Optical Beamforming Techniques for Solid-State Automotive LIDAR

Pavan Bhargava
Vladimir Stojanovic, Ed.

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2023-30
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2023/EECS-2023-30.html

May 1, 2023

Optical Beamforming Techniques for Solid-State Automotive LIDAR

by

Pavan Bhargava

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Vladimir Stojanovic, Chair
Professor Ming Wu
Professor Costas Grigoropoulos

Summer 2021

Optical Beamforming Techniques for Solid-State Automotive LIDAR

Abstract

Optical Beamforming Techniques for Solid-State Automotive LIDAR

by

Pavan Bhargava

Doctor of Philosophy in Electrical Engineering and Computer Science

University of California, Berkeley

Professor Vladimir Stojanovic, Chair

The advent of new products in advanced driver assistance systems, virtual reality headsets, robotics, and biometrics has led to increasing demand for high resolution 3D depth maps. LIDAR (Light Detection and Ranging) has emerged as a strong candidate sensor to supply these applications with depth maps of the environment. In fact, LIDAR is already used as a primary sensor in experimental self-driving vehicles, since it provides high-resolution 3D images that existing RADAR solutions do not. However, the reliability and cost of these automotive LIDAR sensors has precluded their use in high-volume mainstream applications.

This thesis explores and demonstrates beamsteering architectures that significantly reduce the size, weight, power and cost of LIDAR sensors. In particular, we focus on designs which achieve the high-volume, reliability, and resolution that automotive applications demand. We begin with an overview of existing LIDAR sensor architectures, and show that the key to improving cost and scalability is switching from traditional mechanical beam-steering to solid-state beam-steering using silicon photonic platforms. We then model and analyze the performance of different solid-state beam steering choices in the context of automotive applications, and arrive at a design for a thermal optical phased array. This thesis then outlines silicon implementation results from from the first 3D integrated the LIDAR SoC based on the aforementioned thermal optical phased array design. Next, we analyze several key issues that prevent this prototype from scaling to more demanding specifications, and propose a new beamsteering architecture based on focal-plane arrays which circumvents these fundamental issues. Finally, the models and automated electronic-photonic design methodologies developed in this work culminate in the design of a scalable monolithically integrated focal-plane array transmitter in a zero-change 45nm SOI CMOS process.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Background

1.1 LIDAR for Autonomous Systems

Over the last decade, the maturation of research and development in artificial intelligence and machine learning techniques has led to a revolution in autonomous systems. With these new techniques, autonomous systems have grown to be able to process raw video and sensor data in real time and dynamically control complex non-linear systems. These advances have already made significant impacts in fields as diverse as factory automation and consumer robotics [6].

Perhaps the sector most ripe for disruption with new machine learning techniques is the transportation sector. Since the DARPA Grand Challenge in 2004, research and development in autonomous driving technology has been growing substantially [22]. With all of these advances in both algorithms and sensor technology, it is inevitable that fully autonomous driving will become available to the general public. Given that more than 9% of all jobs in the US center are in the transportation industry, this entire sector will be significantly impacted by the advent of automation [64]. The automotive industry has already started taking steps to incorporate these advances through new advanced driver-assistance systems (ADAS). These systems, such as parking assistance, lane departure warnings, adaptive cruise control, and others can make the driving experience less stressful and safer by reducing the impact of human error. As these ADAS systems continue to grow in capability, they will eventually be able to completely take over the task of driving the vehicle and realize the vision of fully autonomous driving.

With the dropping cost of imaging sensors and the increasing demand for ADAS features, the automotive industry is rapidly increasing the number of imaging sensors it installs in its cars. The market for ADAS sensor deployments is expected to increase by more than a factor of 3 by 2030 [8]. These sensors are also critical for enabling the machine learning algorithms themselves. The most popular machine learning algorithms in use today operate by consuming large volumes of data about the environment, and tune themselves to better achieve the objective. For this to work it is important that while the algorithm is being

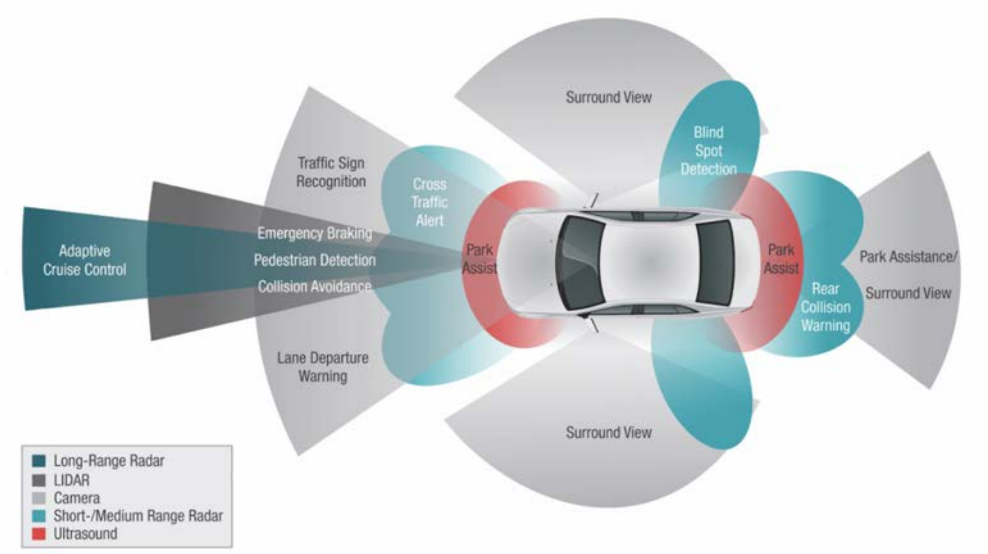


Figure 1.1: Example deployment of sensors in automotive applications to support ADAS and autonomous driving functions

developed that high-quality calibrated data from the environment is available from these sensors [70].

Several different sensors are commonly deployed in cars today to feed ADAS systems. Camera sensors have very high lateral resolution and also provide color information. These sensors are often used to perform object recognition to provide environmental context. For example, cameras enable recognition of signage and traffic lights. One advantage to using cameras is their cost and availability. Camera systems are already regularly deployed in vehicles for other purposes such as theft prevention, and high-resolution cameras are readily available in the market at low cost. In addition, image analysis algorithms have been a primary subject of interest for the machine learning community, enabling many common open-source image recognition tools to be used to bootstrap functionality. On the other hand, cameras are passive sensors that rely on ambient illumination. As a result, recognition of objects from images can be hampered in many corner cases, such as environments with low color contrast or at night when there is no background light.

Active 3D imaging sensors with their own illumination source can compensate for the weaknesses of camera. One example is RADAR, an active imaging sensor that operates in the RF/mm-Wave frequency band. RADAR is capable of much longer range and is not as sensitive to environmental illumination and conditions. However, the lateral resolution of RADAR is much worse than cameras because the carrier wavelength is significantly longer compared to optical bands. One key benefit to RADAR is that it provides velocity information which is often used to distinguish between fast-moving objects on the road and

stationary or oncoming objects. The unit cost of these sensors is also low thanks to highly integrated CMOS implementations [7] now available in the market.

Ultrasound sensors work by sending sound waves into the environment, and measuring the time of flight of the returned signal when it bounces off of a target. These sensors are extremely cheap, and are very robust when detecting distances to short range targets. As a result, they have been the workhorse sensor for most parking proximity applications. Although ultrasound sensors could provide similar 3D imagery as RADAR sensors but with much higher resolution by virtue of the shorter wavelength, ultrasound waves tend to struggle at longer ranges. Their performance at long range is strongly impacted by the characteristics of the air, such as temperature and humidity [16].

While RADAR, camera, and ultrasound sensor data are all important to ADAS systems, they cover orthogonal needs during vehicle operation. Combining these pieces of data together with sensor fusion techniques is often not sufficient to detect and handle long-range small cross-section targets (such as pedestrians or bicyclists) in the environment when traveling at high speed. LIDAR (Light Detection and Ranging) sensors have the potential to fit between the strengths of camera and RADAR sensors and provide long range, high-resolution, and illumination-independent information [70].

LIDAR functions very similarly to RADAR, but uses wavelengths of light in the near-infrared band instead of in the mm-Wave band, leading to significantly improved resolution. Since LIDAR provides direct measurements of distance, but with high resolution, it is very useful when paired with camera data. This pairing allows for each object captured in the camera's field of view to have a measured distance as well. When camera, LIDAR, and RADAR information are combined together, we gain additional velocity information, as well as a robustness to environmental effects that may confound any one sensor. As many self-driving car company's have found, the combination of all of these sensor modalities is critical for high reliability operation [36].

In this thesis, we focus on the development of improved LIDAR sensors. As shown in Figure 1.2, LIDAR is a relatively nascent technology which is expected to grow significantly over the next decade. Currently, LIDAR has not been deployed to vehicles for several reasons: First, LIDAR sensor production has not yet scaled to the high volumes typical of consumer electronics or automotive components. Second, partially due to the low volume and partially due to the complexity of calibration and assembly, LIDAR sensors are extremely expensive. LIDAR sensor suites can sometimes be as expensive as the vehicles they are mounted on. However, with the advent of new silicon photonics platforms, the entirety of a LIDAR system can be integrated onto a single chip, potentially opening up new opportunities for low-cost and high-volume manufacturing. We will explore the feasibility of integrating LIDAR systems onto these new silicon photonics platforms as a part of this thesis.

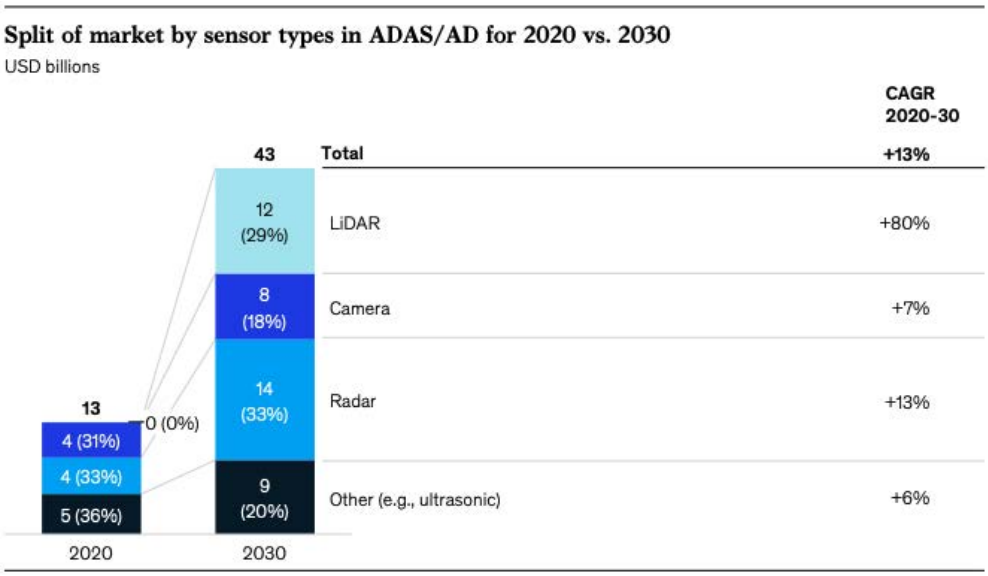


Figure 1.2: Expected growth in automotive sensor market for ADAS from 2020 to 2030, with breakdown of sensor types (McKinsey) [8]

1.2 Anatomy of a LIDAR Sensor

Time-of-flight based LIDAR sensors work by illuminating the environment with a laser light source, and measuring the amount of time it takes for the light to bounce off of targets in the environment and return to the sensor. By knowing the speed of light and the time-of-flight, the distance to the targets in the environment can be measured. An image of the high level operation of a LIDAR sensor is shown in Figure 1.3. First, laser light is generated and modulated with a specific pattern. Then this light is sent to the imaging optics, which illuminates the desired region of the environment with the modulated light. The light reflected off of targets in the environment is collected by a set of imaging optics. Finally, the modulation pattern in the reflected light is measured to extract the time-of-flight. Generally, there are two independent systems, the modulation/detection system, and the imaging optics. While we will briefly cover implementations for modulation and detection in later chapters, the majority of this work will be focused on covering designs for the imaging optics.

Without any imaging optics, entirety of the field of view illuminated by the laser and detected on the receiver would be measured simultaneously. In order to measure the spatial distribution of objects in the field of view (FOV) and form an image, one of two techniques are typically employed. The first possibility, called Flash LIDAR, is where the laser output continues to illuminate the whole FOV, but the receiver is separated into many different pixels, each of which measures a spatially distinct return signal. This is achieved by placing

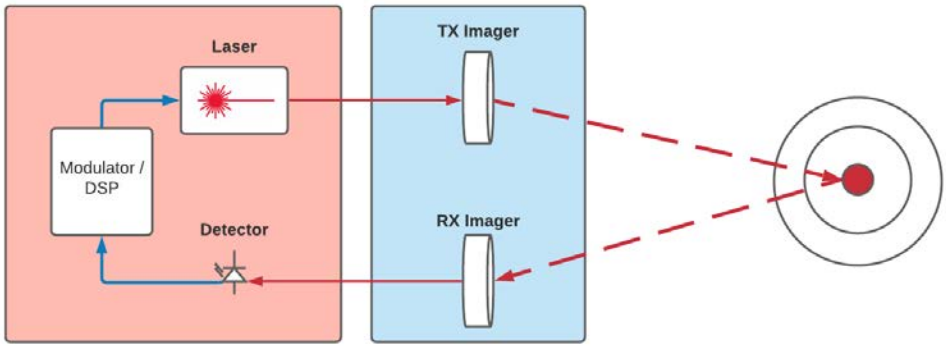


Figure 1.3: Overview of a general LIDAR sensor, with modulation/detection system highlighted in red, and optical beamforming systems highlighted in blue

an array of detectors at the focal plane of a lens system. The lens optics naturally separate each far-field angular component of the return signal to a different detector. In the second technique, beamsteering, a narrow laser beam is formed to concentrate the optical power of the laser to illuminate a single pixel instead of the whole scene. To form an image, the beam is dynamically steered across the FOV. The distance is measured at each beam position, and the image is assembled in software from the series of measurements.

While flash LIDAR implementations are simple, compact, and cheap, they are limited in their ability to deliver optical power to the target. Since there is a fixed upper bound to the amount of power that can be emitted at once from a LIDAR sensor due to eye safety limits, that total power must be divided across the entire FOV. This means that each pixel in the frame receives fewer photons, and as a result has lower signal to noise ratio (SNR). In order to support this lower SNR, current implementations are restricted to the 900nm wavelength range to take advantage of the increased gain from single-photon avalanche detectors (SPADs). While this makes manufacturing in a CMOS-compatible technology easier, the wavelength band is significantly more constrained in output power due to eye-safety. This generally precludes flash LIDAR from supporting the long range typically required for automotive applications. Beamsteering on the other hand focuses the entirety of the laser power into a narrow beam, ensuring that each pixel receives the maximum amount of optical power and SNR [70].

Beamsteering is usually achieved by either rotating the entire optical assembly of lasers, detectors, and optics, or by manipulating a mirror system to redirect a light to and from the fixed laser source and detector. Mechanical beamsteering implementations have been explored extensively in industry. These solutions are often the simplest to implement and offer the fastest time to market. However, scaling these solutions to higher frame-rates, smaller form-factors, and lower cost is very difficult, even at high volume. To reduce the unit cost of the beamsteering system, this thesis will explore various solid-state beamsteering

solutions.

1.3 Thesis Organization

Chapter 2 begins with an overview of the operation principles of different LIDAR beamsteering implementations, and realistic constraints of a system targeted towards autonomous vehicles. We then describe a modeling framework used to evaluate the performance of different beamsteering approaches, and conclude that a thermal OPA based design is a promising first candidate for an integrated LIDAR system. In Chapter 3 we explore trade-offs for different implementations of thermal OPAs, and detail our efforts to demonstrate solid-state LIDAR on a 3D integrated heterogeneous electronic-photonic process. We also examine some shortcomings of this particular design, and of OPAs as a whole. This chapter includes the contents from following publications:

- T. Kim, P. Bhargava, C. V. Poulton, et al., "A Single-Chip Optical Phased Array in a Wafer- Scale Silicon Photonics / CMOS 3D-Integration Platform," IEEE Journal of Solid-State Circuits, accepted for publication.
- T. Kim, P. Bhargava, C. V. Poulton, et al., "A Single-Chip Optical Phased Array in a 3D- Integrated Silicon Photonics/65nm CMOS Technology," in 2019 IEEE International Solid-State Circuits Conference (ISSCC), Feb. 2019, pp. 464–466.
- P.Bhargava, T.Kim, C.V.Poulton, et al., "Fully Integrated Coherent LiDAR in 3D-Integrated Silicon Photonics/65nm CMOS," in 2019 IEEE Symposium on VLSI Circuits, Jun. 2019, pp. 262 – 263.

These works are done in collaboration with Photonic Microsystems Group at MIT and the College of Nanoscale Science and Engineering (CNSE) at SUNY Albany. Christopher V. Poulton at MIT architected, designed and laid out the photonic integrated circuits and Taehwan Kim contributed to the digital control circuit design. Wafer fabrication and 3D integration was done by CNSE. Chapter 4 explores a different idea for solid-state beamsteering, the focal plane array (FPA), and shows how it circumvents some of the fundamental scaling issues with OPAs. We also show a monolithic implementation of the FPA in a zero-change 45nm SOI CMOS process. Finally Chapter 5 describes a scalable and automated design methodology for monotlithically integrated electronic-photonic systems. We present Berkeley Photonics Generator (BPG) as a new open-source Python framework to enable rapid design iteration and verification, and show how it was used to accelerate the design of the FPA described in Chapter 4. The work in this chapter was done together with Sidney Buchbinder, who developed many of the automated DRC cleaning flows, assisted with the design of the software architecture, and enabled integration with external layout verification tools.

Chapter 2

LIDAR Beamsteering Analysis

As illustrated in Chapter 1, the optical beam scanner is an essential part for long-range LIDARs. It is in fact widely believed that successful commercialization of consumer-grade long-range LIDAR largely depends upon whether it is possible to find a low-cost option for reliable optical beam scanner.

In this chapter we introduce the key system metrics by which LIDAR beamsteering systems are measured, and elaborate on how these system metrics impact their performance in an autonomous system. Following this discussion, we derive a set of target specifications that need to be achieved in order to meet the needs of automotive LIDAR applications. Finally, we cover a methodology for analyzing beamforming performance of a LIDAR system, and investigate the effects of several common impairments seen in beam scanner systems. This methodology will be leveraged in upcoming chapters when analyzing specific implementations pursued in this work.

2.1 LIDAR System Metrics

2.1.1 Lateral Resolution

Lateral resolution is the angular size of the smallest feature that can be resolved in the final LIDAR image. For an ideal diffraction-limited beam, there are two main components of the design which contribute to this specification: transmit/receive beam width and the spacing between adjacent beam positions.

In many cases for existing mechanical solutions, the beam can be continuously steered to an arbitrary position in 1 dimension. This means that for this dimension, resolution is completely dominated by the beam width. It is often true that for mechanically steered LIDAR the direction perpendicular to the mechanical rotation is not steered. Instead, multiple lasers emit beams simultaneously in that dimension. As a result, the resolution in this dimension is limited by the number lasers and the angular spacing between the lasers.

The other primary impact to lateral resolution is the width of the beam. As mentioned in

Chapter 1, traditional LIDAR works by making a beam that illuminates the target in only a very narrow angular extent. This narrow beam is then scanned across the field of view, taking measurements at every location to form an image. If multiple objects in the environment are illuminated simultaneously by the beam, and cannot be separated into different pixels on the receiver, they generally cannot be spatially distinguished in the final image. Thus, to improve the resolution, the beam width must be made smaller. The fundamental limit to the width of the beam is dependent on both the wavelength of light used, and the size of the aperture that is used to illuminate or collect the light. For diffraction-limited optics, the final beam width is given by:

$$\theta_{beam} \propto \frac{2\lambda}{\pi w_{aperture}} \tag{2.1}$$

Since the wavelength of light used in LIDAR is on the order of $1\mu m$, a LIDAR scanner can theoretically have 1000x better resolution than a mm-Wave RADAR sensor for the same aperture size.

For many LIDAR sensors, simple measurements of beam-width and space between beam positions is sufficient to quote a resolution specification. It accurately accounts for the total dimensions of the post-processed 3D point cloud to be expected by any following software. However, this resolution metric often fails to appropriately convey aspects of image quality that may degrade the minimum resolvable feature provided by the sensor. In section 2.2, we will explore more holistic metrics which better account for non-idealities commonly found in LIDAR sensors.

2.1.2 Field of View

Field of View (FOV) is defined as the minimum and maximum bounds of the image provided by the sensor. The limits to FOV are strongly dependent on the sensor type. For static lens-based imaging systems, FOV may be limited by the angular extent after which the image quality begins to significantly degrade, or by the physical dimensions of the sensor. In mechanically steered systems, FOV may be limited by range of motion of the mechanical devices. And in solid state sensors, the FOV may be limited by the pitch between antenna, which will be discussed in further detail in Chapter 3.

2.1.3 Frame Rate

Frame rate is the frequency by which every point in the image is updated. For automotive LIDAR applications it is critical that frame rate is high enough that fast-moving objects in the environment can be accurately tracked. Similarly, frame rate sets a lower-bound on latency by which an object can be first detected in the environment. Note that frame rate, resolution, and FOV specifications together place a constraint on how long a beam can spend at a given location. Generally, the SNR of the distance measurement is proportional to both the amount of laser power emitted by the source, and to the amount of time spent

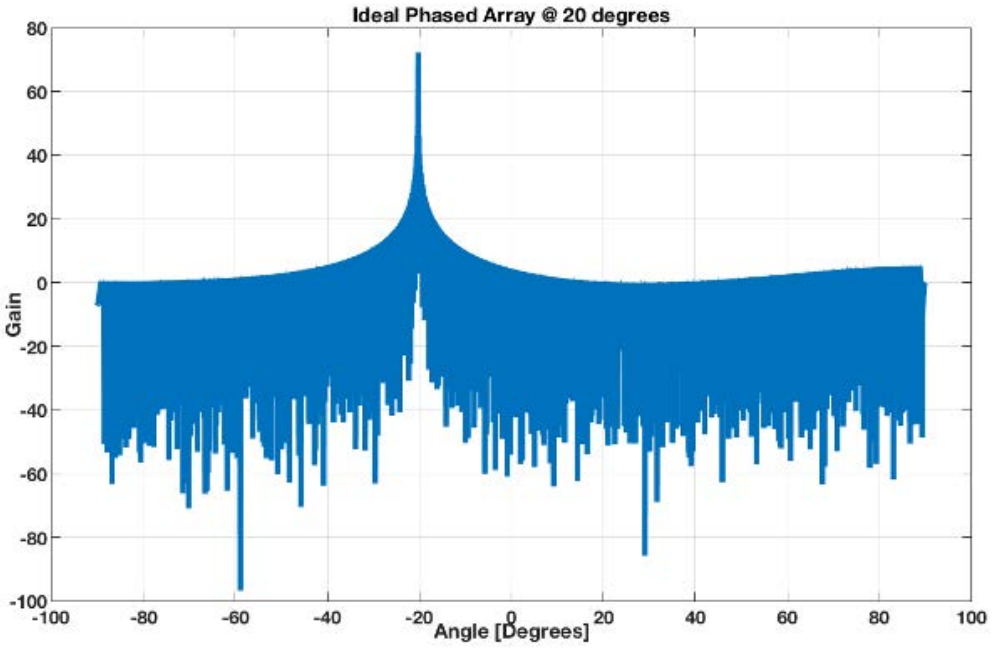


Figure 2.1: Far-field beam pattern of an ideal optical phased array

integrating the reflected light. As a result, choices regarding frame rate, resolution, and FOV of the imaging system strongly impact the performance of the detection system.

2.2 Far-Field Analysis of Beamforming Systems

Beam performance can be qualitatively evaluated by looking at the far-field angular beam pattern. An example of one such beam pattern is shown in Figure 2.1. These plots show the power emitted or received by the beamformer as a function of angle. For high quality narrow beams, the pattern will be tightly distributed around the main lobe and rapidly drop in power as we move away from the center of the beam. The beam width mentioned in section 2.1 can be measured from this plot by extracting the angle at which the power drops by 3dB.

One common impairment present in the beam pattern that is not accounted for in metrics such as beam width or resolution are the presence of side-lobes. Side-lobes are features in the far-field angular beam pattern where a significant amount of power is unintentionally emitted or received outside of the main beam. In Figure 2.1, we show an example beam pattern commonly seen in solid-state LIDAR sensors. The beam pattern follows a $sinc^2$ behavior, leading to periodic peaks and troughs outside of the main beam.

The presence of these side-lobes means that signals from locations outside of the main

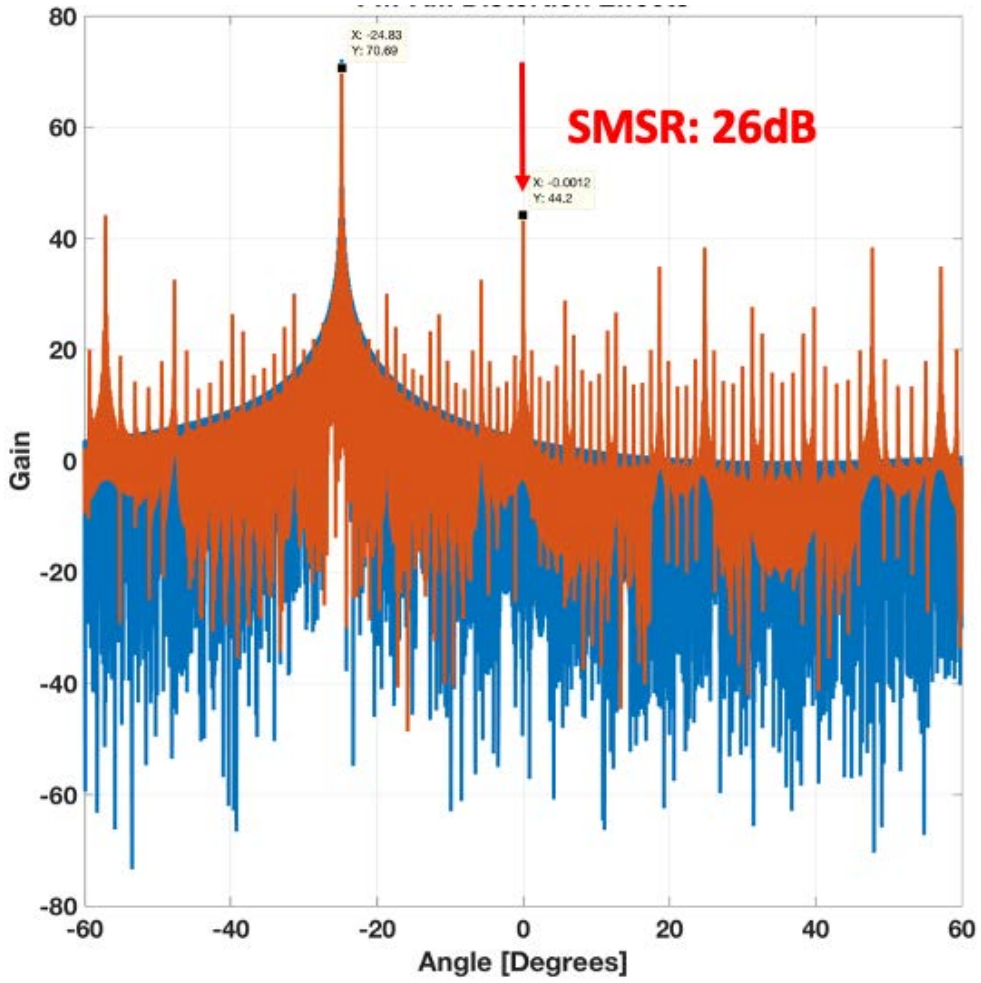


Figure 2.2: Example optical phased array beam pattern with side lobes. SLSR is the ratio between the peak power and the power in any of the side lobes outside of the main beam.

beam appear as if they were in the main beam, effectively aliasing the final image. In automotive applications these side-lobes could result in objects in one lane erroneously appearing as if they were in a different lane, leading to issues with the downstream autonomous driving algorithms. In order to quantify the impact of these side-lobes, it is common to quote side-lobe suppression ratio (SLSR), which is the ratio between the peak power in the main beam and the peak power in any of the side lobes as shown in Figure 2.2.

While this metric is useful for some solid-state beamformers, it is not generally applicable to other solutions which emit Gaussian beams without defined side-lobes. It also does not reflect the position or number of side-lobes. Side-lobes that are adjacent to the main beam simply result in a small degradation in effective resolution. On the other hand, in some beamforming systems side-lobes may appear very far away from the main beam, and have a

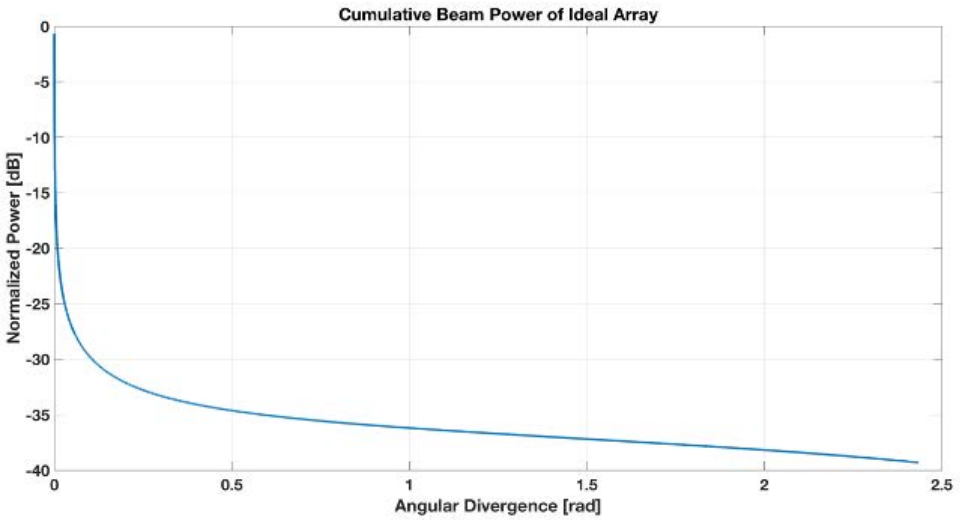


Figure 2.3: Example measured CBP of an ideal optical phased array.

much more significant impact on the final algorithm.

In an effort to more holistically represent the performance of a beamformer, we will focus on the required contrast and the cumulative beam power (CBP). Contrast refers to the measured difference in brightness between two targets that are spatially separated. If the difference in brightness between the two adjacent targets is small (i.e. the contrast is low), then the light from one target has effectively aliased into the measurement for the other target, and the beamformer cannot spatially resolve one target from another. However, if when the targets are spatially farther from each other, one target is very bright and the other is dim enough to be in the noise (i.e. the contrast is high), then it can be said that the beamformer has spatially resolved the two targets. Contrast in this use case is defined as the ratio of power between the two targets, often expressed in dB.

By integrating the power outside of the beam at different angular offsets, we get the CBP. This represents the sum of any power outside of the main beam which may reflect off of interferers and spatially alias in the final image. An example CBP derived from an idealized optical phased array is shown in Figure 2.3. From this plot, and the specification for required contrast, we get a holistic measurement of effective resolution that includes all of the non-idealities present in the beam pattern. Note that the specification for required contrast is strongly dependent on the modulation/detection scheme, and any following algorithms which may be used to improve the image quality. An example case study for calculating and specifying the contrast and the CBP is covered in Section 2.3.

Figure 2.4 shows one example procedure for measuring contrast and CBP in a LIDAR sensor. The sensor is surrounded by a circular reference target which alternates between two different distances. The difference in distances marked as d should be large enough that

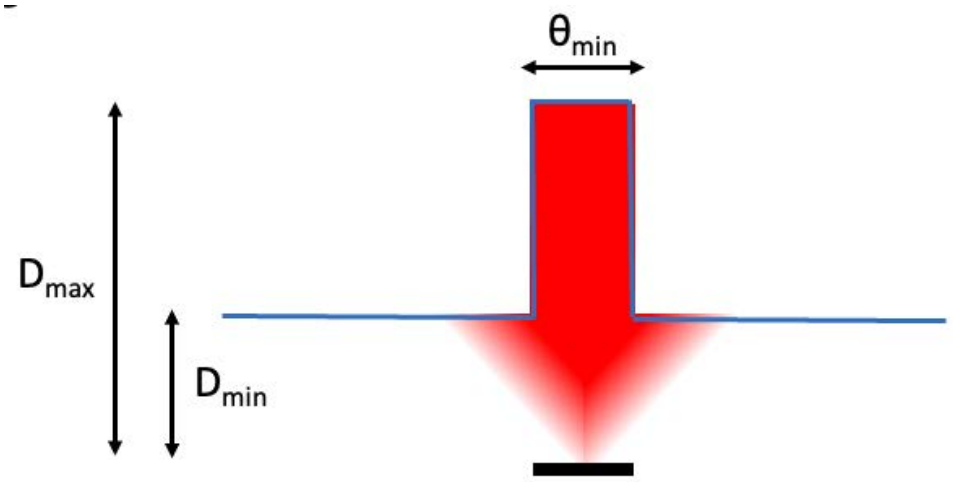


Figure 2.4: Methodology for measuring CBP with a real sensor

the LIDAR sensor can distinguish the two targets purely by the difference in the measured depth, but small enough that the difference in reflected power between the two is negligible. The sensor then performs a measurement when pointing the beam at 0 degrees and records the ratio in measured power between the intended target at distance d_{max} and the interfering target at distance d_{min}. This ratio is the measured contrast at that angle. Note that in order for this measurement technique to work, the LIDAR sensor must be capable of distinguishing between two return signals at the same time. In order to capture the entire CBP, several reference targets with different angular sizes for the target d_{max} should be used. Also note that the performance of most imaging systems degrades when measuring off axis, so repeating the CBP measurement with the beam pointed at the max FOV and at 70% of the FOV can be informative.

2.3 Specifications for Automotive LIDAR Systems

In order to be able to effectively evaluate the relative performance of different beamforming architectures, it is critical to create a set of target specifications which must be met for the desired application. In this section, we will focus on setting a concrete set of specifications for mid-range automotive LIDAR.

2.3.1 Determining Effective Resolution

There are no widely agreed upon requirements for true resolution. It is commonly stated that the spacing between points should be as low as 0.1 degrees, but the quality of the image

is rarely ever mentioned.

It is difficult to set a specification on LIDAR resolution and image quality without holistically considering its impact together with an autonomous driving algorithm. For the purpose of comparing the relative merits between LIDAR sensors, we will propose a simple specification based on a common automotive driving scenario.

Consider a LIDAR sensor deployed on a US highway. We would like to detect the presence of a vehicle up to 100m away. The image quality should be high enough such that the vehicle can be classified and unambiguously located in one lane vs an adjacent lane. Since the average US highway lane width is 3.7m, this means that the resolution must be better than $atan2(\frac{3.7m}{100m}) \approx 40mrad$.

2.3.2 Setting Target Contrast

Now that we have a target resolution, we need to determine the contrast required to achieve the specified resolution. Leveraging the same US highway example, we can assume that aside from the vehicle 100m away, we can imagine that the worst case scenario is that the sensor is surrounded by a spherical reflector at some minimum distance d_{min}. In this scenario, all light not illuminating the target vehicle is reflected back to the sensor, and results in an erroneous interfering signal at short range. Although this scenario is extreme, it serves as a consistent upper bound on image aliasing. A somewhat less extreme version of this commonly occurs during a traffic jam. In these scenarios it is common to have all adjacent lanes filled with vehicles, and to have the vehicle ahead of you in your own lane at a significant distance away.

The exact requirement for target contrast will be strongly dependent on the exact details of the detection hardware, but for the purpose of placing basic bounds on the beamformer performance we will make some simple assumptions. In the example highway scenario, we would like to detect our target that is 100m away while we are surrounded by interferers as close as 1m outside of the main beam. For isotropic scattering, a target at 100m would return 40dB lower power than the interferers that are 1m away. If we assume that the detection system needs better than 10dB of signal to interference ratio to successfully detect the target, this means that the total beam contrast must be 50dB. Assuming that the TX and RX beamformer systems are symmetric, this means that they must each have 25dB of contrast at the target resolution in order to meet this specification.

CBP can be computed by taking the integrating the normalized power within a given angular divergence:

$$CBP(\Delta\phi) = 1 - \frac{1}{P_{total}} \int_{\frac{-\Delta\phi}{2}}^{\frac{\Delta\phi}{2}} I(\theta) d\theta \tag{2.2}$$

This metric essentially measures how much of the total power in the beam is emitted or captured outside of a given angular divergence. By measuring the CBP at the target resolution, we can capture the contrast in the worst case condition, and as a result check that the specification is met. From figure 2.5, we can see that an example idealized optical phased array based beamforming system can meet the criteria if it has > 4096 antenna

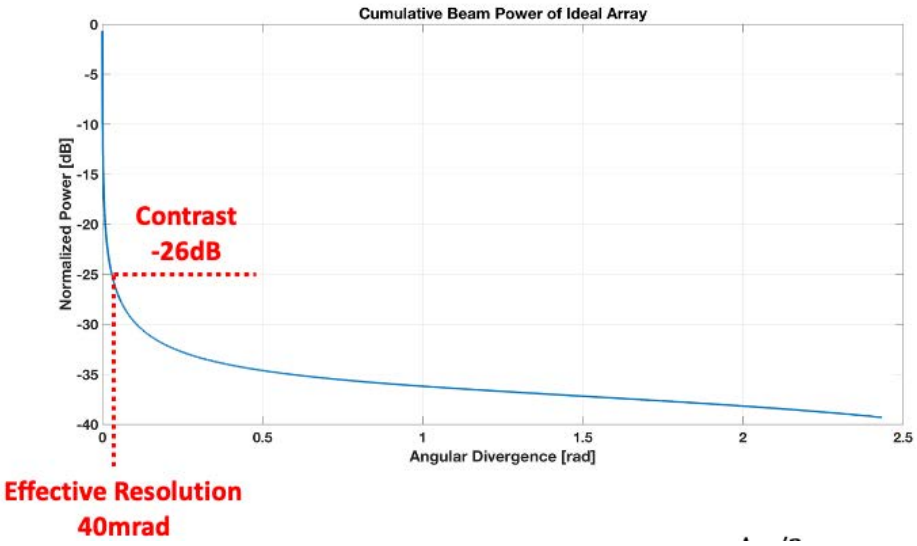


Figure 2.5: Example measured CBP of an ideal optical phased array, with annotated resolution and contrast targets.

2.4 Conclusion

In this chapter we introduced a methodology for evaluating the performance of a beam scanner system. From our analysis, commonly used specifications such as SMSR and 3dB beam-width do not appropriately capture the true impact of beam steering impairments to the final system. We propose using *Cumulative Beam Power* as a metric, which takes into account not only the size of the side lobes but also their location relative to the main beam to give a holistic view of image quality. We performed a case study for automotive applications, and set target specifications based on an assumed overall LIDAR system architecture. While these target specifications may change given the detection methodology and expected use-cases, it sets the stage for holistically analyzing and comparing the relative performance of different beamsteering systems.

Chapter 3

Optical Phased Arrays

Beam scanning modules in existing commercial products are mostly based on mechanical control, such as motor-driven rotating collimation stages/mirrors [48] or galvanometers. In order to make those high-precision mechanical systems work reliably in an unstable environment (e.g. cars), they typically include a number of moving parts and tend to be bulky and slow, which is undesirable as it also leads to large size, increased weight, and high power consumption. More importantly, they require a complex assembly and calibration process, which results in extremely high unit cost. Largely due to this cost issue, LIDAR is currently not considered for use in mass production vehicles. This has motivated active research on optical beam steering techniques that minimize mechanical movements or are completely solid-state during operation. Existing techniques include micro-electro-mechanical systems (MEMS) mirrors [32], lens-assisted emitter arrays [44] [35], liquid crystal waveguides [23], and photonic crystal waveguides combined with diffraction gratings [62].

Optical phased arrays (OPA) are one of the most promising solutions for solid-state beam steering. Radio-frequency/mm-wave phased arrays are already prevalent in RADAR and wireless communications, and it is also possible to realize the same concept in the optical domain. This technique has gained a lot of attention alongside advancements in silicon photonics technology, which enables inexpensive fabrication of a large number of optical components [60]. OPA technology has progressed tremendously within a relatively short period of time, and multiple large-scale implementations with 1000 elements recently reported [51]-[26].

However, it is still challenging to realize a low-cost OPA that can bring the proliferation of chip-scale optical beam scanners to mass markets. For instance, to meet the resolution and field- of-view requirements for automotive LIDARs, the element count should reach 500-1000. The majority of prior OPA demonstrations take a multi-chip approach where photonics and electronics are present on separate chips, and the number and density of I/O connections and electronic circuits clearly exceeds the limits of low-cost packaging options. Several OPA architectures have been proposed to reduce the number of independent electrical signals [20]-[52] by trading off the array control flexibility.

Ultimately, a single-chip solution is desired to completely resolve the I/O and electronics

density problem at a minimum unit cost with guaranteed performance in the presence of process and design-dependent phase uncertainty. 2D monolithic integration of electronics and photonics in an SOI process is one technique of realizing a single-chip OPA [20],[1]. However, physical constraints due to CMOS design rules and limited material/processing steps significantly constrains photonics design and make it hard to meet system requirements. Moreover, thermo-optic phase shifters, often used in such OPAs, consume a significant amount of maximum power and require high voltage swings. Delivering large amounts of power through each of the 1000s of on-chip interconnects tends to cause routing/placement congestion, large die size, and significant circuit power overhead due to limited driver efficiency and voltage droop.

In this chapter, we present our work on the realization of a single-chip OPA on a wafer-scale 3D integration platform which allows for photonics and CMOS electronics to be independently optimized while enabling flexible, dense vertical connections between them [39]. We also introduce key OPA building blocks that leverage the uniqueness of the integration platform, including apodized grating antennas that maximize the array effective aperture and low-voltage L-shaped thermo-optic CMOS-compatible phase shifters connected vertically to pitch-matched pulse density modulated (PDM) switch-mode drivers, completely eliminating the placement/wiring overhead and achieving an area and power-efficient system suited for large-scale OPAs. Experimental validation of our OPA architecture is also presented, including a successful demonstration of two-dimensional wide-range beam steering and with a large array aperture and full array calibration with per-element phase control, which resolves static passive beam pattern distortion.

3.1 Optical Phased Arrays for Robust High-Resolution Beam Scanning

To get a sense of what the performance expectations are for OPAs from a system-level perspective, we can refer to the beam scanner specifications shared across the automotive industry. The main performance metrics and their respective values [67], [66] are as follows.

- Lateral Resolution: 0.1–0.2 degrees (horizontal and vertical)
- FOV: $>$ 90 degrees (horizontal)
- Power Budget: 10–30 W
- System Cost: \$100–\$200

Above all, the scanner must satisfy the resolution requirements so as to recognize small objects (e.g. pedestrians) at a long distance. The relationship between the array parameters and the beam resolution (or full-width at half-maximum (FWHM)) is well known [46],

$$FWHM = cos^{-1}\left(sin\theta_b - \frac{2.78\lambda}{2\pi Nd}\right) - cos^{-1}\left(sin\theta_b + \frac{2.78\lambda}{2\pi Nd}\right) \tag{3.1}$$

where θ_b is the beam angle (0 for upright), d is the antenna spacing, N is the antenna count, and λ is the wavelength, respectively. For a given wavelength, one can note that the lateral resolution is determined solely by the absolute size of the total aperture ($W = Nd$). For example, 0.2 degrees worst case resolution for 90 degrees FOV and $\lambda = 1550$ nm requires W = 0.54 mm.

Meanwhile, the ambiguity-free steering range of a phased array is determined by the antenna spacing.

$$FOV = 2sin^{-1}\left(\frac{\lambda}{2D}\right) \tag{3.2}$$

For example, the spacing should decrease until it meets the range requirement, which results in increased number of antennas. For instance, 90 degrees FOV would require 0.71λ-spacing (d $\approx$ 1.1 μm), which corresponds to 490 antennas for W = 0.54 mm. Alternatively, one can achieve the same FOV with two adjacent OPA channels and relax the spacing to 1.3λ (d $\approx$ 2 μm). The total number of elements is then $2N = 2(W/d) = 540$, slightly higher than the single-channel case. This level of OPA multiplexing would eventually be limited by power and cost constraints. Note that this pitch-FOV trade-off can potentially be relaxed through non-uniform antenna placement [34], [26], which alters the basic relationship of the FOV equation. Nevertheless, to be a compelling alternative to mechanical beam scanners, a clear path to scale the element count to 500–1000 is a must.

3.2 OPA Architectures and Reduced Interface Complexity

In typical OPA implementations, to ensure coherence across the overall array aperture, a single laser source is used and evenly distributed to feed the antenna elements. Given this, the simplest and the most straightforward OPA architecture is the tree architecture (Figure 3.1), where a 1:N optical power splitter distributes the input power into N waveguides each connected to its own phase shifter and an optical antenna. Since the tree architecture mandates N independent phase shifters to address all possible beam directions within the FOV, the I/O and electronics density associated with 500 - 1000 antenna elements at a wavelength-scale pitch has made it challenging to realize a low-cost OPA for practical applications, especially based on a multi-chip integration strategy where photonic and electronic circuits are realized in separate chips and packaged on a cheap substrate.

This has motivated various works to propose alternative OPA architectures that reduce interface and control complexity. One possible option is to change the optical distribution network architecture to a grouped tree structure (Figure 3.1) [20]. In this design, antennas

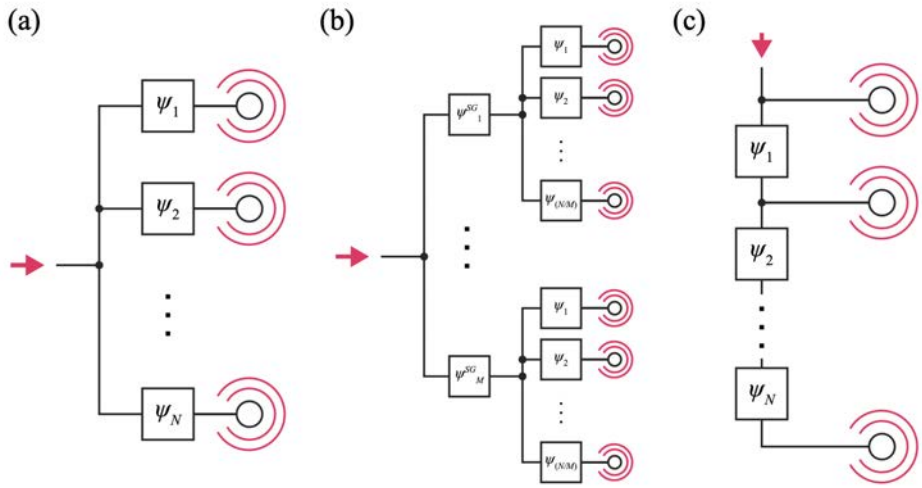


Figure 3.1: OPA distribution network types (a) tree architecture, (b) grouped tree architecture, and (c) cascaded architecture.

and associated phase shifters are divided into M subgroups, and each subgroup has its dedicated phase shifter at its root of the splitter tree to adjust the phase offset between subgroups. For a linear phase ramp, a single set of signals can be shared over the whole array to control the phase shifters within each subgroup. As a result, the overall independent signal count in grouped tree architecture is M + N /M . The level of subgroup hierarchy can also be further increased [20] to improve the granularity of the phase adjustments.

Another solution is the cascaded architecture (Figure 3.1) [52]. In this design, optical power distribution is done through a series of couplers placed along the bus waveguide, and the phase shifters are located between those couplers. Namely, the phase shifter adjusts the relative phase difference, rather than the absolute phase. For a phase ramp required to perform linear scanning, the relative phase between adjacent shifters are always constant across the array. As a result, only one control signal is enough to support all beam positions. Similar to the grouped tree, this architecture can also be segmented (i.e. grouped cascade) and driven by multiple signals to introduce additional flexibility to the phase pattern [52].

Finally, it is worth noting that instead of modifying the optical distribution network, it is also possible to utilize the slow transient response of thermo-optic phase shifters to reduce the interface complexity. In [26], phase shifters are placed in a rectangular array, and the phase shifters within one row are addressed individually through a single row wire in a time-shared fashion using pulse-width modulated (PWM) signal. By synchronizing row PWM signals to the activation signal on the column wire, only 37 wires were used to address 128 phase shifters in [26].

3.3 Process and Design-Dependent Random Phase Fluctuation

One of the well-known issues in silicon photonics that could limit the use of OPAs in mass markets is waveguide coherence [19]. Process variation-induced nanoscale uncertainties in waveguide geometry perturb the effective index and result in random fluctuations of phase [69], disrupting spatial coherence at the array aperture. Still, phase perturbation from process variation is a static error and can be corrected through proper calibration processes. In [20], it was concluded that assuming reported levels of variation for standard strip waveguides fabricated in an SOI process [69], it is possible to maintain decent OPA performance with a grouped tree architecture of subgroup size $N/M = 8$ by adjusting phase offset between subgroups.

However, it must be noted that the waveguide coherence is heavily dependent on the actual fabrication process as well as the geometry of the designed components. This is especially the case in thermal phase shifters where the temperature dependence of the silicon index is utilized to adjust phase. As noted in Section 3.1, the total power budget of the LiDAR system is only on the order of 10s of watts, and it is important to optimize the efficiency of the heater and driver circuits to support large element count. A popular way to enhance the heater phase efficiency is directly embedding the resistive segment of the heater into the waveguide so that the thermal impedance between the heater and the waveguide core is minimized [68]. At the same time, in order to reduce the heater driver circuit complexity, the heater resistance should be decreased to bring the voltage swing down to CMOS-compatible levels, which requires high heater contact density.

As explained later in detail (Section 3.5), we have designed a thermo-optic phase shifter with an L-shaped waveguide where the heater is embedded into a one-sided slab layer to achieve high efficiency and low resistance. However, slab waveguides are known to have much worse waveguide coherence due to additional error sources, including a partial etch step [69]. For example, an analysis based on the coherence of simple strip waveguides can be rather optimistic, especially in OPAs where components are heavily optimized to satisfy the power and area constraints. This eventually motivated us to pursue per-element phase control flexibility, which completely desensitizes OPA performance to the level of device coherence.

3.4 Wafer-Scale 3D Heterogeneous Integration of Silicon Photonics and CMOS Electronics

In [20] and [1], 2D monolithic integration of photonics and electronics on a common SOI CMOS substrate was presented as a low-cost solution to I/O complexity issue, realizing a single-chip OPA and up to 1000 element count [20]. However, 2D monolithic integration has a few critical disadvantages in the context of OPAs. First, photonic device design is

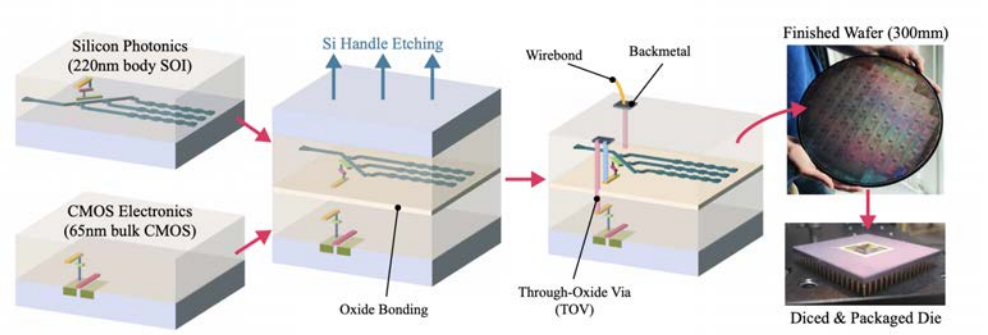


Figure 3.2: Overview of the 3-D heterogeneous integration platform used to construct the single-chip OPA

limited due to CMOS design rules and available materials and processing steps. Moreover, side-by-side placement of photonics and CMOS tends to cause placement/routing congestion in large-scale arrays. Typical thermal phase shifters consume up to 50 - 100 mW of power and require ≈10 V swing, which mandates wiring via thick top-layer metal to minimize the IR loss over the interconnect as well as stacked thick oxide devices, which makes it very challenging to realize compact driver circuits. As a result, OPAs in monolithic platforms often result in a large die footprint for small active array apertures [20], [1].

In this work, we realized single-chip OPAs on a 3D heterogeneous integration platform shown in Figure 3.2, similar to [47]. The integration process starts with two independently optimized 300 mm wafers. Photonic devices are fabricated with 193 nm immersion lithography on a SOI wafer with 220 nm body thickness (typical waveguide loss is 3 dB/cm for the 1500 nm to 1600 nm wavelength range), while electronics are implemented using a standard CMOS technology. A 65 nm low-power bulk CMOS process was used in this work, but in principle, any process can be used as long as it has the matching wafer size to the photonics as well as acceptable transistor and wiring density. The two wafers are then face-to-face oxide-bonded at the wafer scale, and the silicon handle on the photonic wafer is globally etched down to the buried oxide (BOX).

After the wafer bonding and etching, through-oxide vias (TOVs) are formed to establish the electrical connections between CMOS and photonics. TOVs can be densely placed at arbitrary locations with a pitch as small as $7\mu m$ and have extremely low parasitic capacitance (3 fF), which can be treated just like top-level metal vias in the CMOS backend. Lastly, back-metal is placed for pads and TOV-pad connections.

The final result is a single 300 mm wafer, which can be further processed following standard CMOS packaging steps, including dicing and wire-bonding to a ceramic package. Compared to the die-scale vertical integration often used for optical transceivers [10], much higher ($> 7\times$) interconnect density versus copper pillars, flexible via/circuit placement that minimizes wiring/placement overhead, and lower unit cost of wafer-scale 3D integration make

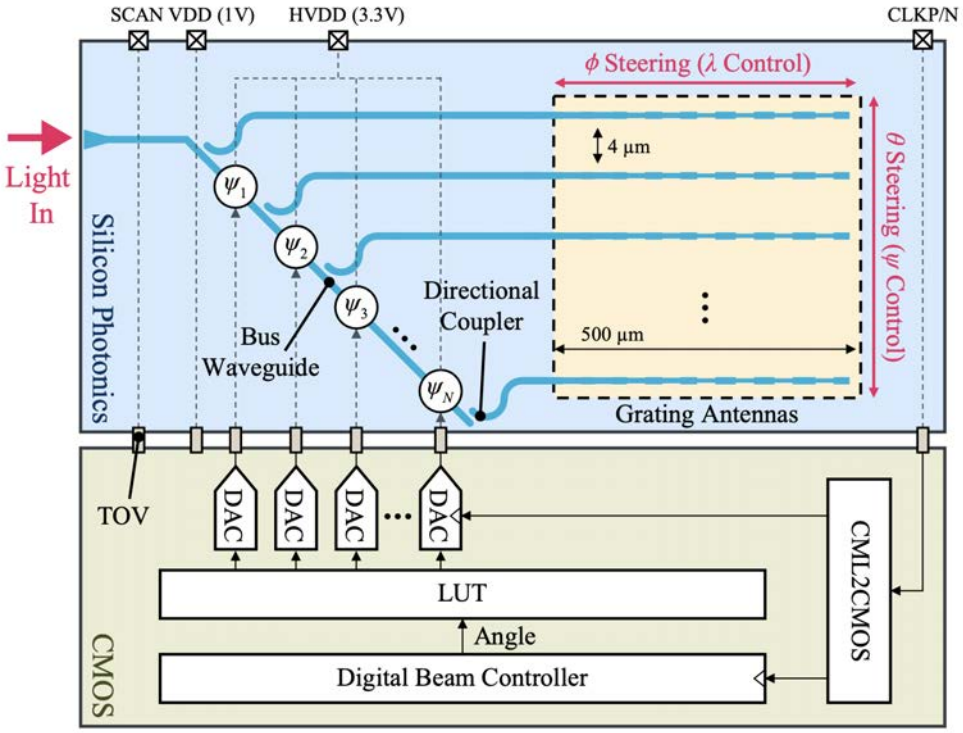


Figure 3.3: Overview of the single-chip OPA architecture

our platform better suited for OPAs.

3.5 Optical Phased Array Implementation

Figure 3.3 shows an overview of the single-chip OPA prototype implemented on our 3D integration platform. The design is based on a cascaded array architecture from Section 3.2. The advantage of the cascaded architecture in our prototype was twofold. First, since it is possible to apply a nominal steering signal through a single wire, it is easier to examine the level of phase uncertainty before packaging using a simple probe. Second, the cascaded architecture is generally less affected by thermal crosstalk [24] since the nominal phase pattern does not have discontinuities coming from limited phase shifter range, unlike the (grouped) tree architecture.

The input laser is coupled into the on-chip bus waveguide from a lensed fiber via an edge coupler. Cascaded directional couplers placed along the bus waveguide distribute the light into each optical antenna element. Thermo-optic phase shifters are embedded into the bus waveguide sections between the couplers to adjust the relative phase offset of adjacent

antenna elements. Ultimately, this enables beam steering along the direction of the array placement (θ in Figure 3.3). The elements are placed at a $4\mu m$ pitch in our prototype, which corresponds to the theoretical steering range ($2\theta_{bmax}$) of 22.3 degrees at $\lambda = 1550$ nm from Equation 3.2. However, the antenna elements used in our work can be placed at a tighter pitch down to the wavelength scale without inter-antenna evanescent coupling [68] (e.g. $1.4\mu m$ pitch was used in Figure 3.5). Each thermal phase shifter is independently driven by a CMOS controller based on a switch-mode digital-to-analog converters (DACs). Such per-element independent control guarantees optimum OPA performance through calibration regardless of design and process-dependent photonic component coherence. One critical source of potential error is variability in the directional coupling coefficients between the bus waveguide and each antenna. Process variation in this coupling parameter can cause non-uniformity in both the phase and amplitude distribution between antenna. Per-element control also mitigates potentially compromised robustness in the cascaded architecture due to the fact that the failure in one phase shifter affects all subsequent antennas: impact of one broken shifter is limited by introducing phase bias to the following shifter. DAC inputs are provided by an on-chip lookup table (LUT) that has the beam position codes after array calibration, enabling rapid steering along arbitrary trajectories.

In addition, our antenna enables the main radiation direction of an individual element to be tunable by the laser wavelength. Namely, the beam steering along the direction orthogonal to the array placement (ϕ in Figure 3.4) is done through wavelength tuning, achieving full two dimensional beam steering. The details of the antenna element, the thermo-optic phase shifter as well as the control circuit design are described in the following subsections.

3.5.1 Apodized Grating Antenna

Figure 3.4 shows our optical antenna concept. The antenna is a long 400-nm-wide waveguide grating formed by fully-etched sidewall perturbations, where the perturbed waveguide section creates a local effective index mismatch. As a result, when the light propagates through this grating, a certain fraction of light is scattered into free space at every perturbed section. The overall length of our grating antenna is $500\mu m$, which enables a large emitting aperture and small divergence angle in ϕ (0.15 degrees). This is unlike previous work where short grating couplers (3.55 μm long) with divergence angles over 10 degrees were used as emitting elements [20].

Note that the light scattering ratio due to the local index contrast is determined by the depth of sidewall etching. To produce a uniform emission profile and maximize the effective aperture of the antenna, the perturbation depth has to increase down the length of the antenna (i.e. the perturbation must be apodized [55]). At the same time, to maintain the same optical path offset between neighboring scattering points, the physical distance between perturbations (δx)should also gradually increase. The 1 nm resolution of the photolithography masks used allows for near infinitesimal changes in perturbation strength (Figure 3.4), again highlighting the benefit of utilizing fully customized photonics.

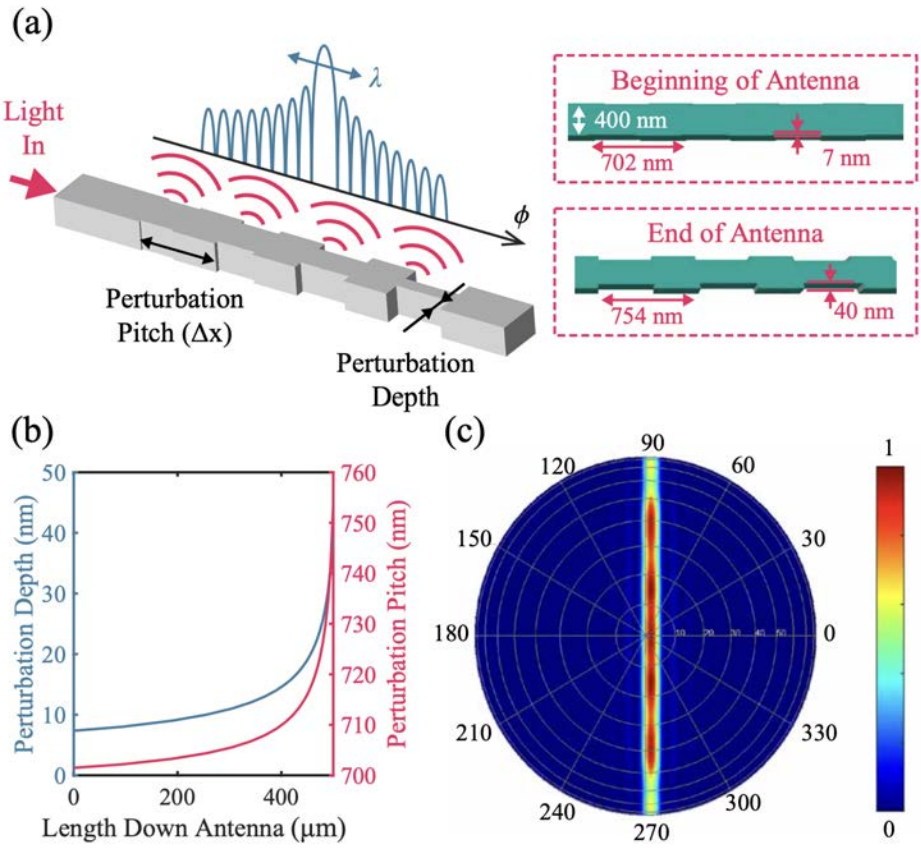


Figure 3.4: (a) Beginning and at the end of the antenna. (b) Perturbation distance and pitch distribution across the antenna element. (c) Antenna emission pattern from an finite-difference time domain (FDTD) simulation.

The reader may notice that a grating antenna composed of repeated scattering points resembles a 1D phased array. In this case, the relative phase difference between emitting elements is set by the optical path between scattering points. The following relationship can be established to determine the position of the main beam in ϕ:

$$\frac{2\pi}{\lambda}\Delta x(sin\phi + n_{eff}) = 2\pi M, M \in Z \tag{3.3}$$

$$\phi_b = sin^{-1}\left(\frac{M\lambda}{\Delta x} - n_{eff}\right) \tag{3.4}$$

where neff is the effective index of the silicon waveguide. We designed the grating to satisfy $n_{eff} = \lambda$, M = 1 at $\lambda = 1550nm$. Since ϕ_b depends always on λ, beam steering along ϕ is possible via input wavelength tuning. The simulated group index at $\lambda = 1550nm$ is 4.43, which results in a steering efficiency of 0.164 degrees per 1 nm wavelength shift.

3.5.2 L-Shaped Thermo-Optic Phase Shifter

We have two key objectives for phase shifter design: high phase efficiency and CMOS-compatible swing. Figure 3.5 shows a perspective view of the layout details around the bus waveguide, the cross-section and top-view of the heater segment, and the layout of directional coupler connected to antenna elements. The bus waveguide is L-shaped, with a partially etched 110-nm-thick slab on one side of a 400-nm-wide 220-nm-thick core (Figure 3.5). A resistive heater is directly embedded into the bus waveguide by N-doping the slab, 350 nm away from the waveguide core (Figure 3.5). This ensures low thermal impedance between the heater and the core, resulting in high thermal efficiency (20 mW/π from COMSOL simulation). At the same time, the waveguide mode and doped region are isolated in the L-shape geometry (Figure 3.5), which ensures minimum optical loss (simulated loss: 0.016 dB). The length of one phase shifter is 32 m and, considering that the thermo-optic coefficient of silicon at 1550 nm is $1.8 \times 10^4 K^1$, 2π shift corresponds to 270 K temperature range.

N-doped slab layer also contains large clearance to form dense contact tethers and lower the resistance. In addition, multiple positive and negative contacts are formed in an interleaved fashion, resulting in an alternating current directions within the heater segment and resistors in parallel (Figure 3.5). This ensures consistent current density (i.e. removal of hot spot and electromigration-limited segments) and low resistance. The resulting resistance of the phase shifter was 270Ω, corresponding to 3.3 V voltage and 12.2 mA current swing.

Between the phase shifters on the bus waveguide, directional couplers are formed by placing strip waveguides on the other side of the bus waveguide core (Figure 3.5). The strip waveguides have a width of 500 nm in order to achieve phase-matching to the L-shaped waveguide and efficient evanescent coupling (simulated loss: 0.015 dB 0.02 dB across 1500 nm 1600 nm). The coupling strength of each directional coupler is gradually increased to ensure uniform distribution of the optical power across the entire aperture (Figure 3.9). After the desired length of evanescent coupling, the strip waveguide is separated away from

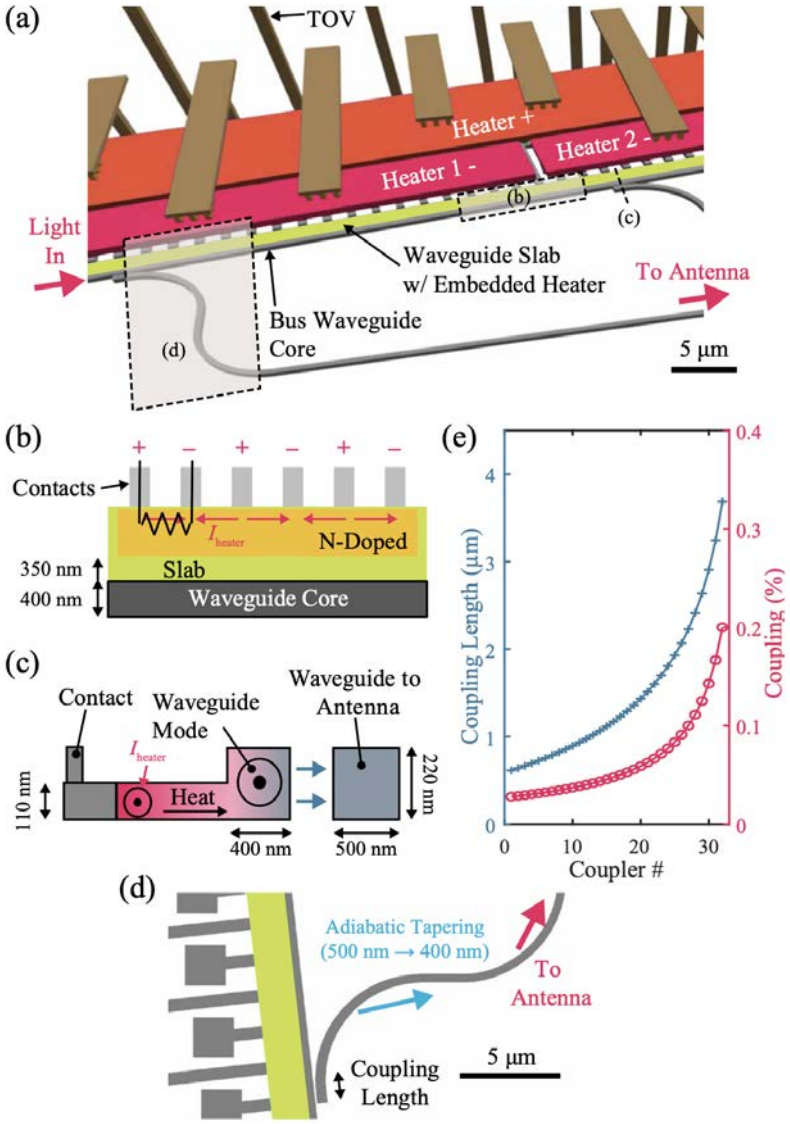


Figure 3.5: (a) Perspective view of the layout details around the bus waveguide section including embedded thermooptic phase shifter and directional coupler. (b) Top view and (c) cross-sectional views of the L-shaped phase shifter. (d) Bus waveguide antenna connection through a directional coupler. (e) Evanescent coupling strength distribution across the bus waveguide for uniform power distribution.

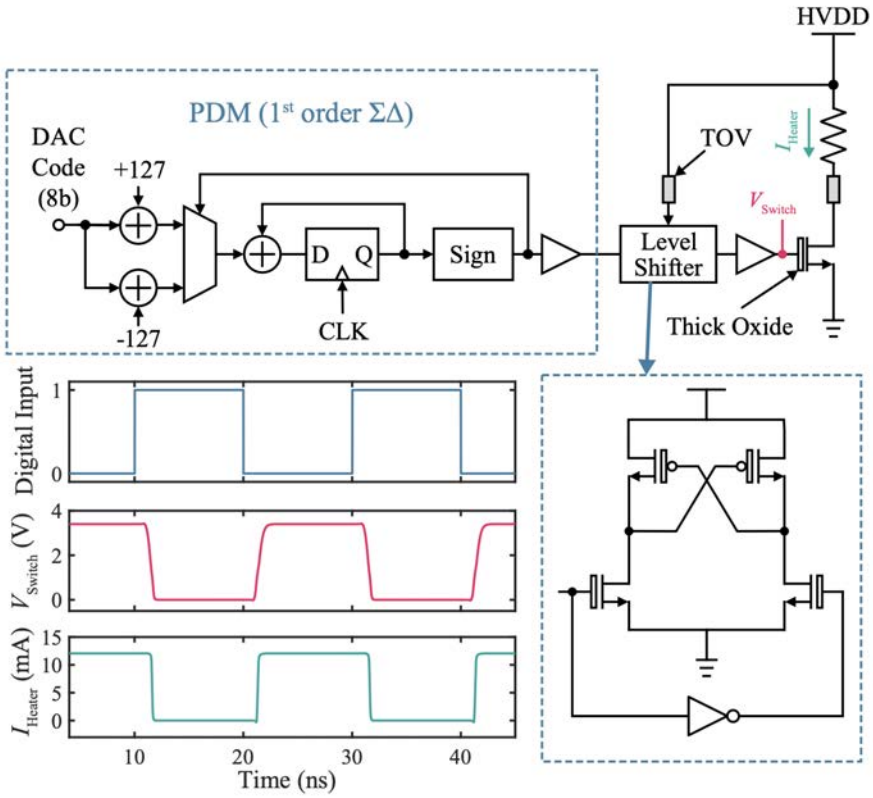


Figure 3.6: Overview of the PDM-driven switch-mode driver connected the photonic heater element, as well as the simulated time-domain waveforms of PDM signal, heater switch gate voltage, and heater current.

the bus ridge waveguide and adiabatically tapered back to a single-mode width of 400 nm for routing to the antenna.

3.5.3 Switch-Mode Heater Driver with PDM Modulator

During beam scanning operation, each heater is expected to consume 20 mW of power (π shift in average across the FOV), which amounts to 10 W for a 500-element OPA. This is already comparable to the total power budget for the entire LiDAR system, which leaves no room for CMOS power. Therefore, it is imperative to ensure that the electrical power consumed in the CMOS circuit is much smaller than the heater itself (i.e. high power efficiency).

Figure 3.6 shows our heater controller circuit design to ensure maximum driver efficiency. We utilized a switch-mode driver comprising of a single NMOS switch connected to the heater.

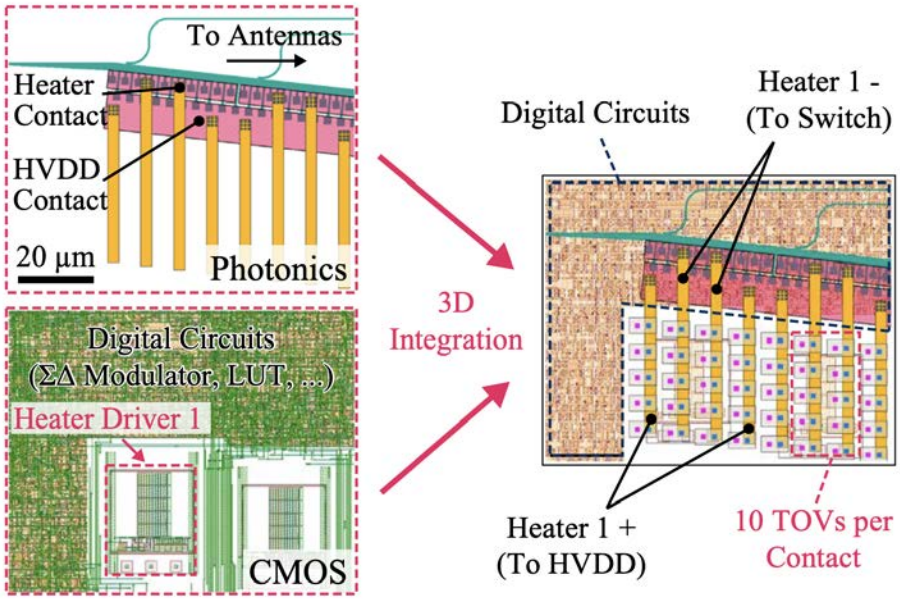


Figure 3.7: Relative placement and TOV-based routing of the thermal phase shifter and the CMOS driver chain, surrounded by digital circuits including PDM modulators and on-chip LUT.

Since the voltage swing of the heater is only 3.3 V, it was possible to directly use a thick-oxide device available in the CMOS process without cascoding (unlike [26]). The 1st-order $\Sigma\Delta$ modulator in the digital domain generates a pulse density modulated (PDM) signal [58] with 8-bit resolution to support an element count up to 128 with an extra bit for calibration, and the PDM signal drives the heater switch through a series of buffers and a level shifter. Modulated heater power is finally low-pass filtered and mapped to the heater temperature by the thermal frequency response. The driver chain was designed to support the maximum clock rate of 400 MHz so that the residual ripple of the PDM signal is sufficiently suppressed. The clock rate was ultimately limited by the current waveform duty cycle distortion caused by the finite edge rate of the switch input voltage. In addition to higher efficiency, another advantage of the switch-mode driver with PDM modulation compared to a current-mode DAC is its inherent linearity. Since an NRZ signal is transferred from the driver input to the heater power instead of an analog signal, it completely circumvents the nonlinear relationship between the current and the heater temperature.

Figure 3.7 shows how the heater and controller circuits are located in 3D. Thanks to its simplicity, the resulting switch-mode driver front end is extremely compact (22 μm $\times$ 22 μm), which is smaller than the size of the phase shifter (32 μm long). This enabled us to place the driver front end directly underneath the heater in a pitch-matched fashion and limit the physical connection between the heater and NMOS switch to direct vertical TOV

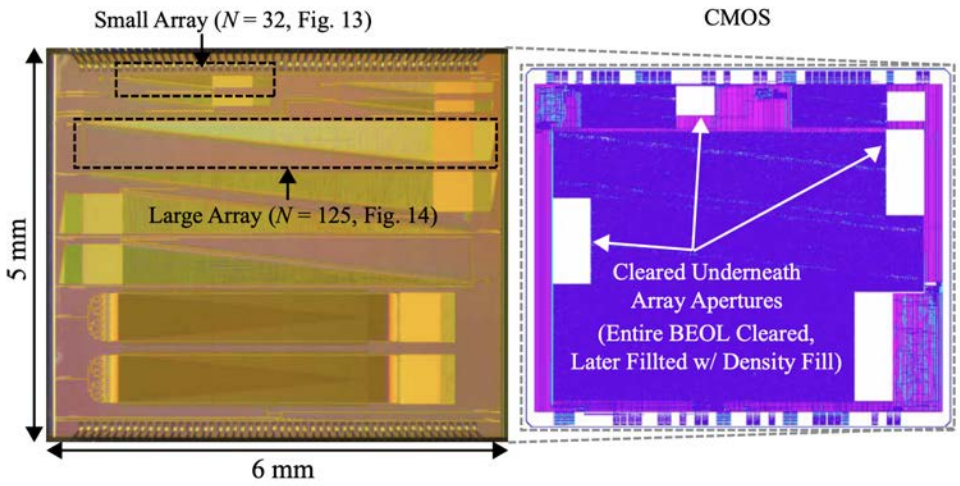


Figure 3.8: Die micrograph of the OPA chip and the GDS image of the CMOS, located underneath the visible photonics layer.

connections. This eliminates potential IR loss from extra wires, which can be a problem considering high heater current (up to 12.12 mA) due to low heater resistance. Ten TOVs per one connection (or 20 TOVs/heater) were used in parallel to ensure electromigration-free operation and further minimize IR loss.

Finally, the remaining CMOS area was simply surrounded by the digital control circuitry through automatic place and route. The highly digital nature of our DAC design and flexible TOV connections also enable efficient utilization of the chip area with an overall floorplan matched to the size of the OPA in the photonic layers. This is particularly important in our 3D integration platform since a significant portion of the silicon area would be wasted if the footprint of the CMOS and photonics are not matched. CMOS power consumption is a function of clock rate and DAC input code. Even for a 400 MHz clock rate and the worst-case DAC code of 0 (100 activity factor), the simulated CMOS power was 1.6 mW (1.4 mW in the driver chain, 234 W in the PDM modulator), much smaller than the heater power. In reality, the measured thermal bandwidth was 32.3 kHz and, even for 11-bit resolution to support 1000 elements, 5 MHz clock rate (or 150 oversampling ratio) is sufficient ($20\mu W$ CMOS power).

3.6 Experimental Results

Figure 3.8 shows the die micrograph of our OPA chip. It includes two types of OPAs composed of identical devices but with different element counts (N = 32 and N = 125) to demonstrate the scalability of the platform. The corresponding top-level layout of the CMOS

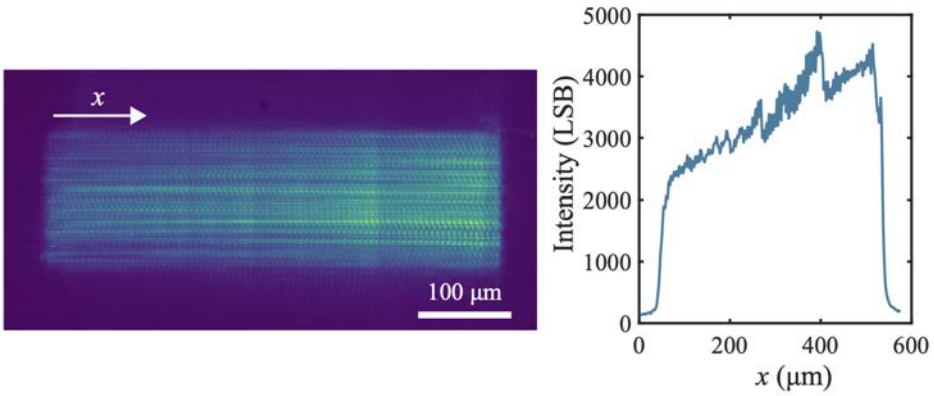


Figure 3.9: (a) Near-field image of the illuminating small array aperture and (b) its cross section along the grating antenna

layer is also shown. Note that the area underneath the array aperture was cleared during the design phase and later filled evenly with the regular density filling structure to eliminate unexpected stray light scattering from the CMOS side.

Figure 3.9 shows the near-field image of the small array variant and the intensity cross section along the grating antenna. The intensity variation over the aperture was kept within 2.25 dB, confirming the effectiveness of the apodization technique illustrated in Section 3.5.

Characterization results for the phase shifter and controller circuit are presented in Figure 3.10. As can be seen from the power versus DAC code plot (Figure 3.10a), the step size is more compressed in the high power regime due to the increased heater resistance from temperature shift. Still, an approximately 40 mW range is achieved, which is sufficient for a 2π phase shift. Also, the heater resistance is close to the expected value of 270Ω even if the resistance is slightly increased in the high power range (Figure 3.10a), again due to the temperature dependence.

The transient response of the heaters is also characterized (Figure 3.10)c using a setup where the heaters are driven by a square wave from a waveform generator, while a photodetector is aligned with the beam direction corresponding to one of the voltage levels, and the photodetector output is monitored on an oscilloscope. Note that the heater temperature and the photodetector output may have a nonlinear relationship depending on the beam shape and the size of the photodetector aperture. To minimize the impact of this potential nonlinearity, the amplitude of the square wave was lowered until it maintained a consistent edge shape while still showing sufficient extinction ratio. The average time constant from the cooling and heating transitions was $4.94\mu s$, corresponding to 32.3 kHz bandwidth assuming a single-pole response.

Figure 3.11a shows the bench-top setup for characterizing the performance of an OPA and

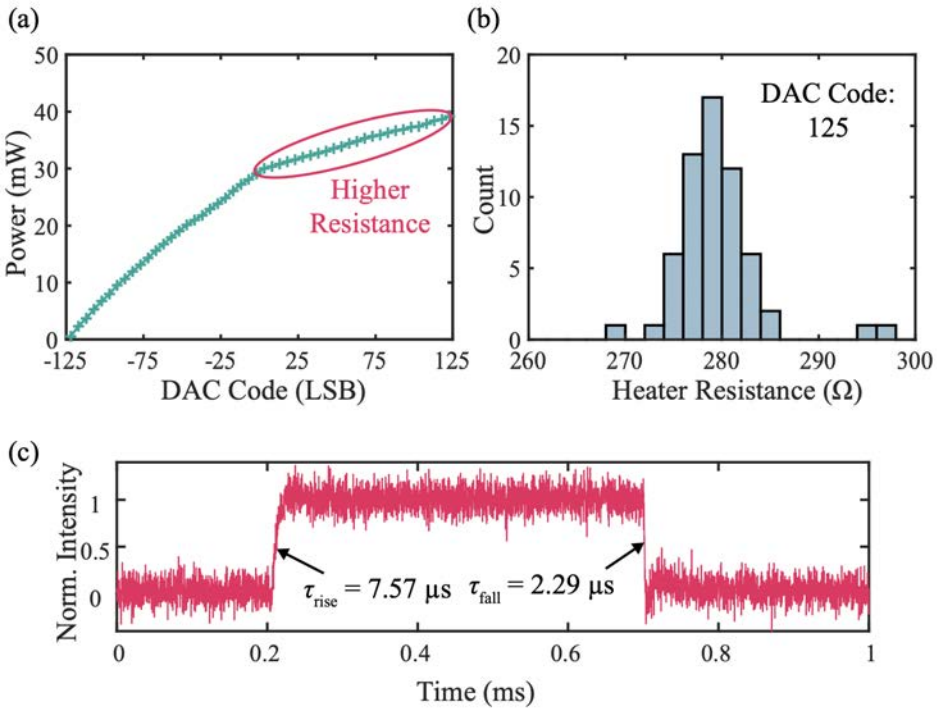
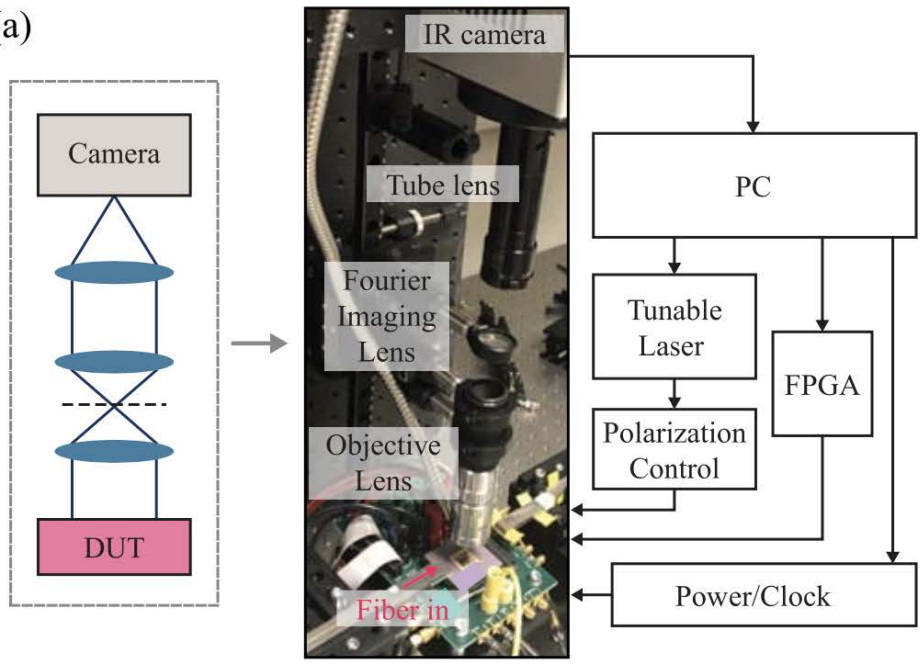


Figure 3.10: Measurement results of the DAC and heater. (a) Power versus DAC code. (b) Statistics of the heater resistance. (c) Thermal transient response.

also for controlling the on-chip phase shifters. The far-field intensity pattern of the emitted light is directly captured on an IR camera (320 × 256 pixels at 30 μm pitch) through a standard 3-lens Fourier imaging setup. The chip surface is placed at the working distance of the 10x or 20x infinity-corrected objective lens (f=20mm or 10mm), and the telecentric setup formed by the Fourier imaging lens (f = 100 mm) focused on the back focal plane of the objective and the tube lens (f = 200mm) forms an image of the Fourier plane at the camera sensor (e.g. 27.5×22 degrees field-of-view and 0.086 degrees per pixel for the 20x objective). The resulting image is then streamed into the PC, which also controls the tunable laser, power/clock source, and the FPGA that programs the on-chip lookup table (LUT) that stores the phase shifter DAC codes via a serial interface.

Using the setup in Figure 3.11a, we can also carry out the array calibration to adjust the DAC LUT for optimum beam quality. Figure 3.11b shows the pseudo-code of the simple local-search algorithm used for OPA calibration. It starts from the first phase shifter in the bus waveguide, adjusts its DAC input until it finds the code that maximizes the target beam quality (e.g. the sidelobe suppression ratio (SLSR) or foreground-background ratio, both of which can be extracted from the IR image), and then moves on to the next phase shifter. The camera image was averaged from multiple shots to minimize the impact of readout noise.

(a)



(b)

```
Pick desired beam position coordinate
Set initial DAC inputs
while i < maximum iteration count do
    for n ← 1 to N do
        for ϵ ∈ {···, −Δ_i, 0, +Δ_i, ···} do
            DAC_{n,i} ← DAC_{n,i−1} + ϵ
            Measure beam quality
        end for
        Pick ϵ_best with maximum beam quality
        DAC_{n,i} ← DAC_{n,i−1} + ϵ_best
    end for
end while
```

Figure 3.11: (a) Experimental setup including far-field imaging optics. (b) Pseudocode of the local search-based beam calibration process.

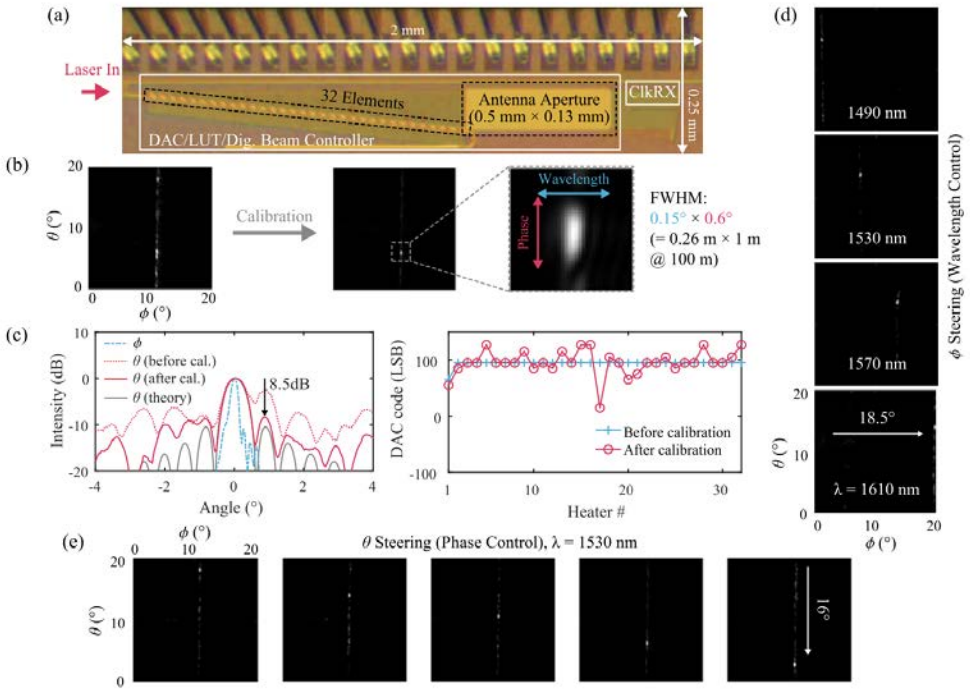


Figure 3.12: Demonstration of beam calibration performance and beam-steering capability for the OPA with 32 elements. (a) Detailed die micrograph of the 32-element phased array. (b) Far-field image of the array, before and after calibration. (c) Cross section of (b) along ϕ and θ, as well as the DAC code distributions. (d) and (e) Beam-steering along ϕ and θ through laser wavelength and on-chip phase shifter control.

Once it reaches the final phase shifter, one iteration cycle is done. Each step took about a minute on average, largely limited by the camera-PC interface and FPGA-PC interface for LUT configuration, which was implemented using a generic API and python script. This can be improved in the future by writing custom software and firmware to streamline the data movement. In the data presented in this work, two iteration cycles were performed using SLSR as the optimization target metric.

The impact of the beam calibration is highlighted in Figure 3.13. In this example, the smaller variant with 32 elements (aperture size: 0.5 × 0.13 mm) is used. The beam is tightly focused in the ϕ direction ($FWHM_\phi \approx 0.15$ degrees) and almost identical to the expected beamwidth for a 0.5 mm aperture and $\lambda = 1530$ nm, again confirming wide effective aperture and the validity of the apodized antenna design. However, without array calibration, significant sidelobe and limited beam contrast in θ is clearly observed from the far-field intensity image as well as the dotted line in the cross-section plot. After calibration, the beam quality is significantly improved, achieving SLSR of 8.5 dB. Figure 3.12 also presents

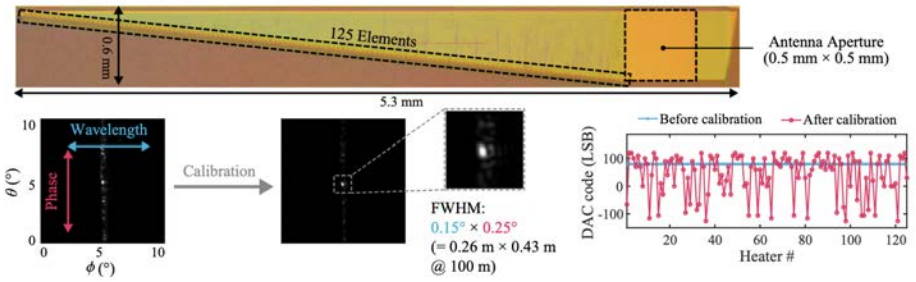


Figure 3.13: View of the larger 125 element array and calibrated beam performance

a full two-dimensional beam steering demonstration based on the same calibration process with two iterations at each beam position, over 16 degrees in θ via phase shifter control and over 18.5 degrees in ϕ by tuning the input laser wavelength across a 120 nm range centered around 1550 nm. Again, the steering range in θ can potentially be extended by reducing the antenna pitch (e.g. 56 degrees was demonstrated in [20]). A final beamwidth of $FWHM_\phi \times FWHM_\theta = 0.15$ degrees $\times$ 0.6 degrees is achieved, which corresponds to a lateral resolution of 0.26 m $\times$ 1 m at 100 m range.

Since our OPA is completely integrated into a single chip through a 3D integration process and enables a pitch-matched layout design of phase shifters and control circuitry, further OPA scaling implies nothing other than additional silicon area and comes with zero overhead for I/O support or electronics routing/placement, as shown in the large array demonstration result shown in Figure 3.13. The antenna aperture size of this large array demonstration is four times larger (0.5 $\times$ 0.5 mm), which results in a measured FWHM beamwidth after phase shifter calibration of 0.15 $\times$ 0.25 degrees (matching the approximately 0.1 to 0.2 degrees beamwidth requirements of long-range automotive LiDAR systems).

3.7 Remaining Challenges and Potential Solutions

Based on the discussion so far, at least in terms of beam resolution and steering range, we conclude that an OPA-based low-cost solid-state beam scanner is becoming a reality. However, to be successfully deployed in real systems, one should also consider metrics related to reliability and maximum emission power.

Ensuring robust operation against temperature variation is particularly important for automotive applications. According to the industry standard [21], consistent operation across 40 to 105 degrees C is a minimum requirement. Due to the same thermo-optic effect, we utilized to build phase shifters, on-chip waveguides will undergo significant index shift in the presence of ambient temperature variation. Fortunately, the impact of ambient temperature shift is common to on-chip waveguides, and the relative phase difference between antennas,

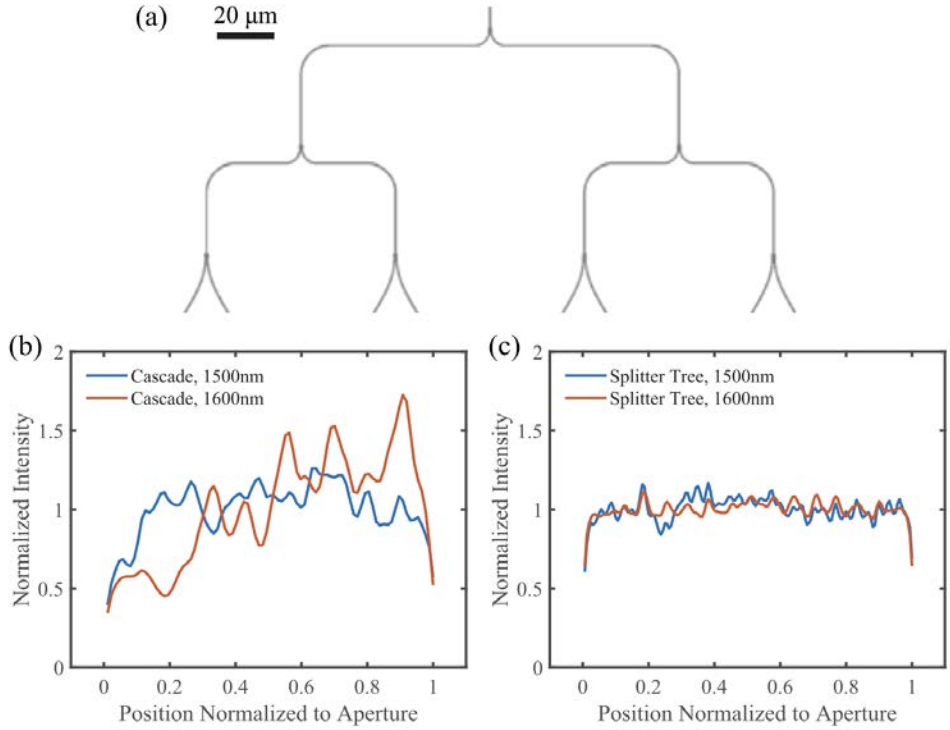


Figure 3.14: (a) Layout of 1x2 splitter-based optical distribution tree available in our process. (b) Measured power distribution across the cascaded directional coupler-based OPA at 1500 and 1600-nm wavelength. (c) Same data as (b), but from the OPA with the distribution using the design in (a).

which affects the actual beam pattern, is largely unaffected.

One notable exception is the case where the optical distribution network introduces physical path length mismatch. For instance, distribution through cascaded evanescent coupling used in this article or for the sub-row distribution in [20] can cause phase slope bias when temperature changes, which results in beam direction offset. Path mismatch can be primarily removed through the careful layout. For the cascaded coupler structure, the optical path after the directional coupler up to the antenna should gradually decrease so that it cancels the additional delay in the bus waveguide. Alternatively, an inherently symmetric distribution architecture can be used. For example, a tree of 1×2 splitters shown in Fig. 3.14 [20], [53] ensures precise match from the OPA input to individual antennas. Moreover, a splitter tree can maintain even power distribution over a wider range of wavelengths compared to the cascaded directional coupler, as shown by the measurement results from our test structures.

One may also note that the emission angle of the grating antenna is also subject to absolute effective index, which makes it temperature sensitive. A change in material can

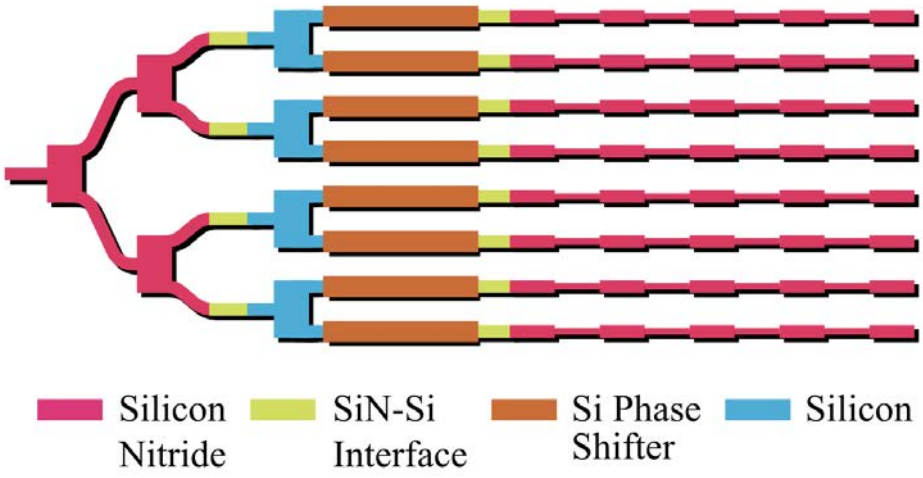


Figure 3.15: Proposed temperature-insensitive, high-emission power OPA architecture realizable in our platform through process customization.

resolve the issue: since our integration platform enables photonics process customization, it is possible to replace the silicon-based antennas in the OPA with antennas based on a material with a lower thermo-optic coefficient, such as silicon nitride (a similar grating design based on SiN was demonstrated in [53]).

To be used in long range (200–300 m) applications, it is also desired that the OPA can support the maximum permissible radiation power set by the laser safety regulation (10 dBm for IEC60825-1 Class 1 in C-band [67]). In our prototype, the fiber-to-chip edge coupler has 3 dB loss, and the theoretical emission efficiency (η = Pbeam/Pinput) for the 2-D array (N = 125, d = 4 μm, dgrating = 700 nm, lgrating = 500 μm) calculated from the basic array directivity equation in [18] and single antenna gain is 9 dB. Considering 3 dB/cm waveguide loss and a 5-mm- long distribution network (Fig. 3.13), the overall insertion loss is expected to be at least 13.5 dB. Even with a reduced antenna spacing (e.g., d = 2 m) and unidirectional antenna design [55] to improve the efficiency by 6 dB, 20 dBm of optical power is needed at the OPA input to support 10 dBm beam despite extra loss from couplers and phase shifters. It is well known that silicon waveguides are unable to handle more than 10dBm without exhibiting significant loss due to two-photon absorption (TPA) [65]. SiN integration can also resolve this issue since SiN can handle much higher power density [53].

Fig. 3.15 shows a high-emission power OPA architecture with reliable operation against temperature drift, realizable in our platform. High power density at the input and first few stages of splitter tree are handled by SiN waveguides.

Meanwhile, although a bench-top laser was used in this article, a deployable system would eventually include an integrated laser that can support sufficient output power and wide wavelength tuning range (50–100 nm) to enable 2-D steering. Fortunately, a variety of

lasers based on compound materials originally developed for C-band fiber optic communications can be re-branded for free-space optics [37]. In particular, III/V semiconductor lasers with large tuning range [63] and high output power [33] are promising because they can be heterogeneously integrated with silicon photonics, removing fiber-to-waveguide coupler loss. Alternatively, rare-earth-doped lasers [49] with high output powers over 300 mW [42] and wide tunability over 46nm [43] can be utilized to further minimize the unit cost. Although it involves an additional optical pump, the integration process is much simpler and done at the wafer scale.

Finally, the simple local search-based calibration algorithm used in this article may not be practical for larger element counts and may fail to find the global optimum within a reasonable time. In fact, characterizing random phase fluctuations in OPAs, is an equivalent problem to acquiring a phase image at the array aperture. Using a near-field image of the intensity distribution at the aperture and a far-field image showing the Fourier image of the wavefront as two inputs, standard phase retrieval algorithms, such as the Gerchberg–Saxton algorithm [29], can be applied to potentially enable one-shot calibration.

3.8 Chapter Summary

In order for integrated optical phased arrays to be a reliable solid-state alternative to mechanical beam scanners, the technology must support thousands of elements at wavelength pitch in spite of stringent area/power/cost constraints. To resolve I/O and electronics density problem, we proposed an OPA design based on a wafer-scale 3D photonics/CMOS integration platform. With device and circuit designs that actively utilize flexible, dense TOV connections between photonics and CMOS, we have demonstrated a compact single-chip OPA with wide-range 2D beam steering and array scalability beyond 100s of elements. Our array calibration result also highlights the importance of flexible phase control in achieving robust OPA performance, especially for thermo-optic phase shifter designs that can facilitate compact, efficient control circuits suitable for large-scale OPAs. Furthermore, 3-D integration allows the photonics to be highly customized independent of electronics for target applications (e.g., C-band coherent optical receiver utilizing Ge photodiodes [9] and hybrid laser/optical gain [49], [43]), opening up unlimited opportunities for integrated free-space system design.

Chapter 4

Focal Plane Arrays

As described in the previous chapter, OPAs represent a promising approach to integrating beamsteering functionality into a CMOS-compatible solid-state form factor. However, our analysis indicates that scaling OPAs to larger and larger sizes to support automotive and other high precision applications poses significant challenges that may preclude their use in high-volume applications. Specifically, maintaining coherence across large apertures, reducing power consumption, and enabling active on-chip calibration are key improvements that need to be made. This chapter will describe a new approach which circumvents these issues by using integrated focal plane arrays (FPA) to perform solid-state beamsteering.

4.1 FPA System Concepts

FPAs (shown in Figure 4.1) operate by leveraging a lens to separate light from different far-field angles into different physical locations at the focal plane. By placing the sensor chip at the focal plane of the lens, the lens effectively maps different far-field beam locations to different locations on the chip according to equation 4.1, where Δx is the distance between the center of the lens and the emitter, and f is the focal length of the lens.

$$\theta_{out} = tan^{-1}(\frac{\Delta x}{f}) \approx \frac{\Delta x}{f} \tag{4.1}$$

This effect can be used to make a solid-state beamsteering transmitter by designing a chip which has an array of emitters that can be electronically turned on or off. Each enabled emitter will form a discrete beam at a different angle in the far-field, and active steering can be achieved by changing which emitter is enabled over time.

Table 4.1 below shows a performance comparison between OPAs and FPAs. The addition of the focal length as a design parameter relaxes hard tradeoffs that were constrained in OPA implementations. For OPAs, in order to narrow the beam width, the size of the aperture needs to be increased (L_{array}), necessitating large, power-hungry arrays that need to maintain coherence across a long distance. On the other hand, with FPAs, the effective "aperture"

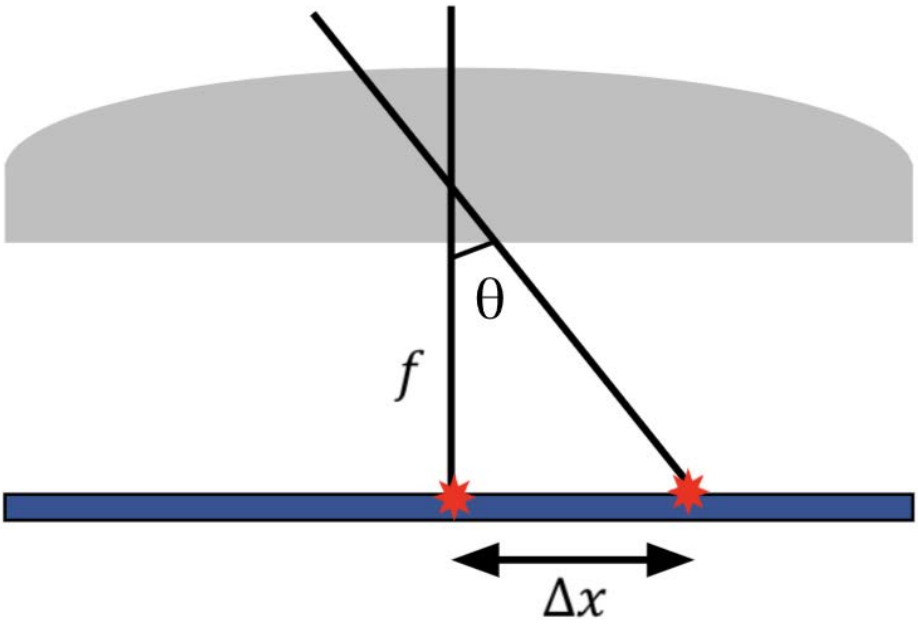


Figure 4.1: Ray optics perspective of far-field beam angle as a function of emitter location on the chip relative.

is determined by the illuminated area of the lens. As a result, with the same chip size, a narrower diffraction-limited beam width can be achieved simply by selecting a larger lens and illuminating a larger area.

For angular resolution, an FPA behaves differently than an OPA since it essentially has a discrete number of beam positions that can be selected from. The minimum distance between these beam positions is set by the ratio of the distance between antenna on the chip and the focal length. For an OPA, the beam center position can effectively be tuned to arbitrary locations, only limited by the phase shifter resolution. In practical use cases, setting the distance between beam positions to be narrower than the beam width is not useful, and so the effective resolution is limited by the beam width rather than the distance between beam center locations.

Another key metric is field of view (FOV). For an OPA, the alias-free FOV is limited by the ratio of the antenna spacing and the wavelength. As a result, FOV of a device is fixed by the design upon manufacturing and is strongly tied to the design rules of the process. In addition, since the required element spacing is often smaller than the size of phase shifters and waveguide routing for OPAs, 2D beamsteering must be accomplished with laser wavelength tuning. This significantly complicates laser design, and places constraints on the feasible FOV. With FPAs, FOV is dependent on the ratio between the focal length and the size of

	focal plane array	optical phased array
Beam Width	$\propto \frac{\lambda}{\pi D_{lens}}$	$\propto \frac{\lambda}{2L_{array}}$
Angular Resolution	$\frac{\Delta x}{f}$	$\propto DAC_{resolution}$
FOV	$2tan^{-1}(\frac{L_{array}}{2f})$	$2sin^{-1}(\frac{\lambda}{2\Delta x})$

Table 4.1: Performance comparison between OPAs and FPAs

the array. As a result, the resolution and FOV can be smoothly traded off using the same chip simply by selecting lenses with different focal lengths and tiling multiple sensor chips side-by-side. With this flexibility, a single FPA silicon design can be scaled to support any application.

An example implementation of an FPA transmitter is shown in Figure 4.2. In this case, the electronically controlled emitter array is implemented by taking a single input laser source and routing it to one of many optical antenna on the chip. This configuration forms a single output beam which can be controlled in 2 dimensions. The $1 \times N$ row decoder optical switch controls the beam angle in one dimension, and the $1 \times M$ column decoder switch controls the beam angle in another dimension.

4.1.1 Optical Switch Network

Figure 4.3 shows an implementation of the $1 \times N$ decoder switch. This switch is made up of a binary tree of 1×2 thermal Mach-Zehnder (MZ) switches. Each switch decides whether its single input source is routed to the top output or the bottom output. Note that unlike in OPAs, where each antenna has some fraction of the input light and must have a specific phase applied by the heater, the FPA implementation routes all of the input light into only one of the output waveguides. As a result, heaters which actuate switches that are not on the optical path can be disabled. This is a key benefit of the FPA approach, since it means that with a binary tree implementation, only $log_2 N$ heaters (where N is the number of output antenna) need to be enabled, leading to significant advantages in power savings as the array

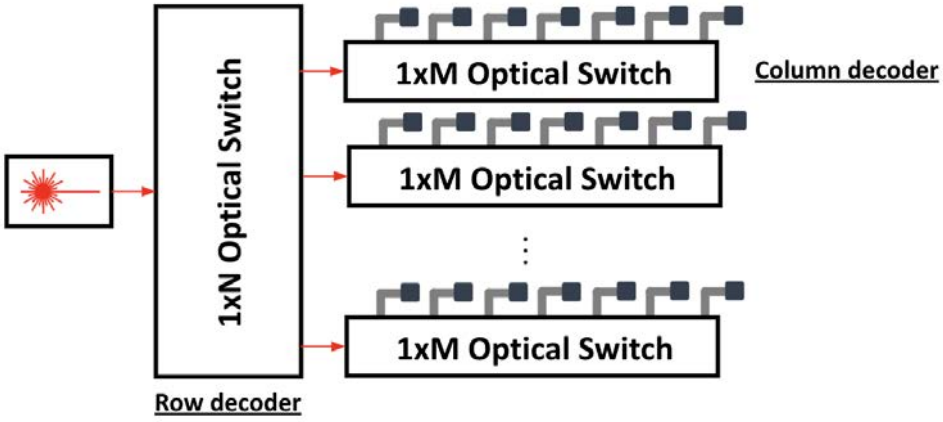


Figure 4.2: Example implementation of a FPA-based solid state beamforming transmitter

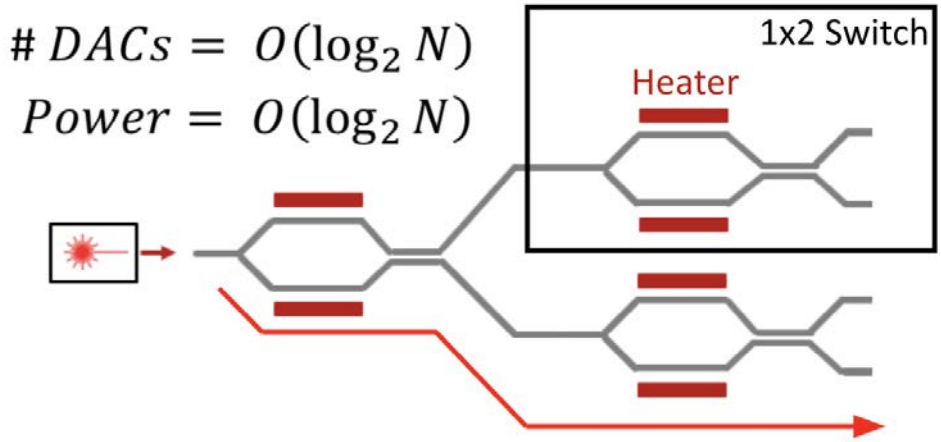


Figure 4.3: FPA optical switch network consisting of a binary tree of thermal MZ switches used to route the input laser source to the desired emitter

increases in size. Furthermore, since the light only illuminates one antenna at any time, the FPA does not need to maintain coherence across a large distance. This effectively means that array size can continue to scale independent of foundry process quality.

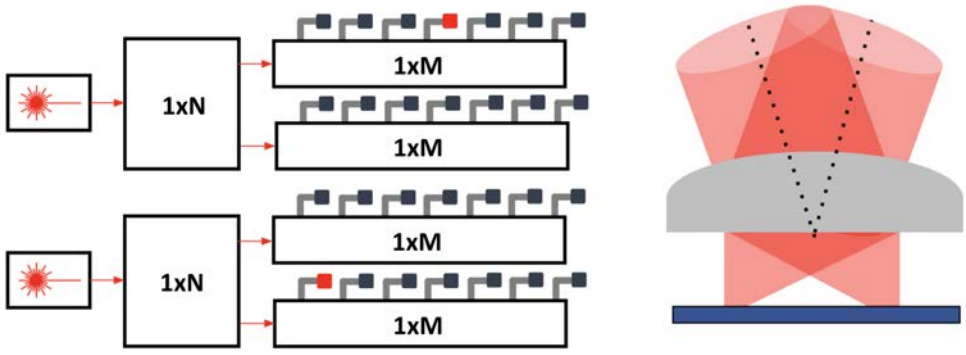


Figure 4.4: Pair of independent arrays sharing a lens to form a 2-beam FPA system

4.1.2 Multi-Beamforming

Both OPAs and FPAs have the capability of forming multiple beams simultaneously, although they achieve this in different ways. OPAs can form multiple beams by linearly summing the phase control signals from two different beams and applying them to the phase shifters. While this is fairly simple for a small number of beams, increasing the number of beams requires significant extra signal processing circuitry at each phase control element, and also can lead to the beam quality degrading as the non-idealities of the phase control for each beam interact with each other due to imperfect gain control. For FPAs on the other hand, multiple beams can be formed simply by placing multiple arrays next to each other at the focal plane of a common lens as show in Figure 4.4. Each array is completely independently controlled, and does not need and additional hardware to maintain coherence between arrays. With this technique, the FOV of an FPA system can be increased while maintaining the same frame rate simply by tiling additional arrays or additional chips side by side and creating additional beams.

4.1.3 Closed-Loop Calibration

One critical issue that plagues OPA designs is maintaining calibration during operation. As mentioned in the previous chapter, due to manufacturing and environmental variation, the exact phase distribution of light across the aperture may not be what is expected during the initial design. This issue is exacerbated as the size of the aperture continues to grow, necessitating calibration. In OPA implementations, this requires active phase measurements across the entire aperture, which is not feasible to do when scaling the aperture size and FOV.

In contrast to this, since FPA designs do not need to maintain coherence across the array, the calibration is a much simpler problem to solve. Manufacturing and environmental

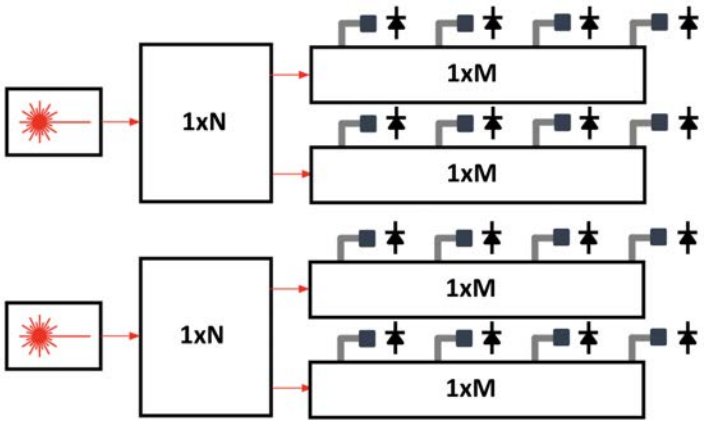


Figure 4.5: Closed-loop on-chip calibration system, using local photodetectors at each antenna to measure distribution of power vs. far-field angle

variation impact the relative phase difference between adjacent arms of the MZ switches, causing some light to leak into other beam locations. However, since the two arms are just a few microns away from each other, process and environmental variation is likely to be very well correlated, leading to reduced impact on the output switch. Furthermore, detecting the non-idealities in the switch configuration is significantly easier in an FPA implementation since instead of measuring phase differences between light in each antenna, we can simply measure the intensity of light in each antenna to determine the current beam pattern. As shown in Figure 4.5, At each antenna, a small amount of power can be tapped off into a local photodetector. The measured current in each photodetector will be proportional to the optical power in the output beam in that given position. Since the entire switch network is implemented as a binary tree, the extinction ratio of any particular switch in the network can be computed with:

$$ER = \frac{\sum P_{onpath}}{\sum P_{offpath}} \tag{4.2}$$

Where $\sum P_{onpath}$ represents the sum of the power measured on all downstream antenna connected to the enabled output on a particular switch, and $\sum P_{offpath}$ represents the sum of the power measured on all downstream antenna connected to the disabled output on a particular switch. Then the heater code for that particular switch can be linearly optimized to maximize this extinction ratio. This measurement and optimization is fairly simple to implement on chip, and can be run live during operation to maintain beam quality under changing conditions.

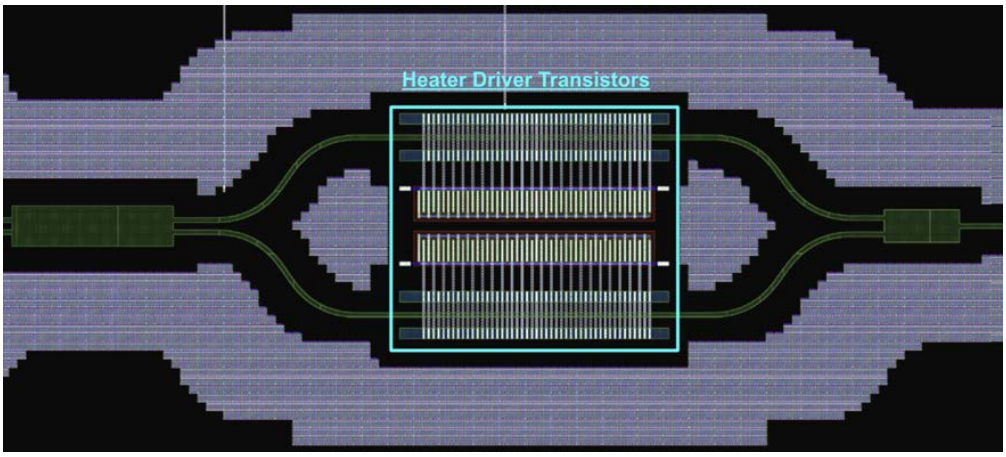


Figure 4.6: Layout image of a generated thermal MZ switch. This is the output of a single generator which creates circuit layouts for both the heater driver transistors and the photonic layouts for the heaters, bends, and MMI splitter/combiners

4.2 Implementation and Results

Using the architecture outlined in the previous section, we implemented a FPA in a zero-change 45nm SOI CMOS process. This process enables SOI transistors to be placed together with photonic devices on the same mask, just a few microns away from each other. This tight integration between CMOS and photonics enables the very high density connection between the control circuitry and the optical switch network. An example of a thermal MZ switch in this process is shown in Figure 4.6.

This thermal MZ switch unit cell is tiled and connected together to form a binary tree for the row decoder, which enables one dimension of beam steering control. In order to enable the 2D beamsteering with high resolution, the column decoder switch must be implemented in a compact form-factor. This is accomplished by creating a similar binary switch tree, but flattening all of the switches to be co-linear, rather in a tree layout as shown in Figure 4.7. In this layout, higher level switches in the tree (shown as the middle switch in the figure), are interleaved between leaf switches (shown as the switches on the left and right), and waveguides are automatically wrapped around and routed between these switches. The 1×4 switch architecture shown in Figure 4.7 can be tiled side by side, with an additional switch in between to extend the design to a 1×8 architecture.

Using the described row and column decoder, we assembled a large FPA design consisting of two parallel 512 element arrays, as well as 6 other smaller 16 element arrays for testing purposes. The large 1024 element FPA consists of two parallel input gratings along with two loopback gratings to enable active alignment of a fiber array. This light then passes through a pair of 32-element row decoders, followed by 16-element column decoders, and finally the output gratings.

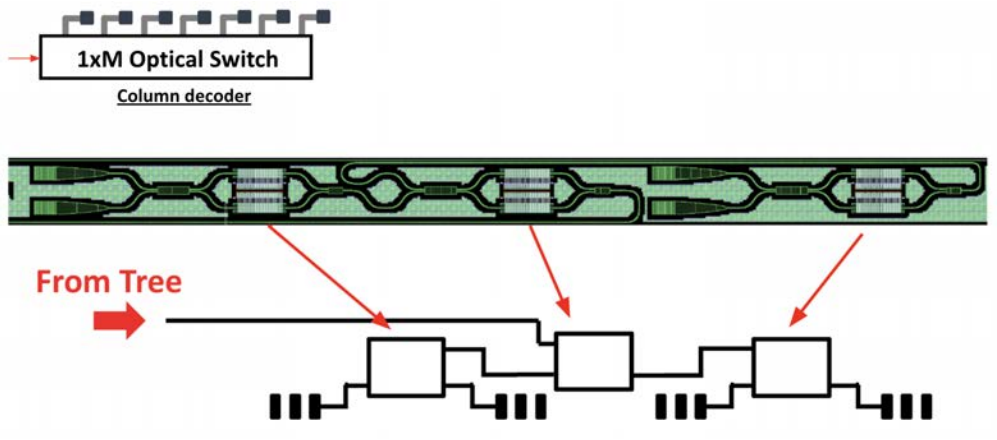


Figure 4.7: Implementation of a 4-element binary switch tree in a flattened form factor, enabling compact layout for maximum resolution

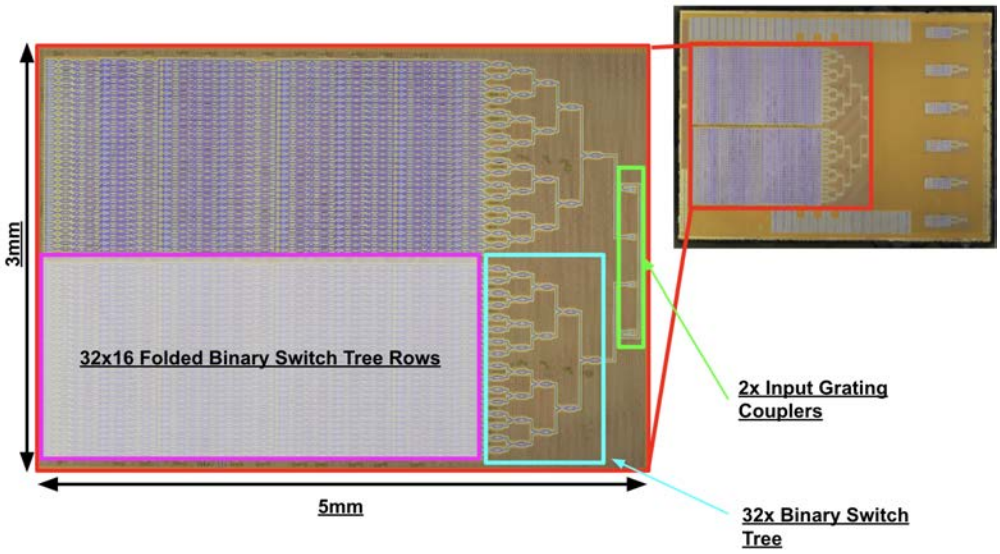


Figure 4.8: Die photo of the FPA chip implemented in 45nm SOI CMOS

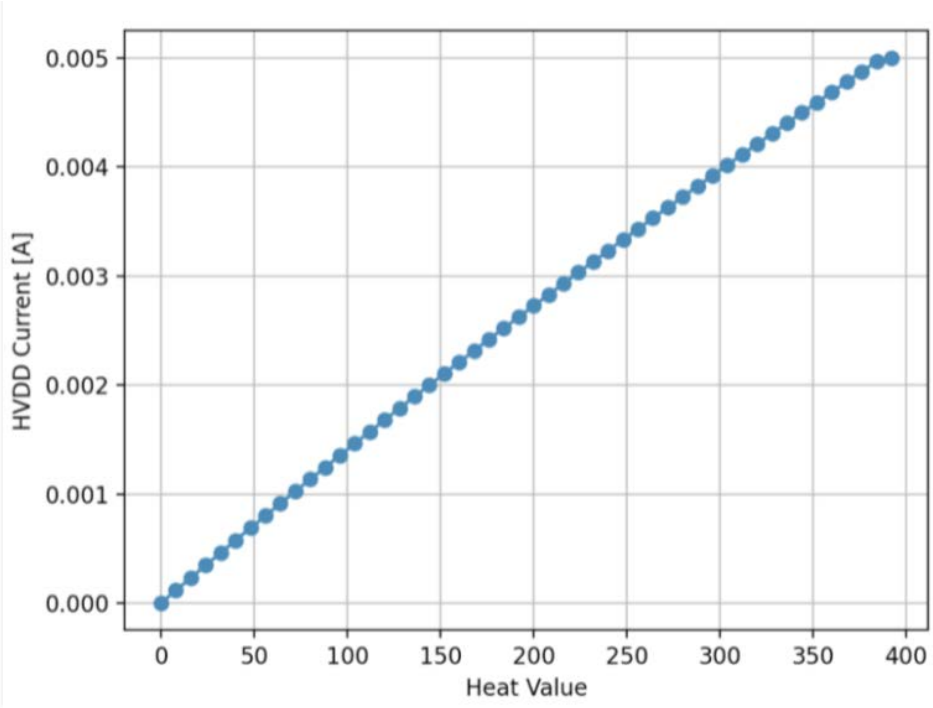


Figure 4.9: Measurement of current draw from the heaters as a function of DAC code

Each array also contains several heater DACs and a steering controller. The steering controller enables the beam to automatically sweep from one position to another after setting a target heat value for each switch configuration.

Figure 4.9 shows characterization results of the heater. In this experiment we sweep the heater DAC code and monitor the heater current. This data confirms that communication to the chip, heater drivers, and heater resistors are all functional. Figure 4.10 shows a near-field optical image of one of the small arrays. This confirms that light is uniformly distributed across all the 16 output gratings. Unfortunately, although we sweep the heater values, we don't observe the optical switching that we expect. The same thermal switch design is fully functional on other chips, but does not seem to function on the FPA implementation. Further debugging is needed in order to diagnose the issue.

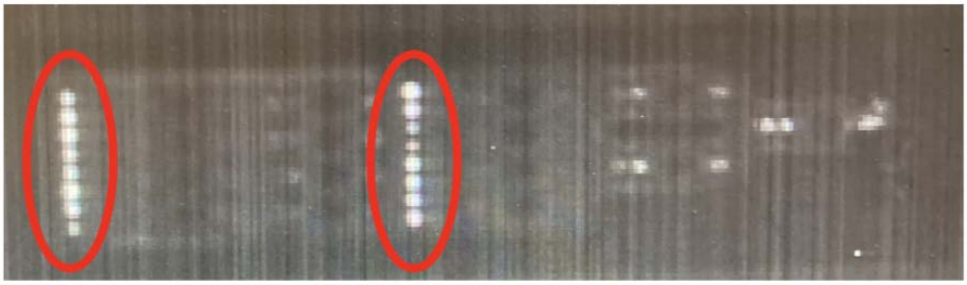


Figure 4.10: Near-field image of the small FPA variant showing relatively uniform emission when control signals are turned off

Chapter 5

Electronic-Photonic Design Methodology

5.1 Introduction

Designing large, scalable, integrated electronic-photonic systems such as the focal plane arrays outlined in Chapter 4 necessitates a design workflow that enables novel device development, scalability to systems with many thousands of optical devices, compatibility with electronic-photonic integration, and a robust verification methodology.

To date, many independent approaches to establishing such a design workflow have been pursued [27], some focused on supporting custom mask layout for research applications [28] and others focused on creating an ecosystem around assembling foundry-made PDK components. While each of these approaches has been successful within their respective domains, the needs of emerging high-volume integrated photonic systems bridge the gap between these approaches.

Many of the major EDA companies offer integrated photonic design environments, which rely on close foundry collaboration for PDK releases and which frequently treat low-level optical devices as black boxes that system designers snap together in a graphical environment [61, 14]. These types of environments are useful in lowering the barrier to entry for designing simple systems with existing foundry made components. However, as competition increases in the photonics space, designers will need to differentiate themselves with novel devices, design techniques, and large-scale photonic designs, which is challenging with these tools.

A significant amount of research in the academic space leverages open source tools such as gdspy or KLayout [28, 40]. These tools allow designers to have fine grain control over the geometry of the primitive devices. However, they lack close integration with standard simulation and verification tools and the higher-level features required to enable easier large-scale photonic system design and integration. As a result, most designers need to combine many different commercial and custom tools to successfully create and verify a new design.

To address these shortcomings, we present the Berkeley Photonics Generator (BPG), an

open-sourced, modular, and easy-to-use framework for end-to-end photonic system design [11]. BPG utilizes a plugin-based architecture with flexible support for third party tools, enabling a full design flow managed from a single source. BPG provides low-level design capabilities, enabling custom device design, as well as high-level integration features, enabling large-scale system design. By encouraging technology-agnostic generator-based design methodologies, BPG enables significant design reuse across different projects and technology nodes. BPG aims to create an agile, robust, and easily adoptable workflow with a rich set of reusable libraries to boost the photonic design ecosystem.

The remainder of the chapter outlines the design workflow from layout to verification. Section 2 discusses the overall architecture and design philosophies of the BPG framework, and Section 3 details the layout generation framework. Section 4 presents an overview of the *Photonic Compiler* and describes how a technology agnostic design is automatically converted into technology specific design. Section 5 details BPG's optical testbench generation capabilities and interface to third party optical simulation tools. Section 6 describes BPG's schematic and behavioral simulation capabilities. Section 7 highlights several examples of large-scale integrated electronic-photonic designs, which were uniquely enabled by the BPG framework. Finally, in section 8, we conclude and discuss future work.

5.2 BPG Design Philosophy

The BPG software platform was designed as a highly modular, plugin-based framework. One of the overarching goals of the BPG framework was to establish a unified ecosystem that brings together the various tools and methodologies required in electro-optic device and system design. The primary features enabled by BPG include: layout generation, technology compilation, optical simulation, behavioral simulation, and verification.

In order to create such a unified ecosystem, BPG utilizes a centralized design manager (Fig. 5.1) to control the flow of a design through the various stages of its development. Each of the primary features provided by BPG corresponds to one of the steps followed during a standard design process. Specialized plugins within each feature category have been developed to provide the required functionality, or manage an interface to a 3rd party tool. The central design manager stores the various elements of the design in a centralized database, and provides an interface between plugins that can be used to pass data through the design flow. Plugins are accessed through simple function calls, so that the top level design flow can be managed through a simple script (Fig. 5.2). BPG's plugin-based design philosophy allows new tools to be easily integrated into the design environment, and enables additional features to be added without disrupting existing workflows. In addition, by separating the design flow into distinct modules, multiple designers can collaborate and specialize in different aspects of design within a common framework. Finally, by storing the design content in a centralized database, BPG obviates the laborious and error prone task of manually porting the same design between various tools.

BPG is written in the Python programming language. Python was chosen for several

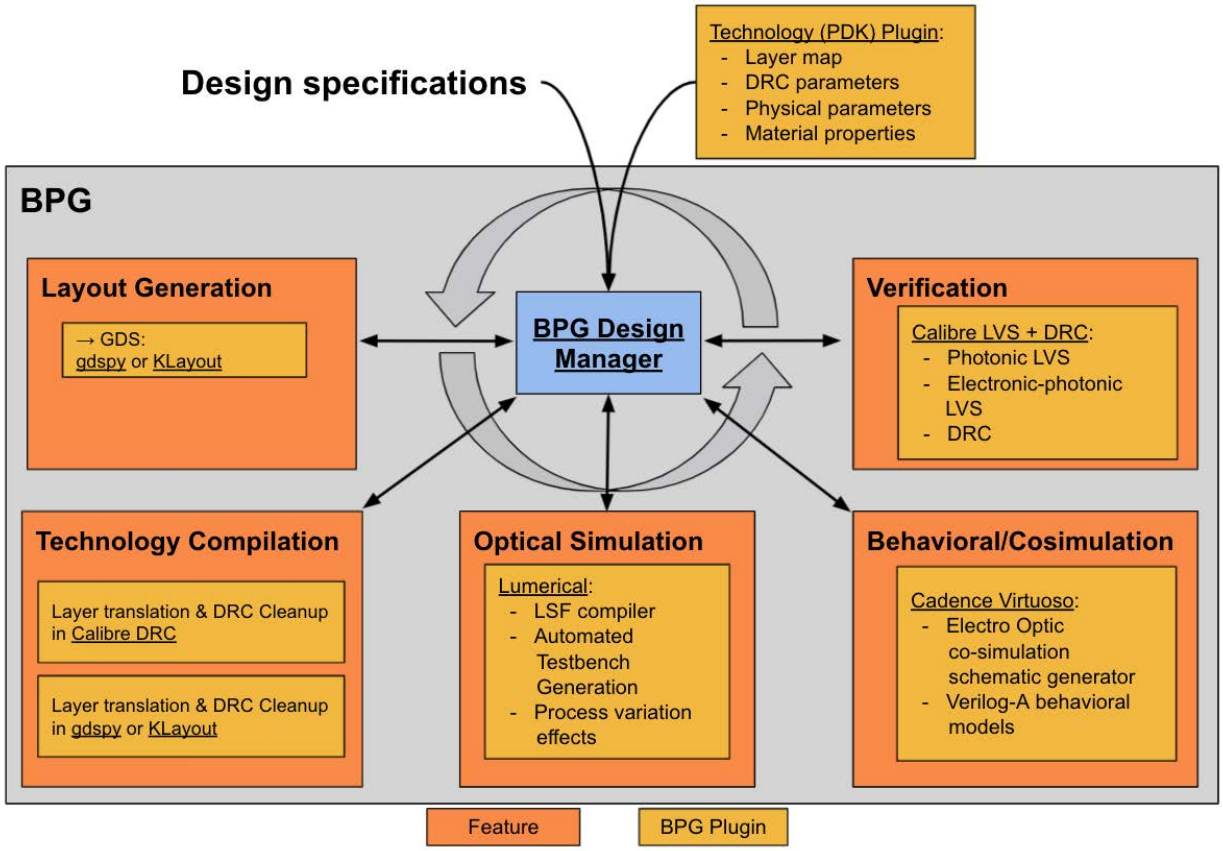


Figure 5.1: Main features provided by BPG, and the plugins that support those features. A central manager coordinates data flow between the various plugins. Designer specifications are provided as input to the design manager. A technology plugin, created once per technology process or provided as part of a PDK, abstracts foundry specific details from the design process.

reasons. First, it is a very popular, easy to learn, and open source language. Python has many established libraries (such as NumPy, SciPy, pandas, and more) that are extensively utilized in the scientific community, that can be used to perform computation, optimization, data import/export, and other operations, eliminating the need for developers to create their own specialized tools. In addition, Python is simple to install on most platforms, enabling easy distribution and adoption of BPG. Finally, Python makes it simple to perform tasks such as dynamic module importing during run time, which are useful when user input specifications can modify the class hierarchy.

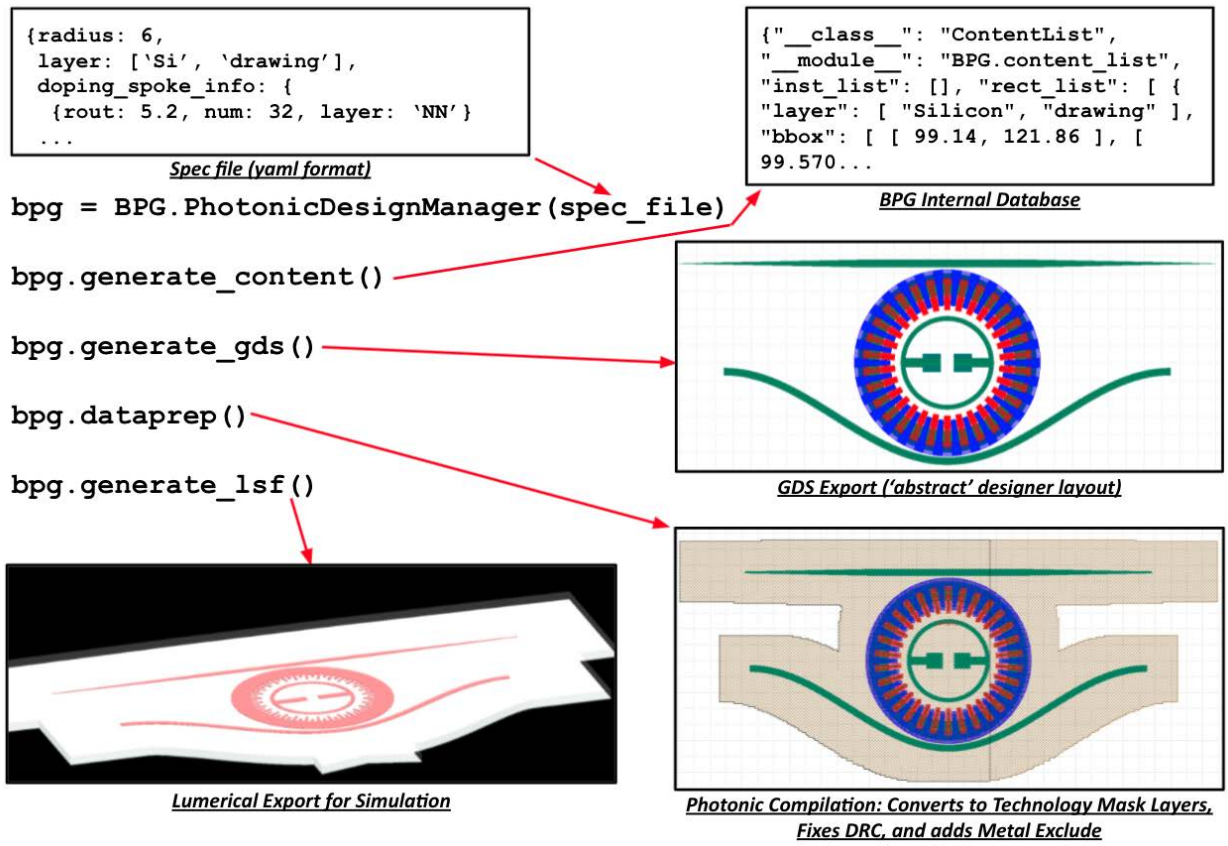


Figure 5.2: Sample design flow within the BPG framework. A top-level script creates a PhotonicDesignManager based on the user-provided specifications. The design manager can then, for example, create the internal representation of the layout, export to GDS, call the *Photonic Compiler* to convert the abstract layout into foundry layers, and export a simulation setup to Lumerical.

5.3 Layout Generation and Generator Based Design

BPG encourages a generator-based approach to designing photonic layouts. Rather than having designers hand-craft a final layout for every possible implementation of a given block through a GUI, the generator design approach requires that designers encode their intent as a generalized algorithm with a set of input user specifications. By doing this, the design flow is reproducible, automatically documented through code, and can be parameterized to be used in a wide variety of circumstances.

While PCells embrace a similar philosophy, we seek to extend them in several key ways. First, the BPG framework encourages and enables technology-agnostic design. By allowing

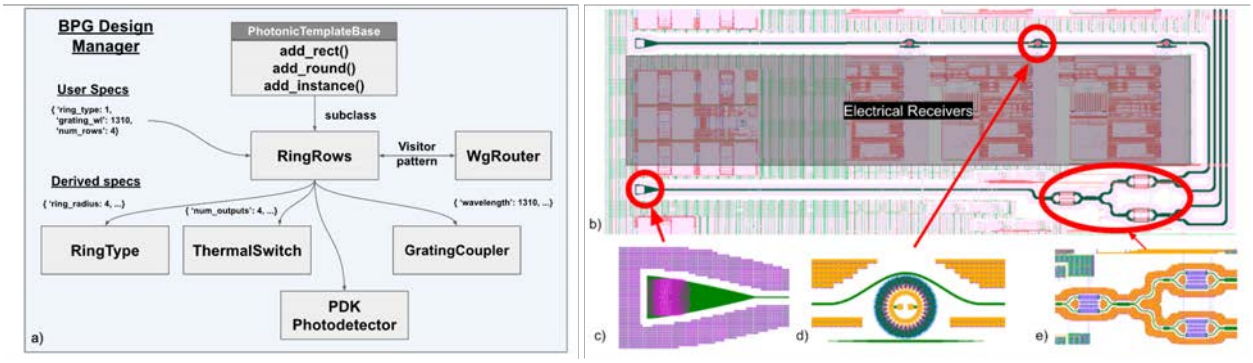


Figure 5.3: a) Example structure for a ring row photonic layout. The top level RingRow class subclasses PhotonicTemplateBase, which by inheritance gives it access to all of the primitive shape and hierarchy creation methods. Using the specifications provided by the user, RingRow then dynamically creates and instantiates all the different component types. These components can themselves be generators which subclass PhotonicTemplateBase such as RingType and ThermalSwitch, or they can be standalone PDK components imported from GDS. Finally, all of these components are assembled and routed together using WgRouter. b) Generated layout implementation of a ring row in GF45RFSOI [71]. c-e) Callouts of (c) grating coupler, (d) ring resonator, and (e) thermal switch tree, emphasizing the complex geometry required in photonic structures.

generator authors to write code for an abstract technology, it generalizes the applicability of a single generator to any number of platforms. This approach is discussed further in Section IV. Next, through BPG, generators have a very close connection with any optical simulator of choice. By leveraging this close connection, it is possible to automatically adjust and optimize the parameters for a specific technology that are not known when the generator is written.

BPG is built on top of Berkeley Analog Generator (BAG) [17], a generator-based framework for creating schematics, layouts, and testbenches for analog and mixed signal electrical circuits. BPG adapts and extends BAG to provide the underlying hierarchy management framework as well as new procedures required for photonic design. One of the key additions made by BPG is support for complex, free-form, polygon-based geometries used in photonic components (Fig. 5.3b-e). BPG is fully compatible with electronic circuit layout generators written for BAG, enabling easy co-design of circuits and photonics in monolithically integrated processes from within a single environment. This co-design of electronics and photonics has been leveraged to create components such as compact thermal optical switches [71]. In addition, the integration with BAG enables use of the complex wiring tools that are available and commonly used for analog circuit designs.

PhotonicTemplateBase is BPG's primary Python abstract base class that provides a common interface for photonic layout generators, and encapsulates concrete methods for

creating and instantiating designs.

The PhotonicTemplateBase low level API contains methods that allow the designer to manipulate primitive geometry and layout objects. Methods such as add_polygon, add_rect, and add_round can be used to draw unrestricted primitive geometry objects. BPG supports polygons with free-form edges, and vertices on a quantized user-defined resolution grid. The add_instance method can be used to instantiate other PhotonicTemplateBase objects within the current layout at a desired location and orientation. BPG supports any-angle instancing (as opposed to restricting to 90 degree cardinal orientations), with intelligent caching that re-computes the generator for an instance only when needed, reducing memory and computation.

Designers can utilize the add_photonic_port method to create a PhotonicPort object within the layout. A PhotonicPort is a BPG primitive object that records, among other information, a location, angle, orientation, width, layer, and name. PhotonicPorts can be strategically placed at important locations in the photonic devices, such as terminals of the input and output waveguides, to demarcate locations that can later be used for automated connectivity and abutment. PhotonicPorts can also serve as location markers during photonic testbench generation and can be extracted from the design to provide I/O location and device information during testing.

The higher-level API provided by PhotonicTemplateBase allows designers to easily assemble subcomponents into a larger, hierarchical design. Methods like add_instance_port_to_port can be used to abut two photonic ports, automatically performing any required translation or rotation. Finally, PhotonicTemplateBase provides an abstract method draw_layout which is implemented by subclasses to define the layout for a particular system or component.

To enable easy connectivity and scalability of designs, BPG provides a semi-automated waveguide router class, WgRouter. This router is capable of connecting two ports using straight waveguide routes, parameterized turns (including circular and Euler), and 'S'-bends. The router keeps track of the overall route length as the route is drawn, so that length matching can be performed. In addition, a waveguide bus fanout router utility is provided to automatically fan out arrays of waveguides to target offsets, ensuring maximal spacing between waveguides and that no waveguide crossings occur. Since this functionality is used and extended in many generator classes, WgRouter is implemented as a plugin, rather than additional methods of PhotonicTemplateBase. This allows WgRouter to use the visitor design pattern to access PhotonicTemplateBase methods and create complex assembled layouts, without increasing complexity of PhotonicTemplateBase (requiring multiple inheritance), and without requiring large specification dictionaries to be passed around to complete the routing tasks.

Foundry provided or pre-existing IP blocks can be encapsulated using the GDSImport subclass of PhotonicTemplateBase. By adding PhotonicPorts at the inputs/outputs of the block, the device can be treated within the hierarchy as any native BPG-created block, and is able to be automatically instantiated, abutted, and rotated using BPG's assembly methods.

These features allow designers to manage the complexity of large-scale systems, such as a WDM ring row (Fig. 5.3). With this methodology, it is simple for users to change high

level specifications, like ring types, number of rings, and wavelength targets, and have the generator automatically design and assemble the layout.

5.4 Technology Compilation and Process Portable Design

5.4.1 Process Portability

To enable process portability, BPG provides an abstract PhotonicTechInfo class. This class contains technology and process parameters required to abstract away hard-coded technology-related constants from layout generators. The PhotonicTechInfo class is implemented once per process by the foundry, PDK provider, or designer. In addition to basic design rules such as minimum width and space, via sizes and enclosures, and area constraints, PhotonicTechInfo also supports default design-related parameters, such as default waveguide routing layers and widths. Layout generators can then access these technology parameters using a standardized API. In this way, designers can abstract away basic technology-specific parameters (for example, by calling PhotonicTechInfo.wg_layer, rather than hard coding a layer), and can re-run the same layout generator using different PhotonicTechInfo technology plugins to regenerate the layout in multiple processes.

To further enable process portability, designers are encouraged to use 'abstract' design layers, such as 'waveguide', 'nitride', 'undoped silicon', or 'n_doping', when implementing the design. In this way, the layout generator captures the designer's intent, rather than the obfuscating minutia of mask design required in a particular process. Thus, by using technology agnostic intent layers, a layout generator can be reused across processes, with subsequent operations 'compiling' the abstract design into a particular technology.

5.4.2 Photonic Compiler

BPG utilizes a *Photonic Compiler* to convert the layout geometry from the internal database representation to external formats, for example GDS. During the compilation process, the abstract designer layers are mapped to mask layers that the foundry requires to fabricate the structure. The compilation procedure utilizes a series of ordered Boolean layer operations (and, or, xor) and polygon manipulation (growth, shrink, snap, Manhattanize) steps, called *Dataprep* (Fig. 5.4). This *Dataprep* routine is standardized per technology, and is provided as part of the technology plugin.

The *Dataprep* routine can provide significant functionality beyond direct layer mapping. First, *Dataprep*'s layer and polygon operations can automatically add utility layers required by the foundry, such as fill-exclude, marker layers around photonic shapes, or create negative mask regions for partially etched silicon, removing these implementation level concerns from the designer. *Dataprep* geometry manipulation can also be utilized for automatic removal of DRC violations [3]. Sequential 'over-of-under' and 'under-of-over' sizing operations remove

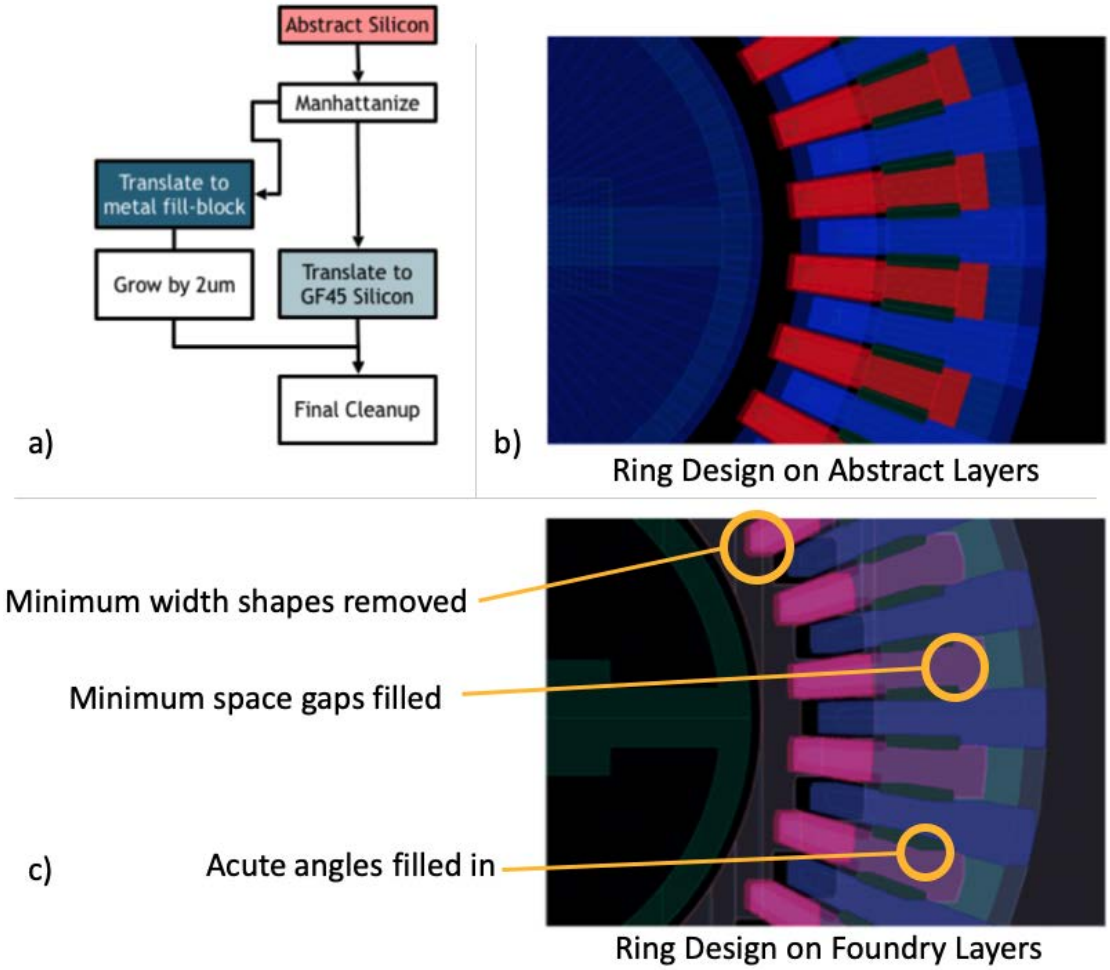


Figure 5.4: a) Sample *Dataprep* routine, showing the operations performed on shapes on the 'silicon' designer layer. In this example process, the silicon shapes are Manhattanized (de-angled) to a user specified grid. In parallel, a 2μm larger fill exclude is generated around the silicon geometry. Finally, these shapes are mapped onto the appropriate foundry layers. b) Sample of a ring modulator in GF45RFSOI, showing the 'abstract' designer layers before *Dataprep*. c) Sample of a ring modulator in GF45RFSOI, showing the foundry layers after *Dataprep*. Callouts indicate several types of DRC violations that are resolved in the *Dataprep* process.

free-form minimum width and space violations, respectively [3]. BPG provides a template for *Dataprep* routines, with a standardized geometry 'instruction set' that includes the add, subtract, rough-add, Manhattanize, and, xor, and snap operations, and regex layer parsing. Finally, specific regions in the design (such as foundry IP cells) can be excluded from standard

Dataprep routine through the use of marker layers and appropriate Boolean subtraction operations.

Utilizing the *Photonic Compiler* to automate DRC correction enables vastly more complex and process portable layout generators. 'Correct by construction' methodologies, which force the designer to produce DRC clean layouts within the layout generator, are not sustainable in advanced process nodes, where the 1000's of DRC rules make even simple curvilinear photonics elements challenging to clean. As an example, for a ring modulator designed in GF45RFSOI [59], a direct mapping of the 'designer intent' geometry to the foundry layers results in tens of thousands of non-waivable DRC violations. However, applying the *Dataprep* routine results in zero remaining non-waivable DRC violations.

5.4.3 Photonic Compiler Implementations

Through BPG's plugin architecture, the computation backend can be customized as desired; however, an open source implementation using gdspy [28] as the computation backend is included, and an ultra-fast backend based on Mentor Graphics's (Siemens EDA) Calibre engine [56] has been implemented and used for wafer-scale photonic compilation. As an example, a 5.5 mm x 8.9 mm electro-optic transceiver IC in GF45RFSOI featuring over 160 ring modulator and detector devices, as wells as 100's of mm of curvilinear waveguide routing and many grating couplers, was generated in less than 2 minutes, and was compiled into a fully DRC clean output using the Calibre *Dataprep* backend in 30 minutes [59].

5.5 Optical Simulation

BPG enables robust and convenient optical device design and characterization through photonic testbench generators. While current work has focused on integration with Lumerical [4], one of the industry leading photonic device design tools, BPG's open plugin framework can be adapted to support other third-party software by simply implementing a standard Python interface between these tools and BPG.

BPG's LumericalTB class provides a framework for generating re-usable testbenches. The DUT's generator and specifications are passed to the LumericalTB, which automatically instantiates the device. The DUT's PhotonicPorts can be accessed to correctly place access waveguides to ensure the proper mode is launched. BPG supports multiple simulation objects, such as simulation boundaries, mode sources, and various optical monitors. In addition, direct Lumerical Script commands can be injected into the testbench using code objects. Within the testbench generator, designers specify which results should be saved from Lumerical for post processing within BPG or using an external tool. Because the testbench generator can access properties of the DUT and of the technology plugin, the same testbench generator can be reused across multiple devices and technologies; for example, a generic MxN transmission characterization testbench can be used for multiple devices such as AWGs, MMIs, or directional couplers. Testbench reuse enables more robust and standard-

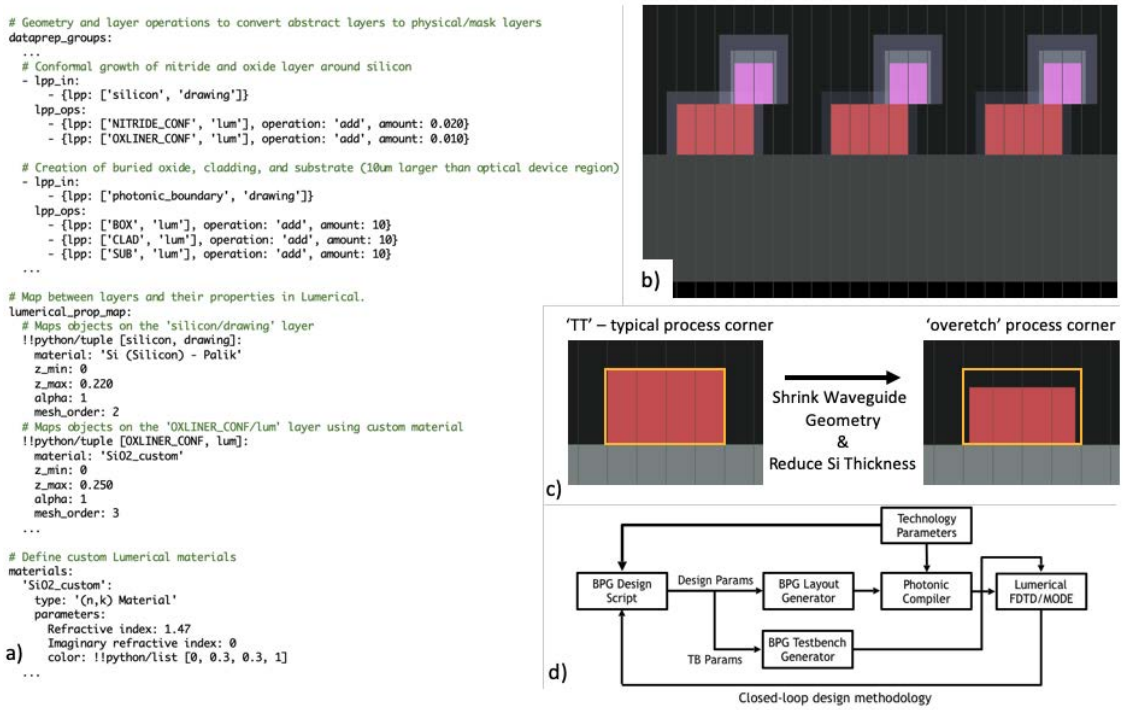


Figure 5.5: a) Sample Lumerical *Dataprep* routine for a dummy process. The dataprep_groups section defines geometry manipulation and layer operations. The lumerical_prop_map section defines how layers map to Lumerical materials, thicknesses, and mesh ordering. The materials section allows custom materials for the technology to be defined. b) Sample output of a grating coupler that models a foundry process which includes conformal liners around the silicon (red) and polysilicon (pink) grating teeth. The only user-defined geometry are the silicon and polysilicon grating elements. All liners and background material are automatically generated through the *Dataprep* routine. c) Example of two 'photonic process corners'. Different *Dataprep* recipes, with adjusted parameters reflecting 'typical' and 'overetched' process corners, run on the same designer geometry to produce different simulation outputs. d) Closed loop design procedure.

ized optical device verification prior to fabrication. Furthermore, codifying the testbench as a generator reduces the possibility of introducing mistakes compared to manually recreating the supporting geometry and simulation objects.

Upon testbench generation, the *Photonic Compiler* is used to convert the DUT and testbench geometry into Lumerical Script code, which can be run via a standard Lumerical workflow. The DUT geometry is converted from the same internal database that creates the tapeout mask outputs; thus, the designer can ensure there is no mismatch in version or parameters between what is simulated and what is fabricated. For export into Lumerical (or

other optical simulators), rather than producing layers that correspond to tapeout masks, the *Dataprep* routine can instead be customized to output layers that model the physical structures present upon device fabrication. This conversion again allows arbitrary geometric and Boolean layer manipulation, which enables complex foundry flows to be modeled accurately.

For example, in some process flows, conformal layers (such as oxides or nitrides) are grown over lithographically defined geometry (such as silicon). Figure 5.5 demonstrates how a Lumerical *Dataprep* flow can accurately model such fabrication processes as a geometric oversize of the silicon layer and Boolean addition onto an oxide liner layer. A technology plugin mapping file defines the set of layers and their properties that finally get exported into Lumerical script (Fig. 5.5a). This ease of mimicking process flow allows designers to be confident they are simulating an accurate and complete view of their fabricated devices, without needed to manually specify all the additional non-designed material regions (such as background materials and conformal growth layers) (Fig 5.5b).

In addition to allowing for complicated process modeling, BPG allows customized Lumerical *Dataprep* routines to support 'photonic process corners', in which growth or shrink operations and changes to thickness or material parameters can be used to model process drift, over- and under-etch, and other fabrication imperfections (Fig. 5.5c). As photonic foundries develop further statistical models, BPG's framework can enable robust, cross-corner photonic verification through PDK-provided variation parameters.

In a standard optical design workflow, BPG's testbench generators can be used to quickly and confidently produce complex Lumerical simulation geometries. Designers can easily update layout parameters for their DUT, and by regenerating the testbench, all monitors, sources, and domains are automatically moved to match the new geometry. Finally, the results from the optical simulation can be fed back into Python to create a closed-loop design methodology (Fig. 5.5d). Many Python libraries exist (such as SciPy) to enable intelligent design space exploration and optimization.

5.6 Electro-optic Co-Simulation and Verification

Co-simulation is a critical aspect of system design to verify the behavior of the entire electro-optic system. In the electrical circuit domain, designers create visual schematics or netlists out of circuit primitives such as resistors, capacitors, and transistors. For photonic integrated circuits, the optical primitives comprise compact models of devices that contain extracted behavioral parameters. BPG uses schematic optical device primitives from a Verilog-A electro-optic co-simulation toolkit [57]. These primitives model the time-domain behavior of optical components (both forward and reverse waves), and can be co-simulated with standard electrical components within the Cadence Virtuoso simulation tool [15]. BPG utilizes the underlying BAG framework's schematic generator [17] to create Virtuoso schematics that are capable of simulating both photonic and electronic devices. A schematic generator, similar to a layout generator, is a codified methodology for describing a circuit's schematic

connectivity and device parameters. In BPG, physical parameters from the layout generator, in addition to other user or technology provided parameters, can be passed into the schematic generator class to manipulate a template schematic into a particular schematic instance. BPG can again utilize BAG's underlying methodologies to create an electro-optic testbench that instantiates the schematic, and simulates the combined behavior.

As a motivating example, Figure 5.6 shows the electro-optic simulation flow for a ring-based WDM link. At the end of the layout generator, derived physical quantities, which may not be known apriori (such as routing distance between rings or around bends), are combined with user-specified parameters into a python dictionary (Fig. 5.6a). The central design manager then passes these parameters to the schematic generator (Fig. 5.6b), which manipulates the schematic template (Fig. 5.6c) to implement the design. The final schematic can be hierarchically combined with an electronic schematic to produce the final electro optic schematic (Fig. 5.6d). The waveguide lengths in this schematic accurately reflect the physical lengths, which are not known apriori, and which are extracted from the layout generator. A final simulation output of the optical eye diagram is shown in (Fig. 5.6e).

Furthermore, the generated schematic can be combined with custom device and netlist extraction routines to perform photonic-LVS (layout versus schematic) verification. While specific implementations of the layout extraction routines are process specific and often proprietary, such tools have been implemented by the authors in GF45RFSOI, a next-generation GlobalFoundries electronic-photonic 45nm process, and Lionix Triplex processes to verify the layout against a simulatable optical schematic. BPG provides plugins to dispatch the LVS routine to the industry standard Mentor Graphics (Siemens EDA) Calibre LVS engine.

In addition to LVS dispatch plugins, BPG provides plugins to support DRC routines using the Mentor Graphics (Siemens EDA) Calibre DRC or KLayout backends.

5.7 Results and Motivating Examples

BPG has been adopted by several research and industry groups, and has been used to design devices and systems in at least five processes: GlobalFoundries' 45RFSOI [50], GlobalFoundries' 45nm next-generation electronic-photonic process [54], LioniX TriPleX [41], Applied Nanotools NanoSOI [5], and AIM Photonics [25]. Designers have used BPG to produce layouts for more than 10 full wafer runs and dozens of individual chips. In this section, we highlight several examples of systems and devices that were enabled by the BPG framework.

BPG's technology agnostic framework and *Photonic Compiler* allows the same design codebase to be reused across multiple processes. For example, the rapid adiabatic coupler introduced in [12] has been fabricated in GF45RFSOI, Applied Nanotools NanoSOI [13], a next-generation GlobalFoundries 45nm electronic-photonic process [38], and LioniX TriPleX (Fig. 5.7a). The same layout generator was used for all devices; the only change to produce these two designs was providing a separate Technology Plugin for each process, and providing appropriate design parameters for each process.

BPG's generator methodology allows for highly flexible and reusable design components. Properly coded generators can take advantage of conditional and looping logic to enable extremely scalable designs. For example, the same binary-tree based 1xN optical switch generator (Fig. 5.7b) can be used to produce 1x4, 1x16, and even 1x512 switch networks. Once coded, multiple instantiations of the parameterized structure can be used without additional design effort.

Seamless integration with electronic devices is enabled through BPG's compatibility with the underlying BAG framework. BAG is a generator-based tool for analog and mixed signal electrical circuits that has been used to produce complex circuits such as ADCs and SerDes transceiver frontends [17]. Using the methods provided by BAG, transistor, resistor, and capacitor based electrical circuits can be generated in a scalable and process portable manner. As example, we show a waveguide heater integrated with its transistor heater drivers only microns away (Fig. 5.7d), with both the electronics and optics generated using a single layout generator. This was achieved by calling a BAG transistor generator class directly from within a BPG generator class. The close integration between BAG and BPG enabled wiring generators to be created to make the connection between the transistors and the photonics. This combined electronic and photonic layout was then verified with Calibre LVS tools before being assembled into a larger design.

BPG has been used for a variety of applications, from large-scale electronic-photonic system-on-chip designs to novel device fabrication. Large scale system-on-chip applications include the design and layout of microring based biomedical and ultrasound detector systems [71, 2]. Novel devices include cryogenic ring modulators and ultra-high shift and FSR rings [31]. Process portable layout generators, such as spoked ring modulator/photodetectors (Fig. 5.7c), have been used to produce highly performant devices, such as 30 GHz/V vertical junction modulators [30]. BPG has been used in novel applications for electro-optic systems, such as cryogenic egress links. Finally, BPG has not just been used in the context of academic research, but has been trusted and verified for use in volume production commercial electro-optic systems [59].

BPG was also a critical tool in implementing the FPA design discussed in the previous chapter. Automated routing tools with low-loss Euler bends enabled the very tight waveguide arrangement required for flattened binary tree switches that can scale to arbitrary numbers of elements. Close integration between circuit generators and photonic generators enabled a single design flow that integrates heater drivers together with photonic components to form a thermal MZ switch shown in figure 5.8. Finally, BPG's simulation tools enabled easy design space exploration for critical sub components such as the MMI splitter and combiners. With this tool, we were able to generate a script which would sweep both the layout and the testbench to evaluate the performance of a given MMI design, and select the optimal generator parameters to meet performance metrics across photonic process corners.

5.8 Conclusion

We presented BPG, a photonic design framework that enables users to craft a workflow encompassing layout, electromagnetic simulation, system simulation, and verification into a robust, automated, and reproducible script. BPG's architectural focus on modularity and extensibility helps support the diverse needs of photonic designers. Whether the goal is to experiment with novel devices that push the boundaries of complex design rules or to design and assemble large electronic-photonic systems for high-volume manufacturing, BPG has the power and flexibility to be customized to each user's specific needs. This has encouraged its use across both research and industry applications, and was a key component to supporting the implementation of the monolithic focal plane array.

Given BPG's ability to easily create technology-agnostic generators, it serves as an ideal platform for creating a universal PDK. In the future, we plan to release a generator library for BPG which can be automatically compiled to a wide range of newly emerging photonic platforms.

The BPG framework is fully open source under the BSD-3 license and is available to anyone now [11]. We include several examples of design flows with layout generators and testbenches to jumpstart new users. Plugins to enable interfaces with other tools such as gdspy, KLayout, and Lumerical are also open source and freely available.

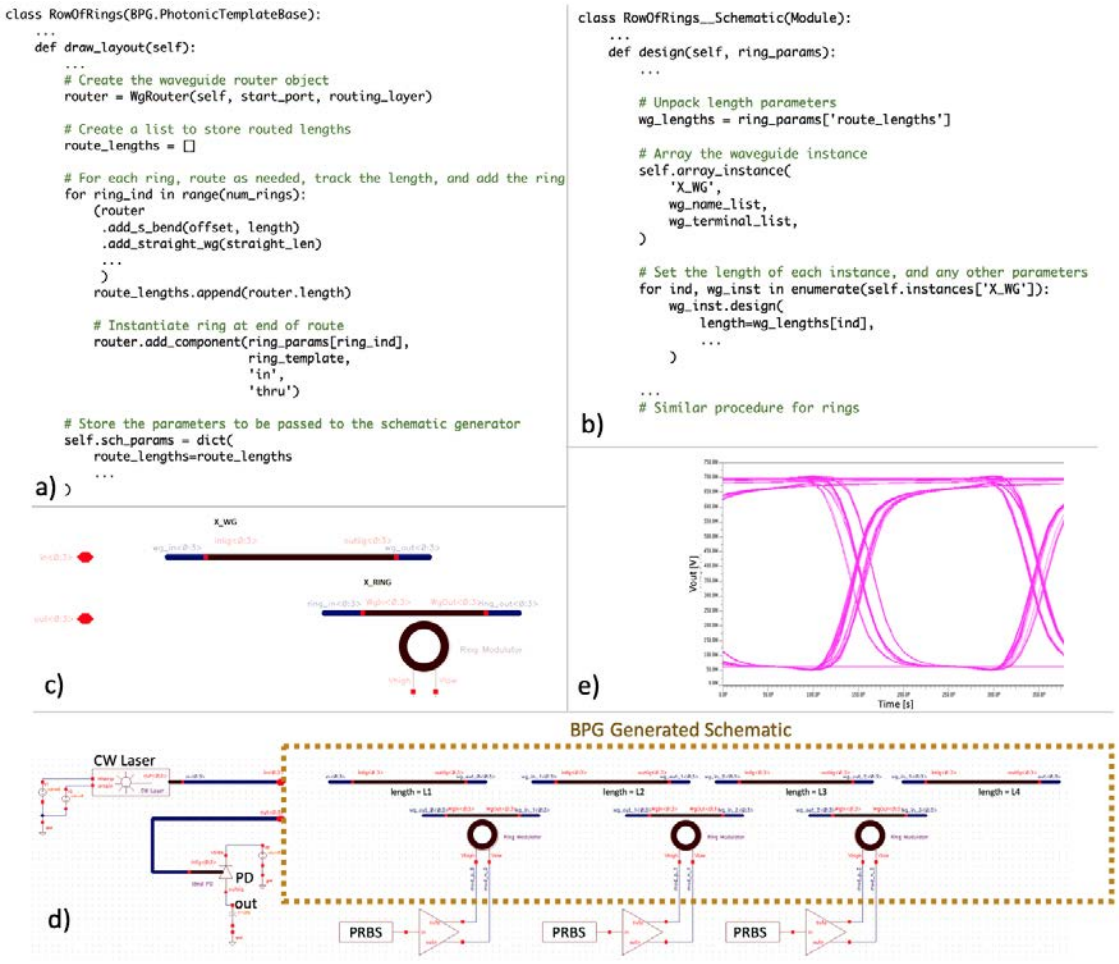


Figure 5.6: a) Layout generator code outline for a row of rings. In the code snippet, a waveguide router is initialized and all the rings are routed and added in-line using the WgRouter class. Note that the route lengths, which might not be known apriori, are computed by the router. The routing lengths, along with other pertinent design parameters, are stored in a sch_params property at the end of the generator code. b) Schematic generator for a row of rings. The contents of sch_params are passed into the design procedure of the schematic generator. The schematic generator arrays the elements of the schematic template, sets the components' terminal connections, and sets the schematic parameters for each element using the provided design parameters. c) Schematic template for a row of rings. The schematic template consists of the waveguide and ring primitives, which will be arrayed and whose parameters are specified by the schematic generator. d) A BPG generated row of rings schematic, embedded in an electro-optic testbench. The electro-optic components such as CW laser, rings, and photodetector are modeled in Verilog-A. The PRBS generator and electrical drivers are standard electrical elements. e) Sample output eye diagram from the output of the photodetector, showing a functional electro-optical system.

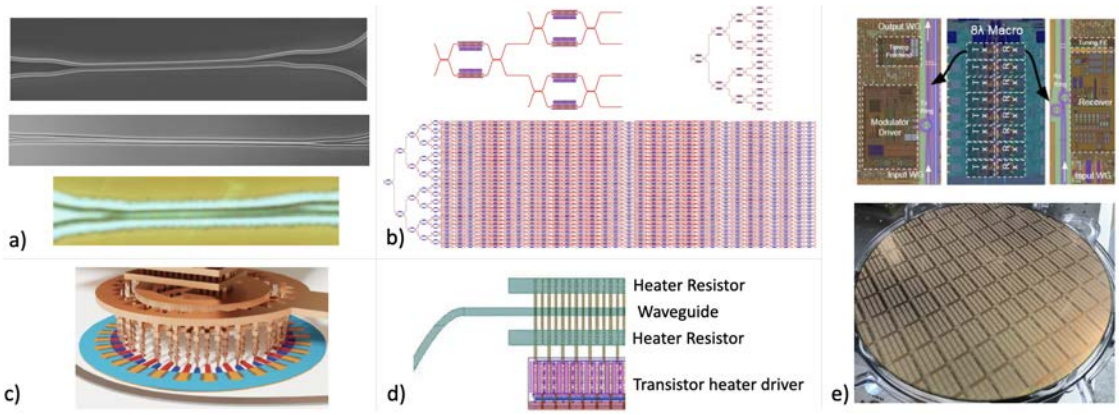


Figure 5.7: a) Three implementations of a rapid adiabatic coupler, implemented in (top) Applied Nanotools NanoSOI [13], (middle) LioniX TripleX, and (bottom) GF45RFSOI. All designs were produced using the same BPG generator. b) 1xN binary switch tree. The same generator is used to generate a 1x4, 1x16, and 1x512 switch. c) High modulation efficiency ring modulator with 30 GHz/V shift [30], implemented in GF45RFSOI. d) Thermal phase shifter with tightly integrated photonics and electronics. A single BPG generator produces both the waveguide and heater geometry, as well as the transistor layout and wiring. The distance between the photonics and electronics is below 3 microns. e) (top) High-speed photonic modulators and ring photodetectors as part of a WDM link [59]. (bottom) BPG generated all photonics components (1000s) on the 300 mm full-wafer run in GF45RFSOI [45].

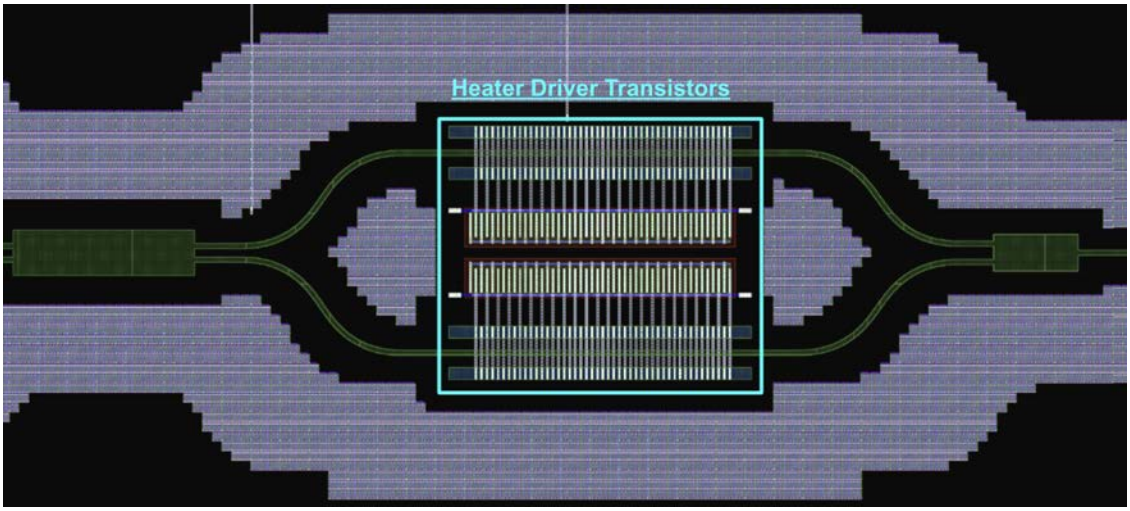


Figure 5.8: Layout image of a generated thermal MZ switch. This is the output of a single generator which creates circuit layouts for both the heater driver transistors and the photonic layouts for the heaters, bends, and MMI splitter/combiners

Chapter 6

Conclusion

As part of this thesis, we made the following key contributions:

- Identified key requirements and issues associated with designing and manufacturing LIDAR beamforming systems for self-driving car applications.
- Implemented first ever OPA-based LIDAR system on a chip using a 3D integrated platform, including optical beamforming, circuit control, and coherent detection.
- Identified specific issues which impede the scaling of OPAs to satisfy the much larger apertures needed for automotive applications.
- Implemented a brand new approach using FPAs on a monolithic silicon photonic platform
- Wrote a new software stack for generator-based electronic-photonic design, focusing on productivity and automation necessary for taping out complex systems in 300mm processes.

Over the course of this work, it has become clear that silicon photonics is maturing very quickly. Designing complex silicon photonics systems is now possible in a 300mm fabrication environment, which can enable future scaling to high volume and low cost. However, through our experience, it has also become clear that silicon photonic chips on their own are not sufficient to enable breakthrough products in the solid-state LIDAR space. We've learned that whether you use OPAs or FPAs, tight integration between electronic control circuitry and photonics is critical to achieving the promised performance in the real world. Without this integration capability, whether 3D or monolithic, scaling to large arrays is nearly impossible without hitting a I/O bottleneck. Also, enabling calibration is clearly a requirement for ensuring beam quality across typical automotive environments.

Although not all of the implemented designs showed the full expected capability, the fundamental merits and intermediate measurements show significant promise for future research and development along the lines described in this thesis.

Bibliography

[1] Hooman Abediasl and Hossein Hashemi. "Monolithic optical phased-array transceiver in a standard SOI CMOS process". en. In: *Opt. Express, OE* 23.5 (Mar. 2015), pp. 6509–6519.

[2] Christos Adamopoulos et al. "Fully Integrated Electronic-Photonic Sensor for Label-Free Refractive Index Sensing in Advanced Zero-Change CMOS-SOI Process". In: *2021 IEEE Custom Integrated Circuits Conference (CICC)*. IEEE. 2021, pp. 1–2.

[3] Luca Alloatti et al. "Photonics design tool for advanced CMOS nodes". In: *IET Optoelectronics* 9.4 (2015), pp. 163–167. ISSN: 17518768. DOI: 10.1049/iet-opt.2015.0003.

[4] Ansys/Lumerical. *High-Performance Photonic Simulation Software - Lumerical*. URL: https://www.lumerical.com/.

[5] Applied Nanotools Inc. *NanoSOI Fabrication*. URL: https://www.appliednt.com/nanosoi-fabrication-service/.

[6] *Artificial Intelligence: The Next Digital Frontier?* Tech. rep. McKinsey.

[7] *Automotive short range and medium range radar integrated circuits and reference designs*. https://www.ti.com/solution/automotive-medium-short-range-radar. Accessed: 2021-6-18.

[8] *Automotive Software and Electronics 2030*. Tech. rep. McKinsey.

[9] P Bhargava et al. "Fully Integrated Coherent LiDAR in 3D-Integrated Silicon Photonics/65nm CMOS". In: *2019 Symposium on VLSI Circuits*. June 2019, pp. C262–C263.

[10] F Boeuf et al. "A multi-wavelength 3D-compatible silicon photonics platform on 300mm SOI wafers for 25Gb/s applications". In: *2013 IEEE International Electron Devices Meeting*. Dec. 2013, pp. 13.3.1–13.3.4.

[11] *BPG*. https://github.com/BerkeleyPhotonicsGenerator/BPG. URL: https://github.com/BerkeleyPhotonicsGenerator/BPG.

[12] Josep M. Fargas Cabanillas and Miloš A. Popović. "Fast adiabatic mode evolution based on geometry-induced suppression of nearest-mode crosstalk". In: *2018 Conference on Lasers and Electro-Optics (CLEO)*. 2018, pp. 1–2.

[13] Josep M. Fargas Cabanillas, Bohan Zhang, and Miloš A. Popović. "Demonstration of 30.12 dB Power Splitting over 145 nm Optical Bandwidth in a 31-m Long 3-dB Rapid Adiabatic Coupler". In: *2020 Optical Fiber Communications Conference and Exhibition (OFC)*. 2020, pp. 1–3.

[14] Cadence. *Integrated Design Flows for Photonics Circuits*. URL: https://www.cadence.com/en%7B%5C_%7DUS/home/solutions/photonics.html.

[15] Cadence Design Systems. *Circuit Design*. URL: %7Bhttps://www.cadence.com/en%5C_US/home/tools/custom-ic-analog-rf-design/circuit-design.html%7D.

[16] Sean Campbell et al. "Sensor Technology in Autonomous Vehicles : A review". In: *2018 29th Irish Signals and Systems Conference (ISSC)*. June 2018, pp. 1–4.

[17] Eric Chang et al. "BAG: A Process-Portable Framework for Generator-based AMS Circuit Design". In: *Proceedings of the Custom Integrated Circuits Conference* (2019). ISSN: 08865930. DOI: 10.1109/CICC.2019.8780349.

[18] Jiashu Chen. "Advanced Architectures for Efficient mm-Wave CMOS Wireless Transmitters". PhD thesis. UC Berkeley, 2013.

[19] L Chrostowski et al. "Impact of fabrication non-uniformity on chip-scale silicon photonic integrated circuits". In: *OFC 2014*. Mar. 2014, pp. 1–3.

[20] Sungwon Chung, Hooman Abediasl, and Hossein Hashemi. "A Monolithically Integrated Large-Scale Optical Phased Array in Silicon-on-Insulator CMOS". In: *IEEE J. Solid-State Circuits* 53.1 (Jan. 2018), pp. 275–296.

[21] Automotive Electronics Council. "Failure Mechanism Based Stress Test Qualification for Integrated Circuits, AEC-Q100". In: (2014).

[22] *Darpa Grand Challenge*. https://www.darpa.mil/news-events/2014-03-13. Accessed: 2021-6-18. Mar. 2014.

[23] Scott R Davis et al. "Liquid crystal clad waveguide laser scanner and waveguide amplifier for LADAR and sensing applications". In: *Integrated Optics: Devices, Materials, and Technologies XIX*. Vol. 9365. International Society for Optics and Photonics, Feb. 2015, 93650N.

[24] Christopher T DeRose et al. "Thermal crosstalk limits for silicon photonic DWDM interconnects". In: *2014 Optical Interconnects Conference*. May 2014, pp. 125–126.

[25] Nicholas M. Fahrenkopf et al. "The AIM Photonics MPW: A Highly Accessible Cutting Edge Technology for Rapid Prototyping of Photonic Integrated Circuits". In: *IEEE Journal of Selected Topics in Quantum Electronics* 25.5 (2019), pp. 1–6. DOI: 10.1109/JSTQE.2019.2935698.

[26] Reza Fatemi, Aroutin Khachaturian, and Ali Hajimiri. "A Nonuniform Sparse 2-D Large-FOV Optical Phased Array With a Low-Power PWM Drive". In: *IEEE J. Solid-State Circuits* 54.5 (May 2019), pp. 1200–1215.

[27] Martin Fiers et al. "Improving the design cycle for nanophotonic components". In: *Journal of Computational Science* 4.5 (2013), pp. 313–324. ISSN: 1877-7503. DOI: 10.1016/j.jocs.2013.05.008. URL: http://dx.doi.org/10.1016/j.jocs.2013.05.008.

[28] Lucas H. Gabrielli. *gdspy*. https://github.com/heitzmann/gdspy. URL: https://github.com/heitzmann/gdspy.

[29] R Gerchberg. "A practical algorithm for the determination of phase from image and diffraction plane pictures". In: *undefined* (1972).

[30] H. Gevorgyan et al. "High Shift Efficiency O-band Spoked-Ring Modulator Allowing Fully Electro-Optic Channel Tuning in a 45nm CMOS Platform". In: *2021 Conference on Lasers and Electro-Optics (CLEO)*. 2021.

[31] H. Gevorgyan, A. Khilo, D. Van Orden, D. Onural, B. Yin, M. T. Wade, V. M. Stojanović, and M. A. Popović. "4Cryo-compatible, silicon spoked-ring modulator in a 45nm CMOS platform for 4K-to-room-temperature optical links". In: *Optical Fiber Communication Conference (OFC) 2021*. Optical Society of America, 2021, M3B.3.

[32] Ulrich Hofmann et al. "Biaxial resonant 7mm-MEMS mirror for automotive LIDAR application". In: *2012 International Conference on Optical MEMS and Nanophotonics*. Aug. 2012, pp. 150–151.

[33] Duanni Huang et al. "High-power sub-kHz linewidth lasers fully integrated on silicon". en. In: *Optica, OPTICA* 6.6 (June 2019), pp. 745–752.

[34] David N Hutchison et al. "High-resolution aliasing-free optical beam steering". en. In: *Optica, OPTICA* 3.8 (Aug. 2016), pp. 887–890.

[35] Daisuke Inoue et al. "Demonstration of a new optical scanner using silicon photonics integrated circuit". en. In: *Opt. Express, OE* 27.3 (Feb. 2019), pp. 2499–2508.

[36] *Introducing the 5th-generation Waymo Driver: Informed by experience, designed for scale, engineered to tackle more environments*. https://blog.waymo.com/2020/03/introducing-5th-generation-waymo-driver.html. Accessed: 2021-6-18.

[37] Brandon J Isaac et al. "Indium Phosphide Photonic Integrated Circuit Transceiver for FMCW LiDAR". In: *IEEE J. Sel. Top. Quantum Electron.* 25.6 (Nov. 2019), pp. 1–7.

[38] J. F. Cabanillas, D. Kita, A. Khilo, F Sedgwick, J. Fini, M. Popovic, M. Wade. "Demonstration and Fabrication Tolerance Study of a Low-Loss,Ultra-broadband Rapid Adiabatic 3-dB Coupler in a Next-Generation 45 nm Monolithic Electronic-Photonic Platform". In: *Optical Fiber Communication Conference (OFC) 2021*. Optical Society of America, 2021, Tu5B.4.

[39] Taehwan Kim et al. "29.5 A Single-Chip Optical Phased Array in a 3D-Integrated Silicon Photonics/65nm CMOS Technology". In: *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*. Feb. 2019, pp. 464–466.

[40] Matthias Köfferlein. *KLayout*. https://www.klayout.de/. URL: https://www.klayout.de/.

[41] A. Leinse et al. "TriPleX waveguide platform: low-loss technology over a wide wavelength range". In: *Integrated Photonics: Materials, Devices, and Applications II*. Vol. 8767. International Society for Optics and Photonics. SPIE, 2013, pp. 86–98. DOI: 10.1117/12.2020574. URL: https://doi.org/10.1117/12.2020574.

[42] Nanxi Li et al. "High-power thulium lasers on a silicon photonics platform". en. In: *Opt. Lett.* 42.6 (Mar. 2017), pp. 1181–1184.

[43] Nanxi Li et al. "Monolithically integrated erbium-doped tunable laser on a CMOS-compatible silicon photonics platform". en. In: *Opt. Express, OE* 26.13 (June 2018), pp. 16200–16211.

[44] Josué J López et al. "Planar-lens Enabled Beam Steering for Chip-scale LIDAR". en. In: *Conference on Lasers and Electro-Optics*. San Jose, California United States: Optical Society of America, May 2018, SM3I.1.

[45] M. Wade et al. "TeraPHY: A Chiplet Technology for Low-Power, High-Bandwidth In-Package Optical I/O". In: *IEEE Micro* 40.2 (2020), pp. 63–71. DOI: 10.1109/MM.2020.2976067.

[46] R. J. Mailloux. "Phased Array Antenna Handbook". In: (2017).

[47] Colin McDonough et al. "Heterogeneous integration of a 300mm silicon photonics-CMOS wafer stack by direct oxide bonding and via-last 3D interconnection". en. In: *Int. Symp. Microelectron.* 2015.1 (Oct. 2015), pp. 000621–000626.

[48] Cristiano Niclass et al. "A 0.18-μ m CMOS SoC for a 100-m-Range 10-Frame/s 200 $\times$96-Pixel Time-of-Flight Depth Sensor". In: *IEEE J. Solid-State Circuits* 49.1 (Jan. 2014), pp. 315–330.

[49] Jelena Notaros et al. "CMOS-Compatible Optical Phased Arrays with Monolithically-Integrated Erbium Lasers". In: *2018 Conference on Lasers and Electro-Optics (CLEO)*. May 2018, pp. 1–2.

[50] P. Agnello et al. "High Performance 45-nm SOI Technology with Enhanced Strain, Porous Low-k BEOL, and Immersion Lithography". In: *2006 International Electron Devices Meeting*. 2006, pp. 1–4. DOI: 10.1109/IEDM.2006.346879.

[51] Christopher Vincent Poulton et al. "Long-Range LiDAR and Free-Space Data Communication With High-Performance Optical Phased Arrays". In: *IEEE J. Sel. Top. Quantum Electron.* 25.5 (Sept. 2019), pp. 1–8.

[52] Christopher V Poulton et al. "Coherent solid-state LIDAR with silicon photonic optical phased arrays". en. In: *Opt. Lett., OL* 42.20 (Oct. 2017), pp. 4091–4094.

[53] Christopher V Poulton et al. "Large-scale silicon nitride nanophotonic phased arrays at infrared and visible wavelengths". en. In: *Opt. Lett.* 42.1 (Jan. 2017), pp. 21–24.

[54] Michal Rakowski et al. "45nm CMOS - Silicon Photonics Monolithic Technology (45CLO) for next-generation, low power and high speed optical interconnects". In: *Optical Fiber Communication Conference (OFC) 2020*. Optical Society of America, 2020, T3H.3. DOI: 10.1364/OFC.2020.T3H.3. URL: http://www.osapublishing.org/abstract.cfm?URI=OFC-2020-T3H.3.

[55] Manan Raval, Christopher V Poulton, and Michael R Watts. "Unidirectional waveguide grating antennas with uniform emission for optical phased arrays". en. In: *Opt. Lett.* 42.13 (July 2017), pp. 2563–2566.

[56] Siemens EDA. *Calibre Design Solutions*. URL: https://eda.sw.siemens.com/en-US/ic/calibre-design/.

[57] Cheryl Sorace-Agaskar et al. "Electro-optical co-simulation for integrated CMOS photonic circuits with VerilogA." In: *Optics express* 23.21 (2015), pp. 27180–203. ISSN: 1094-4087. DOI: 10.1364/OE.23.027180. URL: http://www.osapublishing.org/viewmedia.cfm?uri=oe-23-21-27180%7B%5C&%7Dseq=0%7B%5C&%7Dhtml=true.

[58] Chen Sun et al. "A Monolithically-Integrated Chip-to-Chip Optical Link in Bulk CMOS". In: *IEEE J. Solid-State Circuits* 50.4 (Apr. 2015), pp. 828–844.

[59] Chen Sun et al. "TeraPHY : An O-band WDM Electro-optic Platform for Low Power , Terabit / s Optical I / O". In: 2020, pp. 1–2. ISBN: 9781728164601.

[60] Jie Sun et al. "Large-scale nanophotonic phased array". en. In: *Nature* 493.7431 (Jan. 2013), pp. 195–199.

[61] Synopsys. *Photonic Solutions*. URL: https://www.synopsys.com/photonic-solutions.html.

[62] Goro Takeuchi et al. "Thermally controlled Si photonic crystal slow light waveguide beam steering device". en. In: *Opt. Express, OE* 26.9 (Apr. 2018), pp. 11529–11537.

[63] Minh A Tran et al. "A Widely-Tunable High-SMSR Narrow-Linewidth Laser Heterogeneously Integrated on Silicon". In: *2018 Conference on Lasers and Electro-Optics (CLEO)*. May 2018, pp. 1–2.

[64] *Transportation Employment*. Tech. rep. Bureau of Labor Statistics.

[65] H K Tsang et al. "Optical dispersion, two-photon absorption and self-phase modulation in silicon waveguides at 1.5 μm wavelength". In: *Appl. Phys. Lett.* 80.3 (Jan. 2002), pp. 416–418.

[66] "Velodyne Alpha Puck Datasheet". In: (2019).

[67] Mial E Warren. "Automotive LIDAR Technology". In: *2019 Symposium on VLSI Circuits*. June 2019, pp. C254–C255.

[68] Ami Yaacobi et al. "Integrated phased array for wide-angle beam steering". en. In: *Opt. Lett.* 39.15 (Aug. 2014), pp. 4575–4578.

[69] Yisu Yang et al. "Phase coherence length in silicon photonic platform". en. In: *Opt. Express* 23.13 (June 2015), pp. 16890–16902.

[70] De Jong Yeong et al. "Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review". en. In: *Sensors* 21.6 (Mar. 2021).

[71] Panagiotis Zarkos et al. "Fully Integrated Electronic-Photonic Ultrasound Receiver Array for Endoscopic Imaging Applications in a Zero-Change 45nm CMOS-SOI Process". In: *2021 Symposium on VLSI Circuits*. IEEE. 2021, in press.