 101].

The use of nanoLED arrays for high resolution imaging has been discussed previously, mostly based on simulation due to the technological limitations described above [102, 103].

Experimental demonstrations of lensless imaging using an array of microLEDs have been reported in the literature, but their resolution is limited due to the large size of the emitters, which is about $5 \mu\text{m} \times 5 \mu\text{m}$ [34, 104, 105].

Recap

In this section we have introduced lensless imaging, a technique to achieve high resolution and large field of view imaging without the use of lenses. In these schemes, sub-pixel movement of the light source can overcome the resolution limitations set by the finite pixel size of the imaging sensor and enable diffraction limited resolution. Realizing arrays of individually addressable sub- μm size emitters with sub- μm spacing arises as a promising approach to achieve sub-pixel sampling without moving parts and in a compact form factor. Integrated optics arises as a natural technology to realize such arrays due to the sub-100 nm resolution, low cost and scalability.

In the next section we switch gears and discuss the use of integrated optical systems for a different application: optical beamforming, i.e, the generation of optical beams with tunable orientation and shape. As we will see, some similarities arise in the requirements for lensless imaging and optical beamforming - mainly the need for large arrays of closely spaced emitting apertures. The main difference is that while imaging arrays do not require coherence between individual emitters (although spatial and temporal coherence is required for each emitter), beamforming requires high spatial and temporal coherence to be established between all emitters.

1.4 Optical beamforming with OPAs

The ability to control the phase (an optionally amplitude) of an array of closely spaced radiators enables the control of its emission properties, mainly emission direction and beam shape. This is commonly termed phased array beamforming, and it has been

widely explored in the radiofrequency (RF) domain for applications spanning wireless communications, radar, weather research and human - machine interfaces.

Since the realization of high performance phased arrays requires radiator spacing on the order of the wavelength, the realization of *optical* phased array beamformers (dealing with wavelengths on the micrometer level) has been challenging due to the high fabrication resolution required.

Nevertheless, the advent of integrated photonics and its associated high fabrication resolution (well in the sub-micrometer regime) has made the realization of OPAs a possibility. Extending phased arrays to the optical domain has a plethora of applications including free space optical communications, LIDAR (an extension of radar to optical frequencies that enable much higher resolution), imaging and sensing [48]. All these applications take advantage of the fast and accurate control of the shape and direction of the beam generated by an OPA. This is in contrast with currently available beamforming systems which rely on bulky, slow, expensive and failure-prone mechanical beam steering.

1.4.1 OPA theory and operating principle

The operating principle of OPAs is exactly the same as for RF phased arrays, in which controlling the relative phase between radiating elements controls the emission properties of the array. We can denote the electric field emitted by each element in the array (indexed arbitrarily by (p, q)) in phasor notation as: $E_{p,q} = |E_{p,q}|e^{j\psi_{p,q}}$. With this, the angular power radiation pattern of the array in the far field $|U(\theta, \phi)|^2$ can be expressed as [48, 106]:

$$|U(\theta, \phi)|^2 = |A(\theta, \phi)|^2 |F(\theta, \phi)|^2 \quad (1.8)$$

Where $F(\theta, \phi)$ is the radiation pattern of a single antenna (we are assuming all the

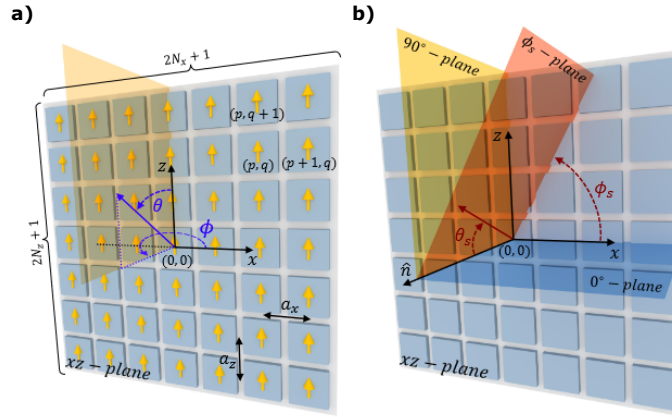


Figure 1-10: A nominal 2D OPA. (a) Coordinate system used for the array factor A (Eq. 1.10). (b) Coordinate system used to describe beam steering. Figure reproduced from [106].

antennas are identical) and $A(\theta, \phi)$ is the so called array factor, which is given by:

$$A(\theta, \phi) = \sum_{p,q} E(p, q) e^{-jkx_{p,q} \sin\theta \cos\phi} e^{-jkz_{p,q} \cos\theta} \quad (1.9)$$

Above, $x_{p,q}$ ($z_{p,q}$) is the x position (z position) of the antenna in the array plane and ϕ and θ are the polar coordinates of the array (Fig. 1-10(a)).

For a uniformly spaced 2D OPA with spacings a_x and a_z in the x and z directions respectively we can write:

$$A(\theta, \phi) = \sum_{p,q} E(p, q) e^{-jkpa_x \sin\theta \cos\phi} e^{-jkqa_z \cos\theta} \quad (1.10)$$

A nominal 2D OPA with uniformly spaced elements is depicted in Fig. 1-10 along with the coordinate system we are using. To describe the beam steering we use another coordinate system denoted by (θ_s, ϕ_s) as shown in Fig. 1-10(b).

To achieve beam steering towards a particular direction, a phase gradient along that

direction needs to be generated. For example, to achieve beam steering in the 90° plane in Fig. 1-10(b) we need to synthesize a linear phase gradient in the z direction.

For simplicity we will study beam steering along one direction, corresponding to the 0° plane in Fig. 1-10(b). This is equivalent to studying a 1D OPA, and extension to a 2D OPA is straightforward by adding an additional dimension.

The steering angle θ_s can be found from the generalized law of reflection as [106]:

$$\sin\theta_s = \frac{\Psi'(x)}{k} = \frac{\Delta\Psi}{ak} \quad (1.11)$$

Above, $\Psi'(x)$ is the phase gradient along the x direction and $k = 2\pi/\lambda$ is the wavevector of the emitted beam. Since the array is made of discrete elements, we can approximate $\Psi'(x) = \Delta\Psi/a$, where $\Delta\Psi$ is the phase difference between adjacent radiating elements.

Achieving perfect beam steering with perfect directionality (i.e, an infinitely high angular directivity) would require an infinite array ($n \rightarrow \infty$) of infinitely close elements ($a \rightarrow 0$) with perfect phase and amplitude control. This is of course not practically achievable, which results in non-idealities and performance limitations.

The first one of such limitations is the existence of side lobes: the finite emitter array size results in relatively strong emission in other angular directions besides the desired one due to diffraction from the square aperture. This is analogous to windowing effects, where applying a square window to a signal results in a *sinc* pattern that generates side lobes. It is important to notice that the strength of such side lobes is not affected by the number of elements or size of the array (Fig. 1-11(a)). Circular and/or Gaussian apodization of the array (where different elements emit a different power) can be used to reduce the side lobe strength by 2 orders of magnitude [106], but require either amplitude control or the addition of transmission masks.

The finite pitch between adjacent emitters also generates radiation lobes due to aliasing and are usually termed "grating lobes". Such lobes are of a similar strength as the main lobe, and as such limit the field of view of the OPA. We can use the grating diffraction equation to describe the FoV of the OPA:

$$\sin\theta_s^{(m)} = \sin\theta_s + m\frac{\lambda}{a} \quad (1.12)$$

Above, m is the order of the grating lobe. The FoV of the OPA is usually defined as twice the angle $\theta_{s,max}$ at which there are no grating lobes in the desired FoV. We thus need to ensure that the first positive and negative lobes ($m = \pm 1$) are outside the FoV: $|\sin\theta_{s,max} \pm \lambda/a| > |\sin\theta_{s,max}|$. This basically reduces to a requirement in the maximum element spacing a_{max} to achieve a given FoV $\theta_{s,max}$:

$$a_{max} = \frac{\lambda}{2|\sin\theta_{s,max}|} \quad (1.13)$$

$$FoV = 2\theta_{s,max} \quad (1.14)$$

To achieve $\theta_{s,max} = 90^\circ$ we need $a = 0.5\lambda$ (Fig. 1-11(b)). It is important to mention here that the FoV of the OPA is also affected by the radiation pattern of a single emitter. Even if the FoV associated to the array factor A allows for large θ_{max} , if the emitter cannot radiate in that direction (i.e, $F(\theta_{max})$ is small) then there won't be efficient emission from the array (Eq. 1.8).

Another important practical limitation is on the achievable angular resolution. A limited aperture size results in the generated beam having a finite angular beam width, which is given by [82]:

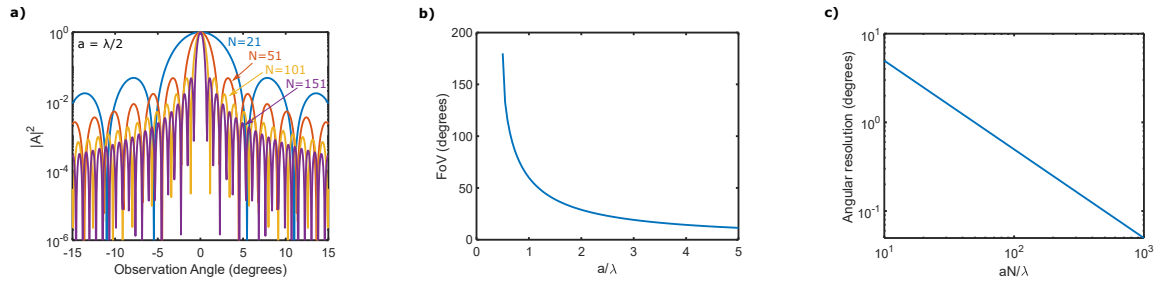


Figure 1-11: OPA beam steering performance. (a) Array factor magnitude as a function of observation angle for a 1D OPA with different number of elements spaced by $\lambda/2$. (b) The OPA field of view as a function of a/λ . (c) The OPA angular resolution as a function of aN/λ .

$$\Delta\theta_s \approx \frac{0.866\lambda}{aN\cos\theta_s} [rad] \quad (1.15)$$

We can see how to achieve high angular resolution (small $\Delta\theta_s$) we need a large aperture extent $a \cdot N$ (N is the total number of emitters in one direction). From Fig. 1-11(c) we can see how for an angular resolution below 0.1 degree we need $aN/\lambda > 500$, which for $a = \lambda/2$ (for 90 degree FoV) requires $N > 250$. The fact that larger arrays result in smaller beamwidth is also observed in Fig. 1-11(a).

The non idealities described above are summarized in Table 1.4. These are the main performance limitations for practical OPAs. Limitations in the phase and amplitude control of individual emitters can also result in radiation lobes that can compromise operation [106].

It is important to note that our theoretical discussion has been limited to beam steering along 1 axis (1D). As already mentioned, 2D beam steering can easily be achieved by using a 2D phased array (such as the one depicted in Fig. 1-10) and imparting a different phase change $\Delta\Psi$ in each direction (Fig. 1-12(a)). Nonetheless, it is worth noting that there are alternative approaches to the realization of 2D beam steering with

Performance metric	Formula
Field of View	$a_{max} = \frac{\lambda}{2 \sin\theta_{s,max} }$
Angular resolution	$\Delta\theta_s \approx \frac{0.866\lambda}{aN\cos\theta_s}$
Side lobe	-10 dB square array -20 dB circular array -30 dB Gaussian window [106]

Table 1.4: Summary of OPA performance metrics and formulas.

1D phased arrays.

The most popular approach relies on the use of emitters that have a wavelength-angle dispersion characteristic in the direction perpendicular to the 1D array, such that steering in that other dimension is achieved by tuning the wavelength of the light source [107] (Fig. 1-12(b)). This approach significantly reduces the number of phase shifters required for beam steering and also relaxes the requirements on the size of the individual emitters, but requires high power, stable tunable light sources and broadband components (e.g., wavelength insensitive beam splitters), which can be hard to realize in integrated platforms.

An alternative approach is the use of optical switches to steer the light to emitters with a different emission angle in the direction perpendicular to the array [108]. While successful, the scalability of this approach is of concern since it requires an OPA unit cell with as many emitters as emission angles in the perpendicular direction. Dynamic configuration of the emitter emission angle in the perpendicular direction through the use of heaters has also been demonstrated, but thermal crosstalk between neighboring emitters is of concern [109]. Finally, 2D beam steering can also be achieved by solely tuning the wavelength (without an active phase controller) although fabrication imperfections result in large side lobes that compromise operation [110].

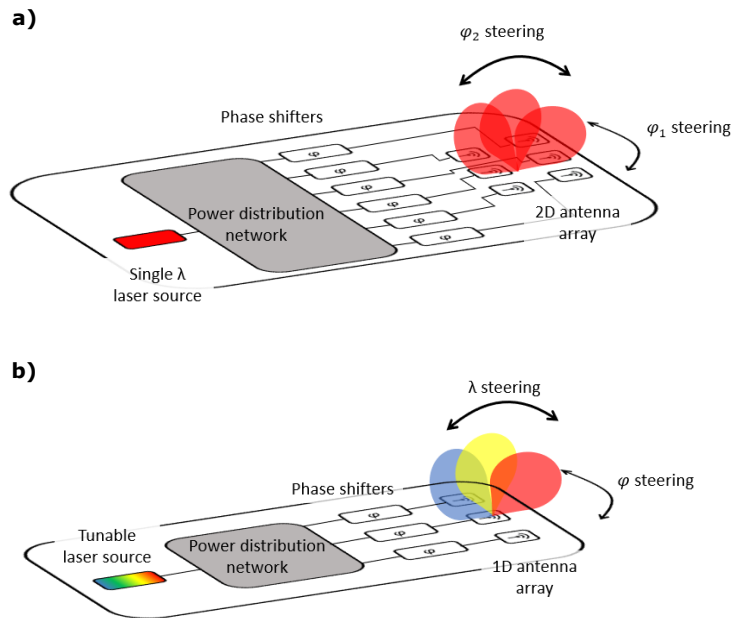


Figure 1-12: OPA 2D beam steering configurations. (a) 2D steering is achieved by using a 2D array of emitters and controlling the phase difference in the two directions. (b) 2D steering is achieved with phase difference in one direction and by wavelength tuning in the other.

1.4.2 OPA beamformers for LIDAR and Free Space Optical Communications

As mentioned above, optical beamforming systems have potential applications in a wide range of fields, with the two most promising arguably being LIDAR and free space optical communications (FSOC). It is therefore useful to describe how OPAs can be used in these two applications and to derive some performance requirements for each scenario.

Beamforming for LIDAR

LIDAR is the natural extension of RADAR from RF frequencies to optical frequencies. As such, the goal of LIDAR is to accurately identify and locate targets or obstacles. Due to the smaller wavelength, LIDAR offers a much higher resolution compared to

RADAR, and is considered a necessary technology for the realization of autonomous driving [111, 112]. Other applications include robotics, topography measurements and environmental monitoring.

Here, we will focus on automotive LIDAR because it is the most promising application and arguably the one with the most demanding performance requirements. The basic idea of automotive LIDAR is depicted in Fig. 1-13(a), where a light pulse is sent in a specific direction, bounces back from the scene/obstacle and is detected. By measuring the time it takes the light to travel to the object and back we can measure the distance. Repeating this measurement in different directions allows to reconstruct a full scene in the form of a point cloud (bottom of Fig. 1-13(a)).

From this, it is clear that beamforming is required to be able to steer the light beam in different directions. Current commercial LIDAR products use mechanical beam steering which is costly, bulky, slow and failure-prone. As such, the use of integrated OPA beamformers is highly attractive due to their compactness, (potentially) low cost and the fact that they don't rely on mechanical steering. This has been long recognized and there are several companies pursuing the realization of OPA beamformers for automotive LIDAR [113–115] with commercial products already in the market.

Several groups have attempted to derive a set of requirements for LIDAR beamformers in terms of FoV and resolution, with differing conclusions as summarized in [111]. We will use the requirements in [116] (Table 1.5), which try to ensure correct detection of cars, motorcycles, bicycles and pedestrians at about 200 m range (which is considered the required range for LIDAR systems).

Besides FoV and resolution, another important performance metric is the speed at which the beamforming system can operate. For LIDAR applications this is mainly set by the desired frame rate of the systems. It is generally assumed that a frame rate of at least 10 fps is necessary, with 25 fps being a safe target [111, 116]. The most

	LIDAR system requirements [116]
Field of View	14° × 3.5°
Angular resolution	0.1° × 0.1°
Frame rate	25 fps
Sampling rate	122.5 kHz

Table 1.5: Requirements for OPA-based LIDAR systems for automotive applications.

demanding LIDAR scheme in terms of sampling frequency is a point scanning LIDAR (where a beam is synthesized at every pixel of the scene), which has a speed requirement given by $f * (FoV_x/res_x) * (FoV_y/res_y)$ (where f is the frame rate, FoV is the angular field of view and res is the angular resolution). The necessary sampling frequency for the OPA LIDAR system we are considering is shown in Table 1.5.

Besides these considerations, power consumption, cost and robustness are all important performance metrics.

Unfortunately, none of the companies developing OPA-based LIDAR systems have performance characteristics publicly available. A summary of the performance of commercially available LIDAR systems can be found in [117]. Commercial solid state LIDAR solutions (which do not rely on OPAs, but on a variety of other techniques including MEMS-based scanning and focal plane arrays) with FoV as large as 90°x20° are available, but have resolution limited to 0.2° or above (Table 3 in [117]).

Beamforming for FSOC

Free space optical communications, where an optical beam transmitted through the atmosphere is used to establish communication between a receiver (RX) and a transmitter (TX) terminal (Fig. 1-13(b)), is an established solution used mostly in deep space communication and inter-satellite links [119] due to the significantly higher bandwidth compared to RF wireless solutions. In the last few years, FSOC has also been explored for

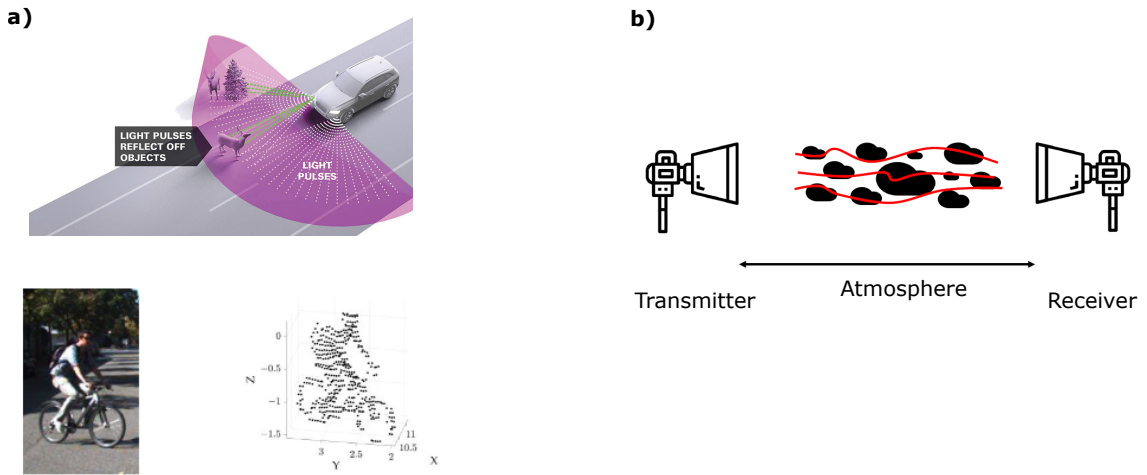


Figure 1-13: OPA beamforming applications. (a) Automotive LIDAR. Top image is from [118], and bottom is from [111]. (b) Free space optical communications.

use in terrestrial links, which poses additional challenges due to the strong atmospheric effects [120].

While highly advantageous, FSOC links require line of sight (LoS) between the transmitter and receiver to be maintained at all times. Due to the small sizes of optical beams, such LoS needs to be actively maintained through what is generically known as pointing and tracking (PaT) [121]. Conventional PaT systems employ mechanically moving elements such as gimbals or mirrors, which are bulky, slow, expensive and significantly increase the size, weight and power (SWaP) of the system, an important metric especially for space systems.

OPA-based beam steering arises as a natural, highly attractive alternative to mechanical PaT with significantly reduced SWaP and power consumption [50, 122, 123]. Not only this, but OPAs can also be used to compensate for distortions of the optical beam due to atmospheric effects, something which is usually addressed with the use of adaptive optics via deformable mirrors, which are slow and have limited resolution [124, 125].

To the best of my knowledge, there is no work discussing specific performance require-

ments for OPA-based FSOC systems in terms of FoV and angular resolution. We will try to derive them here, and we will limit our discussion to terrestrial FSOC systems due to their novelty and their potential for high impact by delivering last-mile connectivity [126].

Assuming that the RX and TX are initially aligned when installed, the necessary FoV should be limited to the maximum relative movement that can occur between RX and TX. There are two main sources of such relative movement:

1. The twist and sway of the pole or structure that holds the terminal, which is caused mostly by wind and thermal fluctuations. Unfortunately, accurate reports of twist and sway movements in structures likely to hold terrestrial FSOC systems, which include light poles, traffic lights and poles on top of buildings are challenging to find. A study of 20 m tall lighting poles found that displacements up to 0.43° from vertical are possible [127], and in buildings it is up to 0.28° [128].
2. Beam wander due to atmospheric effects. Propagation of an optical beam results in a variety of effects, including optical attenuation, beam scintillation (fluctuations in the optical power at the receiver aperture) and beam wander (random displacement of the centroid of the propagated beam) [129,130]. The FoV of the OPA should be large enough to compensate for beam wander. The variance of the beam centroid position due to beam wander in the case of a collimated beam is given by [130,131]:

$$\langle r_c^2 \rangle = 0.97 C_n^2 D_{TX}^{-1/3} L^3 \quad (1.16)$$

Above, D_{TX} is the transmitter aperture diameter, L is the link distance and C_n^2 is the refractive index structure parameter, which is a measure of the turbulence strength of the atmosphere and is between $10^{-13} \text{ m}^{-2/3}$ (strong turbulence) and $10^{-17} \text{ m}^{-2/3}$ (weak turbulence).

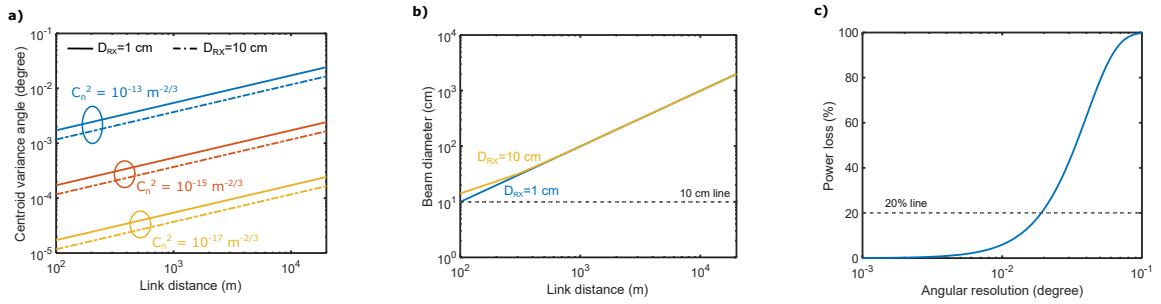


Figure 1-14: (a) Angular variance of the beam centroid due to beam wander as a function of link distance for different turbulence strengths and receiver aperture diameters. (b) Beam diameter at the receiver plane as a function of link distance for different receiver aperture diameters. (c) Worst case power loss with respect to perfect alignment of a Gaussian beam as a function of the angular resolution of the OPA.

Using $\theta_{FOV} = 2 \operatorname{atan}(\sqrt{2 \langle r_c^2 \rangle} / L)$ we can extract the necessary angular FoV⁶. This is plotted in Fig. 1-14(a) for different turbulence strengths and transmitter aperture sizes. We can see how the required FoV increases with increasing link distance and turbulence strength, and decreases as the aperture size increases. An upper bound for the required FoV is around 0.1° .

From this, it is clear that the sway of the structures that hold the TX and RX terminals is the dominant effect that sets the required FoV. **FoV below 3° should be enough to compensate for such movements.**

The requirements on angular resolution are mostly set by the required positioning accuracy of the beam with respect to the receiver. Maximum received power will be achieved when the generated beam is exactly aligned to the receiver aperture. We can get an estimate of the required angular resolution by setting a limitation on the maximum power loss we are willing to tolerate compared to perfect alignment.

Assuming that the OPA-based FSOC system uses a lens that effectively collimates

⁶Notice we are using the 2σ variance in the centroid radius, which allows for misalignment compensation over 95% of the time.

the beam generated by the OPA (which is in general true when the OPA is placed at the focal point of the lens), the beam radius as a function of distance can be obtained from [132]:

$$R^2(L) = R_{TX}^2 + \frac{L^2}{k^2} \quad (1.17)$$

Above, R is the beam radius at a distance L from the transmitter, $k = 2\pi/\lambda$ is the wavevector and R_{TX} is the radius of the transmitter aperture lens.

As shown in Fig. 1-14(b), for typical apertures in the order of 1 - 10 cm and link distances above 100 m the beam diameter at the receiver is significantly larger than the receiver aperture. If we assume that the beam is Gaussian with a beam waist at a distance L given by Eq. 1.17, we can calculate the loss in the power integrated by the RX aperture as a function of misalignment using:

$$\eta_{misalignment} = \frac{e^{-\frac{2r_0^2}{w^2}} \int_0^{R_{RX}} \int_0^{2\pi} r \exp(-2r^2/w^2) \exp(4 r_0 r \cos(\theta)/w^2) d\theta dr}{(\pi/2)w^2(1 - e^{-\frac{2r_0^2}{w^2}})} \quad (1.18)$$

r_0 is the distance between the center of the Gaussian beam and the center of the RX aperture, w is the beam waist radius at the RX plane and R_{RX} is the radius of the receiver aperture. We can easily relate the worst case misalignment distance r_0 to angular resolution through $r_0 = \tan(\Delta\theta/2)L$ where $\Delta\theta$ is the angular resolution.

Joining the above equations we can calculate the fraction of power lost due to the finite angular resolution. This is shown in Fig. 1-14(c). We can see how to limit the losses to 20% (10%) we need to ensure an angular resolution of 0.02° (0.01°). It is worth mentioning that several approaches have been reported to reduce the sensitivity of FSOC systems to angular and positional misalignment. Of particular interest is the

	FSOC system requirements
Field of View	$3^\circ \times 3^\circ$
Max misalignment loss	1 dB (20%)
Angular resolution	$0.02^\circ \times 0.02^\circ$
Sampling rate	100 kHz

Table 1.6: Requirements for OPA-based FSOC systems.

use of multimode fibers for capturing the light at the receiver, where it has been shown that the mode diversity can increase robustness to fluctuations caused by atmospheric effects [133–135]. Such multimode fiber can be combined with a photonic lantern (a device that converts a multimode input to an array of single mode outputs) [136, 137] to realize coherent FSOC systems with increased sensitivity and robustness [138–140].

Finally, the necessary speed at which the beamforming system needs to operate is set by (1) the vibration frequency of the twist / sway of the structure that holds the terminal (which is slow and requires speeds on the order of 1 - 20 Hz [141]), and (2) the coherence time of the atmosphere. The coherence time of the atmosphere is defined as the time period τ at which we can consider that the atmosphere is *still*, i.e, we can consider that the properties of an optical beam propagating through the atmosphere (phase and amplitude) are constant for a time τ . Atmospheric coherence times are between 1 ms and 10 ms [119, 142], therefore requiring beamforming systems with low speeds on the order of 10 - 100 kHz.

The discussion above allows us to derive a full set of performance requirements for an OPA system for FSOC applications, which are summarized in Table 1.6.

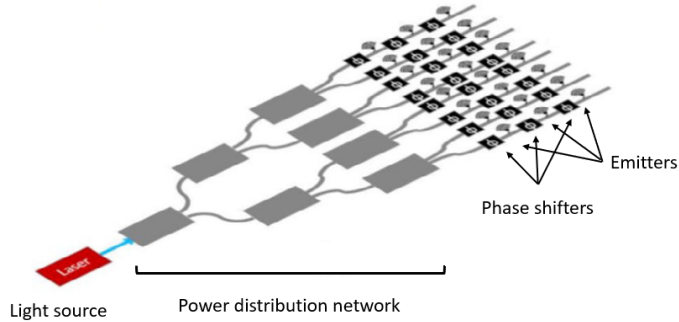


Figure 1-15: The generic architecture of an OPA. Adapted from [112].

1.4.3 Realization of integrated OPAs

As we have reviewed above, the main goal of an OPA is to impart a relative phase change to the (optical) signal being fed to the different emitters. Figure 1-15 depicts a generic OPA architecture with its main components:

1. Light source: provides the coherent signal that will subsequently be manipulated to achieve the desired phase difference between the elements.
2. Power distribution network: evenly splits the power coming from the light source to the different individual emitters.
3. Phase shifters: impart the desired phase to the different signals to be fed to the individual emitters.
4. Emitters/Antennas: Radiate the power coming from the phase shifters to achieve the desired radiation pattern.

Below we discuss each component separately and derive necessary performance metrics for the realization of practical OPAs for a variety of applications, with a focus on LIDAR and FSOC.

Light Source

A laser source is necessary to generate the coherent optical signal that will be fed into each of the emitters. The main requirements for such sources are high output power, small linewidth and high stability. In schemes relying on wavelength tuning for steering in one direction, then an added requirement is for the source to be tunable over a broad wavelength range with accurate and stable control of the emission wavelength.

To achieve small form factor and low cost OPA systems it is necessary for the source to be integrated. In OPA systems based on Si or SiN it is not possible to have a monolithically integrated light source, so the light generated in a separate chip (usually fabricated in a III-V material platform) is coupled to the OPA chip through grating or edge couplers [88, 143] or hybrid integration [144]. Another option that has been explored is the use of erbium doping of silicon waveguides to realize on chip lasers [145]. Truly monolithic OPA systems with an integrated light source have been demonstrated in InP [146].

Power distribution network

The light generated by the laser source needs to be distributed to the N different emitter elements. This is usually achieved by one of three different approaches: (1) directional couplers, where evanescent coupling between two waveguides achieves the power splitting [86]; (2) multimode interference (MMI) splitters, which consist of a multimode waveguide section and rely on the self-imaging principle to achieve even power splitting [147] and (3) star couplers, which are similar to MMIs but rely on curved waveguides [148].

Each approach has its associated advantages and drawbacks. Directional couplers can only achieve 1x2 power splitting, which makes it necessary to serially cascade multiple couplers to achieve a large splitting ratio and therefore increase the required area. MMI splitters can in principle achieve any desired power splitting ratio, but have relatively

large footprints. Star couplers are the most compact devices when a large splitting ratio is required (i.e, for an OPA with a large number of emitters) but have in general a higher insertion loss than MMI or directional couplers.

The main requirements for the power distribution network are low insertion loss, compactness and accurate splitting ratio. In schemes relying on wavelength tuning it is necessary to ensure that the power distribution network can work over the entire operating wavelength range.

Emitters/Antennas

Along with phase shifters, the properties of the individual emitters as well as the array characteristics are the most important considerations affecting the performance of an OPA.

As we reviewed in Section 1.4.1, the spacing and aperture size of the emitter array determine the FoV and angular resolution of the system. In Section 1.4.2 we derived requirements for OPA-based systems with applications in LIDAR and FSOC systems. We can use Eq. 1.13 to derive the necessary antenna spacing (we assume a wavelength of 1550 nm, which is the typical wavelength used in these systems) and Eq. 1.15 to obtain the necessary array size. These are shown in Table 1.7 for the different system requirements we derived in Section 1.4.2.

We can see how large arrays with thousands of elements are required. Realizing such large arrays is challenging in a range of aspects, including system size and power consumption (which we will analyze in detail in the next section).

In terms of spacing, similarly to what we discussed in the context of grating arrays for imaging (Section 1.3.3), two main concerns arise:

1. Crosstalk between neighboring elements: Optical waveguides do not fully confine the optical mode, which results in the presence of an evanescent field outside of

	LIDAR [116]	FSOC
Field of View	$14^\circ \times 3.5^\circ$	$3^\circ \times 3^\circ$
Angular resolution	$0.1^\circ \times 0.1^\circ$	$0.02^\circ \times 0.02^\circ$
Sampling rate	122.5 KHz	100 kHz
Emitter spacing	$6.3 \mu\text{m} \times 25 \mu\text{m}$	$29 \mu\text{m} \times 29 \mu\text{m}$
Array size	120×31 $N = 3,720$	132×132 $N = 17,424$

Table 1.7: Requirements in emitter spacing and array size for OPA-based FSOC and LIDAR systems.

the waveguide core. This can cause unwanted coupling between optical structures spaced by distances on the order of a few wavelengths. This means that the emission properties of each emitter are not independent, causing a degradation in the radiation pattern of the OPA [149]. This problem is even stronger in grating structures due to grating-assisted evanescent coupling [150]. It is therefore necessary to design emitters with reduced crosstalk characteristics, as well as ensuring no crosstalk in the optical routing network that delivers the light to the emitters.

2. Achieving efficient radiation efficiency: Achieving efficient emission is necessary for maximizing the sensitivity of the system. As we saw in Section 1.3.3, a tradeoff exists between antenna size and radiation efficiency (Fig. 1-9), requiring antenna sizes on the order of a few wavelengths for close to unity emission efficiency. Additionally, we need to ensure that most of this radiation is directed upwards, which requires gratings with broken vertical symmetry [50]. Otherwise, unwanted reflections and crosstalk result in distorted radiation patterns and lower sensitivity.

Besides these considerations, it is important to note that the radiation pattern of a single antenna affects the radiation pattern of the OPA (Eq. 1.8). This is most important when looking at large FoV OPA systems: even if the array parameters allow for a large

FoV, no efficient emission at a particular angle will be obtained if the individual emitter cannot radiate in that direction (in other words, $F(\theta) \approx 0$ results in $U(\theta) \approx 0$ regardless of the array factor value $A(\theta)$, Eq. 1.8). It is therefore necessary to design emitters with efficient radiation over a broad angular range.

Finally, one additional consideration when using an OPA as a receiver (which can be done both in LIDAR and FSOC applications) is the achievable fill factor, which is defined as the ratio of area occupied by an emitter/receiver antenna to the full area of the OPA. Clearly, a less than unity fill factor results in a loss in the received signal by an amount equal to the fill factor (since we cannot collect light impinging in areas of the OPA where there is no receive antenna). Thus, maximizing OPA fill factor results in increased collection efficiency and higher sensitivity. Nevertheless, this has to be balanced so that a limited crosstalk between antennas is achieved.

Phase shifters

Phase shifters are arguably the most important element in OPAs, since they impart the relative phase difference between elements necessary to synthesize the desired beam. Since in typical OPA architectures a phase shifter is needed for each individual emitter, limitations in the performance of current state of the art phase shifters are the main bottleneck for the realization of large scale OPA systems, as they are usually the component that occupies the most area and consumes the most power.

The most important performance metrics for phase shifters used in OPA systems are:

- Phase change range: To achieve arbitrary beam steering, we need to ensure that the phase shifter can impart a phase shift in the full range between 0 and 2π in a continuous and highly linear fashion.
- Power consumption or V_π : Minimizing power consumption is paramount to realize scalable OPA systems that can be deployed in typical scenarios. Some phase shifters

(thermo-optic and carrier injection based, for example) have an associated DC power consumption, in which case the power required to achieve a π phase shift P_π is usually used as performance metric. For phase shifters without an associated DC power consumption (such as carrier depletion or capacitor based), the important figure of merit becomes V_π , which is the voltage required to achieve a π phase shift. Phase shifters without DC power dissipation still have dynamic power dissipation, which can be approximated as $P_{dyn} = fCV_\pi^2$, where f is the operation frequency. As we saw, operation frequencies for typical OPA systems are on the order of 100 kHz, which results in very low dynamic power consumption for typical devices. In these cases, V_π is more informative as a performance metric.

For LIDAR applications we can derive a limitation for P_π based on current commercially available systems, which have power consumption on the order of 10 W [151]. Assuming about 30% of the power is consumed by the phase shifters in the OPA system (the rest is consumed in the laser source, the control electronics and processing), this results in a total power consumption limit of about 3 W for the phase shifter ensemble. For FSOC applications we will use a total power consumption limit of 71 W, which corresponds to the maximum power that can be delivered by Power over Ethernet (PoE). This basically ensures that our FSOC terminal can operate without a dedicated power supply, something which is highly desirable for ease of deployment. As in the case of LIDAR, we will assume that 30% of the total power is consumed by the phase shifters, resulting in a total power consumption of 21.3 W. Table 1.8 summarizes the phase shifter P_π requirements for the different systems we have analyzed.

The limitations in V_π come mostly from the maximum voltage that can be easily generated in modern, deeply scaled CMOS processes (which have transistors with

low supply voltages $V_{DD} < 2.5 \text{ V}$). This is typically considered to be between 5 - 10 V [152].

- Area: Minimizing the area occupied by the phase shifter is essential to realize practical OPA systems. Lower area devices result in smaller chips and therefore lower cost, and also minimize device capacitance which has an advantage in terms of power dissipation.

Typical reticle sizes in CMOS processes are limited to below 3 cm x 3 cm, set by the maximum field of view of lithography tools used for fabrication. If we allocate 30% of the chip to the phase shifter network, this translates into a maximum area of 2.7 cm². Table 1.8 shows the maximum area occupied by a single phase shifter for the different systems we have analyzed. This includes both the physical area of the device as well as any spacing between phase shifters (which is necessary, for example, to limit crosstalk).

- Insertion loss: Minimizing the optical power loss of the signal going through the phase shifter ensures efficient systems with high sensitivity.
- Residual amplitude modulation: Besides a phase change, phase shifters can affect the amplitude of the input optical signal. This is an undesired behavior that results in the generation of side lobes (as different elements will be radiating signals with a different amplitude), so it should be minimized.
- Speed: As we have reviewed in Section 1.4.2, we require phase shifters that can operate at a minimum speed of 100 kHz both for LIDAR and FSOC applications.
- Crosstalk: The limitations in chip size usually result in phase shifters located very close to each other. This can cause unwanted crosstalk, where applying a phase

	LIDAR [116]	FSOC
Field of View	14° x 3.5°	3° x 3°
Angular resolution	0.1° x 0.1°	0.02° x 0.02°
Sampling rate	122.5 KHz	100 kHz
Emitter spacing	6.3 μm x 25 μm	29 μm x 29 μm
Array size	120 x 31 N = 3,720	132 x 132 N = 17,424
Phase shifter power	800 μW (3 W total)	1.22 mW (21.3 W total)
Phase shifter area	0.07 mm ² (≈ 250 μm x 250 μm)	0.015 mm ² (≈ 125 μm x 125 μm)

Table 1.8: Requirements in phase shifter power and area for OPA-based FSOC and LIDAR systems.

shift to a given device results in phase shifts in neighboring devices as well (this is particularly true for thermo-optic phase shifters). Of course, crosstalk needs to be minimized for high performance OPA systems.

As discussed in detail in Appendix A, no experimentally demonstrated phase shifter is currently able to deliver the performance requirements in Table 1.8, pinpointing an important area of research focus for the practical realization of OPA-based beam steering systems.

Packaging and control

While not the main focus of the community working on integrated OPAs, packaging considerations are of the utmost importance for the realization of practical beamforming systems. Integrating the light source, the OPA chip, the control electronics and any additional components (such as lenses or thermo-electric coolers (TECs) for temperature control) in a small form factor with low cost and high robustness is challenging and requires careful consideration of multiple aspects that are usually overlooked in academic works, such as operating temperature range and thermal fluctuations or operation under

vibrations.

Related to the point above, and as already discussed, the need for accurate control and calibration of the phase shifters poses a multifaceted challenge, both in terms of scalability and integration. In general, an OPA with N elements requires N phase shifters and, in principle, N dedicated controllers (although there are architectures that only require $2\sqrt{N}$ phase shifters [153]). In hybrid approaches, where the control electronics are fabricated in a different chip as the OPA system, the need to establish a separate electrical connection between each individual controller and phase shifter poses stringent requirements on the density of I/O connections between the two chips. This favors the use of platforms where control electronics can be monolithically integrated with the OPA system.

It is also important to acknowledge that calibration of OPA systems is unavoidable due to fabrication imperfections and drifts that result in non-ideal phase and power differences between the emitter elements. This poses an additional challenge to the realization of practical OPA beamformers and requires careful consideration as well. Multiple approaches for OPA calibration have been reported in the literature including direct measurement of phase and amplitude variations and gradient-descent optimization approaches [48].

Another important consideration when designing OPA systems is the optical power handling limitations of the different components and materials. High output powers are required to achieve high performance systems based on OPA beamformers, but there are practical limitations in the power levels that integrated optical devices can handle before nonlinear effects result in significant degradation of performance. For silicon this limit is on the order of 100 mW, while it is significantly higher for SiN due to its lower optical absorption.

1.4.4 OPA demonstrations: state of the art

Many demonstrations of integrated optical OPAs have been reported in the literature with some variation of the basic architecture we just discussed. As shown in Tables 1.9, 1.10 and 1.11 and Fig. 1-16, OPAs have been demonstrated in a variety of material platforms (with Si and SiN being the most popular) relying on a variety of phase shifting mechanisms (with thermo-optic phase shifters being the most popular due to its simplicity and its compatibility with any material platform / fabrication process).

As is apparent from Tables 1.9, 1.10 and 1.11, no demonstrated solution is capable of simultaneously achieving the required FoV and resolution for both LIDAR or FSOC applications (Table 1.8). It is worth making a few observations:

- As discussed above, thermo-optic phase shifters are the default choice, resulting in power hungry systems with poor scalability.
- Most of the demonstrations are still limited to laboratory settings, which do not require full packaging (an important challenge for real world OPA beam steering systems).
- Only two of the OPA demonstrations reported in the literature to date ([154, 155]) have more than 3,000 elements (which is the minimum required for our real world applications). Most of the demonstrations have array sizes consisting of 50 emitters or less, highlighting the scalability challenges of current OPA designs (in terms of power consumption, area and control complexity).
- Most demonstrations in the literature rely on phase shifters to achieve steering in one direction, and wavelength tuning to achieve steering on the other direction. This is because such an architecture only requires N phase shifters for an $N \times N$

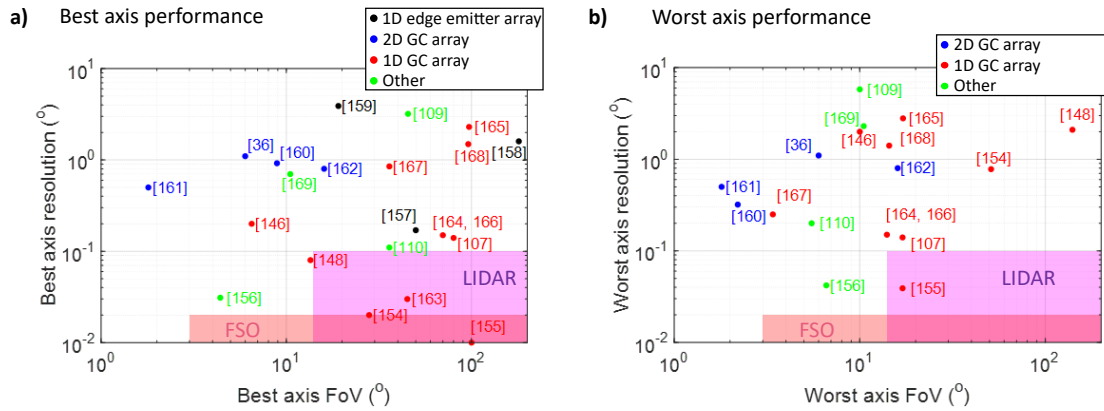


Figure 1-16: (a) Best axis and (b) worst axis performance of OPAs reported in the literature. The red shaded area shows acceptable metrics for FSO applications ($FoV > 3^{\circ}$, resolution $< 0.02^{\circ}$), and the magenta shaded area for LIDAR applications ($FoV > 14^{\circ}$, resolution $< 0.1^{\circ}$). The best steering axis is defined as the axis with a larger ratio of FoV/res .

phase shifter array, which translates into significant savings both in chip area, control electronics and power consumption. On the flip side, such a scheme requires very stable and accurate high power tunable lasers. Furthermore, fabrication variations result in uneven behavior of the emitting apertures as the wavelength is tuned, which results in unwanted side lobes that compromise operation. Schemes relying on phase shifters for steering in both directions, while more complex, can compensate for such variations and have greater flexibility in the generated radiation pattern.

1D array of edge emitters (1D beam steering)

Ref. (year)	λ (μm)	Array size	Emitter pitch (μm)	FOV ($^\circ$)	Angular resolution ($^\circ$)	Material	Phase shifter mechanism	Shifter power (mW/π)
[156] (2020)	0.488	64	> 1.95 (non uniform)	$> 50^\circ$	0.17°	SiN	TO	30
[157] (2018)	1.55	64	0.775	180°	1.6°	Si	TO	NR
[158] (2020)	1.55	8	4.5	19.1°	3.9°	polymer	EO	$V_\pi = 12.5 \text{ V}$

2D array of grating couplers, 2D beam steering with phase control

Ref. (year)	λ (μm)	Array size	Emitter pitch (μm)	FOV ($^\circ$)	Angular resolution ($^\circ$)	Material	Phase shifter mechanism	Shifter power (mW/π)
[86] (2013)	1.55	8×8	9×9	$6^\circ \times 6^\circ$	1.1°	Si	TO	8.5
[159] (2020)	1.55	8×8	10×40	$8.9^\circ \times 2.2^\circ$	$0.92^\circ \times 0.32^\circ$	Si	EO (injection)	17
[160] (2015)	1.55	4×4	50×50	$1.8^\circ \times 1.8^\circ$	$0.5^\circ \times 0.5^\circ$	Si	EO (injection)	$I_\pi = 10 \text{ mA}$
[161] (2019)	1.55	Nonuniform $N = 128$	> 5.6	$16^\circ \times 16^\circ$	$0.8^\circ \times 0.8^\circ$	Si	TO	10.6

Table 1.9: A survey of OPA-based beam steering systems reported in the literature (Part 1). The survey is not exhaustive, and selects works with best performance in its category and with an experimental implementation (i.e., no theoretical/simulation only works are included). TO = thermo-optic. EO = electro-optic. GC = grating coupler. QCSE = Quantum Confined Stark Effect. NR = not reported.

1D array of grating couplers, 2D beam steering (phase-controlled in one direction and wavelength-controlled in the other)

Ref. (year)	λ (μm)	Array size	Emitter pitch (μm)	FOV ($^\circ$)	Angular resolution ($^\circ$)	Material	Phase shifter mechanism	Shifter power (mW/π)
[107] (2016)	1.55	128	7.25	$80^\circ \times 17^\circ$	$0.14^\circ \times 0.14^\circ$	Si	TO	8
[146] (2013)	1.55	8	5.5	$10^\circ \times 6.5^\circ$	$2^\circ \times 0.2^\circ$	InP	EO (injection)	$I_\pi = 10 \text{ mA}$
[162] (2017)	1.55	1024	2	$45^\circ \times \text{NR}$	$0.03^\circ \times \text{NR}$	Si	TO	48
[163] (2018)	1.55	512	1.3	$70^\circ \times 14^\circ$	$0.15^\circ \times 0.15^\circ$	Si	TO	2.6
[164] (2019)	1.55	32	2	$51^\circ \times 28^\circ$	$0.78^\circ \times 0.02^\circ$	Hybrid (III-V/Si)	EO (QCSE)	1.5×10^{-6}
[165] (2020)	1.55	32	1	$97^\circ \times 17.1^\circ$	$2.3^\circ \times 2.8^\circ$	Si/SiN	TO	13
[166] (2020)	1.55	512	1.3	$70^\circ \times 6^\circ$	$0.15^\circ \times 0.08^\circ$	Si	TO	1.7
[154] (2020)	1.55	8192	1	$100^\circ \times 17^\circ$	$0.01^\circ \times 0.039^\circ$	Si	EO	0.15
[148] (2022)	1.55	64	NA	$140^\circ \times 13.5^\circ$	$2.1^\circ \times 0.08^\circ$	Si	TO	2.56
[167] (2022)	1.55	64	2.5	$36^\circ \times 3.4^\circ$	$0.85^\circ \times 0.25^\circ$	Si	EO (depletion)	0.45×10^{-6}
[168] (2022)	1.55	64	1	$96^\circ \times 14.4^\circ$	$1.49^\circ \times 1.41^\circ$	Si/SiN	TO	NR

Table 1.10: A survey of OPA-based beam steering systems reported in the literature (Part 2). The survey is not exhaustive, and selects works with best performance in its category and with an experimental implementation (i.e., no theoretical/simulation only works are included). TO = thermo-optic. EO = electro-optic. GC = grating coupler. QCSE = Quantum Confined Stark Effect. NR = not reported.

Other architectures

Ref. (year)	Brief description	Array size	Emitter pitch (μm)	FOV ($^\circ$)	Angular resolution ($^\circ$)	Material	Phase shifter mechanism	Shifter power (mW/π)
[109] (2019)	GC 1D array, TO steering one dir, EO steering in the other, $\lambda = 1.55 \mu\text{m}$	16	2	$10^\circ \times 45.4^\circ$	$5.8^\circ \times 3.2^\circ$	Si	EO (injection) & TO	2.3 & 150
[110] (2020)	GC 1D array, wav steering both dirs, $\lambda = 1.55 \mu\text{m}$	64	16	$36^\circ \times 5.5^\circ$	$0.11^\circ \times 0.2^\circ$	Si	NA	NA
[155] (2019)	2D GC array MEMS phase shifter $\lambda = 1.55 \mu\text{m}$	160×160	19.1×20	$6.6^\circ \times 4.4^\circ$	$0.042^\circ \times 0.031^\circ$	Si	MEMS	$V_\pi = 10.5 \text{ V}$
[169] (2019)	1D GC array LC phase shifter $\lambda = 632.8 \text{ nm}$	16	2	$10.5^\circ \times \text{NR}$	$0.7^\circ \times 2.3^\circ$	SiN	LC	$V_{op} = \pm 3.5 \text{ V}$

Table 1.11: A survey of OPA-based beam steering systems reported in the literature (Part 3). The survey is not exhaustive, and selects works with best performance in its category and with an experimental implementation (ie., no theoretical/simulation only works are included). MEMS = Micro-electro-mechanical system. LC = liquid crystal. NR = not reported.

1.5 Summary

In this chapter we have discussed the potential of using integrated optics for applications in imaging systems, offering advantages in terms of compactness, cost, system simplicity and reconfigurability.

Nevertheless, the exploitation of integrated optics for imaging has been limited to specific functionalities, typically within bulk optics systems. This is mostly due to (1) lack of cross-domain expertise, where the community working on imaging and bulk optics is not comfortable with integrated optics and vice versa, (2) challenges in the in- and out-coupling of the light from the chip to free space and (3) the lack of scalable, "full functionality" photonic platform that allows for generation, processing and launching of visible and IR light within the same substrate.

We have identified two fields where integrated photonic systems can have an important impact. The first is lensless microscopy, which allows for the realization of large FoV imaging systems with sub-micrometer resolution without the use of lenses, resulting in very compact, low cost systems. Here, we have seen how the realization of a large array of individually addressable, sub-micrometer sized emitting apertures spaced by half the emission wavelength and with a narrow emission spectrum could enable diffraction limited resolution imaging. Integrated photonics arises as a promising technology for the fabrication of these arrays due to the fine fabrication resolution (sub-100 nm), low cost and scalability. The main challenge we face are the realization of efficient, sub-wavelength sized emitting apertures and the individual addressing of a large number of apertures with sub-micrometer spacing.

The second field is the realization of OPA-based beamforming systems, with specific applications in FSOC and LIDAR. Integrated optics offers a natural avenue towards the realization of all the necessary components to realize an OPA (power distribution network,

phase control and emitter array) in the same chip in a very compact form factor without any moving parts. The main challenges here are (1) the realization of small area, low power phase shifters (with power consumption below 1 mW and about $100 \times 100 \mu\text{m}$ in size), (2) fabricating very large arrays of optical antennas with μm spacing and (3) accurate control of all these phase shifters.

It is interesting to notice how both fields have in common the requirement of realizing large arrays of small, efficient and closely spaced emitting apertures. In the first part of this thesis we will deal with these arrays for imaging purposes, which require closer emitter spacing (sub-wavelength) but do not require coherence between the different emitting apertures. In the second part of this thesis we will address the realization of these arrays for beamforming, where the spacing requirements are relaxed ($4 - 10 \lambda$) but coherence between all the emitting apertures needs to be ensured and carefully controlled.

The goal of this thesis is to demonstrate that integrated photonics is a useful technology for imaging applications that can enable improved and new functionality in a compact and low cost manner.

1.6 Thesis outline

The rest of this thesis is organized in 5 chapters, each describing a self-contained project with a specific set of goals.

- Chapter 2 discusses the realization and application of silicon integrated surface emitting LED arrays for the realization of compact lensless microscopes.
- Chapter 3 deals with the modeling and experimental demonstration of waveguide coupled, native silicon light emitting diodes fabricated in commercial silicon photonic processes.

- Chapter 4 demonstrates the applicability of the waveguide coupled sources presented in Chapter 3 to realize fully monolithic refractive index sensors.
- Chapter 5 presents the modeling and system-level study of OPA architectures using non-uniform emitter spacing to ease the requirements in the size of traditional, uniformly spaced OPAs, which become intractable for advanced beamforming applications.
- Lastly, Chapter 6 summarizes the conclusions of this thesis and lays out future work that builds on the findings of this work.

Chapter 2

Lensless Holography with Near-Point CMOS Light Sources

In the first chapter of this thesis we have identified lensless imaging and microscopy as one of the fields where the advantages of integrated photonics can enable the realization of compact and low cost systems with similar or improved performance when compared to traditional bulk optics approaches. In this chapter, we use integrated silicon surface emitting light sources made in CMOS processes to demonstrate such a potential.

First, in Section 2.1 we demonstrate the use of a single subwavelength silicon LED to realize a lensless holographic microscope that circumvents the need of a pinhole to ensure spatial coherence (which is required in traditional lensless microscopes) and therefore results in a very compact system that could be miniaturized to a size of about 1 cm^3 .

In Section 2.2 we then demonstrate an array of subwavelength silicon LEDs with subwavelength spacing, which could enable subpixel sampling imaging and further improve the resolution of our lensless microscope. Such a small emitter size and spacing are enabled by the high fabrication quality of CMOS photonics processes, proving once again

the potential of integrated photonics for applications in imaging.

Lastly, in Section 2.3 we demonstrate the use of the array to realize subpixel sampling holographic imaging. To fully exploit the benefits of our integrated sources, we do so in a novel reflection configuration where both emitter and detector array are integrated in the same chip. Such an approach could enable the realization of lensless microscopes in any system that uses CMOS imaging arrays, such as smartphones.

2.1 Compact lensless holography with a single near-point CMOS light source

As we have discussed in Section 1.3, lensless holography is a promising technique to realize low cost, high resolution, and large FoV microscopes without requiring bulky and expensive optical components such as objectives and lenses.

As we reviewed, one of the main factors limiting the compactness and flexibility of lensless holographic microscopes is the requirement for spatial coherence. Since holographic microscopes usually employ a large area LED as the illumination source, a pinhole is usually required to reduce the effective illumination source dimensions and ensure a large enough degree of spatial coherence [170]. To reduce alignment tolerances and optical losses, big pinholes with diameters in the range of 20 -50 μm are used, resulting in a low NA illumination ($\text{NA} \approx 0.62\lambda/r$), which requires a source-to-sample separation on the order of 10 cm to ensure that the full FoV of the camera is illuminated. A further reduction in the form factor of lensless holographic microscopes could result in increased adoption of the this technique by enabling integration in portable devices and possibly even personal devices such as watches or phones.

2.1.1 Near-point CMOS surface emitting LEDs

Achieving a further reduction in size of lensless microscopes requires rethinking the illumination source and moving away from the ubiquitous LED + pinhole configuration. To this end, we propose the use of near-point surface emitting silicon LEDs realized in commercial CMOS processes as the illumination source for lensless holography. The design, modeling and characterization of these sources was the subject of Dr. Jin Xue's PhD thesis [171] and continued by Dr. Zheng Li [172]. Here, we summarize their characterization results as they are relevant for the subsequent work that was carried out in this thesis.

Fig. 2-1(a) shows the silicon LED structure, which is fabricated in GlobalFoundries 55BCD CMOS microfabrication process. It consists of a vertical n+ poly-Si / crystalline Si n-well / crystalline Si p-substrate junction. The application of high electric field to the overlap area between the gate poly-Si (≈ 100 nm thick) and the crystalline Si n-well results in the formation of a very thin Si filament (a process usually referred to as hard breakdown [173,174]). The cross-section of the breakdown site is usually sub-100 nm, and the Si filament becomes a resistor in series with the gate poly-Si and the crystalline Si n-well [174].

Light generation is realized through phonon-assisted radiative recombination of electron-hole pairs. As shown in Fig. 2-1(b), a hole accumulation layer forms near the gate oxide due to the negative bias applied to the n+/n/p junction. Electrons injected into the accumulation region through the Si filament quickly recombine [175], leading to a highly localized emission spot as shown in the micrograph in Fig. 2-1(c).

From the previous work of Dr. Zheng Li and Dr. Jin Xue [172], these LEDs have a subwavelength-sized emission area below $0.14 \mu\text{m}^2$ (≈ 400 nm emission radius) and a high spatial intensity (>50 mW/cm²) with the emission spectrum centered around

1130 nm (Fig. 2-1(d)).

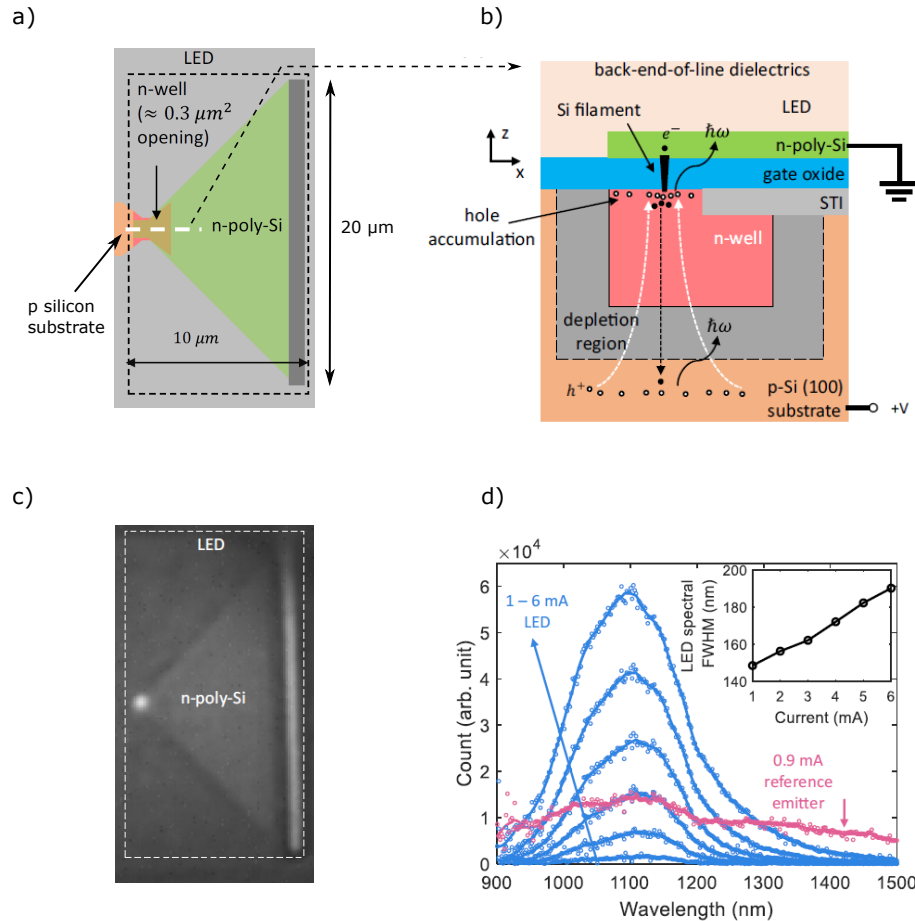


Figure 2-1: Near point CMOS surface emitting LEDs. (a) Layout of the LED, consisting of a vertical junction formed by an n poly-silicon taper, an n doped crystalline silicon region and the lightly p doped silicon substrate. (b) Cross section of the LED in the light generation area. Electrons injected into the crystalline silicon n-well through the poly-silicon taper + filament radiatively recombine with accumulated holes due to the negative bias applied to the junction. (c) Micrograph of the LED emission when biased at 6 mA current. (d) LED emission spectrum measured with an InGaAs spectrometer, with a center wavelength around 1130 nm. The inset shows the FWHM of the spectrum. All the data in this figure is work by Dr. Zheng Li and Dr. Jin Xue [172].

Considerations for lensless holography

We can use such a source for the simplification and further miniaturization of lensless holographic microscopes by taking advantage of its subwavelength emission size, which ensures very high spatial coherence without the need of a pinhole.

Building on the discussion in Section 1.3.2, we can estimate the improvement on compactness from the spatial coherence requirements $\lambda z_1/Dz_2 > 10$, where D is the emission size of the LED. For the CMOS LED source, $\lambda = 1130$ nm and $D = 400$ nm, resulting in a requirement $z_1/z_2_{CMOS_LED} > 3.54$. Traditional lensless holography microscopes use $\lambda \approx 600$ nm and $D = 20$ μm , resulting in $z_1/z_2_{typ} > 330$. Clearly, we can achieve a significant reduction in total size for the same z_2 .

Not only this, but the large NA of the LED (due to its small emission area) also allows for a significant reduction in the required source-sample distance to achieve a given FoV. From simple geometrical considerations it is easy to see that for a source with a given $NA = \sin(\theta)$, a source to sample distance $z_1 = FoV/(2\tan(\theta))$ is required to illuminate a given FoV. Typical lensless microscopes have an $NA = 0.62\lambda/r \approx 0.037$, requiring a source to sample distance $z_{1typ} = 1.3$ cm for a 1 mm FoV. In contrast, the CMOS LED has a $NA \approx 0.66$ ¹, which reduces the required source to sample distance to $z_{1CMOS_LED} = 500$ μm .

While the spatial coherence of the CMOS source is very high, temporal coherence is a concern due to the large spectral bandwidth of about 180 nm. From our discussion in Section 1.3.2, we need to ensure $z_2 < 2\ln(2)\lambda^2/(5\pi\Delta\lambda)$ to achieve diffraction limited resolution. A $\Delta\lambda = 180$ nm translates into a requirement $z_2 < 620$ nm, which is very hard to achieve in practical setups due to the finite size of the glass slide holding the sample and the presence of a protecting layer in the CMOS imager.

¹We measured this by comparing the magnification of a hologram acquired with the CMOS LED to that of a hologram acquired with a conventional LED + pinhole while keeping z_1 and z_2 constant.

An approach to achieve relaxed requirements on z_2 is to filter the emitted light so that the temporal coherence of the source is improved. This of course comes at the expense of a lower emitted power and therefore decreased SNR. Additionally, it requires the addition of the filter element in the holographic microscope, increasing overall complexity and size.

An alternative to light filtering is the use of computational hologram reconstruction techniques that account for the broad emission spectrum of the source, an approach that has been successfully demonstrated in X-ray imaging [176, 177] and deep neural network-based hologram reconstruction [178].

In this sense, in lensless holography having a source with low temporal coherence is preferable to having a source with low spatial coherence: while low temporal coherence can be compensated through computational methods, low spatial coherence increases the necessary z_1 and z_2 and therefore results in large systems (≥ 10 cm). Of course, there is no computational approach to compensate for the larger z_1 required to achieve a given FoV.

2.1.2 Lensless holography setup

In the previous section we have discussed how the near point-source LED with very high spatial coherence developed by Dr. Zheng Li and Dr. Jin Xue is an attractive illumination source for the realization of compact lensless microscopes.

To validate this claim, I built a lensless holography setup using the CMOS LED as an illumination source. A schematic of the setup is shown in Fig. 2-2(a), along with a picture of the actual setup which is shown in Fig. 2-2(b). As discussed, the small emission area of the CMOS LED ensures reasonable spatial coherence without a pinhole, as well as a large NA that allows the source-to-sample distance to be decreased below

5 mm while maintaining a large FoV. This results in a very compact system with a total source-to-camera distance below 1 cm. We use a ZWO ASI 1600MM CMOS camera with a pixel pitch of $3.8 \mu\text{m}$ and a total imaging area of $17.7 \times 13.4 \text{ mm}$.

A sample hologram acquired with the setup, corresponding to a dandelion fuzz, is shown in Fig. 2-2(c). The illuminated FoV in the camera for the $\approx 5 \text{ mm}$ source to sample distance is about $11.5 \times 11.5 \text{ mm}$. For a typical lensless holography setup with a $20 \mu\text{m}$ diameter pinhole the required source to sample distance would be $25 \times$ larger, which would increase z_1 to 12.5 cm .

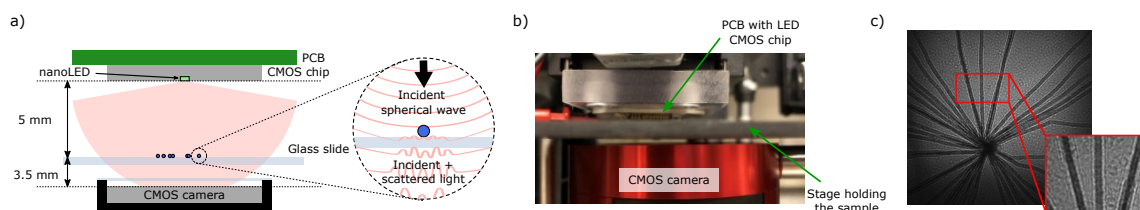


Figure 2-2: Lensless holography setup with CMOS LED as illumination source. (a) Schematic of the setup, where the source to sample distance is about 5 mm. (b) A picture of the constructed setup. (c) An example of an acquired hologram corresponding to a dandelion fuzz sample. The inset shows a closeup where the characteristic interference pattern of holograms is visible.

2.1.3 Computational reconstruction: increasing quality with Deep Image Prior

The acquired holograms need to be computationally reconstructed to recover the image sample. As we discussed earlier, the broadband nature of the CMOS LEDs makes traditional computational reconstruction methods imperfect because they assume monochromatic illumination [59, 179].

Some reconstruction models that account for broadband illumination have been reported in the literature: in [176, 177], a modification of monochromatic iterative phase

retrieval using a priori known source spectrum was introduced for applications in X-ray imaging. Additionally, broadband diffraction models to generate object-diffraction pairs have been used in end-to-end deep neural networks for object retrieval [178]. One important limitation of these implementations is that they require prior knowledge of the exact emission spectrum of the source, which requires an extra spectroscopic calibration step.

To overcome these limitations, we used a novel broadband computational reconstruction model based on a Deep Image Prior (DIP) neural network implementation developed by Dr. Iksung Kang with some assistance of my own [180]. Since it was not the focus of my work, here I will only briefly summarize its main characteristics. Detailed information can be found in [181], and the code can be found in https://github.com/iksungk/simultaneous_recovery.

The basic framework of a DIP architecture is illustrated in Fig. 2-3(a), where we want to recover a complex object f from an experimental measurement g , where measurement and object are related through a known matrix H and we account for the presence of noise so that $g = Hf + n$. Through the DIP architecture, an estimation of the object \hat{f} is obtained by feeding a neural network with random noise. In the forward path, the estimated object \hat{f} is used to generate an estimation of the measurement through $\hat{g} = H\hat{f}$, and then compared to the experimental measurement g through a loss metric (for example, the mean square error). By backpropagating the loss through the neural network weights, an increasingly better estimate of the object we want to recover is obtained.

In the case of using DIP for holographic reconstruction, the experimental measurement g is the acquired hologram, and the matrix H describes the diffraction of light from the sample plane to the image plane. Figure 2-3(b) shows a schematic of the DIP-based computational reconstruction algorithm, which was adapted to the use a broadband

illumination source.

Two different DIP neural networks (NNs) with random noise z as the input are used, one representing the amplitude α (absorption) and another one the phase φ response of the imaged sample, so that the complex transmission function of the sample is $f = \alpha \exp(i\varphi)$. A sigmoid activation function is applied at the end of the NNs to limit α between $[0, 1]$ and φ between $[-\pi, \pi]$.

To account for the spectral shape of the illumination, the wavelength axis is uniformly sampled at points $\lambda_1, \lambda_2, \dots, \lambda_N$, and the electric field propagated after interacting with the sample through free space for each of these wavelengths. A non-paraxial free-space propagation kernel H is used to model the diffraction of the field at wavelength λ_n as it propagates (for a total length corresponding to the sample-camera distance z_2) [182]:

$$H_{\lambda_n}(f) = \mathcal{F}^{-1} \left[\mathcal{F}[f] \circ \exp \left(-iz_2 \left(k_n - \text{Re} \left[\sqrt{k_n^2 - k_x^2 - k_y^2} \right] \right) \right) \right], \quad (2.1)$$

$$k_n = \frac{2\pi}{\lambda_n}.$$

To account for the non-uniform emission spectrum of the source, weighing parameters γ_n ($n = 1, 2, \dots, N$) are used so that the estimation of the recorded hologram at the camera plane is defined as:

$$\hat{I}_{\text{det}} = \sum_{n=1}^N \gamma_n |H_{\lambda_n, \Delta z}(f)|^2. \quad (2.2)$$

In contrast to the methods presented in the literature, this algorithm does not require detailed prior knowledge of the emission spectrum of the source, ie., does not require prior knowledge of the weighing parameters γ_n . Instead, it tries to estimate the contribution of each wavelength by including the spectral weights γ_n as optimizable parameters in the

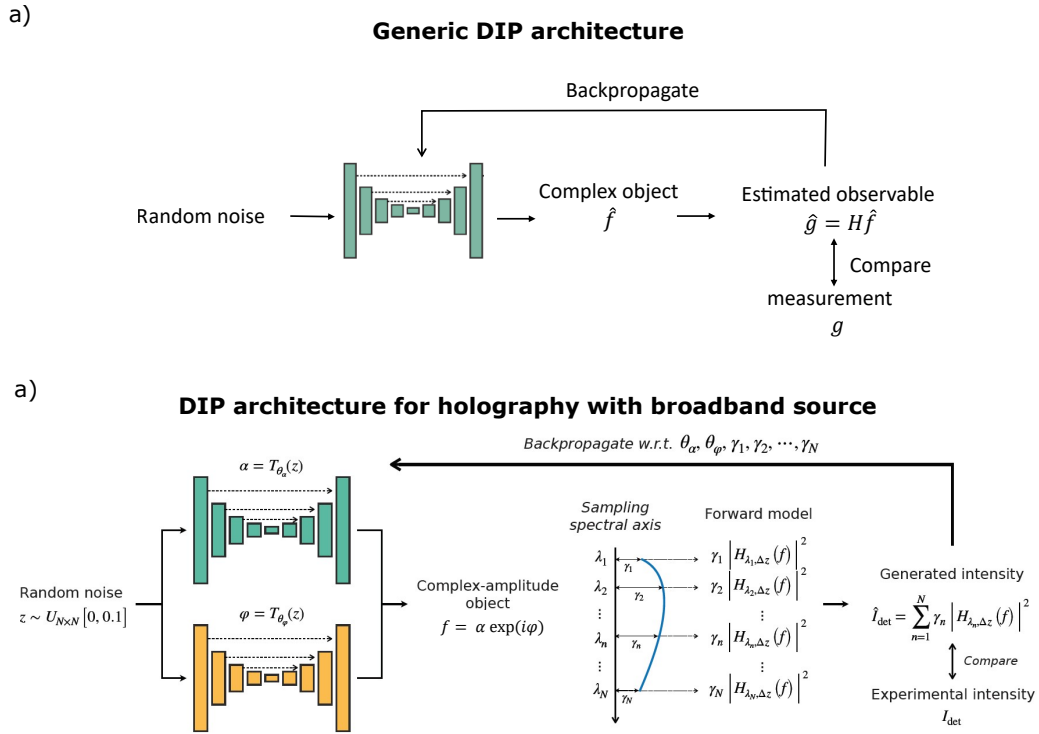


Figure 2-3: (a) Generic DIP architecture, which tries to minimize the difference between an experimental measurement g and the estimation of that measurement generated with a neural network $\hat{g} = H\hat{f}$. (b) DIP architecture for holographic reconstruction with a broadband source. Two untrained neural networks are used to represent the object phase and amplitude response, and the contribution of each wavelength in the source spectrum is accounted through the weights γ_n ($n = 1, 2, \dots, N$). This DIP architecture was developed by Dr. Iksung Kang.

DIP framework.

With this setup, finding the complex transmission of the sample f and the source spectrum reduces to finding the ensemble of parameters $\theta_\alpha, \theta_\varphi$ of the DIP NNs and $\gamma_1, \gamma_2, \dots, \gamma_N$ so that the \hat{I}_{det} is as close as possible to I_{det} (the experimentally recorded hologram).

To solve the inverse problem, an Adam optimizer [183] is used to minimize the loss function:

$$\begin{aligned}
\mathcal{L}(\hat{I}_{\text{det}}, I_{\text{det}}; \theta_{\alpha}, \theta_{\varphi}, \Delta z, \gamma_1, \gamma_2, \dots, \gamma_N) = & \text{MSE}(\hat{I}_{\text{det}}, I_{\text{det}}) + \kappa_1 \text{NPCC}(\hat{I}_{\text{det}}, I_{\text{det}}) \\
& + \kappa_2 \left(1 - \sum_{n=1}^N \gamma_n\right) + \kappa_3 \sum_{n=1}^N \text{ReLU}(-\gamma_n) + \kappa_4 (\text{TV}(\alpha) + \text{TV}(\varphi)),
\end{aligned} \tag{2.3}$$

where MSE stands for mean-squared error, NPCC for negative Pearson correlation coefficient, and TV for total variation. NPCC is used as a first regularizer because it has been shown to achieve higher reconstruction fidelity in phase retrieval applications with convolutional neural networks [184, 185]. The second regularizer is essentially a Lagrange multiplier imposing the constraint on γ_n 's to be normalized with their sum being 1 and non-negative. Lastly, TV is applied to both absorption and phase delay profiles of the object f to smooth out fringe pattern artifacts.

Figure 2-4 shows acquired holograms and the reconstructed samples for three different samples: an ensemble of 20 μm diameter beads (Fig. 2-4(a)), a dandelion fuzz (Fig. 2-4(b)) and a housefly mouth (Fig. 2-4(c)). For each sample, five different elements are shown:

1. The acquired hologram, I_{det} .
2. The approximation to the hologram obtained with the DIP architecture after optimization, \hat{I}_{det} .
3. A micrograph obtained with a regular microscope using a $5\times$ magnification objective.
4. The sample reconstruction using a baseline method employing the steepest gradient descent algorithm and the *a priori* known information of the source spectrum (ξ_n):

$$\hat{f}_{\text{baseline}} = \operatorname{argmin}_f \left[\frac{1}{2} \left| \sum_{n=1}^N \xi_n |H_{\lambda_n, \Delta z}(f)|^2 - I_{\text{det}} \right|^2 \right], \quad (2.4)$$

5. The sample reconstruction using the DIP architecture.

We can clearly see how the DIP algorithm results in significantly better contrast and a background with fewer artifacts, demonstrating the superiority of this approach. For the 20 μm diameter bead sample, we quantified the reconstruction quality by looking at the Pearson Correlation Coefficient (PCC) between the micrograph shown in Fig. 2-4(a3) and the sample reconstructions in Fig. 2-4(a4, a5)². The PCC for the baseline method is $PCC_{\text{baseline}} = 0.5281$, while our method achieves a 40% improvement to $PCC_{\text{DIP}} = 0.7483$. Also importantly, our DIP algorithm is capable of accurately recovering the emission spectrum of our source, as shown in Fig. 2-4(d).

2.2 Exploiting CMOS: realizing sub-wavelength spacing emitter arrays

In the previous section we have demonstrated how the use of near-point CMOS LEDs as illumination sources for lensless holography is attractive because it achieves a large FoV in a very compact form factor below 1 cm.

While interesting, one could argue that the use of a III-V nano-LED (such as the ones presented in Table 1.3), while more expensive, would result in better performance due to

²Preparing a paired dataset to perform a quantitative comparison between our method and the baseline method is fundamentally challenging because the ground truth (the micrograph) is taken in a different imaging setup with a different magnification. Therefore, for quantitative evaluation careful registration, rotation and scaling of the ground truth needs to be performed. Notice how the elimination of such a time consuming step is one of our main motivations to use an untrained NN model as opposed to a training-based NN.

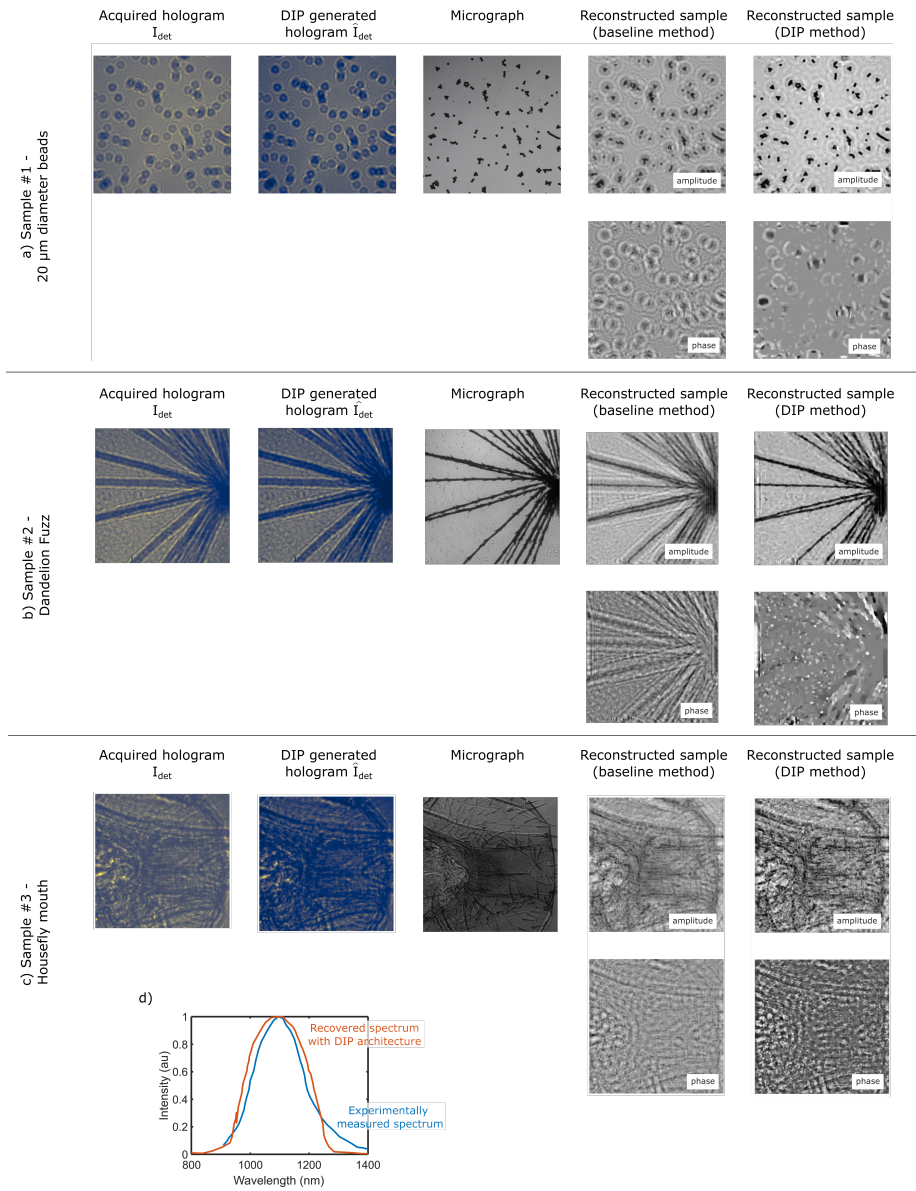


Figure 2-4: DIP-based holographic reconstruction. (a-c) Recorded hologram (column 1), generated hologram with the DIP architecture (column 2), sample micrograph (column 3), image reconstruction with baseline gradient-based algorithm (column 4) and image reconstruction with DIP-based algorithm; for a sample consisting of 20 μm diameter beads (a), a dandelion fuzz (b) and a housefly mouth (c). (d) Comparison of the experimentally measured emission spectrum of the CMOS LED (blue) and the recovered spectrum with our DIP algorithm (orange).

their larger efficiency and generated power.

While this is true for single emitter sources, the picture changes when we try to realize large emitter arrays for use in subpixel shift imaging. As discussed in detail in Sec. 1.3.3, subpixel shift imaging is an attractive approach to overcome the resolution limitation set by the finite pixel size and spacing in the imager that acquires the hologram.

Building imaging sources for subpixel shift imaging that do not require very precise mechanically moving stages is technically challenging because of the need to realize large arrays of submicron spaced, submicron sized emitters. Along with the concerns on the decrease in efficiency and output power as emitter size decreases, one of the main challenges is obtaining reliable electrical contact and individual addressing of submicron spaced, submicron sized emitters - (1) without occluding nearby emitters, (2) without causing crosstalk (where addressing one emitter causes another emitter to also partially turn on) and (3) without shortcircuiting other contact lines [97].

This is particularly challenging in III-V material platforms, which have a limited number of interconnect layers and require 3D integration with CMOS driving electronics. This limits practical emitter pitches to $5\ \mu\text{m}$ and above for large scale systems [98, 99].

On the contrary, realizing these emitter arrays monolithically in a CMOS process allows to take advantage of the very high precision and large number of metal interconnect layers available (> 7 , with the first 3 - 5 layers having a resolution below 100 nm) as well as realize the driving electronics on-chip, offering a viable way of achieving large scale, addressable sub- μm apertures. Furthermore, our CMOS LED configuration uses a transparent top contact (the thin poly-Si contacting the Si filament does not absorb the emitted light), thus relaxing the restrictions to avoid occlusion of the generated light.

Therefore, our near-point CMOS LED sources are attractive candidates for the realization of such a large, tightly spaced emitter array. As a proof of concept I built a four element emitter array, depicted in Fig. 2-5(a, b). The same CMOS LED configuration

presented in the last section (Fig. 2-1) is used, but now instead of a single poly-Si taper we have 4 distinct bars overlapping slightly different regions of a single crystalline silicon (c-Si) island (where the holes are accumulated due to the application of a negative bias). As in the case of the single LED, we rely on the formation of a silicon filament through the gate oxide for the injection of electrons into the c-Si. Since the 4 poly-Si regions overlap different regions of the c-Si, four spatially distinct filaments are formed, and therefore 4 distinct emission spots are generated. Importantly, each of these emission spots can be individually addressed since each poly-Si bar is independently contacted (while the p Si substrate is shared as there is a single c-Si island).

By leveraging the high fabrication accuracy of our CMOS fabrication process, we can make the distinct poly-Si regions as close as 180 nm (Fig. 2-5(b)), limited by the minimum spacing allowed by the process design rules. As detailed in Fig. 2-5(b), the spacing between elements in the x and y directions is slightly different due to the particularities of the process design rules. It is also important to notice that, while the minimum spacing between poly-Si bars is 180 nm, we expect the spacing between emission spots to be in the range between 180 nm - 1,000 nm as the thin silicon filament can be formed anywhere in the c-Si + poly-Si overlap region (red dashed outlined regions in Fig. 2-5(b)).

We performed a detailed characterization of the LED array characteristics, which is summarized in Fig. 2-5. Fig. 2-5(c) shows the IV curve of each individual LED in the array. As can be seen, a relatively large bias voltage of around 18 V is required to achieve a noticeable current. This is because the substrate contact is located about 1 mm away from the LED array, which increases the series resistance significantly. The variability in the measured IV curves of the distinct LEDs comes from the differences in the series resistance of the wirebonds providing electrical access to the chip.

Figure 2-5(d) shows the LI curve of the 4 LEDs, where we measure the power coupled

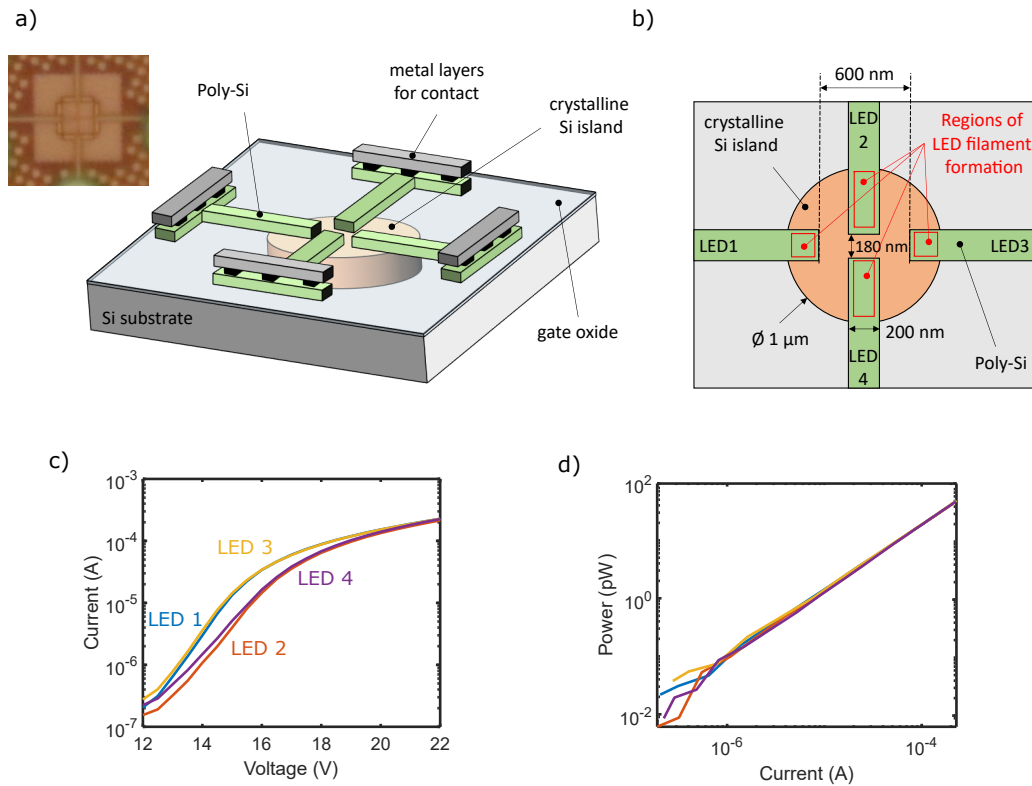


Figure 2-5: A 4 element array of CMOS near-point source LEDs. (a) 3D rendering, showing the 4 distinct poly-Si access region and the single c-Si island. The inset shows a micrograph of the fabricated device. (b) Closeup of the light generation area, where the poly-Si and c-Si overlap. The minimum spacing between LEDs is 180 nm. (c) IV curves of the 4 LEDs. (d) LI curves of the 4 LEDs.

into single mode fiber (SMF) through a high NA objective. At 200 μA bias current we measure an optical power of about 22 pW. Accounting for the transmission loss of the microscope objective, which is about 50%, and the coupling efficiency into SMF, which we measure to be around 40% [172], the power generated per LED is about 100 pW at 200 μA bias current.

We also imaged the generated illumination spots when the different LEDs are turned on using a 100x magnification microscope and extracted the spot size and emission

spot location from them. The results are summarized in Fig. 2-6. As depicted in Fig. 2-6(a), we obtained the generated spot sizes by fitting a gaussian profile to the imaged illumination spot when turning on each individual LED at a time. As shown, the generated emission spots have an average $(FWHM_x, FWHM_y) = (790 \text{ nm}, 735 \text{ nm})$ and a standard deviation of $(\sigma_{FWHM_x}, \sigma_{FWHM_y}) = (41 \text{ nm}, 32 \text{ nm})$ at $200 \mu\text{A}$ bias. The spot size does not change significantly in the $50 \mu\text{A} - 200 \mu\text{A}$ bias current range³. Since these devices have an emission spectrum centered at $\lambda_0=1130 \text{ nm}$, the spot size corresponds to $\approx 0.67\lambda_0$.

To obtain the position and spacing of the different spots we took pictures of the generated emission with two LEDs turned on at the same time. As shown in Fig. 2-6(b), this generates an illumination pattern with two distinct peaks from which we can extract the emission spot spacing by fitting a double gaussian shape. The extracted spot localizations are shown in Fig. 2-6(c). The closest spacing (achieved between LED 3 and 4) is 410 nm (0.36λ), while the largest spacing is 678 nm (0.6λ) between LEDs 1 and 3. This is, to the best of our knowledge, the first time an array of subwavelength-sized emitters with subwavelength spacing has been demonstrated in a CMOS platform, and the smallest spacing ever demonstrated in an array of individually and electrically addressable nanoLEDs when compared to its emission wavelength (Table 2.1).

It is also important to note that there is no measurable crosstalk in the LED addressing (i.e, turning on a specific LED does not result in another one of them also turning on), as the generated emission when a single LED is addressed shows a very distinct single peak (Fig. 2-6(a)) while we would expect to see a double peak if crosstalk was an issue (Fig. 2-6(b)).

³We calibrated the pixel to distance conversion factor by measuring the pixel size of a chip feature of known dimensions.

Ref (year)	Material	Array Size	Emitter size (nm)	Emitter pitch (μm)	Optical power (nW)	Optical intensity ($\text{nW}/\mu\text{m}^2$)	Electrical power (μW)	Emission λ (μm)	Pitch/ λ
[97] (2020)	GaN	6 × 6	400 × 400	0.8	1,000 †	6,250 †	25	0.5	1.6
[97] (2020)	GaN	2 × 32	200 × 200	0.4	100 †	2,500 †	4	0.5	0.8
This work (2023)	Si	2 × 2	790 × 735	0.41 - 0.67	0.1 ‡	0.17 ‡	1,500	1.13	0.36 - 0.6

Table 2.1: Comparison of our CMOS nano-LED array with other nano-LED arrays reported in the literature.

† Estimated from internal quantum efficiencies mentioned in [97] assuming 100% extraction efficiency.

‡ Accounts for limited extraction efficiency.

2.3 Subpixel shift lensless holography with sub-wavelength CMOS emitter arrays

In the previous section we have experimentally demonstrated that we can realize an array of subwavelength-sized emitters with subwavelength spacing exploiting the advantages of CMOS fabrication processes. As presented above, such an array is a good candidate for the realization of small form factor and low cost subpixel shift lensless holographic setups that do not require precise mechanically moving parts.

The fact that our emitter array is fabricated in a CMOS process has an additional advantage besides the ones we have already discussed (scalability, low cost, ease of individual addressing of emitters, monolithic driving circuits...): we can realize both the emitter array and the detector array in the same chip, as schematically shown in Fig. 2-7(a). This opens up an opportunity to realize *single chip*, compact lensless holographic microscopes, in contrast with traditional lensless microscopes which rely on separate

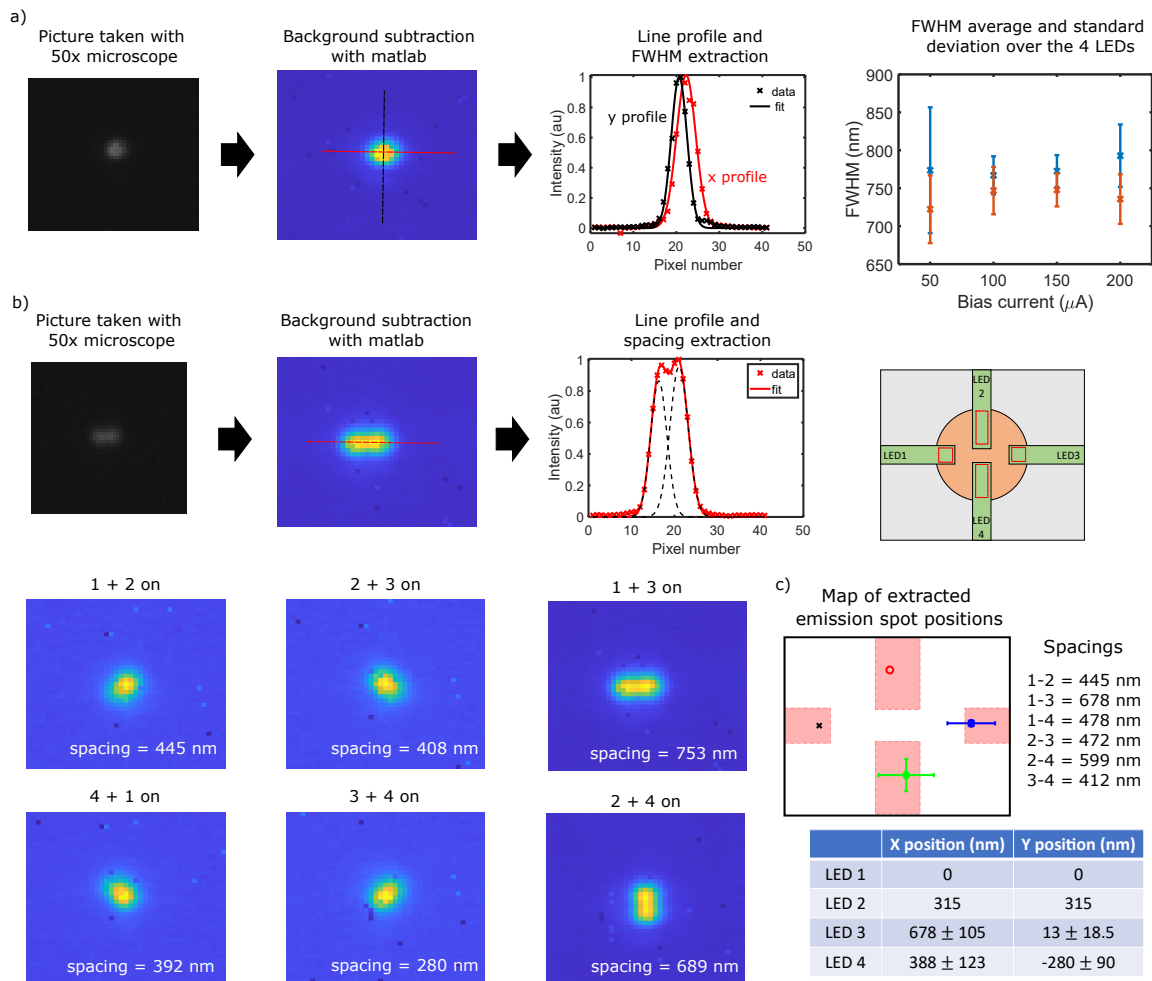


Figure 2-6: CMOS LED array spot size and spacing. (a) Process to obtain the FWHM of each individual spot, and measured FWHM as a function of bias current. The average and standard deviation are taken over the 4 LEDs of the array. (b) Process to obtain the spot spacing from an image with two LEDs turned on at the same time. (c) The obtained map of emission spot positions from the data shown in (b). The red shaded areas correspond to the overlap regions between each poly-Si bar and the c-Si island. Tables showing the (x, y) positions and spacing between elements are also included.

chips / components to realize the illumination source(s) and the detector. Of course, an additional component is required to reflect the light that interacts with the sample back to the detector - a mirror.

A prototypical setup using such a concept, which we name *reflection lensless holography*, is shown in Fig. 2-7(b). The main strength of this approach is that, in principle, any already existing CMOS imaging array could easily be modified to include our CMOS emitter array on the side (Fig. 2-7(a)), and then a lensless holographic microscope could be very easily realized by using a mirror (or any other reflecting element) at the other side of the sample. One could imagine a scenario where CMOS imaging arrays used in smartphone cameras incorporate our CMOS emitter array, allowing to easily turn them into lensless microscopes.

To the best of our knowledge, this is the first time lensless holography in a reflection configuration without the need for additional optical elements has been proposed. The use of lens and lensless holography in reflection configuration has been discussed before by a few research groups [186–188], but always in the context of off-axis holography (where a separate reference beam needs to be generated) and with the addition of a bulk optics beam splitter or diffraction grating (inset of Fig. 2-7(b)) so that the light generated by the illumination component can be routed to the detector. By taking advantage of the emitter and detector array being in the same chip, our configuration eliminates the need for this extra element, maintaining the simplicity and compactness of transmission-based inline holography setups.

2.3.1 Reflection lensless holography - Geometrical considerations

While reflection lensless holography is attractive to simplify holographic microscopes there are some additional considerations stemming from the unique imaging configuration that we need to take into account.

The most important is the fact that in our reflection configuration the generated

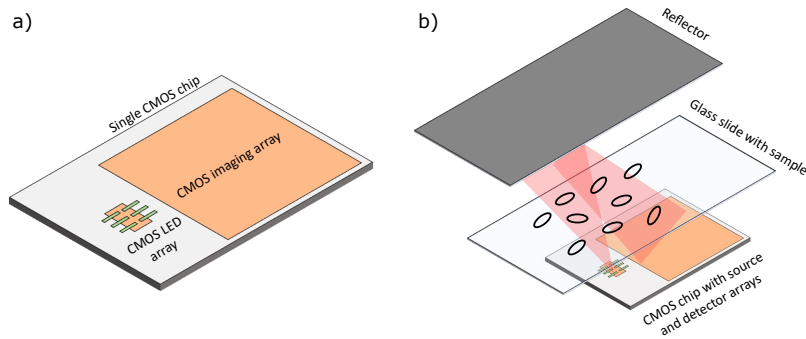


Figure 2-7: Reflection holography. (a) We can leverage our CMOS LEDs to realize both the source and detector arrays in the same chip. (b) To realize lensless microscopy with such a chip we need a reflector to guide the generated light back into the chip.

illumination field interacts with the sample twice: one in the direct path from the source to the sample and another one in the path from the reflector to the detector. This results in the formation of two different holograms at the image plane, as shown in Fig. 2-8(a).

These two holograms have slightly different characteristics:

- Magnification: The fact that our CMOS point sources generate a high NA illumination results in the hologram at the imager plane being magnified with respect to the hologram formed right after sample by a factor $M = (d_{source-sample} + d_{sample-imager})/d_{source-sample}$. As apparent from Fig. 2-8(a), these two values are different for the two images: for Image 1 (where the hologram is formed after reflecting off of the mirror) we have $d_{source-sample} = z_1 + 2z_2$, $d_{sample-imager} = z_1$; for Image 2 (where the hologram is formed before reflecting off of the mirror) we have $d_{source-sample} = z_1$, $d_{sample-imager} = z_1 + 2z_2$.

Therefore, the two holograms have a magnification ratio given by:

$$\frac{M_1}{M_2} = \frac{\frac{2(z_1+z_2)}{z_1+2z_2}}{\frac{2(z_1+z_2)}{z_1}} = \frac{z_1}{z_1 + 2z_2} = \frac{1}{1 + 2\frac{z_2}{z_1}} \quad (2.5)$$

- Shift when using array illumination: When using an array of sources for illumination, an image shift is obtained at the image plane (Fig. 2-8(b)). It is easy to see from geometrical optics how the hologram shift not only depends on the distance between the two sources but also on the geometrical configuration of the setup. In particular, $\Delta_{i,im.plane} = -d \frac{d_{sample-imager}}{d_{source-sample}}$, where d is the spacing between illumination spots at the source.

With this, we can conclude that the ratio in the hologram shift in direction i for the two sub-images is given by (below, d_i is the spacing between illumination sources in direction i):

$$\frac{\Delta_{i,1}}{\Delta_{i,2}} = \frac{d_i \frac{z_1}{z_1+2z_2}}{d_i \frac{z_1+2z_2}{z_1}} = \frac{z_1^2}{(z_1 + 2z_2)^2} \quad (2.6)$$

The formation of these double hologram sets up a limitation on the achievable FoV since for correct sample reconstruction we want to ensure that the two generated holograms do not overlap at the image plane. As shown in Fig. 2-8(a), the position of the two holograms depends on the geometry of the imaging setup. From geometrical considerations, to ensure non-overlapping holograms ($x_{1,1} < x_{2,0}$ in Fig. 2-8(a)) we need to fulfill:

$$\frac{z_1 + 2z_2}{z_1} > \frac{x_{s,max}}{x_{s,min}} \quad (2.7)$$

Above, $x_{s,max}$ ($x_{s,min}$) is the maximum (minimum) x position of the sample (with $x = 0$ defined as the position of the LED, see Fig. 2-8(a)). Besides the above, we also need to ensure that $x_{s,min} > 0$ and that the source illuminates the entire sample (i.e., $\alpha_{2,1} < asin(NA)$ in Fig. 2-8(a), where NA is the numerical aperture of the illumination source).

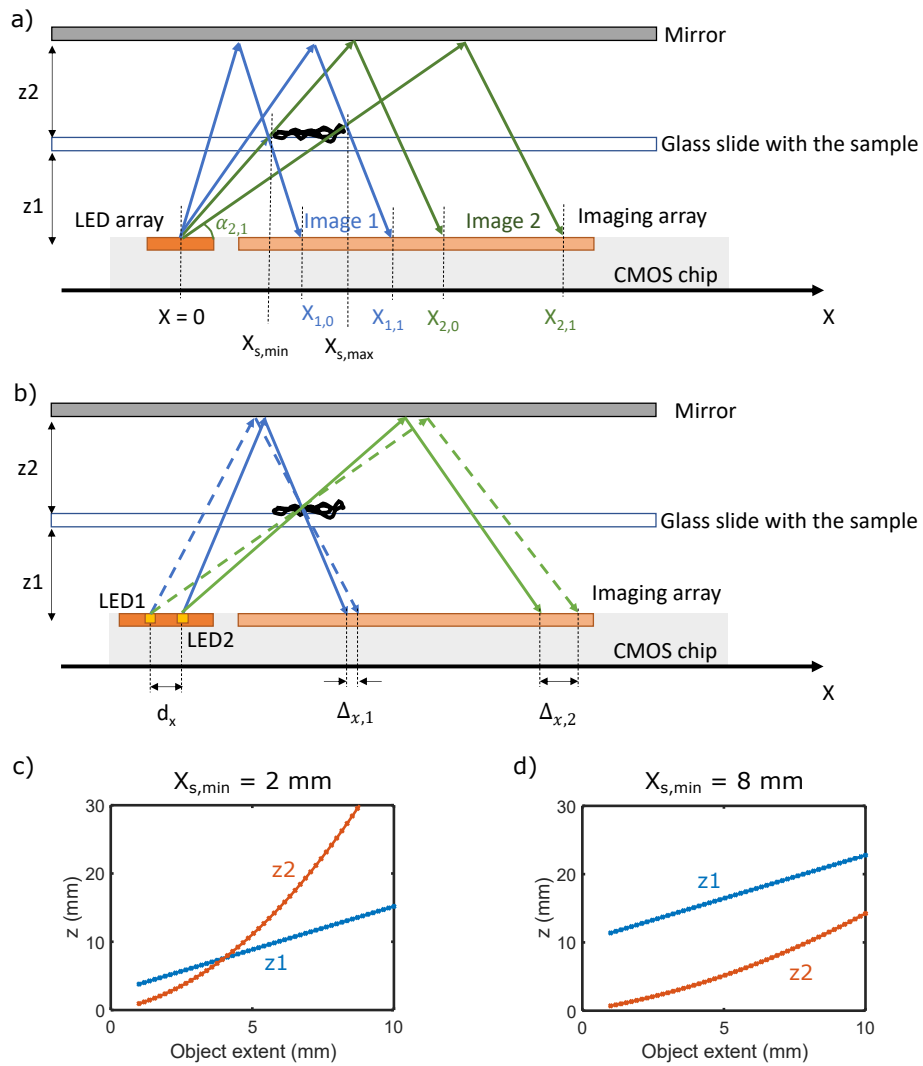


Figure 2-8: Geometric considerations for reflection holography. (a) The illumination interacts with the sample twice, generating two distinct holograms of the same sample at the imaging plane. (b) Illuminating with different LEDs causes a shift in the acquired holograms. (c, d) Required source - sample distance z_1 and sample - reflector distance z_2 as a function of the sample extent when the lower edge of the sample ($X_{s,min}$) is 2 mm away from the LED (b) and when $X_{s,min}=8$ mm (c)

Clearly, a given geometrical arrangement of the reflection holography setup (z_1 , z_2 fixed) sets a limitation on the maximum extent of the sample we can image. This can also be applied in the reverse manner: a desired object extent determines the necessary z_1 and z_2 to avoid overlapping holograms.

Fig. 2-8(c,d) show the z_1 and z_2 required to fully image an object with a given extent without having overlapping holograms (Fig. 2-8(c) corresponds to a minimum position of the object $X_{s,min}=2$ mm and Fig. 2-8(d) to $X_{s,min}=8$ mm). As expected, increasing the object extent requires increasing z_1 and z_2 . By comparing Fig. 2-8(c) and Fig. 2-8(d) we can see how increasing $X_{s,min}$ reduces the required sample - mirror spacing z_2 but increases the required source - sample spacing z_1 . This is because, for a given object extent, increasing $X_{s,min}$ (and therefore $X_{s,max}$) requires increasing the FoV of the LED source, which can be achieved by increasing z_1 . As is apparent from both Fig. 2-8(c) and Fig. 2-8(d), sample sizes above 5 mm can be imaged while still maintaining a compact form factor below 20 mm.

Using reflection lensless holography for accurate depth estimation and proximity sensing

From our analysis above it is important to realize that the ratio of magnifications between the two holograms allows us to directly extract the ratio z_2/z_1 . This is equivalent to stereo imaging [189], but instead of using two cameras separated by a known distance to obtain image disparity, in our case such disparity comes from the path length difference between the two imaging paths. As in stereo imaging, it is unfortunately not possible to de-embed z_2 and z_1 directly - a calibration step is required to do so.

The fact that our light sources are made in a CMOS process offers an easy solution to add the required calibration step, which we depict in Fig. 2-9(a). We can take advantage

of the availability of metals in the CMOS dielectric stack above the silicon and polysilicon layers to fabricate a small calibration structure with a known distance to the LED (and thus a known $z_{1,cal}$)⁴. It is important to note that the uncertainty in $z_{1,cal}$ for this structure is very small ($<1 \mu\text{m}$) thanks to the high reproducibility and control of the CMOS fabrication process.

With this, all reflection holograms will contain, besides the double hologram of the sample, a small double hologram from the calibration structure (Fig. 2-9(b)). By measuring the magnification ratio of the calibration holograms we can accurately measure the calibration structure - mirror distance $z_{2,cal}$, and with this we can accurately extract the LED - mirror distance $z_1 + z_2$, which is fixed regardless of the position of the imaged sample. Knowing $z_1 + z_2$ and the ratio of magnifications for the sample of interest we can determine very accurately the values of $z_{1,sample}$ and $z_{2,sample}$. Thus, we can obtain very accurate depth estimation using our reflection lensless holography configuration.

As it is apparent from Fig. 2-9(a), the small distance between the LED and the calibration structure makes capturing the two calibration holograms challenging, requiring a large imaging array. An approach to overcome such a limitation is depicted in Fig. 2-9(c): use a configuration where only hologram 2 of the calibration structure is captured. We do not need to capture hologram 1 of the calibration structure because we can easily calculate its extents by realizing that (1) the dimensions of the calibration structure are well known (again, thanks to the high fabrication quality of the CMOS process), and (2) the propagation distance of hologram 1 is very small ($z_{1,cal} \approx 5 \mu\text{m}$). This makes it so that there is almost no diffraction of hologram 1, and therefore we can assume that its dimensions are the dimensions of the calibration structure itself. With this, we can extract the magnification ratio of the calibration structure and proceed as described

⁴It is of course necessary to make sure that the calibration structure is small enough to not occlude a large part of the FoV of the imaging system.

above.

The proposed structure can also be utilized to realize cheap proximity sensors capable of measuring small distances. This is depicted in Fig. 2-9(d), where we only use the calibration structure to measure the distance from the LED to a reflective surface. The main advantage of such a configuration is its accuracy: since we have small uncertainty in the value of z_1 thanks to the high CMOS fabrication process control, the only source of error is the uncertainty in extracting the magnification ratio. Figure 2-9(d) plots the relative error in the extraction of z_2 assuming a z_1 of $10 \mu\text{m} \pm 0.5 \mu\text{m}$ (a reasonable distance and uncertainty if we were to use the top metal of the dielectric stack to realize the calibration structure) and different uncertainties in the estimation of the magnification ratio between the two holograms. Relative errors on the order of 10% are achievable for reasonable magnification ratio uncertainties of $\pm 5\%$. Thus, we can locate a sample $100 \mu\text{m}$ away from the LED with an accuracy of $\pm 10 \mu\text{m}$.

The distance measurement range is only limited by the FoV of the imaging sensor, since in order to calculate the magnification we need hologram 2 of the calibration structure to be fully contained within the camera FoV. If we assume a $2 \text{ cm} \times 2 \text{ cm}$ FoV, we get a maximum measurement range of 2 cm. Our reflection holography distance sensor can thus achieve a $\pm 5\%$ measurement accuracy over a $\sim 0 \text{ cm} - 2 \text{ cm}$ range.

Compared to alternative distance sensors with comparable measurement range (interferometer-based laser sensors or capacitive sensors), our approach offers several advantages:

1. Since our sensor is self-calibrated by virtue of our knowledge of z_1 , a separate calibration step or structure is not necessary. In contrast, interferometric sensors require a well-known calibration path and capacitive sensors generally require calibration with a known target for accurate measurements.
2. It does not require complex readout instrumentation - readout is achieved with

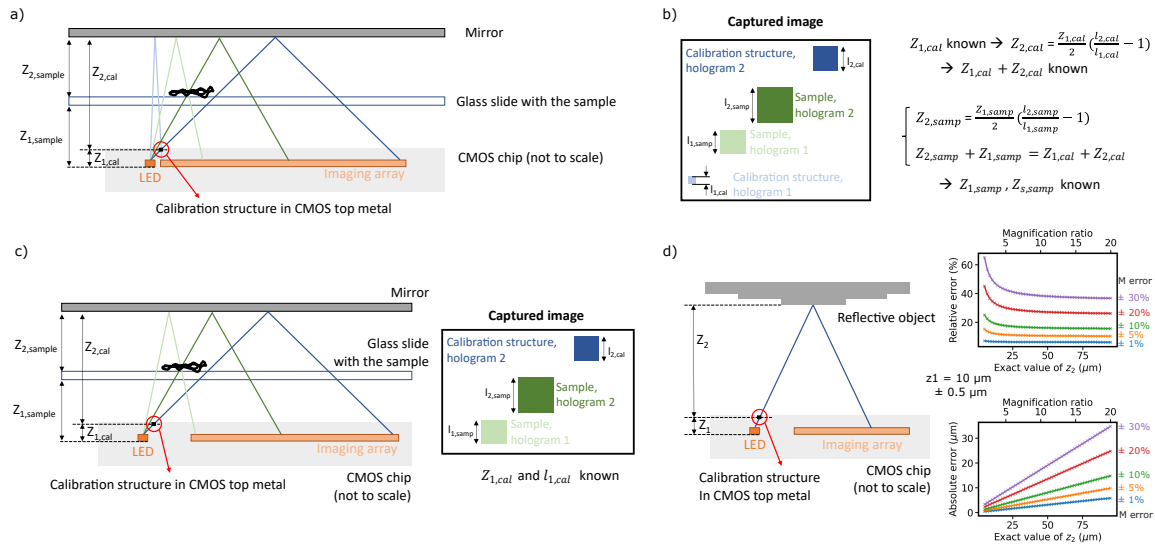


Figure 2-9: Depth estimation with reflection holography. (a) Experimental configuration, where a small calibration structure is added in the top metal of the CMOS stack with a known $z_{1,cal}$. (b) Besides the sample holograms, the captured reflection hologram contains the two subholograms of the calibration structure, which allows us to extract $z_{1,samp}$ and $z_{2,samp}$. (c) An alternative configuration to ease the geometric requirements is to not capture subhologram 1 of the calibration structure, whose dimensions we know. (d) The same approach without a sample can be used as a proximity sensor. On the right, the absolute and relative uncertainties in the estimation of z_2 are shown as a for different magnification ratio uncertainties. We assume $z_1 = 10 \mu m$ with an uncertainty of $\pm 0.5 \mu m$.

a simple CMOS camera. In contrast, both capacitive and interferometric sensors require very low noise voltage sensors for accurate readout.

3. It is a single-component sensor: it only requires a CMOS chip, where LED, calibration structure and camera are integrated. In contrast, capacitive sensors require complex readout electronics, and interferometric sensors require free-space optical components such as beam splitters. Thus, our sensor can achieve lower cost and form factor.
4. These advantages come at the cost of measurement accuracy. High end interferometric and capacitive sensors can achieve measurement resolutions below $1 \mu\text{m}$, something that is hard for our sensor to achieve.

We note that our proposed distance sensor works with a single LED source. The use of an LED array with closely spaced emitters does not have an inherent advantage with respect to the use of a single LED, beyond the fact that we can turn all of them at the same time to increase the emitted power and thus the image SNR. Nevertheless, having multiple LEDs at different edges of the CMOS imaging array could allow distance measurements at different locations and therefore the extraction of relative angle between the reflective surface and the distance sensor.

While we will not explore it in this thesis, it is likely that the presence of two slightly different holograms of the same sample at the image plane can be combined with computational techniques to increase resolution and/or extract additional information about the sample and/or experimental setup beyond depth estimation.

2.3.2 Demonstration of subpixel shift reflection holography with CMOS LED arrays

As discussed, the CMOS LED array we designed and characterized in Section 2.2 allows for the realization of subpixel shift reflection holography in a compact and low cost manner. Figure 2-10 shows the necessary steps to realize subpixel shift holography:

1. Acquisition of holograms with each LED of the array turned on.
2. Extraction of the subpixel shift between the acquired images.
3. Generation of a high resolution hologram from the acquired holograms.
4. Reconstruction of the sample from the high resolution hologram.

We will describe each of these steps in detail over the next sections.

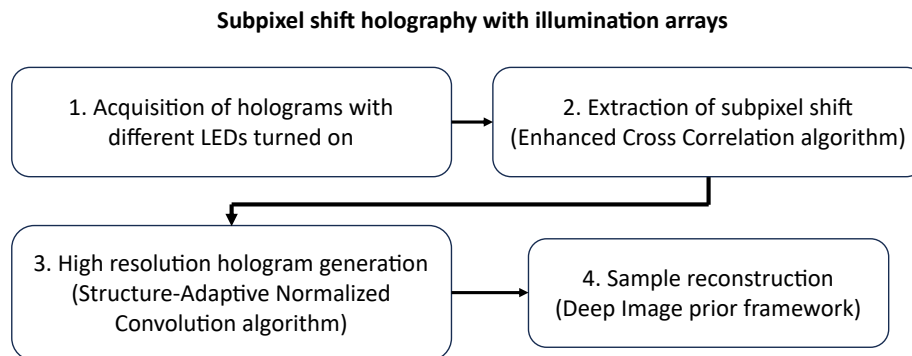


Figure 2-10: Process to realize subpixel shift holography with illumination arrays.

Experimental setup

Figure 2-11 shows the reflection holography setup that we built. While the LED and imaging arrays can be fabricated in the same chip, for our proof of concept we use

two distinct CMOS chips, one containing the LED array (presented in Section 2.2) and another one containing the CMOS imaging array (in particular, the commercially available Panasonic MN34230ALJ CMOS imager). A schematic of the chip arrangement is shown in Fig. 2-11(a), along with pictures in Fig. 2-11(b, c): we mounted the LED array on a transparent 1 mm thick glass carrier to place it on top of the imaging array without occluding its FoV.

It is worth noting that our preliminary configuration is slightly different to the one we presented in the previous section (Fig. 2-8), as there is an additional geometric parameter z_3 corresponding to the vertical separation between the imaging and LED arrays (Fig. 2-11(a)). This difference does not change the requirements to achieve a given FoV (Eq. 2.7 and Figs. 2-8(c, d)), but it does affect the magnification and shift of the two holograms formed at the image plane. Now we have:

$$\frac{M_1}{M_2} = \frac{\frac{2(z_1+z_2)+z_3}{z_1+2z_2}}{\frac{2(z_1+z_2)+z_3}{z_1}} = \frac{z_1}{z_1 + 2z_2} \quad (2.8)$$

$$\frac{\Delta_{i,1}}{\Delta_{i,2}} = \frac{d_i \frac{z_1+z_3}{z_1+2z_2}}{d_i \frac{z_1+2z_2+z_3}{z_1}} = \frac{z_1(z_1 + z_3)}{(z_1 + 2z_2)(z_1 + 2z_2 + z_3)} \quad (2.9)$$

If we compare these equations to the ones we obtained for the fully integrated case (Eqs. 2.5, 2.6) we can see how the individual magnifications and shifts M_1 , M_2 , $\Delta_{i,1}$, $\Delta_{i,2}$ and the shift ratio $\Delta_{i,1}/\Delta_{i,2}$ are different, but the magnification ratio M_1/M_2 is the same.

Fig. 2-11(d) shows a sample hologram of a droplet of 20 μm diameter beads acquired with our reflection setup, as well as an optical micrograph taken with a regular microscope at $5\times$ magnification. As discussed, the two distinct holograms of the same sample are clearly visible, spatially separated and have a different magnification.

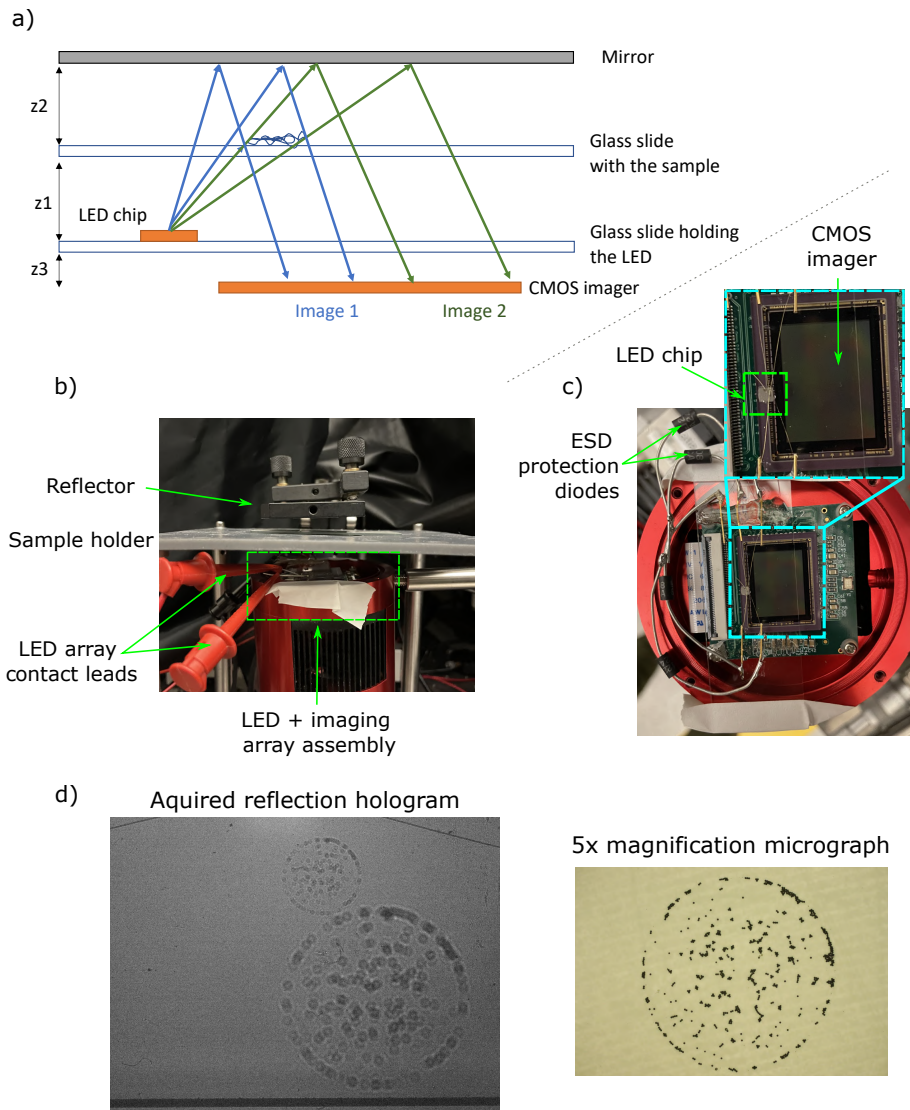


Figure 2-11: CMOS LED array reflection holography setup. (a) A schematic of the setup, where the LED and imaging chips are in different substrates. (b) Picture of the setup. The LED + imager assembly, the sample holder and the reflector are visible. (c) Picture of the LED + imager assembly. A transparent glass slide holds the CMOS chip with the LED array on top of the CMOS imager. (d) A sample hologram acquired with the setup, corresponding to $20\ \mu\text{m}$ diameter beads. A $5\times$ magnification micrograph taken with a regular microscope is also shown.

Estimating subpixel shift

An important step to exploit the benefits of subpixel shift holography is to accurately estimate the shift between the holograms acquired with the different LEDs in the array, a concept usually referred to as image registration.

As we have discussed above, the magnitude of the shift does not only depend on the spacing between elements in the source array, but also on the geometric arrangement of the holographic microscopy setup. While it is in principle possible to calculate the shift by accurately measuring all the geometrical parameters of the setup (z_1 , z_2 and z_3), such an approach requires calibration every time a sample is switched and is prone to experimental measurement uncertainties.

It is therefore desirable to resort to computational techniques that can extract the shift by directly comparing the acquired holograms. Since this is a problem that appears in multiple research and commercial applications, a plethora of techniques have been reported in the literature [190].

Focusing on algorithms for applications in subpixel shift holography, three main approaches have previously been used in the literature:

- Cross-correlation between shifted images: The shift is defined as the location of the peak of the upsampled cross-correlation between the two images. Such upsampled cross-corelation is usually calculated by means of a Fourier transform for computational efficiency [191, 192].
- Maximum A Posteriori (MAP) registration: The subpixel shift is obtained by maximizing the likelihood of observing the recorded holograms according to a given observation model [65, 193].
- Feature-based image registration: The shift is calculated by obtaining displacement

vectors between *key subregions* in the recorded holograms [32].

Of course, each approach has its advantages and drawbacks, which usually involve a tradeoff between accuracy, computation time and robustness to noise. While traditional holographic microscopy approaches record high SNR images (at the expense of increased complexity, size and cost), our recorded holograms have a limited SNR due to the relatively low power of our CMOS LEDs. Thus, our choice of image registration algorithm needs to be robust to the presence of noise.

We evaluated the accuracy of a variety of image registration algorithms under different noise scenarios using synthetic data simulating our experimental conditions. To do so, we took a subregion of a hologram acquired with our reflection holography setup using a single CMOS LED in our array, and computationally generated a fixed, known subpixel shift. We then synthetically added noise of increasing variance to the shifted image (Fig. 2-12(a)) and evaluated the difference between the shift obtained by the various algorithms and the known values. The results are shown in Fig. 2-12(b)⁵. As we can see, the enhanced cross correlation (ECC) method [194] is the one that performs the best under our experimental noise conditions, which we estimated by evaluating the Peak Signal to Noise Ratio (PSNR) between a raw image taken without a sample and the same image with heavy denoising applied to it⁶. We can see how the ECC algorithm achieves a shift accuracy better than 0.02 pixels regardless of the noise strength. Mutual information and optical flow algorithms perform well under high SNR conditions, but start to fail for noisy images.

We therefore used the ECC algorithm [194] (as implemented in the *opencv* python

⁵We note that we do not evaluate the use of MAP registration due to the long computation time required, especially for large images.

⁶Note that applying heavy denoising to such an image does not distort it since the illumination pattern of the LED without a sample should be really smooth. Any high frequency content is due to noise.

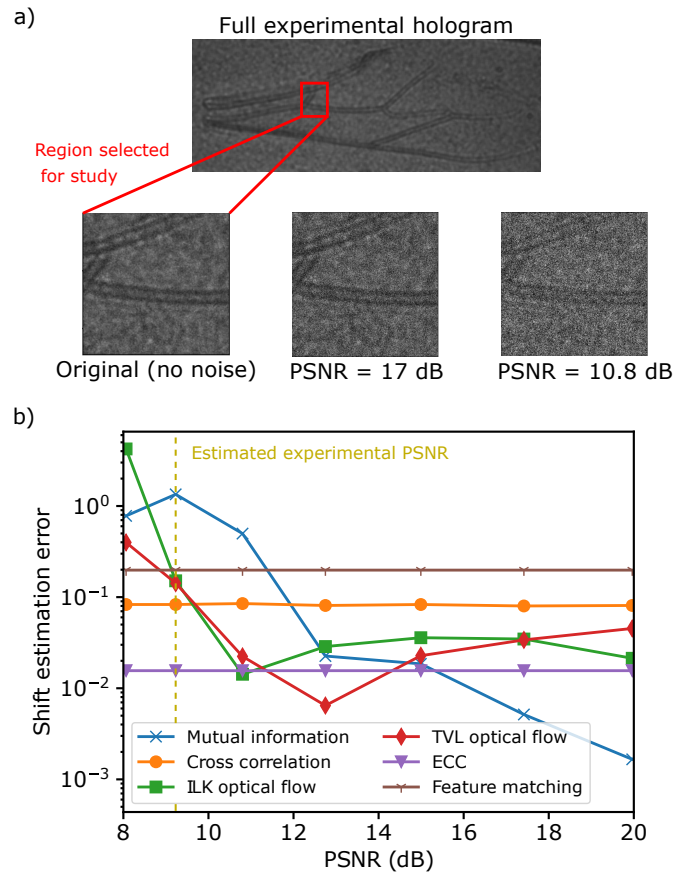


Figure 2-12: Robustness to noise of different image registration algorithms. (a) We use a region of an experimentally acquired hologram of a honeybee wing and synthetically add noise. (b) The error in the extracted shift as a function of the Peak SNR (PSNR) of the image. Each data point includes the average registration error over 10 repeats of the algorithm with 10 distinct synthetic noise addition processes of the same variance.

package [195]) to perform image registration for our four different acquired holograms (one per each CMOS LED in the array, as shown in Fig. 2-13(a) for a mantis wing sample). Figure 2-13 shows the results for five different samples.

A few observations need to be made here:

- As shown in the second row of Fig. 2-13, the obtained subpixel shifts are different for each sample. This is expected, since as we described in detail in Section 2.3.1,

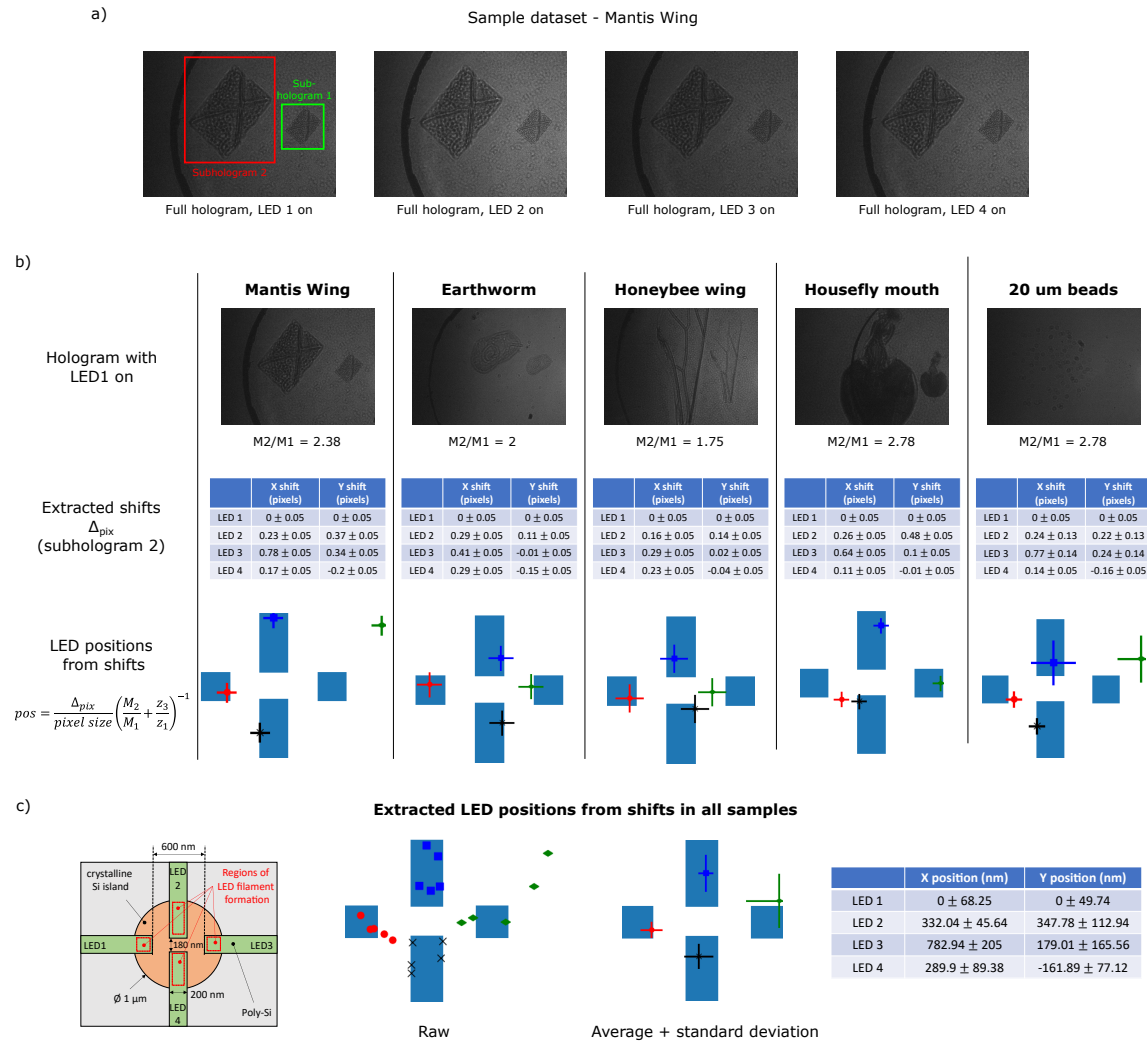


Figure 2-13: Evaluation of subpixel shift in five different samples acquired with our CMOS LED array reflection holography setup. (a) Sample dataset of a mantis wing, which contains four different holograms, each acquired with a different CMOS LED turned on. (b) Extracted subpixel shifts for subhologram 2 for the five different samples studied. The first row shows the hologram acquired with LED 1 turned on, and lists the magnification ratio between subholograms 1 and 2. The second row shows the extracted subpixel shifts between the different holograms in the dataset. The third row shows the LED positions calculated from the subpixel shift, superimposed into the regions of LED filament formation set by the overlap between the poly-Si and the c-Si island in the fabricated device. (c) The calculated LED positions using the extracted subpixel shifts for all the studied samples. A schematic of the LED array layout is shown for reference.

the shift does not only depend on the spacing between LEDs in the array but also on the geometrical parameters of the setup, which change depending on the sample. In our experiments we maintained both the source - sample distance z_1 and the imager - source distance z_3 fixed, and increased the sample - mirror distance z_2 for larger samples (which require a larger FoV).

As expected, samples that have a larger z_2 (which are the samples that have a larger magnification ratio M_2/M_1 - Eq. 2.8) show larger subpixel shifts.

- As described in Section 2.3.1 and shown in Fig. 2-13(a, b), our acquired images contain two holograms of the same sample, each with a different magnification and shift. Nevertheless, the geometrical configuration we used experimentally (with a total imager to mirror distance $z_1 + z_2 + z_3$ below 1 cm) results in subhologram 1 having shifts that are below 0.03 pixels and therefore not detectable. This is why we only show subpixel shift values corresponding to subhologram 2.
- We can easily calculate the LED positions from the extracted hologram shifts by rearranging Eqs. 2.8 and 2.9 and realizing that:

$$\Delta_{i,2} = \frac{d_i}{T} \left(\frac{M_2}{M_1} + \frac{z_3}{z_1} \right) \quad (2.10)$$

Above, besides our previously defined variables, $\Delta_{i,2}$ is the extracted shift in units of pixels and T is the pixel size (3.8 μm in our setup). The magnification ratio between subholograms M_2/M_1 can easily be extracted from the acquired data, and its value is shown in the first row of Fig. 2-13(b).

The third row of Fig. 2-13(b) shows the calculated LED positions for each sample given the extracted shifts and magnification ratios. These are superimposed on the

regions of overlap of the poly-Si access areas and the c-Si island, which as described earlier are the possible areas of LED filament formation (Fig. 2-5(b)).

- Figure 2-13(c) shows the extracted LED positions using the hologram shift data from all the samples we studied. Along with the data shown in the third row of Fig. 2-13, we can see how the extracted LED positions are reasonably consistent between samples. Deviations are due to the inaccuracies of the cross-correlation algorithm under the limited SNRs of our experimental samples, as well as possible small movements of the stage / camera between hologram acquisitions with different LEDs turned on.

It is also worth noticing how our simple reflection imaging configuration (consisting of a camera, a mirror and an LED array, with no lenses or any bulk optical element) allows us to determine the location of the emitters with an accuracy below 200 nm. Doing so with a conventional imaging setup would require an expensive, bulky and high magnification microscope (such as the one we used to directly characterize the LED array in Section 2.2).

High resolution image generation

Once we have extracted the subpixel shift between the acquired images we need to construct the High Resolution (HR) image, which will hopefully capture more features than the individual images and therefore lead to better object reconstruction.

There exist a great variety of methods for generating such an HR image, with their own set of advantages and drawbacks (Ref. [196] contains a good review of the main approaches). Similarly to what we did for the subpixel shift registration methods in the previous section, we evaluated the performance of various algorithms for high resolution image generation using a synthetically generated hologram of beads (Fig. 2-14(a)). From

that image we generated a set of $N=100$ subpixel shifted images with 2x lower resolution - one set without added noise (Fig. 2-14(b)) and one set with gaussian noise added (Fig. 2-14(c), closer to our experimental conditions).

We then used different methods to generate the high resolution image (and based our code on that provided in [197, 198]) ⁷:

- Interpolation: It performs simple bicubic interpolation assuming that the intensity at the position based on the subpixel shift is the recorded pixel value.
- Projection Onto Convex Sets (POCS): The problem of finding the HR image is formulated as that of finding the intersection of a set of convex constraints based on the observed low resolution images [199].
- Iterated Backprojection: One of the most common methods used for superresolution image generation, it iteratively generates a better approximation of the original HR image by modeling the imaging process and comparing the experimentally acquired images to the corresponding low resolution images generated from the current HR approximation [200].
- Regularized Superresolution: A similar approach to iterated backprojection, but it uses an adaptive regularizer with better robustness to noise and edge preserving properties [201].
- Structure-Adaptive Normalized (SAN) Convolution: Uses the normalized convolution framework (where a local approximation of the signal is made by projection into a set of basis functions, similar to a Taylor expansion). The implementation

⁷Detailed description of the different methods is out of the scope of this thesis. The reader is pointed to the references included for each method for a detailed description. We note that this is not an exhaustive list of superresolution image generation approaches.

we use employs a set of basis functions that are robust against noise and discontinuities [202].

We used the described methods with a varying number of subpixel shifted images and compared the resulting high resolution image to the ground truth image (Fig. 2-14(a)) through the use of the Pearson Correlation Coefficient. The results are shown in Fig. 2-14(d) for the no noise case, and in Fig. 2-14(e) for the case with noise.

A few conclusions can be drawn from Figs. 2-14(d,e):

- As expected, all algorithms perform better when there is no noise present in the acquired low resolution images.
- The SAN convolution method performs the best regardless of the number of images and the presence of noise.
- POCS, iterated backprojection and regularized superresolution methods have a similar performance in the noiseless case, but iterated backprojection performs significantly better in the presence of noise.
- All methods but interpolation benefit from having a larger number of low resolution, subpixel shifted images, specially in the case where there is noise present in the acquired images. As is apparent from Figs. 2-14(d,e), on the order of 10 - 20 images are required to improve resolution (or equivalently, improve the correlation coefficient). This is consistent with previously published work on subpixel shift holographic superresolution. For example, reference [192] uses 32 low resolution images, ref. [32] uses 49 of them, [51] uses 23 and [203] uses 80.

Figure 2-15 shows the high resolution holograms reconstructed using the SAN con-

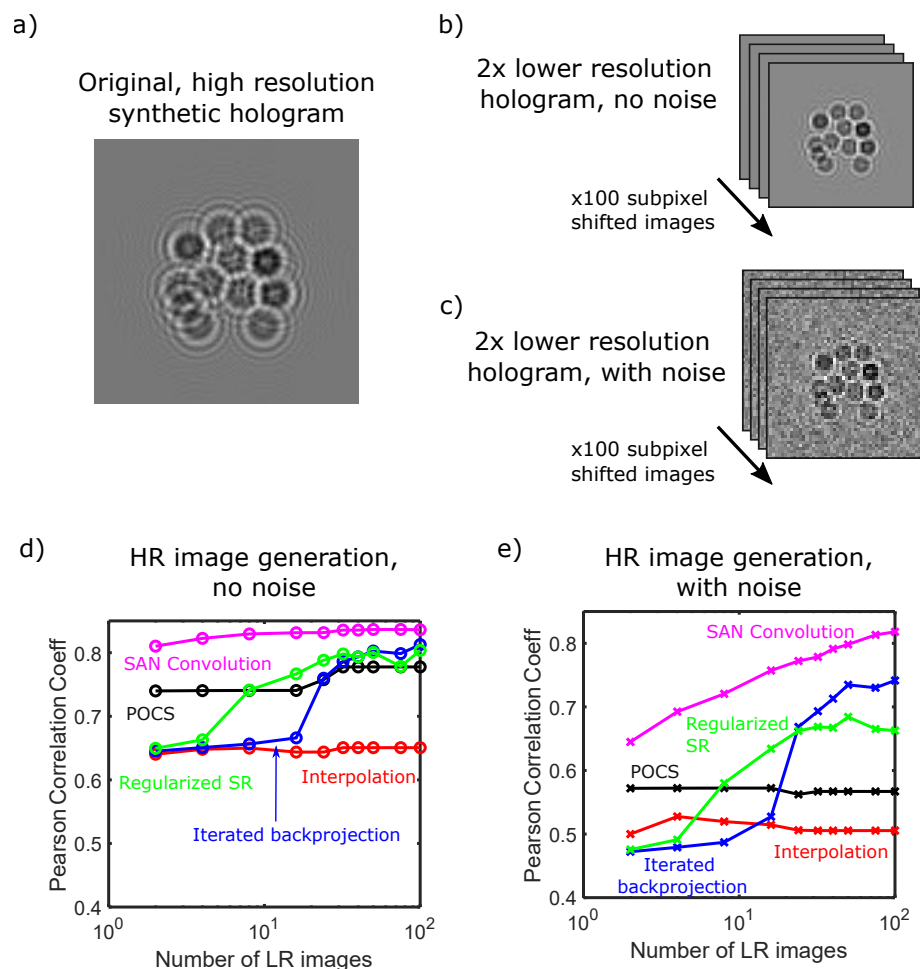


Figure 2-14: Evaluation of superresolution image generation algorithms with a synthetic hologram. (a) The high resolution synthetically generated hologram that is the ground truth. From the ground truth we generated datasets of 100 subpixel shifted holograms with 2x lower resolution, without noise (b) and with noise (c). We used different super-resolution image generation methods with increasing number of subpixel shifted images and compared the results to the ground truth image shown in (a). The Pearson correlation coefficients between the ground truth and the generated superresolution images are shown in (d) for the noiseless dataset and in (e) for the case with added noise.

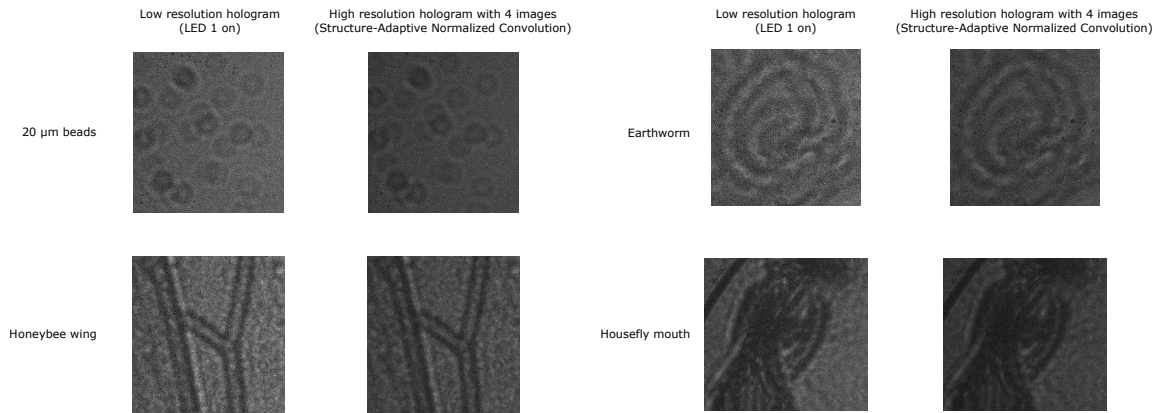


Figure 2-15: High resolution holograms generated using the Structure-Adaptive Normalized (SAN) Convolution algorithm for our experimental samples.

olution method for our experimentally acquired samples ⁸. From our last point above, 4 subpixel shifted images are not enough to recover higher resolution features (we would need a larger array for that), but as can be observed in the second column of Fig. 2-15, our super-resolution image generation method removes some of the noise present in the acquired holograms. Such noise removal is better than if we use the SAN convolution algorithm on a single image or if we just average the 4 acquired images. Thus, even though with 4 images we cannot perform superresolution, we can achieve a higher SNR (Fig. 2-16) ⁹.

Sample reconstruction

We are now ready to perform sample reconstruction of our acquired holograms, which are shown in Figs. 2-17 and 2-18.

For each sample we have three different holograms:

⁸It is worth remembering that we can only perform high resolution hologram generation for subhologram 2, since subhologram 1 has negligible subpixel shift

⁹We estimated the PSNR of each hologram by extracting the noise variance from a region of the image that does not contain sample features.

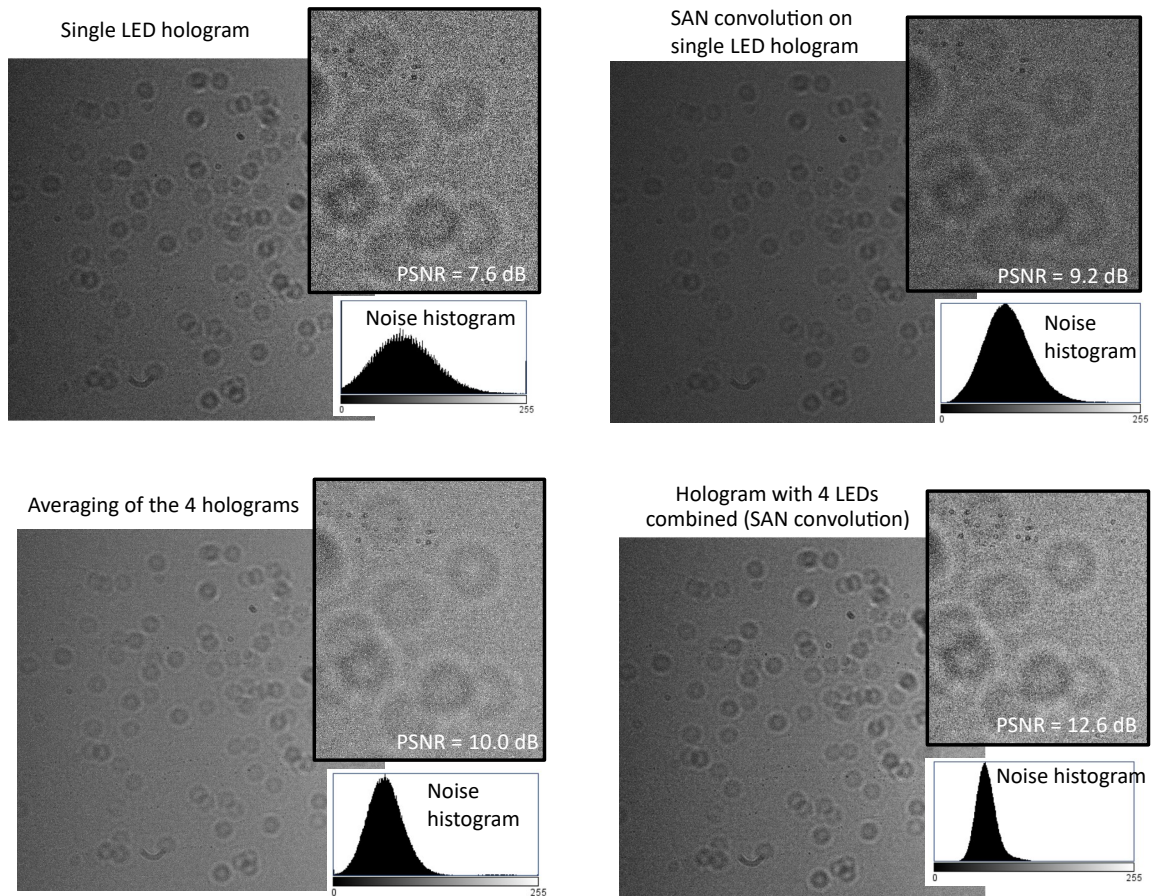


Figure 2-16: Peak Signal to Noise Ratio (PSNR) of different approaches for noise suppression. For each approach the full hologram, a closeup and the noise histogram are shown. The noise histogram is obtained by focusing on a region of the image without sample features. For the 4 LED SAN convolution approach (bottom right corner) a 2x reduction of image size through linear interpolation is performed so it has the same size as the other three approaches.

1. Subhologram 1, with the corresponding reconstructions shown in Fig. 2-17.
2. Low resolution subhologram 2 (corresponding to a single hologram acquired with one of the LEDs turned on), with the corresponding reconstructions shown in the center column of Fig. 2-18.
3. High resolution subhologram 2, constructed using the SAN convolution algorithm (as discussed in the previous section) using the 4 images acquired with different LEDs turned on and the extracted subpixel shifts. The corresponding reconstructions are shown in the right column of Fig. 2-18.

For each hologram we perform two reconstructions:

- A baseline reconstruction using the angular spectrum method [59], where the recorded hologram is transformed to the frequency domain through fast Fourier transform (FFT), back-propagated to the image plane in the frequency domain and then converted back to the space domain with inverse FFT.

As discussed earlier, while simple this approach has several limitations including the assumption of a monochromatic source (which is not true in our case), and the presence of a twin image.

- A reconstruction using a modification of the DIP algorithm developed by Dr. Ik-sung Kang, which as discussed in Section 2.1.3 can overcome the limitations of traditional reconstruction methods.

Since our sources have a large NA, as the source - sample distance increases the power density at the CMOS imaging array decreases, resulting in a lower power integrated over a single pixel and therefore a lower SNR.

Subhologram 1

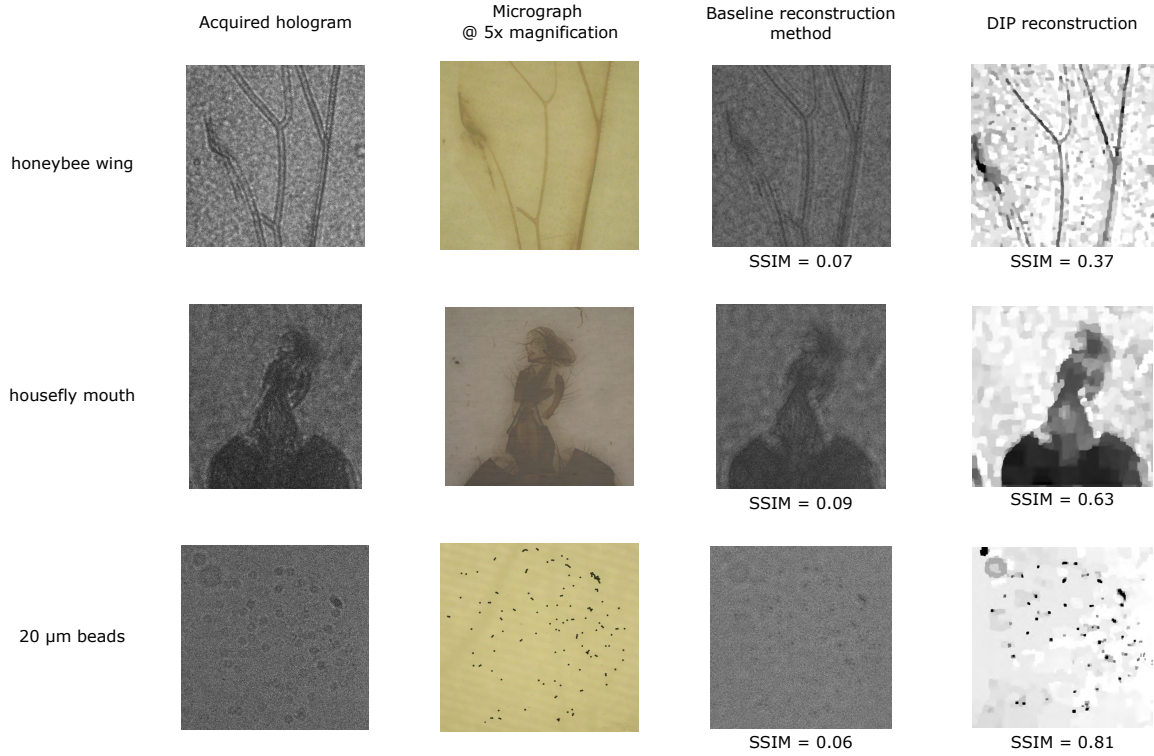


Figure 2-17: Hologram reconstructions of subhologram 1 for three different samples. The acquired hologram (first column), a micrograph taken with a 5x magnification objective (second column), the reconstruction using a baseline backpropagation algorithm (third column) and the reconstruction using our DIP algorithm (fourth column) are shown. The Structural Similarity (SSIM) index between the micrograph and each of the reconstructions is also shown.

Since in our reflection configuration the total source - sample distance is significantly larger than in a transmission lensless setup, the achieved hologram SNRs are lower (compare Fig. 2-4 to Fig. 2-13) and therefore less interference fringes are resolved. This results in the original DIP algorithm producing spurious sample features to "smear out" these lost interference fringes (see reconstruction on the top row of Fig. 2-19). Essentially, the original DIP algorithm takes the absence of interference fringes as a real sample feature, and therefore produces a sample reconstruction that results in a hologram without the unresolved fringes even in the absence of noise in the hologram acquisition.

To prevent the generation of spurious features and improve the robustness of the DIP algorithm to low SNRs I included the presence of noise in the generation of the approximated hologram: after generating the estimated hologram \hat{I}_{det} , additive, white gaussian noise (AWGN) is generated and added to the hologram so that the loss function is computed on $\hat{I}_{det,noise} = \hat{I}_{det} + N(0, \sigma)$ (bottom row of Fig. 2-19). This step prevents the DIP algorithm from overfitting to the noise present in the acquired holograms and improves significantly the reconstruction quality as seen in the bottom row of Fig. 2-19.

Such noise addition step also makes the use of the total variation regularizers used in the original DIP architecture unnecessary. As a result, the number of regularizers (and associated hyperparameters) is reduced. Our new loss function is the following:

$$\begin{aligned}
\mathcal{L}(\hat{I}_{\text{det}}, I_{\text{det}}; \theta_\alpha, \theta_\varphi, \Delta z, \gamma_1, \gamma_2, \dots, \gamma_N) &= \text{MSE}(\hat{I}_{\text{det}}, I_{\text{det}}) + \kappa_1 \text{NPCC}(\hat{I}_{\text{det}}, I_{\text{det}}) \\
&+ \kappa_2 \left(1 - \sum_{n=1}^N \gamma_n\right) + \kappa_3 \sum_{n=1}^N \text{ReLU}(-\gamma_n)
\end{aligned} \tag{2.11}$$

And the hyperparameters are fixed for all the samples to $\kappa_1 = 10$, $\kappa_2 = 10$ and $\kappa_3 = 25$. Notice how we eliminated the TV regularizers, which allowed us to drop the hyperparameter κ_4 which had to be tuned for each sample in the original DIP framework presented in Section 2.1.3 [181]. Thus, our new DIP framework with noise addition avoids the need to tune any the hyperparameters to the specific sample being reconstructed.

We set the variance σ of the AWGN in the novel DIP framework to the estimated noise variance of the experimental hologram. Such estimated variance can easily be obtained by looking at the pixel variance in regions of the experimental hologram that do not have sample features (i.e., the background of the hologram).

To quantitatively evaluate the quality of the resulting reconstructions we use the structural similarity (SSIM) index [204] between the obtained computational reconstruction and a micrograph of the sample taken with a 5x magnification microscope.

Several important conclusions can be drawn from Figs. 2-17 and 2-18:

- As expected, subhologram 1 (Fig. 2-17) has a lower magnification than subhologram 2 (Fig. 2-18). Therefore, for the same reconstruction area, a larger portion of the sample is contained in subhologram 1 compared to subhologram 2.
- The use of the DIP algorithm results in significantly higher quality reconstructions

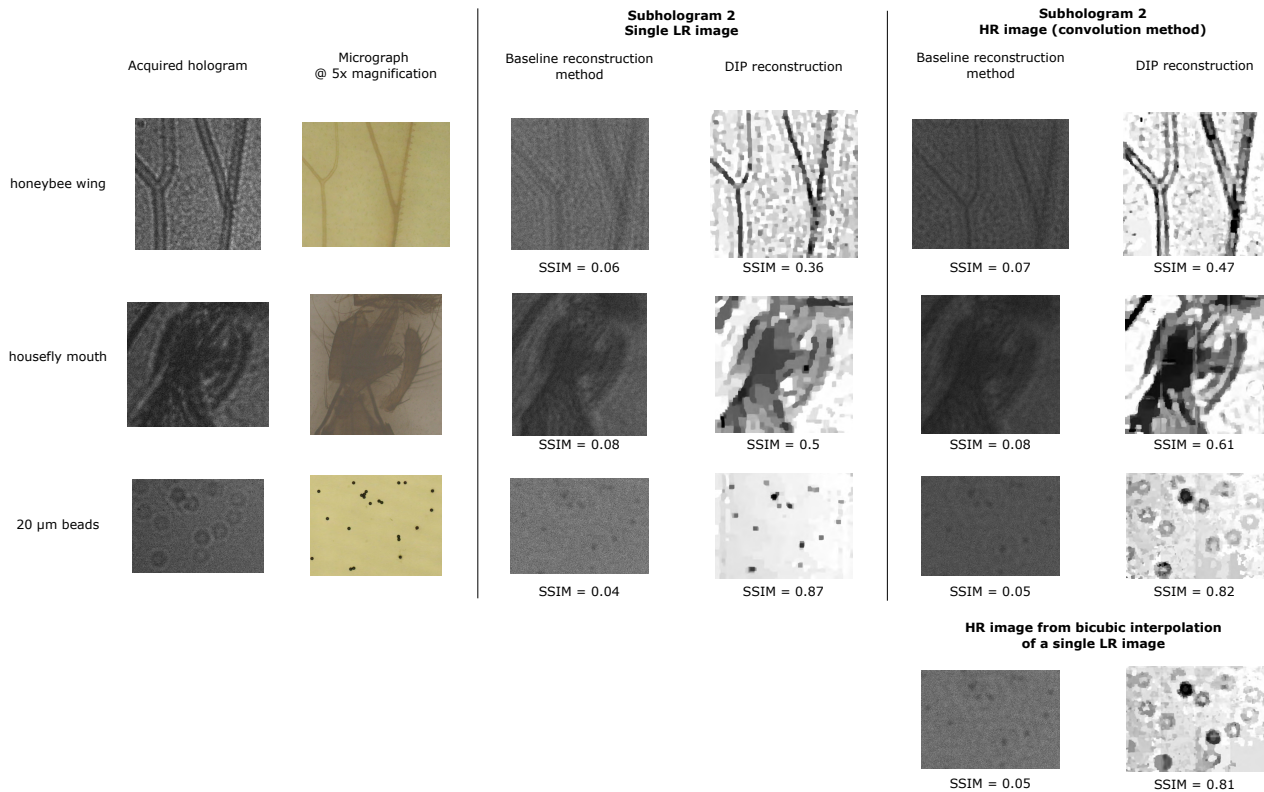


Figure 2-18: Hologram reconstructions of subhologram 2 for three different samples. Only a sub-region of the full hologram reconstruction is shown. Reconstructions for a single low resolution hologram (center) and for a high resolution hologram generated through the Structure-Adaptive Normalized Convolution algorithm (right) are shown. For each sample we show reconstructions using a baseline backpropagation method and our DIP algorithm. For the 20 μm beads sample, we also show reconstruction of a high resolution hologram obtained through bicubic interpolation of a single low resolution hologram. The Structural Similarity (SSIM) index between the micrograph and each of the reconstructions is also shown.

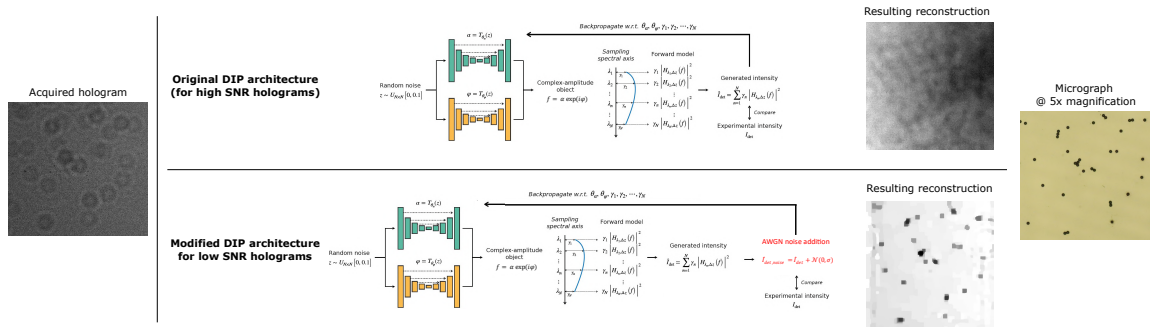


Figure 2-19: Modified DIP architecture for noise robustness, where we include the presence of noise as an extra step to the hologram generation. This significantly improves reconstruction quality.

(higher SSIM index) compared to the baseline backpropagation method, including higher contrast, enhanced resolution and the removal of the twin image. This is expected as discussed in detail in Section 2.1.3.

- The use of HR holograms results in increased resolution of the reconstruction (Fig. 2-18). For example, in the case of the 20 μm diameter beads, the use of a HR hologram allows differentiation of the transparent center of the bead from the dark edges (which is what is observed in the optical micrograph). Similarly, in the case of the honeybee wing, the HR reconstruction correctly shows the transparent center of the veins and their dark edge.
- While the use of HR holograms results in higher resolution reconstructions, we cannot observe an advantage when we use the HR image generated from the subpixel shifted holograms compared to a HR image generated by simple bicubic interpolation of a single LR hologram. This is shown for the 20 μm diameter beads sample in the bottom right of Fig. 2-18: no difference in resolution or reconstruction quality is observed between the two reconstructed samples.

This is expected: as we discussed in Section 2.3.2, we need on the order of 20

subpixel shifted images to observe a resolution increase, whereas we only have 4 due to the limited size of our CMOS LED array.

2.4 Conclusion

In this chapter we have demonstrated for the first time the use of a subwavelength size silicon surface emitting LED to realize lensless holography in a highly compact form factor and without the need of a pinhole.

To address the limitations posed by the limited temporal coherence of the LED we have used a novel DIP algorithm developed by Dr. Iksung Kang capable of blindly recovering the illumination spectrum and simultaneously improve the reconstruction quality and resolution.

Since the silicon LED is made in a CMOS process, we can take advantage of the high fabrication control and the availability of high resolution metals with feature sizes below 100 nm to realize arrays of such LED sources.

In this regard, we experimentally demonstrated, to the best our knowledge, the first array of subwavelength-sized emitters with subwavelength spacing realized in a CMOS platform, and the smallest spacing ever demonstrated in an array of individually and electrically addressable nanoLEDs when compared to its emission wavelength (Table 2.1). While our LED array size is limited, our architecture is highly scalable thanks to the high scalability and fabrication quality of CMOS fabrication processes.

Another advantage of our sources being made in a CMOS process is that we can realize the imaging array required to record the hologram in the same chip. To take advantage of this, we proposed, analyzed and demonstrated a new lensless holography configuration that relies on the reflection off of a mirror to allow for light generation and hologram acquisition in the same chip, resulting in a simpler and cheaper system.

With this, we first discussed how this configuration can be used to realize self-calibrated proximity sensors capable of measuring distances below a $100\ \mu\text{m}$ with high accuracy. We then coupled the reflection configuration with our illumination source array to demonstrate the ability to obtain subpixel shifted holograms, which can enable resolutions beyond the limit set by the finite camera pixel size.

We also proposed a modification to the DIP-based sample reconstruction that significantly increases robustness to low SNR holograms by direct addition of noise to the approximated measurement generated by the neural network. This is, to the best of our knowledge, the first time that the direct addition of noise is used in a DIP architecture.

All in all, in this chapter we set up the foundations to exploit CMOS fabrication processes to realize large scale arrays of sub-wavelength size, sub-wavelength spaced illumination arrays. We also demonstrated how these arrays can enable new imaging configurations and novel functionalities.

Chapter 3

Native Single-mode Waveguide-coupled LEDs in Unmodified Silicon Photonics Fabrication Platforms

As discussed in detail in the introduction, integrated photonics holds a great potential to reduce complexity, cost and enable new functionality for imaging applications. In Chapter 2 we have exploited CMOS photonics to realize arrays of near-point surface emitting LEDs and demonstrated how they can enable subpixel shift holography in a very compact and simple configuration.

While very useful, the sources we studied in Chapter 2 are surface emitting, so we cannot take advantage of the capability of photonic chips to guide and process the generated light *within* the chip.

Generating waveguide coupled light in monolithic silicon photonics systems is very

challenging due to the indirect bandgap of silicon. To overcome this limitation, hybrid integration of III-V materials is possible [205], as well as coupling of the light generated by an external light source into the silicon chip (usually through optical fiber attach) [22]. While effective, these approaches require either complex fabrication processes or costly and time-consuming packaging techniques that reduce system yield and robustness and increase manufacturing costs.

In contrast, the development of a monolithic waveguide-coupled light source in silicon photonic fabrication platforms could enable the realization of integrated optical systems that do not require an external light source. Such a light source will intrinsically have a limited efficiency and output power due to the indirect bandgap of silicon, but it could be useful for a subset of applications where high temporal coherence, high signal to noise ratio or high speed are not critical.

This chapter deals with the modeling and experimental characterization of such native waveguide-coupled light sources in silicon photonics processes. We fabricate waveguide coupled LEDs in two different configurations (linear and resonant) and in three different CMOS photonic platforms (GlobalFoundries 45RFSOI, AIM photonics and GlobalFoundries 45SPCLO) generating waveguide-coupled optical powers in the hundreds of pW power levels. In the first part of this chapter we present the device designs (Section 3.1) and develop a detailed theoretical model accounting for both electrical and optical behavior (Sections 3.2 and 3.3). Then, in Section 3.4 we experimentally characterize the performance of these sources and demonstrate on-chip detection of the generated light using an integrated photodetector.

This is, to the best of our knowledge, the first time that a waveguide-couple light source has been realized in a CMOS photonic process without any modification to the process flow.

3.1 Device designs

Previous work on silicon waveguide-coupled LEDs relies on non-standard fabrication processes to achieve light emission. Libertino et al. demonstrated a waveguide-coupled LED with emission at 1550 nm based on Erbium doped silicon pn junctions [206]. More recent work from NIST has demonstrated narrow emission from waveguide-coupled Si LEDs at a wavelength of 1220 nm, both in straight waveguide (linear structures) [207] and resonant structures [208]. These devices are based on defect implantation of W centers and require operation at cryogenic temperatures, which greatly hinders the usability of these sources. Furthermore, both approaches suffer from increased waveguide loss due to the required defect/ion implantation, and show a degradation of performance at room temperature due to fast detrapping of carriers into the conduction band as thermal energy increases¹.

Instead, we pursue an alternative approach: realize silicon waveguide-coupled LEDs in commercial silicon photonics platforms without any modification to the fabrication process. Since our sources rely on light generation through processes that are intrinsic to crystalline silicon, the most important benefit of our approach is that our sources are *universal* in the sense that they can be realized in any CMOS photonic process that uses crystalline silicon optical waveguides.

To prove the universality of our approach, we demonstrate both linear and resonant silicon LEDs fabricated in three different silicon photonics fabrication platforms (GlobalFoundries 45RFSOI [20], a preliminary version of the currently commercially available AIM Photonics process [209], and GlobalFoundries 45SPCLO [210]). The three native silicon LEDs we designed and fabricated are shown in Fig. 3-1:

¹The waveguide-coupled optical power generated by these devices is not reported in any of the cited works.

1. Resonant strip waveguide LED: The first design (Fig. 3-1(a)) is a microring resonator LED based on interleaved pn junctions, similar to devices we have previously characterized for use as optical modulators [211]. The device has a $10\ \mu\text{m}$ outer diameter and is $1.2\ \mu\text{m}$ thick, and has a T-junction doping profile interleaved along the azimuthal direction [212]. The purpose of such a junction profile is to maximize the overlap of the p-n junction with the fundamental whispering gallery mode propagating along the ring (shown in the bottom left in Fig. 3-1(a)). Electrical contacts to the different sections are achieved through access spokes at the inner edge of the ring, so that metals are kept far away from the optical mode.

The device is fabricated in the commercially available GlobalFoundries 45RFSOI microelectronics process [20], which has a crystalline silicon layer with a thickness below 100 nm. Despite being a microelectronics process, we have previously demonstrated high performance optical and electro-optical devices in this process [213] and have shown monolithic integration with VLSI circuits [214].

2. Linear rib waveguide LED: The second design (Fig. 3-1(b)) is a linear LED based on a lateral pn junction fabricated at the SUNY Poly Albany NanoTech Complex, in a preliminary version of the current commercially-available AIM Photonics foundry process [209].

Since this is a fabrication process tailored for photonics, it has a thicker silicon (220 nm) and it includes a partial etch step which allows for the realization of rib waveguides. We therefore fabricated a linear rib waveguide with a total length of 2 mm, and used full etch sections at the edges of the waveguide to provide electrical contact while minimizing metal losses.

3. Resonant rib waveguide LED: The third design (Fig. 3-1(c)) is a resonant LED

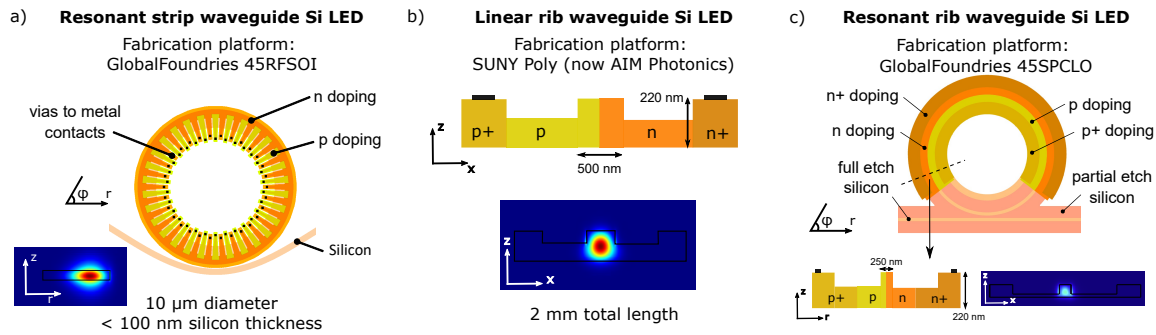


Figure 3-1: Native waveguide coupled silicon LED designs. (a) Ridge waveguide resonant LED design fabricated in the GlobalFoundries 45RFSOI process. The ring outer diameter is $10\ \mu\text{m}$, the ring width is $1.2\ \mu\text{m}$ and the silicon thickness is below $100\ \text{nm}$. The pn junctions are interleaved along the azimuthal direction. The lowest order optical mode propagating along the ring at a wavelength of $1130\ \text{nm}$ is shown in the bottom left corner. (b) Rib waveguide linear LED design fabricated in a preliminary version of the current AIM Photonics process based on a lateral pn junction. The total length of the device is $2\ \text{mm}$. The propagating optical mode at a wavelength of $1130\ \text{nm}$ is shown in the bottom. (c) Rib waveguide resonant LED fabricated in GlobalFoundries 45CLO process, which is based on a lateral pn junction. The ring diameter is $10\ \mu\text{m}$ measured from the outer edge of the full etch silicon. The propagating optical mode at a wavelength of $1130\ \text{nm}$ is shown in the bottom right corner.

based on a rib waveguide with a lateral pn junction (similar to the linear LED described above). The width of the full etch section of the rib waveguide is $250\ \text{nm}$, designed to be single mode at the emission wavelength of the LED (which as we will see is centered around $1130\ \text{nm}$). As shown in Fig. 3-1(c), doping is applied along the circumference of the ring except at the ring - bus coupling region. The ring radius is $5\ \mu\text{m}$ measured from the outer edge of the full etch section.

This device is fabricated in GlobalFoundries 45SPCLO [210], a fabrication platform tailored specifically for silicon photonics applications monolithically integrated with high performance electronics.

3.2 Operation principle and main physics

The main physical mechanisms affecting the performance of our waveguide-coupled silicon LEDs are depicted in Fig. 3-2.

We can break down the overall efficiency of our waveguide-coupled silicon LEDs into two terms:

$$\eta = \eta_i \eta_c \quad (3.1)$$

η_i is the efficiency of converting injected carriers into emitted photons (i.e., the efficiency of generating light from the injected electrical power or the internal quantum efficiency) and η_c is the efficiency with which these generated photons couple into the optical mode supported by the waveguide (i.e., the efficiency of extracting the generated photons into the waveguide mode of interest).

η_i is set by the competition between radiative and non-radiative mechanisms in the active region (the silicon waveguide). In our device, radiative recombination is achieved through phonon-assisted band-to-band recombination, while Shockley-Read-Hall (SRH) recombination both in the bulk of the silicon and at the silicon-oxide surfaces as well as Auger recombination all contribute to non-radiative recombination. Such competition can be easily understood using the *ABC* model [215], which approximates:

$$\eta_i = \frac{BN^2}{A^*N + BN^2 + CN^3} \quad (3.2)$$

Above, N is the number of injected carriers, B is the radiative recombination coefficient, which for silicon is about $4 \cdot 10^{-15} \text{ cm}^3/\text{s}$ at room temperature [216,217], C is the Auger recombination coefficient (which is about $1 \cdot 10^{-31} \text{ cm}^6/\text{s}$ for Si [218]) and A^* is the SRH recombination coefficient, which can be broken down as $A^* = A + S \cdot \frac{\text{area}}{\text{volume}}$

and includes both bulk (first term) and surface (second term) recombination. A is the bulk SRH recombination coefficient and S is the surface recombination velocity.

Equation 3.2 clearly shows that achieving high efficiency not only requires a high B coefficient (i.e. having fast radiative recombination) but also requires suppressing non-radiative recombination mechanisms (i.e. having small A^* and C). Since silicon is an indirect bandgap material, its B coefficient is significantly smaller than that of typical III-V materials used for the realization of LEDs ($4 \cdot 10^{-15} \text{ cm}^3/\text{s}$ for Si compared to $1 \cdot 10^{-11} - 10 \cdot 10^{-11} \text{ cm}^3/\text{s}$ for III-V). Nevertheless, the high quality material and surface passivation achievable in silicon results in significantly lower SRH recombination, an effect that is even more noticeable when shrinking the device to sub- μm dimensions [219].

The mode coupling efficiency η_c also has an important impact in the device efficiency.

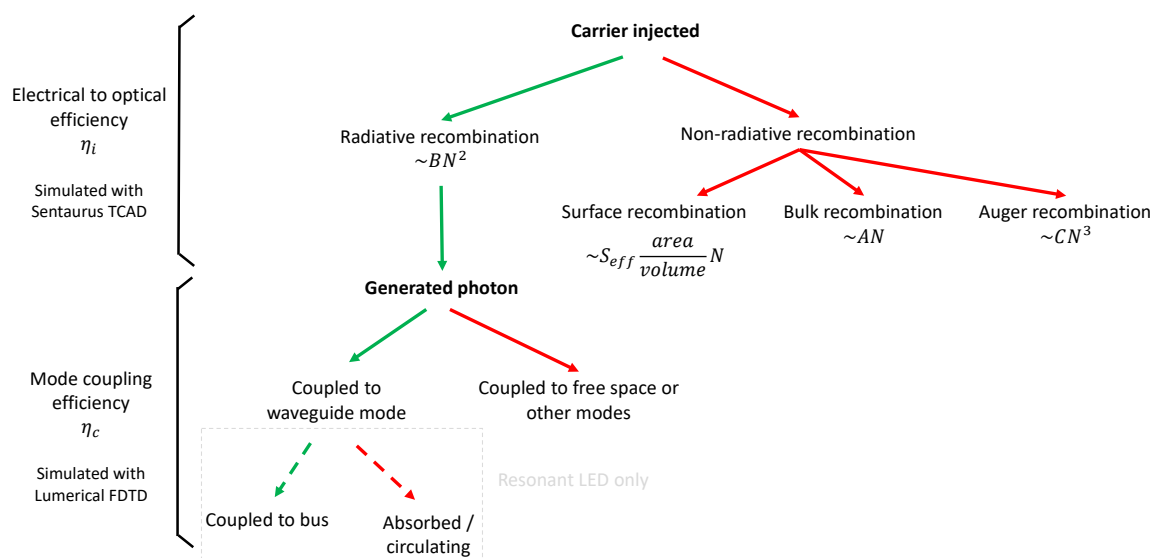


Figure 3-2: Main physical mechanisms affecting the generated optical power and efficiency of our silicon light sources. The competition between radiative and non-radiative recombination mechanisms sets the electrical to optical efficiency η_i . The total efficiency is also affected by how much of the generated light couples into the relevant optical mode of the waveguide η_c .

Ideally, we want every generated photon to be emitted into the optical mode supported by the waveguide. If we consider every radiative electron-hole pair as a dipole source, η_c is related to the overlap integral between the optical mode supported by the waveguide and the electric field generated by the emitting dipole. In the case of the resonant device, we also need to maximize the fraction of power coupled from the ring (where light is generated) to the bus waveguide, which can be adjusted by changing the gap between the ring and the bus.

3.3 Device modeling

To better understand the tradeoffs and main factors limiting the efficiency of our devices, we developed two different simulation models to capture the different physical mechanisms affecting the LED behavior. As indicated in Fig. 3-2, we use Sentaurus TCAD to obtain the electrical to optical efficiency η_i by modeling the carrier transport mechanisms in our silicon p-n junction (Section 3.3.1), and we use Lumerical FDTD to model the optical behavior of our system and obtain the mode coupling efficiency η_c (Section 3.3.2).

3.3.1 Electrical to optical efficiency: Sentaurus TCAD model

Firstly, we used Sentaurus TCAD [220] to model the electrical transport and recombination mechanisms in our devices and obtain the electrical to optical conversion efficiency η_i .

The simulation set up is depicted in Fig. 3-3. For both the resonant and linear LEDs we used the default silicon and silicon dioxide materials provided by Sentaurus, except that we modified silicon's radiative recombination parameter B from its standard value of 0 to $B = 4.7 \cdot 10^{-15} \text{ cm}^{-3}/\text{s}$ as measured by Nguyen [216]. The surface recombination

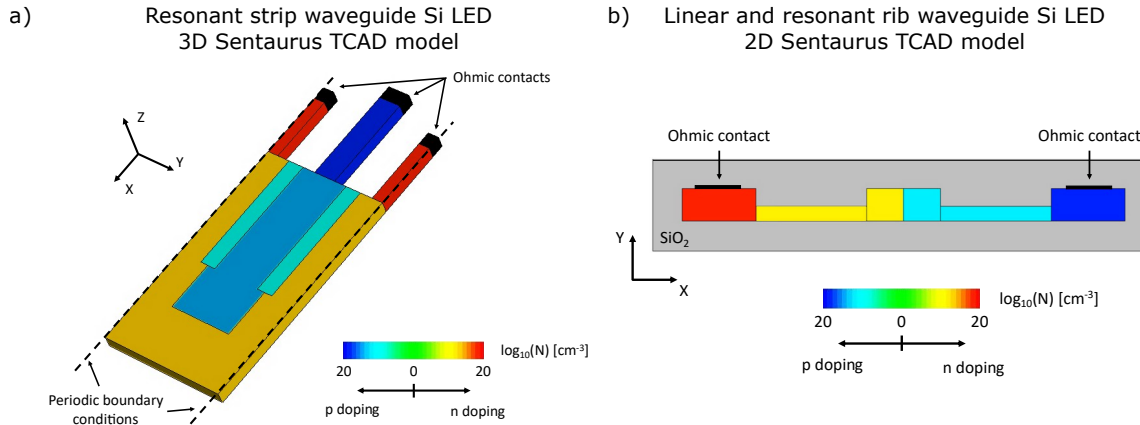


Figure 3-3: Sentaurus TCAD models of (a) the resonant strip waveguide Si LED, and (b) the rib waveguide Si LEDs (both linear and resonant).

velocity at the silicon - oxide interface is set to $S = 1000 \text{ cm/s}$ [221]².

As shown in Fig. 3-3(a), for the strip waveguide resonant silicon LED (based on interleaved pn junctions) we developed a 3D model of a single p-n spoked junction and applied periodic boundary conditions in the y direction to correctly capture the ring configuration, an approach that we have used and validated previously [222]. The whole simulation region depicted in Fig. 3-3(a) is surrounded by silicon dioxide (except in the y direction due to the periodic boundary conditions) so that the effect of surface recombination is correctly captured. We assume that the doping distribution is uniform in the z direction. For the rib waveguide LEDs (both linear and resonant) we developed a 2D model of the waveguide cross section as depicted in Fig. 3-3(b)³. Similarly to the case of the resonant LED model, silicon dioxide surrounds the whole silicon region so that surface recombination effects are properly captured.

²While there exists a range of surface recombination velocities for the Si - oxide interface reported in the literature, as we will discuss in detail below its exact value does not affect significantly the achievable efficiency of our LEDs at relevant operating currents.

³Figure 3-3(b) shows the model for the linear rib waveguide LED. The model for the resonant rib waveguide LED is identical except that the dimensions are adapted.

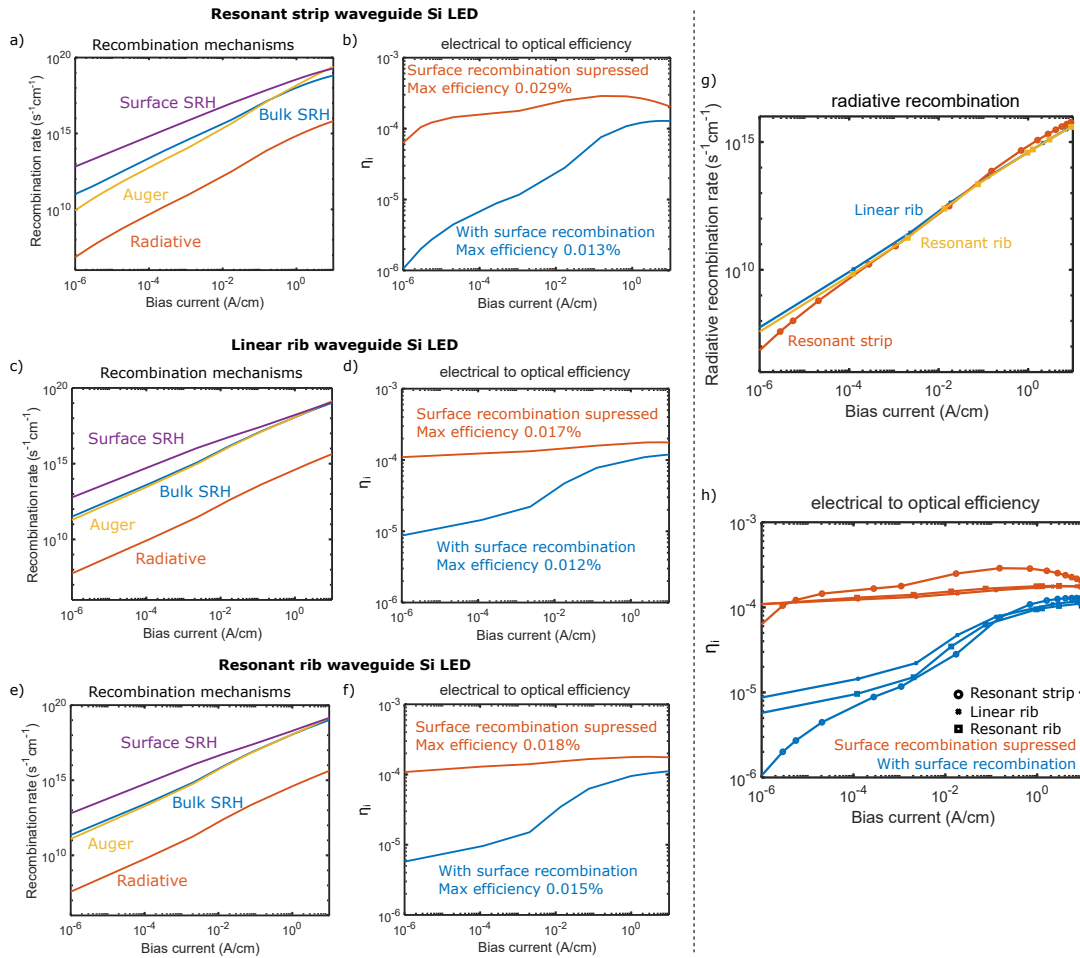


Figure 3-4: Sentaurus TCAD simulation of resonant strip waveguide (a, b) linear rib waveguide (c, d) and resonant rib waveguide (e,f) silicon LEDs. (a, c, e) Recombination rate as a function of bias current for the different recombination mechanisms present in the devices obtained with Sentaurus TCAD. (b, d, f) Electrical to optical efficiency η_i as a function of bias current with suppressed surface recombination (orange, SRV = 0) and with surface recombination (blue, SRV = 1000 cm/s) obtained from Sentaurus TCAD simulations. (g) Radiative recombination rate for the three designs. (f) Electrical to optical efficiency η_i for the three designs, with (blue) and without (orange) surface recombination.

The simulation results for the three designs are shown in Fig. 3-4. Figs. 3-4(a,c,e) show the rate of the different recombination mechanisms (integrated over the whole silicon volume) as a function of bias current for the resonant strip waveguide (Fig. 3-4(a)), linear rib waveguide (Fig. 3-4(c)) and resonant rib waveguide (Fig. 3-4(e)) LEDs. We can see how for all devices surface recombination is the fastest mechanism at lower bias currents and then Auger recombination takes over. Radiative recombination is the slowest (3 to 4 orders of magnitude lower) of the processes due to the indirect bandgap of silicon.

Figures 3-4(b,d,f) show η_i for all designs. As expected, the slow radiative recombination results in low efficiencies on the order of 0.01%. It is interesting to notice that suppressing surface recombination (orange curves in Figs. 3-4(b,d,f)) has a relatively low impact on the maximum achievable efficiency. This is because at the maximum efficiency point Auger recombination is comparable or higher than surface recombination. This points to two important facts: (1) maximum efficiency is achieved when the device is in the high carrier injection regime and (2) to maximize device efficiency we would need to ensure not only low surface recombination but also limited Auger recombination.

While suppressing surface recombination can be achieved by improving surface passivation and using techniques to keep carriers away from the surfaces, reducing Auger recombination in the bulk of the silicon while keeping a high carrier injection so that the radiative recombination rate is high is challenging. Recently, the use of strain engineering has been proposed as a means to modify the configuration of the different conduction bands in silicon and therefore reduce Auger recombination [223], which would allow for increased efficiency.

Of course, another avenue to achieve a higher efficiency is increasing the radiative recombination coefficient B (typically through defect implantation) but this requires device postprocessing and results in other performance tradeoffs as discussed earlier in

this text.

Figure 3-4(g) shows the radiative recombination rate (which is proportional to the generated optical power) as a function of bias current for our three different LED designs, and Fig. 3-4(h) the electrical to optical efficiency. We can see how the radiative recombination rate is almost the same for the three designs. At high current injection (bias currents above 0.1 A/cm) the strip waveguide resonant LED has a slightly larger radiative recombination rate. This can be explained because of the smaller cross-sectional area of that design, that results in a higher carrier concentration N (as in high injection, where the carrier concentration is uniform in the silicon, $N \propto I/LA$ where I/L is the bias current density and A is the cross-sectional area) and therefore a higher radiative recombination coefficient BN^2 . This explains why, as shown in Fig. 3-4(h), the strip waveguide resonant LED has a larger efficiency for current densities where Auger recombination is not too high (from about 10^{-2} A/cm to about 1 A/cm), an effect that is most noticeable when surface recombination is suppressed.

3.3.2 Mode coupling efficiency: Lumerical FDTD model

Secondly, we used Lumerical FDTD [224] to obtain the mode coupling efficiency η_c and the ring-bus coupling in the case of the resonant LEDs.

We model each radiative photon emission process as a dipole source and use a mode expansion monitor to calculate the overlap between the electric field generated by the dipole and that of the supported optical waveguide mode (Fig. 3-5).

It is important to notice that the coupling efficiency depends strongly on the location of the dipole (i.e, where the electron-hole recombination event occurs) and its polarization: maximum coupling will occur for locations with the best overlap with the optical mode and with the same polarization as the optical mode. It is for this reason that

we designed the p-n junction profiles for maximum overlap with the propagating optical mode. Nevertheless, in the regime where substantial radiative recombination is achieved the silicon is intrinsic due to the high number of injected carriers, resulting in the radiative recombination rate being almost uniform over the whole waveguide. This means that the exact location of the p-n junction doesn't affect the spatial profile of the radiative recombination significantly, and therefore its effect in η_c is minimal.

Due to the spatial dependence of the coupling efficiency, we repeat the simulation with the dipole at different locations within the cross section of the waveguide (shown in Fig. 3-5) and perform the average over all the simulated spatial locations. Notice how the fact that radiative recombination is uniform over the whole waveguide allows us to take the simple average of all the simulated spatial locations ⁴. As shown in Figs. 3-5(b,c), for the rib waveguide LEDs we sample 25 points per each section (full etch waveguide, partial etch and contact regions). Since each section has a different cross sectional area, we do the averaging of all the simulated positions weighted by the area of each region:

$$\eta_c = \frac{\overline{\eta_{c,full\ etch}}A_{full\ etch} + \overline{\eta_{c,partial\ etch}}A_{partial\ etch} + \overline{\eta_{c,contact}}A_{contact}}{A_{full\ etch} + A_{partial\ etch} + A_{contact}} \quad (3.3)$$

where $\overline{X_{region}}$ denotes spatial and polarization average over *region* and A_{region} is its area.

Since photon generation occurs through spontaneous emission in an isotropic material the polarization of the generated photons is random and uniformly distributed. To

⁴If the radiative recombination were non-uniform over the silicon waveguide the averaging weight of each simulated position should be related to the radiative recombination rate at that location (in other words, if more photons are generated in a specific location, then it should have a larger weight in the spatial averaging).

account for this, at every spatial location we repeat the simulation for the three possible dipole polarizations.

We note that, in the case of the resonant LEDs, Purcell enhancement should result in increased emission into the resonant optical mode by a factor equal to the Purcell number at the resonant wavelengths of the ring [225]. In our devices this does not have a strong effect in the total power coupled into the bus waveguide for two reasons: (1) the large free carrier losses at the high carrier injection required for strong light emission results in low resonator quality factors and therefore a low Purcell factor ($F_p < 4$); (2) since we care about the power integrated over the whole wavelength range, Purcell enhancement has a limited effect because it only occurs at the resonance wavelengths. We can estimate the increase in the total power due to the Purcell effect as $F_p \cdot FWHM/FSR$, where $FWHM$ is the full-width at half-maximum of the resonance and FSR is the free spectral range. Using the values for our devices $F_p = 2$, $FWHM = 750 \text{ pm}$ (corresponding to a quality factor $Q = 1,500$) and $FSR = 16 \text{ nm}$, we get a limited enhancement of 10%.

Figure 3-6 shows η_c for all designs obtained with Lumerical FDTD. The dashed lines show the coupling efficiency averaged over the whole silicon waveguide area for the three possible dipole polarizations, and the solid line shows the average over the three polarizations. We can see how the coupling to the fundamental mode of the optical waveguides is limited to about 2.5% for the resonant strip waveguide LED, and to about 1% for the both the linear and resonant rib waveguide LEDs. This is because, as expected, only one dipole polarization couples strongly to the propagating optical mode (the one corresponding to the TE polarization of the propagating optical mode). A larger coupling efficiency could potentially be achieved if we were to use multimode waveguides and use the light coupled to the multiple modes of such a waveguide. For rib waveguides, the coupling efficiency could also be increased by minimizing the area of the contact regions, since any light emission in those regions is not coupled into the propagating waveguide

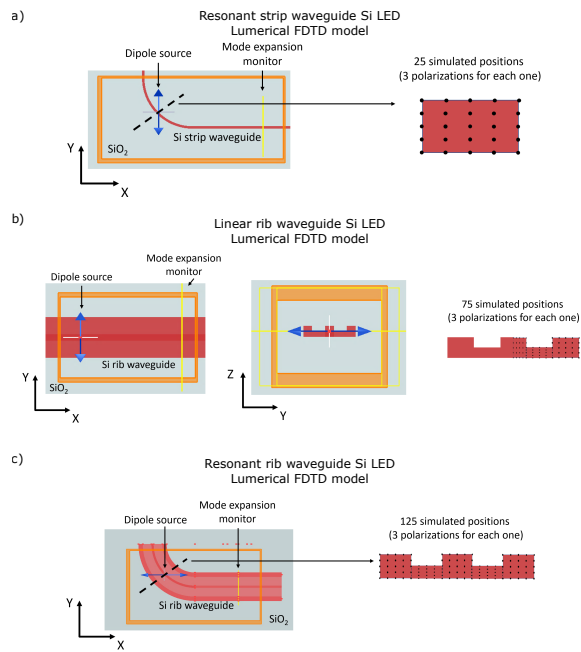


Figure 3-5: Lumerical FDTD simulation models for (a) the strip waveguide resonant Si LED, (b) the linear rib waveguide Si LED and (c) the resonant rib waveguide Si LED.

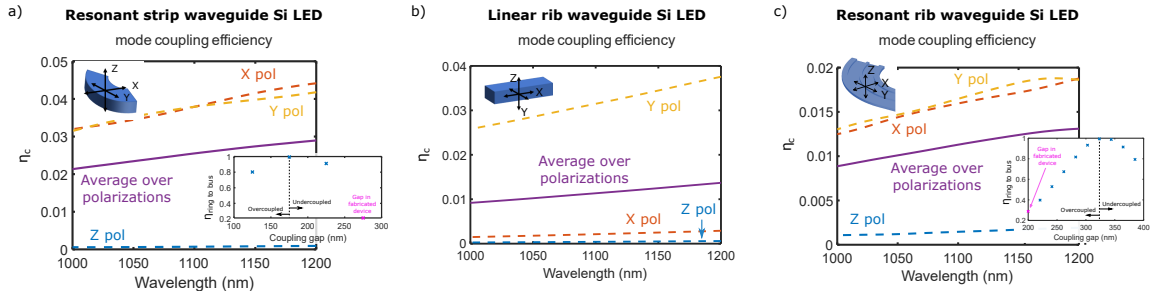


Figure 3-6: Optical mode coupling efficiency η_c obtained with Lumerical FDTD as a function of wavelength for the resonant strip waveguide LED (a), the linear rib waveguide LED (b) and the resonant rib waveguide LED (c). In (a,c), the ring to bus coupling coefficient for different coupling gaps is also shown as an inset.

mode. Decreasing the width of the contact regions from 500 nm (which is what we used for our devices, both linear and resonant) to 250 nm would increase the coupling efficiency to 1.2% (a 20% increase).

As already mentioned, in the case of the resonant LEDs we also need to account for the coupling between the ring and bus waveguides. We want the bus waveguide to be overcoupled to the ring, such that most (if not all) the light generated in the ring is coupled into the bus waveguide within a single round trip. As shown in the insets of Fig. 3-6(a, c), this can be achieved by adjusting the gap between the ring and bus waveguides. We also show in purple the gaps that we used in our fabricated devices.

3.4 Experimental demonstration

3.4.1 Characterization results

The experimental characterization results for the three devices are shown in Fig. 3-7. As described earlier, when we apply a forward bias voltage to these devices light is generated through phonon-assisted radiative recombination. This is easily seen by imaging the

devices with an InGaAs camera (top of Fig. 3-1(a, c)). While part of the generated light is emitted into free space ("back side emission" in Fig. 3-1(a), emission at the center in Fig. 3-1(c)), the rest is coupled into the propagating modes of the optical waveguide (with an efficiency η_c as discussed above). In the case of the linear device, this light propagates along the waveguide and reaches edge couplers at the chip end, where we can collect the light with a single mode fiber.

The power coupled into the single mode fiber as a function of bias current for the linear strip waveguide LED is shown in the top of Fig. 3-1(b), and the spectrum of such coupled light measured with an InGaAs spectrometer (custom-built by Dr. Zheng Li) is shown in the bottom ⁵. The center wavelength of the emitted light is 1.135 μm (close to the silicon bandgap) and has a full-width at half-maximum (FWHM) of 80 nm. The power coupled into the SMF at 20 mA bias current (corresponding to a bias voltage of 1.3 V) is 65 pW, which accounting for the edge coupler loss of 6 dB translates into a total of 260 pW of optical power propagating in the waveguide in each direction. We note that the generated light has a high degree of spatial coherence since it is propagating in a single mode waveguide.

In the case of the resonant devices, only light at the resonance wavelengths of the ring is coupled to the bus waveguide and emitted into free space through vertical grating couplers ("front side emission" in Fig. 3-7(a), emission in the green boxes in Fig. 3-7(c)). As such, we expect the light coupled into the waveguide to have a spectrum with multiple resonances spaced by the free spectral range (FSR) of the ring.

This is the case of the resonant strip waveguide LED (Fig. 3-7(a)), which shows resonances spaced by an FSR of 17 nm and with a FWHM of about 2 nm, limited by the high free carrier loss due to the high current injection. At a bias current of 10 mA

⁵The operating range of the InGaAs spectrometer is from 850 nm to about 1.5 μm with approximately 1 nm resolution at 1.3 μm . The spectrometer was calibrated using a Kr-Ar wavelength calibration source.

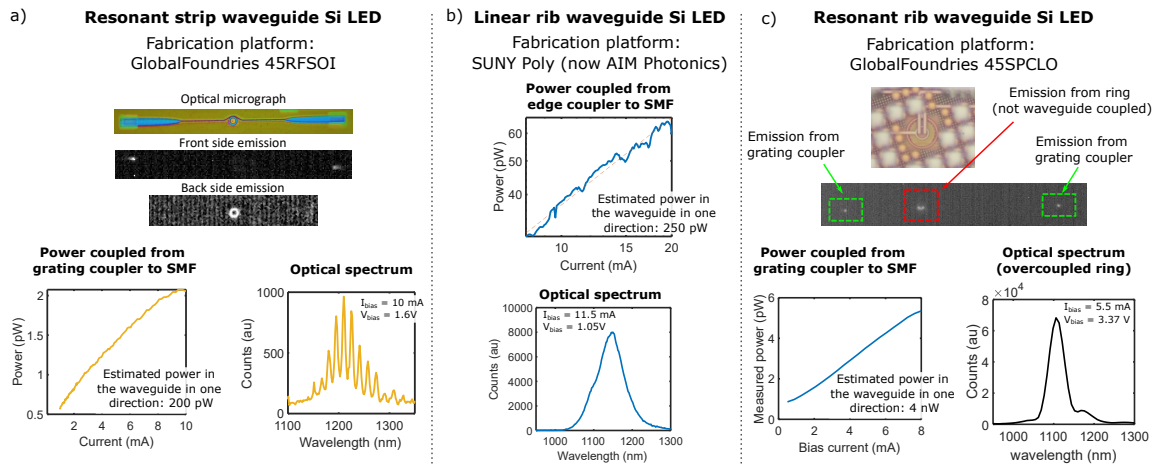


Figure 3-7: Experimental characterization of waveguide coupled LEDs. (a) Optical power coupled into Single Mode Fiber (SMF) as a function of bias current and optical spectrum of the resonant strip waveguide silicon LED. Images of the emission from the grating couplers and from the ring itself acquired with an InGaAs camera are also shown, along with an optical micrograph of the fabricated device. (b) Optical power coupled into Single Mode Fiber (SMF) as a function of bias current and optical spectrum of the linear rib waveguide silicon LED. (c) Optical power coupled into Single Mode Fiber (SMF) as a function of bias current and optical spectrum of the resonant rib waveguide silicon LED. Images of the emission from the grating couplers and from the ring itself acquired with an InGaAs camera are also shown, along with an optical micrograph of the fabricated device.

(corresponding to a bias voltage of 1.4 V), the power coupled from the vertical grating couplers into SMF is 2 pW. Accounting for the grating coupler loss and the limited optical bandwidth of the vertical grating couplers, we can estimate that there is a total of 200 pW of optical power propagating in the bus waveguide in each direction.

We do not observe resonances in the spectrum of the resonant rib waveguide LED (Fig. 3-7(c)). This is because the ring is heavily overcoupled to the bus waveguide (inset of Fig. 3-6(c)), which prevents the formation of sharp resonances. At a bias current of 8 mA (corresponding to a bias voltage of 4.2 V), the power coupled from the vertical grating couplers into SMF is 5 pW, which results in an estimated total of 4 nW of optical power propagating in the bus waveguide in each direction.

To estimate the power propagating in the bus waveguide for the two resonant LEDs we need to account for the fact that we are using vertical grating couplers to couple the light off chip. Vertical grating couplers have a limited optical bandwidth which is smaller than the range of wavelengths generated by the LED, and the coupling efficiency and spectral bandwidth depend on the angle between the optical fiber and the chip. It is therefore necessary to account for such "filtering" for a correct estimation of the total generated power. To do so, our approach was the following (Fig. 3-8):

1. We assume that the optical bandwidth of the light generated by the resonant LEDs is the same as that of the linear LED. This is a reasonable assumption since the physical mechanisms of light generation are the same in both devices.
2. Similarly, we assume that the optical bandwidth of the generated light is independent of the fabrication platform, i.e., that we can consider that the optical bandwidth is the same for an LED fabricated in the GlobalFoundries 45RFSOI or the GlobalFoundries 45CLO platforms and an LED fabricated in the SUNY Poly fabrication platform. Since the light generation mechanism is bimolecular recom-

bination in the bulk of the crystalline silicon, there is no reason why the optical spectrum should be different.

3. With this, we can take the envelope of the measured resonant LED spectrum and normalize it so that the power at the peak wavelength is the same as that of the linear LED. We can then compare the area under the curve for both spectra: if the resonant LED spectrum has an area that is $X\%$ of the area of the linear LED spectrum, then we can conclude that we are only capturing $X\%$ of the total power propagating in the waveguide.
4. Finally, we need to measure the efficiency of the grating coupler at the peak wavelength of the resonant LED spectrum, which we can easily measure through loop-back test structures in the chip. In our case, the peak efficiency of the grating couplers at the wavelength of interest is about $\eta_{GC} = -11$ dB for the 45RFSOI platform and $\eta_{GC} = -18$ dB for the 45CLO platform .

With this, we can estimate the power propagating in the waveguide in each direction as: $P_{wg} = P_{SMF}/(X\eta_{GC})$, resulting in the stated 200 pW (4 nW) of optical power propagating in the bus waveguide in each direction for the resonant strip (rib) waveguide LED.

In the case of the resonant strip waveguide LED, while our experimental device has a ring-bus gap of 275 nm (which results in $\eta_{ring\ to\ bus} = 0.2$), reducing this gap to 175 nm so that $\eta_{ring\ to\ bus} = 1$ (inset of Fig. 3-6(a)) would allow for a $\approx 5\times$ improvement in the power coupled into the bus waveguide, increasing the generated power to 1 nW. It is worth noting that the light generated by the resonant strip waveguide LED not only has a high degree of spatial coherence due to it propagating through a single mode waveguide, but it should also have some degree of temporal coherence thanks to the spectral filtering imparted by the ring resonator.

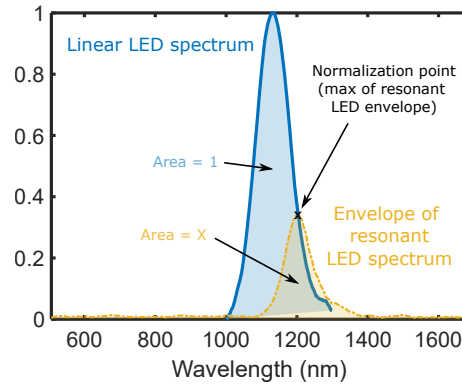


Figure 3-8: Illustration of the procedure to estimate the power in the waveguide for the resonant LED.

A comparison between the experimental measurements and the predictions derived from the Sentaurus TCAD + Lumerical FDTD simulation models is shown in Fig. 3-9. If we use the nominal value for the radiative recombination coefficient $B_{nom} = 4 \cdot 10^{-15} \text{ cm}^3/\text{s}$ measured by Nguyen [216], our simulation model predicts an approximately $10\times$ higher efficiency and generated optical power than what we measured experimentally for all the devices. The fact that the difference between simulation and experiment is the same for both designs likely points to differences between the recombination coefficients A, B and C used in the simulation and the processes in our fabricated devices. If we use the B coefficient as a fitting parameter to improve the agreement between simulation and experimental results we obtain $B_{fit} = 0.33 \cdot 10^{-15} \text{ cm}^3/\text{s} = 0.07B_{nom}$. Such a reduction in B has previously been attributed to bandgap narrowing at high carrier density: as the device goes into the high injection regime its intrinsic carrier concentration increases and its bandgap decreases, and therefore the radiative recombination coefficient decreases (since $B \propto 1/n_i^2$ [217, 226]). This has been previously discussed in silicon in the context of optimal doping of silicon solar cells [227].

	Simulation $B = 4.7 \cdot 10^{-15} \text{ cm}^{-3}/\text{s}$ (nominal value [24])		Simulation $B = 0.33 \cdot 10^{-15} \text{ cm}^{-3}/\text{s}$ (fit to experimental data)		Experiment	
	$\eta = \frac{\text{opt power}}{\text{el power}}$	Optical power *	$\eta = \frac{\text{opt power}}{\text{el power}}$	Optical power *	$\eta = \frac{\text{opt power}}{\text{el power}}$	Optical power †
Resonant strip waveguide Si LED @ 50 μA bias / spoke (1.5 mA total bias)	2e-6	100 pW	1.4e-7	7.66 pW	1e-7	5.5 pW
Linear rib waveguide Si LED @ 10 mA bias	2.5e-7	3.5 nW	1.8e-8	320 pW	2.6e-8	300 pW
Resonant rib waveguide Si LED @ 8 mA bias	1.2e-6	38 nW	0.84e-7	2.66 nW	1.2e-7	4 nW

* Accounting for mode coupling efficiency

† Accounting for coupler loss and spectral selectivity

Figure 3-9: Comparison of the simulated and measured optical power and efficiency for both the resonant and linear Si LEDs.

3.4.2 On-chip detection of the generated light

While the waveguide-coupled optical power generated by our native silicon LEDs is low (hundreds of pW), it can be detected by monolithically integrated photodetectors (PDs) fabricated in the same chip.

As the quality and maturity of silicon photonics fabrication processes and devices improves, high sensitivity photodetectors with very low noise have become widely available [222, 228–231]. The advent of such sensitive detectors enables the detection of on-chip optical powers that were previously not measurable, which opens up the avenue to the use of our low power light sources that would otherwise not be considered suitable for photonic systems.

To illustrate this, we demonstrate on-chip detection of the light generated by our waveguide coupled LEDs using two different monolithically integrated photodetectors:

1. A silicon-only resonant photodetector, which achieves photodetection through surface state absorption (SSA): the termination of the crystal structure at the silicon-oxide interface distorts the band structure and creates intra-gap states, which result in sub-bandgap light absorption [232–234]. While the responsivity of these detectors is limited ($R=26$ mA/W in our case), they can be realized in any silicon photonics process without the need of additional materials like SiGe or Ge.
2. A Ge linear photodetector, which has a much higher responsivity (we experimentally measured $R=0.63$ A/W at 0 V bias), but is not available in all CMOS photonic platforms (as it requires the addition of Ge to the fabrication process).

Resonant Si LED - Resonant Si PD (GlobalFoundries 45RFSOI)

We optically connected two identical strip waveguide silicon ring resonators through a common bus waveguide (Fig. 3-10(a)). The two rings have the same design, described in Fig. 3-1(a) and were fabricated in the GlobalFoundries 45RFSOI process. We operated one of the rings as a resonant LED (i.e, in forward bias), and the other as a photodetector (i.e, in reverse bias).

As mentioned above, since this is a silicon-only photodetector light detection occurs mainly through surface state absorption (SSA), where the termination of the crystal structure at the silicon-oxide interface distorts the band structure and creates intra-gap states, which result in sub-bandgap light absorption [232–234]. In our devices, this process results in a responsivity of about 26 mA/W. While this is a relatively limited responsivity, these devices feature a high sensitivity due to their low dark currents below 100 fA. This is because the high quality of the CMOS fabrication process results in very low defect density in the silicon.

We use a resonant PD instead of a linear PD for two main reasons: (1) resonant

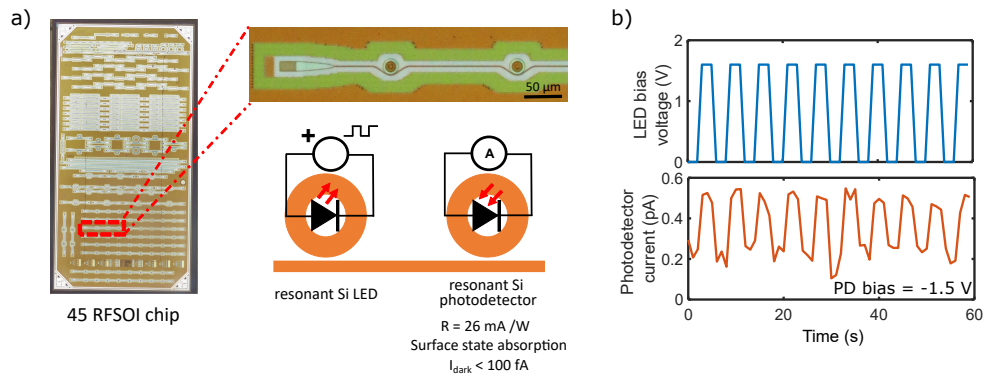


Figure 3-10: On-chip detection of the light generated by the resonant strip waveguide Si LED through a Si resonant PD. (a) Link configuration. Two identical silicon ring resonators are optically connected through a bus waveguide. The first resonator is operated as an LED (i.e, forward biased), while the second resonator is operated as a photodetector (i.e, reverse biased). Photodetection at these wavelengths is achieved through surface state absorption. The high fabrication quality results in dark currents below 100 fA. (b) On the top, the driving voltage applied to the LED is shown as a function of time. On the bottom we show the photocurrent detected by the detector as a function of time. Clearly, the photodiode can detect the light generated by the LED.

enhancement results in a larger responsivity at the resonant wavelength compared to a linear detector with the same length and (2) the inherent wavelength selectivity of resonant detectors can be exploited for certain applications such as wavelength division multiplexing (WDM) based communications [235] or sensing [236] as we will discuss in the next Chapter.

To demonstrate light detection, we modulated the voltage applied to the LED and monitored the photocurrent generated by the PD (Fig. 3-10(b)). Clearly, the detected photocurrent follows the voltage applied to the LED, demonstrating successful detection of the light generated by our native waveguide-coupled silicon LED. The speed of the applied signal is limited by the long integration time required to lower the noise floor of the external source meter we use to monitor the detector photocurrent.

As shown in Fig. 3-10(b), the generated photocurrent when the LED is turned on is

about 500 fA, suggesting a total power of 20 pW coupled into the detector ring. This is an order of magnitude lower than the power generated by the LED (as we characterized in the previous section, a 10 mA, 1.5 V bias results in 200 pW of generated power in each direction), which can be explained by (1) an unoptimized coupling gap between the bus waveguide and the detector ring, which results in undercoupling of the light and (2) the fact that the resonances of the LED are not aligned to the resonances of the photodetector combined with the fact that the photodetector resonances are narrower than those of the Si LED due to lower free carrier loss. Switching from a Si-only resonant photodetector to a higher responsivity resonant photodetector based on SiGe (which we have previously demonstrated in the same fabrication platform [222]) would result in a photocurrent increase of about 15 \times . Optimizing the coupling gaps for both the LED and the detector would result in a further increase in signal of about 25 \times (5 \times higher power coupled from the LED to the bus waveguide, and 5 \times higher coupling from the bus waveguide into the PD resonator). This would result in a total signal increase of 375 \times , which translates into a photocurrent of 180 pA. A further photocurrent increase could be obtained by adding thermal tuning to the resonant detector so that the LED and detector resonances are perfectly aligned.

One possible concern with these results is that the photocurrent increase observed in the detector is due to device heating and not due to the detection of the light generated by the LED: It is well known that the dark current in a pn junction photodetector increases with increasing temperature, and there is a significant amount of electrical power (\approx 15 mW) being dissipated in the LED when it is turned on, which translates into an increase in temperature in the vicinity of the LED.

Since the LED ring resonator and the photodetector ring resonator are close to each other (\approx 90 μ m apart), there is a chance that the increase in the PD photocurrent observed when the LED is turned on is not due to light absorption but due to an increase

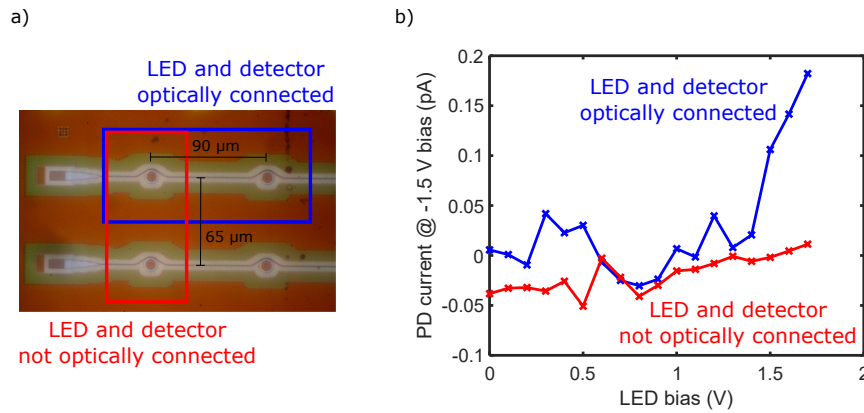


Figure 3-11: Ruling out heating as the source of photocurrent. (a) Micrograph of the chip showing the two distinct measurements. In the blue box, we see the LED and the PD optically connected through a bus waveguide. In the red box, the LED and the PD are not optically connected but a similar distance away so that the heating due to the LED bias is similar. (b) Photodetector current as a function of LED bias voltage for the detector optically connected to the LED (blue) and the detector not optically connected to the LED (red). The bias voltage for both photodetectors is -1.5 V.

in its temperature.

To rule out this possibility, we performed the same light detection experiment described earlier (where we apply a forward bias voltage to the LED and monitor the photocurrent in the PD), but this time monitoring the photocurrent in two different resonant PDs: one optically connected to the LED through a bus waveguide and another one which is not optically connected to the LED (Fig. 3-11(a)). Notice that the distance between the LED and the two PDs (the one optically connected and the one not optically connected) is similar so the temperature increase due to the power dissipated at the LED should be similar. Figure 3-11(b) shows the measured photocurrent for both photodetectors (at a bias voltage of -1.5 V) as a function of the LED bias voltage. We can clearly see how light is detected in the optically connected detector (blue curve in Fig. 3-11(b)) for LED bias voltages above 1.5 V, while no significant increase in photocurrent is observed in the optically disconnected photodetector (red curve in Fig. 3-11(b)). This

shows that the increase in photocurrent in the connected PD is not due to heating but indeed comes from absorption of the light generated by the LED.

Resonant Si LED - Ge linear PD (GlobalFoundries 45CLO)

While the use of a Si resonant PD is attractive because it can be realized in any modern microelectronics process, it comes at the cost of a limited responsivity $R = 26 \text{ mA/W}$. We can alternatively use a PD made of Ge, which has a much higher absorption at our operation wavelengths due to its lower bandgap, and therefore higher responsivity. The main disadvantage of using Ge PDs is that they are only offered in photonics-oriented CMOS fabrication platforms. For example, GlobalFoundries 45RFSOI (a purely microelectronics process) does not have Ge available.

We instead used GlobalFoundries 45CLO process, which has linear Ge PDs with an experimentally measured responsivity $R = 0.63 \text{ A/W}$ at our wavelengths of interest, a $24\times$ larger responsivity than the resonant Si LED used above.

As shown in Fig. 3-12(a), we optically connected a resonant rib waveguide Si LED (the device presented in Fig. 3-1(c)) to a linear Ge PD through a $100 \mu\text{m}$ long SiN waveguide section. Such SiN waveguide section serves two purposes:

1. Electrically isolate the LED from the PD. The use of a Si waveguide would result in an electrical connection between the LED and PD contacts through the Si waveguide itself, as the undoped Si forming the waveguide has a finite resistivity. As a result, a non-negligible electrical crosstalk between the PD and LED would occur which could bury the real signal coming from the absorbed light.

We experimentally measured the LED-PD crosstalk in a test structure where the two devices are connected by a $30 \mu\text{m}$ long Si waveguide. If we apply a voltage between the LED and PD anodes (i.e, p doping) or cathodes (i.e, n doping) we

measure a resistance of $2\text{ M}\Omega$. Thus, a 2 V differential results in a current of about $1\ \mu\text{A}$, which is larger than the signal due to the LED generated light. Similarly, if we apply a voltage between the LED anode and the PD cathode or viceversa a parasitic diode is formed with a series resistance of $2\text{ M}\Omega$.

2. Achieve low loss transport of the light generated by the LED by using a low loss transition from a Si waveguide to a SiN waveguide. Since the propagation loss of the light generated by the LED in a Si waveguide is large (about 14 dB/cm), transitioning to a SiN waveguide (which has no absorption at the generated wavelengths, and therefore significantly lower propagation loss) is attractive for applications that require long routing within the chip. We experimentally measured the transition loss from the Si to the SiN waveguide to be 0.3 dB , which makes the transition to a SiN waveguide beneficial for routing lengths longer than $400\ \mu\text{m}$.

Figure 3-12(b) shows the experimentally measured IV curve of the LED, along with the photocurrent generated by the Ge PD as a function of the voltage applied to the LED. We can see how at a 8 mA , 3.4 V bias the generated photocurrent is about 500 pA , a $1,000\times$ increase over the demonstration with the Si resonant PD. It is important to note that we bias the Ge PD at 0 V to eliminate dark current, which would bury the light signal (at -1 V bias we measure a dark current of 50 nA). 500 pA of photocurrent translate into 0.8 nW of optical power reaching the PD, which accounting for the 0.6 dB loss of the two Si - SiN transitions translates into 0.91 nW of optical power generated by the rib waveguide LED in each direction. This is in good agreement with the experimental characterization of the device (Table 3-9).

Such an increase in photocurrent with respect to the resonant Si PD demonstration can be broken down as follows:

- A $24\times$ larger PD responsivity.

- A $5\times$ increase in ring to bus coupling efficiency as the rib resonant LED is over-coupled.
- A $5\times$ larger power reaching the PD since here there is no bus to ring coupling as we are using a linear PD.
- A $2\times$ lower waveguide coupling efficiency (Fig. 3-6(c)).
- The fact that, since we are using a linear PD, there are no losses due to misalignment between the resonances of the generated LED light and the resonances of the PD, nor losses from the fact that the PD resonances are narrower than those of the LED (an effect that was present in the demonstration of the previous section). To match the measured $1,000\times$ increase in photocurrent, this effect results in a $3.3\times$ increase in the power absorbed by the PD.

We could achieve a 30% larger photocurrent by increasing the PD absorption length, which would improve the responsivity from our experimental $R = 0.63$ A/W (quantum efficiency $\eta = 0.68$) to $R = 0.8$ A/W ($\eta = 0.87$)⁶. While the increase in responsivity would be accompanied by an increase in dark current, this is not of great concern because we operate our PD at 0 V bias.

3.5 Conclusion

In this chapter we have, to the best of our knowledge, demonstrated for the first time the availability of native silicon waveguide-coupled LEDs in wafer-scale, CMOS-compatible silicon photonics processes.

⁶We used an off-the-shelf PD design optimized for high speed operation in the O band ($\lambda=1310$ nm).

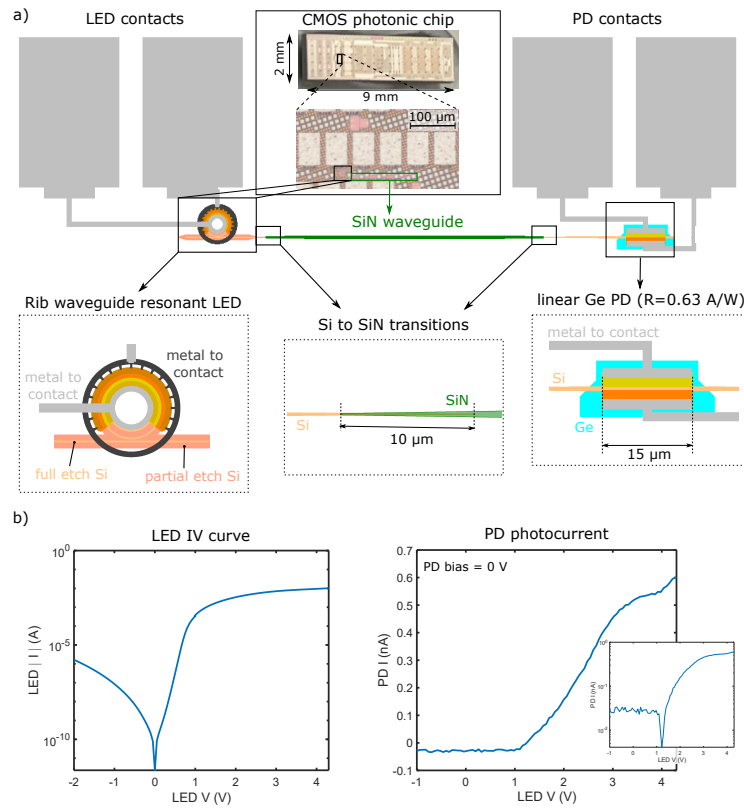


Figure 3-12: On-chip detection of the light generated by the resonant rib waveguide Si LED through a Ge linear PD. (a) Link configuration, where the rib waveguide LED is connected to a Ge PD through a SiN waveguide for electrical isolation. In the top center, a micrograph of the CMOS photonic and a closeup on the link are shown. (b) IV curve of the rib waveguide resonant LED (left) and PD photocurrent as a function of LED bias voltage (right). The inset shows the same data but with the photocurrent in logarithmic scale.

Light emission is achieved through phonon-assisted radiative recombination, with generated powers up to 4 nW and a central wavelength around 1130 nm. These are enough to be detected by monolithically integrated photodetectors.

We have also modeled the electrical and optical behavior of these devices and shown that both surface recombination and Auger recombination are the dominant factors limiting the achievable efficiency. Thus, some improvement in the generated optical power can be obtained by minimizing the interaction of injected carriers with the silicon - oxide interface to minimize surface recombination. Strain engineering has recently been proposed as a means to reduce Auger recombination in silicon [223], which could also increase the achievable quantum efficiency of our devices.

Such low power, low cost, monolithic light sources in silicon photonics processes can find applications in spaces where the use of external light sources (which require complicated packaging and/or coupling) is not feasible or desirable. This includes on-chip optoisolators [237, 238], random number generation [239], the realization of low speed intra-chip communication links or generation of illumination for imaging. These sources could also be used in photonic test structures for quality control of, for example, foundry fabrication processes, enabling wafer-scale testing of optical properties by fully electronic means without the need for external light coupling.

One could also consider the use of such waveguide coupled sources for imaging applications, where out-coupling of the light from the chip to free space can be achieved through the use of a grating coupler array (Fig. 1-8(a)). Compared to the use of surface emitting LEDs (the approach we took in Chapter 2), the use of a grating array cannot achieve as small a spacing or emitter size, but it allows for processing of the generated light before launching it into free space. This could enable functionality that is not possible using surface emitting arrays, such as beam steering or phase control of the illumination field.

Chapter 4

Fully integrated, monolithic refractive index sensors in CMOS photonic fabrication processes

In Chapter 3 we have successfully demonstrated native generation of light in a silicon single-mode waveguide, and also shown that the generated light is enough to be detected by an on-chip photodetector. Such low power, low cost, monolithic light sources in silicon photonics processes can find applications in spaces where the use of external light sources is not feasible or desirable. This is particularly appealing for applications that can benefit from the high scalability and easy replication of our sources into large arrays, overcoming an important bottleneck present in typical silicon photonics systems that require external light generation, which need to either use multiple light sources and coupling points in the chip (which is expensive and raises robustness concerns) or split the light inside the chip (which takes a considerable area).

Massive replication of compact monolithic light sources, waveguides and photode-

tectors within a single chip enables the low cost realization of large arrays of individual "processing units". In this Chapter, we explore the use of these "processing units" as optical refractive index sensors, which would enable the realization of truly monolithic, massively multiplexed optical refractive index sensors without external light coupling. We first present a brief overview of refractive index sensing, including its applications, promises and current limitations in Section 4.1. Then, in Section 4.2 we present a novel integrated refractive index sensor that uses the broadband silicon waveguide coupled LEDs we presented in Chapter 3 as the light source, build a detailed model of the whole sensing system and derive realistic performance limitations. Finally, in Section 4.3 we experimentally demonstrate refractive index sensing using ring resonators in our silicon photonics platform, showcasing the availability of all the components required to realize our proposed sensing system. We conclude that our sensor can achieve comparable performance to state of the art instruments but at a fraction of the complexity, cost and size.

4.1 Refractive index sensing: a brief overview

Optical refractometric sensing has proven to be one of the most successful approaches for the realization of compact sensors, mostly targeted to the analysis of biological samples with applications in diagnostics, drug and chemical development, environmental monitoring and biological research amongst others [240–242].

Refractometric optical sensors can be functionalized for the targeted detection of a wide range of analytes (see for example [240–244] for a list of literature demonstrations) including DNA sequences, cytokines, microRNA, cancer biomarkers and bacteria. Multiple companies are commercializing this technology [245–250], promising better performance than the gold standards in their target fields (such as ELISA or gas chromatography

with mass spectrometry in the case of protein detection and quantification [251, 252]).

The basis of this sensing approach is simple: the sensor measures small changes in the refractive index of its environment, which is correlated to some metric of interest such as the concentration of a specific antibody or protein, or the presence of a hazardous contaminant. In optical refractometric sensing, the change in refractive index is measured through optical means, usually by inferring the change in effective index of an optical mode supported by the sensor.

In contrast with most alternative technologies, refractive index sensing is a label-free technique, which greatly simplifies sample preparation and the expertise required to operate these sensors. Historically, such sensors were made out of bulk optics components, resulting in large and expensive instrumentation and thus limiting its use to highly specialized environments such as clinical laboratories.

Integrated optics in general, and silicon photonics in particular, have been proposed as an alternative approach to solve the limitations of bulk optics-based refractometric optical sensors [241, 253] due to its potential for miniaturization. Multiple technical approaches have been pursued in integrated refractometric sensing, including Surface Plasmon Resonance (SPR) based sensors [254], sensors based on optical waveguide gratings [255] and grating couplers [256] and sensors based on interferometry using different geometries (Mach-Zehnder [257], ring resonator [258], bimodal waveguides [259]. . .).

In this thesis we will consider the realization of integrated refractive index sensors based on ring resonators [244] ¹. The basic principle of ring resonator based optical refractometric sensing is depicted in Fig. 4-1. As an analyte of interest comes close to the vicinity of the ring resonator, the dielectric environment of the ring changes, which

¹A discussion of every single sensing approach is out of the scope of this thesis. The interested reader is pointed to references [240–242] as a good starting point, as well as the individual references next to each technical approach.

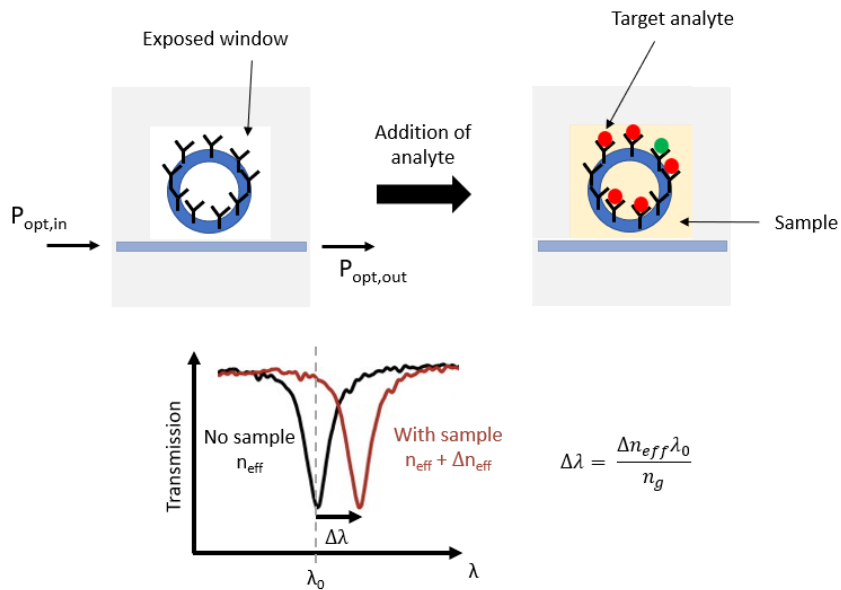


Figure 4-1: Refractive index sensor with optical ring resonators. The addition of analyte causes a change in the effective refractive index of the optical mode propagating around the ring, which changes the resonance wavelength of the ring.

causes a shift in the effective index of the optical mode propagating around the ring and as a result shifts the resonance wavelength of the ring λ_0 , according to [260]:

$$\Delta\lambda_0 = \frac{\Delta n_{eff} \lambda_0}{n_g} \quad (4.1)$$

Above, Δn_{eff} is the change in effective index of the optical mode due to the addition of the analyte, and n_g is the group index of the optical mode.

Thus, by measuring the change in resonance wavelength of the ring the concentration of the target analyte can be inferred. Usually, to ensure that the response of the sensor is due to the presence of a specific target analyte the ring resonator is functionalized - the silicon surface is treated and an additional material is deposited that preferentially responds to the presence of the target analyte that we want to detect. A variety of

materials can be used to functionalize the surface depending on the nature of the target analyte, including polymers and antibodies [261, 262]. Automated functionalization of individual resonators in the same chip with different functionalization agents has been demonstrated, opening up the avenue for the detection of multiple analytes with a single sensor [263].

While such an approach should in principle allow for the integration and miniaturization of the whole sensor, the reality is that most refractometric sensors (and all commercial refractometric optical sensors) are not fully integrated: the sensing chip (the part of the system that is exposed to the analyte) is separate from the light source and readout components, thus requiring complicated and expensive packaging. In the particular case of ring-based refractive index sensors, this is because the standard approach to infer the change in resonance wavelength is by measuring the transmission spectrum of the ring, which requires expensive, non-integrated tunable lasers or spectrometers.

The use of an external light source and/or detector poses an additional limitation to the miniaturization and cost reduction of these sensors, as it requires a means to efficiently couple the light into the sensing chip, which is usually done through the use of cameras and mechanical components [258].

Some work trying to integrate the detector and/or source into the same substrate as the sensing element has been reported in the literature (and is an approach being actively pursued by the company Siphox [247]). In [264], a silicon ring resonator based sensor with integrated germanium photodetectors was presented. While useful to eliminate the need for out-coupling the light and the use of an external detector, this sensor still required the use of an external tunable laser, resulting in a large, high cost system. A miniaturized refractive index sensor with integrated source and photodetector was reported in [265], based on a Mach-Zehnder interferometer as the sensing element followed by an integrated spectrometer based on an arrayed waveguide grating (AWG) fabricated in silicon nitride.

Ref (year)	Brief description	Level of integration	LoD ($\Delta\lambda_0$)	LoD (RIU)	Size	Cost
[258] (2010)	RR external PD external TLS	Low	0.22 pm	$7.6 \cdot 10^{-7}$	Desktop	\$120,000
[264] (2018)	RR integrated PD external TLS	Medium	11.3 pm	$5.2 \cdot 10^{-5}$	Desktop	NR
[265] (2019)	MZI integrated PD integrated LED	High	NA	$2 \cdot 10^{-6}$	37 mm ²	NR

Table 4.1: Performance of optical refractive index sensors with different levels of integration. RIU = refractive index unit. RR = Ring resonator. TLS = Tunable Laser Source. PD = photodetector. MZI = Mach-Zehnder Interferometer. NR = Not Reported. NA = Not Applicable.

The integrated light source was made on a silicon chip, and the individual silicon and silicon nitride substrates were heterogeneously integrated. While successful, this approach requires a custom fabrication process not directly compatible with CMOS platforms, which results in increased cost and concerns for the scalability, yield and repeatability of these sensors. Table 4.1 lists the performance metrics of these two works (which to the best of the author’s knowledge are the only works reporting optical refractometers with integrated sources and/or detectors) as well as that of the most sensitive, non-integrated ring resonator-based refractive index sensor reported in the literature [258]. Notice how the sensitivity of a ring resonator refractometer can be measured in two ways: (1) by the minimum detectable change in resonance wavelength $\Delta\lambda_0$; and (2) the minimum change in the refractive index of the analyte being sensed (measured in Refractive Index Units - RIU).

All in all, these limitations still result in sensors with a similar cost and/or size to

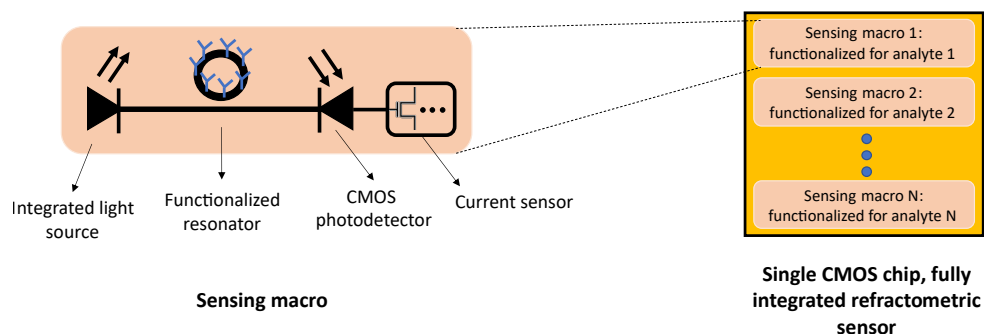


Figure 4-2: Monolithically integrated refractive index sensor. A single sensing unit consists of a light source, sensing ring, photodetector and current sensor, all integrated in the same CMOS chip. Multiple sensing macros, each functionalized for a different analyte, can be fabricated in the same chip.

sensors based on bulk optics, facing the same obstacles to the wide adoption and democratization of refractometric sensors. In the next section we propose the use of our native waveguide coupled LEDs to realize a fully integrated, monolithic refractometric sensor with the potential to overcome the limitations we just described.

4.2 Refractive index sensing with native waveguide coupled LEDs

Our native silicon light sources offer an avenue towards the realization of fully integrated, monolithic refractive index sensors including the light source, the sensing element (a ring resonator in our case) and the photodetector all in the same substrate as schematically depicted in Fig. 4-2.

Such fully integrated biosensor offers many advantages:

1. As discussed earlier, conventional refractive index sensors use an external light source and a photodetector or camera, which are not monolithically integrated

with the sensing chip. This requires a mechanism to couple the light from the light source to the sensing chip and from the sensing chip to the photodetector/camera, either through faulty and aging mechanical setups [258] or through costly and time expensive optical fiber packaging [22].

In contrast, our proposed sensing platform in which both the light source and photodetector are monolithically integrated circumvents this need, resulting in a simpler, cheaper system.

2. Most ring-based refractive index sensors use a tunable source to obtain the transmission spectrum and extract the change in the resonance wavelength. While this scheme is simple and more robust to noise, tunable laser sources are expensive, bulky, and require very careful control and monitoring of the generated wavelength to achieve accurate measurements.

On the other hand, our approach takes advantage of on-chip, low-cost and broadband light sources. This not only results in more compact and cheaper systems, but also allows for the realization of multi-channel sensors that can detect multiple analytes at the same time (Fig. 4-2). This is in contrast with the approach followed by sensors based on tunable sources, which require rerouting the generated light to the different sensing sites sequentially.

3. Since we can integrate all the components in a single CMOS chip we can exploit all of its associated advantages: high fabrication quality with very high resolution and repeatability, low cost, easy packaging and high performance on-chip electronics.

Two main challenges arise for the use of our integrated light sources: (1) the low generated power poses a limitation in the achievable SNR and therefore in the LoD; and

(2) the broadband illumination requires an alternative sensing scheme that does not rely on tunable sources.

Several approaches that use broadband illumination for the realization of refractive index biosensors have been proposed in the literature, including the use of a spectrometer to monitor multiple ring resonances [266] and the use of a second ring that allows one to "track" the resonance of the sensing ring by maximizing the photocurrent at the detector [267–269].

For our fully integrated biosensor we will use two different variations of the second approach (the use of a tracking ring), since it does not require a spectrometer and therefore results in a much simpler system.

Schematics of the two different sensing schemes we propose are shown in Fig. 4-3. Both are based on a feedback system that keeps the transmission spectrum of the sensing and tracking rings aligned so that the current at the photodetector is maximized. When an analyte is added the resonance of the sensing ring shifts, misaligning the spectra and therefore reducing the current at the photodetector. The feedback system then tries to recover the maximum photocurrent by shifting the resonance of the tracking ring by means of a heater. In this approach, the change in resonance wavelength of the sensing ring (and therefore the concentration of the analyte of interest) is proportional to the heater current required to maintain alignment between both rings.

The two schemes differ in what is used as the tracking ring. In Scheme 1 (Fig. 4-3(a)) we use a resonant LED both as the light source and the tracking ring. This results in a very compact and simple system, but limits the minimum achievable width of a single resonance to about 2 nm (as we obtained experimentally, due to the high free carrier loss), which as we will discuss is not optimal for maximum sensitivity. Scheme 2 (Fig. 4-3(b)) uses a linear LED and an additional passive tracking ring. While this scheme requires one more element, it has flexibility on the Q factor of the passive tracking ring,

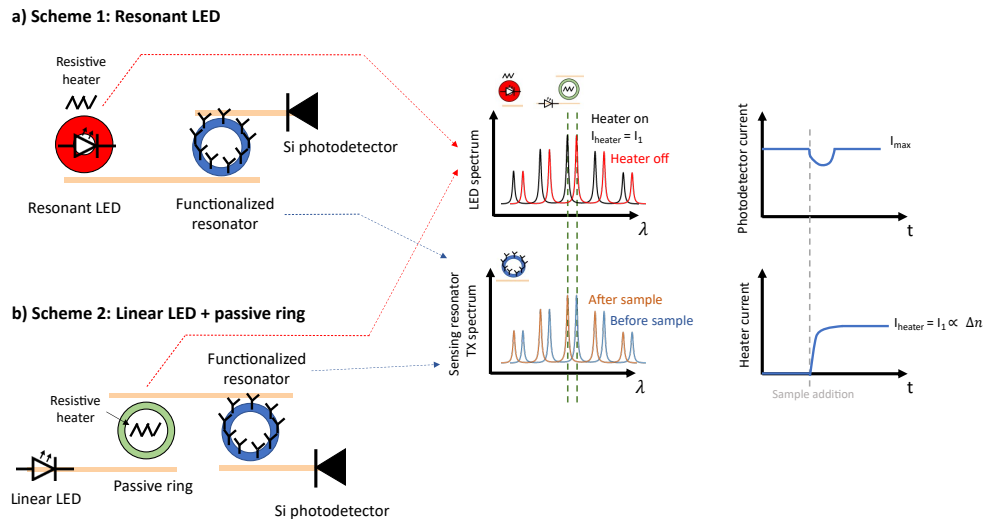


Figure 4-3: Sensing schemes for our fully integrated refractometric sensor, based on a feedback system that maintains alignment between a sensing and a tracking ring. (a) In this scheme we use a resonant LED both as the light source and tracking ring. (b) We use a linear LED as the light source and a passive tracking ring.

which can then be optimized for maximum sensitivity.

4.2.1 Achievable performance

Individual component modeling

To fully describe the two sensing schemes we are considering we need to model the behavior of 4 components, each with a set of parameters:

1. Integrated LED: The integrated LED generates the light that probes the alignment between the tracking and sensing rings. To model the linear LED we use two parameters: (1) the total emitted power integrated over the whole emission bandwidth; and (2) the FWHM of the emission spectrum. The resonant LED has two additional LED parameters which are the FWHM of a single resonance and the resonance FSR.

2. Passive add-drop ring: As described earlier, a passive add-drop ring resonator [260] is used as the sensing element whose properties change with the addition of analyte. We will model the ring resonator according to the theory presented in reference [260], and we will consider that we have flexibility to control the FSR (related to the ring radius) and FWHM (related to the ring-bus coupling coefficients) of the ring - within the limitations set by bend radius losses (which limit the maximum FSR) and linear absorption losses (which set the minimum achievable FWHM). Note that in sensing scheme number 2 (Fig. 4-3(b)) we use two distinct passive add-drop rings: one for sensing and one for tracking.
3. Integrated photodetector: The photodetector converts the optical power filtered by the sensing and tracking rings into a current signal. We will assume we are using a Ge-based integrated photodetector, which is a standard component in modern CMOS photonics processes. We will assume a responsivity of 1 A/W, which is the typical performance of these devices [20,270]. Another important performance metric of these detectors is their dark current, which sets a limitation in the achievable SNR. Nevertheless, since we do not require fast operation of our photodetector we will assume we are operating at 0 V bias, where there is no dark current and therefore the sensitivity of our system is maximized.
4. Current sensing system: The minimum detectable change in effective index (and thus analyte concentration) is ultimately set by the minimum change in photocurrent that we can detect through the integrated current sensing system. We will consider three different performance corners for our CMOS current sensor:
 - High performance, with a noise floor of $0.7 \text{ fA}/\sqrt{\text{Hz}}$. This corresponds to the noise floor of the gold standard, non-integrated current sensor currently






	 Linear LED	 Resonant LED	 Passive add-drop ring	 Photodetector	 Current sensor
Fixed parameters	Spectrum FWHM 100 nm Center wavelength 1130 nm	Spectrum FWHM 100 nm Single resonance FWHM 2 nm Center wavelength 1130 nm	Effective index of optical mode 2.27 Group index of optical mode 4.62	Responsivity 1 A/W Dark current 0 (0 V bias)	Operating bandwidth 100 Hz Noise floor High perf: $0.7 \text{ fA}/\sqrt{\text{Hz}}$ Med perf: $10 \text{ fA}/\sqrt{\text{Hz}}$ High perf: $100 \text{ fA}/\sqrt{\text{Hz}}$
Design parameters		FSR	Single resonance FWHM FSR		

Figure 4-4: Model parameters for each of the components in the fully integrated refractive sensor.

used in commercial systems, the Axon Axopatch 200B [271].

- Medium performance, with a noise floor of $10 \text{ fA}/\sqrt{\text{Hz}}$.
- Low performance, with a noise floor of $100 \text{ fA}/\sqrt{\text{Hz}}$.

Reference [272] reviews the performance of integrated CMOS current sensors reported in the literature, showing achievable noise floors below $1 \text{ fA}/\sqrt{\text{Hz}}$.

Figure 4-4 summarizes the different modeling parameters for each component and shows which of them are fixed and which can be controlled through design choices. The code used for modeling the system and extracting the performance can be found in https://github.com/mdecea/CMOS_photonic_biosensor.

Ideal case: matching FSRs

An important assumption of the sensing scheme we are using is that the FSR of the tracking and sensing rings is the same, so that alignment of a single ring resonance between the tracking and sensing rings ensures alignment of all the other resonances as well. This is in general hard to achieve in real systems due to fabrication variations (see the next section), but we will start by analyzing this situation so we can obtain the best overall achievable performance.

Optimizing Scheme 1

In Scheme 1 (Fig. 4-3(a)) we have two design parameters:

- The FSR of the resonant LED. The spacing between the individual resonance lines generated by the LED does not affect the achievable sensitivity of the system (as long as such spacing is the same as in the sensing ring), so the system level performance is independent of the FSR. Nevertheless, it is likely that a resonant LED with a lower FSR generates more optical power for the same bias current due to Purcell enhancement affecting a larger fraction of the generated wavelengths (as discussed earlier, the overall increase in power due to Purcell enhancement is proportional to FWHM/FSR).
- The FWHM of the sensing ring. Choosing the optimal FWHM is not immediate: a narrower FWHM results in increased sensitivity because we are *averaging* a smaller section of the LED resonance spectrum, but also results in a smaller amount of optical power reaching the detector (and therefore a smaller generated photocurrent and decreased sensitivity). Also, as shown in Fig. 4-5(a), a narrower FWHM (i.e., a larger Q factor) results in a lower maximum transmission to the drop port due to larger propagation losses as the light travels multiple round trips in the ring.

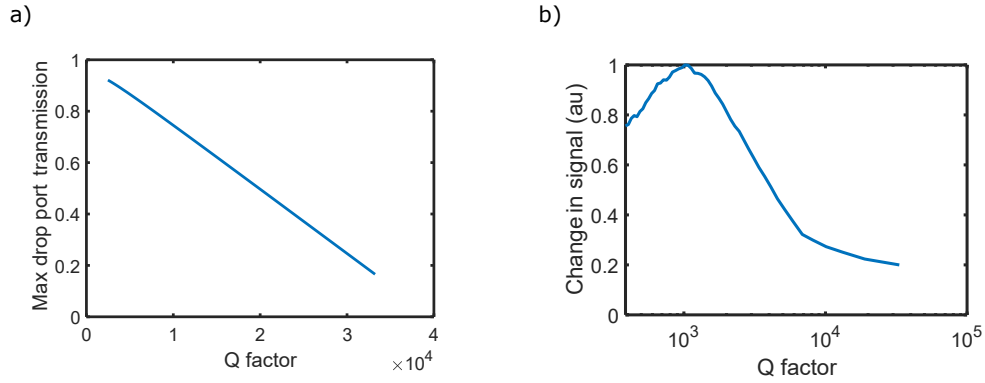


Figure 4-5: Optimizing the FWHM of the add-drop sensing ring in Sensing Scheme 1. (a) Maximum drop port transmission as a function of the ring loaded Q factor. (b) Change in readout signal as a function of the loaded Q factor of the sensing ring.

Figure 4-5(b) shows the readout signal change as a function of the Q factor of the add/drop ring for $\Delta\lambda = 80$ pm (an almost identical curve is obtained for a different value of $\Delta\lambda$). There is clearly an optimum point at which the current change is optimized, which in our case is $Q_{optim} = 1,000$.

Note that the FSR of the sensing ring is not a design parameter since we will design it to match the FSR of the resonant LED.

Now that we have the optimum parameters, we can evaluate the minimum detectable change in resonance wavelength and refractive index of the analyte as a function of the total power generated by the LED, which are shown in Fig. 4-6. To obtain the sensitivity to the change in refractive index of the analyte flowing on the sensor (commonly called the bulk refractive index sensitivity) from the sensitivity to the resonance wavelength change $\Delta\lambda$, we first need to calculate the sensitivity in the effective refractive index of the optical mode, which we can easily obtain using Eq. 4.1: $\Delta n_{eff} = n_g \Delta\lambda / \lambda_0$. From this, the bulk refractive index sensitivity can be obtained by calculating the change in effective index of the mode as the refractive index of the medium surrounding the silicon waveguide changes. We obtained this relation using Lumerical MODE simulations,

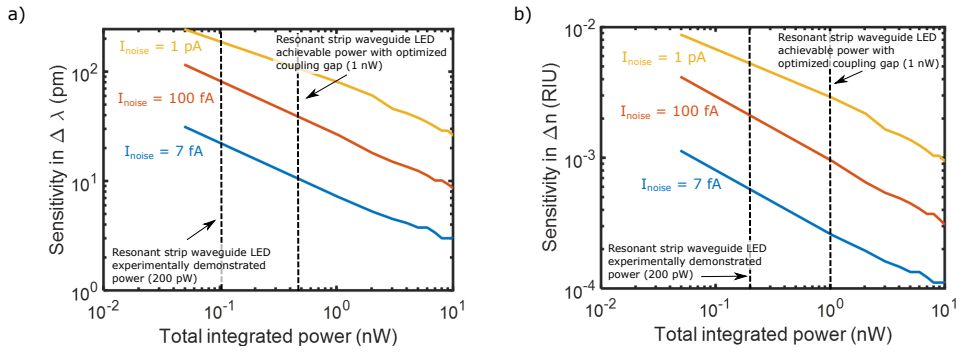


Figure 4-6: Sensitivity of Sensing Scheme 1 as a function of the total power generated by the LED in (a) resonance wavelength change, and (b) change in refractive index of the analyte.

resulting in $\Delta n_{\text{eff}} = 0.1137\Delta n$. For this simulations we considered the refractive index of the medium surrounding the waveguide to be close to the refractive index of water ($n=1.33$) except on the bottom side of the waveguide, where we assumed the presence of silicon dioxide (for supporting the waveguide).

Looking at Fig. 4-6, as expected a lower current sensor noise results in an increased sensitivity, and a higher total generated power also increases the sensitivity. For 200 pW of total optical power (corresponding to our experimental demonstration for the resonant strip waveguide LED in Section 3.4) the maximum achievable sensitivity in resonance wavelength change is $\Delta \lambda_0 = 16 \text{ pm}$, which corresponds to a bulk refractive index sensitivity of $5.75 \cdot 10^{-4} \text{ RIU}$. As discussed earlier, optimizing the ring-bus coupling gap of the resonator would result in a $5\times$ increase in generated power, which increases the sensitivity to $\Delta \lambda_0 = 7.3 \text{ pm}$ and a bulk refractive index sensitivity of $2.6 \cdot 10^{-4} \text{ RIU}$. Further increasing the generated power by $10\times$ (which could be achieved by minimizing Auger and surface recombination and increasing the bias current to the LED) would result in increased sensitivities of $\Delta \lambda_0 = 2.5 \text{ pm}$ ($9 \cdot 10^{-5} \text{ RIU}$ bulk sensitivity).

Optimizing Scheme 2

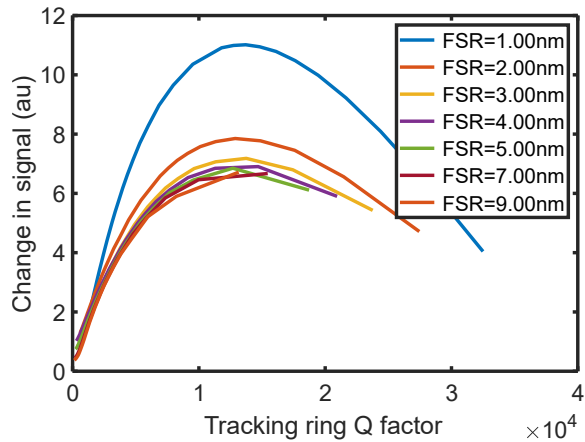


Figure 4-7: Change in readout signal as a function of the Q factor of the tracking ring for different FSRs. The Q factor of the sensing ring is optimized for every point.

Scheme 2 offers more design flexibility than Scheme 1, as it gives freedom to choose both the FSR and FWHM of the tracking ring as well as the FWHM of the sensing ring (once again, the sensing ring FSR is fixed to match the FSR of the tracking ring).

This poses an interesting optimization problem: similar to the case of Scheme 1, a smaller FWHM for the tracking ring results in a sharper resonance, which means that the fractional change in transmitted power is larger and therefore increases sensitivity. But, as we have seen in Fig. 4-5(a), decreasing FWHM also decreases the total transmitted power, resulting in a lower photocurrent signal. In terms of FSR, it is easy to see that we want to make it as small as possible so that as much of the power generated by the linear LED arrives at the sensing ring, of course with the constraint that such FSR needs to be at least $\approx 2\times$ the FWHM so that the spectrum of the light after the tracking ring shows resonances. It is also worth noting that, once the FWHM of the tracking ring is set, there exists an optimum FWHM for the sensing ring (as this is the same optimization problem we had in Scheme 1, which has a clear optimum as shown in Fig. 4-5(b)).

Figure 4-7 shows the change in signal as a function of the Q factor of the tracking ring

for different FSRs ². As discussed earlier, the smallest FSR results in the largest signal change because more of the power generated by the LED is transmitted through the first ring (the tracking ring in our case). We have limited the smallest FSR to 1 nm so that our rings have a radius below 50 μm , which is important to ensure the compactness of a single sensing macro so that a large number of analytes can be detected in the same chip (Fig. 4-2). For the optimum FSR of 1 nm, we get an optimal Q factor for the tracking ring $Q_{track} = 13,700$ ($FWHM_{track} = 82.5$ pm). Interestingly, the optimal Q factor for the sensing ring is the same as that of the tracking ring, $Q_{sens} = 13,700$ ($FWHM_{sens} = 82.5$ pm) and we know that $FWHM_{sens} = FWHM_{track} = 1$ nm.

The resonance wavelength and effective index sensitivities for the optimum parameters are shown in Fig. 4-8. Similarly to Scheme 1, increasing receiver sensitivity and LED power result in better refractive index sensitivity. For 300 pW of total optical power (corresponding to our experimental demonstration of the linear rib waveguide LED in section 3.4) the maximum achievable sensitivity in resonance wavelength change is $\Delta\lambda_0 = 2.4$ pm, which corresponds to bulk refractive index sensitivity of $\Delta n = 9.1 \cdot 10^{-5}$ RIU. If we use the 3.5 nW of output power we experimentally demonstrated with our overcoupled resonant rib waveguide LED (Section 3.4), then the maximum achievable sensitivity increases to $\Delta\lambda_0 = 0.5$ pm ($\Delta n = 2.6 \cdot 10^{-5}$ RIU).

Table 4.2 summarizes the obtained results for our proposed sensing schemes and compares the performance to the integrated and non-integrated ring resonator based refractometric sensors introduced earlier. It's worth noting a few points:

- Sensing Scheme 2 achieves better sensitivity than Sensing Scheme 1. This is because even though in Scheme 1 we are using all of the generated power to *feed* the sensor, the large FWHM of the generated light greatly limits the change in

²As discussed above, once these are set, both the FSR and Q factor of the sensing ring are also fixed.

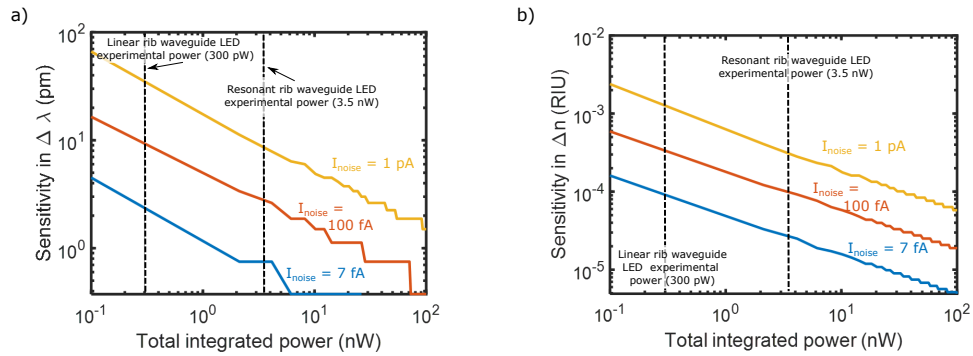


Figure 4-8: Sensitivity of sensing Scheme 2 as a function of the total power generated by the LED in (a) resonance wavelength change, and (b) change in refractive index of the analyte.

signal for small changes in the resonance wavelength. On the other hand, Sensing Scheme 2 can leverage the much sharper resonances of the tracking ring to generate a larger signal despite throwing away part of the light generated by the linear LED.

- Sensing Scheme 2 offers a path towards achieving resonance wavelength sensitivities that are close to the conventional, desktop size and \$100k+ non-integrated solutions. Interestingly, while we achieve a similar wavelength sensitivity than the literature demonstrations, our scheme shows an order of magnitude lower bulk refractive index sensitivity. This means that the group index of the optical mode in the literature demonstrations are much larger than in our resonators. For example, in the case of the commercial solution [258] the required group index of the optical for the wavelength sensitivity to be consistent with the bulk refractive index sensitivity comes out to be $n_g = 5.35$, and in the semi-integrated work [264] it is $n_g = 7.13$, which are much larger than our $n_g = 4.62$ ³.

³It's doubtful that these are the real group indices in the optical waveguides realized in works [258, 264].

Ref (year)	Brief description	Level of integration	LoD ($\Delta\lambda_0$)	LoD (RIU)	Size	Cost
This work	Sensing Scheme 1 resonant strip LED experimental dem	Fully monolithic	16 pm	$5.75 \cdot 10^{-4}$	0.25 mm ²	$\approx 1\$/\text{mm}^2$
This work	Sensing Scheme 1 resonant strip LED achievable power	Fully monolithic	2.5 pm	$9 \cdot 10^{-5}$	0.25 mm ²	$\approx 1\$/\text{mm}^2$
This work	Sensing Scheme 2 linear rib LED experimental dem	Fully monolithic	2.4 pm	$9.1 \cdot 10^{-5}$	0.25 mm ²	$\approx 1\$/\text{mm}^2$
This work	Sensing Scheme 2 resonant rib LED experimental dem	Fully monolithic	0.5 pm	$2.6 \cdot 10^{-5}$	0.25 mm ²	$\approx 1\$/\text{mm}^2$
[258] (2010)	RR external PD external TLS	Low	0.22 pm	$7.6 \cdot 10^{-7}$	Desktop	\$120,000
[264] (2018)	RR integrated PD external TLS	Medium	11.3 pm	$5.2 \cdot 10^{-5}$	Desktop	NR
[265] (2019)	MZI integrated PD (Si) integrated LED (SiN)	High	NA	$2 \cdot 10^{-6}$	37 mm ²	NR

Table 4.2: Performance of optical refractive index sensors with different levels of integration. RIU = refractive index unit. RR = Ring resonator. TLS = Tunable Laser Source. PD = photodetector. MZI = Mach-Zehnder Interferometer. NR = Not Reported. NA = Not Applicable.

Accounting for FSR variability

One of the basic assumptions in all our analysis above is that the FSR of the sensing and tracking rings are exactly the same. This ensures that when sensing and tracking rings are on resonance the maximum amount of power reaches the detector, therefore maximizing the readout signal. Any deviation from perfectly matched FSRs will result in a smaller readout signal when the rings are on resonance, which will decrease the achievable sensitivity. We therefore need to account for the effects that such FSR variability can have in the achievable performance of our sensors.

Achieving perfectly matched FSRs is challenging even for identical rings due to fabrication variations. The FSR of a ring resonator is given by $FSR = \lambda^2 / (n_g L)$, where L is the round trip length of the ring and n_g the group index of the optical mode propagating in the ring [260]. From this, it's clear that any variations in the group index of the waveguide will cause a change in FSR. In particular, $\Delta FSR = -(\Delta n_g / n_g) FSR$.

The main source of variability in n_g comes from fabrication variations that result in slightly different waveguide dimensions (both in thickness and width) for different positions in the fabricated chip [273]⁴.

The variability in n_g depends on the particular foundry process being used. In this sense, the fact that our sensors are fabricated in commercial CMOS photonics processes is highly advantageous, since these are arguably the most precise large scale semiconductor fabrication processes in the world, leveraging decades of research and expertise.

We studied the effect of the n_g variability (and consequent FSR mismatch) on the achievable sensitivity of both our proposed schemes. Figure 4-9 shows the achievable sensitivity for the experimentally demonstrated LED output powers as a function of the

⁴A detailed discussion on the sources and effects of fabrication variations in silicon photonics processes is out of the scope of this thesis. The interested reader is pointed to reference [273] as a good starting point.

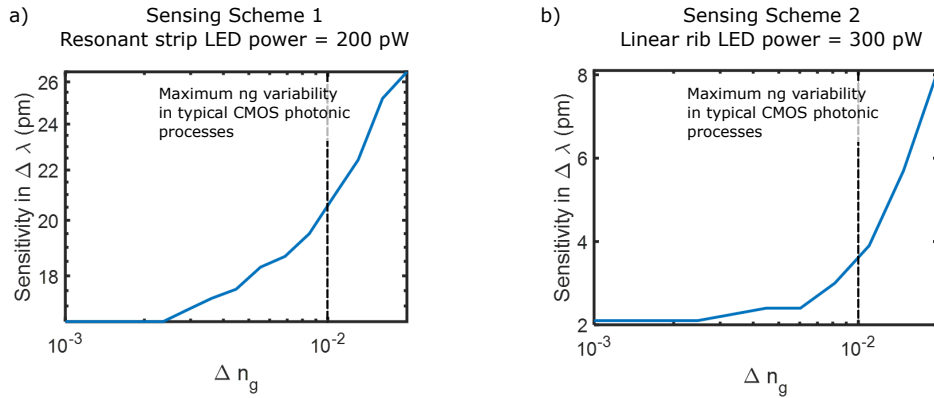


Figure 4-9: Sensitivity of Sensing Scheme 1 (a) and Sensing Scheme 2 (b) as a function of the group index mismatch between sensing and tracking ring. The vertical line shows worst case variability for waveguides in the same chip for typical CMOS photonics processes [273].

mismatch in group index between sensing and tracking rings. The vertical line indicates the group index variability of $\Delta n_g \approx 0.01$ observed in [273], which we consider the typical maximum variation in group index of waveguides fabricated in the same chip.

As expected, in both cases the sensitivity decreases as the variation in the group index of both rings increases. For $\Delta n_g = 0.01$ the sensitivity decreases about 25% for Sensing Scheme 1 (from $\Delta \lambda_0 = 16$ pm to $\Delta \lambda_0 = 20.6$ pm) and about 80% for Sensing Scheme 2 (from $\Delta \lambda_0 = 2$ pm to $\Delta \lambda_0 = 3.6$ pm). Sensing scheme 2 is more sensitive to FSR mismatch because the resonances are much narrower. Fig. 4-9 also shows how keeping the n_g variability below ≈ 0.002 basically eliminates the FSR mismatch effects, which could be achieved by using proper device design and layout techniques [273].

4.3 Experimental demonstration of resonator based refractive index sensing in a CMOS photonic platform

Throughout this chapter we have introduced the concept of a fully monolithic integrated refractive index sensor and done a thorough theoretical analysis of the achievable performance from a system-level perspective. We have seen how our systems can theoretically achieve LoDs similar to state of the art systems at a much lower cost and size.

As we showed in Fig. 4-2, the basic building blocks for our monolithic sensor are (1) an integrated LED, (2) a sensing resonator and (3) an integrated PD capable of detecting the light generated by the LED. It is important to notice that Chapter 3 experimentally demonstrated the realization of both the integrated source and the integrated PD (Section 3.4.2) on the same chip.

Thus, the only component left to demonstrate experimentally is the sensing resonator. A great number of publications have demonstrated refractive index sensing using ring resonators, both in custom and foundry-based silicon photonics processes [244]. As a result, the modeling, design, fabrication and post-processing of such sensing resonators is well established.

We performed a preliminary experimental demonstration of refractive index sensing using a silicon ring resonator, and the results are shown in Fig. 4-10.

Figure 4-10(a) shows a schematic of the tested device as well as the packaging approach to enable delivery of the analyte to the sensing area. A poly-silicon ring resonator with input and output grating couplers was fabricated in the same platform as the linear rib waveguide LED presented in Section 3 (SUNY Poly silicon photonics process, a preliminary version of the now commercially available AIM photonics base active silicon

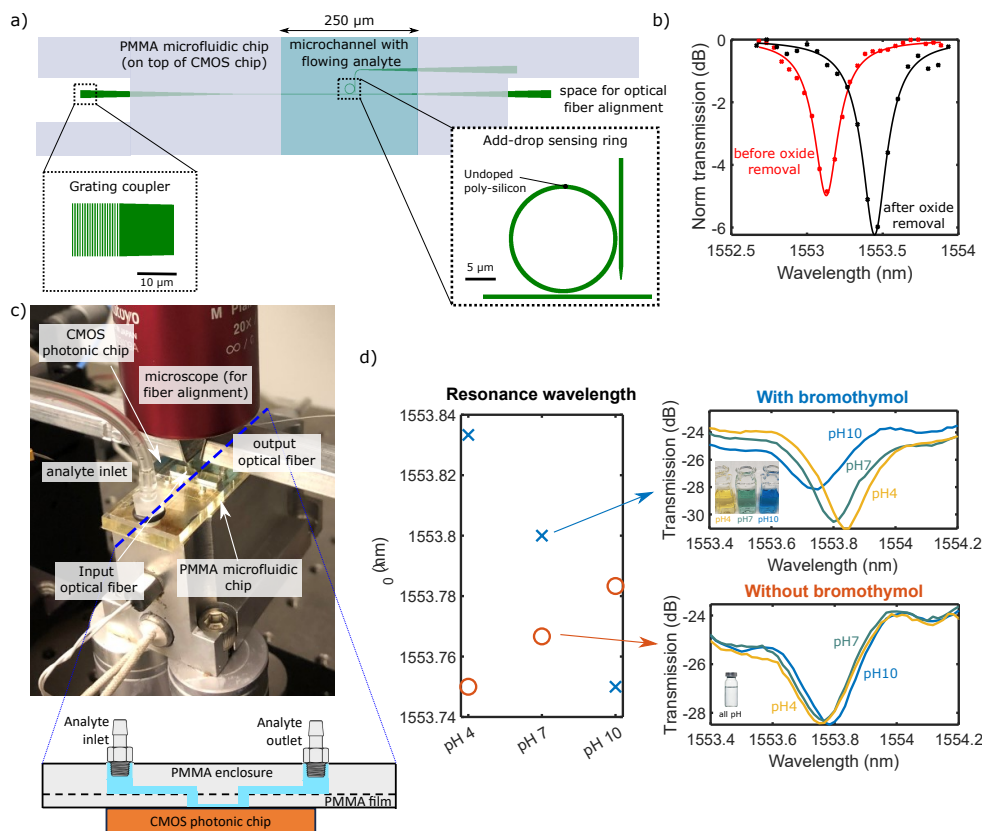


Figure 4-10: Demonstration of refractive index sensing with an integrated poly-silicon ring resonator. (a) A schematic of the tested device, which consists of an add-drop poly-silicon ring resonator fabricated in the SUNY Poly silicon photonics process. Grating couplers provide optical input and outputs, and a PMMA microfluidic chip bonded on top allows for targeted delivery of the analyte to the sensing ring. (b) The normalized transmission spectrum of the ring before (red) and after (black) BOE etching of the oxide cladding to expose the polysilicon waveguide to the analyte. (c) A picture of the experimental setup. The PMMA microfluidic chip, the CMOS photonic chip and the input and output optical fibers are visible. A schematic of the microfluidic package for analyte delivery is shown on the bottom. (d) Measured resonance wavelength of the device for solutions with different pH (left), and transmission spectrum for the same solutions (right). Two sets of data are shown, one with the solutions *as is* (orange circles) and one where bromothymol blue has been added to the solutions (blue crosses). The insets show the color of the different pH solutions.

photonics platform). It consists of a 650 nm wide, 18 μm outer diameter add-drop ring resonator designed for operation at a wavelength around 1550 nm. To expose the silicon waveguide, we removed the oxide cladding from the entire chip using 7:1 Buried Oxide Etch (BOE). As shown in Fig. 4-10(b), before the etch the ring resonator had an extinction ratio of about 5 dB and a loaded Q factor of 10,300. After etching, the resonance shifted about 0.32 nm, the ER increased slightly to 6.2 dB and the loaded Q factor decreased to 7,750, showing that the ring was indeed exposed after etching.

As depicted in Figs. 4-10(a,c), after the oxide etching a PMMA microfluidic chip with a 250 μm wide microchannel was bonded on top of the CMOS photonic chip so that the analyte could be delivered to the sensing ring resonator. The PMMA chip was fabricated by Dr. Jaehwan Kim, and the approach and specifications can be found in his PhD thesis [274]. The microfluidic chip was fabricated by laser-cutting and consisted of two layers: (1) an enclosure acrylic layer that provided the tubing interface and holes for input and output optical fibers access to the grating couplers, and (2) an acrylic film with pressure-sensitive adhesive that hosted the microfluidic channel and provided the adhesive for reliable bonding to the CMOS chip. The bonding of the microfluidic chip to the CMOS chip was done under a semiconductor inspection microscope using a microprobe arm that was modified to hold the microfluidic chip and align the channel above the sensing region.

Figure 4-10(c) shows a picture of the experimental setup: The PMMA microfluidic chip allowing delivery of the liquids is clearly visible, as well as the optical fibers for in- and out-coupling of the light coming from a tunable laser. Notice how our monolithic refractive index sensor would fully eliminate the need for input and output optical fibers, greatly simplifying the packaging requirements and resulting in a much more compact system.

To show the feasibility of using our platform for refractive index sensing we performed

a pH sensing experiment, where we flowed solutions with different pH and measured the resonance wavelength of the sensing resonator by performing wavelength sweeps with a tunable laser. The results are shown in Fig. 4-10(d). Two different sets of experiments are shown:

- One where the solutions as-is were flown into the resonator (orange circles in Fig. 4-10(d)). We can clearly see how an increase in the pH of the solution is accompanied by an increase in the resonance wavelength of the resonator, denoting an increase in the effective index of the optical mode propagating along the ring. This is somewhat surprising, as the refractive index of the solution is not expected to change significantly with pH. Nevertheless, this effect has been previously recognized in silica-based nanocomposite thin film sensors, and it is attributed to the effect of pH dependent charging and ionic adsorption in the silica surface [275]. To the best of our knowledge, this is the first time that such an effect has been reported in silicon. As apparent from Fig. 4-10(d), the resonance wavelength of the ring shifts about 35 pm going from a solution with pH4 to a solution with pH10. Using the performance specs we derived in our theoretical study (Table 4.2), our fully monolithic refractive index sensors could achieve a resolution as low as 0.086 pH units.
- One where bromothymol blue was added to the solutions (blue crosses in Fig. 4-10(d)). Bromothymol blue is a compound that changes its color according to the pH of the solution it is diluted in (see inset of Fig. 4-10(d)). Bromothymol blue is part of a broader family of compounds known as colorimetric dyes or indicators [276], chemicals that change its color according to their chemical or physical environment. A myriad of colorimetric sensors (sensing anything from pH to the presence of volatile organic compounds or food quality) have been reported in the literature

[276].

The readout of such sensors is done through measuring the color change after exposure, which is usually done through the use of a digital camera. Instead, we indirectly measure the change in color through the change in refractive index. To do this, we exploit the fact that a change in the absorption of a material (i.e, a change in its color) must be accompanied by a change in the real part of its refractive index due to the Kramers-Konig relation [277]. This is clearly observed in Fig. 4-10(d), where we can see how the resonance wavelength of the ring changes as the different pH solutions are flowed. In contrast with the case with no colorimetric sensor, the addition of bromothymol blue results in a decrease in resonance wavelength (i.e, a decrease in refractive index) as the pH increases. The total resonance wavelength shift from pH4 to pH10 is now -83 pm, about 2x larger than the case without a colorimetric compound. We can see how, essentially, the addition of a colorimetric sensor acts as an *amplifier* of the change in refractive index. This is, to the best of our knowledge, the first time that amplification of a refractometric sensing signal through the use of colorimetric compounds has been reported. This is a promising avenue to improve the resolution of refractive index sensors with weak intrinsic response. With the measured wavelength change, our fully monolithic refractive index sensors could achieve a resolution as low as 0.04 pH units using colorimetric amplification ⁵.

⁵Notice how, unfortunately, the intrinsic response due to a pH increase causes an increase in the refractive index, while the colorimetric compound causes a decrease in refractive index. Thus, the amplification provided by the colorimetric compound is partially eliminated by the intrinsic response due to pH. If both responses were in the same direction, the resonance shift from pH4 to pH10 could be as large as 113 pm, which would boost the resolution of our fully monolithic sensor down to 0.025 pH units. This highlights the importance of choosing colorimetric amplifiers that cause a refractive index shift in the same direction as the intrinsic response of the system to the property being measured (if there is an intrinsic response).

With this we have experimentally demonstrated successful operation of all the components necessary to build fully monolithic refractive index sensors: (1) an integrated LED, (2) a sensing resonator and (3) an integrated PD capable of detecting the light generated by the LED.

4.4 Conclusion

In this chapter we have studied in detail the application of the silicon waveguide coupled light sources we developed in Chapter 3 to the realization of fully integrated monolithic refractive index sensors. We have argued how our monolithic sensor would overcome the size and cost limitations of current non-integrated solutions, and offer an approach toward highly multiplexed sensing in a single chip by replication of compact sensing macros. This would allow for ubiquitous, real-time monitoring of a variety of analytes with applications in a plethora of fields, including food and water quality monitoring, hazardous gas sensing, realization of electronic noses, biosensing and pharmaceutical/nutraceutical process monitoring.

After detailed modeling and optimization we have shown that a sensing scheme where we use a linear LED and a passive tracking ring can offer sensitivities approaching the best reported in the literature (Table 4.2) even with the low powers generated by our LED and the FSR mismatch coming from fabrication variations.

Chapter 5

OPA architectures for advanced optical beamforming applications

As explained in Chapter 1, the realization of Light Detection and Ranging (LIDAR) systems is by far the most explored application in integrated optics for imaging, both commercially and in academia. In particular, the use of Optical Phased Arrays (OPAs) to realize non-mechanical beamforming and beamsteering is a promising avenue that necessitates integrated optics due to the small ($\sim \lambda$) emitter size and spacing required to achieve a large field of view (FoV) and angular resolution.

Tremendous progress has been reported over the last decade on the realization of OPAs with increasing number of emitters, reduced emitter spacing and lower power consumption (Tables 1.9, 1.10, 1.11).

As the technology matures, its application space expands and its performance requirements increase. In Section 5.1 of this Chapter we introduce two advanced applications for OPA-based beamforming applications: reconfigurable FSOC networks and advanced LIDAR. We derive performance requirements for these systems and show how the standard

OPA approach based on a uniformly spaced array of emitters is not practical.

With this in mind, in Section 5.2 we propose the use of non-uniformly spaced OPAs as a solution to such a bottleneck. We perform a detailed study of the performance tradeoffs between the two types of OPAs (uniformly and non-uniformly spaced) in Section 5.2.1, and then in Section 5.2.2 we show how non-uniform OPAs can offer superior performance for the realization of large scale, high performance OPAs with reduced number of elements and power consumption in certain scenarios.

5.1 Advanced OPA-based beamforming applications

In Section 1.4.2 we introduced the use of OPAs for beamforming applications in automotive LIDAR and for correcting pole sway and beam wander in FSOC systems. These constitute the target applications for most OPA systems being developed at the moment.

Nevertheless, additional, advanced applications are being considered as the performance and maturity of these systems improves. Here, we will analyze two of these advanced applications:

1. Advanced automotive LIDAR: In Section 1.4.2 we used [116] to derive the specifications for an OPA-based LIDAR system for automotive applications. These specifications constitute a *viable* LIDAR system capable of detecting cars at about 200 m range, which is meant to be used as driver assistance system (commonly known as Advanced Driver Assistance System - ADAS).

Beyond driver assistance, one can also envision the use of LIDAR systems in fully autonomous, self-driving cars. The performance requirements in this scenario are significantly increased, both in terms of resolution and FoV, as now we need to ensure higher quality data to be fed into the Machine Learning (ML) algorithms

that make the driving decisions. We use the requirements derived in [111] for this scenario, summarized in Table 5.1 (along with the performance requirements of the viable performance LIDAR system). These are based on ensuring detection of preceding vehicles in curves and the ability to detect a cyclist at a distance of 200 m with ML algorithms.

2. Reconfigurable FSOC networks: In Section 1.4.2 we introduced OPA systems as a means to maintain line of sight between FSOC terminals in face of displacements of the beam due to pole twist and sway and atmospheric beam wander.

A natural extension of such FSOC systems is in reconfigurable FSOC networks where, besides correcting for beam misalignment, a terminal can switch its pointing direction to establish communication with a different terminal (Fig. 5-1). This can be highly advantageous in a variety of FSOC-based systems, including reconfigurable data center networks [278], 5G and beyond networks [279], and satellite networks [50].

Of course, in this scenario we want an OPA with as large a FoV as possible so we can maximize the link directions that can be covered by a single terminal (which results in a lower required node count). We can derive a requirement for FoV based on typical numerical aperture (NA) values for lenses, since even if the OPA can generate a beam at an angle θ it won't be efficiently collected by the lens if that angle is outside its numerical aperture. The maximum NA of commercially available lenses with apertures diameters in the 5-10 cm range (typical for commercial FSOC systems) is about 0.4 [280], which translates into a maximum covered FoV of 47° . In the vertical direction we can derive a FoV requirement based on the maximum height difference that can exist between two terminals to be connected. Assuming a maximum of 10% degree inclination between two terminals seems rea-

	Viability performance [116]	High performance [111]
Field of View	$14^\circ \times 3.5^\circ$	$32^\circ \times 10.2^\circ$
Angular resolution	$0.1^\circ \times 0.1^\circ$	$0.04^\circ \times 0.04^\circ$
Frame rate	25 fps	25 fps
Sampling rate	122.5 kHz	5.1 MHz

Table 5.1: Requirements for viable (left column) and advanced (right column) OPA-based LIDAR systems for automotive applications.

sonable, which translates into a FoV requirement given by $2\text{atan}(0.1) = 11.42^\circ$. The requirements for angular resolution are the same as for the original FSOC system that only corrects for beam displacement, as the requirements on maximum acceptable coupling loss do not change. The new performance requirements for reconfigurable FSOC networks are summarized in Table 5.2.

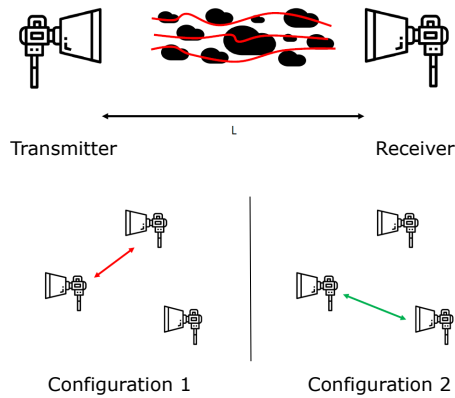


Figure 5-1: OPA-based reconfigurable FSOC networks. Besides correcting for beam wander due to atmospheric effects and pole twist and sway (top), the system is able to point to a different FSOC terminal and reconfigure the communication network.

	FSOC - Beam displacement correction	Reconfigurable FSOC network
Field of View	$3^\circ \times 3^\circ$	$47^\circ \times 11.42^\circ$
Max misalignment loss	1 dB	1 dB
Angular resolution	$0.02^\circ \times 0.02^\circ$	$0.02^\circ \times 0.02^\circ$
Sampling rate	100 kHz	100 kHz

Table 5.2: Requirements for simple (left column) and advanced (right column) OPA-based FSOC systems.

5.1.1 Performance requirements

As is apparent from Tables 5.1 and 5.2, these advanced beamsteering applications require OPAs with significantly increased performance.

We can use the same procedure we followed in Section 1.4.2 to translate the requirements on angular resolution and FoV to required number of elements, spacing, area and power consumption of phase shifters for an OPA based on a uniformly spaced 2D array of emitters.

As a summary of the process ¹:

1. We can translate FoV and angular resolution to required number of elements and spacing between elements using the equations in Table 1.4.
2. Once we know the number of emitter elements we know the number of phase shifters required to control the OPA. With this, we can derive the maximum power consumption per phase shifter - for LIDAR we assume a total of 3 W power consumption for the OPA phase shifter control (30% of the 10 W total power budget of the system, which is based on commercially available LIDAR systems), and for FSOC we assume a total of 21.3 W (30% of 71 W, which corresponds to the

¹The reader is pointed to Section 1.4.2 for a detailed analysis.

maximum power that can be delivered by Power over Ethernet (PoE)).

3. Knowing the total number of required phase shifters we can also calculate the maximum area occupied by a single phase shifter, based on a 30% area budget allocated to the phase shifter network and a total chip area of $3 \text{ cm} \times 3 \text{ cm}$.

Table 5.3 summarizes the performance requirements for both *viable* (derived in Section 1.4.2) and advanced LIDAR and FSOC beam steering systems. As we reviewed in Section 1.4.4, no reported demonstration to date is capable of delivering all the required performance specifications even for the viable (lower performance) systems. Nevertheless, incremental work on already existing architectures (summarized in Tables 1.9, 1.10, 1.11) can arguably deliver the required performance for these applications (left column in Table 5.3).

This is harder to believe for the advanced beamsteering applications we just discussed, which require OPAs approaching a million elements with sub- $5 \mu\text{m}$ spacing and very compact and efficient phase shifters with $< 30 \mu\text{W}$ power consumption and $\approx 20 \mu\text{m} \times 20 \mu\text{m}$ area.

The natural question that arises is then: can we find alternative architectures to deliver the required performance, or are we reaching the limit in scalability of integrated OPAs?

5.1.2 A suitable system-level Figure of Merit (FoM) for comparing OPA architectures

In order to answer the question we just posed we need to define a FoM that allows to compare different OPA architectures from a system-level perspective.

LIDAR

	Viable performance [116]	High performance [111]
Field of View	$14^\circ \times 3.5^\circ$	$32^\circ \times 10.2^\circ$
Angular resolution	$0.1^\circ \times 0.1^\circ$	$0.04^\circ \times 0.04^\circ$
Frame rate	25 fps	25 fps
Sampling rate	122.5 kHz	5.1 MHz
Emitter spacing	$6.3 \mu\text{m} \times 25 \mu\text{m}$	$2.8 \mu\text{m} \times 8.7 \mu\text{m}$
Array size	120×31 N = 3,720	690×221 N = 152,490
Phase shifter power	800 μW	20 μW
Phase shifter area	0.07 mm^2 ($\approx 250 \mu\text{m} \times 250 \mu\text{m}$)	$1,770 \mu\text{m}^2$ ($\approx 42 \mu\text{m} \times 42 \mu\text{m}$)

FSOC

	FSOC - Beam displacement correction	Reconfigurable FSOC network
Field of View	$3^\circ \times 3^\circ$	$47^\circ \times 11.42^\circ$
Max misalignment loss	1 dB	1 dB
Angular resolution	$0.02^\circ \times 0.02^\circ$	$0.02^\circ \times 0.02^\circ$
Sampling rate	100 kHz	100 kHz
Emitter spacing	$29 \mu\text{m} \times 29 \mu\text{m}$	$2 \mu\text{m} \times 7.8 \mu\text{m}$
Array size	132×132 N = 17,424	1920×490 N = 940,800
Phase shifter power	1.22 mW	22.6 μW
Phase shifter area	0.015 mm^2 ($\approx 125 \mu\text{m} \times 125 \mu\text{m}$)	$290 \mu\text{m}^2$ ($\approx 17 \mu\text{m} \times 17 \mu\text{m}$)

Table 5.3: Requirements in phase shifter power and area for OPA-based FSOC and LIDAR systems.

We will distinguish here three sets of applications for OPA-based beam steering systems:

1. Optical power limited applications: These are applications in which the main goal for the OPA is to maximize the optical power contained in the main lobe for a given desired angular resolution, and where electrical power consumption is not a major concern. These applications use the maximum possible input optical power to the OPA (which is usually set by how much optical power can a waveguide sustain before nonlinear effects become relevant) and minimizing the optical loss of the OPA is crucial.

In this case, the relevant figure of merit is simply the optical power contained in the main lobe $FoM = P_{opt,ml}$.

An example application would be long reach FSOC systems. Here, electrical power dissipation is not a major concern, whereas the optical power emitted by the OPA is paramount to maximize the achievable link distance.

2. Total (electrical + optical) power limited applications: these are applications in which the system is not optical power budget limited, but instead we are trying to minimize total power dissipation. In this case, it is worth recognizing that the main goal of an OPA is to deliver the maximum possible link distance L or SNR ² with the minimum total power consumption P_{el} . Thus, a suitable FoM is:

$$FoM = \frac{L}{P_{el}} \quad (5.1)$$

²Notice that these two metrics are equivalent - having a system with a larger achievable link distance for a given receiver sensitivity is equivalent to having a system that achieves a larger SNR for a fixed link distance.

Basically, this figure of merit measures how much power is required per unit of link distance.

We will assume a simple free space propagation model describing the diffraction of a collimated optical beam [132] where the losses are proportional to L^2 (where L is the link distance)³. Thus, we can relate the output optical power in the main lobe emitted by the OPA $P_{opt,ml}$ to achievable link distance L through $L \propto \sqrt{P_{opt,ml}}$.

An example application could be short reach FSOC systems (such as for 6G picocell backhaul [281]) or LIDAR systems with high sensitivity detectors. Minimizing total power consumption in these scenarios is important to minimize operation expenses and maximize battery life.

3. Asymmetric link applications: Above we have considered applications with a symmetric link - where receiver and transmitter have a similar set of constraints in terms of power consumption or required performance.

A different set of applications is that of asymmetric links, where the requirements for the transmitter and receiver are very different. An example of such an asymmetric link is in satellite to earth communications: one of the terminals (the satellite) has very strong limitations in the total power consumption of the transceiver (as it will affect the operating time of the satellite), while the other terminal (the ground receiver) can consume as much power as required (as it has access to a virtually infinite power source).

In this case, the relevant FoMs are different for the two terminals. For the terminal that is power constrained we still want to maximize the achievable link distance for a given power consumption budget, so we can use the same FoM as the scenario

³More realistic models would need to account for atmospheric absorption and distortion, as well as limited reflection efficiency of targets in the case of LIDAR. This is beyond the scope of our analysis.

above: $FoM = \frac{L}{P_{el}}$. For the terminal that is not power constrained we simply want to maximize both the receiver sensitivity (so that it can better receive the signal from the power-starved terminal) and emitted power (so that the signal reaching the power-starved terminal is as large as possible). Thus, the relevant FoM for the non-constrained terminal is again simply $FoM = P_{opt,ml}$.

Other examples of asymmetric links could be the readout of cryogenic signals to room temperature through FSOC links [282] or communication of Internet of Things (IoT) sensor networks to IoT gateways [283].

It is important to note here that these are system-level FoMs. Of course, we need to ensure that the different OPA architectures fulfill other specifications, such as FoV and angular resolution (Table 5.3), and that they can be fabricated and operate reliably. If we have two OPA architectures delivering sufficient FoV and angular resolution, we can then use the above FoMs to compare their performance from a system-level perspective.

5.1.3 What now?

From Table 5.3, we clearly need alternatives to the use of uniformly spaced OPA architectures if we want to enable the advanced beamsteering applications we discussed. The basic goal of these alternative architectures should be the reduction of the required number of elements, as this allows to increase the power consumption and reduce the area requirements of the phase shifters and also reduces packaging complexity. Also importantly, increasing the minimum spacing between emitter elements allows for reduced crosstalk and simplified chip layout.

A variety of approaches arise to tackle the intractable complexity of uniform OPAs for advanced applications:

1. Use of 1D OPA arrays with wavelength steering in the other direction. As discussed in Section 1.4.4, a common OPA architecture to achieve 2D beam steering relies on the use of phase shifters to achieve steering in one direction and on a change in the operation wavelength to achieve steering in the other direction (exploiting the dispersion in the emission angle of grating couplers with wavelength).

This approach greatly reduces the required number of phase shifters as they are only needed to control the emission angle in one direction. For example, for the high performance LIDAR (FSOC) system we would require only 690 (1920) elements, which translates into a $220\times$ ($490\times$) reduction in number of elements and control signals, and a $220\times$ ($490\times$) increase in allowable phase shifter power and area.

Nevertheless, the use of these architectures has several associated drawbacks:

- They require high power tunable lasers (which are expensive and require careful packaging and control). Furthermore, to achieve a high pointing stability and accuracy the emission wavelength has to be very accurately controlled, which is challenging in the typical operation environment of these systems.
- Fabrication variations across the chip result in a slightly different behavior of the emitting aperture as the wavelength is tuned. This causes unwanted side lobes and reduces the side lobe suppression ratio (SLSR), which can compromise operation. Such variations are not easy to compensate or correct since we do not have individual control of the emission phase of each element.
- The limited wavelength tuning range of the laser sets a limit to the maximum achievable steering range in the direction controlled by the emission wavelength. Furthermore, all the optical components need to be designed for high performance operation over the whole wavelength range of the laser, which

can be challenging in integrated optical platforms.

- Maintaining communication between the Tx and Rx terminal as the wavelength is shifted poses an important challenge for the alignment algorithm, as the Rx terminal needs to be able to detect such a change in wavelength accurately and quickly adapt to it.

2. Use of imaging systems: The use of imaging systems allows one to tradeoff angular resolution with FoV - the use of a magnification system increases the effective aperture size (and thus increases angular resolution) but also increases the emitter spacing (and thus reduces the FoV). Alternatively, a demagnification system decreases emitter spacing (and thus increases FoV) at the expense of a smaller aperture size (and thus worse angular resolution).

While the use of imaging systems can be useful to overcome, for example, practical limitations in achievable emitter spacing, it does not constitute a solution to reduce the number of emitters or increase the power consumption and area of the phase shifters.

3. Use of mechanically moving elements: A way to reduce the FoV requirements of advanced beam steering systems is to break down the pointing into two stages: (1) A slow, high angular range movement stage, which is usually implemented with a mechanically moving mirror, and (2) a fast, small range movement stage, which is implemented by the OPA.

Such an approach allows for a great reduction in the required FoV of the OPA, as the heavy lifting in the movement of the beam is done by the mechanically moving mirror:

- For reconfigurable FSOC networks, we can safely assume that the maximum

operation speed of the mechanically moving mirror (usually in the 100 Hz range) is faster than the rate at which we need to reconfigure the network. With this assumption, the OPA only needs to take care of correcting for beam displacement due to atmospheric phenomena and pole twist and sway. Thus, the requirements for the OPA are reduced to the case we analyzed in Section 1.4.2 (first column in Table 5.3).

- The situation is more involved in the case of advanced LIDAR - the fast steering mirror cannot operate at the required 5.1 MHz speed, so advanced schemes where both the OPA and the mirror are moving at the same time need to be employed to achieve a large scan FoV with the required speeds. A crude estimation of the OPA required FoV is to divide the full required FoV by the max operation speed of the mirror (again, around 100 Hz), yielding a $32^\circ \times 10.2^\circ / 100 = 3.2^\circ \times 1.02^\circ$ FoV required with the same angular resolution ($0.04^\circ \times 0.04^\circ$).

It is clear that the use of mechanically moving mirrors offers an approach to decrease the OPA element count and spacing significantly without sacrificing other performance metrics. Nevertheless, besides being slow, mechanically moving mirrors are bulky, prone to failure, require complex assembly (including alignment), and are fairly expensive (ranging from \$3,000 - \$20,000 [284–286]). Migrating away from the use of mechanically moving parts to fully integrated beamsteering systems is required for realizing small form factor products with a sufficiently low cost to make them attractive for mass markets such as LIDAR or reconfigurable FSOC networks for 5G and beyond.

4. Use of non-uniform OPAs: Uniformly spaced OPAs have been the standard approach for the realization of integrated beamsteering systems. Nevertheless, over

the last few years, the use of emitter arrays with non-uniform spacing has been proposed as a means to reduce the number of elements required to achieve a given FoV and angular resolution [80, 161, 287–293]. While, for a fixed FoV, the angular resolution of a uniform OPA system increases linearly with the number of elements, it scales up to quadratically with the number of elements in the case of non-uniform OPAs [80]. This allows for a great reduction in the required number of elements for our advanced OPA applications, which could make the use of fully integrated beam steering systems without mechanically moving parts feasible. Nevertheless, the use of non-uniform OPAs has several associated drawbacks, such as limited side-lobe suppression ratio (SLSR), that have not been studied in detail in the existing literature.

In the next section we study in detail the use of non-uniform OPAs for the realization of large FoV, high angular resolution beam steering systems with a focus on the advanced OPA applications we discussed (Table 5.3).

Ref. (year)	1D / 2D	Number of elements	Min. emitter spacing (μm)	FoV ($^\circ$)	Angular resolution ($^\circ$)	SLSR (dB)	Emitter placement algorithm
[287] (2021)	1D phase 1D λ	128	NR average: 19.7	1: $140^\circ \times 19.23^\circ$ 2: $100^\circ \times 19.4^\circ$	1: $0.021^\circ \times 0.1^\circ$ 2: $0.021^\circ \times 0.029^\circ$	7	NR
[288] (2023)	2D phase	128	5.6	$24^\circ \times 16^\circ$	$0.31^\circ \times 0.49^\circ$	9	Optimize (GA)
[289] (2022)	1D phase	1024	1.24 average: 1.92	120°	NR	15.3	Optimize (PSO)
[80] (2021)	2D phase	127	15	$5.92^\circ \times 5.92^\circ$	$0.043^\circ \times 0.043^\circ$	6.3	Deterministic (Costas array)
[291] (2022)	2D phase	81	12 average: 20	$14^\circ \times 11^\circ$	$0.39^\circ \times 0.41^\circ$	5	Optimize (GA)
[161] (2019)	2D phase	128	5.6	$16^\circ \times 16^\circ$	$0.8^\circ \times 0.8^\circ$	12	Optimize (GA)
[144] (2015)	1D phase 1D λ	32	3.4	$23^\circ \times 3.6^\circ$	$1^\circ \times 0.6^\circ$	8.2	Deterministic (gaussian)
[294] (2022)	2D phase	64	9.2	$23^\circ \times 16.3^\circ$	$0.6^\circ \times 0.6^\circ$	11.3	Deterministic (coprime)
[107] (2016)	1D phase 1D λ	128	5.4 average: 7.24	$80^\circ \times 17^\circ$	$0.14^\circ \times 0.14^\circ$	8.8	Pseudo-random
[290] (2023)	2D phase	81	15	NR	$0.32^\circ \times 0.32^\circ$	8	Deterministic (circular)
[156] (2020)	1D phase	64	1.95 average: 15.4	50°	0.17°	9	Optimize
[295] (2015)	2D phase	64	30	$2.2^\circ \times 1.2^\circ$	$0.31^\circ \times 0.31^\circ$	7	Pseudo-random

Table 5.4: A survey of non-uniform OPAs reported in the literature. GA = Genetic Algorithm. PSO = Particle Swarm Optimization. NR = Not Reported. NA = Not Applicable.

5.2 Non-uniform OPAs

As presented in Section 1.4.1, a uniformly spaced OPA creates a far field radiation pattern with well defined grating lobes that sets a limitation to the achievable FoV, with required spacings on the order of $\lambda/2$ to achieve FoV close to $\pm 90^\circ$.

In contrast, non-uniformly spaced OPAs can achieve the same main lobe radiation width (i.e, the same angular resolution) without generating grating lobes. This is because in uniformly spaced OPAs there is constructive interference at multiple radiation angles, while the unequal inter-element spacing prevents that from happening in non-uniform OPAs (Fig. 5-2).

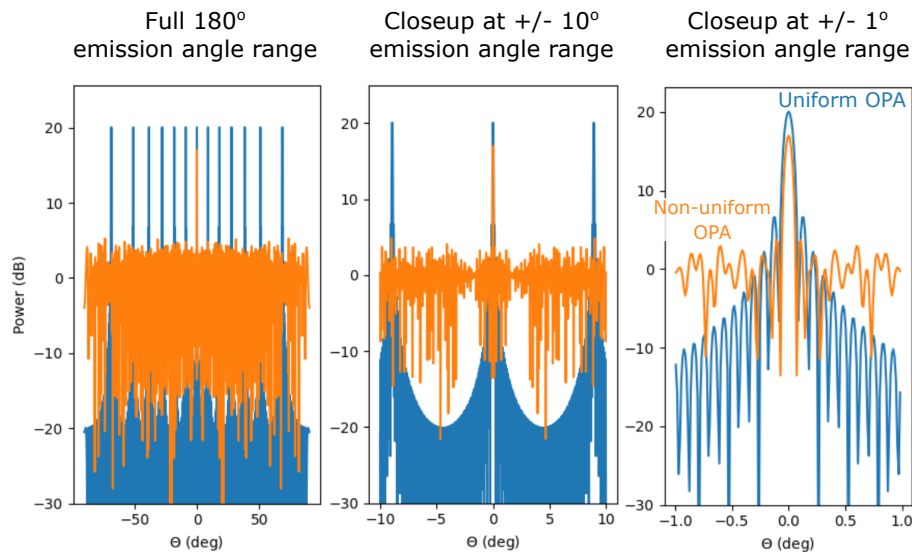


Figure 5-2: The radiation pattern for a uniformly spaced (blue) and a non-uniformly spaced (orange) 1D OPA with the same total extent of 2 mm. The uniformly spaced OPA has $N_u=100$ elements, and the non-uniform OPA has $N_{nr}=50$ elements. The different plots show the same radiation pattern but with a different extent in the angular axis (the x axis). In contrast to uniform OPAs, non-uniform OPAs have no grating lobes.

In practice, this means that non-uniformly spaced OPAs can achieve the same angular resolution as a uniform OPA but with increased FoV and a significantly lower number of

elements. As we will review in detail below, the main drawback of non-uniform OPAs is that they generally have a lower SLSR and lower power contained in the main radiation lobe - essentially they have a lower diffraction efficiency so a higher laser power is required to achieve the same emitted power. It is worth mentioning that non-uniform emitting arrays were studied as early as the 1960s in the context of RF phased arrays [296, 297].

While the concept of non-uniform arrays has been around for a long time, it is only in the last ~ 5 years that it has started to take prominence in the field of OPAs (see Table 5.4). This is likely because small, proof-of-concept OPAs with limited resolution and/or FoV can be realized with uniform OPAs, which are easier to analyze and design. But, as detailed in the previous section, as the community evolves towards the realization of OPAs for practical applications, uniform OPAs cannot deliver the required performance at a reasonable element count (Table 5.3).

Table 5.4 lists most of the non-uniformly spaced OPA demonstrations reported in the literature. As we can see, in general these arrays achieve a similar angular resolution with a larger FoV, lower number of elements and minimum element spacing than traditional, uniformly spaced OPAs (Tables 1.9, 1.10, 1.11).

One added complexity of the design of non-uniform OPAs is the need to determine the position of the emitters. Usually, one defines a uniform grid of possible antenna positions based on the minimum acceptable spacing between elements, and then a given desired number of antennas is placed in such grid based on some design procedure / algorithm. Various procedures have been used to determine the placement of these elements (Fig. 5-3):

- **Deterministic placement:** The different elements are placed according to a given analytical function. A variety of deterministic approaches have been reported in the literature, including circular arrangement [290], gaussian spacing [144], linearly

increasing spacing and spacing following \cos^x functions [293].

Another interesting category of deterministic placement approaches is based on the concept of co-prime arrays [149, 294]. Here, the non-uniform array is composed of subarrays of equally spaced elements with spacings $s_i = q_i \times s_{min}$. If the set of q_i are co-primes (i.e, their greatest common divisor is 1) then the grating lobes generated by each subarray are non-overlapping, therefore eliminating the FoV limitation of uniform arrays. This concept has been proposed in two different configurations: one where an OPA is composed of co-prime subarrays [149] and another where the Tx and Rx OPA are uniformly spaced arrays with a different, co-prime spacing [294].

Another category of deterministic non-uniform arrays is that based on non-redundant spacing, where the element placement is chosen such that the spacing between any

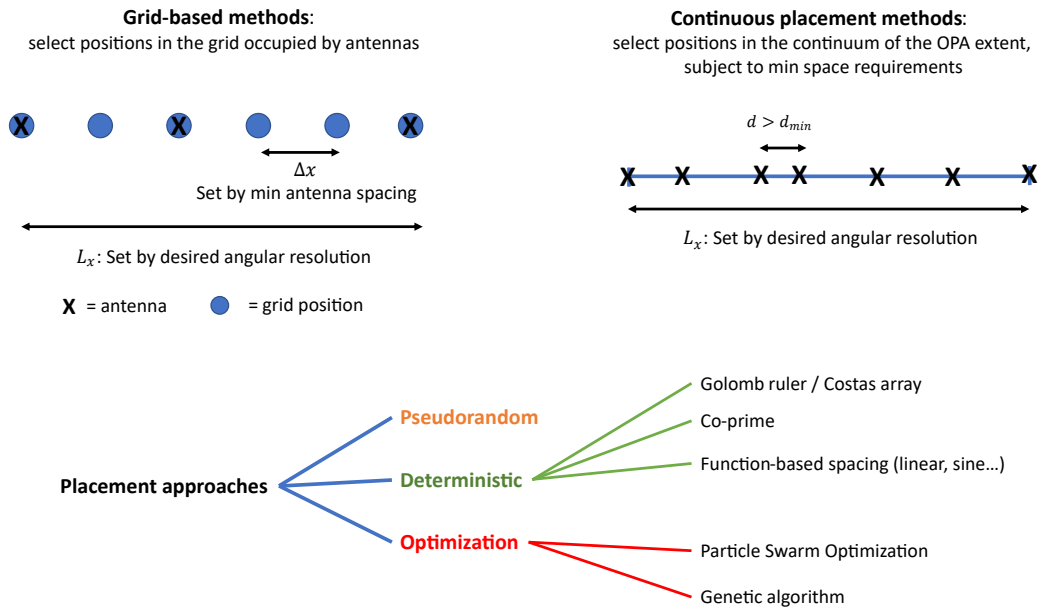


Figure 5-3: Non-uniform OPA emitter placement approaches. Both grid-based and continuous placement methods can be used, and different approaches can be used to determine the optimal position of the antennas.

pair of emitters is mutually distinct, i.e, no two elements have the same spacing [80]. This is a well known mathematical problem, and solutions have been found for both 1D (known as Golomb rulers [298]) and 2D arrays (with a popular set of solutions known as Costas arrays [299]).

It can be shown that non-redundant arrays maximize the flatness of the autocorrelation function of the array layout, which in turn maximizes the flatness of the autocorrelation function of the electric field generated by the OPA [80]. Since the OPA far field radiation pattern is the fourier transform of the electric field autocorrelation [80], a maximally flat autocorrelation ensures a maximally narrow radiation pattern, which is our goal to achieve a high angular resolution and FoV. As shown in [80], the number of resolvable points in non-redundant OPAs (defined as $\text{FoV}/(\text{ang. res})$) scales quadratically with the number of elements ($\propto N^2$), while it has a uniform scaling ($\propto N$) in uniformly spaced OPAs. Any other non-uniform placement approach will have a scaling between linear and quadratic with the number of elements ($\propto N^x, 1 < x < 2$).

- Pseudo-Random placement: Random placement of the emitters has also been studied to realize non-uniform OPAs [107, 295]. Notice how this is a similar approach to the non-redundant array, where the random placement tries to eliminate any correlation between the position of the emitters.

Since random placement can generate arrays with low performance radiation patterns, usually a pseudo-random approach is taken where the random placement is repeated until the simulated radiation pattern of the resulting array achieves a minimum performance (usually determined by the minimum acceptable SLSR) [107].

- Optimization algorithms: As compute power increases, the use of computational

algorithms to optimize the emitter positions becomes feasible. In this case, optimization algorithms are used to obtain an emitter placement that minimizes a given figure of merit (usually a combination of maximizing power in the main lobe and SLSR). In particular, both Genetic Algorithms (GA) [288,291,292] and Particle Swarm Optimization (PSO) [289] have been used for the realization of non-uniform OPAs.

Compared to deterministic placement methods, the use of optimization algorithms allows for more freedom in the design of the desired radiation pattern properties, for example allowing to tune the importance of different performance metrics. On the flip side, optimization algorithms are slow and computationally intensive, and there is in general no guarantee that the global optimum has been reached.

Of course, all these different approaches result in a different array layout, which in turn results in a different radiation pattern and therefore different system performance. In the next section we compare in detail the performance of all these non-uniform placement approaches, as well as the traditional uniformly spaced OPA, in the context of the advanced applications we introduced at the beginning of this chapter. To the best of our knowledge, a comparison between non-uniform placement approaches has only been partially done in [293] for deterministic and PSO-based optimization in a generic context without a target application.

5.2.1 Performance tradeoffs

The thinned array curse

As can be seen in the right-most panel of Fig. 5-2, non-uniform OPAs have lower power emitted in the main radiation lobe synthesized by the array. This is a well known result

in radio-frequency (RF) phased arrays, and is commonly referred to as the *thinned-array curse* or the *sparse array curse*. The term *thinned-array curse* is attributed to Robert Forward, who used it in unpublished Hughes Research Laboratories reports in 1976, and was first published in 1984 [300].

Basically, the thinned-array curse states that if we take an antenna array with N elements and increase the spacing between elements so that the effective aperture size increases (keeping the number of elements constant), the angular resolution of the beam will increase accordingly but the angular power density will keep constant, therefore resulting in a decrease in the total power contained in the main lobe.

Another way of stating the same concept is that if we take an array with N elements that has a main lobe power given by P_{main} and remove n elements, then the power in the main lobe for the new array is at most $p_{main} = P_{main}(N - n)/N$.

While this is a central result in non-uniform phased array theory, a good fraction of the published literature on non-uniform OPAs (Table 5.4) either does not mention the reduction in power or simply mentions it without delving into the consequences for OPA performance.

Figure 5-4 shows the power contained in the main lobe for a uniform OPA with $N_{uniform} = 100$ elements (dashed blue curve) and compares it with non-uniform 1D OPAs with increasing number of elements and the same total extent (and therefore the same angular resolution) as the uniform array. We can clearly see how non-uniform OPAs allow achieving a given angular resolution with a lower number of elements at the expense of reduced power in the main lobe.

It is worth noting that the power contained in the main lobe (a critical parameter that determines the sensitivity of our OPA-based system) is independent of the approach taken to determine the placement of the antennas in the non-uniform array and only depends on the total number of elements. This is an important observation, and is an observation

that can be misinterpreted when reading the published literature in non-uniform OPAs. The only effect of different antenna placements in a non-uniform OPA is in the achievable SLSR, but the power contained in the main lobe is unaffected. We will explore this in more detail in the next section.

The thinned-array curse sets an important limitation on the use of non-uniform OPAs for beam-steering applications. If we go back to the different application scenarios we described in Section 5.1.2:

1. Optical power limited applications cannot benefit from the use of non-uniform OPAs, as the total optical power emitted by the OPA is significantly decreased

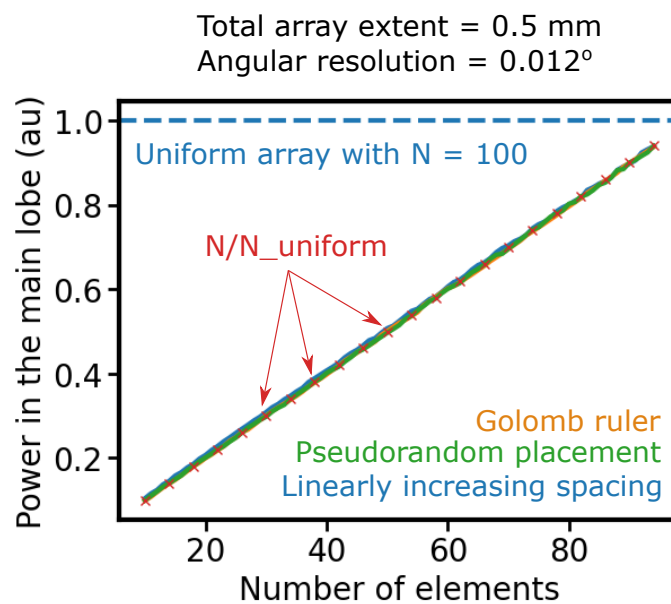


Figure 5-4: The thinned array curse. Power contained in the main lobe as a function of the number of elements for non-uniform OPAs with different placement approaches (linearly increasing spacing (blue), pseudo-random placement (green) and Golomb ruler based (orange)). The dashed line shows the power in the main lobe for a uniform array with $N_{uniform}=100$ elements, and the red crosses plot $N/N_{uniform} = N/100$. All the arrays have a 0.5 mm extent and a 0.012 ° angular resolution.

with respect to a uniform OPA.

2. In total (electrical + optical) power limited applications, non-uniform OPAs could have an advantage as they have decreased power dissipation. We will explore this in detail in Section 5.2.2.
3. In asymmetric link applications, non-uniform OPAs could be beneficial for the power-starved terminal, but are not suitable for the non-starved terminal as in this case we just want to maximize the emitted optical power.

SLSR of non-uniform OPAs

From our discussion above it is clear that the specific placement of the elements in a non-uniform OPA does not have a strong effect in the power emitted in the main lobe or in the angular resolution of the generated beam ⁴. Nevertheless, emitter placement is important in non-uniform OPAs to minimize the SLSR, which is of paramount importance to ensure that the system utilizing the OPA can correctly identify the main lobe and reduce the "noise" coming from light emitted into undesired angles (Fig. 5-5(a)).

To the best of the author's knowledge, while it is usually recognized that having a sufficiently high SLSR is a basic requirement for any OPA-based beam steering system, a detailed quantitative analysis of SLSR requirements for different applications does not exist in the literature. We will try to do so here for our applications of interest:

1. In the case of FSOC links, SLSR matters mostly because a higher SLSR generally translates into a larger power contained in the main lobe and therefore a longer link distance or a higher SNR.

⁴It is worth noting again that this fact is usually not explicitly stated in the published literature on non-uniform OPAs.

The presence of strong side lobes could be a concern, as it could cause the link tracking algorithm to take one of these strong side lobes as the main OPA lobe (Fig. 5-5(b)). Since such side lobe would have a smaller emitted power, this could compromise link performance.

We therefore need to make sure that the tracking algorithm used to establish and maintain alignment between transmitter and receiver accounts for the presence of such side lobes. In this case, there is not a strong requirement in terms of SLSR.

2. SLSR requirements become more important in LIDAR applications, as the presence of side lobes can significantly impact the achievable SNR.

This is depicted in Fig. 5-5(c): the presence of a strong side lobe can cause the system to detect the presence of a target in the direction of the main lobe (θ_0) if a high reflectivity target is located at the side lobe emission angle (θ_{sl}).

Since the reflectivity of common targets in LIDAR applications can range from 98% to about 0.5% [301], a SLSR on the order of 30 dB is required to be able to correctly detect a low reflectivity target in the presence of a high reflectivity target at the side lobe emission angle.

Minimizing the SLSR essentially requires eliminating strong constructive interference between the light emitted by different optical antennas at any angle that is not the desired angle of emission of the OPA. We can derive an upper bound for the achievable SLSR of an OPA with a given efficiency and angular resolution by assuming that all the power that is not emitted into the main lobe is evenly distributed amongst all other angles of emission ⁵. In this case:

⁵Note that this implies that there are no other grating lobes in the radiation pattern, or equivalently that the FoV of the OPA is 180° . If there are additional grating lobes (such as in a uniform OPA) then the achievable SLSR is larger as there is more power contained in the ensemble of grating lobes and therefore there is less power to distribute in the sidelobes.

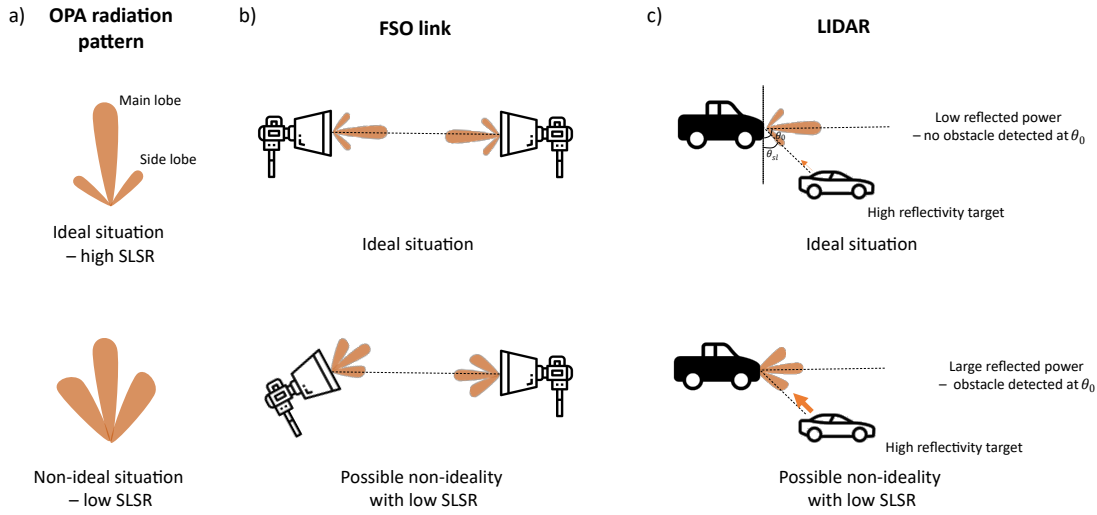


Figure 5-5: The effects of Side Lobe Suppression Ratio (SLSR). (a) Example radiation patterns with high SLSR (top) and low SLSR (bottom). (b) Low SLSR can cause an FSO link to be established between side lobes, which decreases range. (c) In LIDAR, a low SLSR can cause the detection of a target due to light reflected at other angles other than that of the main lobe.

$$P_{sidelobe} = 1 - P_{main\ lobe} \quad (5.2)$$

$$SLSR_{best} = 10 \log_{10} \left(\frac{P_{peak}}{P_{sidelobe}/(180 - \Delta\Theta)} \right) \quad (5.3)$$

Figure 5-6(a) shows such an upper bound for a non-uniform OPA with a total extent of $500 \mu\text{m}$ and an angular resolution of 0.156° as a function of the number of elements: regardless of the number of elements, the upper bound for the SLSR is close to 20 dB.

Achieving such an upper bound with a finite number of emitters is not possible, because the discreteness of the spacing between elements means that we cannot achieve an even distribution of power outside the main lobe. For instance, Fig. 5-6(a) also shows the SLSR achieved by a uniform OPA with $500 \mu\text{m}$ extent and a total of $N=100$ elements,

which is $SLSR_{uniform} = 13.26$ dB.

We can further develop an intuition on the challenges of maximizing SLSR by studying an OPA comprised of N elements, as depicted in Fig. 5-6(b), such that the spacing between elements i and k is d_{ik} . We know there will be constructive interference between the light emitted by two emitters in the OPA at an angle θ if $d_{ik}\sin(\theta) = m\lambda$, where m is an integer and λ is the operation wavelength. Therefore, to avoid constructive interference at any angle θ that is not the desired steering angle θ_{des} we would need to ensure:

$$d_{ik} \neq \frac{m\lambda}{\sin(\theta)}, \quad \forall \theta \neq \theta_{des}, \quad i \neq k \quad (5.4)$$

Figure 5-6(b) plots $m\lambda/\sin(\theta)$ for $m = 1..4$, along with arbitrarily chosen distances between elements in an $N = 3$ OPA. It is easy to see how even for $N = 3$ finding a set of distances that fulfill Eq. 5.4 (which is equivalent to ensuring that the crossing of the dotted lines with the solid lines happens at different angles θ) is challenging, especially at steep emission angles. It is therefore not trivial to synthesize a non-uniform OPA with high SLSR.

We studied the achievable SLSR of the different non-uniform OPA placement approaches described above (Fig. 5-3) as a function of number of elements, and the results are summarized in Fig. 5-6(c). All OPAs have a total extent of $500 \mu\text{m}$, corresponding to an angular resolution of 0.156° , a minimum spacing of $5 \mu\text{m}$ and an operation wavelength of 1550 nm . Several points are worth discussing:

- In general, the achievable SLSR depends on the angle at which we are steering the beam (see inset of Fig. 5-6(c)). In the main plot of Fig. 5-6(c) we show the average SLSR over a $\pm 80^\circ$ steering range, and the error bars show the maximum and minimum SLSR over the same steering ranges. Data points that have no error

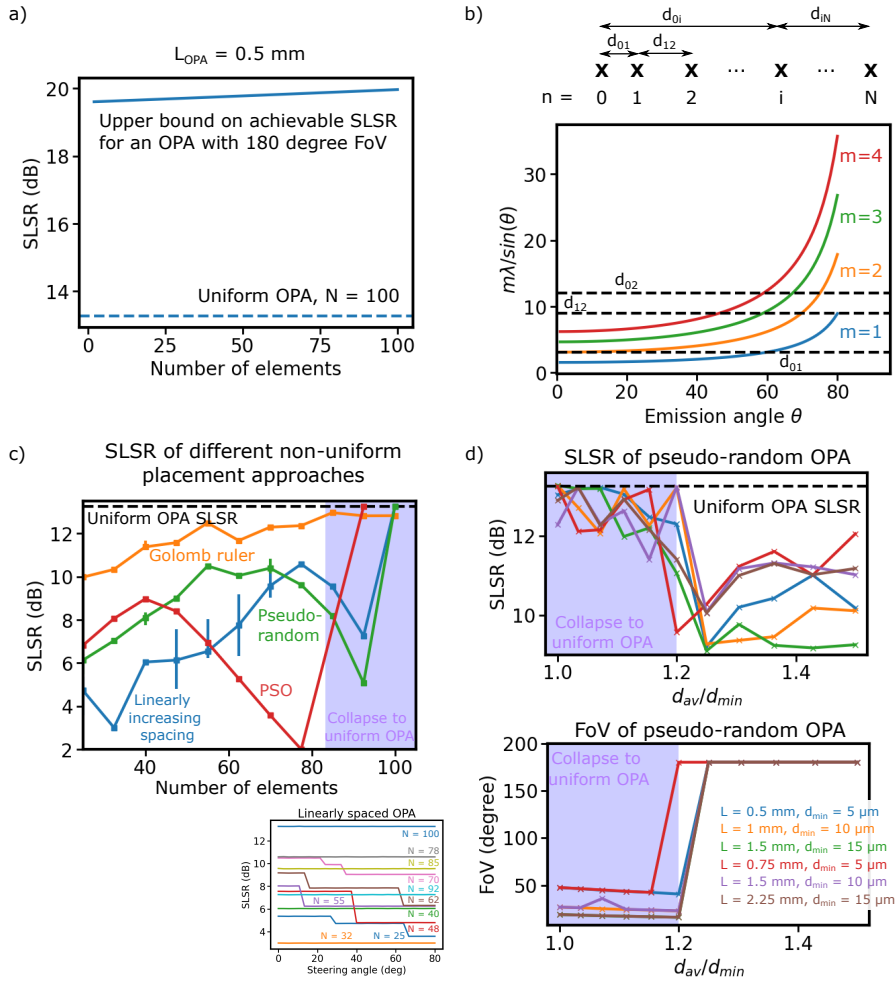


Figure 5-6: Side Lobe Suppression Ratio (SLSR) characteristics of non-uniform OPAs. (a) The upper bound on achievable SLSR (even distribution of power outside the main lobe) as a function of the number of elements of a non-uniform OPA. The SLSR of a uniform OPA with $N=100$ is also shown. (b) Minimum SLSR is achieved when $d_{ik} \neq m\lambda/\sin(\theta)$, $\forall \theta \neq \theta_{des}$, $i \neq k$. This is hard to achieve at steep angles even for $N = 3$, which is what is depicted. (c) SLSR of different non-uniform placement approaches. The minimum antenna spacing is $5 \mu\text{m}$ and the total OPA extent is 0.5 mm. Squares show the average SLSR over a $\pm 80^\circ$ steering range, and error bars show the maximum and minimum SLSR over the steering range. The inset shows the SLSR vs steering angle for a linearly increasing spacing OPA. (d) The SLSR (top) and FoV (bottom) of a pseudo-random OPA as a function of d_{av}/d_{min} . When $d_{av}/d_{min} < 1.2$ the non-uniform OPA collapses to a uniform OPA with the accompanied decrease in FoV.

bars have a constant SLSR with steering angle.

- For optimization-based placement approaches (pseudo-random and PSO) we used a relatively steep steering angle of $\theta_{des} = 60^\circ$ to optimize the placement. This results in a configuration that has smaller variations of SLSR with steering angle, something that has been recognized before [293].
- While PSO results in well-performing configurations for a relatively low number of elements ($N < 50$), it fails for large number of elements. This is because as the number of elements increases the search space of the optimization increases and the PSO algorithm cannot find a good optimum within the allocated number of iterations ($k = 100$ in our case) and population ($m = 10$ in our case).
- The achievable SLSR increases as the number of elements increases, and it converges to that of a uniform OPA (dashed blue line in Fig. 5-6(c)). This is because, for a fixed total extent of the OPA, as the number of elements increases the non-uniform OPA spacing becomes increasingly uniform (due to the limitation set by the minimum spacing between elements). Another way of looking at this is to realize that, for a fixed total extent L of an OPA, the average distance between elements is given by $d_{av} = L/N$. As N increases d_{av} decreases: in particular, as $N \rightarrow N_{unif}$, $d_{av} \rightarrow d_{min}$, at which point the OPA becomes uniform with an inter-element spacing given by d_{min} (the minimum allowed spacing).

This raises an important consideration when designing non-uniform OPAs: to maintain a large FoV with a nonuniform OPA we need to ensure that $d_{av} > d_{min}$. This ensures that the spacing between elements does not converge to the minimum spacing d_{min} , in which case we would have a uniform OPA (which we know results in a limited FoV due to full constructive interference occurring at other angles

besides the desired steering angle, Fig. 5-2).

Fig. 5-6(d) shows the SLSR and FoV of a pseudo-random OPA as a function of the ratio of the average distance to the minimum distance between elements d_{av}/d_{min} (therefore, increasing d_{av}/d_{min} is equivalent to a lower number of elements N) for OPAs with different number of elements and minimum antenna spacing. We can see how for $d_{av}/d_{min} > 1.2$ the non-uniform OPA maintains the full FoV with a large SLSR (albeit lower than that achievable with a uniform OPA). When $d_{av}/d_{min} < 1.2$ grating lobes appear which limit the FoV: essentially, the non-uniform OPA has collapsed into a uniform OPA with an interelement spacing equal to the minimum distance d_{min} . The conclusion here is that to ensure that a non-uniform OPA eliminates the presence of grating lobes we need to ensure that $d_{av}/d_{min} > 1.2$, or equivalently that $N_{nu} < 0.8N_u$.

- Golomb ruler placement seems to be the best performing non-uniform OPA, with the highest achievable SLSR and low variability over steering angle. This is expected, since as explained above Golomb rulers are constructed so as to ensure that $d_{ij} \neq d_{km}$ for $i \neq j \neq k \neq m$, which approaches the requirements to inhibit constructive interference between antennas (Eq. 5.4). Nevertheless, an important limitation of Golomb rulers is that there is a minimum total OPA length L_{min} that can be realized given a total number of elements and a minimum antenna spacing, as a Golomb ruler places the antennas in multiples of such a minimum spacing. For example, for a minimum antenna spacing $d_{min} = 5 \mu\text{m}$, a Golomb ruler with $N = 50$ has a minimum OPA extent of $L_{min} = 10.47 \text{ mm}$, and one with $N = 100$ has $L_{min} = 44 \text{ mm}$.

If we want a large number of elements and a small total OPA extent (which is incompatible with Golomb rulers) pseudo-random placement seems to be the best

performing, achieving a large SLSR across a wide range of element numbers and with small variability across the steering range. Deterministic non-uniform arrays with linearly increasing spacing also achieve good performance but show noticeable SLSR variability with steering angle.

It is important to notice that even for uniform OPAs it is not possible to achieve SLSR above 13 dB. While this is likely enough for FSOC applications, it is far from the 30 dB required for LIDAR applications, an issue that, to the best of the author's knowledge, is not discussed in the literature.

Two main approaches can be used to overcome such a limitation:

1. Use a receiver that is also angularly selective. For instance, using a uniform OPA at the transmitter (with 13 dB SLSR) and another uniform OPA at the receiver (with another 13 dB SLSR) gives a combined angular selectivity of 26 dB, close to the 30 dB required for LIDAR applications.

Nevertheless, such an approach increases system complexity significantly, as using an angularly selective receiver requires a means for the receiver to track the emission angle of the transmitter. Additionally, the signal captured by the receiver is weaker as power reflected at angles outside the receiver FoV is lost.

2. Increase the SLSR of the transmitter OPA by means of amplitude apodization. Until now we have only considered OPAs where all antennas emit the same optical power, which is the standard design in experimental OPA demonstrations.

Nevertheless, it is possible to increase the SLSR of both uniform and non-uniform OPAs by means of amplitude apodization, where different antennas emit a different power depending on their position.

For uniform OPAs, using a gaussian distribution of emitted power (Fig. 5-7(a)) can significantly increase the SLSR, as shown in Fig. 5-7(b). This can be understood by realizing that the far field radiation pattern of the OPA is related to the Fourier transform (FT) of the near field distribution emitted by the antennas. With this, having a uniform distribution of emitted power is equivalent to the application of a square windowing to the near field pattern, which causes the far field radiation pattern to be convoluted with a sinc function (blue curve in the inset of Fig. 5-7(a), since the FT of a square function is a sinc function). Instead, the use of a gaussian distribution of emitted powers is the same as applying a gaussian windowing to the near field, which causes the far field radiation pattern to be convoluted with a gaussian function (orange curve in the inset of Fig. 5-7(a), since the FT of a gaussian function is a gaussian function). The presence of side lobes in the sinc function (which do not exist in a gaussian function) cause a decrease in the SLSR. It is also important to mention that the increase in SLSR is independent of steering angle (inset of Fig. 5-7(b)).

As shown in Fig. 5-7(c), the increase in SLSR comes nevertheless accompanied by a decrease in the angular resolution. This is because, as shown in the inset of Fig. 5-7(a), the main lobe of a gaussian function decays more slowly than that of a sinc. There exists therefore a tradeoff between achievable SLSR and angular resolution, which can be tuned by increasing the sharpness of the gaussian apodization of the emitted power (as shown in the inset of Fig. 5-7(c)).

It is natural to consider if a similar approach can be used to increase the SLSR of non-uniform OPAs. The answer is yes, but with some additional considerations and limitations, which we illustrate in Figs. 5-7(d,e,f) for a linearly increasing spacing non-uniform OPA.

Firstly, it is not enough to apply simple gaussian apodization of the emitted power, but it is necessary to find the optimum amplitude distribution using an optimization framework. We use the approach described in [302, 303], which translates the problem of finding the optimum power distribution to a Second Order Cone Program (SOCP), which can be efficiently solved by convex optimizers. Solving such an optimization problem gives the optimum power distribution to maximize SLSR. An example of the optimum power distribution is shown in Fig. 5-7(d) and the resulting radiation pattern is shown in Fig. 5-7(e).

Second, the improvement in SLSR is not as large as in the case of uniform OPAs. As can be seen in Fig. 5-7(e), the improvement in SLSR is limited to about 3.5 dB (the exact improvement depends on the non-uniform placement approach, as well as the antenna spacing and total extent of the OPA). This is because, as discussed above, the non-uniform spacing between elements makes it very challenging to avoid constructive interference at any angle outside the main lobe. Having amplitude control can help make such constructive interference smaller, but cannot eliminate it (since we would need phase control for that, but the phase is set by the required steering angle).

As shown in Fig. 5-7(f), it is possible to maintain the gain in SLSR over the whole steering range, but for that to happen we need to perform the optimization at a large steering angle (vertical dashed line in Fig. 5-7(f)).

An important observation to make here is that, both in the case of uniform and non-uniform OPAs, we have considered the use of a fixed amplitude apodization, which can be easily implemented by the use of non 50-50 splitters in the power distribution network without increasing the complexity of the OPA control.

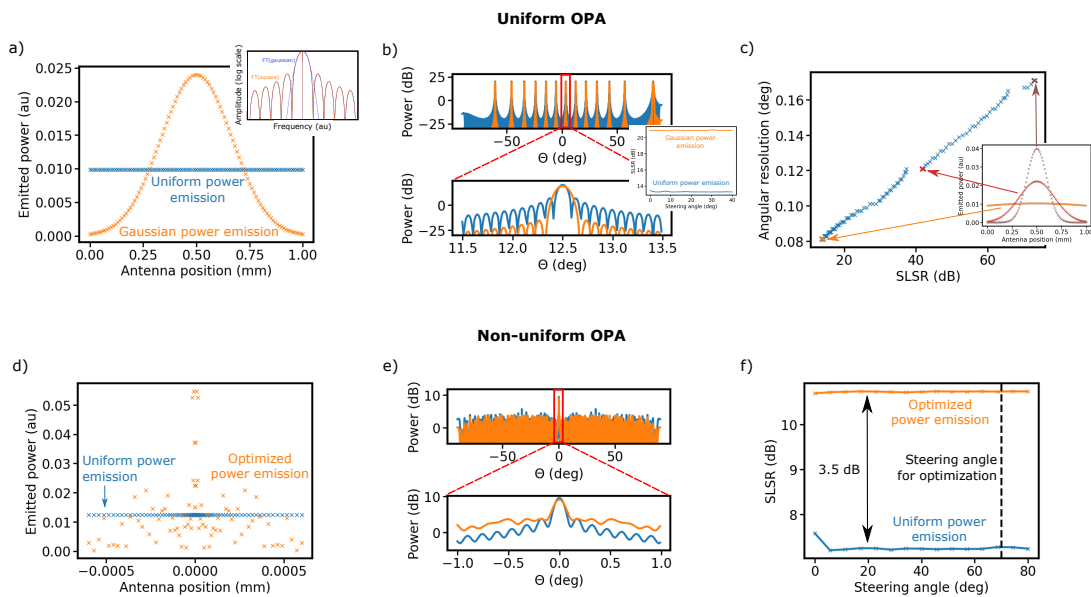


Figure 5-7: Increasing SLSR by amplitude apodization. (a) Power emitted by each antenna in a 1D uniform OPA for uniform (blue) and gaussian (orange) distributions. The inset shows the Fourier Transform (FT) of a square (blue) and a gaussian (blue) window. (b) Radiation pattern of a uniform OPA with uniform antenna power emission (blue) and gaussian distribution. The inset shows the SLSR as a function of steering angle. (c) Angular resolution as a function of achieved SLSR for gaussian power apodization. The inset shows some sample antenna emitted power distributions. (d) Optimum power distribution (orange) and uniform power distribution (blue) of a linearly increasing non-uniform OPA. (e) The radiation pattern of the non-uniform OPA with uniform power distribution (blue) and optimized power distribution (orange). (f) The SLSR as a function of steering angle for uniform (blue) and optimized (orange) power distributions. The dashed black line shows the steering angle at which the optimization was performed.

Non-uniform OPA performance - key takeaways

It is worth summarizing the conclusions we have drawn in this section:

1. The thinned array curse: no matter what placement approach we use, the power emitted into the main lobe by a non-uniform OPA will always be lower than that of a uniform array with the same total extent due to the lower number of emitters. In particular, we can say $P_{nonuniform} = P_{uniform} N_{nonuniform} / N_{uniform}$. This deems the use of non-uniform arrays not suitable for optical power limited applications such as long range FSOC.
2. Different non-uniform placement approaches result in different SLSR levels. Golomb ruler OPAs result in the highest achievable SLSR for a given number of elements, but have a limit in the minimum achievable OPA length. For applications requiring smaller OPA lengths, pseudo-random placement results in the highest SLSR with low variability with desired steering angle.

While optimizing the power emitted by each antenna can increase the achievable SLSR, non-uniform OPAs cannot achieve the 30 dB of SLSR required for automotive LIDAR applications. In this case, an angularly selective receiver needs to be used to achieve enough angular discrimination.

Another point worth mentioning is how, for OPAs with a large number of emitters, removing a small number of elements does not have a large impact in OPA performance - the angular resolution stays the same, the SLSR will slightly decrease and the power emitted into the main lobe will decrease by a factor $n_{removed}/N$ (which is a small number if N is large and $n_{removed}$ is small). Removing some emitters in a uniform OPA can be useful to accommodate for the necessary routing to bring the light into each individual

emitter [161], especially for OPAs with a large N that require a large number of routing optical waveguides.

5.2.2 System-level study

In the previous section we have thoroughly analyzed the tradeoffs of non-uniform OPAs from the point of view of their performance, which can be summarized as a decrease in the number of elements required to achieve a given angular resolution that is accompanied by a decrease in the power emitted in the main lobe. In this section we extend our analysis and quantitatively evaluate how the use of non-uniform OPAs compares to uniform OPAs from a system-level perspective.

We do so from three different angles:

1. Total optical power emitted in the main lobe, which is important to ensure that the beam-steering system has enough *link budget* to detect the emitted signal with a high enough SNR and therefore operate reliably.
2. Total power consumption (electrical + optical), which is important to (1) minimize operating expenses or maximize time of operation in case of battery-powered systems and (2) guarantee that the whole system in general, and the photonic chip in particular, operate at safe temperatures.
3. Area, packaging and control complexity, which is important to minimize cost and ensure scalability.

Optical power emitted in the main lobe

As we have discussed in detail in the previous section, non-uniform OPAs suffer from the thinned array curse, and therefore will always have a lower power emitted in the main

lobe compared to a uniform OPA with the same extent.

We can model the output optical power in the main lobe (denoted by the subscript ml) of an OPA (uniform or non-uniform) with:

$$P_{opt,out,ml} = \eta_{rad} \eta_{GC} \eta_{ps} \eta_{power_{distr}} \eta_{routing} P_{opt,in} \quad (5.5)$$

Above, η_{rad} accounts for the radiation efficiency into the main lobe of the OPA, η_{GC} is the radiation efficiency of the emitting antenna (usually a grating coupler), η_{ps} accounts for the optical loss of the phase shifter controlling the phase emitted by each antenna, $\eta_{power_{distr}}$ accounts for the loss of the optical power distribution network (the ensemble of splitters that divide the input optical power in equal parts for each emitting antenna), $\eta_{routing}$ accounts for the loss of the routing waveguides that deliver the light to each emitting antenna and $P_{opt,in}$ is the input optical power. Note how the total optical power emitted by the OPA integrated over the whole angular emission range is given by $P_{opt,out,total} = P_{opt,out,ml}/\eta_{rad}$.

We can approximate that η_{GC} , η_{ps} and $\eta_{routing}$ are the same for both uniform and non-uniform OPAs⁶. Therefore, we can compare the optical power emitted into the main lobe by a uniform (denoted by the sub-index u) and non-uniform OPA (denoted by the sub-index nu) with:

$$\frac{P_{opt,out,ml,u}}{P_{opt,out,ml,nu}} = \frac{\eta_{rad,u} \eta_{power_{distr},u}}{\eta_{rad,nu} \eta_{power_{distr},nu}} \quad (5.6)$$

As we have seen in the previous section, due to the thinned array curse the ratio of radiation efficiencies is simply the ratio of the number of elements $\eta_{rad,u}/\eta_{rad,nu} = N_u/N_{nu}$.

⁶ $\eta_{routing}$ would likely be higher for a uniform OPA as a larger number of elements requires more complex routing to avoid crossing other waveguides.

The power distribution efficiency $\eta_{power\,distr}$ is different for uniform and non-uniform OPAs due to the different number of elements of each approach, which then requires a different total splitting ratio of the incoming power. Here we will consider the use of nested 1x2 power splitters to achieve the required power distribution, which is the standard approach in experimental OPA demonstrations ⁷. In this case, the required number of 1x2 splitters can be calculated as $n_{splitters} = \lceil \log_2 N \rceil$. If we denote the optical transmission of a single splitter as η_{1x2} we then get to:

$$\frac{P_{opt,out,ml,u}}{P_{opt,out,ml,nu}} = \frac{N_u \eta_{1x2}^{\lceil \log_2 N_u \rceil}}{N_{nu} \eta_{1x2}^{\lceil \log_2 N_{nu} \rceil}} = \frac{N_u}{N_{nu}} \eta_{1x2}^{\lceil \log_2 N_u \rceil - \lceil \log_2 N_{nu} \rceil} \quad (5.7)$$

Figure 5-8(a) shows the ratio of emitted powers according to Eq. 5.7 for different losses of the 1x2 splitter (typical losses for 1x2 splitters in photonics processes are below 0.2 dB). As expected, the uniform OPA has a larger output optical power. We can also see how the ratio of emitted powers is mainly set by the ratio of the number of elements (dashed black line in Fig. 5-8(a)), and that the 1x2 splitter loss has a small effect.

It is interesting to translate the decrease in emitted optical power to a decrease in the maximum distance that our beam-steering system can operate at. For our advanced application scenarios, this would translate to the maximum distance at which an object can be detected for the LIDAR application, and the maximum link distance for the FSO application. As mentioned in Section 5.1.2, we will assume a free space propagation model where the losses are proportional to L^2 (where L is the link distance), which means that the maximum achievable link distance for a given receiver sensitivity S_{rec} is $L \propto \sqrt{P_{out,ml,OPA}/S_{rec}}$.

Here we need to make a distinction between two possible scenarios:

⁷Other approaches are possible, such as the use of star couplers or 1xN splitters with $N > 2$, but these tend to have higher insertion loss, and in the case of star couplers achieving large splitting ratios is very challenging.

- Scenario 1: The receiver system is not an OPA. This is most likely the case in asymmetric links and in LIDAR applications, where the signal reflected from the target is measured by a high sensitivity photodetector array adjacent to the OPA, such as a Single Photon Avalanche Detector (SPAD) array.

In this case we can assume that the sensitivity of the receiver is the same regardless of if we are using a uniform or non-uniform OPA at the transmitter. In this case, we then have $L_u/L_{nu} = \sqrt{P_{opt,out,ml,u}/P_{opt,out,ml,nu}}$. This is shown in the top of Fig. 5-8(b). We can see how a non-uniform OPA with half the number of elements

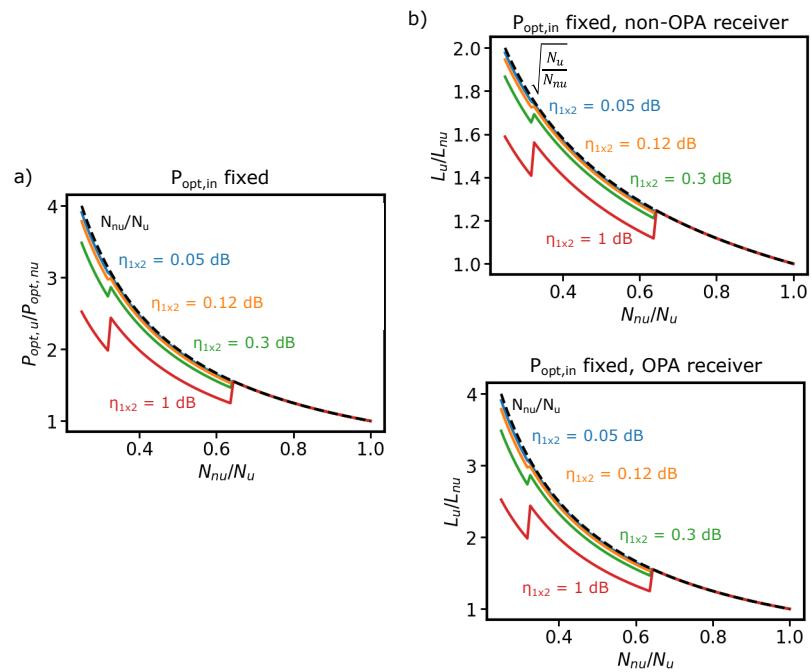


Figure 5-8: System-level optical output power considerations for non-uniform OPAs. (a) The ratio of optical output power and (b) the ratio of achievable link distance as a function of the ratio between number of elements of non-uniform OPAs for different losses of the 1x2 power splitter component. The input optical power to the OPA is the same for the uniform and non-uniform OPAs. The top plot of (b) assumes that the receiver is not OPA based (typical for LIDAR), while the bottom plot of (b) assumes the receiver is OPA based as well (typical for FSOC).

as a uniform OPA has 70% of the reach.

- Scenario 2: The receiver system is also an OPA. This is the typical configuration in a symmetric FSOC link, where both the transmitter and receiver need to be able to steer the beam. In this case we need to account for the loss in sensitivity of the receiver if we are using a non-uniform OPA, which is exactly the same as the reduction in the power emitted in the main lobe when using the OPA as a transmitter (η_{rad}). Therefore, $S_{nu} = S_u \frac{P_{opt,out,ml,nu}}{P_{opt,out,ml,u}}$, and then $L_u/L_{nu} = P_{opt,out,ml,u}/P_{opt,out,ml,nu}$. As shown in the bottom of Fig. 5-8(b), in this scenario a non-uniform OPA with half the number of elements as a uniform OPA has 50% of the reach.

This reinforces our conclusion in the discussion of the thinned array curse: non-uniform OPAs are not a suitable solution in systems where the only concern is the maximization of the link budget (or equivalently the maximization of the link distance).

Albeit straightforward, it is worth mentioning that we can achieve the same output optical power in the main lobe with a non-uniform OPA ($P_{opt,out,ml,nu}$) as in a uniform OPA ($P_{opt,out,ml,u}$) by simply increasing the input optical power to the OPA $P_{opt,in}$ (Eq.5.5). Two limitations are relevant here:

1. Eye safety considerations set a limit to the optical power emitted by an OPA system, which at a wavelength of 1550 nm needs to be below 21.3 dBm (134.9 mW) to be considered class 1M (and therefore safe for outdoor use without additional restrictions).
2. The maximum input optical power that our OPA system can sustain before damage to the waveguides, which is on the order of 1 W if we use a fabrication platform that includes SiN.

Both considerations set a limit in the maximum achievable output optical power integrated over the whole angular range of emission of the OPA system:

$$P_{opt,out,total} = \min(P_{out\ total\ max,eye\ safe}, P_{in\ max} - IL_{chip}) \quad (5.8)$$

Therefore, increasing the input optical power to a non-uniform OPA is only possible if the input optical power is below the waveguide damage threshold and the total output power is below the eye safety limit.

Total power consumption

Output optical power is usually not the only performance specification of a beam-steering system. Total power consumption is of paramount importance to ensure reliable and scalable operation of the system, as it affects:

- The operating expenses of the system, as the operator needs to cover the cost of the required power.
- In the case of battery-operated systems (such as vehicles in the case of LIDAR), the power consumption sets a limitation in the time between charges (for electrical vehicles) or between refuels (for gas vehicles).
- There is a limitation in the maximum power that can be dissipated in the optical chip before it becomes mechanically unstable due to the accompanying temperature increase. Ensuring that such a limit is not reached is thus essential for the correct operation of the beam-steering system.

There are two main sources of power consumption and dissipation in an OPA-based beam steering system, schematically depicted in Fig. 5-9:

1. Optical power consumption: Comes from absorption and scattering losses as the light propagates in the OPA chip, as well as any coupling losses when delivering the light generated by the laser source (which is usually external to the OPA chip) to the OPA chip.

We can quantify the optical losses in the OPA (Fig. 5-9) with:

$$\eta_{opt} = \eta_{in_coup} \eta_{rad} \eta_{GC} \eta_{ps} \eta_{power_distr} \quad (5.9)$$

The first term above (η_{in_coup}) accounts for the coupling loss from the laser into the chip, and the rest account for losses within the OPA chip: η_{rad} is the radiation efficiency into the main lobe of the OPA, η_{GC} the antenna loss, η_{ps} the insertion loss of a single phase shifter and η_{power_distr} is the loss acquired along the routing

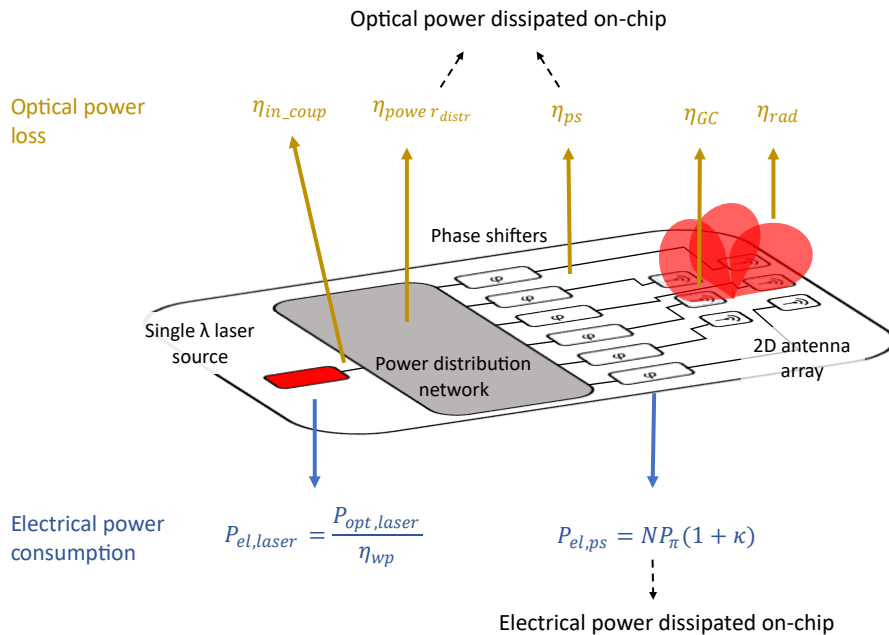


Figure 5-9: OPA optical and electrical power consumption, dissipation and loss breakdown.

network that splits and delivers the light to each antenna.

A total optical power $P_{opt,laser}\eta_{opt}$ is lost in the OPA system, but only a part of it is actually dissipated in the OPA chip:

$$P_{opt,diss} = P_{opt,laser} \eta_{ps} \eta_{power_{distr}} \quad (5.10)$$

Neither the grating coupler efficiency η_{GC} , the main lobe radiation efficiency η_{rad} nor the input coupling efficiency η_{in_coup} appear, as these loss mechanisms radiate the lost power into free space instead of it being absorbed in the chip.

2. Electrical power consumption and dissipation, which comes mainly from:

- The electrical power needed to drive the laser delivering the light to the OPA chip. We will model such power consumption as $P_{el,laser} = P_{opt,laser}/\eta_{wp}$, where η_{wp} is the wall plug efficiency of the laser (usually around 10% for typical Distributed Feedback Lasers (DFB) [304] or External Cavity Lasers (ECL) [305]), and $P_{opt,laser}$ is the output optical power of the laser.
- The power needed to drive the phase shifters controlling the phase emitted by each optical antenna. Here we not only need to account for the power consumption of the phase shifter itself, but also the power consumption of the element that is driving such phase shifter, which is usually a Digital to Analog Converter (DAC).

We will model the electrical power consumption due to the phase shifters as $P_{el,ps} = NP_{\pi}(1 + \kappa)$. Here, N is the number of antennas, P_{π} is the electrical power required to achieve a π phase shift, and we model the DAC power consumption as a fraction of P_{π} (i.e, the power consumption of a single DAC is κP_{π}).

The value of κ depends on the type of DAC used, the need to deliver either current (in case of a heater phase shifter) or voltage (depletion-based phase shifter) and the transistor characteristics. Using back of the envelope calculations and sensible values for all these parameters, a value $\kappa = 0.4$ seems reasonable both for linear and switching DACs.

It is important to notice that there are other sources of electrical power consumption in our OPA-based beam steering system which we will not consider in our analysis. Examples include the power required to operate the control algorithm that determines the phase shift to be applied to each element, as well as power consumption of any active cooling element present in the system.

It is also worth emphasizing the difference between power consumption and power dissipation. Power consumption is the total power that needs to be delivered to the system to operate (which in our model is $P_{cons} = P_{el}$). Power dissipation is the total power dissipated by the system, which includes both the electrical power consumption and the optical losses ($P_{diss} = P_{el} + P_{opt,diss}$).

We can thus summarize our power consumption and dissipation model with:

$$P_{opt,out,ml} = P_{opt,laser} \eta_{opt} = P_{opt,laser} \eta_{in_coup} \eta_{rad} \eta_{GC} \eta_{ps} \eta_{power_distr} \quad (5.11)$$

$$P_{opt,diss} = P_{opt,laser} \eta_{ps} \eta_{power_distr} \quad (5.12)$$

$$P_{el} = P_{el,laser} + P_{el,ps} = \frac{P_{opt,laser}}{\eta_{wp}} + NP_{\pi}(1 + \kappa) \quad (5.13)$$

$P_{opt,out,ml}$ is the emitted output optical power in the main lobe, $P_{opt,diss OPA}$ is the optical power dissipated in the OPA chip, and P_{el} is the total electrical power consumption. For our analysis, we will use a wall plug efficiency $\eta_{wp} = 10\%$, and a DAC driver efficiency $\kappa = 0.4$.

For the sake of completion, it is important to mention that the electrical power dissipation in the OPA chip is given by $P_{el,diss OPA,mono} = NP_{\pi}(1 + \kappa)$ in a monolithic architecture where the phase shifter drivers are integrated in the OPA chip, and by $P_{el,diss OPA,hybrid} = NP_{\pi}$ in a hybrid architecture where the phase shifter drivers are in a different chip.

As we discussed in Section 5.1.2, a suitable FoM in a system that has constraints in total power dissipation is:

$$FoM = \frac{L}{P_{el}} \quad (5.14)$$

In particular, if we want to compare the value of such FoM for uniform and non-uniform OPAs, we have:

$$\frac{FoM_{nu}}{FoM_u} = \frac{L_{nu}/P_{el,nu}}{L_u/P_{el,u}} = \frac{L_{nu}}{L_u} \frac{P_{el,u}}{P_{el,nu}} \quad (5.15)$$

Using the above derived equations, we get for scenario 1 (OPA transmitter to non-OPA receiver system):

$$\begin{aligned}
\frac{FOM_{nu}}{FOM_u} \Big|_{\text{scenario 1}} &= \sqrt{\frac{P_{opt,out,ml,nu} \frac{P_{opt,laser}}{\eta_{wp}} + N_u P_\pi (1 + \kappa)}{P_{opt,out,ml,u} \frac{P_{opt,laser}}{\eta_{wp}} + N_{nu} P_\pi (1 + \kappa)}} = \\
&= \sqrt{\frac{\eta_{rad,nu} \frac{P_{opt,laser}}{\eta_{wp}} + N_u P_\pi (1 + \kappa)}{\eta_{rad,u} \frac{P_{opt,laser}}{\eta_{wp}} + N_{nu} P_\pi (1 + \kappa)}} = \\
&= \sqrt{\frac{N_{nu} \frac{P_{opt,laser}}{\eta_{wp}} + N_u P_\pi (1 + \kappa)}{N_u \frac{P_{opt,laser}}{\eta_{wp}} + N_{nu} P_\pi (1 + \kappa)}}
\end{aligned} \tag{5.16}$$

And similarly for scenario 2 (OPA transmitter and OPA receiver):

$$\frac{FOM_{nu}}{FOM_u} \Big|_{\text{scenario 2}} = \frac{N_{nu} \frac{P_{opt,laser}}{\eta_{wp}} + N_u P_\pi (1 + \kappa)}{N_u \frac{P_{opt,laser}}{\eta_{wp}} + N_{nu} P_\pi (1 + \kappa)} \tag{5.17}$$

Where, in both Eqs. 5.16 and 5.17, in the last equality we have ignored the effect of the additional 1x2 coupler loss (which we have seen is negligible, Fig. 5-8).

Equations 5.16 and 5.17 are relatively complicated, but two simple limit cases can be derived:

- Phase shifter power limited: when $P_{opt,laser}/\eta_{wp} \ll NP_\pi$, we can simplify the expression for Scenario 1 to:

$$\frac{FOM_{nu}}{FOM_u} \Big|_{\text{scenario 1, ps limit}} = \sqrt{\frac{N_u}{N_{nu}}} \tag{5.18}$$

Since $N_{nu} < N_u$, in this limit the use of a non-uniform OPA is always better in terms of link distance achievable for a given electrical power budget.

For Scenario 2:

$$\frac{FOM_{nu}}{FOM_u} \Big|_{\text{scenario 2, ps limit}} = 1 \quad (5.19)$$

Which means that both approaches have the same performance in terms of link distance achievable for a given electrical power budget.

- Laser power limited: in the opposite case where $P_{opt,laser}/\eta_{wp} \gg NP_{\pi}$, we then have for Scenario 1:

$$\frac{FOM_{nu}}{FOM_u} \Big|_{\text{scenario 1, laser limit}} = \sqrt{\frac{N_{nu}}{N_u}} \quad (5.20)$$

Which is the inverse of the phase shifter limit, so the use of a uniform OPA becomes more advantageous.

For Scenario 2:

$$\frac{FOM_{nu}}{FOM_u} \Big|_{\text{scenario 2, laser limit}} = \frac{N_{nu}}{N_u} \quad (5.21)$$

Which also results in uniform OPAs being advantageous in terms of link distance per unit of electrical power consumption.

It is clear from our analysis that, in terms of link distance achievable for a given power consumption, uniform OPAs are always better performing in Scenario 2: a symmetric link where both the emitter and receiver in the beam-steering system are based on OPAs.

The situation becomes more interesting in Scenario 1, where the emitter is OPA-based but the receiver is not. In this case, the use of non-uniform OPAs becomes attractive as we approach the phase shifter limit. Of course, which limit are we closer to will depend on the required laser power $P_{opt,laser}$ and the power consumption of the phase

shifter ensemble $P_\pi N$. To study this, it is useful to define the ratio of electrical power consumption of the uniform OPA phase shifters to the electrical power consumption of the laser:

$$A = \frac{N_u P_\pi (1 + \kappa)}{P_{opt,laser} / \eta_{wp}} \quad (5.22)$$

And we can then rewrite Eq 5.16 as:

$$\left. \frac{FOM_{nu}}{FOM_u} \right|_{scenario\ 1} = \sqrt{\frac{N_{nu}}{N_u} \frac{1 + A}{1 + A \frac{N_{nu}}{N_u}}} \quad (5.23)$$

Fig. 5-10(a) shows the equation above for different values of A and fraction of non-uniform OPA elements, as well as the two limits described above. We can see how for $A > 4$ (which is equivalent to saying that the electrical power consumption of the phase shifters is $4\times$ larger than that of the laser), non-uniform OPAs have an advantage in terms of link distance per unit power.

How does this situation translate to actual power numbers? Fig. 5-10(b) shows the phase shifter P_π required for non-uniform OPAs to be advantageous compared to uniform OPAs as a function of the required number of elements in a uniform OPA. The use of phase shifters with P_π above the line make non-uniform OPAs a better choice in terms of link distance per electrical power spent. We can see how even for the lowest performance application we are considering (viable LIDAR), non-uniform OPAs are better if we use phase shifters with $P_\pi > 10\ mW$ (a good amount of thermal phase shifters have P_π above this number, Table A.1). As the number of elements increases, $P_\pi < 100\ \mu W$ are required for uniform OPAs to be advantageous, a number that can be hard to achieve even for depletion-type phase shifters if we account for their dynamic power dissipation.

We can thus conclude that, in most practical applications requiring a large number of elements with asymmetric links and non-OPA based receiver links (Scenario 1), the use

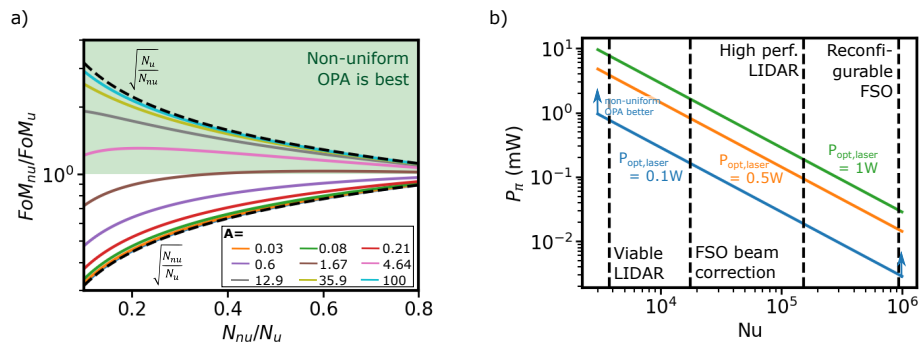


Figure 5-10: System-level electrical power consumption considerations for non-uniform OPAs in Scenario 1 (non-OPA based receiver). (a) Ratio of FoM for a non-uniform OPA over the FoM for a uniform OPA as a function of the fraction of elements in the non-uniform OPA. Different values of the parameter A (defined in the main text) are shown. We can see how for $A > 4$ the use of non-uniform OPAs is better regardless of the number of elements. (b) The phase shifter P_π at which the use of non-uniform OPAs is advantageous as a function of the number of elements in the equivalent uniform OPA for different laser powers. If we use phase shifters with P_π above the line, then non-uniform OPAs are better. Representative element numbers for the applications we have discussed are shown.

of non-uniform OPAs results in a better achievable link distance per amount of electrical power delivered to the system. Nevertheless, it is important to remember here that, as discussed in the previous section, uniform OPAs can achieve a larger total link distance than non-uniform OPAs (at the cost of larger electrical power dissipation) as they can achieve a larger power emitted in the main lobe for the same input optical power due to the thinned array curse.

There is another situation we can encounter when considering the power consumption and dissipation of our OPA, and is that where we are limited by the maximum permissible power dissipation in the chip. There is a limit on how much power can be dissipated in a chip before its temperature starts increasing beyond acceptable levels. Such limits depend on the cooling system being used [306]: passive air cooling can sustain up to 1 kW/m² power dissipation, while active air cooling (e.g, through a fan) can sustain up to 10 kW/m² and liquid cooling can reach 1,000 kW/m².

Assuming a maximum reticle size of 3 cm x 3 cm (limited by the FoV of lithography tools), this translates into a maximum on-chip power dissipation of 0.9 W for passive air cooling and 9 W for active air cooling.

If the limiting factor for the realization of a uniform OPA is on-chip power dissipation, the switch to a non-uniform OPA is basically the only alternative - a reduction in the number of elements will allow for a reduction of the electrical power dissipation in the OPA chip by the same amount.

Area, packaging and control complexity

In the previous section we have discussed how the use of non-uniform OPAs can yield a larger link distance or a better SNR in scenarios where power consumption and/or power dissipation are limiting.

Furthermore, non-uniform OPAs can relax the required area per phase shifter. As we discussed in Section 5.1.1, limitations in the total area of the OPA chip impose a limit in the size that can be occupied by each phase shifter. From Table 5.3, phase shifters with areas as small as $290 \mu m^2$ ($17 \mu m \times 17 \mu m$) are required for our advanced applications if we allocate 30% of the total chip area to the phase shifter network. To the best of our knowledge, no integrated phase shifter reported in the literature is even close to delivering the necessary performance specifications in such a small form factor (Fig. A-2).

The use of non-uniform OPAs results in relaxed requirements for phase shifter area - a lower number of emitters results in a lower number of phase shifters, which therefore allows to increase the area occupied by a single phase shifter by a factor $A_{ps,nu} = A_{ps,u} N_u / N_{nu}$.

Another important consideration is packaging and IO bump density in the chip. In non-monolithic OPAs, where the DAC driving each phase shifter is in a separate chip, at least one IO bump is required per phase shifter assuming a shared ground connection⁸. As the number of phase shifters increases so does the bump density required to escape all the signals. Achievable solder bump pitches are above $130 \mu m$, while copper pillars allow for a decrease down to $30 \mu m$ when using a staggered configuration [307, 308]. Figure 5-11 shows the necessary solder bump density as a function of the number of elements in the OPA assuming we can use 50% of the chip area for the solder bumps associated to the phase shifter network and that we require one bump per phase shifter. We can see how even high density copper pillars are not enough for the most demanding OPA applications. In this case, we can use non-uniform OPAs to reduce the number of phase shifters and therefore the required bump density.

⁸Notice this is the most *optimistic* approximation. Specially for phase shifters with current consumption (thermal or forward biased pn junctions) a single shared ground will not work as there will be a very large amount of current in said ground, which will cause electromigration issues. In general, more than one bump per phase shifter is required in non-monolithic OPAs.

It is worth noting here that the use of monolithic architectures where the DACs driving the phase shifters are integrated in the same chip greatly relaxes the requirements in solder bump density, as in this situation we can time-multiplex the addressing of each phase shifter through the use of scan chains or similar approaches.

Yet another aspect where the use of non-uniform OPAs can be advantageous is in the complexity of the control and calibration of the phase shifters in the OPA. Any practical OPA requires at least a calibration step to initially align the phases at each antenna, which are not uniform even if all the phase shifters are path length matched due to fabrication variations and thermal gradients within the chip, which cause changes in the effective refractive index of the optical mode propagating along the waveguide [309]. In some applications, continuous optimization of the phase shifters is required, for example in FSOC applications where we are trying to correct for atmospheric turbulence effects [310].

A variety of optimization methods can be used to determine the best phase for each shifter, but all of them require, in general, more iterations and/or time to converge as the number of phase shifters increases. This can be problematic if the convergence time approaches the time scale at which the fluctuations we are trying to compensate for occur. It is clear then how non-uniform OPAs can achieve faster convergence as they reduce the number of phase shifters to optimize.

5.3 Conclusion

In this chapter we have first presented two advanced OPA-based beam steering applications - high performance LIDAR and reconfigurable FSOC networks - and derived performance requirements.

The use of uniform OPAs for these applications is highly challenging due to the very large number of elements required ($N > 150,000$). In front of this, we have introduced

the possibility of using non-uniformly spaced OPAs to relax the requirement on the number of elements.

We have analyzed in detail the performance tradeoffs of non-uniform OPAs and shown how the thinned array curse prevents non-uniform OPAs from achieving the same output optical power in the main lobe as uniform OPAs of the same extent. As a consequence, non-uniform OPAs are not suitable for applications that are optical power limited (such as long distance FSOC).

We have then analyzed the total power consumption of non-uniform OPAs and compared it to uniform OPAs. We have seen how in the case of an OPA - OPA link non-uniform OPAs do not present an advantage with respect to uniform OPAs. On the contrary, in the case of an OPA transmitter - non-OPA receiver link (common situation in asymmetric links and in most LIDAR architectures), we have seen how non-uniform OPAs result in a better efficiency in terms of link distance per unit of electrical power spent as long as $A = \frac{N_u P_\pi (1+\kappa)}{P_{opt, laser} / \eta_{wp}} > 4$, a situation which is in general true for the advanced applications we presented. We have also seen how non-uniform OPAs are a good

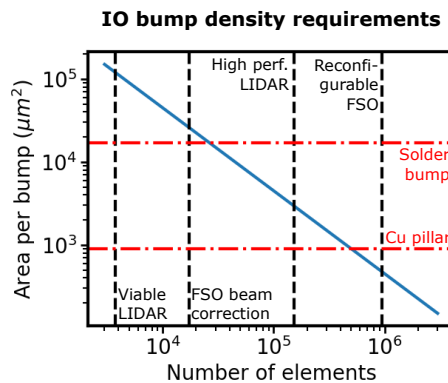


Figure 5-11: Bump density requirements for OPAs as a function of the number of elements. We consider that 50% of a 3x3 cm chip can be occupied by solder bumps for the phase shifters. Horizontal lines show limits for solder bumps and copper pillars, and vertical lines show number of elements for the OPA applications we have discussed.

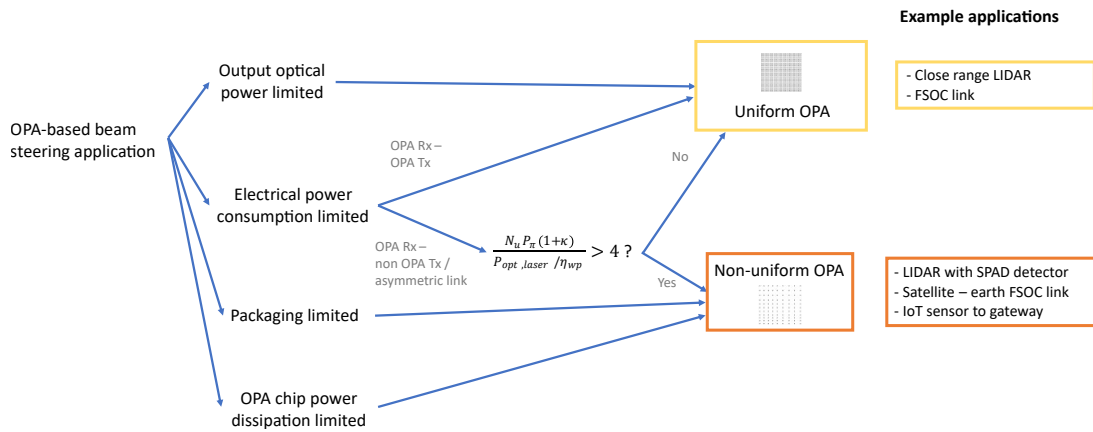


Figure 5-12: OPA configuration decision chart.

approach when our system is limited by power dissipation in the chip or by packaging density.

This is, to the best of our knowledge, the first time a thorough study of the applicability of non-uniform OPAs for realistic applications has been made. Our analysis presents a *recipe* to decide which OPA configuration is best depending on the application and constraints (Fig. 5-12).

Chapter 6

Conclusions and Future Work

In the first chapter of this thesis we argued how integrated optics is a promising avenue towards the realization of low cost, compact and reconfigurable imaging systems capable of overcoming the limitations of traditional free space optics. We identified two main challenges hampering the wide adoption of integrated optics in imaging applications: (1) the challenges of realizing controllable in- and out-coupling of the light in the photonic integrated circuit into free space and (2) the lack of a scalable, "full functionality" photonic platform allowing for generation, processing and launching of visible and IR light within the same substrate.

Within the field of imaging, we identified lensless holography and beamsteering systems as promising applications that can benefit greatly from the possibilities offered by integrated optics, and realized that both call for the realization of fully controllable arrays of closely spaced emitters.

With this in mind, in Chapters 2 and 3 of this thesis we have presented two techniques to realize such arrays of closely spaced emitters:

- In Chapter 2 we demonstrated the realization of surface emitting LEDs in CMOS

microelectronics processes. In particular, we experimentally realized an array of point-like emitters with sub-wavelength spacing (as small as 0.4λ). This is, to the best of the author's knowledge, the smallest spacing ever realized in an array of electronically individually addressable light sources.

We showed how we can use these sources to (1) perform large FoV holographic microscopy in a very compact form factor and (2) obtain subpixel shifted images without any mechanically moving parts, again in a very compact form factor.

We also introduced for the first time the use of a Deep Image Prior (DIP) architecture to overcome the limitations of the low temporal coherence and low generated power of our sources, and also introduced reflection holography as a means to both simplify the experimental setup (we have only one chip generating the illumination and detecting the hologram) and obtain some information about the setup geometry thanks to the generation of a double image.

- In Chapter 3 we used a similar approach to realize integrated waveguide-coupled LEDs in CMOS photonic platforms. We developed a detailed theoretical model to understand their behavior and experimentally characterized their performance. Compared to the surface emitting LEDs in Chapter 2, these sources are single mode, can have increased temporal coherence (in the case of the resonant LEDs) and their properties could be tuned by means of other integrated optical components such as filters, modulators... Thus, these sources have a wider application space than the surface emitting LEDs, although they also have lower power due to the limited waveguide mode coupling efficiency.

The waveguide coupled sources in Chapter 3 can be used for lensless imaging through the realization of grating coupler arrays, but to demonstrate the wider applicability of

these sources we discussed their use in refractive index sensors in Chapter 4. In particular, we showed how such sources, when used in conjunction with sensitive integrated photodetectors and sensing elements, can be used to realize truly monolithic refractive index sensors with performance similar to state of the art, commercial tools at a fraction of the size and cost. This constitutes a first demonstration of the potential of the "full functionality" photonic platforms we discussed above.

Finally, in Chapter 5 we delved into the scalability issues of OPA-based beam steering systems. In particular, we showed how, for advanced applications, the use of uniformly spaced emitter arrays quickly faces a bottleneck in terms of power consumption and area, requiring element numbers above the 100,000s and phase shifters with area and power consumption specifications that are orders of magnitude lower than state of the art devices.

Given this, we studied the potential of non-uniform OPAs to relax some of these bottlenecks while maintaining performance. We identified the performance tradeoffs with respect to uniform OPAs, mainly in terms of reduced SLSR and power in the main lobe, and showed that non-uniform OPAs can be beneficial in power limited scenarios and asymmetric links.

All in all, I believe this thesis constitutes an initial step towards the wide adoption of integrated optics for imaging applications. By means of both theoretical and experimental work, we have demonstrated how silicon photonic integrated circuits can indeed result in high performance imaging systems with reduced cost and size, which could significantly widen the use of these systems from specialized settings to everyday consumer use. There is, nevertheless, a significant amount of work yet to be done, as well as several research directions worth pursuing, which we will discuss briefly in the following sections.

6.1 Lensless Holography with Near-Point CMOS Light Sources

Chapter 2 presented the use of near-point CMOS LEDs as highly spatially coherent light sources for holography and demonstrated the possibility of realizing arrays of really closely spaced emitters. Taking advantage of the fact that our sources are realized in CMOS processes, we also introduced the concept of lensless reflection holography and showed that we can get subpixel shift images without moving parts using our CMOS LED array.

Significant improvements and further exploration should be performed in most of the innovations presented in this chapter.

6.1.1 Improving the performance of near-point CMOS LEDs

Regarding the performance of the near-point CMOS LED, devising approaches to improve the output power (both in terms of electrical to optical efficiency and out-coupling efficiency) is paramount to improve the imaging SNR. Similarly, techniques to improve the temporal coherence of the emitted light would also help for holographic imaging.

In order to devise these approaches it is necessary to get a deep understanding of the carrier transport mechanisms dominating the performance of these near-point LEDs, which is something that requires careful modeling in a computational tool such as Sentaurus TCAD. The biggest challenge to developing such a model is including the presence of the filament into the structure, as both its dimensions and properties (resistance, recombination coefficients, etc.) are unknown. Another important challenge is the complexity of the physics involved, which likely require the inclusion of thermal and high energy effects to correctly capture the full behavior of the device.

Of course, another approach to improve the imaging SNR without improving the

output power of the sources is to use CMOS detector arrays with higher responsivity in the near-IR region (the range of wavelengths where our CMOS LEDs emit). Low cost, high performance cameras at these wavelengths are being developed as they are becoming commercially attractive due to the switch of LIDAR products to operate at these wavelengths (see for example [311]).

6.1.2 Scaling the CMOS LED array to a large number of elements

In this thesis we demonstrated a CMOS LED array with four elements, and discussed how the scalability of CMOS processes guarantees an easy scaling of such arrays to large element counts.

As discussed in Section 2.3, scaling the array to a large number of elements is necessary so that we can realize multiple subpixel shift values and therefore increase the imaging resolution toward the diffraction limit (with a required number of shifts on the order of 50). Having a large emitter array can also help increase the FoV beyond the limit set by the NA of a single emitter, although this is not a significant issue for our sources because they have a really high NA stemming from their very small emission spot.

The greatest challenge for scaling up the number of emitters in our array is being able to individually contact each poly-Si bar (1) without occluding the emission of nearby emitters and (2) maintaining the small emitter spacing required for effective subpixel shift imaging. It is fairly straightforward to realize a 2D array of closely spaced emitters with only two rows, as we can then use a configuration like the one depicted in Fig. 6-1(a). Nevertheless, having more than 2 rows with subwavelength spacing becomes very challenging because the minimum size of the electrical contact to a poly-Si bar is about $100 \text{ nm} \times 100 \text{ nm}$ (which is set by the size of a single electrical via contacting the

poly-Si).

Nevertheless, a 2 row array of emitters is enough to achieve an arbitrary number of subpixel shifts in both the x and y dimensions. It is important to realize that in an emitter array for applications in subpixel shift imaging we do not want a uniform spacing. That's because we do not gain any additional information by taking more than two pictures with the same subpixel shift ¹. We can therefore achieve a large number of distinct subpixel shifts with a configuration like the one shown in Fig. 6-1(a), where each element in the 2 row array is spaced by a different amount. Notice how in principle, with an $N \times 2$ array of emitters we could generate a total of $\binom{2N}{2}$ pixel shifts in each direction (each pair of emitters can be configured to have a different Δx and Δy). As shown in Fig. 6-1(a), we adapt the shape of the c-Si region to minimize the uncertainty on the position of the Si filament and thus minimize the uncertainty on the illumination spot position.

If we want to increase the FoV of our imaging system we can then replicate our 2-row array in different positions of the chip as depicted in Fig. 6-1(b).

It is also worth mentioning that the fact that our emitter array is fabricated in a CMOS process allows for the realization of the driver system that turns on each individual LED at a time monolithically in the same chip. This greatly simplifies the system packaging (and therefore reduces cost), and also overcomes the spacing limitations set by the requirements of 3D integration, which limits practical emitter pitches to $5 \mu\text{m}$ and above [98, 99].

6.1.3 Further exploiting computational techniques

I believe there is a great potential to further exploit computational techniques beyond what we demonstrated in Chapter 2, where we exploited DIP architectures to improve

¹Doing this is equivalent to taking two identical pictures - we could gain in terms of SNR as we could average these two pictures, but we do not gain additional subpixel shifts by doing so

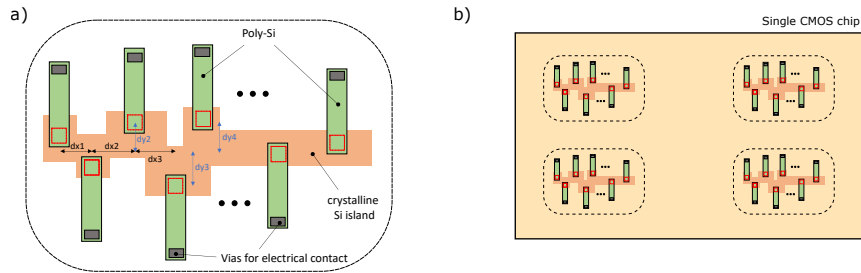


Figure 6-1: Scaling the CMOS LED array to large element number. (a) With a two row array we can achieve an arbitrary number of spacings in both the x and y dimensions. (b) To increase the FoV of the subpixel shift microscope we can replicate the 2-row array shown in (a) in different positions of the CMOS chip.

the quality of the sample reconstructions obtained from our acquired holograms.

For example, it would be interesting to further explore the possibilities that the double hologram formation in reflection holography offers beyond retrieving geometrical information of the setup. For instance, in 3-dimensional structures the double image formation could allow to retrieve relative depth using a similar approach to stereographic imaging.

I also believe that the DIP architecture can unlock novel functionalities for holographic imaging. For example, we could take advantage of the fact that the DIP architecture is capable of recovering the wavelength spectrum of the light hitting the camera to perform spectroscopy as well as imaging.

There also needs to be some further work to theoretically understand the inner workings and the limitations of the AWGN noise addition in the novel DIP architecture we presented in Section 2.3.2. This includes, for example, studying if there is an optimal noise variance that results in the best reconstruction quality or exploring at which SNR the AWGN addition cannot reliably recover the sample features.

6.2 Native Single-mode Waveguide-coupled LEDs in Unmodified Silicon Photonics Fabrication Platforms

In Chapter 3 we presented waveguide coupled silicon LEDs in silicon photonics fabrication platforms, both theoretically through TCAD and FDTD models as well as experimentally.

Future work for these devices can be divided into two main categories: further characterization and device optimization.

6.2.1 Further characterization

In Chapter 3 we showed the DC characterization of the fabricated waveguide coupled LEDs: generated optical power and spectrum.

It is also very important to characterize the dynamic behavior of these devices, as bandwidth is a critical feature for some of the potential applications of these devices such as its use in optoisolators or for intra-chip communication. Since we are using forward biased pn junctions to generate the required bimolecular recombination, we expect the bandwidth of these devices to be relatively slow (on the order of 1 MHz - 500 MHz depending on the bias current) and limited by the high diffusion capacitance of the pn junction.

It would also be of interest to characterize the performance of these devices at low temperatures, since in silicon the radiative recombination coefficient increases [217] and Auger recombination decreases [312]. Operation of waveguide coupled LEDs at low temperatures could enable low power, low thermal budget readout of cryogenic signals to room temperature, an important requirement for potentially disruptive cryogenic technologies such as high performance computing [313,314] or quantum computing [315,316].

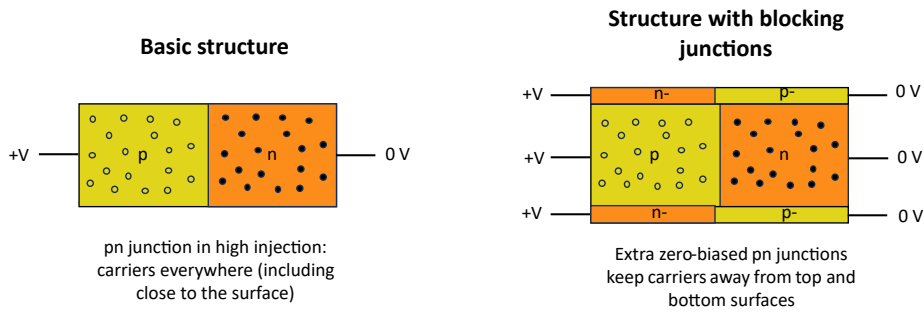


Figure 6-2: Use of blocking junctions to reduce effect of surface recombination in the electrical to optical conversion efficiency of waveguide-coupled LEDs.

This is a challenge that I studied in detail during my Masters thesis [317].

6.2.2 Device optimization

While the power generated by our experimentally demonstrated devices is enough to be detected both by integrated photodetectors and cameras, it would be very appealing to increase the efficiency and generated optical power of our devices. We envision several avenues to realize such improvements.

Increasing electrical to optical efficiency

As discussed in Chapter 3, an approach to improve the electrical to optical efficiency is to minimize surface recombination, which can be achieved by improving the quality of surface passivation or by using some approach to keep the generated carriers away from surfaces. The latter could be achieved through the use of additional zero-biased pn junctions close to the surfaces, which would keep the injected carriers away from the surfaces thanks to the junction's potential barrier. A simple schematic of this concept is shown in Fig. 6-2. From Fig. 6-2 it is obvious that in this configuration care needs to be taken to avoid the formation of a low impedance resistor between the main doping

regions (the ones that are forward biased in high injection) and the blocking regions with the same doping specie.

The LED efficiency can also be increased by reducing Auger recombination in the silicon bulk, which could be achieved by means of stress engineering [223].

Alternatively, it is also possible to increase efficiency by exploiting mechanisms to improve the radiative recombination coefficient of silicon. The simplest approach would be to exploit defect implantation, either as a step in the fabrication process [207] or post-fabrication [318].

Increasing coupling efficiency

There is also a lot of room for improvement in terms of coupling efficiency, which for our experimentally demonstrated devices is between 1% and 2.5%.

We envision two approaches to improve the coupling efficiency:

1. The use of an adiabatic ring configuration [319] as depicted in Fig. 6-3(a). This allows us to use a truly single mode waveguide all along the circumference of the ring and minimize the volume of the contact regions. This way, we can ensure a relatively high coupling efficiency regardless of where the radiative recombination event happens, a property that we could not achieve with our experimentally fabricated devices (which have large contact regions where radiative recombination occurs but that do not couple to the waveguide mode).
2. The highest possible coupling efficiency in our waveguide coupled LEDs would be achieved if all radiative recombination events occurred in the area where the optical mode is the most intense (i.e., right at the center of the optical waveguide). This is really challenging to achieve with our current carrier injection approach, which exploits pn junctions biased in the high injection regime.

An alternative approach to achieve localized carrier injection would be the use of filaments, the same approach we used to achieve small emission spots in our surface emitting LEDs of Chapter 2. The added complexity is that now the structure not only needs to ensure the formation of a filament but also needs to allow for an optical waveguide mode to be supported. A possible configuration that would fulfill both requirements is shown in Fig. 6-3(b), where a single mode waveguide is formed in the lightly p-doped crystalline silicon (which is biased at 0 V) and we use an n-doped poly-Si taper on top (biased at a negative voltage) to form the filament and achieve localized carrier injection at the center of the waveguide. The tapered shape is used to ensure that the filament is formed at the waveguide center as it is the region where the electric field is the strongest.

We note that this filament injection waveguide coupled LED is not based on a forward biased pn junction and thus its bandwidth is not limited by high diffusion capacitance. As a result, fast operation of these devices could be possible: characterization results in similar structures performed by Dr. Zheng Li achieved pulse generation with a FWHM as short as 1 ns [320].

6.3 Fully integrated, monolithic refractive index sensors in CMOS photonic fabrication processes

In Chapter 4 we theoretically studied the application of our waveguide coupled LEDs to realize truly monolithic, highly multiplexed refractive index sensors, showing that performances comparable to state of the art, commercial products is possible.

Of course, the next step is to experimentally demonstrate such a sensor and experimentally evaluate its performance. As discussed, we have demonstrated every single

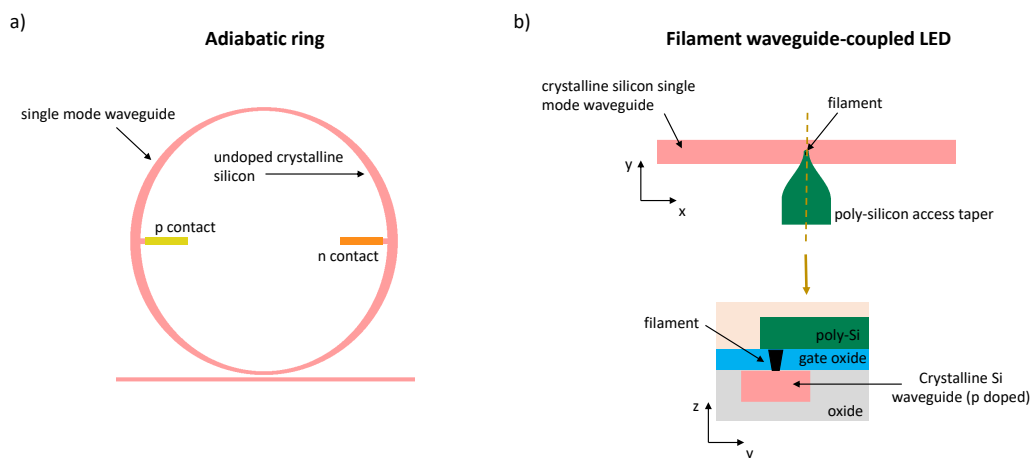


Figure 6-3: Structures to increase coupling efficiency of waveguide-coupled LEDs. (a) An adiabatic ring to minimize the contact area. (b) Localizing carrier injection in the middle of the optical waveguide using a filament through the gate oxide and a poly-silicon access taper.

component required to realize the sensor, but it is still necessary to integrate all of them together. With this in mind I have completed two distinct tape-outs, each with a different focus, which we discuss in the next sections.

6.3.1 NanoSOI tape-out: Focus on learning lessons in refractometric sensing

Before being able to demonstrate our concept of a fully monolithic refractive index sensor it is necessary for us to gather experience in refractometric sensing. With this goal, I taped out some sensing structures in the NanoSOI silicon photonics process [321]. This is a simple process meant for prototyping, with a single silicon layer and up to two metal layers, one of which can be used as a heater. While this process does not have doping implants (and therefore we cannot realize our waveguide coupled LED sources or photodetectors), it offers an oxide opening step which allows us to expose the silicon

waveguide to the environment without any post-processing.

As such, this tape-out gives us the possibility of learning about: (1) the behavior of ring resonators as refractometric sensors, (2) the best approach to calibrate and measure changes in the resonance wavelength of the sensing ring, (3) functionalization of the silicon surface, (4) delivery of the analyte, (5) approaches to maximize sensitivity.

Nevertheless, these structures require external light generation and detection, which prevent us from demonstrating a fully monolithic sensor. On the flip side, the fact that light generation and detection are external allows us to easily study the sensitivity limits of these sensors through the variation of both the input optical power as well as the noise floor of the photodetector.

The test structures we included in this tape-out are the following (Fig. 6-4):

- Single rings for measuring the achievable sensitivity of different functionalization techniques, both in a ridge waveguide configuration and in a slot waveguide configuration.
- Structures with two rings, one exposed to the analyte and one meant for calibration of environmental fluctuations not related to the analyte of interest (for example, changes in temperature or humidity).
- Structures to realize Sensing Scheme 2 presented in Chapter 4, consisting of two rings connected through the drop port. These structures use edge couplers to enable the input and output of broadband light, which is necessary to mimic the use of an LED source in the fully integrated sensor.

Each of these structures was replicated with a variety of coupling gaps, waveguide widths and operation wavelengths.

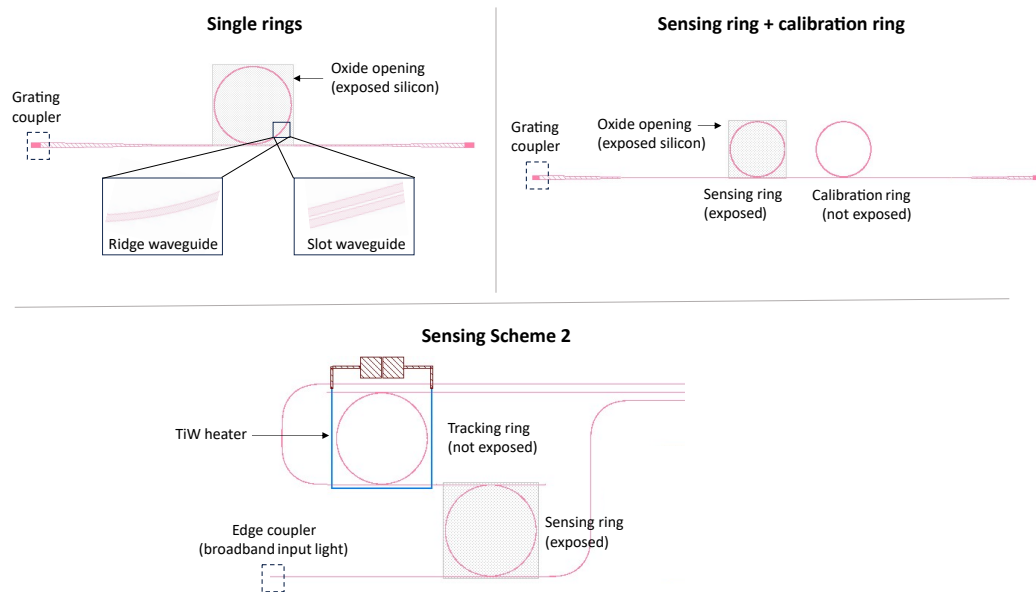


Figure 6-4: Test structures for refractometric sensing included in the NanoSOI tape-out.

6.3.2 GlobalFoundries 45CLO tape-out: Experimental demonstration of fully monolithic refractive index sensors

As mentioned, the NanoSOI tape-out described above lacks doping implants, so it is not possible to realize integrated LEDs or photodetectors. To be able to demonstrate our concept of a fully integrated refractometric index sensor I did an additional tape-out in the GlobalFoundries 45CLO process, the same process we used to demonstrate the resonant rib waveguide LED in Chapter 3.

This process has doping implants (and therefore we can realize integrated silicon LEDs) and Ge (so we can also realize high responsivity photodetectors). I therefore included structures to realize both Sensing Scheme 1 and Sensing Scheme 2 described in Chapter 4 in this tape-out (Fig. 6-5).

These structures should enable the demonstration and characterization of our concept of a fully integrated refractive index sensor based on a broadband source. One added

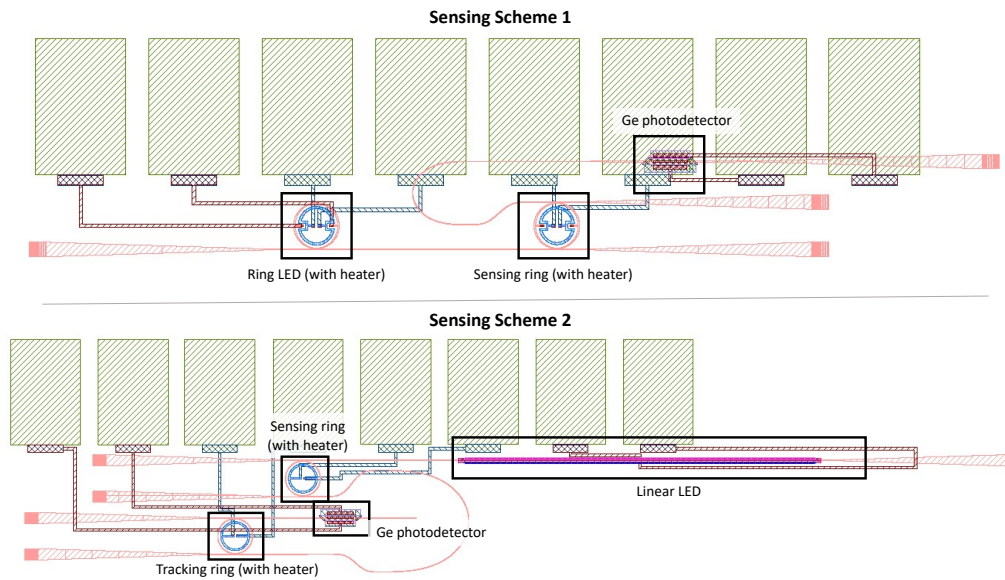


Figure 6-5: Test structures for the demonstration of the fully integrated refractometric sensor included in the 45CLO tape-out.

challenge in these structures is that post-processing is required to expose the sensing ring to the analyte, since an oxide opening step was not included in this tape-out (although it can be added to the 45CLO process we used as a process variation). A possibility to overcome the need for exposing the sensing ring is to mimic the resonance shift caused by the presence of an analyte though the use of the available heaters to shift the ring resonance through the thermo-optic effect.

The only non-integrated device in the taped-out structures is the current sensor, which needs to be external. The next step after demonstrating the correct functioning of the sensors in these tape-out would be to also integrate the current sensor and replicate the structure multiple times to demonstrate the potential for large scale multiplexing.

6.4 OPA architectures for advanced optical beam-forming applications

In Chapter 5 we analyzed in detail the suitability of the use of non-uniform OPAs in advanced beamsteering applications, showing that it can indeed be beneficial in scenarios where total electrical power consumption is the limiting performance requirement.

6.4.1 Round OPAs for optical power limited scenarios

From our analysis in Chapter 5, we have seen how non-uniform OPAs are not a suitable solution for scenarios where we want to achieve maximum output optical power. In this case, uniform OPAs are the only approach to achieve as large an output power as possible and therefore achieve as large a link distance as possible.

It is reasonable to assume that in these scenarios where large link distances are required we will use some free space lens arrangement after the OPA. This can be, for example, a telescope in the case of an OPA for FSOC applications.

Since all lenses are round, the use of uniformly spaced OPAs in a round configuration (as depicted in Fig. 6-6(a)) can offer significant gains in emitted power and therefore link distance.

Such gain comes simply from the fact that we can fit more emitting antennas in a circular lens when using a circularly arranged OPA. As shown in Fig. 6-6(b), a square array inscribed in a circle of a diameter equal to the diagonal of said square only occupies 63% of the circle area. By using a perfectly circular array we can fill the entire aperture. Using the same principle from the thinned-array curse, such perfectly circular array will have a $1/0.63 = 1.6\times$ larger emitted power compared to the square OPA (and it will also have $1.6\times$ the number of emitters).

One could argue that such gain could be obtained by simply scaling the number of elements in a square OPA (with its associated increase in dimensions). While true, such an approach would require a larger diameter lens, which results in a significant cost increase ².

While the realization of a perfectly round OPA might be challenging, even a pixelated approximation to a circle results in significant gains with respect to a square OPA. This is shown in Fig. 6-6(c), where we plot round OPAs with different resolution (which we refer to as granularity ³). It can be shown that the fractional area occupied by a pixelated square of granularity N inscribed in a circle is given by:

$$\frac{A_{pix_circ}}{A_{circ}} = 0.78 + \sum_{i=1}^N 2^{i+1} \left(\frac{0.325}{2^i}\right)^2 \quad (6.1)$$

Figure 6-6(c) shows the gain in link budget for an OPA - OPA link (Scenario 2 discussed in the previous section) with respect to a square OPA for different granularity levels - notice that since this is an OPA - OPA link, the gain is $(A_{pix_circ}/A_{square})^2$. We can see how link gains up to 4 dB are possible for a perfectly circular OPA.

Nevertheless, a circular OPA will have additional losses that prevent reaching the gain levels shown in Fig. 6-6(c):

²Lens cost increase with diameter is highly nonlinear - large diameter lenses are expensive and induce more aberrations.

³We define granularity with the following recursive algorithm:

1. Start with a circle of granularity 0, which is simply a cross as shown in the left of Fig. 6-6(c).
2. A circle of granularity 1 can be constructed by taking each inner edge of the cross (i.e, the granularity 0 circle) and adding a square with a side length corresponding to half of said inner edge length.
3. A circle of granularity 2 can be constructed by taking each inner edge of the granularity 1 circle and adding a square with a side length corresponding to half of said inner edge length.
4. Repeat the process as many times as the desired granularity.

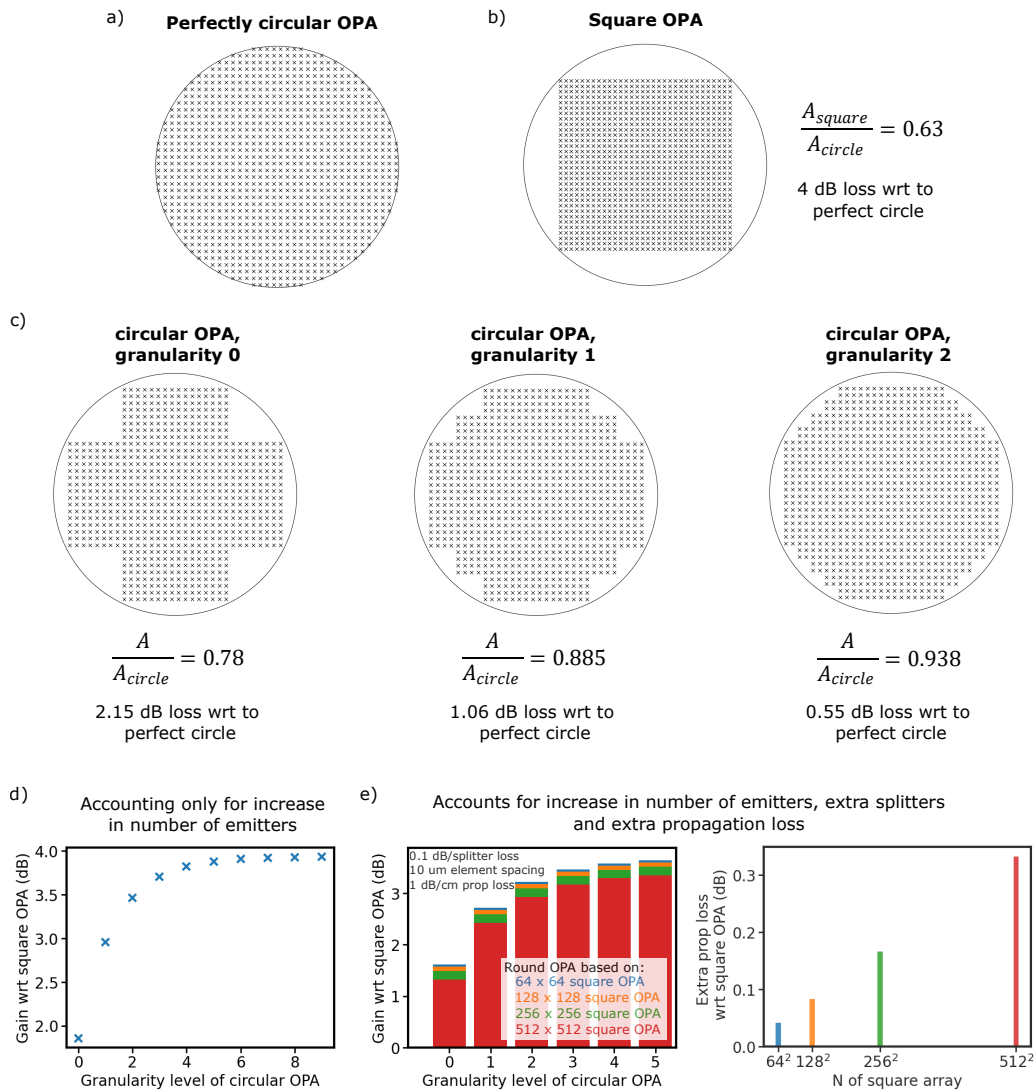


Figure 6-6: Round OPA gains. (a) A uniformly spaced OPA in a perfectly circular configuration. (b) A square OPA, which only occupies 63% of the area of the circle that encompasses it. (c) Pixelated circular OPAs of increasing granularity. Higher granularities result in better approximations to a perfect circle and therefore have less loss. (d) The link gain with respect to a square OPA as a function of circular OPA granularity if we only account for the increase in number of emitters. (e) Link gain with respect to a square OPA as a function of circular OPA granularity, accounting for extra propagation loss and splitter loss. 4 different OPA sizes are shown.

- Since there are more antennas in the OPA we need a larger number of 1x2 splitters in the power distribution network. Additional 1x2 splitters increase the insertion loss of such network (η_{power_distr}). For our analysis we will assume that each additional 1x2 splitter required adds 0.1 dB of extra insertion loss, and we can calculate the number of splitters through $N_{1x2} = \lceil \log_2(N) \rceil$.
- Round OPAs will also have larger routing loss simply because the elements are further away from the center. We can crudely estimate the insertion loss due to routing by using the average distance of all the emitters to the center of the OPA:

$$\bar{r} = \frac{1}{N} \sum_{i=1}^N \sqrt{x_i^2 + y_i^2} \quad (6.2)$$

Where x_i (y_i) is the x (y) position of emitter i .

For a square OPA of side length $L = r\sqrt{2} \rightarrow \bar{r} = \frac{\sqrt{2}r}{6}[\sqrt{2} + \log(1 + \sqrt{2})]$.

And for a circle of radius r it is $\bar{r} = 2r/3$.

So the ratio of average distance between a square OPA and a circular OPA is $\bar{r}_{circ}/\bar{r}_{sq} = 2\sqrt{2}/(\sqrt{2} + \log(1 + \sqrt{2})) = 1.574$. We will therefore assume that the propagation loss in a circular OPA is $1.574 \times$ larger than that of a square OPA.

Figure 6-6(d) shows the link budget gains with respect to a square OPA for different square OPA element counts, assuming a $10 \mu\text{m}$ spacing between adjacent antennas and a propagation loss of 1 dB/cm (which is a reasonable estimate of worst case propagation loss in a SiN single mode waveguide).

Two main conclusions can be extracted from our analysis:

1. As the size of the square OPA increases the gains from switching to a circular OPA decrease. This is because larger OPAs result in longer total propagation distances

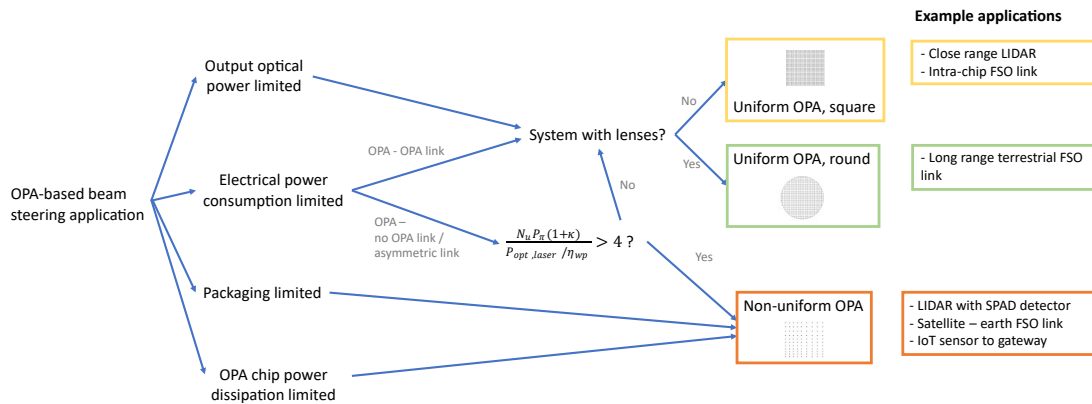


Figure 6-7: OPA configuration decision chart including round OPAs.

(as propagation distance is $\propto r \propto \sqrt{N}$), which has a larger penalty for round OPAs (since $\overline{r}_{round} = 1.574\overline{r}_{square}$).

- The gains of increasing the circular OPA granularity decrease as the granularity increases. Having circular OPA granularities beyond 3 yield almost no improvement in link budget.

We can add the use of round OPAs to our OPA configuration *recipe* we developed in Chapter 5, which is shown in Fig. 6-7

Practical realization of round OPAs through pruned H-trees

We have seen how using circular OPAs can achieve link budget increases (or equivalently SNR increases) above 3 dB. This can be highly attractive for applications that are optical power starved (notice that round OPAs have increased power consumption compared to square OPAs because they have more elements).

An important aspect that needs to be addressed for the realization of practical round OPAs is the routing and optical power distribution - how can we achieve efficient routing of the light to each antenna?

Here, we propose the construction of round OPAs through the use of pruned H-trees. H-trees are an attractive approach for the realization of the power distribution network in square OPAs, as they allow for compact and space efficient power distribution within the same area as the emitter array and only require the use of 1x2 splitters (Fig. 6-8(a))⁴.

We can construct a round OPA from a square H-tree network by simply pruning some of the splitters according to the desired granularity, as shown in Fig. 6-8(b). Two points are worth emphasizing:

1. Pruned H-trees require the use of splitters with a non 50:50 splitting ratio. The location and split ratio depend on the OPA size and the desired granularity.
2. Pruned H-trees can only realize circular OPAs with certain number of elements - it is not possible to realize circular OPAs of an arbitrary number of elements with a pruned H-tree.

One of the biggest concerns with this approach is the use of non-50:50 splitters. The realization of splitters with accurate 50:50 splitting ratio is relatively straightforward, as one can use 1x2 MMI components which by virtue of its symmetry ensure a near-perfect 50:50 splitting [322].

To achieve arbitrary coupling ratios different than 50:50 the standard approach is to use directional couplers, where evanescent coupling between two closely-spaced waveguides allows for power transfer from one waveguide to another. By tuning the interaction length between the two waveguides, the splitting ratio of the directional coupler can be controlled. Nevertheless, these structures are sensitive to fabrication variations, which

⁴Notice how the H-tree approach results in compact routing, but requires the phase shifter to be realized at the end of the H-tree and very close to the emitter. While efficient, this routing approach strongly limits the minimum achievable spacing between antennas to be above the size of one of such phase shifters, which limits the achievable FoV.

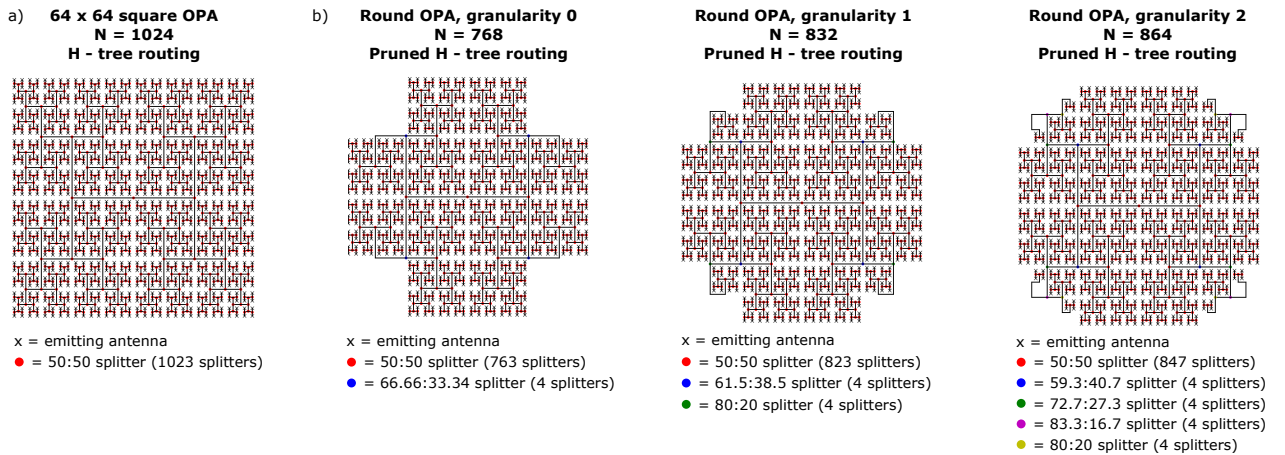


Figure 6-8: Round OPA routing through pruned H-trees. (a) An H-tree configuration for power routing of a 64 x 64 element OPA. (b) Pruned H-tree configurations for round OPAs with increasing granularities. For both (a) and (b), the crosses show positions of the emitting antennas, and circles show splitters. Different colors denote different splitting ratios.

cause asymmetry between the adjacent waveguides and result in variations in the effective refractive index of the optical modes, therefore causing variations in the split ratio. While several approaches have been proposed and demonstrated to improve the robustness to fabrication variations [323–325], even in this case the variations can be significant.

We studied the effects of non-even power emitted by each antenna in the total power emitted in the main lobe for a uniform OPA and found that, as long as the emitted phase is correct, there are no significant effects. For example, a $\pm 40\%$ relative variation in the power emitted by each antenna only results in a 6% decrease in the total power emitted in the main lobe. The requirement of non 50-50 splitters in round OPAs should therefore not be a significant source of performance impairments.

6.4.2 Realization of OPA-based beamsteering systems for real-world FSOC applications

During the last year of my PhD I have been working part time at Project Taara on realizing OPA-based beamsteering systems for terrestrial applications, in particular bringing last-mile connectivity to communities in developing countries [326]. I have participated in the completion of two tape-outs in two different silicon photonics platforms, with two different focus:

1. The experimental evaluation of the gain of round OPAs in practical terrestrial FSOC applications. A round OPA has been designed and fabricated following what we described in Section 6.4.1 above, and it will be tested in the near future.
2. While beneficial, it is not likely that the switch to non-uniform OPAs is enough to solve the scalability problem of OPAs for advanced applications. Advancements in other areas are also necessary, with improvements in the efficiency, power consumption and size of phase shifters arguably being the most critical (see Appendix A). On the context of my work with project Taara I have designed several ring-based silicon phase shifters (using both the thermo-optic and free carrier dispersion effects) with the goal of increasing efficiency and reducing area by exploiting the resonant nature of these devices.

Similarly, I believe that work on architectures requiring less than one phase shifter and/or driver per emitting element is a promising avenue to decrease the performance requirements of individual photonic components for advanced OPAs. Nevertheless, such architectures are usually hard to demonstrate experimentally due to inaccuracies caused by fabrication variations. Work on a scalable, fast and low-cost post-fabrication trimming

of photonic devices is paramount to enable such architectures in practical scenarios [327–330].

Appendix A

A survey of phase shifters for OPA-based beam steering applications

In Section 1.4.3 we introduced the importance of phase shifters for the realization of integrated OPAs. As discussed there, phase change range, power consumption, area, insertion loss, speed, crosstalk and residual amplitude modulation are all important performance metrics that need to be accounted for when designing an integrated phase shifter for an OPA.

We can define a figure of merit (FoM) that includes several of these performance metrics (area, power consumption and loss in particular). For devices with DC power consumption, we can use $FoM = P_{\pi}L\alpha$ (with units of dB·mW), where L is the device length and α is the propagation length. For devices without power consumption we can use $FoM = V_{\pi}L\alpha$ (with units of dB·V).

To realize a phase shifter we need a structure that changes the refractive index of

the medium through which the optical mode is propagating (typically Si or SiN). The acquired phase change in a waveguide of length L with a refractive index change Δn is given by $\Delta\phi = 2\pi L\Delta n/\lambda$. It is clear how to minimize the required device length we want to achieve as large a Δn as possible.

Two main physical mechanisms have been typically used to realize phase shifters in OPA systems: the thermo-optic effect and the plasma dispersion effect (Fig. A-1).

A.1 Thermo-optic phase shifters

Thermo-optic (TO) phase shifters are the most typical phase shifter used in reported demonstrations due to its ease of implementation. TO phase shifters exploit the TO effect, in which a change in the temperature of the material is accompanied by a change in the refractive index. Silicon has a pretty high TO coefficient, which at a wavelength of 1550 nm and at room temperature is $dn_{Si}/dT = 1.87 \times 10^{-4} K^{-1}$ [331]. As a result, relatively compact and efficient TO phase shifters can be realized in silicon photonics processes (Table A.1, Fig. A-2) [332]. As a rule of thumb, a 30 degree change in silicon temperature translates into $\Delta n \approx 5.61 \times 10^{-3}$, which requires a length of about 250 μm to achieve a 2π phase shift. SiN also has an associated thermo-optic effect, albeit about 5 \times weaker than in Si ($dn_{SiN}/dT = 2.45 \times 10^{-5} K^{-1}$ [333]), which results in increased power consumption (Table A.1).

A variety of approaches to achieve a heating of the Si or SiN waveguide have been employed, including the use of metal heaters on top of the waveguide or the use of doped silicon resistors adjacent to the waveguide (Ref. [332] is given to the interested reader for an exhaustive list of the different approaches).

As summarized in Fig. A-1, TO phase shifters are simple to implement both in Si and SiN and typically result in relatively compact devices due to the strong TO effect.

Furthermore, since the TO effect does not have any effect on the propagation loss and does not require doping of the optical waveguide, TO phase shifters have low insertion loss and no residual amplitude modulation.

Nevertheless, TO phase shifters have a pretty high power consumption (since they have an associated DC power dissipation required to maintain the desired temperature). Multiple approaches have been employed to increase the power efficiency of TO phase shifters (including the use of folded and multipass waveguides), but power consumptions are still on the 1-3 mW range, which are too high for large scale OPAs. The exception is when an undercut is used to remove the substrate (usually silicon dioxide) locally around the structure to reduce heat leakage and achieve efficient heating. While this approach can result in power dissipation as low as $50\mu\text{W}$ [334], it also results in very slow devices with operation speeds below 1 kHz and requires careful fabrication and design to avoid the structures collapsing.

Typical time constants for TO phase shifters are 1 - 10 μs (due to the need for the heat to diffuse in and out of the waveguide), limiting the operation speed to about 200 kHz. Finally, TO phase shifters also suffer from crosstalk between neighboring devices due to heat conduction, which makes it necessary to keep adjacent devices relatively far from each other and thus increases the required chip area.

A.2 Plasma dispersion phase shifters

The other typical approach to the realization of silicon integrated phase shifters for OPA applications is the use of the plasma dispersion effect. Here, a change in the concentration of free carriers in the optical waveguide generates a change in its refractive index. The plasma dispersion effect can be explained and approximated using the Drude model [13], but empirical expressions for the generated refractive index change at room temperature

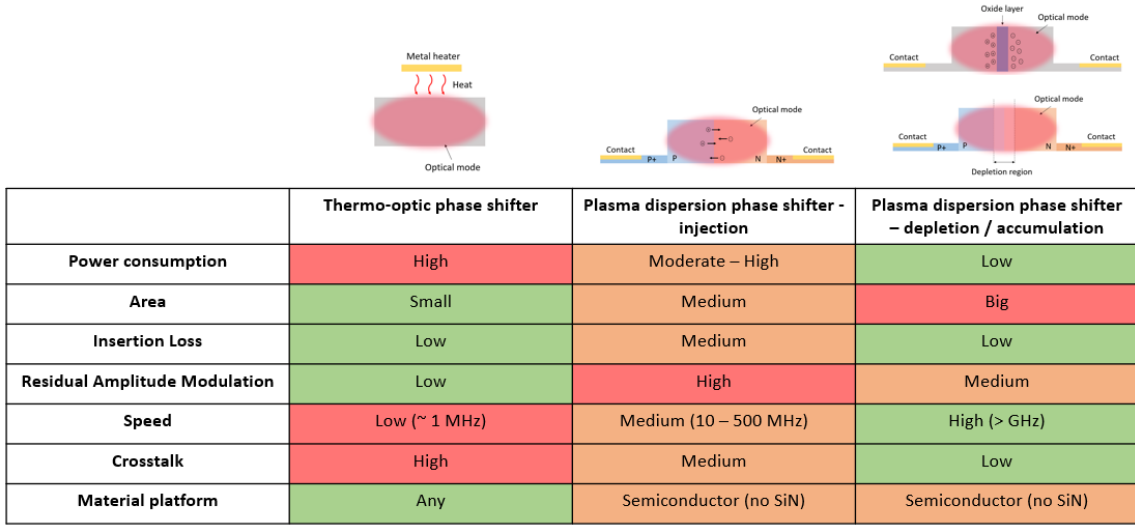


Figure A-1: A comparison of the performance of difference phase shifters used for the realization of OPAs.

exist in what is commonly referred as the Soref's equations [335].

At 1550 nm wavelength:

$$\Delta n = -(8.8 \times 10^{-22} \times \Delta N_e + 8.5 \times 10^{-18} \times \Delta N_h^{0.8}) \quad (\text{A.1})$$

$$\Delta \alpha = 8.5 \times 10^{-18} \times \Delta N_e + 6 \times 10^{-18} \times \Delta N_h \quad (\text{A.2})$$

Where Δn is the change in the refractive index, $\Delta \alpha$ is the change in the absorption coefficient in cm^{-1} and ΔN_e (ΔN_h) is the change in the density of free electrons (holes) in cm^{-3} . As an example, a change in the carrier density of $5 \times 10^{17} \text{ cm}^{-3}$ results in a change in refractive index $\Delta n = -1.66 \times 10^{-3}$, requiring a length of about $100 \mu\text{m}$ for a 2π phase shift.

Compared to TO phase shifters, plasma dispersion phase shifters generally have a lower power dissipation and higher operation speed. Nevertheless, plasma dispersion can

only be exploited in semiconductor materials (therefore we can't use it in SiN) and, as is apparent from Eqs. A.1 and A.2, it has an associated residual amplitude modulation (changing the refractive index n causes a change in the propagation loss α). Furthermore, since plasma dispersion phase shifters rely on the use of doped waveguides to realize pn junctions and achieve a change in carrier concentration they generally have a higher insertion loss (as the presence of doping impurities increases propagation loss due to free carrier induced loss, Eq. A.2).

Within plasma dispersion phase shifters, different approaches can be used to generate the necessary changes in the free carrier density (ΔN_e and ΔN_h), which are schematically depicted in Fig. A-1 [336]. Carrier accumulation (top right in Fig. A-1) uses a capacitor structure and modifies the charge stored in it. Carrier depletion (middle right in Fig. A-1) uses a reversed biased p-n junction, and changes in the carrier density are achieved by modulating the width of the depletion region. Carrier injection (center in Fig. A-1) uses a forward biased p-n or p-i-n junction to control the amount of carriers injected into the optical waveguide.

Each operation mode features a different set of advantages and drawbacks. In terms of power consumption, the use of depletion or accumulation shifters is advantageous since these do not have a DC power dissipation, while injection-based shifters require the flow of a DC current. In terms of efficiency, carrier injection is advantageous since it generates the highest change in refractive index (and thus results in shorter required lengths), followed by carrier accumulation and then carrier depletion, which has the weakest associated Δn and usually requires high V_π on the order of 3 - 6 V. Carrier depletion and accumulation can achieve high operation speeds over 1 GHz, while carrier injection has a limited speed below ≈ 100 MHz due to the high diffusion capacitance of a forward biased p-n junction (notice this speed is enough both for FSO and LIDAR applications). Finally, while injection and depletion devices are easily implemented in

silicon photonics processes, carrier accumulation schemes are more challenging due to the need of a thin insulating layer (usually made of silicon dioxide) to realize the capacitor structure.

Table A.2 lists demonstrations of efficient silicon integrated phase shifters in the three operation modes discussed, and their performance is plotted in Fig. A-2. While promising, more work is required to decrease the driving voltage / device footprint ¹ and minimize insertion loss (as seen in Fig. A-2, no demonstrated plasma dispersion phase shifter has good enough performance for our requirements).

A.3 Other phase shifters

While TO and plasma dispersion are the most common choices to realize phase shifters, a variety of alternative approaches have been investigated for the realization of compact and efficient phase shifters (Table A.3, Fig. A-2).

Some examples include:

- Organic polymers: The use of high electro-optic coefficient polymers exploiting the Pockels effect in silicon slot waveguides can result in phase modulators with no DC consumption and very low driving voltages [8]. Nevertheless, these devices tend to have high propagation losses, require postprocessing to expose the silicon waveguide and deposit the polymer, and there are concerns with the stability of these materials with time and temperature. The efficiency of polymer-based phase shifters can be enhanced with the use of plasmonic structures, but these have even higher propagation losses and are not readily CMOS compatible (gold, which is the typical metal used in plasmonic structures, is not CMOS compatible) [337].

¹Notice that these are complementary variables. A longer device will require a smaller driving voltage.

- Ferroelectric materials - Lithium Niobate (LN) and Barium Titanate (BTO): Lithium Niobate (LN) is a common material used to realize optical modulators due to its low loss and high electro-optic coefficient. While traditionally limited to bulk devices, the advent of thin film lithium niobate (TFLN) has allowed for the realization of integrated optical devices with high performance [338], and phase modulators with low driving voltages and low loss have been reported in the literature [339]. BTO is another ferroelectric material with a higher electro-optic coefficient than LN, therefore resulting in higher efficiency phase modulators [340]. Both BTO and LN are challenging to pattern and not CMOS compatible, raising concerns of scalability, cost and integration with silicon photonics. Additionally, they require the application of high voltage to align all the ferroelectric domains and achieve maximum efficiency (a process commonly known as "poling") and suffer from DC bias drifts that can compromise operation, especially at the low speeds required for our applications [341].
- Microelectromechanical systems (MEMS): Photonic integrated MEMS systems where the application of a control signal (usually a voltage) results in mechanical movement of suspended optical waveguides enables the realization of multiple optical operations, including phase shifters [342]. Such phase shifters can have low insertion loss, MHz speed operation and low $V_{\pi}L$ [343]. Nevertheless, these structures are usually short to reduce concerns with collapsing or breaking of the suspended waveguides, which results in relatively large V_{π} around 10 V. Furthermore, MEMS phase shifters require the "release" of the optical waveguides (so that they are suspended and can mechanically move), a fabrication step which is not standard in photonics fabrication processes.
- III/V materials: As we have already discussed, III/V materials have been widely

studied for integrated photonics applications due to their direct bandgap (that allows for the realization of on-chip sources) and their first order electro-optic effect (where the refractive index changes with applied electric field, something that does not happen in centrosymmetric materials like silicon) [2,3]. As a result, efficient phase modulators have been realized in these materials [344].

- III/V-Si hybrids: As discussed in the point above, III/V materials are advantageous compared to silicon for the realization of phase shifters due to their first order electro-optic effect. Nevertheless, purely III/V phase shifters need to be fabricated in a different substrate and then either transferred to the silicon chip or optically connected through optical fibers or some other method. Instead, the addition of the III/V material directly into the silicon fabrication process results in simpler packaging and potentially reduced cost. Using this approach, very efficient phase shifters with low loss have been demonstrated [345]. The main challenge for the adoption of these structures is the need for the addition of the III/V material into the fabrication process flow, which is not standard and thus results in lower yield and higher cost.
- Liquid crystals (LCs): The use of liquid crystals with strong birefringence can be used to realize phase modulation by using a control signal to reorient the molecules and therefore obtain a change in the refractive index [346]. Traditionally, LC devices are bulky and not integrated, but work on miniaturization and integration with silicon and silicon nitride is under way [347]. The main disadvantage of LCs is their slow bandwidth due to the need for the molecules to reorient, with typical speeds limited to the KHz range.

While highly promising (as seen in Fig. A-2, several devices meet the requirements for FSO and LIDAR applications), these approaches have the important drawback of not

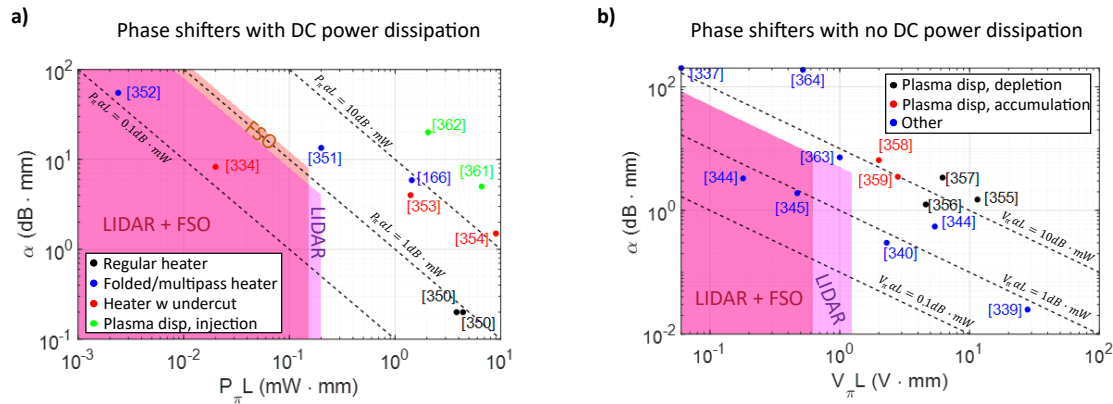


Figure A-2: High performance phase shifters reported in the literature. (a) Phase shifters that have DC power dissipation. (b) Phase shifters with no DC power dissipation. The red shaded area shows acceptable metrics for FSO applications ($P_\pi < 1.22 \text{ mW}$ or $V_\pi < 5 \text{ V}$, $L < 125 \mu\text{m}$, $IL < 1 \text{ dB}$), and the magenta shaded area for LIDAR applications ($P_\pi < 0.8 \text{ mW}$ or $V_\pi < 5 \text{ V}$, $L < 250 \mu\text{m}$, $IL < 1 \text{ dB}$). While some phase shifters can deliver the required performance in terms of $P_\pi L$ and α , none can do so with the required operation speed.

being readily compatible with standard foundry fabrication processes. Thus, either post-processing, hybrid integration or non-standard fabrication is required, raising important concerns in terms of ease of fabrication, cost and scalability.

References [336, 348, 349] are given to the interested reader for a more detailed review of phase shifters in integrated optics platforms.

Thermo-optic phase shifters

Ref. (year)	Brief description	Material	Time constant	Footprint	IL	P_{π}	$P_{\pi L}$	$P_{\pi L\alpha}$
[350] (2021)	TiN metal heater	Si	9 μ s	2 μ m x 200 μ m	NR	19.12 mW	3.82 mW·mm	NR
[350] (2021)	Doped silicon heater	Si	12.8 μ s	2 μ m x 200 μ m	NR	21.96 mW	4.39 mW·mm	NR
[166] (2020)	Multipass waveguide, metal heater	Si	6.5 μ s	25 μ m x 851 μ m	5 dB	1.7 mW	1.44 mW·mm	8.5 dB·mW
[351] (2020)	Folded waveguide, metal heater	Si	4 μ s	28 μ m x 67 μ m	0.9 dB	3 mW	0.2 mW·mm	2.7 dB·mW
[334] (2015)	Folded waveguide, metal heater with substrate undercut	Si	780 μ s	35 μ m x 400 μ m	3.3 dB	0.05 mW	0.02 mW·mm	0.165 dB·mW
[352] (2019)	Ring resonator, nichrome metal heater	Si	NR	6 μ m x 6 μ m	NR	0.4 mW	0.0024 mW·mm	NR
[353] (2022)	Folded waveguide, metal heater with substrate undercut	SiN	590 μ s	300 μ m x 1,200 μ m	4.8 dB	1.16 mW	1.4 mW·mm	5.57 dB·mW
[354] (2020)	Folded waveguide, metal heater with substrate undercut	SiN	20 μ s	5 μ m x 300 μ m	0.45 dB	30 mW	9 mW·mm	13.5 dB·mW

Table A.1: A survey of thermo-optic phase shifters reported in the literature. The survey is not exhaustive, and selects works with best performance in its category and with an experimental implementation (i.e., no theoretical/simulation only works are included). NR = Not reported.

Plasma dispersion silicon phase shifters

Ref. (year)	Brief description	Time constant	Footprint	IL	V_{π}	$V_{\pi}L$	$V_{\pi}L\alpha$
[355] (2017)	Depletion based, lateral pn junction	6.3 ps	10 $\mu\text{m} \times 1,000 \mu\text{m}$	1.5 dB	11.5 V	11.5 V·mm	17.25 dB·V
[356] (2017)	Depletion based, U shaped pn junction	12.2 ps	10 $\mu\text{m} \times 2,000 \mu\text{m}$	2.5 dB	2.3 V	4.6 V·mm	5.75 dB·V
[357] (2017)	Depletion based, interleaved pn junction	12.6 ps	10 $\mu\text{m} \times 1,500 \mu\text{m}$	5.1 dB	4.13 V	6.2 V·mm	21.06 dB·V
[358] (2014)	Accumulation based, vertical capacitor	5.6 ps	10 $\mu\text{m} \times 400 \mu\text{m}$	2.6 dB	5 V	2 V·mm	5.2 dB·V
[359] (2016)	Accumulation based, vertical capacitor	6.36 ps	10 $\mu\text{m} \times 500 \mu\text{m}$	1.75 dB	5.6 V	2.8 V·mm	9.8 dB·V
[360] (2014)	Injection based, lateral pin junction	NR	10 $\mu\text{m} \times 500 \mu\text{m}$	NR (>3.7 dB)	1.16 V P_{π} NR	0.58 V·mm	NR
[361] (2022)	Injection based, lateral pn junction	10 ps	10 $\mu\text{m} \times 500 \mu\text{m}$	2.5 dB	1.4 V $P_{\pi} = 13.2 \text{ mW}$	0.7 V·mm	3.5 dB·V 33 dB·mW
[109] (2019)	Injection based, lateral pin junction	56.4 ns	10 $\mu\text{m} \times 1,000 \mu\text{m}$	NR	NR $P_{\pi} = 2.3 \text{ mW}$	NR	NR
[362] (2020)	Injection based, BJT junction (3 terminal device)	1.6 ns	10 $\mu\text{m} \times 500 \mu\text{m}$	10 dB	0.12 V $P_{\pi} = 4.1 \text{ mW}$	0.06 V·mm	1.2 dB·V 41 dB·mW

Table A.2: A survey of efficient silicon phase shifters based on the plasma dispersion effect. The survey is not exhaustive, and selects works with best performance in its category and with an experimental implementation (i.e., no theoretical/simulation only works are included). NR = Not reported. BJT = Bipolar Junction Transistor.

Other phase shifters

Ref. (year)	Brief description	Time constant	Footprint	IL	V_{π}	$V_{\pi}L$	$V_{\pi}L\alpha$
[363] (2018)	Silicon - electrooptic polymer hybrid	2 ps	25 μm x 1,000 μm	7.2 dB	1 V	1 V·mm	7.2 dB·V
[337] (2019)	Plasmonic - electrooptic polymer hybrid	0.31 ps	50 μm x 20 μm	10 dB	3 V	0.06 V·mm	30 dB·V
[339] (2018)	Lithium Niobate	3.5 ps	10 μm x 20,000 μm	0.5 dB	1.4 V	28 V·mm	0.7 dB·V
[340] (2019)	Barium Titanate	6.3 ps	5 μm x 1,000 μm	0.3 dB	2.3 V	2.3 V·mm	0.69 dB·V
[343] (2021)	MEMS phase shifter	0.16 μs	45 μm x 100 μm	0.33 dB	10.7 V	0.18 V·mm	3.53 dB·V
[364] (2018)	ENZ material (ITO)	1 μs	2 μm x 32 μm	6 dB	16 V	0.52 V·mm	96 dB·V
[344] (2017)	InP	2.8 ps	2 μm x 3,600 μm	2 dB	1.5 V	5.4 V·mm	3 dB·V
[345] (2017)	InGaAsP / Si hybrid	1.6 ns	20 μm x 500 μm	0.95 dB	0.86 V	0.47 V·mm	0.81 dB·V
[347] (2022)	Liquid crystal on SiN	NR	20 μm x 500 μm	0.5 dB	0.15 V	0.075 V·mm	0.075 dB·V

Table A.3: A survey of efficient phase shifters reported in the literature not based in the thermo-optic effect or the plasma dispersion effect. The survey is not exhaustive, and selects works with best performance in its category and with an experimental implementation (i.e., no theoretical/simulation only works are included).

Bibliography

- [1] S. E. Miller, "Integrated optics: An introduction," *The Bell System Technical Journal*, vol. 48, no. 7, pp. 2059–2069, Sep. 1969.
- [2] V. Evtuhov and A. Yariv, "Gaas and gaalas devices for integrated optics ." *IEEE Transactions on Microwave Theory and Techniques*, vol. 23, no. 1, pp. 44–57, Jan 1975.
- [3] M. Smit, X. Leijtens, H. Ambrosius, E. Bente, J. van der Tol, B. Smalbrugge, T. de Vries, E.-J. Geluk, J. Bolk, R. van Veldhoven, L. Augustin, P. Thijs, D. D'Agostino, H. Rabbani, K. Lawniczuk, S. Stopinski, S. Tahvili, A. Corradi, E. Kleijn, D. Dzibrou, M. Felicetti, E. Bitincka, V. Moskalenko, J. Zhao, R. Santos, G. Gilardi, W. Yao, K. Williams, P. Stabile, P. Kuindersma, J. Pello, S. Bhat, Y. Jiao, D. Heiss, G. Roelkens, M. Wale, P. Firth, F. Soares, N. Grote, M. Schell, H. Debregeas, M. Achouche, J.-L. Gentner, A. Bakker, T. Korthorst, D. Gallagher, A. Dabbs, A. Melloni, F. Morichetti, D. Melati, A. Wonfor, R. Penty, R. Broeke, B. Musk, and D. Robbins, "An introduction to InP-based generic integration technology," *Semiconductor Science and Technology*, vol. 29, no. 8, p. 083001, jun 2014. [Online]. Available: <https://doi.org/10.1088%2F0268-1242%2F29%2F8%2F083001>
- [4] J. E. Bowers and A. Y. Liu, "A comparison of four approaches to photonic integration," in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, March 2017, pp. 1–3.
- [5] B. Jalali and S. Fathpour, "Silicon photonics," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4600–4615, Dec 2006.
- [6] T. Sharma, J. Wang, B. K. Kaushik, Z. Cheng, R. Kumar, Z. Wei, and X. Li, "Review of recent progress on silicon nitride-based photonic integrated circuits," *IEEE Access*, vol. 8, pp. 195 436–195 446, 2020.

- [7] D. Zhu, L. Shao, M. Yu, R. Cheng, B. Desiatov, C. J. Xin, Y. Hu, J. Holzgrafe, S. Ghosh, A. Shams-Ansari, E. Puma, N. Sinclair, C. Reimer, M. Zhang, and M. Lončar, "Integrated photonics on thin-film lithium niobate," *Adv. Opt. Photon.*, vol. 13, no. 2, pp. 242–352, Jun 2021. [Online]. Available: <https://opg.optica.org/aop/abstract.cfm?URI=aop-13-2-242>
- [8] I. Taghavi, M. Moridsadat, A. Tofini, S. Raza, N. A. F. Jaeger, L. Chrostowski, B. J. Shastri, and S. Shekhar, "Polymer modulators in silicon photonics: review and projections," *Nanophotonics*, vol. 11, no. 17, pp. 3855–3871, 2022. [Online]. Available: <https://doi.org/10.1515/nanoph-2022-0141>
- [9] F. Xia, H. Wang, D. Xiao, M. Dubey, and A. Ramasubramaniam, "Two-dimensional material nanophotonics," *Nature Photonics*, vol. 8, no. 12, pp. 899–907, Dec 2014. [Online]. Available: <https://doi.org/10.1038/nphoton.2014.271>
- [10] W. Jaffray, S. Saha, V. M. Shalaev, A. Boltasseva, and M. Ferrera, "Transparent conducting oxides: from all-dielectric plasmonics to a new paradigm in integrated photonics," *Adv. Opt. Photon.*, vol. 14, no. 2, pp. 148–208, Jun 2022. [Online]. Available: <https://opg.optica.org/aop/abstract.cfm?URI=aop-14-2-148>
- [11] R. A. Soref and J. P. Lorenzo, "Single-crystal silicon: a new material for 1.3 and 1.6 μm integrated-optical components," *Electronics Letters*, vol. 21, no. 21, pp. 953–954, October 1985.
- [12] G. T. Reed, W. R. Headley, and C. E. J. Png, "Silicon photonics: the early years," in *Optoelectronic Integration on Silicon II*, J. A. Kubby and G. E. Jabbour, Eds., vol. 5730, International Society for Optics and Photonics. SPIE, 2005, pp. 1 – 18. [Online]. Available: <https://doi.org/10.1117/12.596921>
- [13] G. T. Reed and A. P. Knights, *Silicon Photonics: An Introduction*, 1st ed. The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England: John Wiley & Sons Ltd, 2004.
- [14] R. Soref, "The past, present, and future of silicon photonics," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 6, pp. 1678–1687, Nov 2006.
- [15] A. L. Washburn, M. S. Luchansky, A. L. Bowman, and R. C. Bailey, "Quantitative, label-free detection of five protein biomarkers using multiplexed arrays of silicon photonic microring resonators," *Analytical Chemistry*, vol. 82, no. 1, pp. 69–72, Jan 2010. [Online]. Available: <https://doi.org/10.1021/ac902451b>

- [16] J. Leuthold, C. Koos, and W. Freude, "Nonlinear silicon photonics," *Nature Photonics*, vol. 4, no. 8, pp. 535–544, 2010. [Online]. Available: <https://doi.org/10.1038/nphoton.2010.185>
- [17] X. Qiang, X. Zhou, J. Wang, C. M. Wilkes, T. Loke, S. O’Gara, L. Kling, G. D. Marshall, R. Santagati, T. C. Ralph, J. B. Wang, J. L. O’Brien, M. G. Thompson, and J. C. F. Matthews, "Large-scale silicon quantum photonics implementing arbitrary two-qubit processing," *Nature Photonics*, vol. 12, no. 9, pp. 534–539, 2018. [Online]. Available: <https://doi.org/10.1038/s41566-018-0236-y>
- [18] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljacic, "Deep learning with coherent nanophotonic circuits," *Nature Photonics*, vol. 11, no. 7, pp. 441–446, 2017. [Online]. Available: <https://doi.org/10.1038/nphoton.2017.93>
- [19] "Silicon photonics 2022," Yole Intelligence, Tech. Rep., 2022. [Online]. Available: <https://www.yolegroup.com/product/report/silicon-photonics-2022/?cn-reloaded=1>
- [20] J. F. Song, X. S. Luo, J. S. Kee, Q. Liu, K. W. Kim, Y. Shin, M. K. Park, K. W. Ang, and G. Q. Lo, "A novel optical multiplexed, label-free bio-photonics-sensor realized on cmos-compatible optoelectronic integrated circuit (oeic) platform," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 14.4.1–14.4.4.
- [21] N. M. Fahrenkopf, C. McDonough, G. L. Leake, Z. Su, E. Timurdogan, and D. D. Coolbaugh, "The aim photonics mpw: A highly accessible cutting edge technology for rapid prototyping of photonic integrated circuits," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–6, 2019.
- [22] R. Marchetti, C. Lacava, L. Carroll, K. Gradkowski, and P. Minzioni, "Coupling strategies for silicon photonics integrated chips," *Photon. Res.*, vol. 7, no. 2, pp. 201–239, Feb 2019. [Online]. Available: <https://opg.optica.org/prj/abstract.cfm?URI=prj-7-2-201>
- [23] Z. Lin and W. Shi, "Broadband, low-loss silicon photonic y-junction with an arbitrary power splitting ratio," *Opt. Express*, vol. 27, no. 10, pp. 14 338–14 343, May 2019. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-10-14338>
- [24] N. Jovanovic, P. G. Tuthill, B. Norris, S. Gross, P. Stewart, N. Charles, S. Lacour, M. Ams, J. S. Lawrence, A. Lehmann, C. Niel, J. G. Robertson,

- G. D. Marshall, M. Ireland, A. Fuerbach, and M. J. Withford, "Starlight demonstration of the Dragonfly instrument: an integrated photonic pupil-remapping interferometer for high-contrast imaging," *Monthly Notices of the Royal Astronomical Society*, vol. 427, no. 1, pp. 806–815, 11 2012. [Online]. Available: <https://doi.org/10.1111/j.1365-2966.2012.21997.x>
- [25] Y. Huang, M. Badar, A. Nitkowski, A. Weinroth, N. Tansu, and C. Zhou, "Wide-field high-speed space-division multiplexing optical coherence tomography using an integrated photonic device," *Biomed. Opt. Express*, vol. 8, no. 8, pp. 3856–3867, Aug 2017. [Online]. Available: <https://opg.optica.org/boe/abstract.cfm?URI=boe-8-8-3856>
- [26] S. Romero-García, T. Klos, E. Klein, J. Leuermann, D. Geuzebroek, J. van Kerkhof, M. Büscher, J. Krieg, P. Leisching, and J. Witzens, "Photonic integrated circuits for multi-color laser engines," in *Silicon Photonics XII*, G. T. Reed and A. P. Knights, Eds., vol. 10108, International Society for Optics and Photonics. SPIE, 2017, p. 101080Z. [Online]. Available: <https://doi.org/10.1117/12.2250758>
- [27] T. Su, G. Liu, K. E. Badham, S. T. Thurman, R. L. Kendrick, A. Duncan, D. Wuchenich, C. Ogden, G. Chriqui, S. Feng, J. Chun, W. Lai, and S. J. B. Yoo, "Interferometric imaging using si₃n₄ photonic integrated circuits for a spider imager," *Opt. Express*, vol. 26, no. 10, pp. 12 801–12 812, May 2018. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-10-12801>
- [28] E. A. Rank, R. Sentosa, D. J. Harper, M. Salas, A. Gaugutz, D. Seyringer, S. Nevlacsil, A. Maese-Novo, M. Eggeling, P. Muellner, R. Hainberger, M. Sagmeister, J. Kraft, R. A. Leitgeb, and W. Drexler, "Toward optical coherence tomography on a chip: in vivo three-dimensional human retinal imaging using photonic integrated circuit-based arrayed waveguide gratings," *Light: Science & Applications*, vol. 10, no. 1, p. 6, Jan 2021. [Online]. Available: <https://doi.org/10.1038/s41377-020-00450-0>
- [29] N. Yu and F. Capasso, "Flat optics with designer metasurfaces," *Nature Materials*, vol. 13, no. 2, pp. 139–150, Feb 2014. [Online]. Available: <https://doi.org/10.1038/nmat3839>
- [30] A. Yulaev, W. Zhu, C. Zhang, D. A. Westly, H. J. Lezec, A. Agrawal, and V. Aksyuk, "Metasurface-integrated photonic platform for versatile free-space beam projection with polarization control," *ACS Photonics*, vol. 6, no. 11, pp. 2902–2909, 2019. [Online]. Available: <https://doi.org/10.1021/acsp Photonics.9b01000>

- [31] K. Nakamura, "Hetero-integration photonics for imaging and sensing applications," in *Imaging and Applied Optics*. Optica Publishing Group, 2013, p. JW4B.2. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=AIO-2013-JW4B.2>
- [32] A. C. Sobieranski, F. Inci, H. C. Tekin, M. Yuksekkaya, E. Comunello, D. Cobra, A. von Wangenheim, and U. Demirci, "Portable lensless wide-field microscopy imaging platform based on digital inline holography and multi-frame pixel super-resolution," *Light: Science & Applications*, vol. 4, no. 10, pp. e346–e346, Oct 2015. [Online]. Available: <https://doi.org/10.1038/lisa.2015.119>
- [33] R. S. Arvidson, C. Fischer, D. S. Sawyer, G. D. Scott, D. Natelson, and A. Lüttge, "Lateral resolution enhancement of vertical scanning interferometry by sub-pixel sampling," *Microscopy and Microanalysis*, vol. 20, no. 1, p. 90–98, 2014.
- [34] E. Prajapati, S. Kumar, and S. Kumar, "Muscope: a miniature on-chip lensless microscope," *Lab Chip*, vol. 21, pp. 4357–4363, 2021. [Online]. Available: <http://dx.doi.org/10.1039/D1LC00792K>
- [35] R. Konoike, K. Suzuki, S. Namiki, H. Kawashima, and K. Ikeda, "Ultra-compact silicon photonics switch with high-density thermo-optic heaters," *Opt. Express*, vol. 27, no. 7, pp. 10 332–10 342, Apr 2019. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-7-10332>
- [36] M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *Journal of Microscopy*, vol. 198, no. 2, pp. 82–87, 2000. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2818.2000.00710.x>
- [37] Ø. I. Helle, F. T. Dullo, M. Lahrberg, J.-C. Tinguely, O. G. Hellesø, and B. S. Ahluwalia, "Structured illumination microscopy using a photonic chip," *Nature Photonics*, vol. 14, no. 7, pp. 431–438, Jul 2020. [Online]. Available: <https://doi.org/10.1038/s41566-020-0620-2>
- [38] O. Arisev, Q. Deng, D. Kouznetsov, M. Mahmud-Ul-Hasan, P. V. Dorpe, and N. Verellen, "Experimental demonstration of structured illumination microscopy using a photonic integrated circuit," in *Optica Advanced Photonics Congress 2022*. Optica Publishing Group, 2022, p. ITu1B.5. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=IPRSN-2022-ITu1B.5>
- [39] N. Dostart, B. Zhang, M. Brand, D. Feldkhun, M. Popović, and K. Wagner, "Fourier-basis structured illumination imaging with an array of integrated

- optical phased arrays," *J. Opt. Soc. Am. A*, vol. 38, no. 10, pp. B19–B28, Oct 2021. [Online]. Available: <https://opg.optica.org/josaa/abstract.cfm?URI=josaa-38-10-B19>
- [40] D. A. Coucheron, Ø. I. Helle, C. I. Øie, F. T. Dullo, and B. S. Ahluwalia, "Chip-based nanoscopy: towards integration and high-throughput imaging," in *Nanoimaging and Nanospectroscopy V*, P. Verma and A. Egner, Eds., vol. 10350, International Society for Optics and Photonics. SPIE, 2017, p. 103500W. [Online]. Available: <https://doi.org/10.1117/12.2273902>
- [41] I. S. Opstad, D. H. Hansen, S. A. na, F. Ströhl, A. Priyadarshi, J.-C. Tinguely, F. T. Dullo, R. A. Dalmo, T. Seternes, B. S. Ahluwalia, and K. Agarwal, "Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination," *Opt. Express*, vol. 29, no. 15, pp. 23 368–23 380, Jul 2021. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-29-15-23368>
- [42] N. Zhang, K. Cicek, J. Zhu, S. Li, H. Li, M. Sorel, X. Cai, and S. Yu, "Manipulating optical vortices using integrated photonics," *Frontiers of Optoelectronics*, vol. 9, no. 2, pp. 194–205, Jun 2016. [Online]. Available: <https://doi.org/10.1007/s12200-016-0623-2>
- [43] V. Boominathan, J. T. Robinson, L. Waller, and A. Veeraraghavan, "Recent advances in lensless imaging," *Optica*, vol. 9, no. 1, pp. 1–16, Jan 2022. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-9-1-1>
- [44] A. Ozcan and E. McLeod, "Lensless imaging and sensing," *Annual Review of Biomedical Engineering*, vol. 18, no. 1, pp. 77–102, 2016, pMID: 27420569. [Online]. Available: <https://doi.org/10.1146/annurev-bioeng-092515-010849>
- [45] R. Stahl, D. Vercauteren, T. Claes, G. Vanmeerbeeck, V. Mukund, R. Jansen, J. Song, L. Hoffman, X. Rottenberg, A. Lambrechts, and L. Lagae, "Microscope-on-chip: combining lens-free microscopy with integrated photonics," in *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XIII*, D. L. Farkas, D. V. Nicolau, and R. C. Leif, Eds., vol. 9328, International Society for Optics and Photonics. SPIE, 2015, p. 93281C. [Online]. Available: <https://doi.org/10.1117/12.2077484>
- [46] A. White, P. Khial, F. Salehi, B. Hassibi, and A. Hajimiri, "A silicon photonics computational lensless active-flat-optics imaging system," *Scientific Reports*, vol. 10, no. 1, p. 1689, Feb 2020. [Online]. Available: <https://doi.org/10.1038/s41598-020-58027-1>

- [47] H. A. Clevenson, S. J. Spector, L. Benney, M. G. Moebius, J. Brown, A. Hare, A. Huang, J. Mlynarczyk, C. V. Poulton, E. Hosseini, M. R. Watts, R. Dawson, J. P. Laine, and B. F. Lane, "Incoherent light imaging using an optical phased array," *Applied Physics Letters*, vol. 116, no. 3, p. 031105, 2020. [Online]. Available: <https://doi.org/10.1063/1.5130697>
- [48] Y. Guo, Y. Guo, C. Li, H. Zhang, X. Zhou, and L. Zhang, "Integrated optical phased arrays for beam forming and steering," *Applied Sciences*, vol. 11, no. 9, 2021. [Online]. Available: <https://www.mdpi.com/2076-3417/11/9/4017>
- [49] N. Li, C. P. Ho, J. Xue, L. W. Lim, G. Chen, Y. H. Fu, and L. Y. T. Lee, "A progress review on solid-state lidar and nanophotonics-based lidar sensors," *Laser & Photonics Reviews*, p. 2100511, 2022. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.202100511>
- [50] J. He, T. Dong, and Y. Xu, "Review of photonic integrated optical phased arrays for space optical communication," *IEEE Access*, vol. 8, pp. 188 284–188 298, 2020.
- [51] W. Bishara, U. Sikora, O. Mudanyali, T.-W. Su, O. Yaglidere, S. Luckhart, and A. Ozcan, "Holographic pixel super-resolution in portable lensless on-chip microscopy using a fiber-optic array," *Lab Chip*, vol. 11, pp. 1276–1279, 2011. [Online]. Available: <http://dx.doi.org/10.1039/C0LC00684J>
- [52] X. Heng, D. Erickson, L. R. Baugh, Z. Yaqoob, P. W. Sternberg, D. Psaltis, and C. Yang, "Optofluidic microscopy—a method for implementing a high resolution optical microscope on a chip," *Lab Chip*, vol. 6, pp. 1274–1276, 2006. [Online]. Available: <http://dx.doi.org/10.1039/B604676B>
- [53] "Differences between light microscope and electron microscope," <https://microbiologyinfo.com/differences-between-light-microscope-and-electron-microscope/>.
- [54] S. Seo, T.-W. Su, D. K. Tseng, A. Erlinger, and A. Ozcan, "Lensfree holographic imaging for on-chip cytometry and diagnostics," *Lab Chip*, vol. 9, pp. 777–787, 2009. [Online]. Available: <http://dx.doi.org/10.1039/B813943A>
- [55] S. A. Lee and C. Yang, "A smartphone-based chip-scale microscope using ambient illumination," *Lab Chip*, vol. 14, pp. 3056–3063, 2014. [Online]. Available: <http://dx.doi.org/10.1039/C4LC00523F>

- [56] G. Zheng, S. A. Lee, S. Yang, and C. Yang, "Sub-pixel resolving optofluidic microscope for on-chip cell imaging," *Lab Chip*, vol. 10, pp. 3125–3129, 2010. [Online]. Available: <http://dx.doi.org/10.1039/C0LC00213E>
- [57] X. Heng, D. Erickson, L. R. Baugh, Z. Yaqoob, P. W. Sternberg, D. Psaltis, and C. Yang, "Optofluidic microscopy—a method for implementing a high resolution optical microscope on a chip," *Lab Chip*, vol. 6, pp. 1274–1276, 2006. [Online]. Available: <http://dx.doi.org/10.1039/B604676B>
- [58] Y. Wu and A. Ozcan, "Lensless digital holographic microscopy and its applications in biomedicine and environmental monitoring," *Methods*, vol. 136, pp. 4–16, 2018, methods in Quantitative Phase Imaging in Life Science. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1046202317301974>
- [59] T. Latychevskaia and H.-W. Fink, "Practical algorithms for simulation and reconstruction of digital in-line holograms," *Appl. Opt.*, vol. 54, no. 9, pp. 2424–2434, Mar 2015. [Online]. Available: <https://opg.optica.org/ao/abstract.cfm?URI=ao-54-9-2424>
- [60] *Optical Foundations of Holography*. John Wiley & Sons, Ltd, 2004, ch. 2, pp. 9–80. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/3527604154.ch2>
- [61] A. R. Thompson, J. M. Moran, and G. W. Swenson, *Van Cittert–Zernike Theorem, Spatial Coherence, and Scattering*. Cham: Springer International Publishing, 2017, pp. 767–786. [Online]. Available: https://doi.org/10.1007/978-3-319-44431-4_15
- [62] "Coherence," [https://en.wikipedia.org/wiki/Coherence_\(physics\)](https://en.wikipedia.org/wiki/Coherence_(physics)).
- [63] "The photon," <https://cronodon.com/Atomic/Photon.html>.
- [64] M. Wang, S. Feng, and J. Wu, "Multilayer pixel super-resolution lensless in-line holographic microscope with random sample movement," *Scientific Reports*, vol. 7, no. 1, p. 12791, Oct 2017. [Online]. Available: <https://doi.org/10.1038/s41598-017-13134-4>
- [65] W. Bishara, T.-W. Su, A. F. Coskun, and A. Ozcan, "Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution," *Opt. Express*, vol. 18, no. 11, pp. 11181–11191, May 2010. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-18-11-11181>

- [66] W. Luo, Y. Zhang, A. Feizi, Z. Göröcs, and A. Ozcan, "Pixel super-resolution using wavelength scanning," *Light: Science & Applications*, vol. 5, no. 4, pp. e16 060–e16 060, Apr 2016. [Online]. Available: <https://doi.org/10.1038/lssa.2016.60>
- [67] H. Wang, Z. Göröcs, W. Luo, Y. Zhang, Y. Rivenson, L. A. Bentolila, and A. Ozcan, "Computational out-of-focus imaging increases the space–bandwidth product in lens-based coherent microscopy," *Optica*, vol. 3, no. 12, pp. 1422–1429, Dec 2016. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-3-12-1422>
- [68] E. McLeod, W. Luo, O. Mudanyali, A. Greenbaum, and A. Ozcan, "Toward giga-pixel nanoscopy on a chip: a computational wide-field look at the nano-scale without the use of lenses," *Lab Chip*, vol. 13, pp. 2028–2035, 2013. [Online]. Available: <http://dx.doi.org/10.1039/C3LC50222H>
- [69] "Lyncee tec lens-less sensor," <https://www.lynceetec.com/lens-less-sensor/#tab-2>.
- [70] "Lyncee tec transmission digital holographic microscope," https://www.lynceetec.com/wp-content/uploads/2014/02/LynceeTec_DHM_Tseries_Datasheet-2020.pdf.
- [71] "Lyncee tec objectives for digital holographic microscopes," <https://www.lynceetec.com/microscope-objectives/#tab-5>.
- [72] "Ovizio iline-f holographic microscope," <https://ovizio.com/wp-content/uploads/2020/08/ILine-F-product-description.pdf>.
- [73] "Metrolaser, inc. digital holographic microscopes," <https://www.metrolaserinc.com/wp-content/uploads/2022/10/DHM-brochure.pdf>.
- [74] "Phi holomonitor," <https://phiab.com/holomonitor/>.
- [75] "Holmarc digital holographic microscope," <https://holmarc.com/dhm.php>.
- [76] "Dantec dynamics shadow sizing system," https://www.dantecdynamics.com/wp-content/uploads/2022/04/0520_v4_SS_Shadow-Sizing.pdf.
- [77] R. Won, "Eyes on super-resolution," *Nature Photonics*, vol. 3, no. 7, pp. 368–369, Jul 2009. [Online]. Available: <https://doi.org/10.1038/nphoton.2009.103>

- [78] S. Kumar, M. Mahadevappa, and P. K. Dutta, "Lensless in-line holographic microscopy with light source of low spatio-temporal coherence," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 27, no. 4, pp. 1–8, 2021.
- [79] M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, and R. G. Baraniuk, "Flatcam: Thin, lensless cameras using coded aperture and computation," *IEEE Transactions on Computational Imaging*, vol. 3, no. 3, pp. 384–397, 2017.
- [80] T. Fukui, R. Tanomura, K. Komatsu, D. Yamashita, S. Takahashi, Y. Nakano, and T. Tanemura, "Non-redundant optical phased array," *Optica*, vol. 8, no. 10, pp. 1350–1358, Oct 2021. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-8-10-1350>
- [81] H. S. Wasisto, J. D. Prades, J. Gülink, and A. Waag, "Beyond solid-state lighting: Miniaturization, hybrid integration, and applications of gan nano- and micro-leds," *Applied Physics Reviews*, vol. 6, no. 4, p. 041315, 2019. [Online]. Available: <https://doi.org/10.1063/1.5096322>
- [82] K. V. Acoleyen, W. Bogaerts, J. Jágerská, N. L. Thomas, R. Houdré, and R. Baets, "Off-chip beam steering with a one-dimensional optical phased array on silicon-on-insulator," *Opt. Lett.*, vol. 34, no. 9, pp. 1477–1479, May 2009. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-34-9-1477>
- [83] S. A. Miller, C. T. Phare, Y.-C. Chang, X. Ji, O. A. J. Gordillo, A. Mohanty, S. P. Roberts, M. C. Shin, B. Stern, M. Zadka, and M. Lipson, "512-element actively steered silicon phased array for low-power lidar," in *2018 Conference on Lasers and Electro-Optics (CLEO)*, 2018, pp. 1–2.
- [84] W. Xu, L. Zhou, L. Lu, and J. Chen, "Aliasing-free optical phased array beam-steering with a plateau envelope," *Opt. Express*, vol. 27, no. 3, pp. 3354–3368, Feb 2019. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-3-3354>
- [85] Y. Zhang, Y.-C. Ling, K. Zhang, C. Gentry, D. Sadighi, G. Whaley, J. Colosimo, P. Suni, and S. J. B. Yoo, "Sub-wavelength-pitch silicon-photonics optical phased array for large field-of-regard coherent optical beam steering," *Opt. Express*, vol. 27, no. 3, pp. 1929–1940, Feb 2019. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-3-1929>
- [86] J. Sun, E. Timurdogan, A. Yaacobi, E. S. Hosseini, and M. R. Watts, "Large-scale nanophotonic phased array," *Nature*, vol. 493, no. 7431, pp. 195–199, Jan 2013. [Online]. Available: <https://doi.org/10.1038/nature11727>

- [87] R. Fatemi, B. Abiri, A. Khachaturian, and A. Hajimiri, "High sensitivity active flat optics optical phased array receiver with a two-dimensional aperture," *Opt. Express*, vol. 26, no. 23, pp. 29 983–29 999, Nov 2018. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-23-29983>
- [88] L. Cheng, S. Mao, Z. Li, Y. Han, and H. Y. Fu, "Grating couplers on silicon photonics: Design principles, emerging trends and practical issues," *Micromachines*, vol. 11, no. 7, 2020. [Online]. Available: <https://www.mdpi.com/2072-666X/11/7/666>
- [89] C. J. R. Sheppard, "Approximate calculation of the reflection coefficient from a stratified medium," *Pure and Applied Optics: Journal of the European Optical Society Part A*, vol. 4, no. 5, p. 665, sep 1995. [Online]. Available: <https://dx.doi.org/10.1088/0963-9659/4/5/018>
- [90] W. Song, R. Gatdula, S. Abbaslou, M. Lu, A. Stein, W. Y.-C. Lai, J. Provine, R. F. W. Pease, D. N. Christodoulides, and W. Jiang, "High-density waveguide superlattices with low crosstalk," *Nature Communications*, vol. 6, no. 1, p. 7027, May 2015. [Online]. Available: <https://doi.org/10.1038/ncomms8027>
- [91] C. T. Phare, M. C. Shin, J. Sharma, S. Ahasan, H. Krishnaswamy, and M. Lipson, "Silicon optical phased array with grating lobe-free beam formation over 180 degree field of view," in *Conference on Lasers and Electro-Optics*. Optica Publishing Group, 2018, p. SM3I.2. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=CLEO_SI-2018-SM3I.2
- [92] X. Yi, Y. Zhang, H. Zeng, S. Gao, S. Guo, and C. Qiu, "Demonstration of an ultra-compact 8-channel sinusoidal silicon waveguide array for optical phased array," *Opt. Lett.*, vol. 47, no. 2, pp. 226–229, Jan 2022. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-47-2-226>
- [93] S. Wolter, H. Spende, J. Gülink, J. Hartmann, H.-H. Wehmann, A. Waag, A. Lex, A. Avramescu, H.-J. Lugauer, N. v. Malm, J.-J. Drolet, and M. Strassburg, "Size-dependent electroluminescence and current-voltage measurements of blue ingan/gan μ leds down to the submicron scale," *Nanomaterials*, vol. 11, no. 4, 2021. [Online]. Available: <https://www.mdpi.com/2079-4991/11/4/836>
- [94] S. S. Konoplev, K. A. Bulashevich, and S. Y. Karpov, "From large-size to micro-leds: Scaling trends revealed by modeling," *physica status solidi (a)*, vol. 215, no. 10, p. 1700508, 2018. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/pssa.201700508>

- [95] B. Lu, L. Wang, Z. Hao, Y. Luo, C.-J. Chen, M.-C. Wu, J. Tang, J. K. Kim, and E. F. Schubert, "Dennard scaling in optoelectronics: Scientific challenges and countermeasures of micro-leds for displays," *Laser & Photonics Reviews*, p. 2100433, 2022. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.202100433>
- [96] J. M. Smith, R. Ley, M. S. Wong, Y. H. Baek, J. H. Kang, C. H. Kim, M. J. Gordon, S. Nakamura, J. S. Speck, and S. P. DenBaars, "Comparison of size-dependent characteristics of blue and green ingan microleds down to 1 μm in diameter," *Applied Physics Letters*, vol. 116, no. 7, p. 071102, 2020. [Online]. Available: <https://doi.org/10.1063/1.5144819>
- [97] D. D. Bezshlyakh, H. Spende, T. Weimann, P. Hinze, S. Bornemann, J. Gülink, J. Canals, J. D. Prades, A. Dieguez, and A. Waag, "Directly addressable gan-based nano-led arrays: fabrication and electro-optical characterization," *Microsystems & Nanoengineering*, vol. 6, no. 1, p. 88, Oct 2020. [Online]. Available: <https://doi.org/10.1038/s41378-020-00198-y>
- [98] L. Zhang, F. Ou, W. C. Chong, Y. Chen, and Q. Li, "Wafer-scale monolithic hybrid integration of si-based ic and iii-v epi-layers—a mass manufacturable approach for active matrix micro-led micro-displays," *Journal of the Society for Information Display*, vol. 26, no. 3, pp. 137–145, 2018. [Online]. Available: <https://sid.onlinelibrary.wiley.com/doi/abs/10.1002/jsid.649>
- [99] K. Yadavalli, C.-L. Chuang, and H. S. El-Ghoroury, "Monolithic and heterogeneous integration of RGB micro-LED arrays with pixel-level optics array and CMOS image processor to enable small form-factor display applications," in *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR)*, B. C. Kress and C. Peroz, Eds., vol. 11310, International Society for Optics and Photonics. SPIE, 2020, p. 113100Z. [Online]. Available: <https://doi.org/10.1117/12.2561859>
- [100] M. Mikulics, J. Mayer, and H. H. Hardtdegen, "Cutting-edge nano-led technology," *Journal of Applied Physics*, vol. 131, no. 11, p. 110903, 2022. [Online]. Available: <https://doi.org/10.1063/5.0087279>
- [101] S. S. Deka, S. Jiang, S. H. Pan, and Y. Fainman, "Nanolaser arrays: toward application-driven dense integration," *Nanophotonics*, vol. 10, no. 1, pp. 149–169, 2021. [Online]. Available: <https://doi.org/10.1515/nanoph-2020-0372>

- [102] K. Kluczyk-Korch, S. Moreno, J. Canals, A. Diéguez, J. Gülink, J. Hartmann, A. Waag, A. Di Carlo, and M. Auf der Maur, "Individually switchable ingan/gan nano-led arrays as highly resolved illumination engines," *Electronics*, vol. 10, no. 15, 2021. [Online]. Available: <https://www.mdpi.com/2079-9292/10/15/1829>
- [103] N. Franch, J. Canals, V. Moro, A. Vilá, A. Romano-Rodríguez, J. D. Prades, J. Gülink, D. Bezshlyakh, A. Waag, K. Kluczyk-Korch, M. A. der Maur, A. di Carlo, and Ángel Diéguez, "Nano illumination microscopy: a technique based on scanning with an array of individually addressable nanoleds," *Opt. Express*, vol. 28, no. 13, pp. 19 044–19 057, Jun 2020. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-28-13-19044>
- [104] J. Canals, V. Moro, N. Franch, S. Moreno, O. Alonso, A. Vilà, J. D. Prades, J. Gülink, D. D. Bezshlyakh, A. Waag, and A. Diéguez, "A shadow image microscope based on an array of nanoLEDs," in *Unconventional Optical Imaging II*, C. Fournier, M. P. Georges, and G. Popescu, Eds., vol. 11351, International Society for Optics and Photonics. SPIE, 2020, p. 113510D. [Online]. Available: <https://doi.org/10.1117/12.2559394>
- [105] A. Vila, S. Moreno, J. Canals, and A. Diéguez, "A compact raster lensless microscope based on a microdisplay," *Sensors*, vol. 21, no. 17, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/17/5941>
- [106] A. C. Lesina, D. Goodwill, E. Bernier, L. Ramunno, and P. Berini, "On the performance of optical phased array technology for beam steering: effect of pixel limitations," *Opt. Express*, vol. 28, no. 21, pp. 31 637–31 657, Oct 2020. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-28-21-31637>
- [107] D. N. Hutchison, J. Sun, J. K. Doylend, R. Kumar, J. Heck, W. Kim, C. T. Phare, A. Feshali, and H. Rong, "High-resolution aliasing-free optical beam steering," *Optica*, vol. 3, no. 8, pp. 887–890, Aug 2016. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-3-8-887>
- [108] N. A. Tyler, D. Fowler, S. Malhouitre, S. Garcia, P. Grosse, W. Rabaud, and B. Szlag, "Sin integrated optical phased arrays for two-dimensional beam steering at a single near-infrared wavelength," *Opt. Express*, vol. 27, no. 4, pp. 5851–5858, Feb 2019. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-4-5851>
- [109] S.-H. Kim, J.-B. You, Y.-G. Ha, G. Kang, D.-S. Lee, H. Yoon, D.-E. Yoo, D.-W. Lee, K. Yu, C.-H. Youn, and H.-H. Park, "Thermo-optic control

of the longitudinal radiation angle in a silicon-based optical phased array,” *Opt. Lett.*, vol. 44, no. 2, pp. 411–414, Jan 2019. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-44-2-411>

- [110] N. Dostart, B. Zhang, A. Khilo, M. Brand, K. A. Qubaisi, D. Onural, D. Feldkhun, K. H. Wagner, and M. A. Popović, “Serpentine optical phased arrays for scalable integrated photonic lidar beam steering,” *Optica*, vol. 7, no. 6, pp. 726–733, Jun 2020. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-7-6-726>
- [111] Z. Dai, A. Wolf, P.-P. Ley, T. Glück, M. C. Sundermeier, and R. Lachmayer, “Requirements for automotive lidar systems,” *Sensors*, vol. 22, no. 19, 2022. [Online]. Available: <https://www.mdpi.com/1424-8220/22/19/7532>
- [112] C.-P. Hsu, B. Li, B. Solano-Rivas, A. R. Gohil, P. H. Chan, A. D. Moore, and V. Donzella, “A review and perspective on optical phased array for automotive lidar,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 27, no. 1, pp. 1–16, 2021.
- [113] “Quanergy,” <https://quanergy.com/>.
- [114] “Pointcloud,” <https://point.cloud/>.
- [115] “Analog photonics,” <https://www.analogphotonics.com/product/phased-array-lidar/>.
- [116] Y. Takashima and B. Hellman, “Review paper: imaging lidar by digital micromirror device,” *Optical Review*, vol. 27, no. 5, pp. 400–408, Oct 2020. [Online]. Available: <https://doi.org/10.1007/s10043-020-00620-w>
- [117] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, “Sensor and sensor fusion technology in autonomous vehicles: A review,” *Sensors*, vol. 21, no. 6, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/6/2140>
- [118] “The future of lidar for automotive applications,” <https://passive-components.eu/the-future-of-lidar-for-automotive-applications/>.
- [119] V. W. S. Chan, “Free-space optical communications,” *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4750–4762, 2006.
- [120] G. Z. and P. W. O., “Terrestrial free-space optical communications,” in *Mobile and Wireless Communications*, S. A. Fares and F. Adachi, Eds. Rijeka: IntechOpen, 2010, ch. 17. [Online]. Available: <https://doi.org/10.5772/7698>

- [121] S. V. K. Eguri, A. B. R. A, and N. Sharma, "Survey on acquisition, tracking and pointing (atp) systems and beam profile correction techniques in fso communication systems," *Journal of Optical Communications*, 2022. [Online]. Available: <https://doi.org/10.1515/joc-2021-0222>
- [122] C. V. Poulton, M. J. Byrd, P. Russo, E. Timurdogan, M. Khandaker, D. Vermeulen, and M. R. Watts, "Long-range lidar and free-space data communication with high-performance optical phased arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–8, 2019.
- [123] A. C. Foster, M. Kossey, N. MacFarlane, C. Rizk, T. Kott, R. Osiander, and N. Mosavi, "Chip-scale optical phased arrays for inter-spacecraft communications," in *Micro- and Nanotechnology Sensors, Systems, and Applications XI*, T. George and M. S. Islam, Eds., vol. 10982, International Society for Optics and Photonics. SPIE, 2019, p. 109820H. [Online]. Available: <https://doi.org/10.1117/12.2520305>
- [124] R. K. Tyson, "Bit-error rate for free-space adaptive optics laser communications," *J. Opt. Soc. Am. A*, vol. 19, no. 4, pp. 753–758, Apr 2002. [Online]. Available: <https://opg.optica.org/josaa/abstract.cfm?URI=josaa-19-4-753>
- [125] C. E. Max, "Adaptive optics: An introduction," https://www.ucolick.org/~max/289/Assigned%20Readings/Max_Adaptive_Optics_Intro_v1.pdf.
- [126] E. Leitgeb, M. Gebhart, and U. Birnbacher, "Optical networks, last mile access and applications," *Journal of Optical and Fiber Communications Reports*, vol. 2, no. 1, pp. 56–85, Mar 2005. [Online]. Available: <https://doi.org/10.1007/s10297-004-0025-x>
- [127] S. Devi and L. Govindaraju, "A study on wind induced vibration on lighting poles," *International Journal of Research in Engineering and Technology*, vol. 03, pp. 172–178, 2014.
- [128] S. Arnon, "Effects of atmospheric turbulence and building sway on optical wireless-communication systems," *Opt. Lett.*, vol. 28, no. 2, pp. 129–131, Jan 2003. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-28-2-129>
- [129] F. Dios, J. A. Rubio, A. Rodríguez, and A. Comerón, "Scintillation and beam-wander analysis in an optical ground station-satellite uplink," *Appl. Opt.*, vol. 43, no. 19, pp. 3866–3873, Jul 2004. [Online]. Available: <https://opg.optica.org/ao/abstract.cfm?URI=ao-43-19-3866>

- [130] R. Fante, "Electromagnetic beam propagation in turbulent media," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1669–1692, 1975.
- [131] J. H. Churnside and R. J. Lataitis, "Wander of an optical beam in the turbulent atmosphere," *Appl. Opt.*, vol. 29, no. 7, pp. 926–930, Mar 1990. [Online]. Available: <https://opg.optica.org/ao/abstract.cfm?URI=ao-29-7-926>
- [132] R. Manning, "Beam wave considerations for optical link budget calculations," *NASA Technical Memorandum*, Sept 2016.
- [133] M. Arikawa, T. Ishikawa, K. Hosokawa, S. Takahashi, Y. Ono, and T. Ito, "Mitigation of fading caused by atmospheric turbulence with fmf coupling and maximum ratio combining used in 320-m free-space optical transmission of," in *ECOC 2016; 42nd European Conference on Optical Communication*, 2016, pp. 1–3.
- [134] H. Liu, B. Huang, J. C. A. Zacarias, H. Wen, H. Chen, N. K. Fontaine, R. Ryf, J. E. Antonio-Lopez, R. A. Correa, and G. Li, "Turbulence-resistant fso communication using a few-mode pre-amplified receiver," *Scientific Reports*, vol. 9, no. 1, p. 16247, Nov 2019. [Online]. Available: <https://doi.org/10.1038/s41598-019-52698-1>
- [135] D. J. Geisler, T. M. Yarnall, G. Lund, C. M. Schieler, M. L. Stevens, N. K. Fontaine, B. S. Robinson, and S. A. Hamilton, "Experimental comparison of 3-mode and single-mode coupling over a 1.6-km free-space link," in *Free-Space Laser Communication and Atmospheric Propagation XXX*, H. Hemmati and D. M. Boroson, Eds., vol. 10524, International Society for Optics and Photonics. SPIE, 2018, p. 105240H. [Online]. Available: <https://doi.org/10.1117/12.2294039>
- [136] T. A. Birks, I. Gris-Sánchez, S. Yerolatsitis, S. G. Leon-Saval, and R. R. Thomson, "The photonic lantern," *Adv. Opt. Photon.*, vol. 7, no. 2, pp. 107–167, Jun 2015. [Online]. Available: <https://opg.optica.org/aop/abstract.cfm?URI=aop-7-2-107>
- [137] I. Spaleniak, N. Jovanovic, S. Gross, M. J. Ireland, J. S. Lawrence, and M. J. Withford, "Integrated photonic building blocks for next-generation astronomical instrumentation ii: the multimode to single mode transition," *Opt. Express*, vol. 21, no. 22, pp. 27 197–27 208, Nov 2013. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-21-22-27197>
- [138] B. Zhang, R. Yuan, J. Sun, J. Cheng, and M.-S. Alouini, "Free-space optical communication using non-mode-selective photonic lantern-based coherent receiver," *IEEE Transactions on Communications*, vol. 69, no. 8, pp. 5367–5380, 2021.

- [139] I. Ozdur, P. Toliver, and T. K. Woodward, "Photonic-lantern-based coherent lidar system," *Opt. Express*, vol. 23, no. 4, pp. 5312–5316, Feb 2015. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-23-4-5312>
- [140] B. Zhang, J. Sun, C. Lao, C. H. Betters, A. Argyros, Y. Zhou, and S. G. Leon-Saval, "All-fiber photonic lantern multimode optical receiver with coherent adaptive optics beam combining," 2021. [Online]. Available: <https://arxiv.org/abs/2105.09516>
- [141] R. MacVoy, E. Sheffield, and J. Tartaglia, "Wind-Induced Vibration and the Effects on Steel and Aluminum Light Poles," DWM Holdings, Tech. Rep., Jan 2016. [Online]. Available: https://lytepoles.com/wp-content/uploads/2016/06/Wind-Induced-Vibration_WhitePaper_Final.pdf
- [142] J. Davis and W. Tango, "Measurement of the atmospheric coherence time," *Publications of the Astronomical Society of the Pacific*, vol. 108, no. 723, p. 456, may 1996. [Online]. Available: <https://dx.doi.org/10.1086/133747>
- [143] X. Mu, S. Wu, L. Cheng, and H. Fu, "Edge couplers in silicon photonic integrated circuits: A review," *Applied Sciences*, vol. 10, no. 4, 2020. [Online]. Available: <https://www.mdpi.com/2076-3417/10/4/1538>
- [144] J. C. Hulme, J. K. Doylend, M. J. R. Heck, J. D. Peters, M. L. Davenport, J. T. Bovington, L. A. Coldren, and J. E. Bowers, "Fully integrated hybrid silicon two dimensional beam scanner," *Opt. Express*, vol. 23, no. 5, pp. 5861–5874, Mar 2015. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-23-5-5861>
- [145] J. Notaros, N. Li, C. V. Poulton, Z. Su, M. J. Byrd, E. S. Magden, E. Timurdogan, C. Baiocco, N. M. Fahrenkopf, and M. R. Watts, "Cmos-compatible optical phased array powered by a monolithically-integrated erbium laser," *Journal of Lightwave Technology*, vol. 37, no. 24, pp. 5982–5987, 2019.
- [146] W. Guo, P. R. A. Binetti, C. Althouse, M. L. Mašanović, H. P. M. M. Ambrosius, L. A. Johansson, and L. A. Coldren, "Two-dimensional optical beam steering with inp-based photonic integrated circuits," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 4, pp. 6 100 212–6 100 212, 2013.
- [147] D. Kwong, A. Hosseini, J. Covey, Y. Zhang, X. Xu, H. Subbaraman, and R. T. Chen, "On-chip silicon optical phased array for two-dimensional beam steering," *Opt. Lett.*, vol. 39, no. 4, pp. 941–944, Feb 2014. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-39-4-941>

- [148] Y. Liu and H. Hu, "Silicon optical phased array with a 180-degree field of view for 2d optical beam steering," *Optica*, vol. 9, no. 8, pp. 903–907, Aug 2022. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-9-8-903>
- [149] A. Hosseini, D. Kwong, Y. Zhao, Y.-S. Chen, F. Crnogorac, R. F. W. Pease, and R. T. Chen, "Unequally spaced waveguide arrays for silicon nanomembrane-based efficient large angle optical beam steering," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 15, no. 5, pp. 1439–1446, 2009.
- [150] D. Kwong, A. Hosseini, J. Covey, X. Xu, Y. Zhang, S. Chakravarty, and R. T. Chen, "Corrugated waveguide-based optical phased array with crosstalk suppression," *IEEE Photonics Technology Letters*, vol. 26, no. 10, pp. 991–994, 2014.
- [151] J. A. Baxter, D. A. Merced, D. J. Costinett, L. M. Tolbert, and B. Ozpineci, "Review of electrical architectures and power requirements for automated vehicles," in *2018 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2018, pp. 944–949.
- [152] Y. Ismail, "High-Voltage Generation and Drive in Low-Voltage CMOS Technology," Master's thesis, University of California, Los Angeles, 2015. [Online]. Available: <https://escholarship.org/uc/item/7g1014zm>
- [153] F. Ashtiani and F. Aflatouni, " $N \times n$ optical phased array with $2n$ phase shifters," *Opt. Express*, vol. 27, no. 19, pp. 27 183–27 190, Sep 2019. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-19-27183>
- [154] C. V. Poulton, M. J. Byrd, B. Moss, E. Timurdogan, R. Millman, and M. R. Watts, "8192-element optical phased array with 100° steering range and flip-chip cmos," in *Conference on Lasers and Electro-Optics*. Optica Publishing Group, 2020, p. JTh4A.3. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=CLEO_SI-2020-JTh4A.3
- [155] Y. Wang, G. Zhou, X. Zhang, K. Kwon, P.-A. Blanche, N. Triesault, K. sik Yu, and M. C. Wu, "2d broadband beamsteering with large-scale mems optical phased array," *Optica*, vol. 6, no. 5, pp. 557–562, May 2019. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-6-5-557>
- [156] M. C. Shin, A. Mohanty, K. Watson, G. R. Bhatt, C. T. Phare, S. A. Miller, M. Zadka, B. S. Lee, X. Ji, I. Datta, and M. Lipson, "Chip-scale blue light phased array," *Opt. Lett.*, vol. 45, no. 7, pp. 1934–1937, Apr 2020. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-45-7-1934>

- [157] C. T. Phare, M. C. Shin, S. A. Miller, B. Stern, and M. Lipson, "Silicon optical phased array with high-efficiency beam formation over 180 degree field of view," 2018. [Online]. Available: <https://arxiv.org/abs/1802.04624>
- [158] Y. Hirano, Y. Miyamoto, M. Miura, Y. Motoyama, K. Machida, T. Yamada, A. Otomo, and H. Kikuchi, "High-speed optical-beam scanning by an optical phased array using electro-optic polymer waveguides," *IEEE Photonics Journal*, vol. 12, no. 2, pp. 1–7, 2020.
- [159] H. Zhang, Z. Zhang, J. Lv, C. Peng, and W. Hu, "Fast beam steering enabled by a chip-scale optical phased array with 8×8 elements," *Optics Communications*, vol. 461, p. 125267, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S003040182030016X>
- [160] F. Aflatouni, B. Abiri, A. Rekhi, and A. Hajimiri, "Nanophotonic projection system," *Opt. Express*, vol. 23, no. 16, pp. 21 012–21 022, Aug 2015. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-23-16-21012>
- [161] R. Fatemi, A. Khachaturian, and A. Hajimiri, "A nonuniform sparse 2-d large-fov optical phased array with a low-power pwm drive," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 5, pp. 1200–1215, 2019.
- [162] S. Chung, H. Abediasl, and H. Hashemi, "A monolithically integrated large-scale optical phased array in silicon-on-insulator cmos," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 1, pp. 275–296, 2018.
- [163] S. A. Miller, C. T. Phare, Y.-C. Chang, X. Ji, O. A. J. Gordillo, A. Mohanty, S. P. Roberts, M. C. Shin, B. Stern, M. Zadka, and M. Lipson, "512-element actively steered silicon phased array for low-power lidar," in *2018 Conference on Lasers and Electro-Optics (CLEO)*, 2018, pp. 1–2.
- [164] W. Xie, T. Komljenovic, J. Huang, M. L. Davenport, and J. E. Bowers, "Dense iii-v/si phase-shifter based optical phased array," 2019. [Online]. Available: <https://arxiv.org/abs/1904.01104>
- [165] P. Wang, G. Luo, Y. Xu, Y. Li, Y. Su, J. Ma, R. Wang, Z. Yang, X. Zhou, Y. Zhang, and J. Pan, "Design and fabrication of a sin-si dual-layer optical phased array chip," *Photon. Res.*, vol. 8, no. 6, pp. 912–919, Jun 2020. [Online]. Available: <https://opg.optica.org/prj/abstract.cfm?URI=prj-8-6-912>

- [166] S. A. Miller, Y.-C. Chang, C. T. Phare, M. C. Shin, M. Zadka, S. P. Roberts, B. Stern, X. Ji, A. Mohanty, O. A. J. Gordillo, U. D. Dave, and M. Lipson, "Large-scale optical phased array using a low-power multi-pass silicon photonic platform," *Optica*, vol. 7, no. 1, pp. 3–6, Jan 2020. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-7-1-3>
- [167] Z. Zhang, H. Yu, Q. Huang, Z. Zhou, B. Chen, T. Dai, H. Qiu, Y. Wang, and J. Yang, "High-speed and low-power silicon optical phased array based on the carrier-depletion mechanism," *IEEE Photonics Technology Letters*, vol. 34, no. 5, pp. 271–274, 2022.
- [168] L. Zhang, Y. Li, B. Chen, Y. Wang, H. Li, Y. Hou, M. Tao, Y. Li, Z. Zhi, X. Liu, X. Li, Q. Na, Q. Xie, M. Zhang, X. Li, F. Gao, X. Luo, G.-Q. Lo, and J. Song, "Two-dimensional multi-layered sin-on-soi optical phased array with wide-scanning and long-distance ranging," *Opt. Express*, vol. 30, no. 4, pp. 5008–5018, Feb 2022. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-30-4-5008>
- [169] J. Notaros, M. Notaros, M. Raval, and M. R. Watts, "Liquid-crystal-based visible-light integrated optical phased arrays," in *Conference on Lasers and Electro-Optics*. Optica Publishing Group, 2019, p. STu3O.3. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=CLEO_SI-2019-STu3O.3
- [170] Y. Deng and D. Chu, "Coherence properties of different light sources and their effect on the image sharpness and speckle of holographic displays," *Scientific Reports*, vol. 7, no. 1, p. 5893, Jul 2017. [Online]. Available: <https://doi.org/10.1038/s41598-017-06215-x>
- [171] J. Xue, "A small, bright silicon light-emitting diode directly integrated with microelectronics," Master's thesis, Massachusetts Institute of Technology, 2021. [Online]. Available: <https://dspace.mit.edu/handle/1721.1/130779>
- [172] Z. Li, J. Xue, M. de Cea, J. Kim, H. Nong, D. Chong, K. Y. Lim, E. Quek, and R. J. Ram, "A sub-wavelength si led integrated in a cmos platform," *Nature Communications*, vol. 14, no. 1, p. 882, Feb 2023. [Online]. Available: <https://doi.org/10.1038/s41467-023-36639-1>
- [173] J. H. Stathis, "Percolation models for gate oxide breakdown," *Journal of Applied Physics*, vol. 86, no. 10, pp. 5757–5766, 1999. [Online]. Available: <https://doi.org/10.1063/1.371590>

- [174] E. Miranda and J. Suñé, "Electron transport through broken down ultra-thin sio2 layers in mos devices," *Microelectronics Reliability*, vol. 44, no. 1, pp. 1–23, 2004. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0026271403003871>
- [175] N. Akil, V. E. Houtsma, P. LeMinh, J. Holleman, V. Zieren, D. de Mooij, P. H. Woerlee, A. van den Berg, and H. Wallinga, "Modeling of light-emission spectra measured on silicon nanometer-scale diode antifuses," *Journal of Applied Physics*, vol. 88, no. 4, pp. 1916–1922, 2000. [Online]. Available: <https://doi.org/10.1063/1.1305552>
- [176] J. Huijts, S. Fernandez, D. Gauthier, M. Kholodtsova, A. Maghraoui, K. Medjoubi, A. Somogyi, W. Boutu, and H. Merdji, "Broadband coherent diffractive imaging," *Nature Photonics*, vol. 14, no. 10, pp. 618–622, Oct 2020. [Online]. Available: <https://doi.org/10.1038/s41566-020-0660-7>
- [177] B. Abbey, L. W. Whitehead, H. M. Quiney, D. J. Vine, G. A. Cadenazzi, C. A. Henderson, K. A. Nugent, E. Balaur, C. T. Putkunz, A. G. Peele, G. J. Williams, and I. McNulty, "Lensless imaging using broadband x-ray sources," *Nature Photonics*, vol. 5, no. 7, pp. 420–424, Jul 2011. [Online]. Available: <https://doi.org/10.1038/nphoton.2011.125>
- [178] J. Wu, L. Cao, and G. Barbastathis, "Dnn-fza camera: a deep learning approach toward broadband fza lensless imaging," *Opt. Lett.*, vol. 46, no. 1, pp. 130–133, Jan 2021. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-46-1-130>
- [179] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.
- [180] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [181] I. Kang, M. de Cea, J. Xue, Z. Li, G. Barbastathis, and R. J. Ram, "Simultaneous spectral recovery and cmos micro-led holography with an untrained deep neural network," *Optica*, vol. 9, no. 10, pp. 1149–1155, Oct 2022. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-9-10-1149>
- [182] M. D. Feit and J. A. Fleck, "Beam nonparaxiality, filament formation, and beam breakup in the self-focusing of optical beams," *J. Opt. Soc. Am. B*, vol. 5, no. 3, pp. 633–640, 1988.

- [183] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [184] A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Phys. Rev. Lett.*, vol. 121, no. 24, p. 243902, 2018.
- [185] S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," *Optica*, vol. 5, no. 7, pp. 803–813, 2018.
- [186] M. Lee, O. Yaglidere, and A. Ozcan, "Field-portable reflection and transmission microscopy based on lensless holography," *Biomed. Opt. Express*, vol. 2, no. 9, pp. 2721–2730, Sep 2011. [Online]. Available: <https://opg.optica.org/boe/abstract.cfm?URI=boe-2-9-2721>
- [187] F. Wang, D. Wang, J. Cao, B. Que, L. Wang, S. Yuan, X. Wang, F. Peng, and Z. Tong, "Reflection in-line and off-axis hybrid digital holography," *Optics Communications*, vol. 459, p. 124909, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0030401819310132>
- [188] J. A. Picazo-Bueno, J. Garcia, and V. Mico, "Single-element reflective digital holographic microscopy," *Frontiers in Physics*, vol. 9, 2021. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fphy.2021.639607>
- [189] M. Brown, D. Burschka, and G. Hager, "Advances in computational stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 993–1008, 2003.
- [190] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, 2003.
- [191] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Opt. Lett.*, vol. 33, no. 2, pp. 156–158, Jan 2008. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-33-2-156>
- [192] J. Zhang, J. Sun, Q. Chen, J. Li, and C. Zuo, "Adaptive pixel-super-resolved lensfree in-line digital holography for wide-field on-chip microscopy," *Scientific Reports*, vol. 7, no. 1, p. 11777, Sep 2017. [Online]. Available: <https://doi.org/10.1038/s41598-017-11715-x>

- [193] R. Hardie, K. Barnard, and E. Armstrong, "Joint map registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, 1997.
- [194] G. D. Evangelidis and E. Z. Psarakis, "Parametric image alignment using enhanced correlation coefficient maximization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1858–1865, 2008.
- [195] "Opencv enhanced cross correlation implementation," https://docs.opencv.org/3.4/dc/d6b/group__video__track.html#ga1aa357007eaec11e9ed03500ecbcbe47.
- [196] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, 2003.
- [197] P. Vandewalle, S. Süsstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, no. 1, p. 071459, Dec 2006. [Online]. Available: <https://doi.org/10.1155/ASP/2006/71459>
- [198] "Superresolution v2.0 github repository," https://github.com/edwardtoday/PolyU_MScST/tree/master/COMP5422/Lab/superresolution_v_2.0.
- [199] A. Tekalp, M. Ozkan, and M. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," in *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, 1992, pp. 169–172 vol.3.
- [200] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, no. 3, pp. 231–239, 1991. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/104996529190045L>
- [201] B. Jacob Maiseli, N. Ally, and H. Gao, "A noise-suppressing and edge-preserving multiframe super-resolution image reconstruction method," *Signal Processing: Image Communication*, vol. 34, pp. 1–13, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0923596515000363>
- [202] T. Q. Pham, L. J. van Vliet, and K. Schutte, "Robust fusion of irregularly sampled data using adaptive normalized convolution," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, no. 1, p. 083268, Dec 2006. [Online]. Available: <https://doi.org/10.1155/ASP/2006/83268>

- [203] A. Greenbaum and A. Ozcan, "Maskless imaging of dense samples using pixel super-resolution based multi-height lensfree on-chip microscopy," *Opt. Express*, vol. 20, no. 3, pp. 3129–3143, Jan 2012. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-20-3-3129>
- [204] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [205] Z. Wang, A. Abbasi, U. Dave, A. De Groote, S. Kumari, B. Kunert, C. Merckling, M. Pantouvaki, Y. Shi, B. Tian, K. Van Gasse, J. Verbist, R. Wang, W. Xie, J. Zhang, Y. Zhu, J. Bauwelinck, X. Yin, Z. Hens, J. Van Campenhout, B. Kuyken, R. Baets, G. Morthier, D. Van Thourhout, and G. Roelkens, "Novel light source integration approaches for silicon photonics," *Laser & Photonics Reviews*, vol. 11, no. 4, p. 1700063, 2017. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.201700063>
- [206] S. Libertino, S. Coffa, and M. Saggio, "Design and fabrication of integrated si-based optoelectronic devices," *Materials Science in Semiconductor Processing*, vol. 3, no. 5, pp. 375–381, 2000, european Materials Research Society Spring Meeting Symposium. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1369800100000585>
- [207] S. Buckley, J. Chiles, A. N. McCaughan, G. Moody, K. L. Silverman, M. J. Stevens, R. P. Mirin, S. W. Nam, and J. M. Shainline, "All-silicon light-emitting diodes waveguide-integrated with superconducting single-photon detectors," *Applied Physics Letters*, vol. 111, no. 14, p. 141101, 2017. [Online]. Available: <https://doi.org/10.1063/1.4994692>
- [208] A. N. Tait, S. M. Buckley, J. Chiles, A. N. McCaughan, S. Olson, S. P. Rao, S. W. Nam, R. P. Mirin, and J. M. Shainline, "Microring resonator-coupled photoluminescence from silicon W centers," *Journal of Physics: Photonics*, vol. 2, no. 4, p. 045001, jul 2020. [Online]. Available: <https://doi.org/10.1088/2515-7647/ab95f2>
- [209] "Aim photonics base active pic platform," <https://www.aimphotonics.com/base-active-pic>.
- [210] M. Rakowski, C. Meagher, K. Nummy, A. Aboketaf, J. Ayala, Y. Bian, B. Harris, K. Mclean, K. McStay, A. Sahin, L. Medina, B. Peng, Z. Sowinski, A. Stricker, T. Houghton, C. Hedges, K. Giewont, A. Jacob,

- T. Letavic, D. Riggs, A. Yu, and J. Pellerin, "45nm cmos - silicon photonics monolithic technology (45clo) for next-generation, low power and high speed optical interconnects," in *Optical Fiber Communication Conference (OFC) 2020*. Optica Publishing Group, 2020, p. T3H.3. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=OFC-2020-T3H.3>
- [211] M. de Cea, A. H. Atabaki, L. Alloatti, M. Wade, M. Popovic, and R. J. Ram, "A thin silicon photonic platform for telecommunication wavelengths," in *2017 European Conference on Optical Communication (ECOC)*, Sep. 2017, pp. 1–3.
- [212] L. Alloatti, D. Cheian, and R. J. Ram, "High-speed modulator with interleaved junctions in zero-change CMOS photonics," *Applied Physics Letters*, vol. 108, no. 13, p. 131101, 2016.
- [213] V. Stojanović, R. J. Ram, M. Popović, S. Lin, S. Moazeni, M. Wade, C. Sun, L. Alloatti, A. Atabaki, F. Pavanello, N. Mehta, and P. Bhargava, "Monolithic silicon-photonics platforms in state-of-the-art cmos soi processes," *Opt. Express*, vol. 26, no. 10, pp. 13 106–13 121, May 2018. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-10-13106>
- [214] C. Sun, M. T. Wade, Y. Lee, J. S. Orcutt, L. Alloatti, M. S. Georgas, A. S. Waterman, J. M. Shainline, R. R. Avizienis, S. Lin, B. R. Moss, R. Kumar, F. Pavanello, A. H. Atabaki, H. M. Cook, A. J. Ou, J. C. Leu, Y.-H. Chen, K. Asanović, R. J. Ram, M. A. Popović, and V. M. Stojanović, "Single-chip microprocessor that communicates directly using light," *Nature*, vol. 528, no. 7583, pp. 534–538, Dec 2015. [Online]. Available: <https://doi.org/10.1038/nature16454>
- [215] S. Karpov, "ABC-model for interpretation of internal quantum efficiency and its droop in III-nitride LEDs," in *Numerical Simulation of Optoelectronic Devices, 2014*, 2014, pp. 17–18.
- [216] H. T. Nguyen, S. C. Baker-Finch, and D. Macdonald, "Temperature dependence of the radiative recombination coefficient in crystalline silicon from spectral photoluminescence," *Applied Physics Letters*, vol. 104, no. 11, p. 112105, 2014. [Online]. Available: <https://doi.org/10.1063/1.4869295>
- [217] T. Trupke, M. A. Green, P. Würfel, P. P. Altermatt, A. Wang, J. Zhao, and R. Corkish, "Temperature dependence of the radiative recombination coefficient of intrinsic crystalline silicon," *Journal of Applied Physics*, vol. 94, no. 8, pp. 4930–4937, 2003. [Online]. Available: <https://aip.scitation.org/doi/abs/10.1063/1.1610231>

- [218] J. Dziewior and W. Schmid, "Auger coefficients for highly doped and highly excited silicon," *Applied Physics Letters*, vol. 31, no. 5, pp. 346–348, 1977. [Online]. Available: <https://doi.org/10.1063/1.89694>
- [219] J. Xue, J. Kim, A. Mestre, K. M. Tan, D. Chong, S. Roy, H. Nong, K. Y. Lim, D. Gray, D. Kramnik, A. Atabaki, E. Quek, and R. J. Ram, "Low-voltage, high-brightness silicon micro-LEDs for CMOS photonics," *IEEE Transactions on Electron Devices*, vol. 68, no. 8, pp. 3870–3875, 2021.
- [220] "Sentaurus TCAD," <https://www.synopsys.com/silicon/tcad.html>.
- [221] A. W. Stephens, A. G. Aberle, and M. A. Green, "Surface recombination velocity measurements at the silicon–silicon dioxide interface by microwave-detected photoconductance decay," *Journal of Applied Physics*, vol. 76, no. 1, pp. 363–370, 1994. [Online]. Available: <https://doi.org/10.1063/1.357082>
- [222] M. de Cea, D. Van Orden, J. Fini, M. Wade, and R. J. Ram, "High-speed, zero-biased silicon-germanium photodetector," *APL Photonics*, vol. 6, no. 4, p. 041302, 2021. [Online]. Available: <https://doi.org/10.1063/5.0047037>
- [223] K. Bushick and E. Kioupakis, "Phonon-assisted Auger-Meitner recombination in silicon from first principles," *Phys. Rev. Lett.*, vol. 131, p. 076902, Aug 2023. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.131.076902>
- [224] "Ansys Lumerical FDTD," <https://www.ansys.com/products/photonics/fdtd>.
- [225] S. Lin, Y. Kobayashi, Y. Ishikawa, and K. Wada, "Luminescence enhancement by Si ring resonator structures on silicon on insulator," *Applied Physics Letters*, vol. 92, no. 2, p. 021113, 2008. [Online]. Available: <https://doi.org/10.1063/1.2835051>
- [226] A. Schenk, "Finite-temperature full random-phase approximation model of band gap narrowing for silicon device simulation," *Journal of Applied Physics*, vol. 84, no. 7, pp. 3684–3695, 10 1998. [Online]. Available: <https://doi.org/10.1063/1.368545>
- [227] S. Yalamanchili, N. S. Lewis, and H. A. Atwater, "Role of doping dependent radiative and non-radiative recombination in determining the limiting efficiencies of silicon solar cells," in *2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC)*, 2018, pp. 3223–3226.

- [228] L. Alloatti, S. A. Srinivasan, J. S. Orcutt, and R. J. Ram, "Waveguide-coupled detector in zero-change complementary metal-oxide-semiconductor," *Applied Physics Letters*, vol. 107, no. 4, p. 041104, 2015. [Online]. Available: <https://doi.org/10.1063/1.4927393>
- [229] C. T. DeRose, D. C. Trotter, W. A. Zortman, A. L. Starbuck, M. Fisher, M. R. Watts, and P. S. Davids, "Ultra compact 45 GHz CMOS compatible germanium waveguide photodiode with low dark current," *Opt. Express*, vol. 19, no. 25, pp. 24897–24904, Dec 2011. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-19-25-24897>
- [230] M. Piels and J. E. Bowers, "1 - photodetectors for silicon photonic integrated circuits," in *Photodetectors*, B. Nabet, Ed. Woodhead Publishing, 2016, pp. 3–20. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9781782424451000014>
- [231] S. A. Srinivasan, J. Lambrecht, M. Berciano, S. Lardenois, P. Absil, J. Bauwelinck, X. Yin, M. Pantouvaki, and J. Van Campenhout, "Highly sensitive 56 Gbps NRZ O-band biCMOS-silicon photonics receiver using a Ge/Si avalanche photodiode," in *2020 Optical Fiber Communications Conference and Exhibition (OFC)*, 2020, pp. 1–3.
- [232] S. Grillanda and F. Morichetti, "Light-induced metal-like surface of silicon photonic waveguides," *Nature Communications*, vol. 6, no. 1, p. 8182, Sep 2015. [Online]. Available: <https://doi.org/10.1038/ncomms9182>
- [233] A. Gil-Molina, I. Aldaya, J. L. Pita, L. H. Gabrielli, H. L. Fragnito, and P. Dainese, "Optical free-carrier generation in silicon nano-waveguides at 1550 nm," *Applied Physics Letters*, vol. 112, no. 25, p. 251104, 2018. [Online]. Available: <https://doi.org/10.1063/1.5023589>
- [234] M. de Cea, A. H. Atabaki, and R. J. Ram, "Energy harvesting optical modulators with sub-attojoule per bit electrical energy consumption," *Nature Communications*, vol. 12, no. 1, p. 2326, Apr 2021. [Online]. Available: <https://doi.org/10.1038/s41467-021-22460-1>
- [235] A. Joshi, C. Batten, Y.-J. Kwon, S. Beamer, I. Shamim, K. Asanovic, and V. Stojanovic, "Silicon-photonic networks for global on-chip communication," in *2009 3rd ACM/IEEE International Symposium on Networks-on-Chip*, 2009, pp. 124–133.

- [236] E. Luan, H. Shoman, D. M. Ratner, K. C. Cheung, and L. Chrostowski, "Silicon photonic biosensors using label-free detection," *Sensors*, vol. 18, no. 10, 2018. [Online]. Available: <https://www.mdpi.com/1424-8220/18/10/3519>
- [237] B. van Driehuisen and R. Wolffenbuttel, "Optocoupler based on the avalanche light emission in silicon," *Sensors and Actuators A: Physical*, vol. 31, no. 1, pp. 229–240, 1992, proceedings of Eurosensors V. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0924424792801100>
- [238] L. Rebohle, J. von Borany, D. Borchert, H. Frob, T. Gebel, M. Helm, W. Moller, and W. Skorupa, "Efficient blue light emission from silicon: The first integrated si-based optocoupler," *Electrochemical and Solid-State Letters*, vol. 4, no. 7, p. G57, 2001. [Online]. Available: <https://doi.org/10.1149/1.1375005>
- [239] F. Acerbi, N. Massari, L. Gasparini, A. Tomasi, N. Zorzi, G. Fontana, L. Pavesi, and A. Gola, "Structures and methods for fully-integrated quantum random number generators," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 3, pp. 1–8, 2020.
- [240] Y. Chen, J. Liu, Z. Yang, J. S. Wilkinson, and X. Zhou, "Optical biosensors based on refractometric sensing schemes: A review," *Biosensors and Bioelectronics*, vol. 144, p. 111693, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0956566319307729>
- [241] M. Estevez, M. Alvarez, and L. Lechuga, "Integrated optical devices for lab-on-a-chip biosensing applications," *Laser & Photonics Reviews*, vol. 6, no. 4, pp. 463–487, 2012. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.201100025>
- [242] X. Fan, I. M. White, S. I. Shopova, H. Zhu, J. D. Suter, and Y. Sun, "Sensitive optical biosensors for unlabeled targets: A review," *Analytica Chimica Acta*, vol. 620, no. 1, pp. 8–26, 2008. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0003267008009343>
- [243] C. A. Barrios, "Integrated microring resonator sensor arrays for labs-on-chips," *Analytical and Bioanalytical Chemistry*, vol. 403, no. 6, pp. 1467–1475, Jun 2012. [Online]. Available: <https://doi.org/10.1007/s00216-012-5937-3>
- [244] M. C. Cardenosa-Rubio, H. M. Robison, and R. C. Bailey, "Recent advances in environmental and clinical analysis using microring resonator-based sensors," *Current Opinion in Environmental Science & Health*, vol. 10,

- pp. 38–46, 2019, environmental Monitoring Assessment: Biosensors for environmental monitoring and human health. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2468584418300813>
- [245] “Genalyte,” <https://www.genalyte.com/>.
- [246] “Resosens bioassay system,” <http://www.resonantsensors.com/resosens-bioassay-system.php>.
- [247] “Siphox,” <https://siphox.com/>.
- [248] “Optiqua,” <https://optiqua.com/>.
- [249] “Corning epic,” <https://www.corning.com/media/worldwide/cls/documents/CLS-ES-038REV3.pdf>.
- [250] “Creoptix wave,” <https://www.creoptix.com/>.
- [251] M. S. Luchansky and R. C. Bailey, “Silicon photonic microring resonators for quantitative cytokine detection and t-cell secretion analysis,” *Analytical Chemistry*, vol. 82, no. 5, pp. 1975–1981, 2010, pMID: 20143780. [Online]. Available: <https://doi.org/10.1021/ac902725q>
- [252] A. Székács, N. Trummer, N. Adányi, M. Váradi, and I. Szendrő, “Development of a non-labeled immunosensor for the herbicide trifluralin via optical waveguide lightmode spectroscopic detection,” *Analytica Chimica Acta*, vol. 487, no. 1, pp. 31–42, 2003, papers Resulting from the Fifth Workshop on Biosensors and Biological Techniques in Environmental Analysis. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0003267003003027>
- [253] J. Wang, M. M. Sanchez, Y. Yin, R. Herzer, L. Ma, and O. G. Schmidt, “Silicon-based integrated label-free optofluidic biosensors: Latest advances and roadmap,” *Advanced Materials Technologies*, vol. 5, no. 6, p. 1901138, 2020. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/admt.201901138>
- [254] J. Homola, “Present and future of surface plasmon resonance biosensors,” *Analytical and Bioanalytical Chemistry*, vol. 377, no. 3, pp. 528–539, Oct 2003. [Online]. Available: <https://doi.org/10.1007/s00216-003-2101-0>
- [255] G. Quaranta, G. Basset, O. J. F. Martin, and B. Gallinet, “Recent advances in resonant waveguide gratings,” *Laser & Photonics Reviews*, vol. 12, no. 9, p. 1800017, 2018.

- [256] J. Vörös, J. Ramsden, G. Csúcs, I. Szendrő, S. De Paul, M. Textor, and N. Spencer, "Optical grating coupler biosensors," *Biomaterials*, vol. 23, no. 17, pp. 3699–3710, 2002. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0142961202001035>
- [257] B. Sepúlveda, J. S. del Río, M. Moreno, F. J. Blanco, K. Mayora, C. Domínguez, and L. M. Lechuga, "Optical biosensor microsystems based on the integration of highly sensitive mach–zehnder interferometer devices," *Journal of Optics A: Pure and Applied Optics*, vol. 8, no. 7, p. S561, jun 2006. [Online]. Available: <https://dx.doi.org/10.1088/1464-4258/8/7/S41>
- [258] M. Iqbal, M. A. Gleeson, B. Spaugh, F. Tybor, W. G. Gunn, M. Hochberg, T. Baehr-Jones, R. C. Bailey, and L. C. Gunn, "Label-free biosensor arrays based on silicon ring resonators and high-speed optical scanning instrumentation," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 3, pp. 654–661, 2010.
- [259] K. E. Zinoviev, A. B. González-Guerrero, C. Domínguez, and L. M. Lechuga, "Integrated bimodal waveguide interferometric biosensor for label-free analysis," *J. Lightwave Technol.*, vol. 29, no. 13, pp. 1926–1930, Jul 2011. [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-29-13-1926>
- [260] W. Bogaerts, P. De Heyn, T. Van Vaerenbergh, K. De Vos, S. Kumar Selvaraja, T. Claes, P. Dumon, P. Bienstman, D. Van Thourhout, and R. Baets, "Silicon microring resonators," *Laser & Photonics Reviews*, vol. 6, no. 1, pp. 47–73, 2012. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.201100017>
- [261] R. Malik, V. K. Tomer, Y. K. Mishra, and L. Lin, "Functional gas sensing nanomaterials: A panoramic view," *Applied Physics Reviews*, vol. 7, no. 2, p. 021301, 2020. [Online]. Available: <https://doi.org/10.1063/1.5123479>
- [262] M. Oliverio, S. Perotto, G. C. Messina, L. Lovato, and F. De Angelis, "Chemical functionalization of plasmonic surface biosensors: A tutorial review on issues, strategies, and costs," *ACS Applied Materials & Interfaces*, vol. 9, no. 35, pp. 29 394–29 411, 2017, pMID: 28796479. [Online]. Available: <https://doi.org/10.1021/acsami.7b01583>
- [263] J. T. Kirk, G. E. Fridley, J. W. Chamberlain, E. D. Christensen, M. Hochberg, and D. M. Ratner, "Multiplexed inkjet functionalization of silicon photonic

- biosensors,” *Lab Chip*, vol. 11, pp. 1372–1377, 2011. [Online]. Available: <http://dx.doi.org/10.1039/C0LC00313A>
- [264] L. Laplatine, E. Luan, K. Cheung, D. M. Ratner, Y. Dattner, and L. Chrostowski, “System-level integration of active silicon photonic biosensors using fan-out wafer-level-packaging for low cost and multiplexed point-of-care diagnostic testing,” *Sensors and Actuators B: Chemical*, vol. 273, pp. 1610–1617, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0925400518312620>
- [265] K. Misiakos, E. Makarona, M. Hoekman, R. Fyrogenis, K. Tukkiniemi, G. Jobst, P. S. Petrou, S. E. Kakabakos, A. Salapatas, D. Goustouridis, M. Harjanne, P. Heimala, and I. Raptis, “All-silicon spectrally resolved interferometric circuit for multiplexed diagnostics: A monolithic lab-on-a-chip integrating all active and passive components,” *ACS Photonics*, vol. 6, no. 7, pp. 1694–1705, 2019. [Online]. Available: <https://doi.org/10.1021/acsphotonics.9b00235>
- [266] T. Claes, W. Bogaerts, and P. Bienstman, “Experimental characterization of a silicon photonic biosensor consisting of two cascaded ring resonators based on the vernier-effect and introduction of a curve fitting method for an improved detection limit,” *Opt. Express*, vol. 18, no. 22, pp. 22747–22761, Oct 2010. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-18-22-22747>
- [267] L. Dias, E. Luan, H. Shoman, H. Jayatilleka, S. Shekhar, L. Chrostowski, and N. A. Jaeger, “Cost-effective, cmos-compatible, label-free biosensors using doped silicon detectors and a broadband source,” in *2019 Conference on Lasers and Electro-Optics (CLEO)*, 2019, pp. 1–2.
- [268] J. F. Song, X. S. Luo, J. S. Kee, Q. Liu, K. W. Kim, Y. Shin, M. K. Park, K. W. Ang, and G. Q. Lo, “A novel optical multiplexed, label-free bio-photonic-sensor realized on cmos-compatible optoelectronic integrated circuit (oeic) platform,” in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 14.4.1–14.4.4.
- [269] L. Dias, H. Shoman, E. Luan, H. Jayatilleka, S. Shekhar, L. Chrostowski, and N. A. F. Jaeger, “Cost-effective silicon-photonic biosensors using doped silicon detectors and a broadband source,” *Opt. Express*, vol. 31, no. 6, pp. 9135–9145, Mar 2023. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-31-6-9135>
- [270] “Aim photonics apsunny pdk component library,” <https://www.aimphotonics.com/apsunny-component-library>.

- [271] "Axon axopatch 200b," <https://www.moleculardevices.com/products/axon-patch-clamp-system/amplifiers/axon-instruments-patch-clamp-amplifiers#specs>.
- [272] M. Crescentini, M. Bennati, M. Carminati, and M. Tartagni, "Noise limits of cmos current interfaces for biosensors: A review," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 8, no. 2, pp. 278–292, 2014.
- [273] Y. Xing, "Behavioural Models, Parameter Extraction and Yield Prediction for Silicon Photonic Circuits," Master's thesis, Ghent University, 2019. [Online]. Available: <http://photonics.intec.ugent.be/publications/PhD.asp?ID=250>
- [274] J. Kim, "Monolithic Integration of Fluidics, Electronics, and Photonics using CMOS Foundry Processes," Master's thesis, Massachusetts Institute of Technology, 2023. [Online]. Available: <https://dspace.mit.edu/handle/1721.1/150038>
- [275] J. Ohodnicki, P. R. and C. Wang, "Optical waveguide modeling of refractive index mediated pH responses in silica nanocomposite thin film based fiber optic sensors," *Journal of Applied Physics*, vol. 119, no. 6, p. 064502, 02 2016. [Online]. Available: <https://doi.org/10.1063/1.4941103>
- [276] J. Askim and K. Suslick, "8.04 - colorimetric and fluorometric sensor arrays for molecular recognition," in *Comprehensive Supramolecular Chemistry II*, J. L. Atwood, Ed. Oxford: Elsevier, 2017, pp. 37–88. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780124095472126162>
- [277] V. Lucarini, J. J. Saarinen, K.-E. Peiponen, and E. M. Vartiainen, *Kramers-Kronig relations in optical materials research*. Springer Science & Business Media, 2005, vol. 110.
- [278] N. Hamedazimi, Z. Qazi, H. Gupta, V. Sekar, S. R. Das, J. P. Longtin, H. Shah, and A. Tanwer, "Firefly: A reconfigurable wireless data center fabric using free-space optics," *SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 4, p. 319–330, aug 2014. [Online]. Available: <https://doi.org/10.1145/2740070.2626328>
- [279] Z. Gu, J. Zhang, Y. Ji, L. Bai, and X. Sun, "Network topology reconfiguration for fso-based fronthaul/backhaul in 5g+ wireless networks," *IEEE Access*, vol. 6, pp. 69 426–69 437, 2018.
- [280] "Thorlabs plano-convex lenses," https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=3281.

- [281] M. Curran, M. S. Rahman, H. Gupta, K. Zheng, J. Longtin, S. R. Das, and T. Mohamed, "Fsonet: A wireless backhaul for multi-gigabit picocells using steerable free space optics," in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 154–166. [Online]. Available: <https://doi.org/10.1145/3117811.3129239>
- [282] B. Lu, Y.-C. Lu, J. Cheng, R. Schneider, J. Zolper, and G. Goncher, "Gigabit-per-second cryogenic optical link using optimized low-temperature algaas-gaas vertical-cavity surface-emitting lasers," *IEEE Journal of Quantum Electronics*, vol. 32, no. 8, pp. 1347–1359, 1996.
- [283] Q. Zhu, R. Wang, Q. Chen, Y. Liu, and W. Qin, "Iot gateway: Bridging wireless sensor networks into internet of things," in *2010 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing*, 2010, pp. 347–352.
- [284] "Thorlabs fast steering mirrors," https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=15347.
- [285] "Newport fast steering mirror system," <https://www.newport.com/p/FSM-300-01>.
- [286] "Optotune beam steering systems," <https://pacer-usa.com/components/optics-and-lenses/beam-steering/>.
- [287] Y. Li, B. Chen, Q. Na, Q. Xie, M. Tao, L. Zhang, Z. Zhi, Y. Li, X. Liu, X. Luo, G. Lo, F. Gao, X. Li, and J. Song, "Wide-steering-angle high-resolution optical phased array," *Photon. Res.*, vol. 9, no. 12, pp. 2511–2518, Dec 2021. [Online]. Available: <https://opg.optica.org/prj/abstract.cfm?URI=prj-9-12-2511>
- [288] Q. Huang, H. Yu, Z. Zhang, J. Zhao, Z. Zhou, N. Ning, B. Lv, K. Yin, Y. Wang, and J. Yang, "Sparse 2-d optical phased array with large grating-lobe-free steering range based on an aperiodic grid," *Opt. Lett.*, vol. 48, no. 11, pp. 2849–2852, Jun 2023. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-48-11-2849>
- [289] T. Dong, J. He, X. He, Y. Xu, and J. Zhou, "Hybrid design approach of optical phased array with wide beam steering range and low side-lobe level," *Opt. Lett.*, vol. 47, no. 4, pp. 806–809, Feb 2022. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-47-4-806>

- [290] Y. Lei, L. Zhang, Y. Xue, Y. Ren, Q. Zhang, W. Zhang, and X. Sun, "Suppressing grating lobes of large-aperture optical phased array with circular array design," *Appl. Opt.*, vol. 62, no. 15, pp. 4110–4114, May 2023. [Online]. Available: <https://opg.optica.org/ao/abstract.cfm?URI=ao-62-15-4110>
- [291] Y. Xue, Q. Zhang, Y. Ren, Y. Lei, X. Sun, and L. Zhang, "Two-dimensional single-lobe si photonic optical phased array with minimal antennas using a non-uniform large spacing array design," *Appl. Opt.*, vol. 61, no. 24, pp. 7158–7162, Aug 2022. [Online]. Available: <https://opg.optica.org/ao/abstract.cfm?URI=ao-61-24-7158>
- [292] Z. Wang, P. Ma, Y. Yang, L. Cui, L. Yu, G. Luo, P. Wang, J. Pan, Y. Su, and Y. Zhang, "Non-uniform optical phased array with a large steering angle and high side mode suppression ratio," *Appl. Opt.*, vol. 61, no. 36, pp. 10788–10793, Dec 2022. [Online]. Available: <https://opg.optica.org/ao/abstract.cfm?URI=ao-61-36-10788>
- [293] T. Komljenovic, R. Helkey, L. Coldren, and J. E. Bowers, "Sparse aperiodic arrays for optical beam forming and lidar," *Opt. Express*, vol. 25, no. 3, pp. 2511–2528, Feb 2017. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-25-3-2511>
- [294] A. Khachaturian, R. Fatemi, and A. Hajimiri, "Achieving full grating-lobe-free field-of-view with low-complexity co-prime photonic beamforming transceivers," *Photonics Research*, vol. 10, 03 2022.
- [295] K. Sayyah, O. Efimov, P. R. Patterson, J. H. Schaffner, C. White, J.-F. P. Seurin, G. Xu, and A. Miglo, "Two-dimensional pseudo-random optical phased array based on tandem optical injection locking of vertical cavity surface emitting lasers." *Optics express*, vol. 23 15, pp. 19405–16, 2015.
- [296] H. Unz, "Linear arrays with arbitrarily distributed elements," *IRE Transactions on Antennas and Propagation*, vol. 8, no. 2, pp. 222–223, 1960.
- [297] Y. Lo, "A mathematical theory of antenna arrays with randomly spaced elements," *IEEE Transactions on Antennas and Propagation*, vol. 12, no. 3, pp. 257–268, 1964.
- [298] J. Robinson, "Optimum golomb rulers," *IEEE Transactions on Computers*, vol. 28, no. 12, pp. 943–944, dec 1979.
- [299] S. Golomb and H. Taylor, "Constructions and properties of costas arrays," *Proceedings of the IEEE*, vol. 72, no. 9, pp. 1143–1163, 1984.

- [300] R. L. Forward, "Roundtrip interstellar travel using laser-pushed lightsails," *Journal of Spacecraft and Rockets*, vol. 21, no. 2, pp. 187–195, 1984. [Online]. Available: <https://doi.org/10.2514/3.8632>
- [301] D. Tonvall, "Study of reflective and polarization properties of objects found in automotive LiDAR applications," Master's thesis, Umea University, 2020. [Online]. Available: <http://www.diva-portal.org/smash/get/diva2:1453642/FULLTEXT01.pdf>
- [302] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebert, "Applications of second-order cone programming," *Linear Algebra and its Applications*, vol. 284, no. 1, pp. 193–228, 1998, international Linear Algebra Society (ILAS) Symposium on Fast Algorithms for Control, Signals and Image Processing. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0024379598100320>
- [303] L. S. Lasdon, J. Plummer, B. Buehler, and A. D. Waren, "Optimal design of efficient acoustic antenna arrays," *Mathematical Programming*, vol. 39, no. 2, pp. 131–155, Jun 1987. [Online]. Available: <https://doi.org/10.1007/BF02592949>
- [304] "Thorlabs 1555 nm single-frequency dfb lasers," https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=15266.
- [305] "Thorlabs 1550 nm single-frequency ecl lasers," https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=4934.
- [306] S. Sohel Murshed and C. Nieto de Castro, "A critical review of traditional and emerging techniques and fluids for electronics cooling," *Renewable and Sustainable Energy Reviews*, vol. 78, pp. 821–833, 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1364032117305944>
- [307] *Single Chip and Multi Chip Integration*. IEEE Electronics Packaging Society, 2021, ch. 8. [Online]. Available: <https://eps.ieee.org/technology/heterogeneous-integration-roadmap/2021-edition.html>
- [308] "Challenges grow for creating smaller bumps for flip chips," <https://semiengineering.com/challenges-grow-for-creating-smaller-bumps-for-flip-chips/>.
- [309] T. Komljenovic and P. Pintus, "On-chip calibration and control of optical phased arrays," *Opt. Express*, vol. 26, no. 3, pp. 3199–3210, Feb 2018. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-3-3199>

- [310] O. Cierny, B. Moision, and K. Cahoy, "Evaluating the performance of a sensorless wavefront correction algorithm for turbulent horizontal point-to-point links," in *Free-Space Laser Communications XXXV*, H. Hemmati and B. S. Robinson, Eds., vol. 12413, International Society for Optics and Photonics. SPIE, 2023, p. 1241312. [Online]. Available: <https://doi.org/10.1117/12.2651076>
- [311] "Trieve raven sensor," <https://trieve.tech/raven/>.
- [312] K. G. Svantesson and N. G. Nilsson, "The temperature dependence of the auger recombination coefficient of undoped silicon," *Journal of Physics C: Solid State Physics*, vol. 12, no. 23, p. 5111, dec 1979. [Online]. Available: <https://dx.doi.org/10.1088/0022-3719/12/23/019>
- [313] M. L. Laboratory, "Forecasting superconductive electronics technology," *The Next Wave*, vol. 20, no. 3, 2014.
- [314] D. S. Holmes, A. L. Ripple, and M. A. Manheimer, "Energy-efficient superconducting computing—power budgets and requirements," *IEEE Transactions on Applied Superconductivity*, vol. 23, no. 3, Jun 2013.
- [315] J. Clarke and F. K. Wilhelm, "Superconducting quantum bits," *Nature*, vol. 453, p. 1031, Jun 2008.
- [316] J. M. Gambetta, J. Chow, and M. Steffen, "Building logical qubits in a superconducting quantum computing system," *npj Quantum Information*, vol. 3, no. 2, 2017.
- [317] M. de Cea, "Millivolt silicon photonic modulators for cryogenic applications," Master's thesis, Massachusetts Institute of Technology, 2020.
- [318] J. K. Doylend, P. E. Jessop, and A. P. Knights, "Silicon photonic resonator-enhanced defect-mediated photodiode for sub-bandgap detection," *Opt. Express*, vol. 18, no. 14, pp. 14 671–14 678, Jul 2010. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-18-14-14671>
- [319] M. R. Watts, "Adiabatic microring resonators," *Opt. Lett.*, vol. 35, no. 19, pp. 3231–3233, Oct 2010. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-35-19-3231>
- [320] Z. Li and R. J. Ram, "Nanosecond pulsed cmos led for all-silicon time-of-flight ranging," *Opt. Express*, vol. 31, no. 15, pp. 24 307–24 319, Jul 2023. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-31-15-24307>

- [321] "Nanosoi silicon photonics protoyping process," <https://www.appliednt.com/nanosoi-fabrication-service/>.
- [322] C. Vázquez, A. Tapetado, J. Orcutt, H. C. Meng, and R. Ram, "Tolerance analysis for efficient MMI devices in silicon photonics," in *Silicon Photonics IX*, J. Kubby and G. T. Reed, Eds., vol. 8990, International Society for Optics and Photonics. SPIE, 2014, p. 89900A. [Online]. Available: <https://doi.org/10.1117/12.2039730>
- [323] R. K. Gupta, S. Chandran, and B. K. Das, "Wavelength-independent directional couplers for integrated silicon photonics," *J. Lightwave Technol.*, vol. 35, no. 22, pp. 4916–4923, Nov 2017. [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-35-22-4916>
- [324] B. Little and T. Murphy, "Design rules for maximally flat wavelength-insensitive optical power dividers using mach-zehnder structures," *IEEE Photonics Technology Letters*, vol. 9, no. 12, pp. 1607–1609, 1997.
- [325] M. Katzman, Y. Piasetzky, E. Rubin, B. Barenboim, M. Priel, M. Erew, A. Zadok, and H. Suchowski, "Robust directional couplers for state manipulation in silicon photonic-integrated circuits," *J. Lightwave Technol.*, vol. 40, no. 23, pp. 7634–7639, Dec 2022. [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-40-23-7634>
- [326] "Project taara," <https://x.company/projects/taara/>.
- [327] Y. Wu, H. Shang, X. Zheng, and T. Chu, "Post-processing trimming of silicon photonic devices using femtosecond laser," *Nanomaterials*, vol. 13, no. 6, 2023. [Online]. Available: <https://www.mdpi.com/2079-4991/13/6/1031>
- [328] M. M. Milosevic, X. Chen, W. Cao, A. F. J. Runge, Y. Franz, C. G. Littlejohns, S. Mailis, A. C. Peacock, D. J. Thomson, and G. T. Reed, "Ion implantation in silicon for trimming the operating wavelength of ring resonators," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, no. 4, pp. 1–7, 2018.
- [329] B. Chen, X. Yu, X. Chen, M. M. Milosevic, D. J. Thomson, A. Z. Khokhar, S. Saito, O. L. Muskens, and G. T. Reed, "Real-time monitoring and gradient feedback enable accurate trimming of ion-implanted silicon photonic devices," *Opt. Express*, vol. 26, no. 19, pp. 24953–24963, Sep 2018. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-19-24953>

- [330] H. Jayatilleka, H. Frish, R. Kumar, J. Heck, C. Ma, M. N. Sakib, D. Huang, and H. Rong, "Post-fabrication trimming of silicon photonic ring resonators at wafer-scale," *J. Lightwave Technol.*, vol. 39, no. 15, pp. 5083–5088, Aug 2021. [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-39-15-5083>
- [331] J. Komma, C. Schwarz, G. Hofmann, D. Heinert, and R. Nawrodt, "Thermo-optic coefficient of silicon at 1550 nm and cryogenic temperatures," *Applied Physics Letters*, vol. 101, no. 4, p. 041905, 2012. [Online]. Available: <https://doi.org/10.1063/1.4738989>
- [332] S. Liu, J. Feng, Y. Tian, H. Zhao, L. Jin, B. Ouyang, J. Zhu, and J. Guo, "Thermo-optic phase shifters based on silicon-on-insulator platform: state-of-the-art and a review," *Frontiers of Optoelectronics*, vol. 15, no. 1, p. 9, Apr 2022. [Online]. Available: <https://doi.org/10.1007/s12200-022-00012-9>
- [333] A. Arbabi and L. L. Goddard, "Measurements of the refractive indices and thermo-optic coefficients of Si_3N_4 and SiO_2 using microring resonances," *Opt. Lett.*, vol. 38, no. 19, pp. 3878–3881, Oct 2013. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-38-19-3878>
- [334] Z. Lu, K. Murray, H. Jayatilleka, and L. Chrostowski, "Michelson interferometer thermo-optic switch on soi with a 50- μW power consumption," *IEEE Photonics Technology Letters*, vol. 27, no. 22, pp. 2319–2322, 2015.
- [335] R. Soref and B. Bennett, "Electrooptical effects in silicon," *IEEE Journal of Quantum Electronics*, vol. 23, no. 1, pp. 123–129, January 1987.
- [336] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, "Silicon optical modulators," *Nature Photonics*, vol. 4, pp. 518–526, July 2010.
- [337] M. Burla, C. Hoessbacher, W. Heni, C. Haffner, Y. Fedoryshyn, D. Werner, T. Watanabe, H. Massler, D. L. Elder, L. R. Dalton, and J. Leuthold, "500 ghz plasmonic mach-zehnder modulator enabling sub-thz microwave photonics," *APL Photonics*, vol. 4, no. 5, p. 056106, 2019. [Online]. Available: <https://doi.org/10.1063/1.5086868>
- [338] M. Zhang, C. Wang, P. Kharel, D. Zhu, and M. Lončar, "Integrated lithium niobate electro-optic modulators: when performance meets scalability," *Optica*, vol. 8, no. 5, pp. 652–667, May 2021. [Online]. Available: <https://opg.optica.org/optica/abstract.cfm?URI=optica-8-5-652>

- [339] C. Wang, M. Zhang, X. Chen, M. Bertrand, A. Shams-Ansari, S. Chandrasekhar, P. Winzer, and M. Lončar, "Integrated lithium niobate electro-optic modulators operating at cmos-compatible voltages," *Nature*, vol. 562, no. 7725, pp. 101–104, Oct 2018. [Online]. Available: <https://doi.org/10.1038/s41586-018-0551-y>
- [340] F. Eltes, C. Mai, D. Caimi, M. Kroh, Y. Popoff, G. Winzer, D. Petousi, S. Lischke, J. E. Ortmann, L. Czornomaz, L. Zimmermann, J. Fompeyrine, and S. Abel, "A batio₃-based electro-optic pockels modulator monolithically integrated on an advanced silicon photonics platform," *J. Lightwave Technol.*, vol. 37, no. 5, pp. 1456–1462, Mar 2019. [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-37-5-1456>
- [341] J. P. Salvestrini, L. Guilbert, M. Fontana, M. Abarkan, and S. Gille, "Analysis and control of the dc drift in linbo₃-based mach–zehnder modulators," *J. Lightwave Technol.*, vol. 29, no. 10, pp. 1522–1534, May 2011. [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-29-10-1522>
- [342] C. Errando-Herranz, A. Y. Takabayashi, P. Edinger, H. Sattari, K. B. Gylfason, and N. Quack, "Mems for photonic integrated circuits," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 2, pp. 1–16, 2020.
- [343] P. Edinger, A. Y. Takabayashi, C. Errando-Herranz, U. Khan, H. Sattari, P. Verheyen, W. Bogaerts, N. Quack, and K. B. Gylfason, "Silicon photonic microelectromechanical phase shifters for scalable programmable photonics," *Opt. Lett.*, vol. 46, no. 22, pp. 5671–5674, Nov 2021. [Online]. Available: <https://opg.optica.org/ol/abstract.cfm?URI=ol-46-22-5671>
- [344] Y. Ogiso, J. Ozaki, Y. Ueda, N. Kashio, N. Kikuchi, E. Yamada, H. Tanobe, S. Kanazawa, H. Yamazaki, Y. Ohiso, T. Fujii, and M. Kohtoku, "Over 67 ghz bandwidth and 1.5 v $v\pi$ inp-based optical iq modulator with n-i-p-n heterostructure," *J. Lightwave Technol.*, vol. 35, no. 8, pp. 1450–1455, Apr 2017. [Online]. Available: <https://opg.optica.org/jlt/abstract.cfm?URI=jlt-35-8-1450>
- [345] J.-H. Han, F. Boeuf, J. Fujikata, S. Takahashi, S. Takagi, and M. Takenaka, "Efficient low-loss ingaasp/si hybrid mos optical modulator," *Nature Photonics*, vol. 11, no. 8, pp. 486–490, Aug 2017. [Online]. Available: <https://doi.org/10.1038/nphoton.2017.122>
- [346] Z. He, F. Gou, R. Chen, K. Yin, T. Zhan, and S.-T. Wu, "Liquid crystal beam steering devices: Principles, recent advances, and future

- developments,” *Crystals*, vol. 9, no. 6, 2019. [Online]. Available: <https://www.mdpi.com/2073-4352/9/6/292>
- [347] M. Notaros, T. Dyer, M. Raval, C. Baiocco, J. Notaros, and M. R. Watts, “Integrated visible-light liquid-crystal-based phase modulators,” *Opt. Express*, vol. 30, no. 8, pp. 13790–13801, Apr 2022. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-30-8-13790>
- [348] A. Rahim, A. Hermans, B. Wohlfeil, D. Petousi, B. Kuyken, D. V. Thourhout, and R. G. Baets, “Taking silicon photonics modulators to a higher performance level: state-of-the-art and a review of new technologies,” *Advanced Photonics*, vol. 3, no. 2, p. 024003, 2021. [Online]. Available: <https://doi.org/10.1117/1.AP.3.2.024003>
- [349] Y. Kim, J.-H. Han, D. Ahn, and S. Kim, “Heterogeneously-integrated optical phase shifters for next-generation modulators and switches on a silicon photonics platform: A review,” *Micromachines*, vol. 12, no. 6, 2021. [Online]. Available: <https://www.mdpi.com/2072-666X/12/6/625>
- [350] S. Liu, Y. Tian, Y. Lu, and J. Feng, “Comparison of thermo-optic phase-shifters implemented on CUMEC silicon photonics platform,” in *Seventh Symposium on Novel Photoelectronic Detection Technology and Applications*, J. Su, J. Chu, Q. Yu, and H. Jiang, Eds., vol. 11763, International Society for Optics and Photonics. SPIE, 2021, p. 1176374. [Online]. Available: <https://doi.org/10.1117/12.2587413>
- [351] H. Qiu, Y. Liu, C. Luan, D. Kong, X. Guan, Y. Ding, and H. Hu, “Energy-efficient thermo-optic phase shifter with a small footprint based on a silicon spiral waveguide,” in *Conference on Lasers and Electro-Optics*. Optica Publishing Group, 2020, p. STh4O.2. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=CLEO_SI-2020-STh4O.2
- [352] H. Larocque, L. Ranzani, J. Leatham, J. Tate, A. Niechayev, T. Yengst, T. Komljenovic, C. Fodran, D. Smith, and M. Soltani, “Beam steering with ultracompact and low-power silicon resonator phase shifters,” *Opt. Express*, vol. 27, no. 24, pp. 34639–34654, Nov 2019. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-24-34639>
- [353] Z. Yong, H. Chen, X. Luo, A. Govdeli, H. Chua, S. S. Azadeh, A. Stalmashonak, G.-Q. Lo, J. K. S. Poon, and W. D. Sacher, “Power-efficient silicon nitride thermo-optic phase shifters for visible light,” *Opt.*

- Express*, vol. 30, no. 5, pp. 7225–7237, Feb 2022. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-30-5-7225>
- [354] A. Mohanty, Q. Li, M. A. Tadayon, S. P. Roberts, G. R. Bhatt, E. Shim, X. Ji, J. Cardenas, S. A. Miller, A. Kepecs, and M. Lipson, “Reconfigurable nanophotonic silicon probes for sub-millisecond deep-brain optical stimulation,” *Nature Biomedical Engineering*, vol. 4, no. 2, pp. 223–231, Feb 2020. [Online]. Available: <https://doi.org/10.1038/s41551-020-0516-y>
- [355] D. Perez-Galacho, C. Baudot, T. Hirtzlin, S. Messaoudène, N. Vulliet, P. Crozat, F. Boeuf, L. Vivien, and D. Marris-Morini, “Low voltage 25gbps silicon mach-zehnder modulator in the o-band,” *Opt. Express*, vol. 25, no. 10, pp. 11217–11222, May 2017. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-25-10-11217>
- [356] Z. Yong, W. D. Sacher, Y. Huang, J. C. Mikkelsen, Y. Yang, X. Luo, P. Dumais, D. Goodwill, H. Bahrami, P. G.-Q. Lo, E. Bernier, and J. K. S. Poon, “U-shaped pn junctions for efficient silicon mach-zehnder and microring modulators in the o-band,” *Opt. Express*, vol. 25, no. 7, pp. 8425–8439, Apr 2017. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-25-7-8425>
- [357] H. Yu, M. Pantouvaki, J. V. Campenhout, D. Korn, K. Komorowska, P. Dumon, Y. Li, P. Verheyen, P. Absil, L. Alloatti, D. Hillerkuss, J. Leuthold, R. Baets, and W. Bogaerts, “Performance tradeoff between lateral and interdigitated doping patterns for high speed carrier-depletion based silicon modulators,” *Opt. Express*, vol. 20, no. 12, pp. 12926–12938, Jun 2012. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-20-12-12926>
- [358] M. Webster, C. Appel, P. Gothoskar, S. Sunder, B. Dama, and K. Shastri, “Silicon photonic modulator based on a mos-capacitor and a cmos driver,” in *2014 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*, 2014, pp. 1–4.
- [359] J. Fujikata, S. Takahashi, M. Takahashi, M. Noguchi, T. Nakamura, and Y. Arakawa, “High-performance mos-capacitor-type si optical modulator and surface-illumination-type ge photodetector for optical interconnection,” *Japanese Journal of Applied Physics*, vol. 55, no. 4S, p. 04EC01, mar 2016. [Online]. Available: <https://dx.doi.org/10.7567/JJAP.55.04EC01>
- [360] D. Patel, V. Veerasubramanian, S. Ghosh, A. Samani, Q. Zhong, and D. V. Plant, “High-speed compact silicon photonic michelson interferometric modulator,” *Opt.*

Express, vol. 22, no. 22, pp. 26 788–26 802, Nov 2014. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-22-22-26788>

- [361] S. Dev, K. Singh, R. Hosseini, A. Misra, M. Catuneanu, S. Preußler, T. Schneider, and K. Jamshidi, “Compact and energy-efficient forward-biased pn silicon mach-zehnder modulator,” *IEEE Photonics Journal*, vol. 14, no. 2, pp. 1–7, 2022.
- [362] R.-L. Chao, Z. Ahmad, J. Chen, Y. Lai, and J.-W. Shi, “Bjt-type optical phase shifter with small power consumption and fast response time on a silicon photonics foundry platform,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 2, pp. 1–7, 2020.
- [363] S. Wolf, H. Zwickel, C. Kieninger, M. Lauermann, W. Hartmann, Y. Kutuvantavida, W. Freude, S. Randel, and C. Koos, “Coherent modulation up to 100 gbd 16qam using silicon-organic hybrid (soh) devices,” *Opt. Express*, vol. 26, no. 1, pp. 220–232, Jan 2018. [Online]. Available: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-1-220>
- [364] R. Amin, R. Maiti, C. Carfano, Z. Ma, M. H. Tahersima, Y. Lilach, D. Ratnayake, H. Dalir, and V. J. Sorger, “0.52 v mm ito-based mach-zehnder modulator in silicon photonics,” *APL Photonics*, vol. 3, no. 12, p. 126104, 2018. [Online]. Available: <https://doi.org/10.1063/1.5052635>