Realization of Integrated Coherent LiDAR

Taehwan Kim

Electrical Engineering and Computer Sciences
University of California at Berkeley

Realization of Integrated Coherent LiDAR

by

Taehwan Kim

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Vladimir Stojanović, Chair
Professor Ming C. Wu
Professor Costas Grigoropoulos

Summer 2019

Realization of Integrated Coherent LiDAR

Copyright 2019
by
Taehwan Kim

Abstract

Realization of Integrated Coherent LiDAR

by

Taehwan Kim

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Vladimir Stojanović, Chair

LiDAR (Light Detection and Ranging) captures high-definition real-time 3D images of the surrounding environment through active sensing with infrared lasers. It has unique advantages that can compensate the fundamental limitations in camera-based 3D imaging via vision algorithms or RADARs, which makes it an important sensing modality to guarantee robust autonomy in self-driving cars. However, high price tag of existing commercial LiDAR modules based on mechanical beam scanners and intensity-based detection scheme makes them unusable in the context of mass produced consumer products.

The focus of thesis is on the integrated coherent LiDAR with optical phased array-based solid-state beam steering, which has great potential to dramatically bring down the cost of a LiDAR module. It begins with an overview of LiDAR implementation options and system requirements in the context of autonomous vehicles, which leads us to conclude that beam-steering coherent FMCW LiDAR in optical C-band is indeed the best implementation strategy to realize low-cost automotive LiDARs. Motivated by this observation, a quantitative framework for evaluating FMCW LiDAR performance is also introduced to predict the design that satisfies car-grade performance requirements. Then the thesis presents the silicon implementation results from our single-chip optical phased array and integrated coherent LiDAR prototype. Our implementations leverage the 3D heterogeneous integration platform, where custom silicon photonics and nanoscale CMOS fabricated at a 300 mm wafer facility are combined at the wafer-scale to minimize the unit cost without I/O density issues. After discussing remaining challenges and possible ways to enhance the operating range and system reliability, this thesis finally addresses the problem of fundamental trade-off between phase noise and wavelength tuning in FMCW laser source, and present

circuit- and algorithm-level techniques to enable FMCW measurements beyond inherent laser coherence range limit.

Contents

List of Figures

List of Tables

Acknowledgements

I have always felt that the trajectory of my life so far has been just an accumulation of small, random decisions without long-term plans. Yet, somehow I'm here, unable to imagine a better place than where I am right now. I have come to realize that it is because I've been able to observe the *gradient* provided by the people around me, so that I can set the direction of my baby steps with confidence.

First I want like to thank my advisor, Professor Vladimir Stojanović. I have learned so much from his incredible ability to capture the essence of a problem, cutting right through all the noise around it. His uncompromising tenacity in pursuing what he believes in, with regards to both technical and non-technical issues, is simply amazing. I am grateful for the opportunity to witness his commitment and be inspired.

I would also like to thank my undergraduate advisor, Professor Jaeha Kim, for giving me an opportunity to explore the world of research, and for teaching me how to present my ideas. The first digital integrated circuits course from him that I took back in 2010 would always remain as one of the defining moments of my career.

I would also like to thank Professor Ming Wu for his technical guidance and encouragement, and for teaching me how lasers work. Thanks are also due to the rest of my thesis and qual exam committee, Professors Costas Grigoropoulos, and Ali Niknejad, for their invaluable feedback.

I am incredibly proud to be a part of the Integrated Systems Group family, a collection of the smartest people I've ever met. Thanks Chen for making everything work, Ranko for being the coolest, Sen for joining my support for Cory hall group meetings, Sajjad for having such a great taste in everything, Krishna for always being inspirational, Nandish for rich circuit knowledge and comedic reliefs, Pavan for his amazing ability to understand and complete my broken sentences, Sidney for his quest for introducing order in our chaotic world of ISG, Panos and Christos for making me dream about traveling to Greece someday.

Large part of the research described this thesis was done in collaboration with the Photonic Microsystems Group at MIT and the Colleges of Nanoscale Science and Engineering at SUNY

ACKNOWLEDGEMENTS ix

Albany, and I would like to thank Professor Michael Watts and CNSE team for their leadership and fabrication support. A huge thanks is reserved for Chris Poulton; I just feel extremely lucky that I could rely on his brilliance and learn so much from him. I would also like to thank Jelena Notaros for her help throughout the project.

I would also like to thank BWRC and EECS department staffs, especially Candy Corpus and Shirley Salanio, for making my life so much easier, for always be willing to answer whatever stupid questions I may have. I deeply appreciate their ceaseless effort to make sure that everyone's being taken care of, no one's in a blind spot.

I would like to express my gratitude for the staffs at Intel Labs for the hospitality and giving me the great opportunity to learn and explore the possibility of photonics technology from different perspectives.

I would also like to acknowledge the members of the Korean EECS Association (KEECS). I would like to thank the past members, Jaehwa, Yunsup, DJ, Kangwook, Jaeduk, Hokeun, Sangyoon, Seobin, Jaeyoung, for welcoming me into this community and for precious advice. I would also like to thank and wish the best of luck for the current members: Jinkyu, Sangmin, Kyoungtae, Hyungjin, Taesung, Jeongseok, Edward, Dayeol, Kunmo, Suhong, Jichan.

I would like to acknowledge Kwanjeong Educational Foundation for the financial support to my graduate study.

My special thanks go to Seonah for being around me. I am just excited about our next chapter and even more fun we're going to have.

Finally, I would like to express my infinite gratitude and love to my mom, my dad, and my sister. Words cannot express how comforting it is to always remind that, no matter where I am physically or emotionally, there's always a safe room somewhere in this world that I can go back and stay for however long I want. Thank you so much for always keeping that room since the day I was born.

Chapter 1

Introduction

While lamenting over the absence of once enjoyed exponential performance boost from technology scaling, the semiconductor industry is still performing exceptionally well: the record high number of units were shipped in 2018, and the long-term outlook is also strong [1]. Some of the reasons why many are optimistic about the future of semiconductor industry are as follows:

- The cost of having a custom chip in the system keeps dropping due to continued effort to enhance the productivity and openness of silicon engineering [2], [3]. At the same time, deep learning revolution [4] has motivated hardware engineers to gain a better understanding of efficient siliconization of artificial neural networks [5]. As a result, the state-of-the-art machine learning algorithms are quickly penetrating into the consumer electronics in everyday life, enabling a whole new range of user experience.
- To make it possible for the electronic devices to make independent decisions so that they can provide useful functions with minimum human involvement, they must be equipped with various sensors and actuators so that they can perceive and act on the environment. This is fueling the further growth of low-cost sensor/actuator front-ends that are available on the CMOS platform, which is rapidly expanding from traditional capacitive / inductive / diode-based interfaces to MEMS [6] and silicon photonics [7] leveraging heterogeneous integration and advanced packaging techniques.
- As those *intelligent* electronics become prevalent, there must be a matching growth in hardware infrastructure to accommodate the computing and communication needs. Furthermore, the lack of scaling-driven performance gain is forcing hardware engineers to search for workload-dependent optimization opportunities with custom silicon [8].

An epitome of the trend outlined above is found in the automotive industry. There has been

NVIDIA DRIVE: Hyperion Kit   Tesla Motors: Full Self-Driving Chip

Audi: zFAS System
(Mobileye/NVIDIA/Altera/Infinion)
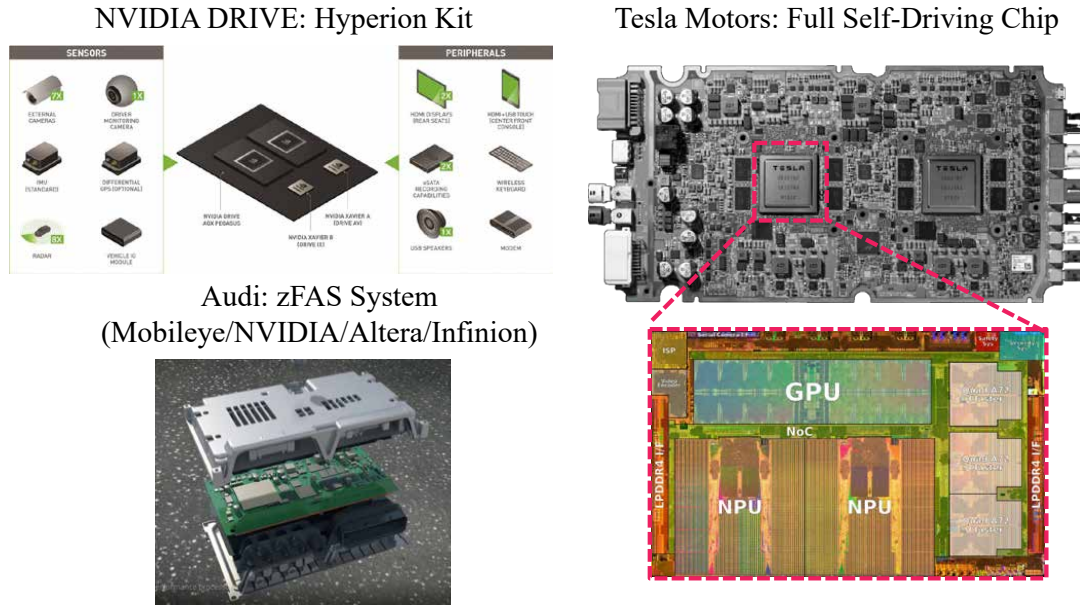
Figure 1.1: Examples of hardware systems for autonomous driving.

a strong demand for advanced driver-assistance system (ADAS), and ultimately autonomous vehicles, in order to make driving experience less stressful and safer by reducing the impact of human error. Due to the dynamic and uncertain nature of the car environment, massive amount of local computing power and advanced algorithms to enable quick adaptation and timely decision-making is deemed essential. At the same time, reliable sensors to feed the driving system are crucial to make a fully informed decision. It was not until recently it made economic sense to build such complex hardware systems for mass production cars, and now a number of semiconductor companies, car part companies, and the car manufacturers themselves are investing heavily in the research and development of autonomous driving system (Figure 1.1) in anticipation of the rise of self-driving cars in the foreseeable future.

The major part of the perception layer of an autonomous driving system is dedicated to simultaneous localization and mapping (SLAM) [9], where the map of unknown environment is constantly updated with respect to the location of the car. SLAM algorithms operate based on inputs from a few different types sensors, and the majority of them are imaging sensors. This is rather obvious considering the fact that humans also rely heavily on the visual information while driving. Imaging sensors commonly considered for self-driving cars are camera, RADAR, and LiDAR [10].

Camera supports excellent diffraction-limited lateral resolution and is also capable of resolv-
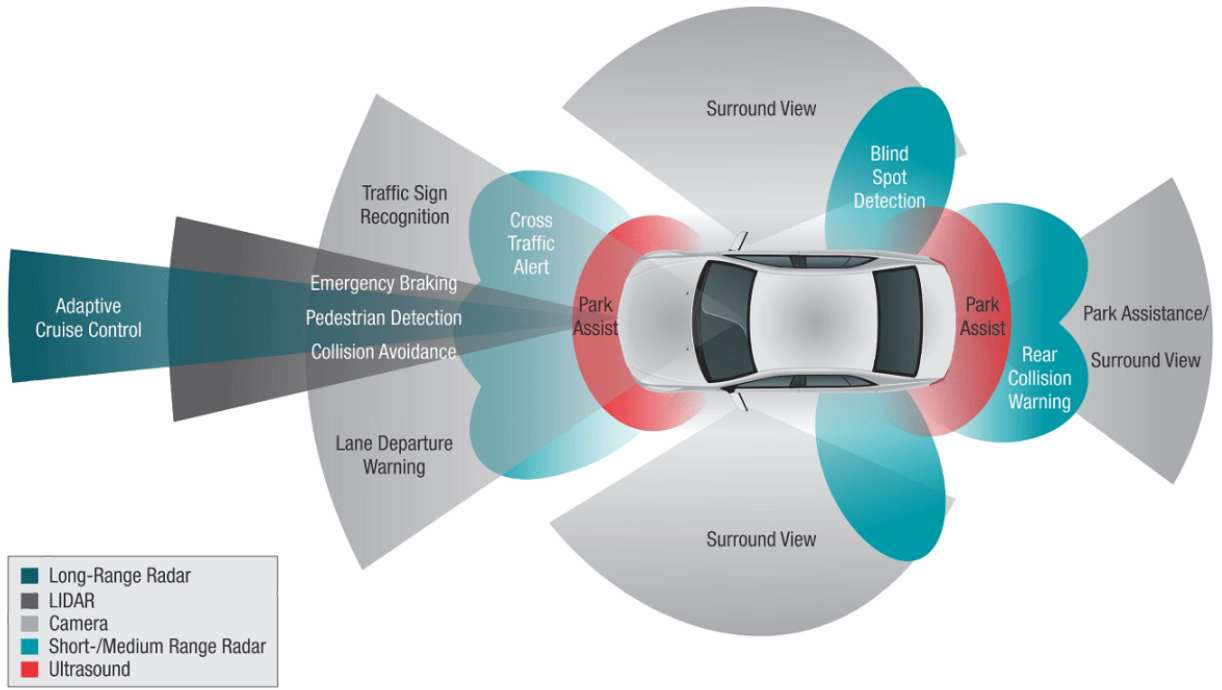
Figure 1.2: An overview of the sensing system for autonomous vehicles/ADAS (Texas Instruments).

ing color or texture of the object. Another big advantage is the cost: a number of cheap options are available in the market, and most modern cars are already being shipped with several cameras which serve as event data recorder (EDR) or a part of parking assistance system. Availability of local computing in cars is making cameras even more powerful, as exemplified by high-quality camera-based 3D imaging reported in [11]. On the other hand, cameras are fundamentally passive sensors relying on ambient light, and thus the imaging is significantly limited at night or under bad weather conditions. In addition, while computational imaging is certainly enhancing the amount of information retrieved from normal 2D cameras, it is hard to define general performance lower bound, or guarantee artifact-free imaging with certain confidence. This is especially a big problem in safety-critical systems such as cars.

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/mmWave frequency band. The lateral resolution of RADAR is much worse than cameras because the carrier wavelength is significantly longer compared to optical bands. However, it is less sensitive to weather conditions and generally capable of long range measurements 100 m ∼ 1 km. The unit

Table 1.1: Comparison of Depth Imaging Sensors

Features	Stereo Camera & Image Processing	Radar	LiDAR
Operation Principle	Passive	Active	Active
Lateral Resolution	🟢	🔴	🟢
Object Color/Texture	🟢	🔴	🔴
3D Imaging Range	🟡	🟢	🟡
Unit Price	🟢	🟢	🔴
Ambient Light Dependency	🔴	🟢	🟢
Weather Dependency	🔴	🟢	🟡
Computation/Data	🔴	🟢	🟢
Error Bound	🔴	🟢	🟢

cost is also low thanks to various CMOS implementations [12] now available in the market.

LiDAR (Light Detection and Ranging), is similar to RADAR, but it uses laser light instead of microwave and thus its diffraction-limited lateral resolution is on par with cameras. Therefore, it can provide the high-resolution ground truth/reference data to the camera-based 3D imaging. Namely, car imaging system design is not about the competition between different sensors, but about the effective way to utilize those sensors collectively in order to ultimately achieve consistent, robust driving performance regardless of ambient light or weather condition. Since each of those sensors has different advantages and limitations (Table 1.1), all three imaging modes are likely to be needed to realize such system, as long as the cost of entire automotive sensing system is kept reasonable for mass production cars.

In this thesis, we focus on LiDAR, an imaging sensor that can play a unique role in any applications that require high-resolution, real-time 3D images, including self-driving cars. LiDAR is an interesting topic not just because it has a huge potential to form a high-volume market, but also because it can be benefited the most from integration. The cost is currently the main

obstacle that is blocking the use of LiDARs in consumer products: for instance, typical cost of existing automotive-grade LiDAR is higher than the car itself [13]. As mentioned above, silicon photonics is gradually becoming a part of extended device portfolio for CMOS technology, and we need a comprehensive study on the feasibility of silicon as the baseline platform for building high-performance LiDAR systems.

1.1 Thesis Organization

Chapter 2 begins with an overview of the operation principle of different LiDAR implementations. I also present the list of major LiDAR metrics and parameters for evaluating the LiDAR performance. Based on this, I illustrate the ideal LiDAR for automotive applications and conclude that beam steering illumination scheme combined with coherent detection-based LiDAR working at optical C-band is desired.

Motivated by the conclusion in Chapter 2, I present comprehensive system-level study of FMCW coherent LiDAR in Chapter 3. Key design parameters for the individual building blocks are defined, and their impact on the overall performance is identified. The chapter is concluded by a system-wise link budget analysis to estimate the performance of coherent LiDAR system.

Chapter 4 and Chapter 5 studies the siliconization of coherent LiDAR concept described in the previous section by illustrating our research effort to demonstrate solid-state FMCW LiDAR system on a 3D heterogeneous electronics-photonics integration platform, including optical phased array-based solid state beam steering module and on-chip coherent receiver. In addition to the details of our design and experimental results from silicon demonstration, practical issues and remaining challenges are extensively discussed. Those chapters include 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 Pavan Bhargava at UC Berkeley contributed to the design of analog circuits. Wafer fabrication and 3D integration was done by the CNSE.

Lastly, Chapter 6 formulates the problem of FMCW source phase noise. It is emphasized that while the maximum range of the coherent LiDAR is significantly limited by phase noise, it is hard to resolve this problem through laser design while satisfying the rest of laser requirements. Alternatively, I present circuit- and algorithm-level solutions to enable coherent LiDAR measurement beyond the native coherent range of the laser. This chapters includes the contents from following publications:

- T. Kim, P. Bhargava, and V. Stojanović, "Optimal Spectral Estimation and System Trade-Off in Long-Distance Frequency-Modulated Continuous-Wave Lidar," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018, pp. 1583–1587.
- T. Kim, P. Bhargava, and V. Stojanović, "Overcoming the Coherence Distance Barrier in Long-Range FMCW LIDAR," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2018, STh3L.7.

Chapter 2

LiDAR: Light Detection and Ranging

This chapter provides an overview of existing LiDARs and the way we evaluate them for different applications. First, we classify LiDAR implementations using two key criteria: modulation / demodulation scheme (i.e. the way the depth is actually inferred from the physical signal) and object illumination method. We also summarize common performance metrics, which act as the bridge between various applications and implementation strategies. Finally, as a case study, the chapter reviews critical requirements for LiDARs in autonomous vehicles. The reasons why coherent LiDARs, the topic of later chapters, are particularly appealing for long-range depth imaging for cars, are also discussed.
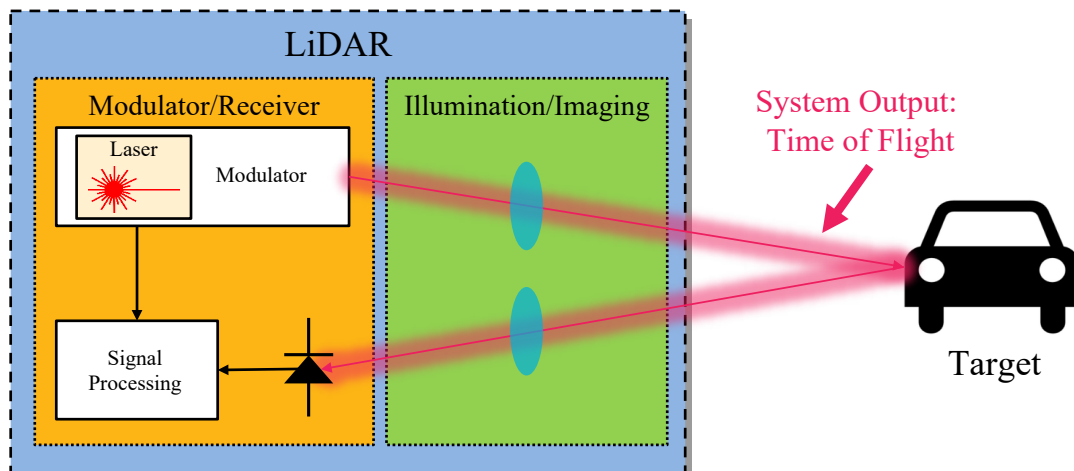
Figure 2.1: Overview of a LiDAR system.

2.1   Types of LiDAR Implementation

Figure 2.1 shows a simplified overview of a LiDAR system. A sensing operation starts at the laser source, which provides amplitude and/or phase modulated light. The modulated light is transmitted to the sensing target through the illumination optics, hits the target, and reflected light from the target is subsequently collected by the imaging optics. Finally, the receiver records the amplitude/phase of the light from the imaging optics, correlates it with the modulation signal, and extracts the time it took for the light to come back from the target, or laser time-of-flight (TOF), to eventually measure the distance to the target. From Figure 2.1, one can also notice that LiDAR system design begins with finding answers to two key questions: how to modulate the laser so as to enable depth estimation at the receiver (modulation and detection scheme), and how to illuminate the target object (illumination method).

2.1.1   Modulation and Detection Schemes

Intensity Modulation, Direct Detection

In intensity-modulation, direct-detection (IM-DD) LiDARs, the intensity of the source laser is modulated, and the TOF is estimated by comparing the intensity patterns of transmitted (TX) and received (RX) light in the time domain. Pulsed LiDAR, or *direct* TOF sensing (Figure 2.2(a)), is the most well-known and intuitive variant where short pulses of light are transmitted and then the time-domain gap between TX and RX pulses are directly measured using an electronic timer. Other types of modulation signals other than a pulse train can also be used as long as the TOF is related to the correlation between TX/RX intensity. For example, sinusoidal intensity modulation or *indirect* TOF sensing shown in Figure 2.2(b) (also called continuous wave intensity modulation) is commonly used in the TOF sensors for consumer electronics [19] since it is simpler
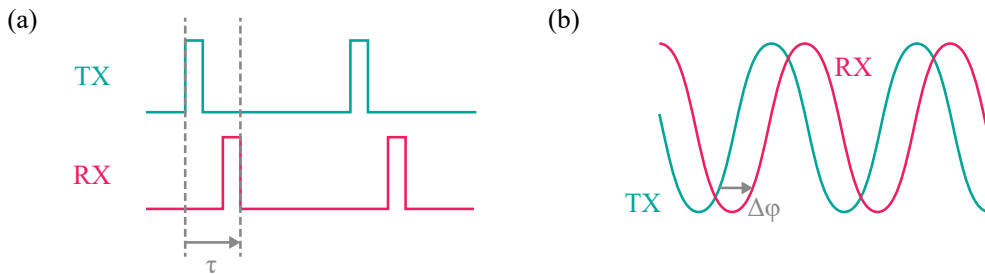
Figure 2.2: (a) Direct and (b) indirect TOF sensing using an intensity-modulated laser source.
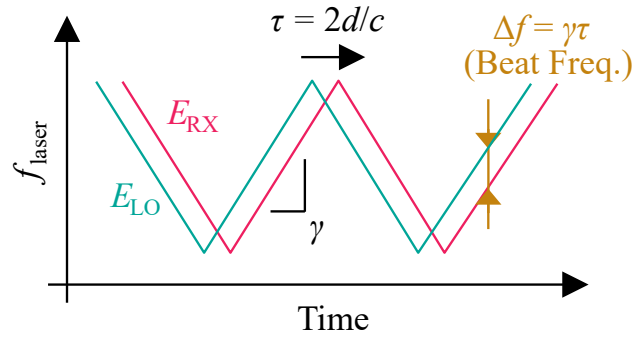
Figure 2.3: Coherent FMCW LiDAR principle.

to implement and does not require pulsed lasers. The distance in an indirect TOF sensor is inferred from the phase difference between TX and RX intensity sinusoids.

Coherent Detection

In coherent LiDARs, modulation is done in the laser phase/frequency domain while the intensity is often kept constant. The reflected light is optically mixed with the local oscillator (LO) laser, which is typically realized by simply tapping-off certain portion of the transmitted laser. Finally, the TOF is inferred from the downconverted electrical signal at the optical mixer output. Figure 2.3 shows the principle of frequency modulated continuous wave (FMCW) LiDAR, a popular coherent LiDAR variant, where the frequency of tunable continuous-wave laser is linearly modulated (a triangular wave is used in this example). One can notice that the TOF causes instantaneous frequency difference between two lasers (Δf). This difference, or *beat* frequency, is linearly proportional to the TOF. Namely, by recording the beat frequency, the distance to the target is measured.

2.1.2 Object Illumination Methods

Flash

As the name suggests, the entire field-of-view (FOV) is "flashed" at once, and the reflection is spatially resolved using an imaging optics followed by a detector array (Figure 2.4(a)). The working principle and the optics design of the Flash LiDAR is essentially the same as standard photographic cameras equipped with a flash light. Naturally, it is a preferred choice for existing cam-
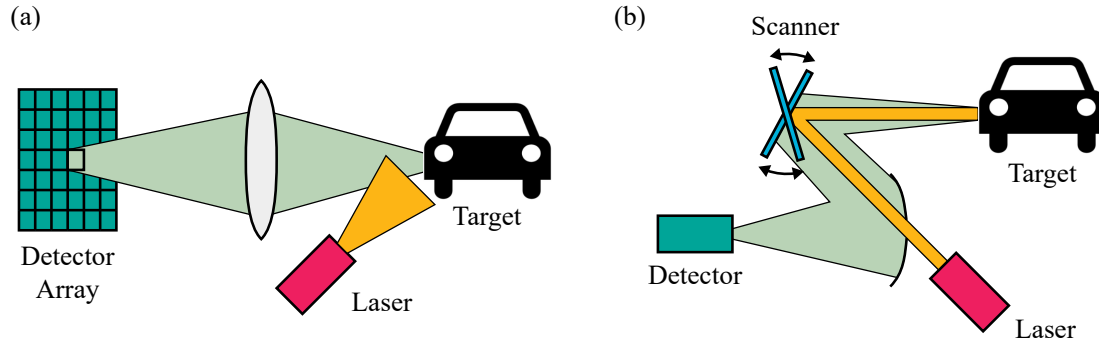
Figure 2.4: (a) Flash illumination-based LiDAR and (b) beam scanning-based LiDAR.

era sensor/imaging system manufacturers, and relatively low-cost options for short- to mid-range ($<$10 m) applications are already widely available.

Beam Scanning

A laser *beam* is formed to concentrate the optical power within single pixel instead of the whole scene, which makes it a point-wise measurement system. To form an image, the beam is scanned through the FOV (Figure 2.4(b)). Namely, a beam scanner-based LiDAR resembles a traditional rangefinder, rather than a camera. Scanning LiDAR achieves higher signal-to-noise ratio (SNR) at the cost of lower points per second (i.e. point throughput) and slower frame rate, and more importantly, at the cost of having a beam scanner. Beam scanner is often realized through mechanical actuation of either the source itself or the discrete optics around the source. While mechanical optical beam scanner design is already an established domain of engineering, there is a fundamental difficulty associated with achieving sufficient control precision and reliability goals for automotive vehicles using a low-cost mechanical system. To reduce the unit cost of a scanner module and make it feasible for consumer electronics, various solid-state beam scanning solutions are also actively being developed. This is discussed in depth in Chapter 4.

Flash/Scanning Hybrid

Using proper optics, scanning and flashing illumination can be combined as a compromise between the SNR and the frame rate. For example, we can scan the beam in one direction while flashing multiple pixels on the other direction at once, and resolve sub-beam pixels using an imaging lens and one-dimensional pixel array [20].

2.2 LiDAR System Metrics

This section presents a summary of key metrics in LiDAR systems. The application of interest ultimately dictates which metric the system architect should prioritize. At the same time, different LiDAR implementations have their own strengths and weaknesses, which also makes it easier for a certain LiDAR to meet some requirements while being limited in the others. Namely, a clear understanding of the linkage between those metrics and the application domain as well as the implementation strategy is crucial to set the right direction at the initial stages of system development.

2.2.1 Ranging Resolution and Precision

The primary performance metric of the depth sensing are the resolution (minimum resolvable distance between multiple targets on single line-of-sight) and the precision (standard deviation of measurements) of ranging. In fact, regardless of the actual modulation format, it is known that the information-theoretic lower bound of the absolute resolution (ΔR) and precision (σ_R) are given as follows [21], [22]:

$$\Delta R \propto \frac{c}{\text{BW}}, \quad \sigma_R \propto \frac{c}{\text{BW}} \frac{1}{\sqrt{\text{SNR}}}. \tag{2.1}$$

c is the speed of light, BW is the spectral bandwidth of the modulation, and SNR is the signal to noise ratio at the LiDAR receiver. For instance, ranging resolution and precision of a pulsed LiDAR are determined by the pulse width: shorter, sharper pulse means wider signal bandwidth, which results in finer ranging resolution.

2.2.2 Operating Range

Assuming diffusive target, the signal power at the receiver is inversely proportional to the square of the distance between the sensor and the target ([23], Section 3.4). Therefore, the maximum operating range of the LiDAR is primarily determined by the sensitivity of the LiDAR receiver, or the lowest permissible SNR to still satisfy the ranging precision requirement (Equation 2.1). Note that in the case of indirect TOF LiDARs, maximum range is often limited by the modulation frequency and associated phase ambiguity [24] rather than SNR.

2.2.3 Lateral Resolution and Field of View

Just like any other imaging system, fundamental lateral resolution limit in LiDAR comes from diffraction. Diffraction-limited (angular) lateral resolution is roughly on the scale of the wave-

length divided by the size of the the entrance pupil of the imaging optics in a flash illumination LiDAR. In a beam scanning LiDARs, the beam divergence (assuming Gaussian beams) is determined by the wavelength divided by the size of beam waist or its equivalent. The fact that the lateral resolution is directly proportional to the wavelength in both cases highlights the major advantage of the LiDAR: it can achieve much finer resolution compared to RADARs for given aperture size (or the same resolution with much smaller aperture) because the wavelength is 1000× smaller. Note that the resolution in a Flash LiDAR may also be limited by the detector array pitch, depending on the effective focal length of the imaging optics. Similar to standard cameras, field of view in a flash illumination LiDAR is a function of the focal length of the imaging optics and the size of the detector array (or any other field stop in the system). On the other hand, the scanning range of the beam steering module of choice determines the FOV in a beam scanning LiDAR.

2.2.4   Frame Rate

Frame rate primarily determines how fast the object can move without inducing significant motion blur. Since higher frame rate and shorter measurement window results in lower SNR, there is a direct trade-off between the frame rate and the maximum detection range in SNR-limited systems.

2.2.5   Background and Interference Suppression

LiDAR is an active imaging system, and any light other than the source laser is a potential source of disturbance. For example, dynamic background noise due to ambient light from lighting or sunlight can degrade the ranging performance by adding low-frequency drift, increasing photodetector shot noise, or saturating the receiver analog front-end. Moreover, laser light from other LiDARs can also be coupled into the system and degrade receiver performance. This is particularly detrimental for intensity modulated flash LiDARs because the nature of illumination/imaging mechanism results in higher chance of interference compared to beam-scanning coherent LiDARs, due to both increased spatial selectivity of the beam and coherence of the received signal. Different than background noise, the interfering signal and the actual signal from the source laser are largely distinguishable if the same modulation scheme is used, which makes it challenging to reject the interference using circuit techniques or detection algorithms.

2.2.6 Sensing Modes

Although the main objective of a LiDAR sensor is to measure the target distance, it is often possible to retrieve other useful information about the remote object from the LiDAR signal and perform multi-dimensional imaging. One example is signal intensity, which can be used to record surface reflectivity or roughness [25]. In the case of coherent LiDARs, it is also possible to record Doppler shift of the reflected laser induced by radial movements of the target [26], enabling velocity sensing.

2.2.7 Laser Wavelength

While the wavelength itself is not a performance metric per se, the choice of operating wavelength has a significant impact on various aspects of the system including lateral resolution, maximum emission power, atmospheric absorption, and the amount of background noise. It also affects the cost to realize a certain type of LiDAR system because the availability of a device-level building block is wavelength-dependent. Most commercial LiDARs use near-infrared (NIR) lasers, and the wavelength around 900 nm is a popular choice because of low sea-level solar radiation intensity at that wavelength band due to water vapor absorption [27] (i.e. low background noise due to sunlight) and the availability of silicon avalanche photodiodes to build low-cost detector array with high sensitivity. Optical C-band around 1550 nm wavelength, used mostly by communication systems due to low fiber loss, is also gaining popularity since it is not absorbed by human eye and therefore still safe at higher power levels. Moreover, cheap solid-state implementation of a LiDAR module is potentially possible as the silicon photonics technology is compatible with C-band laser.

2.2.8 Eye Safety and Maximum Emission Power

The maximum emission power from a LiDAR is primarily restricted by the IEC laser safety standard [28], and most consumer products are designed for class 1 eye safety. Since eye safety requirements are related not only to the absolute beam power density but also to other factors including wavelength, exposure time, and pulse duration in the case of pulsed lasers, the actual maximum number of photons usable for one pixel measurement can significantly vary depending on the actual system and failure mechanism design. Note that the maximum emission power can also be limited by the wall-plug efficiency of the laser source including the insertion loss of the beam scanner. For example, if the system budget is 10 W and the wall-plug efficiency is 1%, maximum optical power is limited to 100 mW.

2.2.9 Reliability

System reliability requirements are also heavily dependent on the application. For example, any electronic device going into a car is subject to AEC stress test qualification [29], which examines the system reliability against temperature variation (during operation/storage), electromigration, mechanical shock/vibration, etc.

2.2.10 Size, Weight, and Power-Cost (SWaP-C)

Finally, it is crucial to consider if the typical size, weight, and power consumption of the system are reasonable in the context of particular application. Moreover, the operation/repair cost over the device lifetime must be taken into account for total cost calculation. If the LiDAR requires periodic involvement of a specialist to maintain its functionality, the actual cost of owning a LiDAR in a automotive/robotic system may become substantially higher than just the manufacturing cost.

2.3 Case Study: Automotive LiDAR

As illustrated in the previous chapter, the major driving force behind LiDAR research and development is currently coming from the automotive industry, where the push for self-driving cars and extensive adoption of advanced driver-assistance systems (ADAS) are calling for high-performing 3D imaging solutions. In this section, we examine the main system requirements for automotive LiDARs with the metrics introduced in Section 2.2. We also discuss the implementation strategy suitable for meeting such requirements.

Although there is still no consensus on the exact role of LiDARs in self-driving cars, it is likely that the car sensor system is going to evolve into the direction where unique strength of each sensing mode is fully exploited. Again, the key benefit of LiDAR compared to other existing active depth sensors is its superior lateral resolution. Note that the main goal of the depth imaging system while driving is recognition and localization of the surrounding objects (i.e. bounding box formation [30]), rather than capturing exact details of the object itself. Since the absolute size of the commonly encountered objects (e.g. pedestrians, other cars, debris, poles, etc.) are fixed, we can reason that high lateral resolution becomes critical when objects are placed at a relatively far distance. In other words, LiDARs can play an important role in detecting objects on the front side of the car during mid- to high-speed driving, which is crucial for functions such as pedestrian detection, emergency breaking, and collision avoidance.

Table 2.1: Suggested Long-Range Automotive LiDAR Performance [13], [31]

Metric	Value
Ranging Precision	~1 cm
Operating Range	200 m~300 m
Lateral Resolution	0.1°~0.2°
Field of View	~90°
Frame Rate	>25 Hz
Eye Safety	IEC60825-1 Class 1
Reliability	AEC-Q100
Size	$100\,\mathrm{cm}^3$~$200\,\mathrm{cm}^3$
Weight	<5 kg
Power	10 W~30 W
Cost	$100~$200

Table 2.1 presents a list of performance requirements for long-range automotive LiDAR shared among manufacturers [13], which is in line with the preceding discussion and includes most metrics covered in Section 2.2. Maximum operating range of 300 m covers the stopping sight distance for driving speeds up to 80 mph (930 ft [32]). At the same time, it order to recognize a pedestrian ($\sim$ 1 m size) at 300 m distance, < 0.2° angular resolution is needed. Reasonable frame rate is also important for limiting in-frame ranging ambiguity while driving (for example, 25 Hz frame rate with 80 mph driving speed induces radial motion blur over $\sim$ 1.5 m range). What is missing on the list is background/interference suppression. Automotive LiDAR, by definition, assumes outdoor operation. It is likely that the interference associated with the laser directly emitted from cars on the opposite side of the road is much more powerful than the actual signal from the backscattered light from the target. Moreover, different than short-distance TOF cameras where the

signal power is still high and the background is treated as low-frequency noise (and therefore often removed via circuit techniques [33]), long-distance LiDARs are strictly SNR limited, and the SNR is directly degraded by the background through additional shot noise. In sum, the main goal is to achieve the highest receiver sensitivity while being immune to the interference.

With those basic requirements in mind, we can now move on to the optimal implementation strategy. As described in Section 2.1, the two major design choices to make are target illumination method (flash vs. beam scanning) and the modulation/detection scheme (direct detection vs. coherent). The former is easier to decide, because it is clear that obtaining enough optical power at the receiver through flash illumination, when the target is at multiple hundreds of meters away from the sensor, is extremely hard, if not impossible.

Once we rule out the flash illumination, the benefit of direct-detection LiDAR becomes less obvious. Since the measurement is done point-by-point, receiver complexity is less of an issue in beam scanning LiDARs than traditional imaging systems with a detector array and parallel readout circuitry. On the other hand, coherent detection is appealing because it can guarantee optimum quantum-limit sensitivity and inherent rejection of incoherent interferers [34], both of which are desperately desired in long-range LiDARs. While the sensitivity of direct-detection receivers are also often enhanced using Geiger-mode single-photon avalanche diodes [35], it still does not provide means to distinguish the photons from source laser and those from interference. While recent works have suggested DSP-based techniques to suppress ambient light [20], rejecting signals from other LiDAR sensors is still an issue. Moreover, algorithm-level solutions will add extra latency to the overall control loop in self-driving cars, which is very costly considering that the total response time is typically limited from a few seconds and ultimately down to a fraction of a second [36]. Furthermore, point-wise velocity sensing in coherent LiDARs is also very attractive feature for automotive applications.

Another decision to be made is laser wavelength. As mentioned above, it is essentially a choice between 900 nm and 1550 nm. 900 nm has been the only feasible choice for detector array-based TOF sensors because of the availability of cheap silicon avalanche diodes. This also becomes non-critical in LiDARs based on coherent detection. In fact, 1550 nm has clear advantages for long-range free-space applications due to higher eye-safe emission power and lower absorption by water vapors. In addition, as hinted in Section 2.2.7, it is compatible with existing photonics integration platform and devices originally developed for fiber-optic communication systems including silicon photonics. Note from Table 2.1 that automotive LiDARs are subject to extremely tight size, weight, power, and cost constraint, and solid-state implementation is the only way to realize such system with high reliability.

In summary, to reach targets that are multiple hundreds of meters away from the sensor, it

is obvious that beam scanning is the right illumination method. We also conclude that coherent detection is a better way to achieve both the highest sensitivity and interference immunity. Lastly, solid-state implementation based on integrated photonics at 1550 nm wavelength band is the only way to ultimately realize long-range automotive LiDARs at the price point realistic for mass production cars. The remainder of this thesis is dedicated to the realization of this vision and identifying/addressing new challenges.

Chapter 3

Coherent LiDAR Performance Analysis

In this chapter, we introduce a simple framework for the quantitative analysis of beam scanning-based coherent LiDAR performance. To limit the scope of discussion, we focus on FMCW LiDARs with linear frequency modulation (similar to Figure 2.3). We begin this chapter by reviewing the operation principle of FMCW LiDAR and basic definitions. Then we move on to more detailed discussions on the FMCW laser modulator and the coherent receiver, and their impact on the system performance. After a brief discussion on free-space loss, we finally perform a link budget analysis to evaluate if the target performance listed in Table 2.1 is realistic.
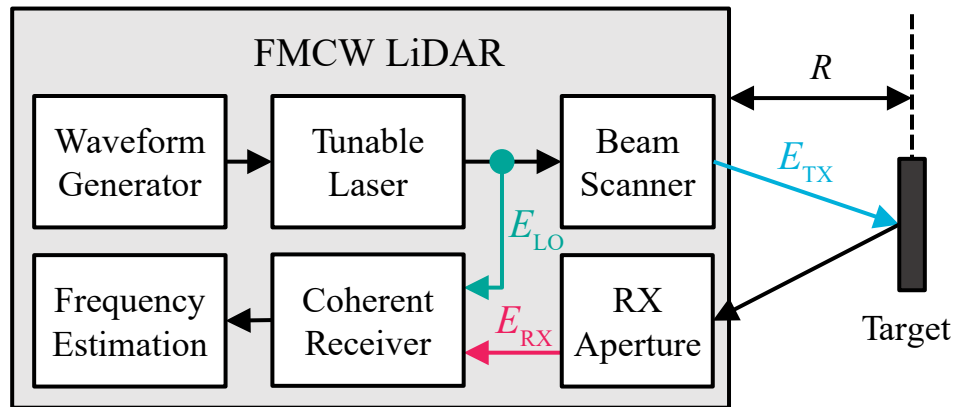
Figure 3.1: Overview of frequency-modulated continuous-wave LiDAR system.
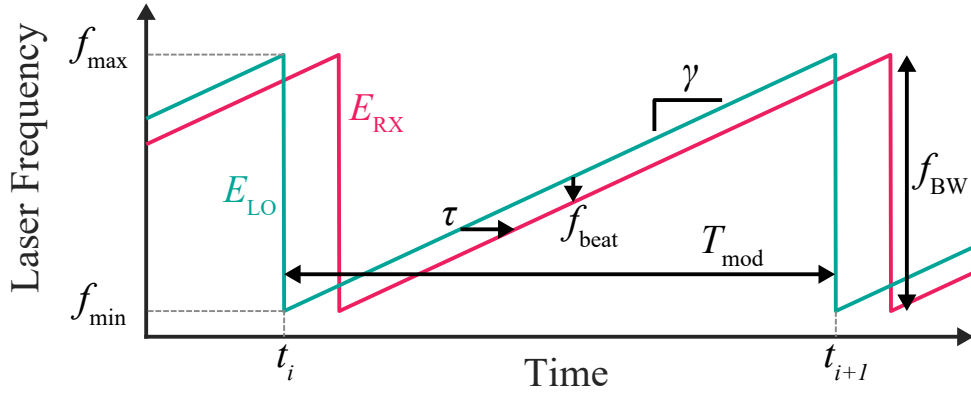
Figure 3.2: Time-domain laser frequency waveform in FMCW LiDAR.

3.1 FMCW LiDAR Fundamentals

Figure 3.1 shows an overview of the FMCW LiDAR system. A continuous-wave wavelength-tunable laser is used as the source, where the tuning input is driven by a waveform generator to perform frequency modulation. If the source laser frequency at time t is given as $f_{\mathrm{src}}(t)$, the phase of the electrical field of the laser source $\phi_{\mathrm{src}}(t)$ is expressed as follows.

$$\phi_{\mathrm{src}}(t) = \int_0^t 2\pi f_{\mathrm{src}}(u)du + \phi_{\mathrm{src}}(0) + \phi_{\mathrm{n}}(t) \tag{3.1}$$

$\phi_{\mathrm{n}}(t)$ represents laser phase noise. Figure 3.2 shows an example of modulation waveform often used in FMCW LiDAR (sawtooth) where $f(t)$ is given as follows:

$$f_{\mathrm{src}}(t) = \frac{f_{\mathrm{BW}}}{T_{\mathrm{mod}}}(t - t_i) + f_{\mathrm{min}}, \; t_i \leq t < t_{i+1}, \; i \in \mathbb{Z}. \tag{3.2}$$

The modulation period is T_{mod} (i.e. $t_i = iT_{\mathrm{mod}}$) and the modulation bandwidth is f_{BW}. As a result, the laser is linearly "chirped" by a fixed rate of $\gamma = f_{\mathrm{BW}}/T_{\mathrm{mod}}$ at each cycle.

Once the FMCW laser source is generated, it is split into two beams as seen in Figure 3.1. One of the beams, or local oscillator (LO) laser (E_{LO}), is directly forwarded to one of the coherent receiver inputs, while the other beam (E_{TX}) is sent to the beam scanning module and emitted to the free space, towards the target. The light is backscattered from the target, and subsequently collected by the sensor receiving aperture, and finally sent to the other input of the coherent receiver (E_{RX}). This free-space propagation and backscattering process comes with significant optical loss, and we can define L_{FS} as the free-space loss for the target distance R (i.e. $P_{\mathrm{RX}} =$
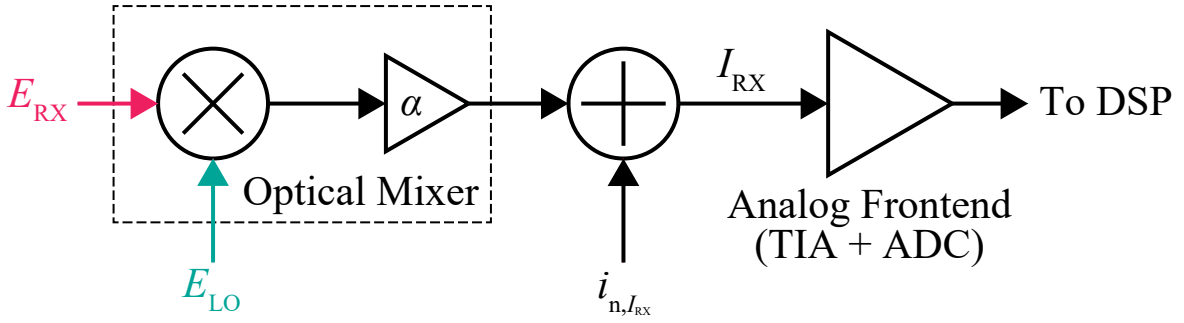
Figure 3.3: Overview of coherent detection frontend.

$L_{\mathrm{FS}}(R)P_{\mathrm{TX}})$. Note that the phase of LO/RX beams ($\phi_{\mathrm{LO}}, \phi_{\mathrm{RX}}$) simply become copies of ϕ_{src} but with different delay. For simplicity, we can neglect the delay within the sensor system, which makes $\phi_{\mathrm{LO}}(t) = \phi_{\mathrm{src}}(t)$ and $\phi_{\mathrm{RX}} = \phi_{\mathrm{src}}(t - \tau)$ where τ is the time-of-flight (TOF) $\tau = 2d/c$, and c is the speed of light.

The inside of a coherent receiver is shown in Figure 3.3. It is composed of an optical mixer and an electrical frontend. At the optical mixer, E_{RX} is downconverted by E_{LO} and outputs the baseband signal as an electrical current I_{RX}. Consequently, the amplitude of I_{RX} becomes proportional to both E_{RX} and E_{LO}, and the phase is the difference between ϕ_{LO} and ϕ_{RX} [34].

$$A_{I_{\mathrm{RX}}} = \alpha \sqrt{P_{\mathrm{RX}} P_{\mathrm{LO}}} \tag{3.3}$$

$$\phi_{I_{\mathrm{RX}}}(t; \tau) = \phi_{\mathrm{LO}}(t) - \phi_{\mathrm{RX}}(t) = \phi_{\mathrm{src}}(t) - \phi_{\mathrm{src}}(t - \tau) \tag{3.4}$$

α is a constant determined by the optical mixer design. We can plug Equation 3.1 into Equation 3.4 and get the following:

$$\phi_{I_{\mathrm{RX}}}(t; \tau) = \int_{t-\tau}^{t} 2\pi f_{\mathrm{src}}(u) du + \phi_{\mathrm{n}}(t) - \phi_{\mathrm{n}}(t - \tau). \tag{3.5}$$

Let's ignore the phase noise terms for now. Then we can notice that the instantaneous frequency of the baseband signal is expressed as follows:

$$f_{I_{\mathrm{RX}}}(t; \tau) = \frac{1}{2\pi} \frac{d\phi_{I_{\mathrm{RX}}}(t; \tau)}{dt} = f_{\mathrm{src}}(t) - f_{\mathrm{src}}(t - \tau). \tag{3.6}$$

Using Equation 3.2, $f_{I_{\mathrm{RX}}}(t)$ is simplified as follows:

$$f_{I_{\mathrm{RX}}}(t; \tau) = \gamma(t - t_i) - \gamma(t - \tau - t_i) = \gamma\tau. \tag{3.7}$$

Namely, the baseband signal becomes a sinusoidal tone (or *beat* tone) of frequency $f_{I_{RX}} = \gamma\tau$. Resulting time-domain expression and one-sided power spectral density (PSD) of the signal at the input of the electrical frontend is expressed as follows (T_{meas} is the length of the measurement, and here we assume that rectangular window is used):

$$I_{RX}(t; \tau) = A_{I_{RX}} \cos\left(2\pi f_{I_{RX}}(\tau)t + \phi_0\right) + i_{n,I_{RX}}(t),$$ (3.8)

$$S_{I_{RX}}(\omega; \tau) = \frac{A_{I_{RX}}^2}{2} T_{meas} \operatorname{sinc}^2\left(\frac{(\omega - 2\pi f_{I_{RX}}(\tau))T_{meas}}{2}\right) + S_{i_n,I_{RX}}(\omega).$$ (3.9)

$i_{n,I_{RX}}(t)$ is the additive noise term, which includes the noise from the optical mixer as well as the input-referred noise of the electrical frontend. Both of them are typically modeled as white noise (i.e. $S_{i_n,I_{RX}}(\omega) = N_{I_{RX}}$).

Finally, $f_{I_{RX}}$ is estimated from I_{RX} using spectral estimation method of choice, and finally yields the estimated distance to the target:

$$\hat{R} = \frac{c\hat{\tau}}{2} = \frac{c\hat{f}_{I_{RX}}}{2\gamma}.$$ (3.10)

It can be seen from Equation 3.7 that the electrical bandwidth requirement is determined by the signal frequency corresponding to the maximum detection range (R_{max}), which is expressed as follows:

$$\mathrm{BW}_{elec} = \frac{2R_{max}}{c}\gamma = \frac{2R_{max}}{c}\frac{f_{BW}}{T_{mod}}.$$ (3.11)

Note that the range estimation problem in FMCW LiDAR is reduced to classic line spectral estimation problem, whose resolution is known as follows [37]:

$$\Delta f = \frac{0.89}{T_{meas}}.$$ (3.12)

Moreover, theoretical lower bound of the single-frequency estimation precision (Cramér-Rao lower bound, CRLB) in the presence of additive white noise is also known as the following [38]:

$$\mathrm{CRLB}_{\hat{f}_{I_{RX}}} = \frac{1}{2\pi T_{meas}}\sqrt{\frac{3N_{I_{RX}}}{(A_{I_{RX}}^2/2)T_{meas}}} = \frac{1}{2\pi T_{meas}}\sqrt{\frac{3}{\mathrm{SNR}}}.$$ (3.13)

Note that this lower bound is actually achievable with efficient algorithms based on maximum-likelihood estimation as shown in [38]. Equation 3.12 and Equation 3.13 shows that the frequency estimation resolution is solely set by the signal length T_{meas} and the precision is determined by T_{meas} and also by the ratio of the signal peak spectral density ($A_{I_{RX}}^2 T_{meas}/2$) to the noise floor density ($N_{I_{RX}}$), or signal to noise ratio ($\mathrm{SNR} = A_{I_{RX}}^2 T_{meas}/(2N_{I_{RX}})$).

From Equation 3.10, Equation 3.12, and Equation 3.13, ranging resolution and precision lower bound is given as the following:

$$\Delta R = \frac{0.89}{T_{\text{meas}}} \frac{cT_{\text{mod}}}{f_{\text{BW}}}, \quad \text{CRLB}_{\hat{R}} = \frac{1}{2\pi T_{\text{meas}}} \sqrt{\frac{3}{\text{SNR}}} \frac{cT_{\text{mod}}}{2f_{\text{BW}}}. \tag{3.14}$$

Looking at Equation 3.14, it tempting to think that it is possible to improve the ranging resolution and precision simply by reducing T_{mod}. However, it should be noted that when T_{mod} is smaller than T_{meas}, a piece of time-domain baseband signal of length T_{meas} will inevitably include discontinuities at every t_i. These time domain discontinuities cause spectral leakage and broadens the signal spectrum, degrading the spectral estimation performance. Namely, once modulation bandwidth f_{BW} and total measurement time T_{meas} is fixed, reducing T_{mod} below T_{meas} does not improve the ranging performance and actually lowers the signal spectrum density peak due to leakage. Moreover, smaller T_{mod} increases the chirp rate γ and therefore requires higher modulator/receiver bandwidth (Equation 3.11). Therefore, we can simply assume that T_{meas} is equal to T_{mod} in practical systems. Resulting ranging resolution and precision lower bound is expressed as follows:

$$\Delta R = \frac{0.89c}{f_{\text{BW}}}, \quad \text{CRLB}_{\hat{R}} = \frac{c}{4\pi f_{\text{BW}}} \sqrt{\frac{3}{\text{SNR}}}. \tag{3.15}$$

Again, the ranging resolution (ΔR) is the minimum spacing between multiple targets on the same line-of-sight direction to be recognized as separate objects, and the ranging precision (represented by its Cramér-Rao lower bound $\text{CRLB}_{\hat{R}}$) is the standard deviation of single ranging measurement. This agrees with the expression in Equation 2.1 and confirms that ranging performance is always fundamentally limited by the modulation bandwidth and SNR. It is worth noting that it is much easier to achieve high *absolute* modulation bandwidth in FMCW LiDARs using tunable lasers where the laser cavity is directly modulated, compared to direct-detection LiDARs relying on an explicit intensity modulators. At the same time, Equation 3.11 shows that receiver electrical bandwidth is not directly determined by f_{BW} itself (which is the case in direct-detection LiDARs) but indirectly through the factor of $(2R_{\text{max}}/c)/T_{\text{mod}}$, which is typically much smaller than 1. In sum, required electrical bandwidth in both modulator and receiver are significantly lower in FMCW LiDARs for given resolution and precision target. This is the main reason why similar sensing principle is often utilized in applications where ultrahigh-resolution depth sensing is needed, such as the optical coherent tomography (OCT).

Equation 3.15 also makes it seem like it is possible to maintain the same ranging performance with lower SNR by increasing the modulation bandwidth f_{BW}, which would be great for SNR-limited long-range LiDARs. In reality, the threshold effect in frequency estimation problem [38]
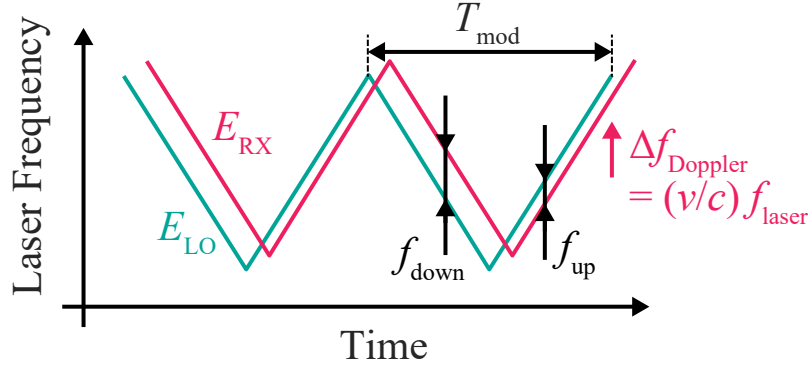
Figure 3.4: Principle of velocity sensing via Doppler shift in FMCW LiDAR.

eventually sets the SNR lower bound. Once the SNR goes below certain *threshold* (typically 20 ~ 30), Equation 3.13 becomes no longer valid, and the estimation performance essentially falls apart. From Equation 3.15, estimation precision on the verge of the SNR threshold (θ) is given as follows:

$$\text{CRLB}_{\hat{R}} = \frac{c}{4\pi f_{\text{BW}}}\sqrt{\frac{3}{\theta}}, \quad \text{SNR} = \frac{A_{I_{\text{RX}}}^2 T_{\text{meas}}}{2N_{I_{\text{RX}}}} = \theta. \tag{3.16}$$

Equation 3.16 is useful in our case since it shows achievable ranging performance with the worst-case SNR, which is in fact the relevant operation mode in long-range LiDARs.

Lastly, as mentioned in Section 2.2.6, coherent LiDAR enables velocity detection by recording Doppler shift. Figure 3.4 shows how velocity detection is actually done with triangular frequency modulation. Because of the Doppler effect, the frequency of E_{RX} is shifted by $\Delta f_{\text{Doppler}} = (v/c)f_{\text{laser}} = v/\lambda_{\text{laser}}$ when the radial velocity of the target is v. As a result, when the target is approaching the sensor, the beat frequency for downward chirping period (f_{down}) is increased by $\Delta f_{\text{Doppler}}$ and decreased by the same amount for upward chirping period (f_{up}). By taking the difference and the sum of those two frequencies, we can capture both distance and radial movement of the target:

$$f_{\text{down}} + f_{\text{up}} = 2\gamma\tau, \tag{3.17}$$

$$f_{\text{down}} - f_{\text{up}} = 2\Delta f_{\text{Doppler}} = 2v/\lambda_{\text{laser}}. \tag{3.18}$$

We can notice that the precision of the difference and the sum frequency estimation is the same. Therefore, the velocity estimation precision of an FMCW LiDAR is derived directly from Equation 3.13:

$$\text{CRLB}_{\hat{v}} = \frac{\lambda_{\text{laser}}}{2\pi T_{\text{meas}}}\sqrt{\frac{3}{\text{SNR}}}. \tag{3.19}$$

Equation 3.19 shows that, different than distance estimation, velocity measurement through Doppler shift is independent of chirping modulation. The electrical bandwidth requirement should also now take into account the maximum Doppler shift, which can go up to a few 10s of MHz in a high-speed car driving scenario:

$$\mathrm{BW}_{\mathrm{elec}} = \frac{2R_{\max}}{c}\frac{2f_{\mathrm{BW}}}{T_{\mathrm{mod}}} + \frac{|v_{\max}|}{\lambda_{\mathrm{laser}}}. \tag{3.20}$$

Note that the first term is multiplied by 2 from the original expression in Equation 3.11 to reflect the fact that one modulation cycle comprises up- and down-chirp, doubling the chirp rate γ for given f_{BW} and T_{mod}.

So far, we have established the basic relationship between physical parameters (e.g. laser modulation bandwidth/period, measurement time window length, laser power, maximum range) and signal amplitude, receiver bandwidth, and ranging resolution and precision. Looking back at Figure 3.1, we can identify three key elements in the system: FMCW laser source, optical beam-scanning transmitter, and coherent receiver. For the next two sections, we will take a closer look at the source and receiver (beam scanner is covered extensively in the next chapter and therefore assumed to be ideal for now). We will also briefly discuss the source of free-space loss (L_{FS}) and expected loss in realistic scenarios.

3.2 FMCW Laser Source

As described in the previous section, an FMCW laser source would include a tunable continuous wave laser and a modulation waveform generator circuit to drive the tuning input (a realistic tunable laser diodes also includes temperature and driving current regulation circuitry). Keeping in mind the fact that the ranging precision on the verge of SNR threshold is solely determined by the modulation bandwidth (Equation 3.16), in order to achieve the ranging precision target from Section 2.3, required minimum tuning bandwidth of the laser is calculated as follows, for $\theta = 30$:

$$f_{\mathrm{BW}} \geq \frac{1}{4\sqrt{10}\pi}\frac{c}{\sigma_R} = 0.0251\frac{3\times10^8\ \mathrm{m/s}}{1\ \mathrm{cm}} \approx 0.75\ \mathrm{GHz}. \tag{3.21}$$

We can easily implement a CW source with few GHz of tuning range with a semiconductor laser and bias current modulation utilizing cavity chirping [39], and various low-cost options in the eye-safe C-band (1550 nm) are already available.

Unfortunately, any realistic FMCW source contains both deterministic and stochastic modulation errors (Figure 3.5). Wavelength tuning of a laser is often limited by two major issues:
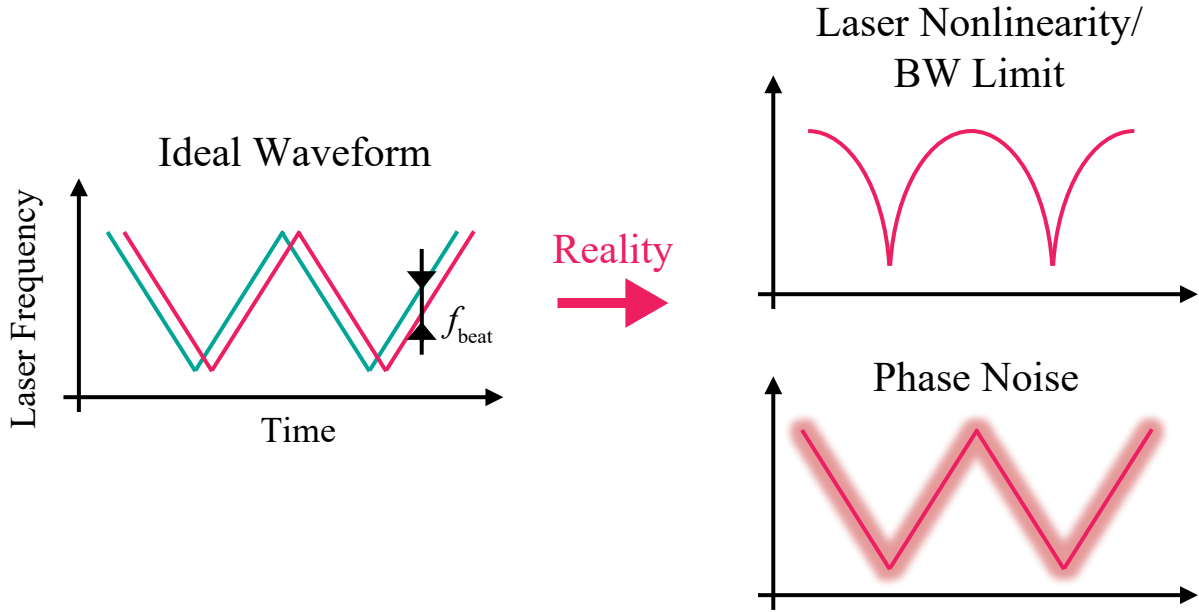
Figure 3.5: Deterministic and stochastic errors in FMCW source.

nonlinear relationship between the drive current and wavelength [40], and insufficient speed of the tuning response. Both of them can contribute to the chirp rate deviation $\Delta\gamma(t)$ from the nominal value. From Equation 3.7, it is clear that $\Delta\gamma$ will directly manifest itself as deterministic beat frequency deviation ($\Delta f_{I_{RX}}$), which is expressed as the following (assume $\gamma > 0$):

$$f_{I_{RX}}|_{\Delta\gamma\neq0} = (\gamma + \Delta\gamma)\tau = f_{I_{RX}}|_{\Delta\gamma=0}\left(1 + \frac{\Delta\gamma}{\gamma}\right) \tag{3.22}$$

$$\Delta f_{I_{RX}} = f_{I_{RX}}|_{\Delta\gamma=0}\frac{\Delta\gamma}{\gamma}. \tag{3.23}$$

Dynamic frequency deviation will inevitably broaden the spectrum of the baseband signal, reducing the signal spectral density peak from $A_{I_{RX}}^2 T_{\text{meas}}/2$, which eventually lowers the SNR. Interestingly, the amount of frequency deviation is proportional to both relative chirp rate error ($\Delta\gamma/\gamma$) and the error-free baseband frequency ($f_{I_{RX}}|_{\Delta\gamma=0} = \gamma\tau$). Therefore, chirp rate deviation is the most detrimental when the target is at the maximum range (R_{max}). To mitigate non-linearity, pre-distortion [40], [41] or feedback control leveraging optical phase-locked loop [22] can be used, while standard bandwidth extension techniques implemented in the driver circuitry can compensate tuning bandwidth limitation.

Let's also revisit the assumption that the laser phase noise term in Equation 3.5 is negligible. While the Q factor of the cavity itself is generally higher in lasers compared to the typical

resonators used in electrical oscillators, absolute linewidth (or 3dB spectral bandwidth) of the laser spectrum is quite broad as the carrier frequency itself is much higher. For example, C-band (1550 nm) semiconductor tunable lasers often used in FMCW LiDARs typically have a few 100 kHz of 3 dB linewidth. Although analyzing the impact of phase noise on FMCW LiDAR performance with exact formula in general case is rather involved [22], at least in the case where laser frequency noise is assumed to be white, a simple expression for the signal power degradation is known as follows (will be derived in Section 6.1):

$$P_{I_{RX}} = P_{I_{RX,0}} e^{-4\pi \Delta v R/c}, \quad R < R_{coh} = \frac{c}{2\pi \Delta v}. \tag{3.24}$$

$P_{I_{RX,0}} = A_{I_{RX}}^2 T_{meas}/2$ is the signal power when the phase noise is absent, and Δv is the 3 dB linewidth of the laser. R_{coh} is called coherence distance, which represents the maximum target range where we can still call E_{RX} and E_{LO} to be "coherent". Similar to the case of chirp rate error, LiDAR suffers more when the target distance is larger, which is obvious since the power of baseband signal phase noise ($\phi_{n,I_{RX}}(t) = \phi_n(t) - \phi_n(t - \tau)$) goes up for larger τ as $\phi_n(t)$ and $\phi_n(t - \tau)$ become less correlated. For example, when the linewidth is 200 kHz and the target distance is 300 m, signal power goes down by 10.9 dB from phase noise alone.

3.3 Coherent LiDAR Receiver

Figure 3.6 shows how the optical mixer in Figure 3.3 is actually implemented. It includes a 2×2 3 dB coupler, which is mathematically modeled as a unitary matrix, and two photodetectors, each modeled as an envelope detector with an intrinsic gain equal to its responsivity (R_{PD}), in a "balanced" configuration [34]. By expanding this model, we can express I_{out} in terms of two input fields:

$$I_{RX} = I_{out} = 2R_{PD}\sqrt{P_{LO}P_{RX}} \cos{(\phi_{LO}(t) - \phi_{RX}(t))} + i_{n,shot}(t). \tag{3.25}$$

Compared to Equation 3.3, we can notice that mixer coefficient α is equal to $2R_{PD}$ assuming the ideal mixer with no excess insertion loss. $i_{n,shot}$ represents the shot noise added by photodetectors, and its one-side PSD is expressed as follows:

$$S_{i_{n,PD}}(\omega) = N_{I_{RX}} = 2qR_{PD}P_{LO} + 2qR_{PD}P_{RX} + 2qR_{PD}P_{interf.} + 2qI_{dark}. \tag{3.26}$$

$P_{interf.}$ is the power of the interfering light from other sensors or sun, and I_{dark} is the dark current of the diode. Typically, P_{LO} is much bigger than P_{RX}, P_{interf}, and I_{dark}/R_{PD}, and the impact of last three terms in negligible.

In addition to the diode noise, input-referred noise of the electrical analog frontend (AFE) including transimpedance amplifier (TIA) and also the quantization noise at the analog-to-digital converter (ADC) will be added. Given this, we would like to calculate the sensitivity of the receiver to meet the performance target (i.e. minimum P_{RX}) and associated P_{LO} level as a function of measurement length $T_{meas} = T_{mod}$. Note that in Equation 3.25 and Equation 3.26, both signal power ($P_{I_{RX}} = A_{I_{RX}}^2/2 = 2R_{PD}^2 P_{RX} P_{LO}$) and the LO shot noise power are proportional to P_{LO}. Therefore, the power of the LO laser does not affect the SNR in shot noise-limited case. The signal to noise ratio in such case is called *quantum-limit* SNR, which is the fundamental SNR upper bound in any optical detection system [34]. Still, the absolute signal and noise power at the output of the optical mixer are proportional to the LO power, and we can essentially treat the LO power level as a pre-amplification gain before the electrical frontend. In other words, by increasing P_{LO}, we can ensure that LO shot noise is always the dominant noise source and the receiver is shot noise limited.

Assuming that the receiver is indeed operating in the shot noise limited regime, we can plug in the signal amplitude in Equation 3.25 and the noise density in Equation 3.26 to Equation 3.16
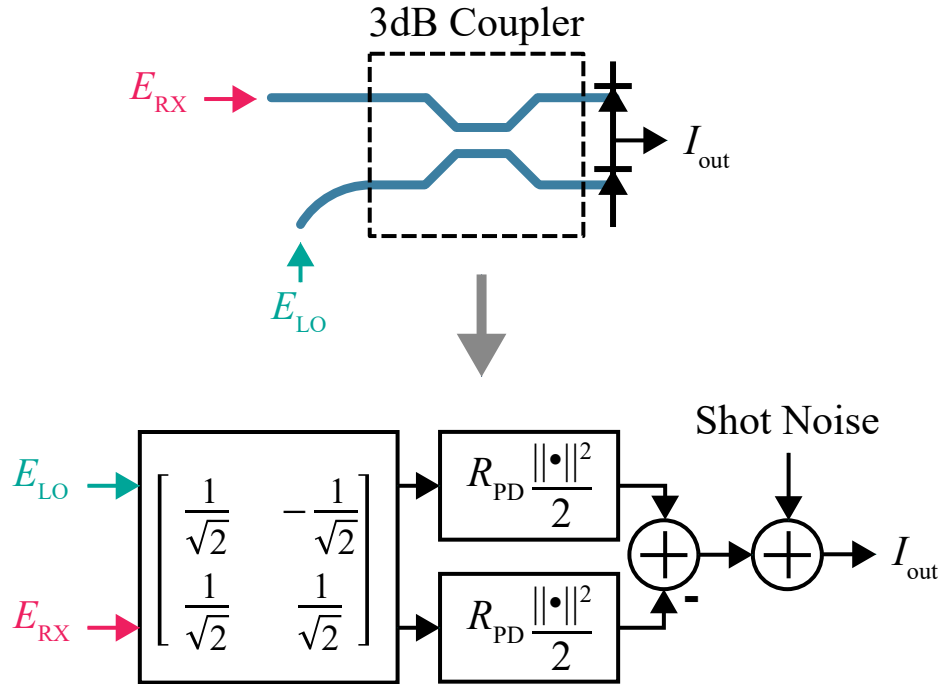
Figure 3.6: Optical mixer realization with photonic devices and equivalent model.

to calculate the minimum P_{RX} that avoids the threshold effect.

$$P_{\text{RX}} \geq \frac{q\theta}{R_{\text{PD}}} \frac{1}{T_{\text{meas}}} \tag{3.27}$$

For instance, P_{RX} should at least be larger than $0.48\,\text{pW}$ or $\approx -103\,\text{dBm}$ if the SNR threshold is 30, photodetector responsivity is $1\,\text{A/W}$ and T_{meas} is $100\,\mu s$.

Meanwhile, the receiver bandwidth requirement is calculated using Equation 3.20. Picking f_{BW} of $\sim 1\,\text{GHz}$ to satisfy Equation 3.21 with some margin, signal bandwidth is $60\,\text{MHz}$ for T_{meas} of $100\,\mu s$ ($R_{\text{max}} = 300\,\text{m}$, $v_{\text{max}} = 70\,\text{mph}$). Based on this, we can now predict the level of input-referred electrical noise including both quantization and thermal noise, and derive required P_{LO} to guarantee shot noise limited operation. As for the quantization noise, the input-referred one-side quantization noise density for an ADC of N-bit effective number of bits (ENOB) and bandwidth $f_{\text{BW,ADC}}$ is expressed as follows:

$$S_{i_{\text{n}},\text{in},Q}(\omega) = \frac{\Delta^2}{12} \frac{1}{f_{\text{BW,ADC}}} \left(\frac{1}{R_{\text{TIA}}}\right)^2 = \frac{\Delta_{\text{in}}^2}{12} \frac{1}{f_{\text{BW,ADC}}} \tag{3.28}$$

$$\Delta_{\text{in}} = \frac{\Delta}{R_{\text{TIA}}} = \frac{A_{\text{in,ADC}}}{(2^N - 1)R_{\text{TIA}}} \tag{3.29}$$

R_{TIA} is the TIA gain. Δ_{in} is the input-referred LSB of the ADC, which is dependent on N and the input range of the ADC ($A_{\text{in,ADC}}$) as well as R_{TIA}. If we pick a reference ADC with known performance numbers, input-referred quantization noise density is then purely a function of the TIA gain. For a typical pseudo-differential electrical frontend for optical receiver with resistor-feedback TIA and moderate bandwidth, transimpedance, bandwidth and input-referred thermal noise power spectral density are expressed as follows [42]:

$$\text{BW}_{\text{RX}} = \frac{1}{2\pi} \frac{g_{\text{m}} r_{\text{o}}}{R_{\text{FB}} C_{\text{PD}}} \tag{3.30}$$

$$R_{\text{TIA}} = \frac{g_{\text{m}} r_{\text{o}}}{1 + g_{\text{m}} r_{\text{o}}} R_{\text{FB}} \sim R_{\text{FB}} \tag{3.31}$$

$$S_{i_{\text{n}},\text{in},\text{therm}}(\omega) = \frac{8 k_{\text{B}} T}{R_{\text{FB}}} + \frac{8 k_{\text{B}} T \gamma_{\text{MOS}}}{g_{\text{m}} R_{\text{FB}}^2} \tag{3.32}$$

T is the temperature and γ_{MOS} is the MOS transistor noise coefficient. We can increase R_{FB} just until the bandwidth requirement in Equation 3.11 is exactly met so that total input-referred noise spectral density $S_{i_{\text{n}},\text{in},\text{AFE}} = S_{i_{\text{n}},\text{in},Q} + S_{i_{\text{n}},\text{in},\text{therm}}$ is minimized. Lastly, to avoid harmonic generation due to the distortion from clipping, maximum range of the ADC should be larger than the sum of the amplitude of the signal tone and the standard deviation of the instantaneous shot noise:

$$\frac{\Delta_{\text{in}}(2^N - 1)}{2} \geq A_{I_{\text{RX}}} + \sigma_{\text{in,PD}} = A_{I_{\text{RX}}} + \sqrt{N_{I_{\text{RX}}} \cdot \text{BW}_{\text{elec.}}}. \tag{3.33}$$

Based on ADC performance extracted from modern continuous-time sigma delta ADCs [43], let's assume that the ADC has 12-bit ENOB, 100 MHz bandwidth and 1 V input swing, and T_{meas} = 100 μs. The intrinsic gain of the modern nanoscale process is ~ 10. For C_{PD} = 0.1 pF, maximum R_{FB} to meet the bandwidth requirement is then 268 kΩ from Equation 3.30. For this much transimpedance, total noise power spectral density is roughly 0.35 pA/$\sqrt{\text{Hz}}$. In order for the shot noise to be dominant by 10 dB margin, P_{LO} should be at least 5 μW or −24.1 dBm. If T_{meas} is reduced to 10 μs, this goes up to −15.47 dBm.

3.4 Free Space Loss

Now that we have found out the receiver sensitivity to perform ranging as a function of measurement time, let's briefly discuss how much power can actually reach the receiver in a realistic system. One thing to notice is that, as mentioned in Section 2.2.8, the maximum power from the beam scanner is primarily limited by the eye safety regulation. For instance, IEC60825-1 Class 1 laser in 1550 nm band is limited to ~ 10 dBm.

The loss associated with the free space propagation and backscattering in a laser beam scanning-based remote sensing system is calculated using standard LiDAR/RADAR equation [44]:

$$P_{\text{RX}} = \frac{P_{\text{TX}}G_{\text{TX}}}{4\pi R^2}\sigma\frac{A_{\text{RX}}}{4\pi R^2}\eta_{\text{atm}}, \quad G_{\text{TX}} = \frac{16}{\theta_{\text{B}}^2}, \quad \sigma = 4\rho\cos\phi A_{\text{S}} \approx \pi R^2\theta_{\text{B}}^2\rho\cos\phi \quad (3.34)$$

G_{TX} is the transmitting antenna gain, θ_{B} is the transmitting beamwidth, σ is the backscatter cross-section area assuming typical diffusive target surface following Lambertian model. A_{S} is the target receiving area, A_{RX} is the receiver aperture area, η_{atm} is the atmospheric efficiency, ρ is the target reflectivity, and ϕ is the target incident angle (i.e. the angle between the beam direction and the target surface normal). It is also assumed that the size of the target surface is large enough to completely include the spot formed by the beam scanner. The free-space loss L_{FS} for target range R is finally expressed as follows:

$$L_{\text{FS}}(R) = \frac{P_{\text{RX}}(R)}{P_{\text{TX}}} = \frac{A_{\text{RX}}}{\pi R^2}\rho\cos\phi\eta_{\text{atm}}. \quad (3.35)$$

For instance, for maximum target distance of 300 m, the loss is −94 dB with 1 cm^2 aperture size, even with perfect reflectivity, no atmospheric loss, and normal incidence. At the same time, Equation 3.35 highlights that recording intensity of the LiDAR signal in addition to the frequency can provide useful information about the target surface ($\rho\cos\phi$), assuming that the characteristic of the atmosphere does not change rapidly during the measurement.

Table 3.1: Summary of FMCW LiDAR System Formula (SNR Threshold = θ, $T_{\text{meas}} = T_{\text{mod}}$)

Parameter	Expression
Point Generation per Second	$\dfrac{1}{T_{\text{meas}}}$
Ranging Resolution (ΔR)	$\dfrac{0.89c}{f_{\text{BW}}}$
Ranging Precision at Threshold (σ_R)	$\dfrac{c}{4\pi f_{\text{BW}}}\sqrt{\dfrac{3}{\theta}}$
Velocity Precision at Threshold (σ_v)	$\dfrac{\lambda_{\text{laser}}}{2\pi T_{\text{meas}}}\sqrt{\dfrac{3}{\theta}}$
Electrical Bandwidth Requirement	$\text{BW}_{\text{elec}} \geq \dfrac{2R_{\text{max}}}{c}\dfrac{2f_{\text{BW}}}{T_{\text{meas}}} + \dfrac{|v_{\text{max}}|}{\lambda_{\text{laser}}}$
Shot Noise-Limited SNR	$\dfrac{R_{\text{PD}}P_{\text{RX}}T_{\text{meas}}}{q}$
Receiver Sensitivity	$P_{\text{RX}} \geq \dfrac{q\theta}{R_{\text{PD}}T_{\text{meas}}}$
Worst-Case Free Space Loss (Lambertian)	$\dfrac{A_{\text{RX}}}{\pi R_{\text{max}}^2}(\rho\cos\phi\,\eta_{\text{atm}})$

3.5   Link Budget Analysis

Table 3.1 summarizes the FMCW LiDAR system formulas so far. Assuming the worst-case SNR operation on the verge of threshold effect, both ranging resolution and threshold are solely determined by the laser modulation bandwidth (f_{BW}). Once f_{BW} is given, the shot noise-limited receiver sensitivity and the receiver electrical bandwidth requirement are determined by the measurement time. Lastly, the LO power to guarantee shot noise-limited operation is determined by the input-referred noise of the analog frontend, which depends mostly on the electrical bandwidth requirement (which is dependent on T_{meas}) and ADC resolution and sampling rate. Again, the maximum power emission ($P_{\text{TX,max}}$) is limited by the eye safety regulation.

One can notice from Table 3.1 that parameters over which a system designer actually has control are T_{meas} and A_{RX}: one can either increase P_{RX} by increasing receiving aperture size or decrease the sensitivity requirement by increasing measurement time. From Equation 3.27, Equation 3.35 and assuming perfect target and system/atmospheric efficiency, we can directly see how
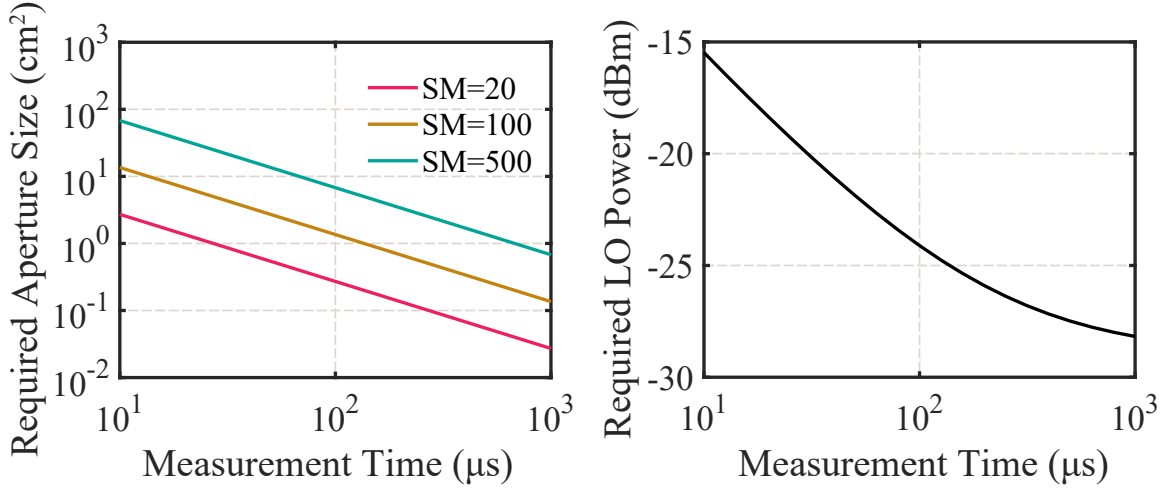
Figure 3.7: Required receiving aperture size and LO power for satisfying LiDAR receiver sensitivity requirement and ensuring shot-noise limited operation mode ($R_{max} = 300$ m).

much system margin (SM) do we have.

$$SM = \left(\frac{R_{PD}P_{TX,max}}{q\theta\pi R_{max}^2}\right) A_{RX}T_{meas} \tag{3.36}$$

This margin should provide enough buffer to absorb any additional non-idealities in the system, including not limited to:

- Low reflectivity (Equation 3.35)
- Oblique incidence (Equation 3.35)
- Atmospheric absorption (Equation 3.35)
- Effective fill factor of the receiving aperture
- Device insertion loss
- State of polarization misalignment between E_{LO} and E_{RX}
- Deterministic/stochastic modulation error in FMCW source (Section 3.2).

Figure 3.7 shows required A_{RX} and P_{LO} corresponding to the following parameters: $R_{PD} = 1$ A/W, $P_{TX,max} = 10$ dBm, $f_{BW} = 1$ GHz, $R_{max} = 300$ m, $v_{max} = 30$ m/s, $\theta = 30$, $C_{PD} = 0.1$ pF, $A_{in,ADC} = 1$ V, $ENOB_{ADC} = 12$, $f_{BW,ADC} = 100$ MHz. For example, 100 μs measurement time, which corresponds to 10 000 pts/s or 400 points per frame if the target frame rate is 25 Hz, requires 1.36 cm^2 aperture size to have 20 dB margin. Considering the angular resolution ($\sim 0.2°$) and FOV ($\sim 90°$) specifications from Table 2.1, only one horizontal line scan already exhausts this
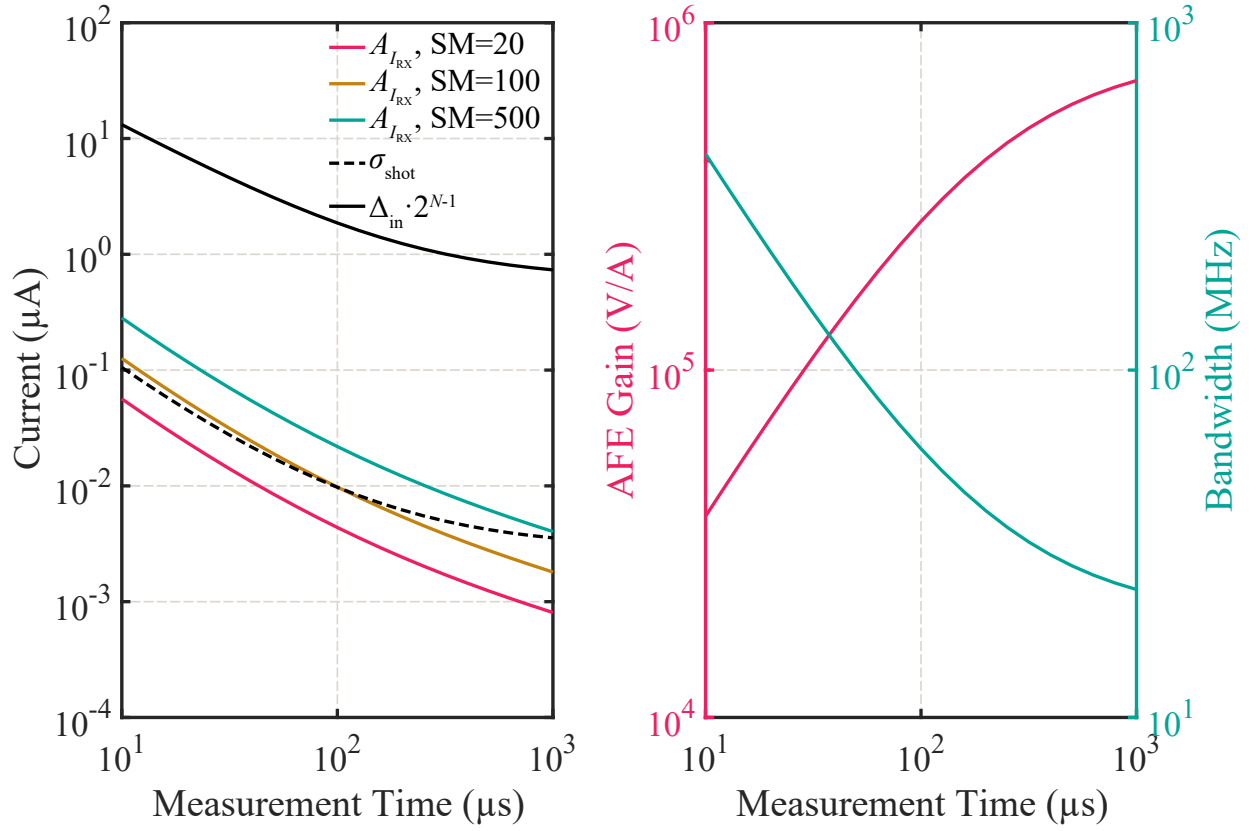
Figure 3.8: Photocurrent signal swing for different SNR margin targets as well as input-referred ADC range (left) and analog frontend gain as well as bandwidth (right) for operation scenario illustrated in Section 3.5.

many points. In sum, to have sufficient margin for a realistic system, at least $> 1\,\mathrm{cm}^2$ receiving aperture size is needed. Even with that much aperture, to support multiple lines on the vertical axis (or 2D imaging), the system is likely to employ multiple laser/receiver channels. Fortunately, required LO power corresponding to the T_{meas} of interest in Figure 3.7 is quite low, which makes it potentially possible for us to implement a multi-channel coherent detection system on a integrated photonics. If lower TIA gain than what is determined by $\mathrm{BW}_{\mathrm{elec}}$ and Equation 3.20 is used, P_{LO} should further increase to ensure shot noise-limited performance.

Figure 3.8 examines if the maximum amplitude range requirement in Equation 3.33 is satified. For all cases, input-referred maximum range of the ADC is around two orders of magnitude larger than the sum of the signal and the noise standard deviation, and as the free space loss is inversely proportional to the square of the range, the receiver does not saturate for the target range beyond
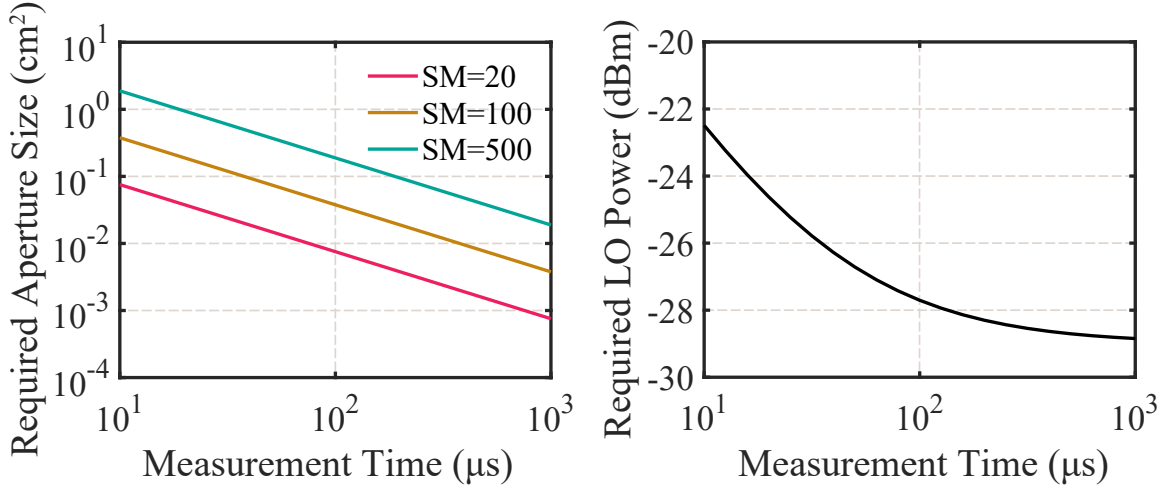
Figure 3.9: Required receiving aperture size and LO power for $R_{\mathrm{max}} = 50\,\mathrm{m}$.

1% of R_{max} ($\sim 3\,\mathrm{m}$).

Once the framework is in place, it is easy to get the same set of results for different operation scenarios. Figure 3.9 shows the same result as Figure 3.7, but now assuming the maximum range of 50 m, which is relevant for relatively close targets in low- to mid-speed driving scenarios or side- and backward-looking installations. As smaller range means less free-space propagation loss (Equation 3.35), required aperture size to guarantee the same SNR margin is much smaller (e.g. $T_{\mathrm{meas}} = 100\,\mu\mathrm{s}$ now requires a few millimeter square-sized aperture). At the same time, lower electrical bandwidth due to shorter range (Equation 3.20) allows higher TIA gain, which suppresses analog frontend noise and relaxes the LO power requirement.

3.6 Chapter Summary

In this chapter, a quantitative framework for beam-scanning FMCW LiDAR performance analysis is presented. Based on information-theoretical performance lower bound of line spectral estimation, the sensitivity of optical coherent frontend is derived assuming shot-noise limited operation and for given laser modulation bandwidth. By taking into account the free-space propagation loss associated with diffusive detection target, a set of LiDAR requirements such as size of the receiver aperture, LO laser power, and ADC dynamic range are easily identified. It is concluded that to achieve reasonable frame rate, at least centimeter square-scale receiver aperture is needed for long-range (up to 300 m) applications.

Chapter 4

Solid-State Optical Beam Scanning with Optical Phased Array[1]

As illustrated in Chapter 2, optical beam scanner is an essential part for long-range LiDARs, although we have been assuming that we somehow have an ideal scanner. 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 fact, there are a number of other applications beyond LiDAR: high spatial resolution of the optical beams can also enable fast and secure wireless communication [45], and can even potentially serve as the next-generation propulsion mechanism for spacecraft [46].

Beam scanning modules in existing commercial products are mostly based on mechanical control, such as motor-driven rotating collimation stages/mirrors [20] 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 [47], lens-assisted emitter

---

[1]The content of this chapter was derived from [14], a work 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 used in this work. Pavan Bhargava at UC Berkeley contributed to the design of analog CMOS circuits. Wafer fabrication and 3D integration was done by the CNSE.

arrays [48], [49], liquid crystal waveguides [50], and photonic crystal waveguides combined with diffraction gratings [51].

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 [52], [53]. OPA technology has progressed tremendously within a relatively short period of time, and multiple large-scale implementations with ∼1000 elements recently reported [54]–[58].

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 [57]–[59] 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 [57], [60]. 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, I 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 [15]. I 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 vali-

datation 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.

The remainder of this chapter is organized as follows. Section 4.1 starts by describing the basic operation principle of OPA-based beam steering. Then it introduces requirements for optical phased arrays (OPA) in the context of automotive LiDAR and presents an overview of practical challenges associated with meeting the desired performance. Section 4.2 introduces the wafer-scale 3D photonics-electronics heterogeneous integration platform used for our single-chip OPA. Section 4.3 provides the overview of our OPA architecture as well as building block design considerations. Then, experimental demonstration of the single-chip OPA prototype is presented in Section 4.4. Section 4.5 revisits the system requirements to discuss remaining challenges and illustrates future directions. Finally, Section 4.6 summarizes the chapter.

4.1 Optical Phased Arrays for Robust High-Resolution Beam Scanning

4.1.1 OPA Fundamentals and Key Metrics

Figure 4.1(a) shows a one-dimensional, uniform-spacing optical phased array (OPA) often used as a beam-steering transmitter. It includes an array of N antennas separated by a fixed spacing d, as well as the same number of phase shifters driven by electrical signals, which introduce a configurable phase offset (ψ_n) to each antenna. When the phase difference between adjacent antennas is constant for the whole array ($\psi_n - \psi_{n-1} = \Delta\psi$) and assuming that each individual antenna is an isotropic emitter, the resulting far-field intensity pattern for wavelength λ and wavenumber $k = 2\pi/\lambda$, normalized by the per-element intensity, is given as follows [61], [62]:

$$\hat{I}(\theta, \phi) = \left| \sum_{n=0}^{N-1} e^{jn(kd\sin\theta + \Delta\psi)} \right|^2 = \left| \frac{\sin\left[\frac{N}{2}(kd\sin\theta + \Delta\psi)\right]}{\sin\left[\frac{1}{2}(kd\sin\theta + \Delta\psi)\right]} \right|^2. \tag{4.1}$$

From Equation 4.1, one can notice that the peak intensity of N^2 is achieved when the following condition is met:

$$kd\sin\theta + \Delta\psi = 2M\pi, \ M \in \mathbb{Z}. \tag{4.2}$$

In other words, the *beam* is formed at an angle θ_b that satisfies Equation 4.2. Moreover, since θ_b is dependent on $\Delta\psi$, we can *steer* the beam by adjusting $\Delta\psi$. Figure 4.1(b) shows $\hat{I}$ for different
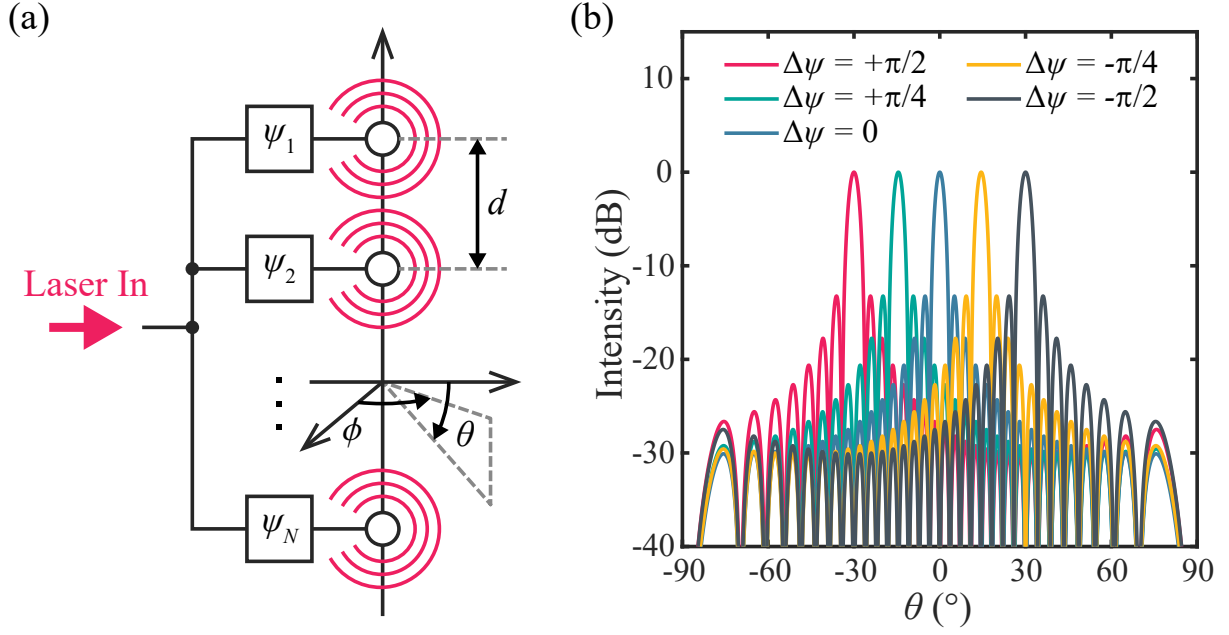
Figure 4.1: (a) Overview of the one-dimensional optical phased array (b) far-field intensity pattern for different phase difference $\Delta\psi$ ($N = 32$, $d = \lambda/2$).

$\Delta\psi$ values ($N = 32$, $d = \lambda/2$). Note that $\hat{I}$ depends on θ through a sine function. Due to this nonlinear mapping, the overall beam pattern is broadened for larger $|\theta_b|$.

From the system design perspective, the main performance metrics of the OPA are as follows: beam width, maximum steering range, emission efficiency (η), and side-lobe suppression ratio (SLSR).

Beam width ($\Delta\theta_b$) determines the lateral resolution of the imaging system based on the OPA. From Equation 4.1, the expression for the full-width at half-maximum (FWHM) beam width (in radians) is derived as follows [62]:

$$\text{FWHM} = \cos^{-1}\left(\sin\theta_b - \frac{2.78}{kNd}\right) - \cos^{-1}\left(\sin\theta_b + \frac{2.78}{kNd}\right). \tag{4.3}$$

For instance, the beam FWHM at $\theta_b = 0$ is $\frac{2.78\lambda}{\pi Nd}$. This implies that the lateral resolution is mainly determined by the inverse of the overall size of the array aperture ($W = N \cdot d$). As previously mentioned, beam width broadens as $|\theta_b|$ increases, and therefore the worst-case beam resolution is obtained at the extremes of the steering range.

We can also calculate the maximum steering range of an OPA from Equation 4.2 assuming

that a phase shifter is capable of supporting full 2π range:

$$|\theta_{\rm b}| < 2\sin^{-1}\left(\frac{\lambda}{2d}\right). \tag{4.4}$$

This is equal to $\pi/2$ when $d = \lambda/2$, which allows the array to address the entire hemisphere. Namely, smaller antenna spacing as close as possible to the half wavelength is desired to maximize the field-of-view (FOV) of the array.

Transmitter emission efficiency (η), defined as the ratio between the amount of power focused within the main beam ($P_{\rm b}$) and the total input power ($P_{\rm in}$) (i.e. $\eta = P_{\rm b}/P_{\rm in}$), is particularly important in our context as it is effectively added into the wall-plug efficiency of the source laser and affects the total system power consumption. As an example, we can calculate the theoretical emission efficiency of a 1D array of N antennas with Equation 4.1 and Equation 4.3:

$$I_{\rm element}(\theta, \phi) = G(\theta, \phi)\frac{P_{\rm in}}{4\pi N}, \quad G|_{\theta=0} = \alpha \tag{4.5}$$

$$\begin{aligned}
\eta|_{\theta_{\rm b}=0} &= \frac{1}{P_{\rm in}} \int_0^{2\pi} \int_{-\pi/2}^{+\pi/2} I_{\rm element}(\theta, \phi)\hat{I}(\theta, \phi)|_{\Delta\psi=0} \cos\theta d\theta d\phi \\
&\approx 2\pi\alpha({\rm FWHM})\left(\frac{N^2}{4\pi N}\right) = \left(\frac{2.78\lambda}{\pi N d}\right)\frac{\alpha N}{2} = \frac{2.78\lambda}{2\pi d}\alpha.
\end{aligned} \tag{4.6}$$

α is the antenna gain of single element at the direction of the beam. Note that the efficiency is independent of the antenna count and depends solely on the antenna spacing.

Lastly, sufficient sidelobe suppression is required in LiDARs to ensure that we only receive clean signal from the target aligned with the main beam at $\theta_{\rm b}$ and not spurious signals from sidelobes. For a uniform-spacing array with constant intensity across the array, nominal SLSR is equal to the ratio between the main beam and the first sidelobe, which be calculated from Equation 4.1.

$$\text{SLSR} = \left|N\sin\left(\frac{3\pi}{2N}\right)\right|^2 \tag{4.7}$$

This quickly converges to ~13.4 dB as N increases beyond ~20. Note that it is potentially possible to increase the SLSR limit by using non-uniform intensity pattern, such as Gaussian pattern.

4.1.2 OPA LiDAR System Requirements

We can now look back at Section 2.3 to get a sense of what the performance expectations are for optical phased arrays (OPAs) from a system-level perspective. From the requirement list in Table 2.1, we can notice that lateral resolution and FOV specifications are directly related to the

beam scanner design. In addition, the beam scanner design will also be limited by power and cost constraints:

- Lateral Resolution: $0.1° \sim 0.2°$
- Field of View (FOV): >90°
- Power Budget: $10\,\text{W} \sim 30\,\text{W}$
- System Cost: $100~$200

Above all, the scanner must satisfy resolution requirements so as to recognize small objects (e.g. pedestrians) at long distance, and we saw from Equation 4.3 that the beamwidth/resolution is determined by the overall array aperture size. For example, 0.2° worst-case resolution for 90° FOV and $\lambda = 1550\,\text{nm}$ requires $W = 0.54\,\text{mm}$. Meanwhile, the ambiguity-free steering range of a phased array was determined by the antenna spacing (Equation 4.4). Namely, the spacing should decrease until it meets the range requirement, which results in increased number of antennas. For instance, 90° FOV would require 0.71λ-spacing ($d \sim 1.1\,\mu\text{m}$), which corresponds to 490 antennas for $W = 0.54\,\text{mm}$.

Alternatively, one can achieve the same FOV with two adjacent OPA channels and relax the spacing to 1.3λ ($d \sim 2\,\mu\text{m}$). 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 [55], [58], which alters the basic relationship of Equation 4.4. Nevertheless, to be a compelling alternative to mechanical beam scanners, a clear path to scale the element count to $500 \sim 1000$ at a spacing of $\sim \lambda$ is a must.

4.1.3   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 most simple and straightforward OPA architecture is the tree architecture (Figure 4.2(a)), 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 $500 \sim 1000$ 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.
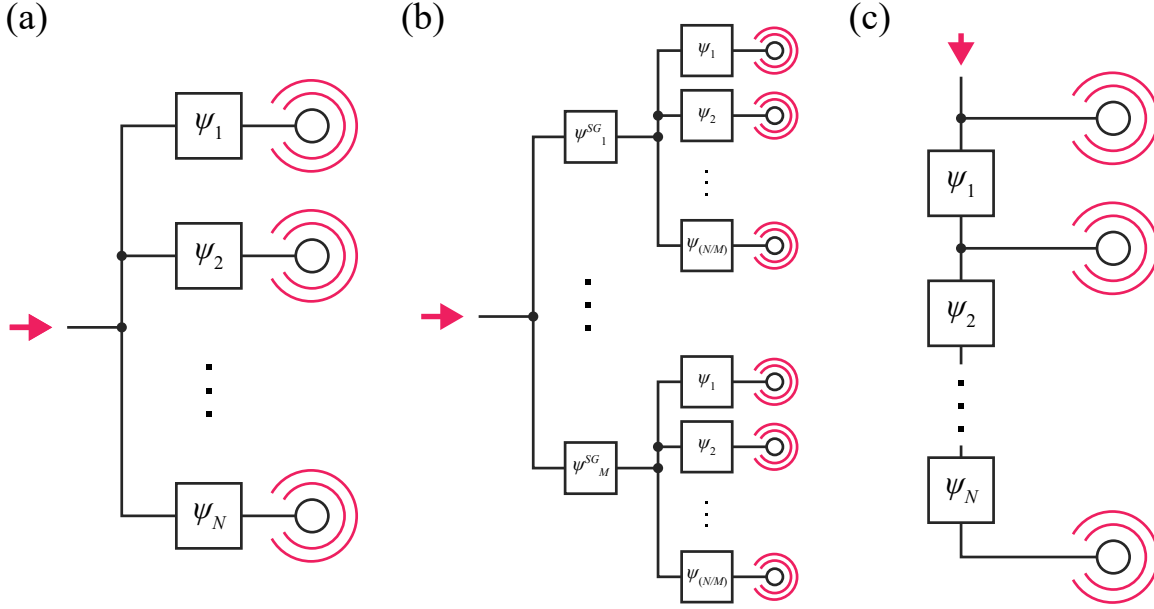
Figure 4.2: Optical phased array distribution network types: (a) tree architecture, (b) grouped tree architecture, and (c) cascaded architecture.

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 4.2(b)) [57]. In this design, antennas 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 [57] to improve the granularity of the phase adjustments.

Another solution is the cascaded architecture (Figure 4.2(c)) [59]. 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 [59].

Lastly, 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 [58], 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 [58].

4.1.4   Process and Design-Dependent Random Phase Fluctuation

In theory, alternative OPA architectures in Section 4.1.3 should resolve the complexity bottleneck in large-element count OPAs. Unfortunately, their effectiveness could be limited in a mass production scenario. Any realistic silicon photonics process suffers from process variation [63] which manifests itself as nanoscale structural uncertainties. It results in waveguide geometry variation and perturbs the effective index and the phase of the field within the waveguide [64]. Since the local phase fluctuation accumulates as the field progresses, the total phase error at the output of the waveguide is often modeled as a random walk. If there are two waveguides of length L_1 and L_2 and the physical separation between two waveguides is Δx, the expression for the variance of the relative phase offset observed at the outputs of the waveguides is as follows [64]:

$$\sigma^2_{\psi_1-\psi_2}(L_1, L_2) = \frac{2|L_1 - L_2|}{L_{\text{coh}}}, \quad \Delta x < \sim 700\,\mu\text{m} \tag{4.8}$$

L_{coh} is called coherence length, and typical value for standard strip waveguide and a SOI process with 220 nm-thick body is ~ 4 mm [64]. Equation 4.8 provides a simple model for the relative phase uncertainty between adjacent antennas:

$$\sigma^2_{\Delta\psi} = \frac{2d_\psi}{L_{\text{coh}}}. \tag{4.9}$$

d_ψ as the spacing between neighboring phase shifters, which is on the order of tens of microns in typical designs.

Using Equation 4.9, we can simulate the impact of the random phase fluctuations on the emission pattern. Figure 4.3 (a), (b) shows the simulated phase fluctuation and corresponding beam pattern ($N = 128, d = \lambda/2$) from a 1-dimensional array. d_ψ is assumed to be equal to the size of 30 μm-long phase shifter. It can be noted that while the width and the direction of the main beam is largely unchanged, the intensity of the spurious beams and overall noise floor go up significantly. Figure 4.3 (c), (d) shows the Monte-Carlo simulation results from different antenna counts ($N = 32, 128, 512$), where side lobe suppression ratio (SLSR) and emission efficiency
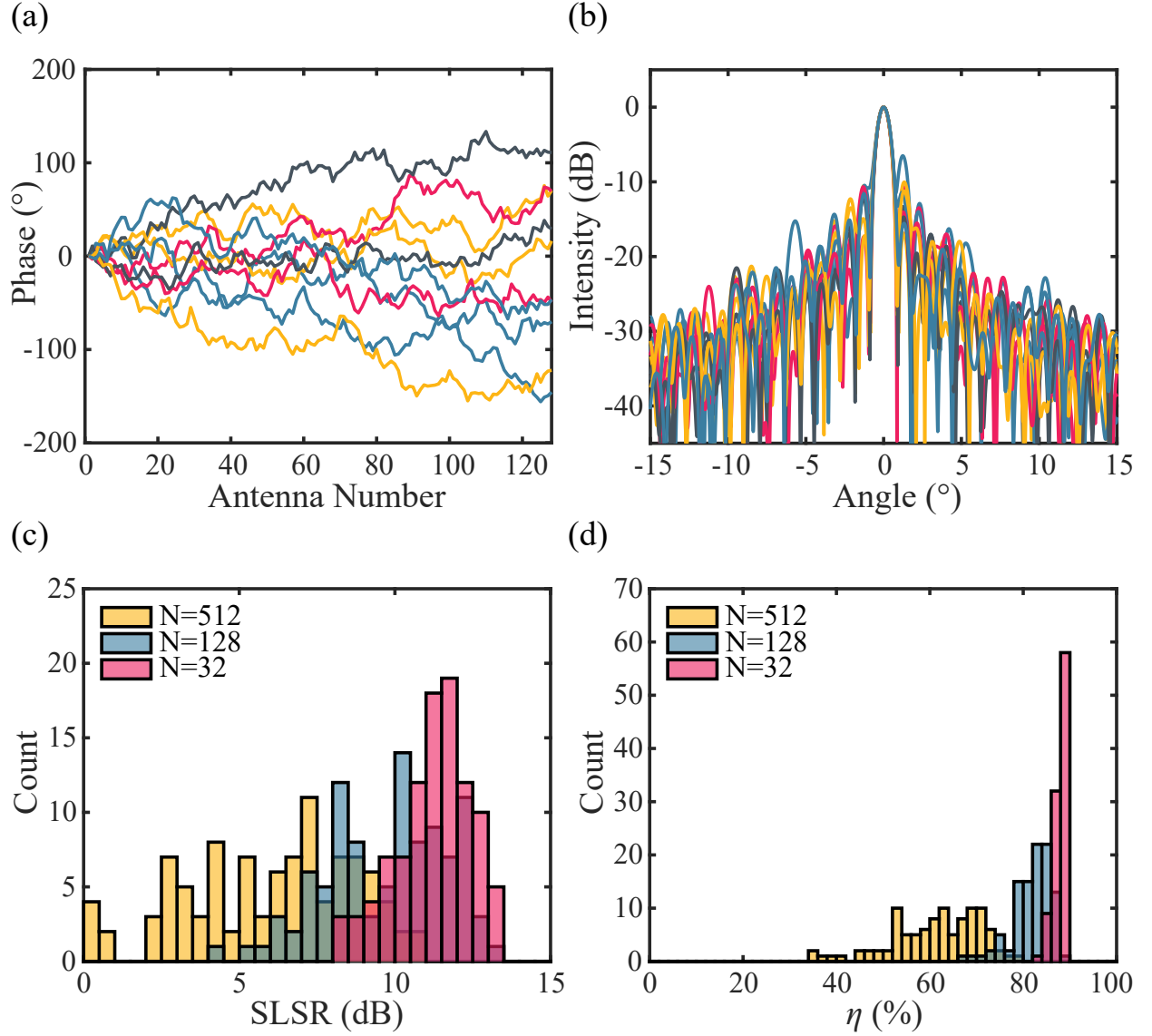
Figure 4.3: Monte Carlo simulation results for (a) phase fluctuation error pattern and (b) corresponding far-field intensity assuming 128 antennas, (c) side-lobe suppression ratio histogram for 512, 128, and 32 antennas, and (d) transmitter radiation loss histogram for 512, 128, and 32 antennas.

($\eta = P_{\text{beam}}/P_{\text{input}}$) was used as the evaluation metrics. It is clear that the degradation is much more severe in OPAs with increased element count. Increased element count/bigger aperture is equivalent to finer far-field angular resolution, which results in higher phase precision requirement [57].
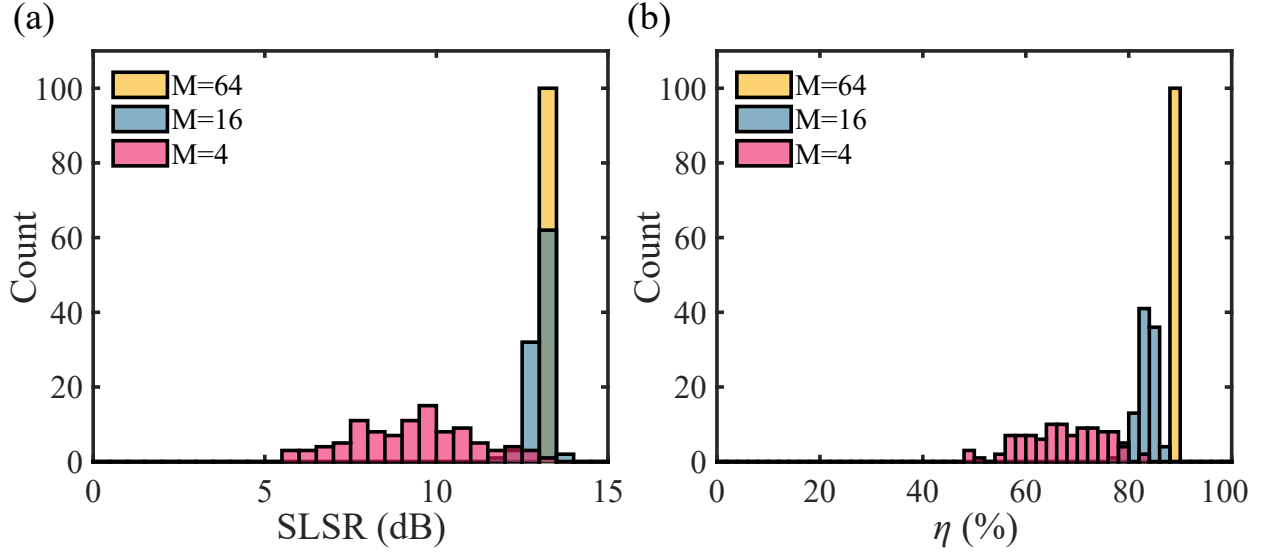
Figure 4.4: Monte Carlo simulation results after calibration for a 512 element array with different subgroup sizes (M = 64, 16, and 4): (a) side-lobe suppression ratio histogram and (b) transmitter radiation loss histogram.

Still, phase fluctuation from process variation is a static error and can be corrected with one-time calibration. Figure 4.4 shows post-calibration results from 100 Monte-Carlo simulations using 512-element OPAs with grouped tree architecture for three different subgroup size ($M = 4, 16, 64$). It is clear that the mean and variation of the beam quality improves for higher level of calibration granularity, and approaches ideal performance for the case of $N/M = 8$. Indeed, the subgroup size of the grouped tree architecture in [57] was set to be 8.

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 4.1.2, 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 [65]. 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.
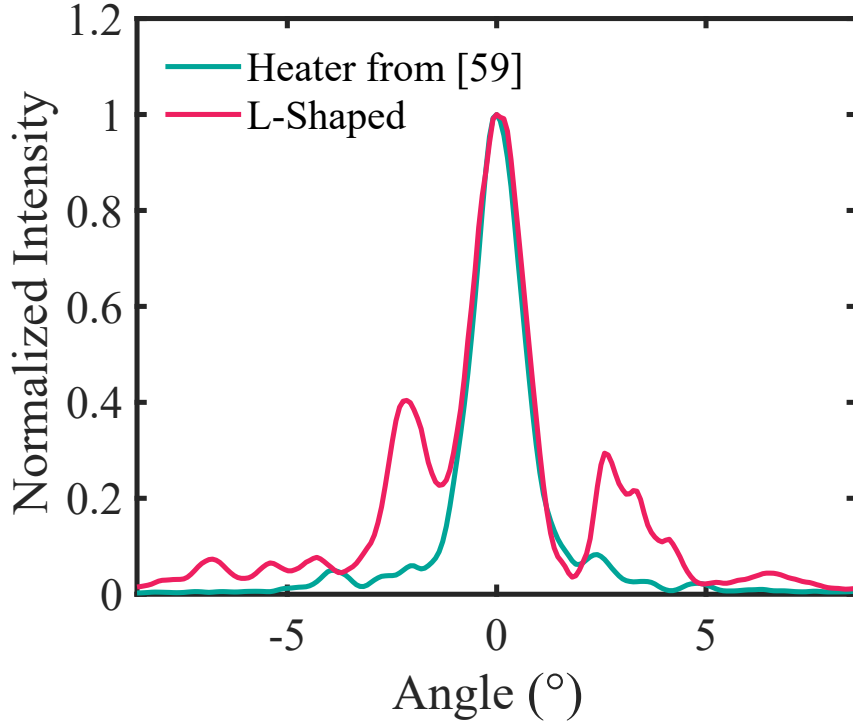
Figure 4.5: Far-field emission pattern measurement results from two OPAs with different heater designs (54 elements, 1.4 μm spacing), one with heaters from [59] and the other with L-shaped heaters from this work.

As explained later in detail (Section 4.3.2), 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 [64]. Figure 4.5 shows the measured far-field beam pattern of two OPAs ($N = 54$, $d = 1.4$ μm). The two OPAs are identical except for the phase shifter: one uses the same heater design as in [59] (11 V swing) and the other uses the CMOS-compatible L-shaped design proposed in this work. It is clear that the the coherence is substantially worse in the L-shape heaters. Namely, an analysis based on the coherence of simple strip waveguides can be rather optimistic, especially in OPAs where components are heavily optimized to satisfy power and area constraints. As explained in Section 4.3, this eventually motivated us to pursue per-element phase control flexibility, which completely desensitizes OPA performance to the level of device coherence.
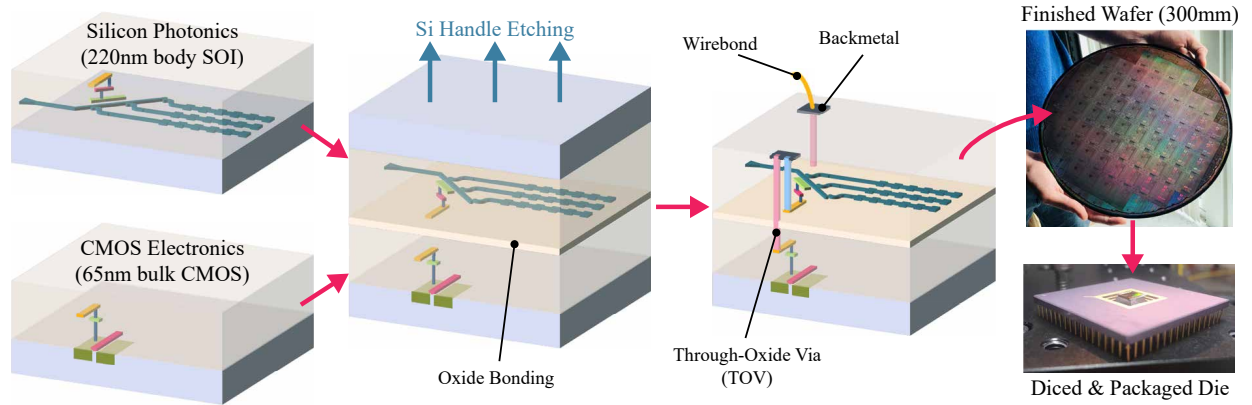
Figure 4.6: Overview of the 3D heterogeneous integration platform used to construct the single chip OPA.

4.2 Wafer-Scale 3D Heterogeneous Integration of Silicon Photonics and CMOS Electronics[2]

In [57] and [60], 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 [57]. However, 2D monolithic integration has a few critical disadvantages in the context of OPAs. First, photonic device design is 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 mW ∼ 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 [57], [60].

In this work, we realized single-chip OPAs on a 3D heterogeneous integration platform shown in Figure 4.6, similar to [66]. 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

---

[2]Wafer fabrication and 3D integration were done at the CNSE SUNY Albany.

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 μm 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 [67], much higher (> 7×) interconnect density versus copper pillars, flexible via/circuit placement that minimizes wiring/placement overhead, and lower unit cost of wafer-scale 3D integration make our platform better suited for OPAs.

4.3 Optical Phased Array Implementation[3]

Figure 4.7 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 4.1.3. 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 (as was done in Figure 4.5). Second, the cascaded architecture is generally less affected by thermal crosstalk [68] since the nominal phase pattern does not have discontinuities coming from limited phase shifter range, unlike the (grouped) tree architecture. Nevertheless, any architecture can eventually be supported once the functionality of the system is verified, and we will revisit this in Section 4.5.

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 4.7). The elements are placed at a 4 μm pitch in our prototype, which corresponds to the theoretical steering range ($2|\theta_b|_{\max}$) of 22.3° at $\lambda = 1550$ nm from Equation 4.4. However, the antenna ele-

---

[3]Photonic circuits illustrated in this section was architected, designed, and laid out by C. V. Poulton at MIT.
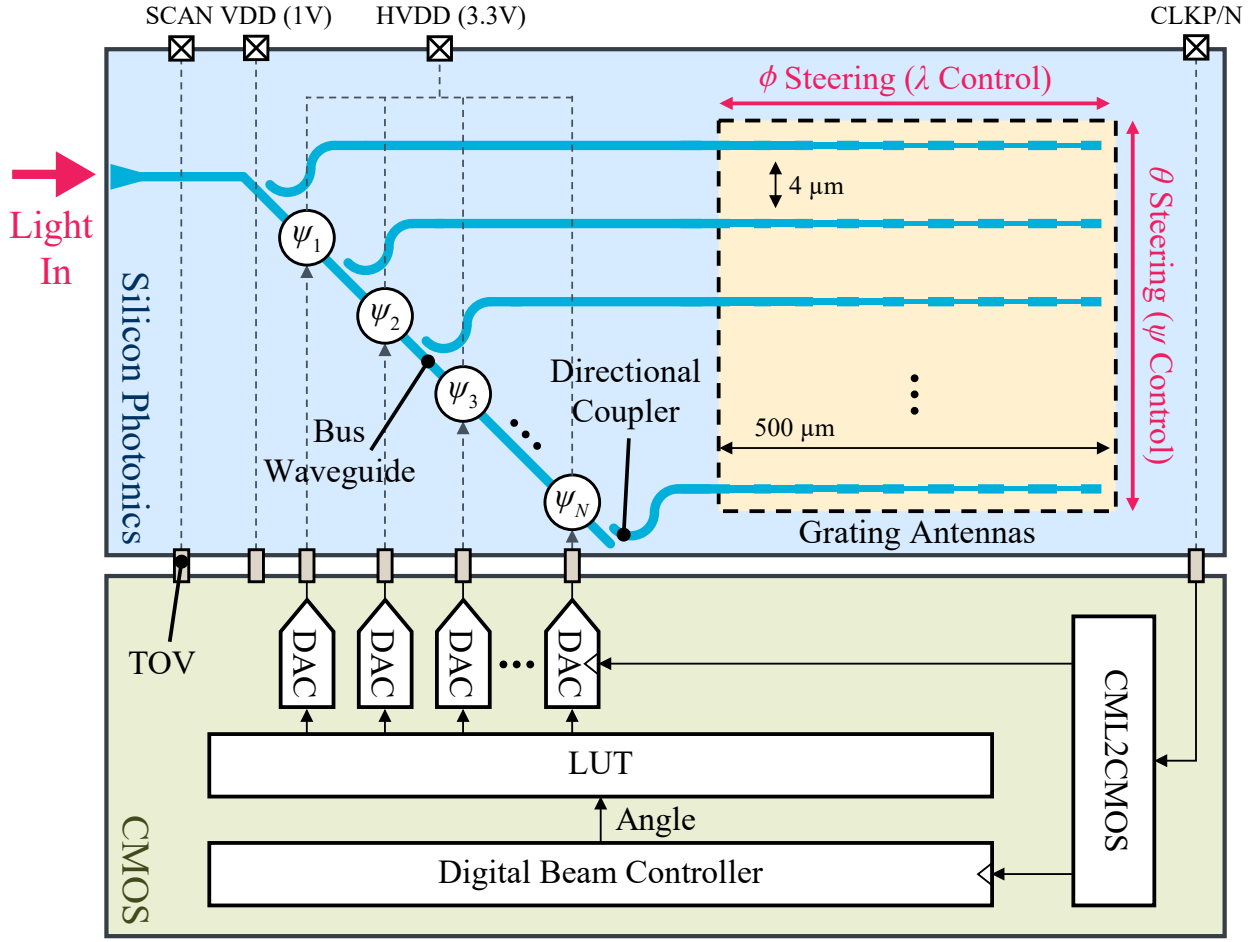
Figure 4.7: Overview of the single-chip OPA architecture.

ments used in our work can be placed at a tighter pitch down to the wavelength scale without inter-antenna evanescent coupling [65] (e.g. 1.4 μm pitch was used in Figure 4.5). 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. This 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 4.7) 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.

4.3.1 Apodized Grating Antenna

Figure 4.8(a) 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 μm, which enables a large emitting aperture and small divergence angle in ϕ (~0.15°). This is unlike previous work where short grating couplers (3.55-μm-long) with divergence angles over 10° were used as emitting elements [57].

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* [69]). 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 4.8(a), (b)), again highlighting the benefit of utilizing fully customized photonics.

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:

$$\frac{2\pi}{\lambda}\Delta x(\sin\phi + n_{\text{eff}}) = 2\pi M, \ M \in \mathbb{Z}, \tag{4.10}$$

$$\phi_{\text{b}} = \sin^{-1}\left(\frac{M\lambda}{\Delta x} - n_{\text{eff}}\right), \tag{4.11}$$

where n_{eff} is the effective index of the silicon waveguide. We designed the grating to satisfy $n_{\text{eff}}\Delta x = \lambda$, $M = 1$ at $\lambda = 1550$ nm. Since ϕ_b depends always on λ, beam steering along ϕ is possible via input wavelength tuning, where steering efficiency is expressed as follows (n_{g}:
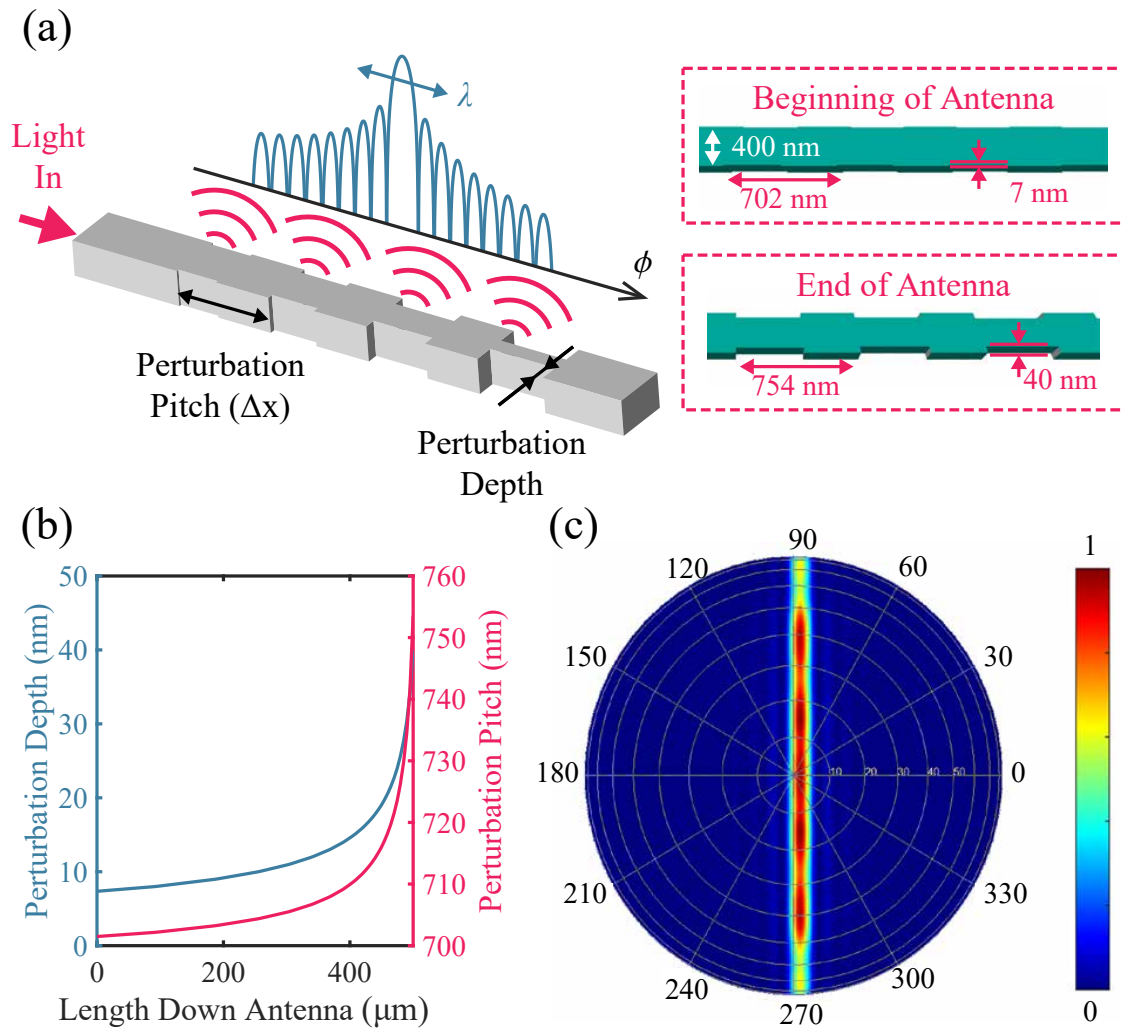
Figure 4.8: (a) Apodized grating antenna overview and the dimensions at the beginning and at the end of the antenna (b) Perturbation distance and pitch distribution across the antenna element (c) Antenna emission pattern from an FDTD simulation.

waveguide group index):

$$\frac{d\phi_\text{b}}{d\lambda} = \frac{\frac{1}{\Delta x} - \frac{dn_\text{eff}}{d\lambda}}{\sqrt{1 - \left(\frac{\lambda}{\Delta x} - n_\text{eff}\right)^2}} = \frac{n_\text{eff}}{\lambda} + \frac{n_\text{g} - n_\text{eff}}{\lambda} = \frac{n_\text{g}}{\lambda}. \tag{4.12}$$

The simulated group index at $\lambda = 1550\,\text{nm}$ is 4.43, which results in a steering efficiency of $0.164°$ per 1 nm wavelength shift.

4.3.2 L-Shaped Thermo-Optic Phase Shifter

As introduced in Section 4.1.4, we had two key objectives for phase shifter design: high phase efficiency and CMOS-compatible swing. Figure 4.9 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 4.9(b), (c)). A resistive heater is directly embedded into the bus waveguide by N-doping the slab, 350 nm away from the waveguide core (Figure 4.9(b)). This ensures low thermal impedance between the heater and the core, resulting in high thermal efficiency ($20\,\text{mW}/\pi$ from COMSOL simulation). At the same time, the waveguide mode and doped region are isolated in the L-shape geometry (Figure 4.9(c)), which ensures minimum optical loss (simulated loss: 0.016 dB). The length of one phase shifter is $32\,\mu\text{m}$ and, considering that the thermo-optic coefficient of silicon at 1550 nm is $1.8 \times 10^{-4}\,\text{K}^{-1}$, 2π shift corresponds to $270\,\text{K}$ 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 4.9(b)). 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 $270\,\Omega$, 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 4.9(c), (d)). 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: $0.015\,\text{dB} \sim 0.02\,\text{dB}$ across $1500\,\text{nm} \sim 1600\,\text{nm}$). The coupling strength of each directional coupler is gradually increased to ensure uniform distribution of the optical power across the entire aperture (Figure 4.9(e)). After the desired length of evanescent coupling, the strip waveguide is separated away from the bus ridge waveguide and adiabatically tapered back to a single-mode width of 400 nm for routing to the antenna.
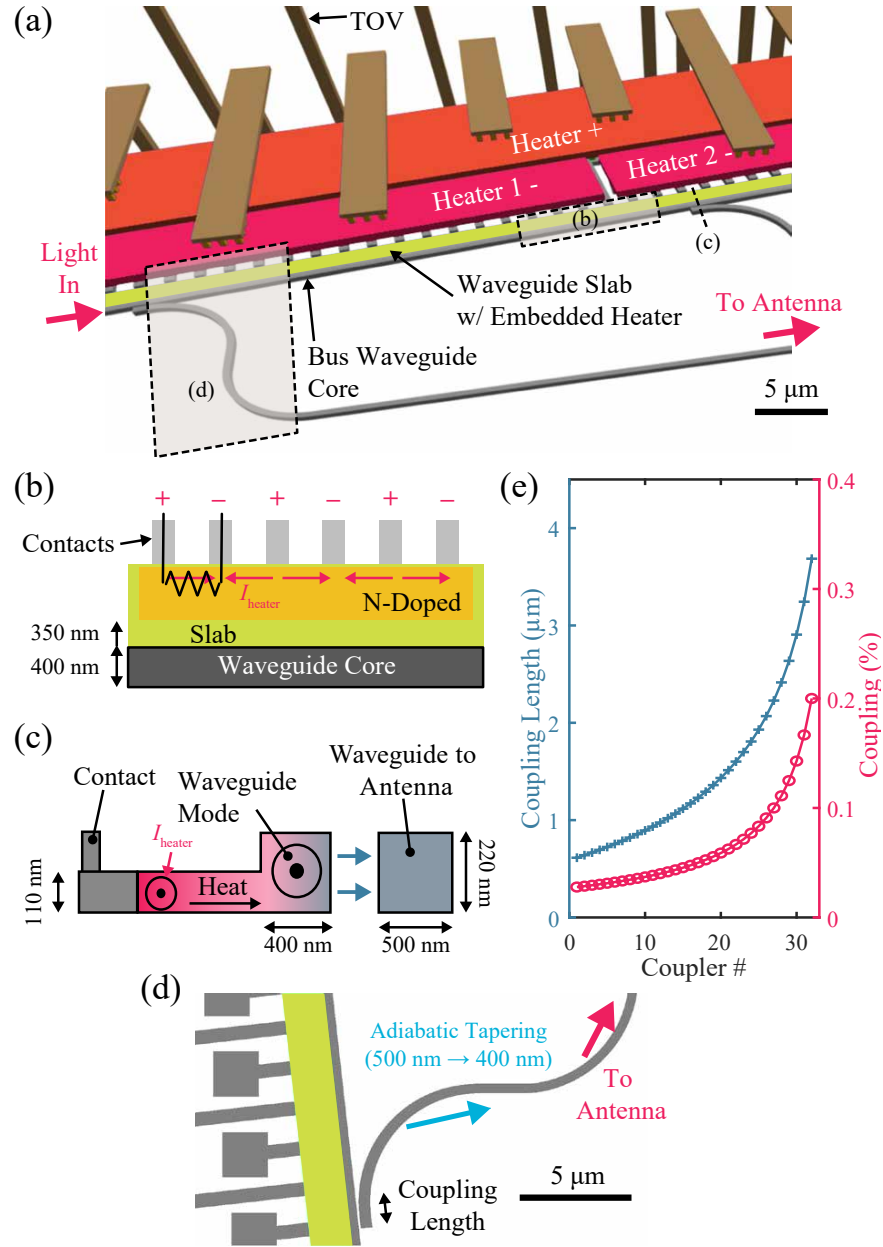
Figure 4.9: (a) Perspective view of the layout details around the bus waveguide section including embedded thermo-optic phase shifter and directional coupler. (b) The top view and (c) the cross-section views of the L-shaped phase shifter. (d) Bus waveguide-antenna connection through a directional coupler. (e) Evanescent coupling strength distribution across the bus waveguide for uniform power distribution.
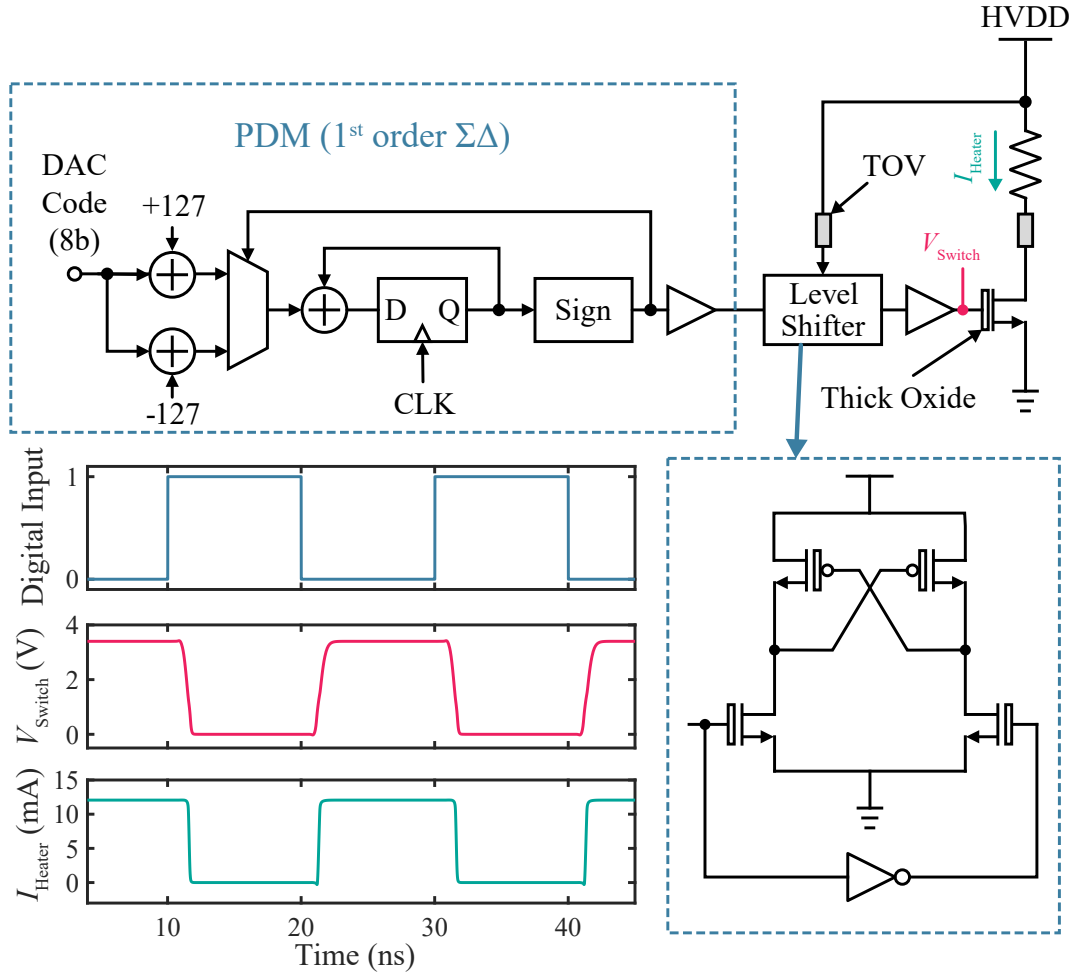
Figure 4.10: Overview of the PDM-driven switch-mode driver connected the photonic heater element, as well as the simulated time-domain waveforms of PDM signal, heater switch gate voltage, and heater current.

4.3.3 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 (Section 4.1.2), 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 4.10 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.
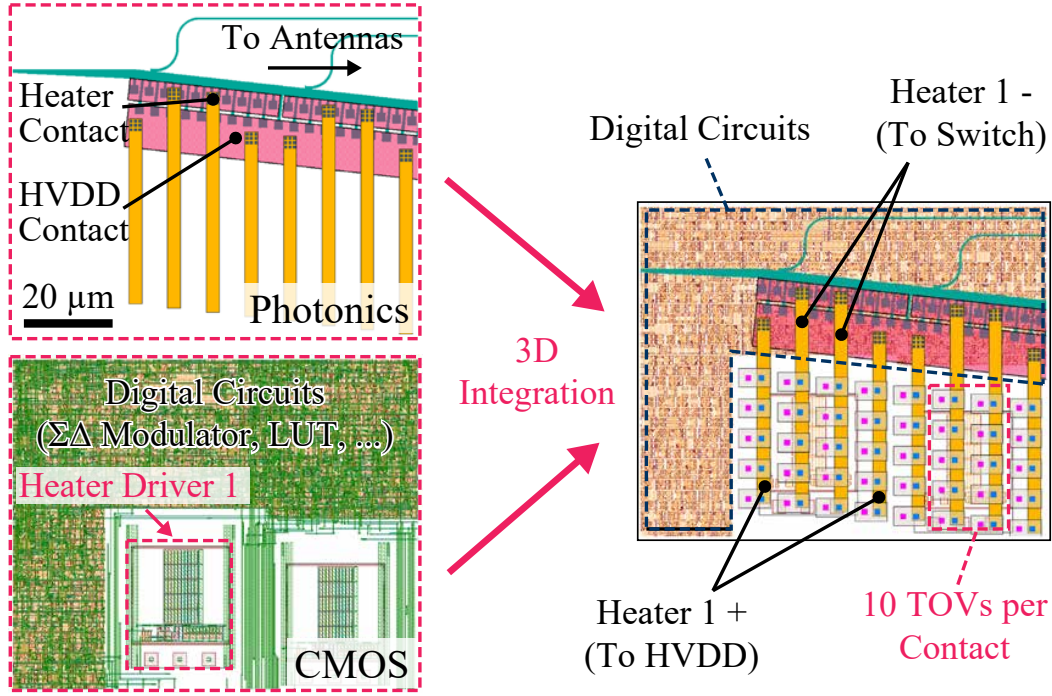
Figure 4.11: Relative placement and TOV-based routing of the thermal phase shifter and the CMOS driver chain, surrounded by digital circuits including PDM modulators and on-chip LUT.

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 [58]). The 1st-order $\Delta\Sigma$ modulator in the digital domain generates a pulse density modulated (PDM) signal [70] 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 (Figure 4.10). In addition to higher effiency, 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 4.11 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 × 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 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 μW in the PDM modulator), much smaller than the heater power. In reality, the measured thermal bandwidth was 32.3 kHz (Section 4.4) and, even for 11-bit resolution to support 1000 elements, $\sim$ 5 MHz clock rate (or $\sim$ 150 oversampling ratio) is sufficient (20 μW CMOS power).

4.4 Experimental Results

Figure 4.12 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 (further details and dimensions are shown in Figure 4.16 and Figure 4.17). The corresponding top-level layout of the CMOS 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 4.13 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 4.3.1.

Characterization results for the phase shifter and controller circuit are presented in Figure 4.14. As can be seen from the power versus DAC code plot (Figure 4.14(a)), 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
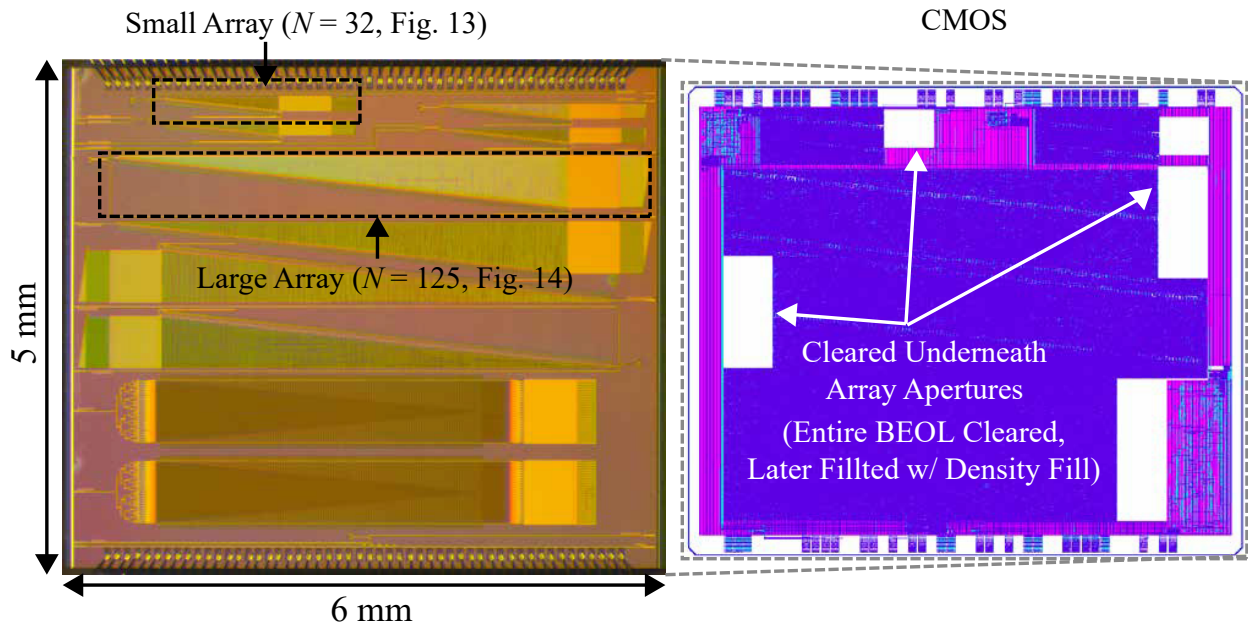
Figure 4.12: Die micrograph of the OPA chip and the GDS image of the CMOS, located underneath the visible photonics layer.
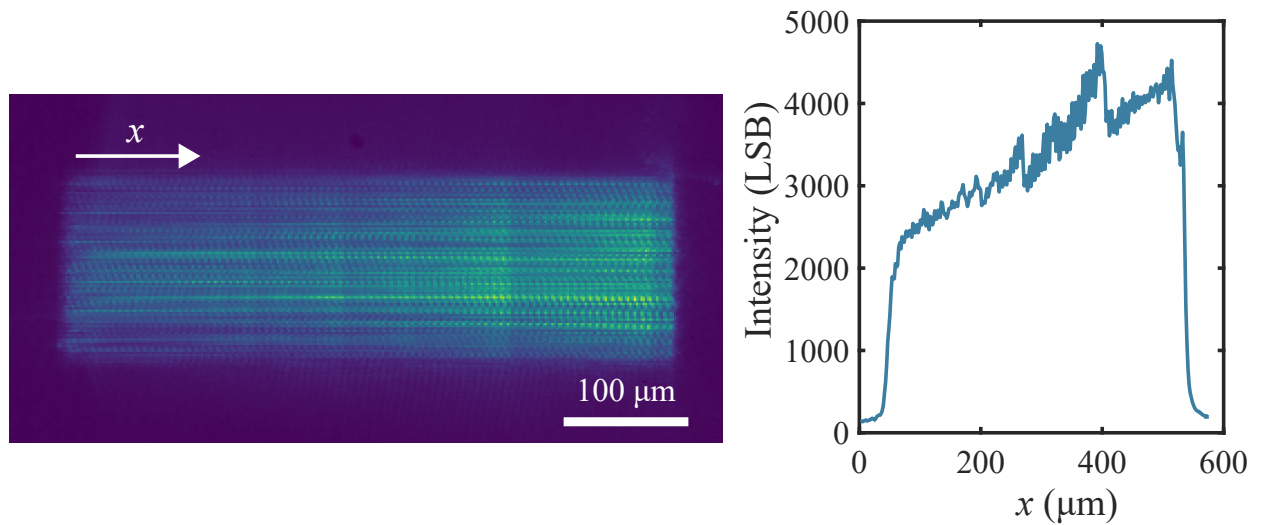
Figure 4.13: Near-field image of the illuminating small array aperture and its cross-section along the grating antenna.
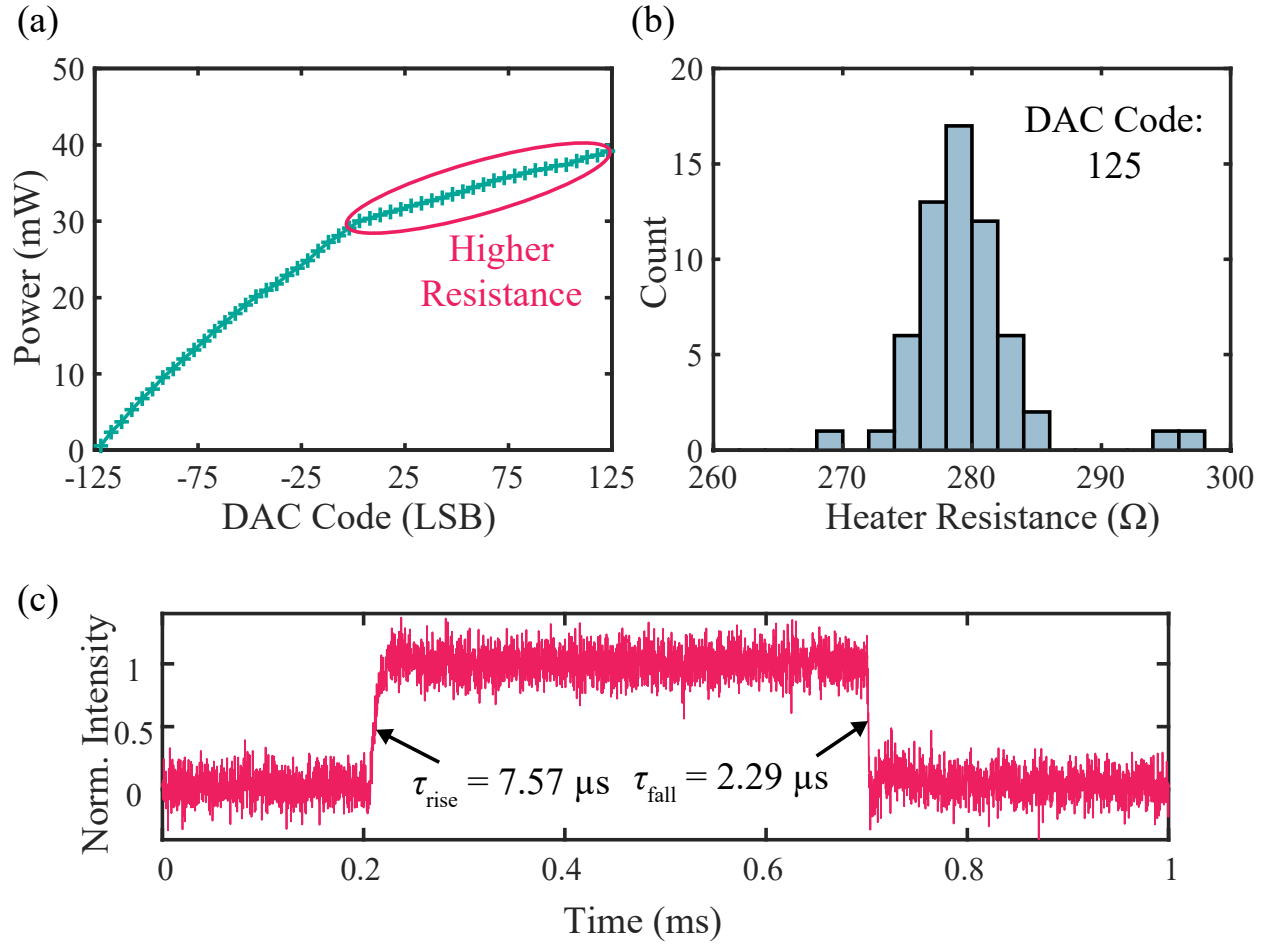
Figure 4.14: Measurement results of the DAC and heater: (a) Power vs. DAC Code, (b) statistics of the heater resistance, and (c) thermal transient response.

heater resistance is close to the expected value of 270 Ω even if the resistance is slightly increased in the high power range (Figure 4.14(b)), again due to the temperature dependence.

The transient response of the heaters is also characterized (Figure 4.14(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 µs, corresponding to 32.3 kHz bandwidth assuming a single-pole response. In thermal phase shifters, asymmetrical time constants for heating and cooling are often observed due to a non-ideal heat-sink: normally the total cooling time constant is expected to be larger.

Figure 4.15(a) shows the bench-top setup for characterizing the performance of an OPA and 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 10× or 20× infinity-corrected objective lens (f = 20 mm or 10 mm), 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 = 200 mm) forms an image of the Fourier plane at the camera sensor (e.g. $27.5° \times 22°$ field-of-view and 0.086° per pixel for the 20× 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 4.15(a), we can also carry out the array calibration to adjust the DAC LUT for optimum beam quality. Figure 4.15(b) 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. 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 4.16. In this example, the smaller variant with 32 elements (aperture size: 0.5 mm × 0.13 mm) is used. The beam is tightly focused in the ϕ direction (FWHM$_\phi \approx 0.15°$) and almost identical to the expected beamwidth for a 0.5 mm aperture and λ = 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 4.16 also presents a full two-dimensional beam steering demonstration based on the same calibration process with two iterations at each beam position, over 16° in θ via phase shifter control and over 18.5° in ϕ by tuning the input laser wavelength across

(a)

(b)

1: Pick desired beam position coordinate
2: Set initial DAC inputs
3: **while** $i <$ maximum iteration count **do**
4:     **for** $n \leftarrow 1$ to N **do**
5:         **for** $\epsilon \in \{\cdots, -\Delta_i, 0, +\Delta_i, \cdots\}$ **do**
6:             $\text{DAC}_{n,i} \leftarrow \text{DAC}_{n,i-1} + \epsilon$
7:             Measure beam quality
8:         **end for**
9:         Pick ϵ_{best} with maximum beam quality
10:         $\text{DAC}_{n,i} \leftarrow \text{DAC}_{n,i-1} + \epsilon_{\text{best}}$
11:     **end for**
12: **end while**
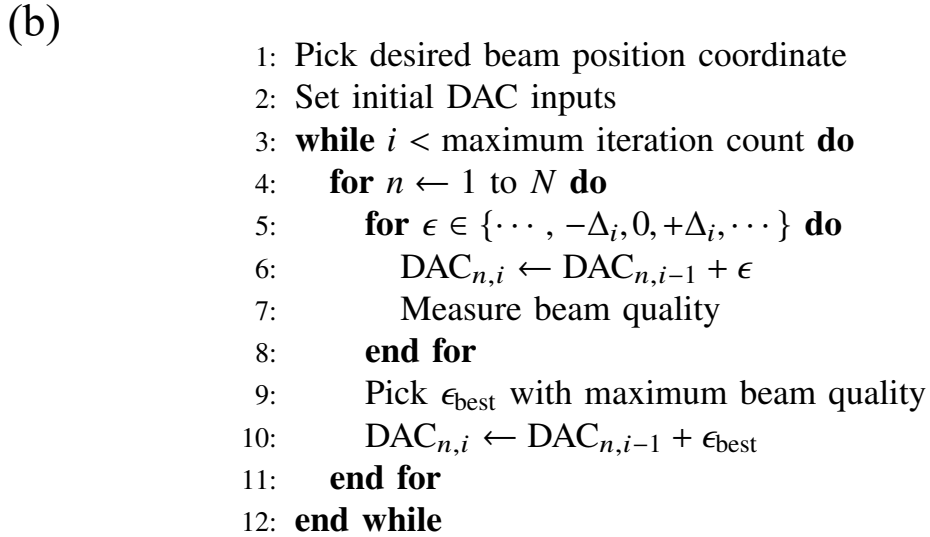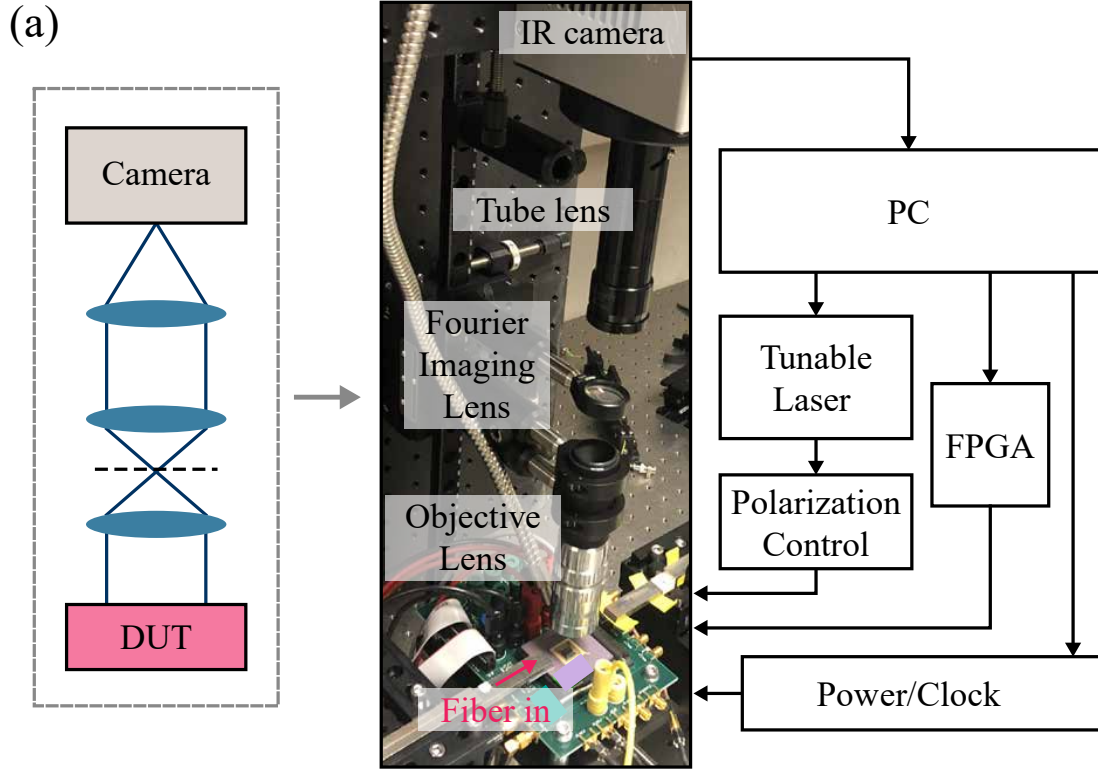
Figure 4.15: (a) Experimental setup including far-field imaging optics. (b) Pseudo-code of the local search-based beam calibration process.
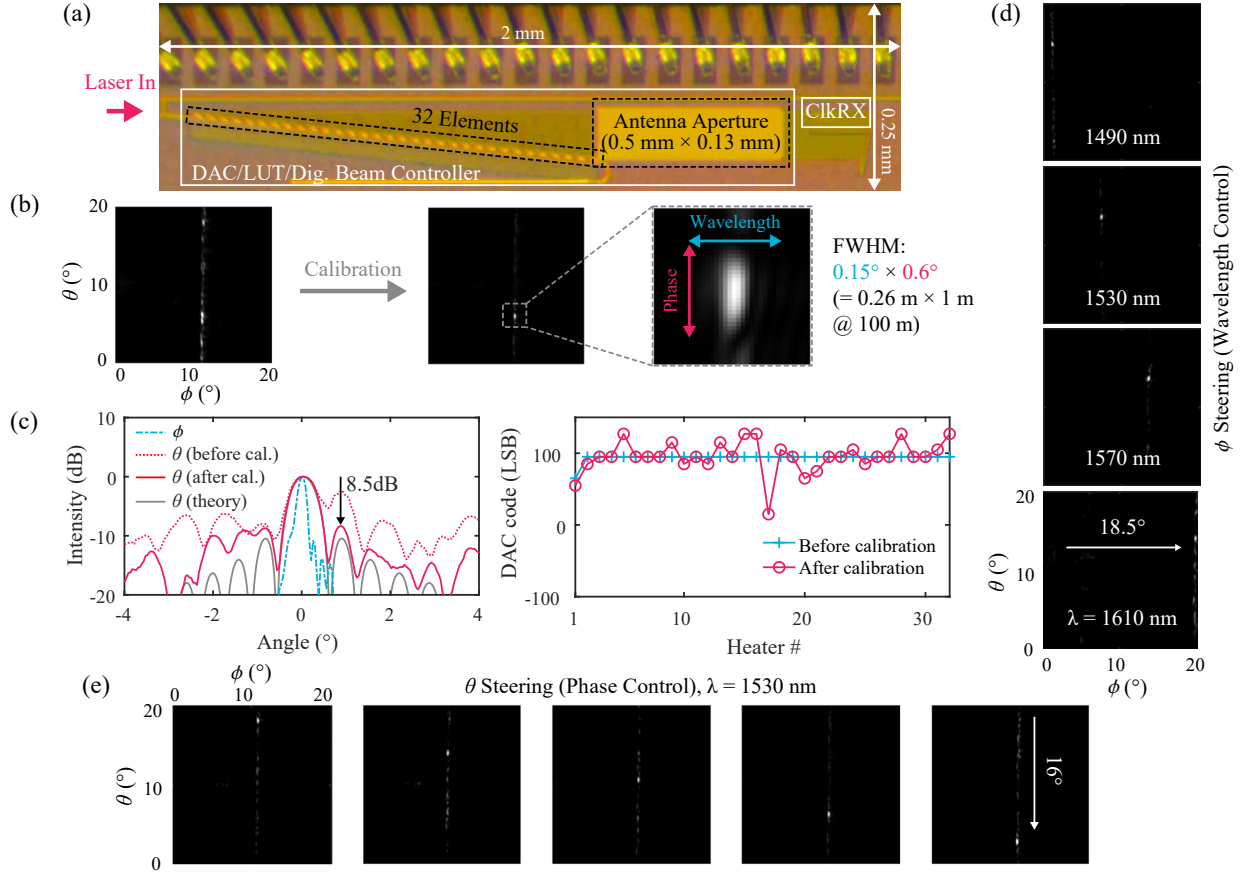
Figure 4.16: Demonstration of beam calibration performance and beam steering capability for the OPA with 32 elements. (a) Detailed die micrograph of the 32-element phased array. (b) Far-field image of the array, before and after calibration. (c) Cross-section of (b) along ϕ and θ, as well as the DAC code distributions. (d, e) Beam steering along ϕ and θ through laser wavelength and on-chip phase shifter control.

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° was demonstrated in [71]). A final beamwidth of $\text{FWHM}_\phi \times \text{FWHM}_\theta = 0.15° \times 0.6°$ is achieved, which corresponds to a lateral resolution of $0.26\,\text{m} \times 1\,\text{m}$ 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 4.17. The antenna aperture size of this large array demonstration is four times larger
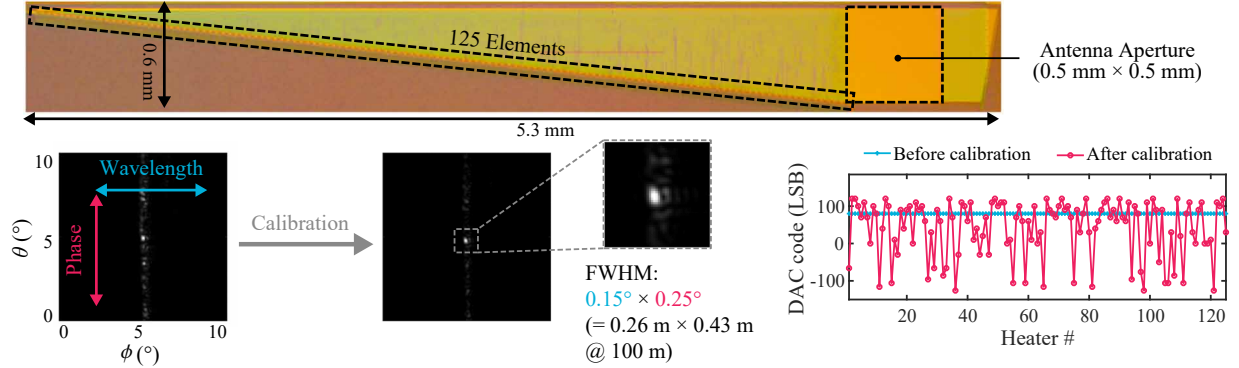
Figure 4.17: View of the larger 125 element array and calibrated beam performance.

(0.5 mm×0.5 mm), which results in a measured FWHM beamwidth after phase shifter calibration of 0.15° × 0.25° (matching the approximately 0.1° to 0.2° beamwidth requirements of long-range automotive LiDAR systems).

Table 4.1 summarizes our results and compares them with state-of-the-art OPA implementations designed for beam-steering. The first four columns of the table compare our results with the single-chip OPA tailored for wide-range beam-steering [57]. The next five columns present the latest multi-chip or photonics-only examples. Among single chip demonstrations, our work is the first that can do full 2D steering with reasonable range via phase and wavelength control and with high beam resolution in both dimensions.

4.5 Remaining Challenges

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, the basic requirements discussed in Section 4.1.2 tell only one side of the story; 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, as we discussed in Section 2.2. According to the industry standard [29], consistent operation across −40 °C ~ 105 °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

Table 4.1: Summary and Comparisons with Prior Beam-Steering Optical Phased Arrays

	This Work		[57]		[71]	[56]	[55]	[72]	[58]
	Small Variant	Large Variant	Small Variant	Large Variant					
Technology	Silicon Photonics/65nm CMOS		0.18 µm SOI CMOS		Silicon Photonics	Silicon Photonics	Silicon Photonics	Hybrid III-V/Si-Photonics	Silicon Photonics
Electronics	Integrated DAC		Integrated DAC		External	External	External	External	External CMOS
Beamwidth (ϕ/θ)	0.15°/0.6°	0.15°/0.25°	>10°/1.2°	>10°/0.03°	0.04°/0.04°	0.15°/0.15°	0.14°/0.14°	0.6°/1°	0.8°/0.8°
Aperture Size	500 µm× 130 µm	500 µm× 500 µm	3.55 µm× 64 µm	3.55 µm× 2046 µm	2000 µm× 2000 µm	1000 µm× 665 µm	910 µm× 910 µm	N/A	N/A
Steering Range (ϕ/θ)	18.5°/16°	18.5°/-	0°/45°	0°/-	15°/56°	14°/70°	17°/80°	3.6°/23°	16°/16°
Sidelobe Suppression	8.5 dB	7.4 dB	9 dB	9 dB	12 dB	8 dB	9 dB	5.5 dB	12 dB
Number of Elements	32	125	32	1024	512	512	128	32	128
Independent Phase Controls	32	125	36	136	512	512	128	32	128
Phase Shifter Type	Thermal		Thermal		Electro-Optic	Thermal	Thermal	Thermal/Electro-Optic	Thermal
Phase Efficiency	20 mW/π		34.7 mW/π		<2 µW/π	2.6 mW/π	80 mW/π	160 mW/π	10.6 mW/π
Required Voltage Supply	3.3 V		8 V		N/A	N/A	N/A	N/A	10 V

(a) 20 μm

(b)

(c)
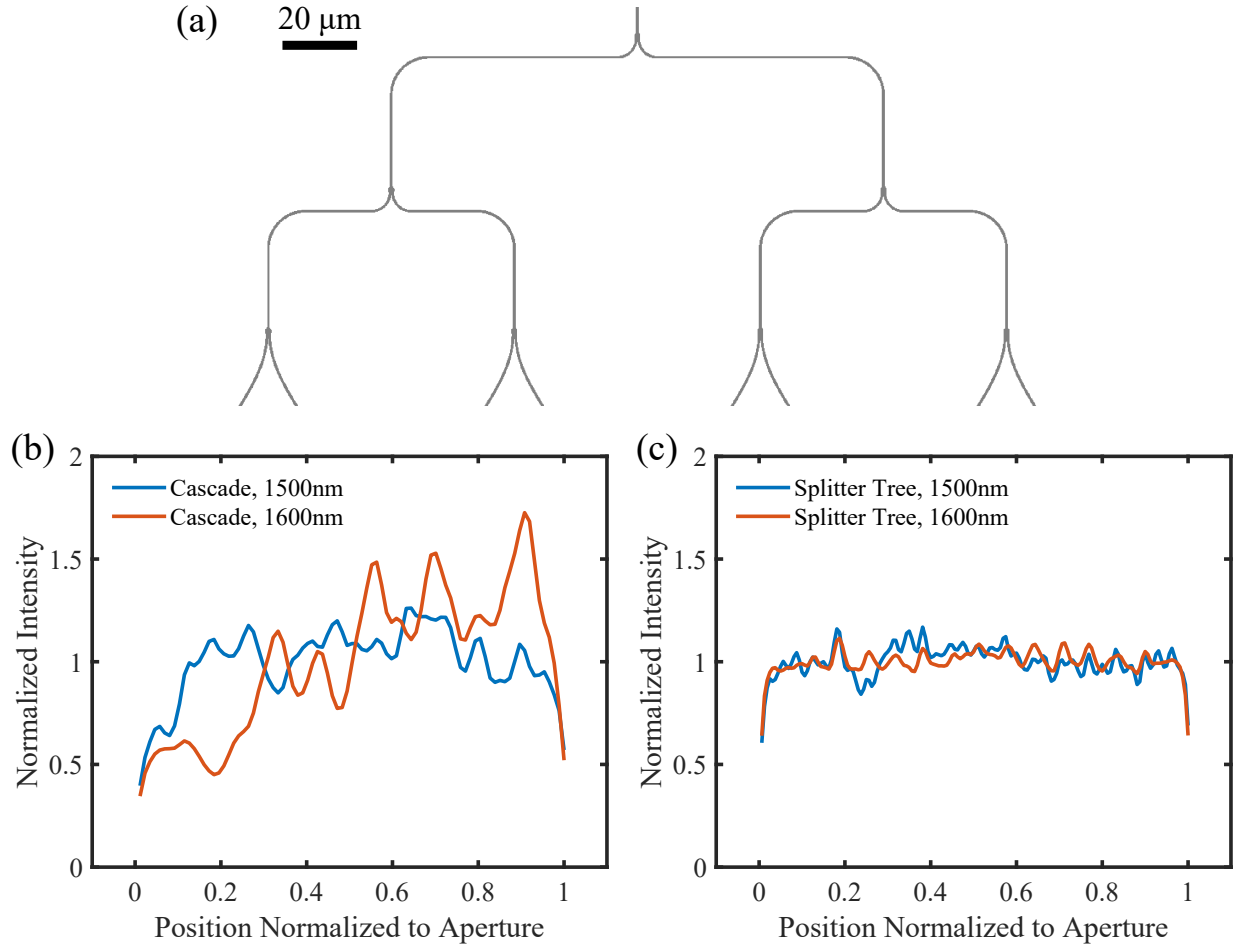
Figure 4.18: (a) The layout of 1×2 splitter-based optical distribution tree available in our process. (b) Measured power distribution across the cascaded directional coupler-based OPA at 1500 nm and 1600 nm wavelength. (c) Same data as (b), but from the OPA with the distribution using the design in (a).

antennas, 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 work or for the sub-row distribution in [57] can cause phase slope bias when temperature changes, which results in beam direction offset. Path mismatch can be primarily removed through 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
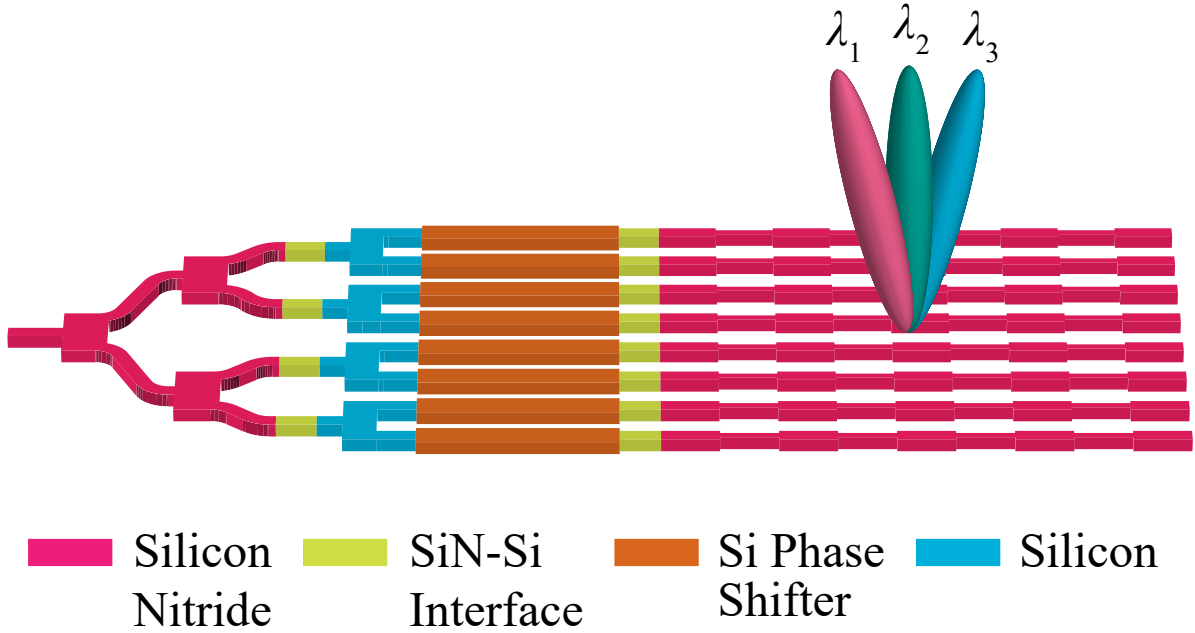
Figure 4.19: Proposed temperature-insensitive, high emission power OPA architecture realizable in our platform through process customization.

example, a tree of 1×2 splitters shown in Figure 4.18(a) [57], [73] 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 (Figure 4.18(b)).

From Equation 4.11, one can also notice 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 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 [73]).

We also discussed in Chapter 3 that to be used in long-range applications, it is also desired that the OPA can support the maximum permissible radiation power set by the laser safety regulation ($\sim 10\,\mathrm{dBm}$ for IEC60825-1 Class 1 in C-band [13]). In our prototype, the fiber-to-chip edge coupler has $\sim 3\,\mathrm{dB}$ loss, and the theoretical emission efficiency ($\eta = P_{\mathrm{beam}}/P_{\mathrm{input}}$) for the 2-dimensional array ($N = 125$, $d = 4\,\mu\mathrm{m}$, $d_{\mathrm{grating}} = 700\,\mathrm{nm}$, $l_{\mathrm{grating}} = 500\,\mu\mathrm{m}$) calculated from the basic array directivity equation in [74] and single antenna gain extracted from Figure 4.8(c) is $\sim -9\,\mathrm{dB}$. Considering $3\,\mathrm{dB/cm}$ waveguide loss and a 5-mm-long distribution network (Figure 4.17), the

overall insertion loss is expected to be at least 13.5 dB. Even with a reduced antenna spacing (e.g. $d = 2\,\mu m$, ~ 3 dB extra array gain) and unidirectional antenna design [69] (another ~ 3 dB gain) to improve the efficiency, ~ 20 dBm of optical power is needed at the OPA input to support ~ 10 dBm beam ignoring the extra loss from couplers and phase shifters. It is well known that silicon waveguides are unable to handle more than ~ 10 dBm without exhibiting significant loss due to two-photon absorption (TPA) [75]. SiN integration can also resolve this issue since SiN can handle much higher power density [73]. Figure 4.19 presents 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 work, a deployable system would eventually include an integrated laser that can support sufficient output power and wide wavelength tuning range (50 nm $\sim$ 100 nm) to enable 2D 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 [76]. Especially, III/V semiconductor lasers with large tuning range [77] and high output power [78] are promising because they can be heterogeneously integrated with silicon photonics, removing fiber-to-waveguide coupler loss. Alternatively, rare-earth-doped lasers [79] with high output powers over 300 mW [80] and wide tunability over 46 nm [81] can be utilized to further minimize the unit cost. Although it involves an additional optical pump, integration process is much simpler and done at the wafer scale.

Lastly, the simple local search-based calibration algorithm used in this work 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 [82], can be applied to potentially enable one-shot calibration. It must be noted that in the case of thermo-optic phase shifters, phase shifter crosstalk is another major source of disturbance. Since thermal isolation between heaters is never perfect in reality, adjusting one phase shifter will also affect the phase of adjacent shifters. As a result, the input phase pattern is spatially low-pass filtered. It is still possible to compensate for the crosstalk by using a spatially *pre-emphasized* input phase pattern. Note that implementation of such convolution necessitates per-element phase control access regardless of the OPA architecture.

4.6 Chapter Summary

In order for integrated optical phased arrays to be a reliable solid-state alternative to mechanical beam scanners, the technology must support $500 \sim 1000$ 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.

Chapter 5

Realization of Integrated Coherent LiDAR[1]

As briefly mentioned at the end of the previous chapter, with our wafer-scale 3D heterogeneous silicon photonics-CMOS integration technology, it is possible to implement not only the optical phased array but the entire eye-safe optical C-band LiDAR system on a single chip, at the minimum possible unit cost. In other words, it is an ideal platform to fulfill the vision of integrated beam-scanning coherent LiDAR for cars, which was illustrated in Section 2.3. In this chapter, I present our work on the integrated coherent LiDAR system as well as the experimental results from the prototype.

5.1 Integrated Coherent LiDAR Overview

An overview of the integrated coherent LiDAR system is shown in Figure 5.1. It combines two optical phased arrays (OPAs) for beam transmission/reception and a coherent optical front-end with CMOS electronics for FMCW LiDAR signal detection. External FMCW laser source is coupled into the input waveguide through an edge coupler and firstly split into the transmitter (TX) OPA and LO path. The TX OPA emits the laser beam into free space, which in turn hits the target and comes back to the receiver (RX) OPA. Finally, the received laser from the RX OPA is mixed with the LO laser and detected at the coherent receiver. The receiver includes the electrical ana-

---

[1]The content of this chapter was derived from [16], a work 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 used in this work. Pavan Bhargava at UC Berkeley contributed to the design of analog CMOS circuits. Wafer fabrication and 3D integration was done by the CNSE.
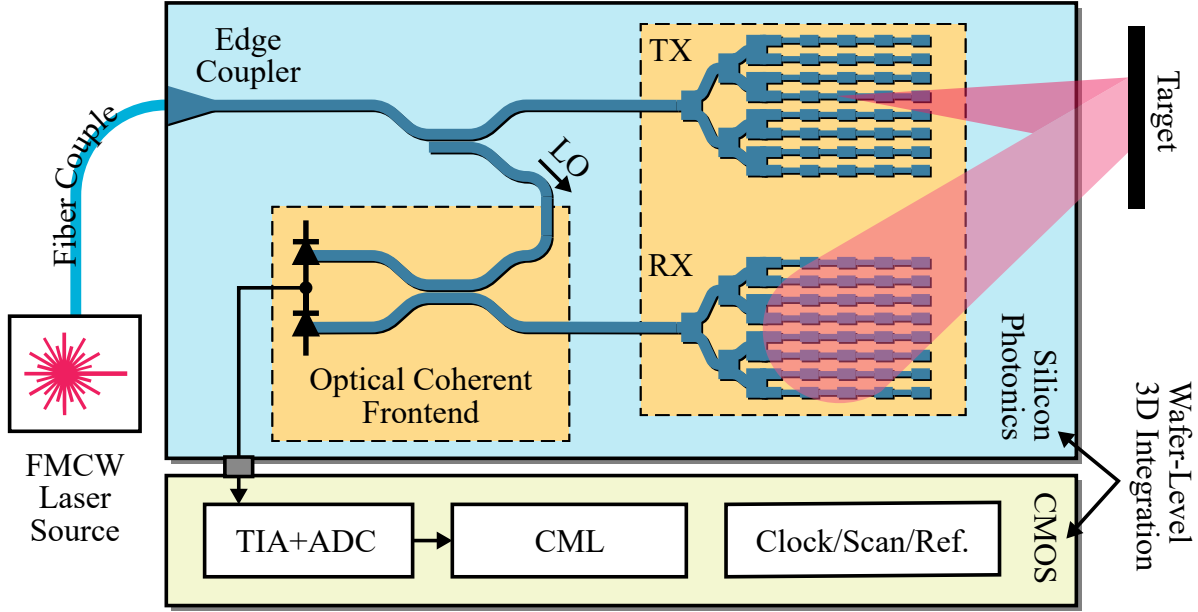
Figure 5.1: Integrated coherent LiDAR system overview.

log frontend (TIA and ΣΔ ADC), and the output of the ADC is directly forwarded to the off-chip via CML interface. CMOS layer also includes clock distribution logic, reference circuits and the scan interface to configure the circuit.

5.2 Optical Phased Array for Beam Formation

Figure 5.2 shows the die micrograph of the 128-element passive OPA used in our system. The power from the input waveguide is evenly distributed into each array element through a 1×2 multi-mode interference (MMI) coupler tree. A 500 μm-long grating with apodized perturbations described in Section 4.3.1 was used as our antenna element. 128 elements are placed at 2 μm pitch, resulting in an array aperture size of 500 μm × 254 μm. Tighter pitch of the antenna elements in this OPA compared to the standalone OPA prototype in Chapter 4 enables higher array gain and rejection of interfering beams through spatial filtering in both transmit/receive arrays due to superior fill factor and grating lobe suppression.

The far-field imaging setup and beam pattern of our OPA is shown in Figure 5.3. A beam full-width at half-maximum (FWHM) of 0.15°/0.3° was achieved, equivalent to 26 cm/52 cm diffraction-limited resolution at 100m range. Our grating design can support ~ 18.5° vertical steering range via laser wavelength sweep over 120 nm, and full two-dimensional wide-range beamsteering can
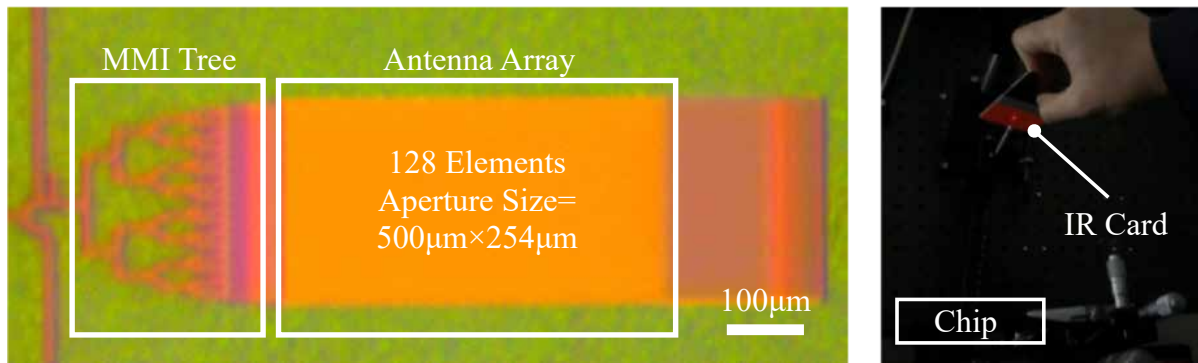
Figure 5.2: Die micrograph of the OPA and output beam captured on an IR card.
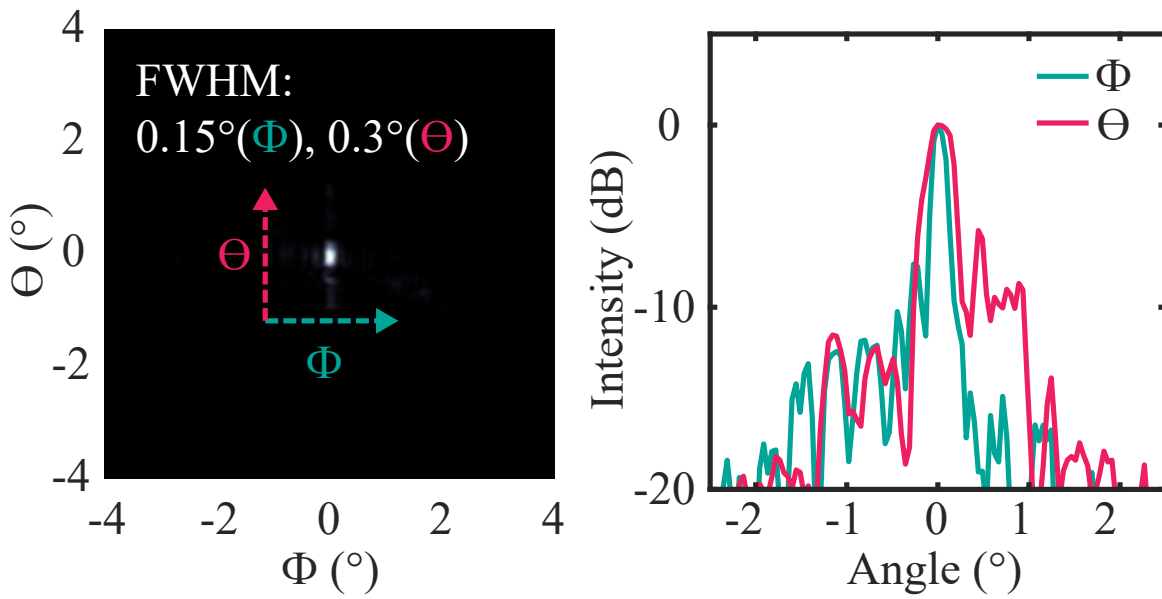
Figure 5.3: Measured OPA beam pattern and cross-section.
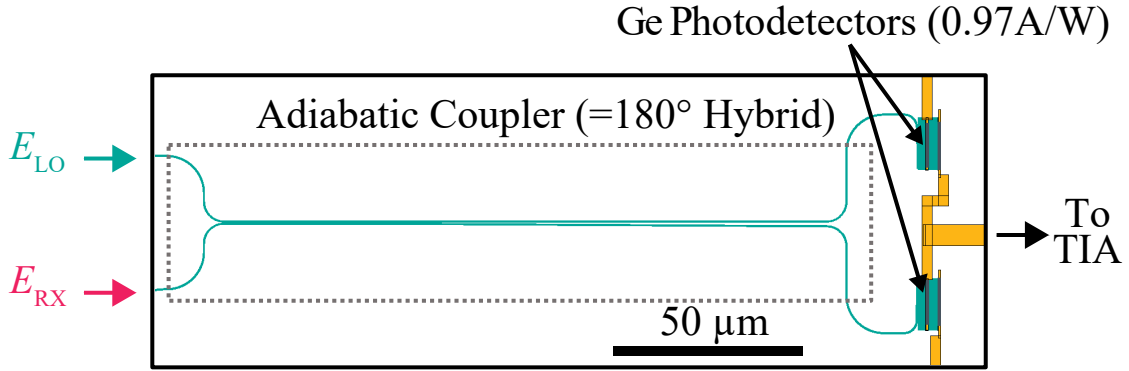
Figure 5.4: Layout of the optical coherent detection frontend in EPHI platform.

be achieved with the addition of element-wise phase shifters as shown in Chapter 4.

5.3   Integrated Optical Coherent Detection

Figure 5.4 shows the layout of the optical mixer circuit implementation described in Section 3.3. 2×2 coupler is realized using an adiabatic coupler which can maintain consistent coupling ratio over wider wavelength range. The outputs of the adiabatic coupler are then connected to a balanced detector comprising two $4\,\mu\text{m} \times 12\,\mu\text{m}$ mode-evolution Germanium-on-Silicon photodetectors [83] to capture the power envelope of the mixed fields and output the difference as the photocurrent. The measured responsivity of the photodetector at 1550 nm is $\sim 1\,\text{A/W}$.

Figure 5.5 shows the architecture of our integrated coherent LiDAR receiver. The output of the balanced detector is converted to voltage using an inverter-based pseudo-differential transimpedence amplifier (TIA) with the gain of 150 kΩ, and then digitized by the first-order ΣΔ analog-to-digital converter (ADC), which includes a switched-capacitor integrator, and a preamplifier followed by a quantizer. ΣΔ feedback was implemented with a current DAC connected to the TIA input. Note that ΣΔ modulation can be deactivated if the interfering tones are not present.

As explained in Section 3.3, the sum of the input-referred electrical noise including thermal noise and ADC quantization noise should be dominated by shot noise originated from the LO laser to ensure quantum-limit SNR, within the bandwidth of interest given by Equation 3.20. Input-referred thermal noise of the analog frontend is estimated from Equation 3.32, and the
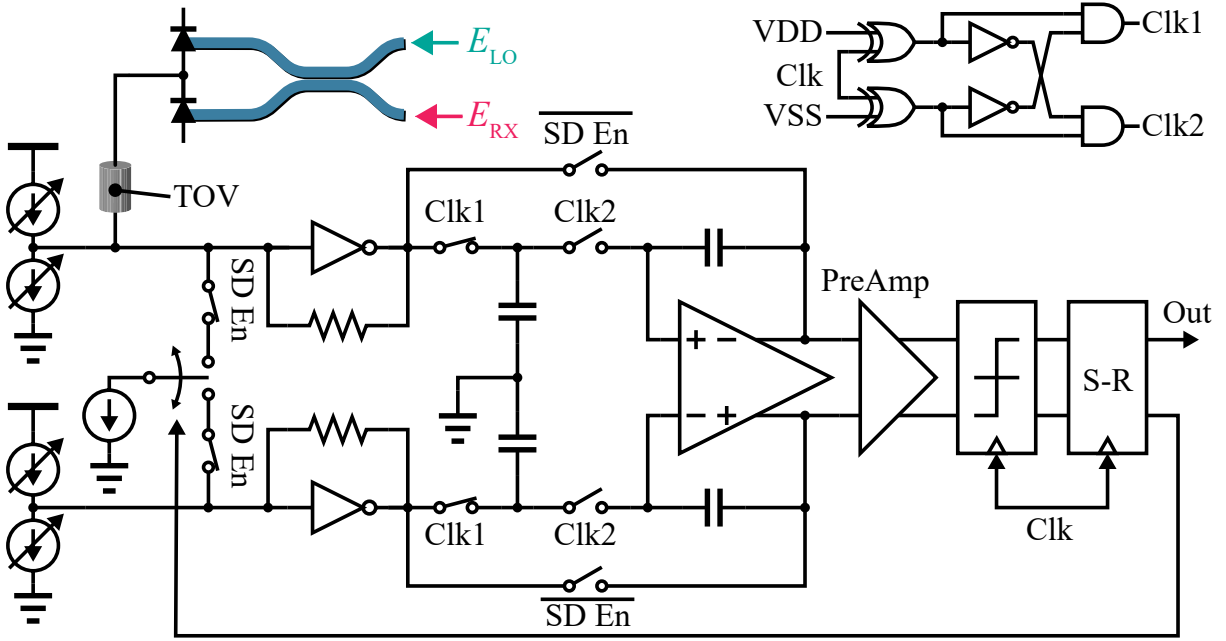
Figure 5.5: Coherent LiDAR receiver architecture.

input-referred quantization noise from 1st-order $\Sigma\Delta$ ADC is expressed as follows:

$$S_{i_\text{n},\text{in},Q}(\omega) \approx \frac{\Delta^2}{6} \frac{1}{f_\text{s}} \left(\frac{\omega}{R_\text{TIA} f_\text{s}} \right)^2 . \tag{5.1}$$

Note that the quantization noise density becomes the highest for the maximum range measurement. Then, following conditions should be met to guarantee shot-noise limited performance:

$$S_{i_\text{n,PD}}(\omega_\text{max}) \geq S_{i_\text{n},\text{in},Q}(\omega_\text{max}) + S_{i_\text{n},\text{in},\text{therm}}(\omega_\text{max}), \tag{5.2}$$

$$2q R_\text{PD} P_\text{LO} \geq \frac{\Delta^2}{6} \frac{1}{f_\text{s}} \left(\frac{\omega_\text{max}}{R_\text{TIA} f_\text{s}} \right)^2 + \frac{8 k_\text{B} T}{R_\text{FB}} + \frac{8 k_\text{B} T \gamma_\text{MOS}}{g_\text{m} R_\text{FB}^2}, \quad \omega_\text{max} = \frac{4\pi R_\text{max}}{c} \gamma. \tag{5.3}$$

In short, the main change is the factor $(\omega_\text{max}/f_\text{s})$ from noise shaping instead of $\sim 1/(2^N - 1)$ in Equation 3.29. For example, to maintain the same LO power level as 8-bit successive-approximation-register (SAR) ADC case, used as an example in Section 3.5, $f_\text{max}/f_\text{s} = (2R_\text{max}\gamma/c)/f_\text{s}$ should be ~ 40.

5.4 Coherent LiDAR Ranging Demonstration

Figure 5.6 shows the fabricated device and die micrograph of the integrated coherent LiDAR. The system is implemented using the same 3D integration technology described in Section 4.2.
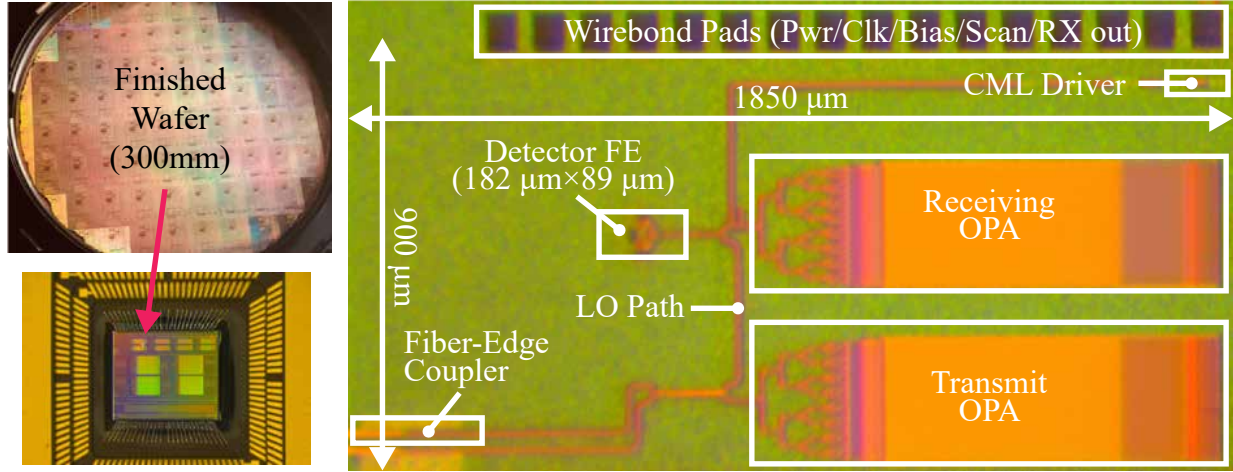
Figure 5.6: Completed 300mm wafer and packaged device, along with the die micrograph of the integrated LiDAR system.

We firstly characterized the LiDAR receiver using the setup shown in Figure Figure 5.7. An external-cavity laser centered at 1550nm was used as our FMCW source, where injection current modulation was performed with a triangular wave from an arbitrary waveform generator (chirp rate: 9 THz/s). Free-space propagation delay was emulated by a 30 m-long single-mode fiber. The output of the ADC was forwarded off-chip via CML driver and directly monitored with an oscilloscope, along with the output from a separate off-the-shelf detector as a reference. The resulting signal in the spectral domain is shown in Figure 5.8 (f_{clock}=250 MHz, P_{LO}=2 dBm, P_{RX}=−30 dBm). Signal length (T_{meas}) was 50 μs. We observed that the noise floor is dominated by the shot noise up to ∼ 20 MHz, which is equivalent to 333 m range for given chirp rate. The required LO power was rather high and mostly limited by the input-referred quantization noise. Considering that the signal bandwidth was much lower than what is supported by the current TIA, LO power can go down in the future by using higher TIA gain, and also by increasing ADC sampling frequency.

Figure 5.9 shows the setup used for the ranging demonstration using prototype LiDAR. The LiDAR measurement was done using a target (made with reflective tape) placed at various positions on a rail. The same FMCW source used for the receiver testing was also used for this demonstration. The output waveforms corresponding to three different target locations are shown in Figure 5.10, and the dependence of the signal frequency on the target distance is clearly seen. The absence of interfering tones within the receiver signal confirms that the main transmit beam was well confined within the target of interest with no back-reflections from grating lobes. The ranging resolution was limited by the finite sample length (0.5 ms) or FFT bin size (2 kHz) as

mentioned in Equation 3.15, which is equivalent to 3 cm ranging resolution.
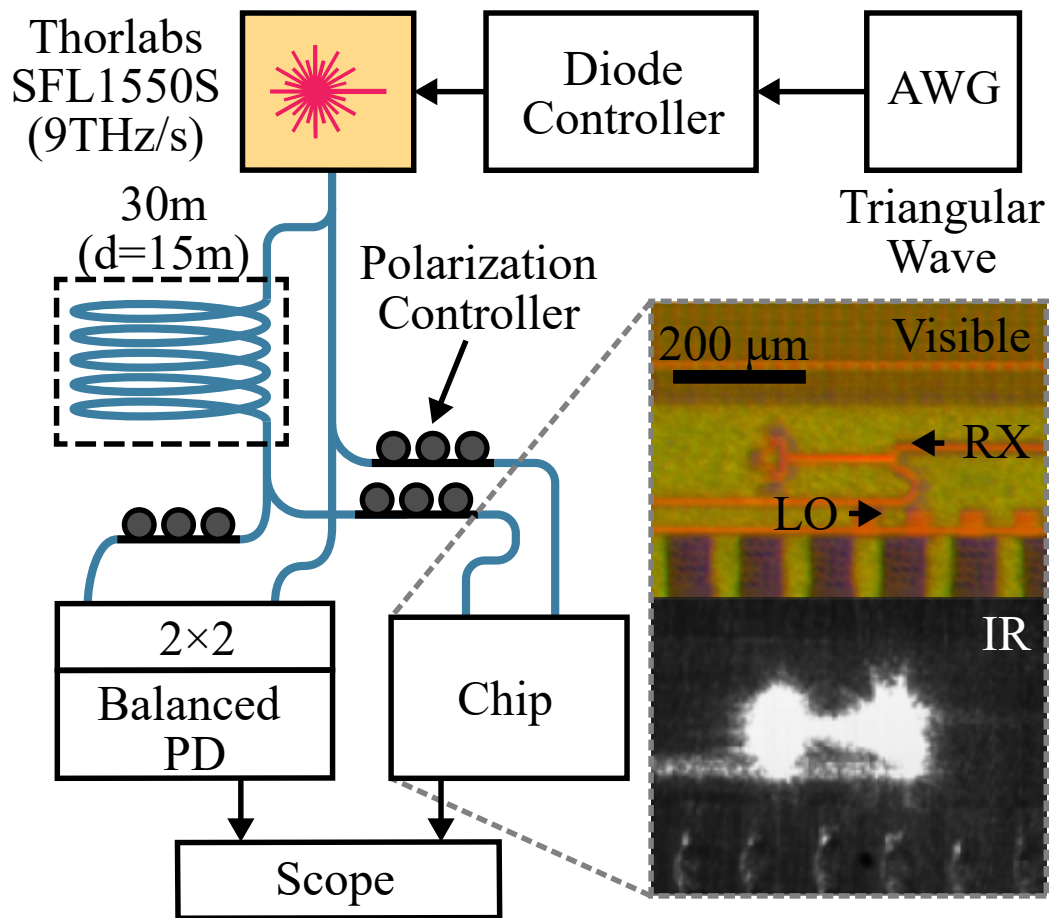
Figure 5.7: Receiver characterization setup with emulated LiDAR measurement.
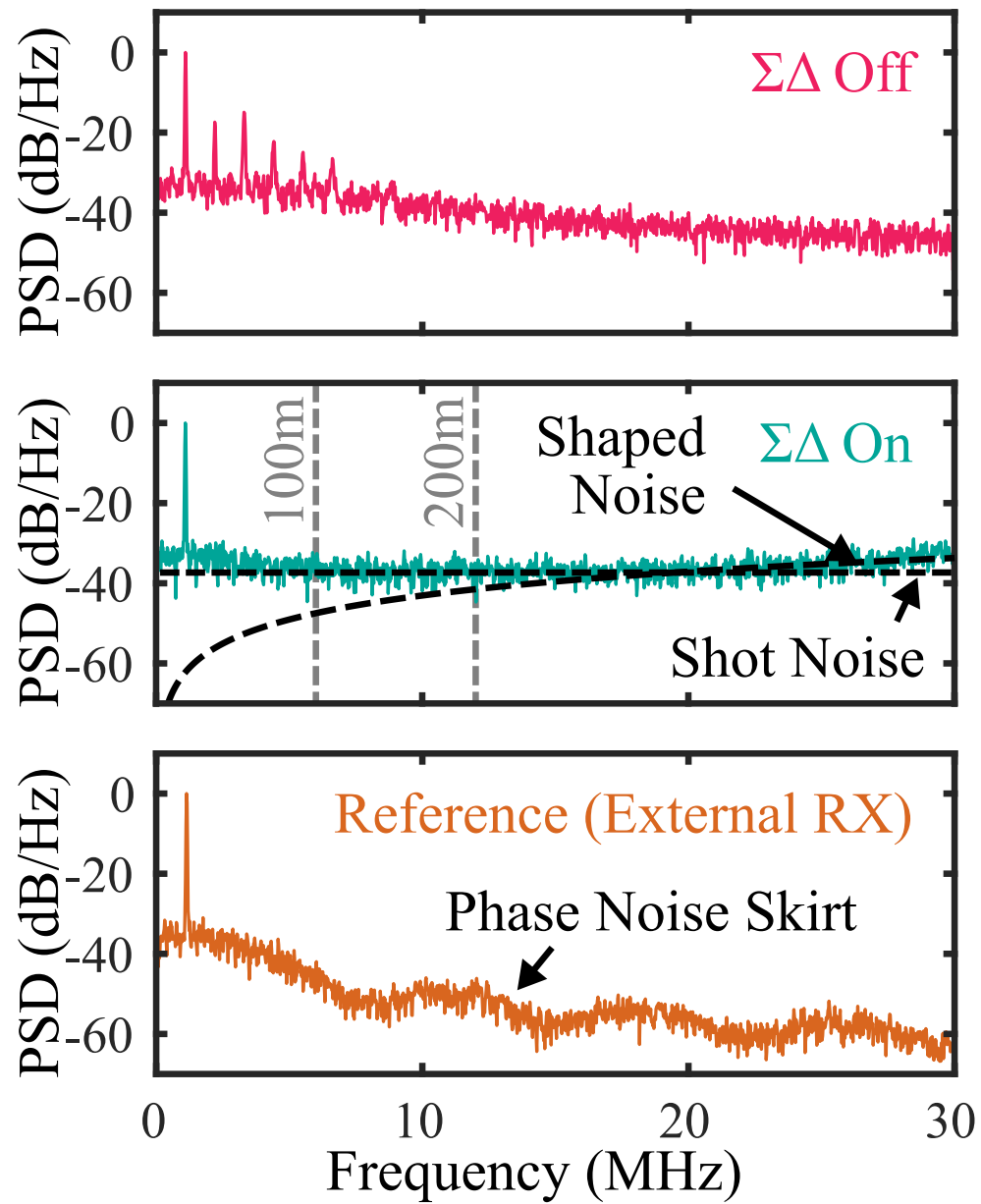
Figure 5.8: Measured spectrum of the integrated coherent receiver output with $\Sigma\Delta$ on/off, compared with off-the-shelf coherent receiver as a reference.
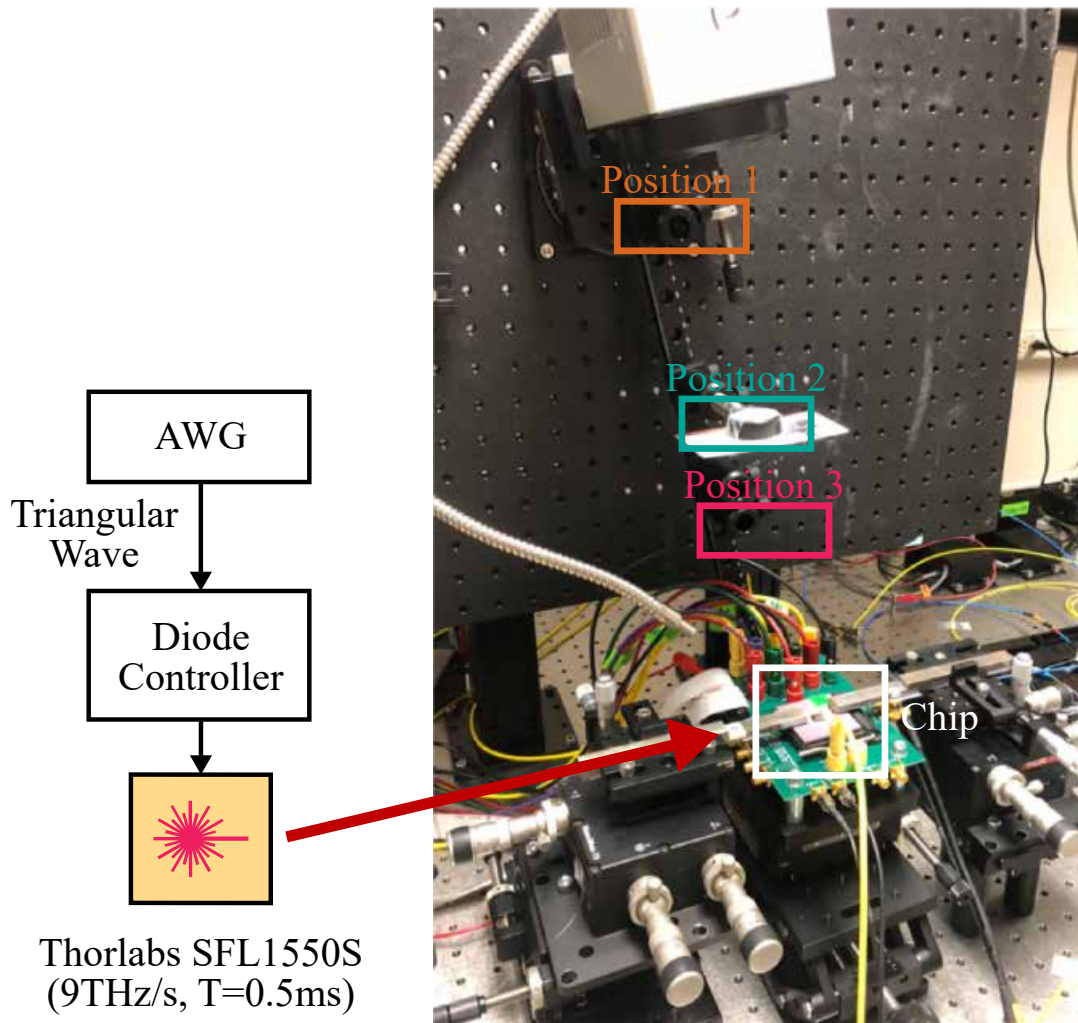
Figure 5.9: Integrated FMCW LiDAR ranging demonstration setup.
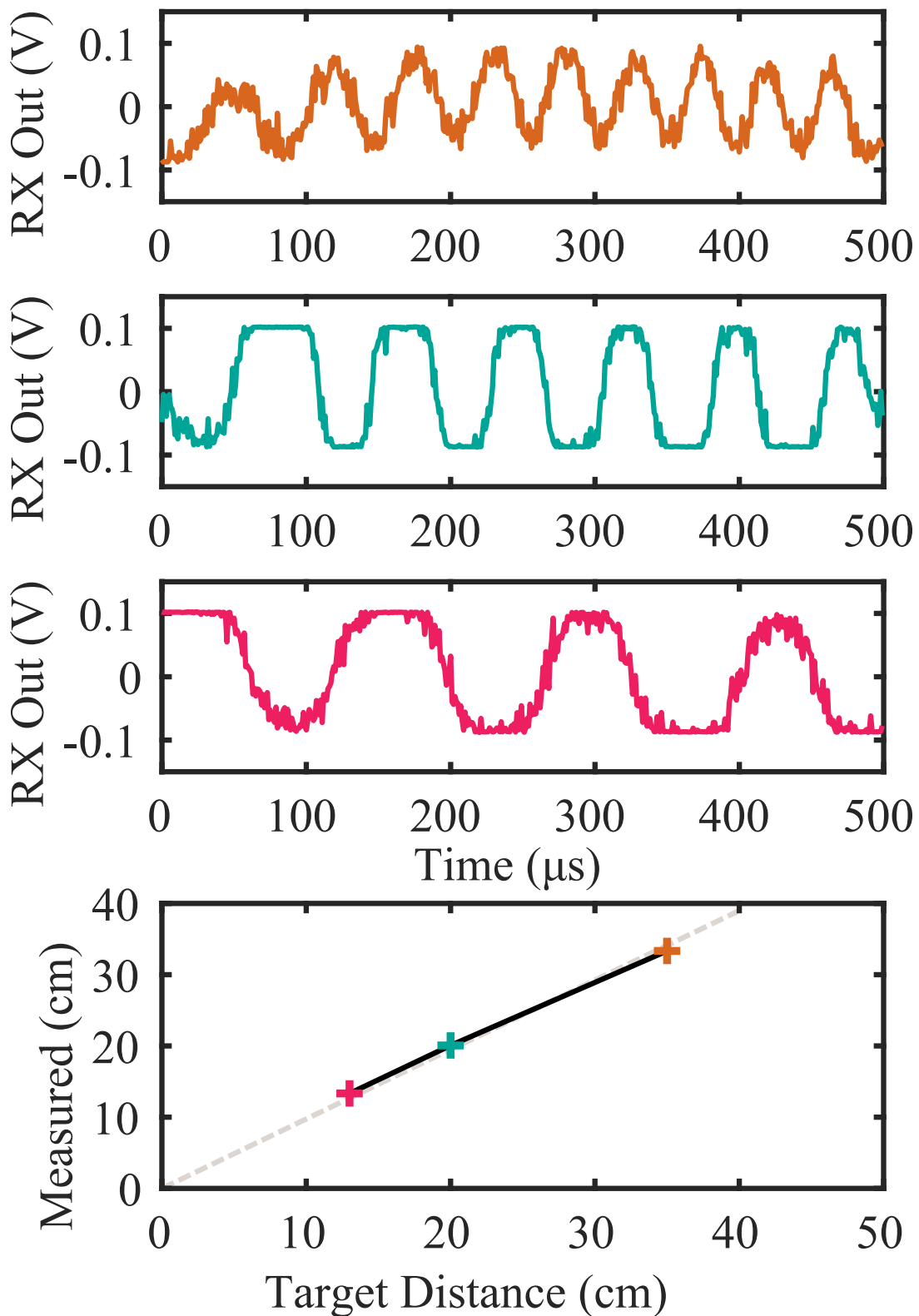
Figure 5.10: Integrated FMCW LiDAR ranging demonstration result. Coherent receiver output signal corresponding to three different object positions, showing linear relationship between the object distance and the receiver signal frequency.

5.5 Future Directions

Now that we have shown that it is indeed possible to build a fully integrated coherent LiDAR with the state-of-the-art silicon photonics / CMOS technology, we can go back to the discussion in Chapter 2 and Chapter 3 and see what is the current gap between our prototype and the ultimate system that satisfies the requirements in Table 2.1. As the remaining issues in OPA-based beam scanning were discussed extensively in Section 4.5, let's focus on the rest of the system.

The most obvious issue is the size of the receiving aperture. In our prototype, we simply used a copy of transmitting OPA as the receiving aperture, and its physical size was only one eighth of $1\,\text{mm}^2$. What is worse, the effective fill factor of the antenna aperture is roughly 25%, deduced from the array emission efficiency ($\sim -6\,\text{dB}$) and assuming reciprocity. Resulting effective aperture size is more than $3 \sim 4$ orders of magnitude smaller than $> 1\,\text{cm}^2$ aperture size suggested in Section 3.5. In fact, the effective aperture size of some of the largest reported silicon photonics-based OPAs (Table 4.1) are still only a few $1\,\text{mm}^2$, and even with single-chip integration technology, it is unlikely that single OPA is going to scale up by more than two orders.

One way of increasing the size of the receiving aperture is to tile N receiving OPAs, each including its own coherent receiver, and averaging N output signals in the spectral domain. This is in fact equivalent to increasing the measurement length T_{meas} by N. While this is indeed an attractive solution, especially considering low unit cost of silicon-based OPAs, scaling this to > 100 chips is still a serious challenge because each receiving OPA+receiver chip requires individual fiber connection to bring in the LO laser, and the packaging complexity and associated cost rapidly increases.

Alternatively, we can use traditional imaging lens/detector array-like receiver architecture, where the detector array is replaced with optical antennas [84]. The key advantage of such imager-type receiver is the fact that the size of the receiving aperture (i.e. entrance pupil of the imaging system) is no longer directly coupled with chip-scale dimensions, and can be easily extended to $> 1\,\text{cm}^2$ scale. On the other hand, achieving large field of view can be challenging because lens aberrations such as field curvature can significantly degrade the efficiency of the pixels associated with off-axis antennas [49]. Co-design of the lens system and lens-to-chip interface (e.g. grating antenna or end-fire edge coupler shown in Figure 5.11) can potentially mitigate field curvature issue.

Figure 5.12 shows a possible receiver system architecture that can be combined with OPA-based beam scanner illustrated in Figure 4.19. Here we assume that multi-wavelength tunable laser is coupled into the transmitting OPA and illuminates multiple vertical pixels simultaneously. The imaging lens collects the backscattered light from the target, and subsequently activates the
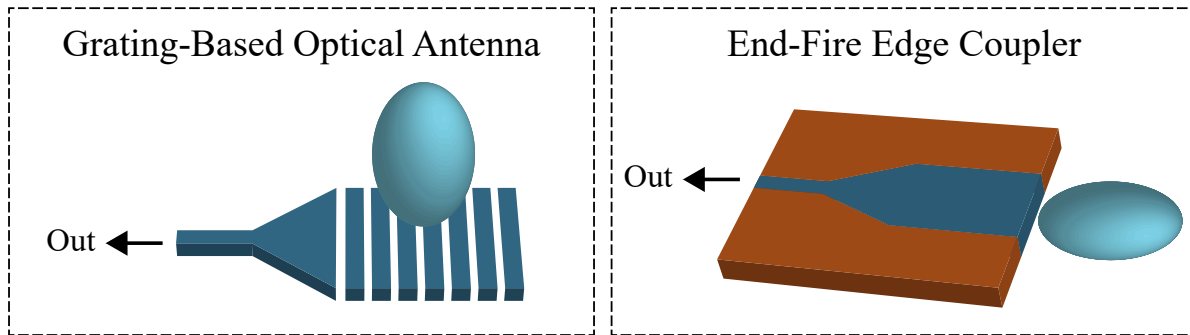
Figure 5.11: Examples of optical antenna for a lens-to-chip interface.

pixels associated with the direction addressed by the transmitting OPA. To alleviate the chip complexity by reducing the number of coherent receivers while minimizing the insertion loss faced by backscattered light, optical switch networks based on thermo-optic optical multiplexer/demultiplexer circuits can make a connection between the LO input and the active pixels at the dedicated coherent receiver. Multiple pixels on the vertical axis are addressed simultaneously by having one mux/demux structure per each wavelength.

Another problem which has not been discussed so far is depolarization [23]. It is possible for the polarization state of the transmitted beam (e.g. aligned with on-chip waveguide linear TE polarization at the transmit OPA aperture) to be altered during backscattering or propagation through the atmosphere comprising inhomogeneous particles. Namely, the receiver should be insensitive to return polarization state to maintain overall SNR. This can potentially be achieved by utilizing polarization insensitive optical antennas [85] or tiling two receiver chips with orthogonal orientation and taking the average of two output spectrums. The latter can actually enable another imaging mode based on depolarization ratio [23], which can uncover useful information about the atmospheric composition or the surface characteristics.

5.6 Chapter Summary

In this chapter, the demonstration of the first coherent solid-state LiDAR system integrated on a single chip fabricated in a 300 mm wafer facility is presented, proving that high-resolution LiDAR systems can indeed be realized at low cost, leveraging state-of-the-art silicon photonics-CMOS integration technology. 3D integration allows the photonics to be highly customized independent of electronics, and further system cost reduction can also be achieved through hybrid laser/optical
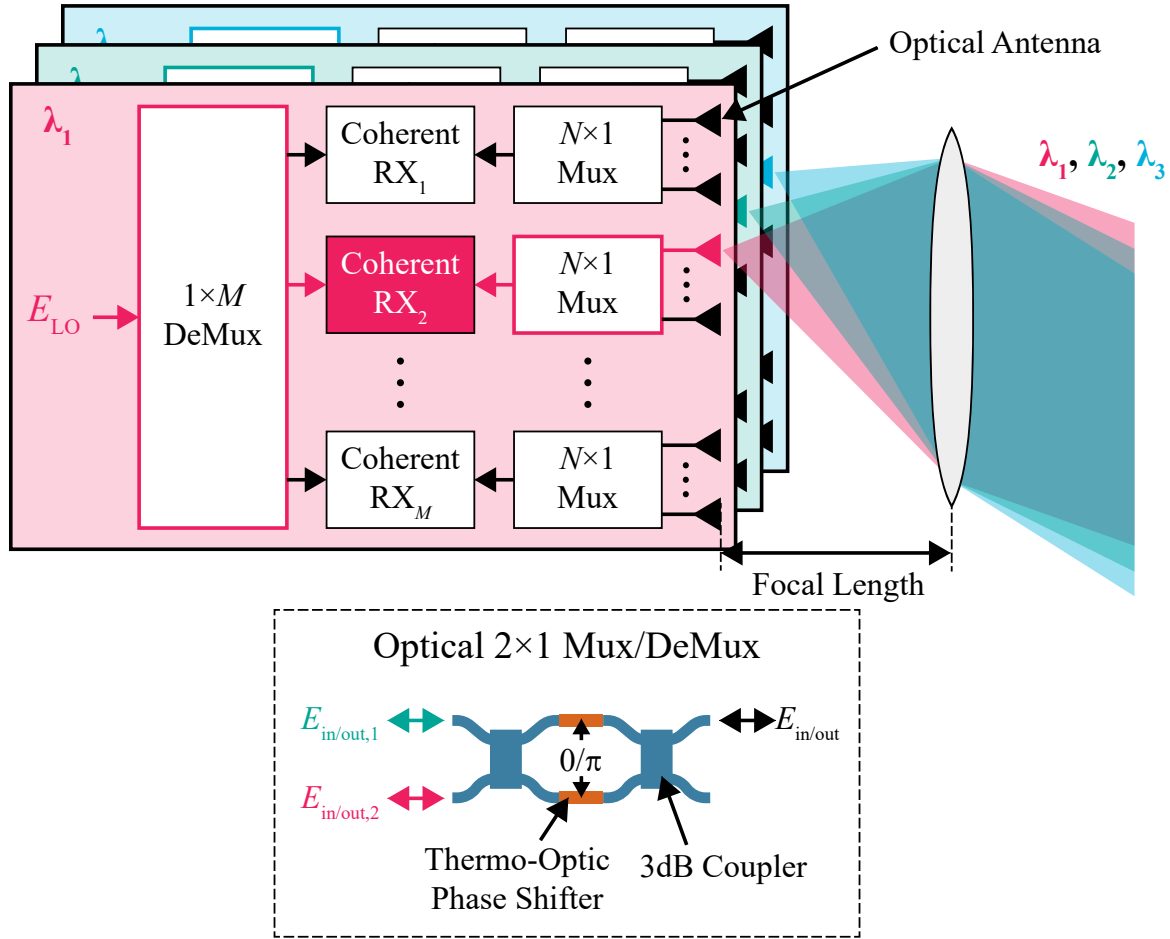
Figure 5.12: Possible large-aperture coherent LiDAR receiver architecture based on optical antenna array placed at the focal plane of an imaging lens. Input LO laser and backscattered light from the active antenna *pixel* are combined at the dedicated coherent receiver through integrated optical multiplexer/demultiplexer. Assuming grating-based OPA (Section 4.3.1) at the transmitter, 2D imaging is enabled by having dedicated mux/demux tree per each wavelength.

gain integration [79], [81]). I also point out that to eventually address long-range applications, it is necessary to extend the receiver-side aperture to centimeter-scale. Centimeter-scale aperture may be achieved with an imager-like receiver architecture through discrete optics and on-chip photonics co-design, where each imaging pixel is replaced by an optical antenna.

Chapter 6

Overcoming Coherence Distance Limit in FMCW LiDAR

As mentioned previously, the depth information in a FMCW LiDAR is captured by the frequency of the beating tone at the coherent receiver. As discussed briefly in Section 3.2, this implies that the quality of the received signal is also a direct function of the spectral purity of the laser source. A popular metric to quantify the phase noise performance of a laser in the context of interferometric measurement system with coherent detection is coherence time/distance, which is a measure of how much time/path mismatch you should introduce to two laser beams, both of which are originated from the same source, until they lose temporal coherence. Formally speaking, the coherence time τ_{coh} and the coherence distance d_{coh} are defined as the following:

$$\tau_{\mathrm{coh}} = \frac{1}{\pi \Delta \nu}, \quad d_{\mathrm{coh}} = \frac{c}{\pi \Delta \nu}. \tag{6.1}$$

$\Delta \nu$ is the laser linewidth, which is defined as the 3 dB bandwidth around the nominal laser frequency. In the case of FMCW LiDAR, coherence *range* is one half of the coherence distance to take into account the round trip (Equation 3.24). For example, the coherence range of a laser with 150 kHz linewidth is ~ 318 m, which is close to the desired maximum detection range of an automotive LiDAR (Table 2.1).

The presence of laser phase noise certainly degrades overall system performance, but the actual way how degradation happens can vary depending on the nature of the system. Simply speaking, depth measurement systems based on FMCW modulation can be classified into two categories:

- Long-range (> 100 m), moderate ranging precision (~ 1 cm), SNR-limited systems: Automotive LiDAR belongs to this category. As illustrated in Equation 3.24, laser phase noise

causes the signal peak in the PSD domain, which determines if the LiDAR sensitivity requirement is met (Equation 3.27), to go down by the factor of $e^{-4\pi\Delta\nu R_{max}/c}$ at the maximum range, even within the coherence range defined in Equation 6.1.

- High ranging precision ($< 10\,\mu m \sim 100\,\mu m$), short~moderate range ($< 1\,m$): Applications such as microimagers [86] or optical coherent tomography (OCT) require high-resolution imaging of relatively static objects. As the measurement is done at relatively short range and in a controlled environment, the SNR of the measurement is often well above the threshold discussed in Section 3.1. On the other hand, to achieve high ranging resolution, desired continuous tuning bandwidth of the laser is very wide (Equation 3.15) and often goes beyond 100 GHz [86]. In this case, the phase noise degrades the performance by causing the measured signal to deviate from the signal model commonly assumed in the spectral estimation algorithms that does not take into account the phase noise term [87].

In this chapter, we formulate the problem of laser phase noise in FMCW LiDAR, and discuss possible solutions to improve the system performance in both long-range and high-resolution applications. Considering stringent tunable laser requirements in terms of tuning range, tuning speed, output power, and wall-plug efficiency, relaxing linewidth requirement can dramatically expand the range of *usable* lasers and eventually reduce the overall cost of the system. In fact, how phase noise actually makes the performance worse also hints at the way to make improvements. For the case of long-range LiDARs, the phase noise level of the FMCW source should physically go down to mitigate the phase noise-induced SNR degradation, and this calls for a circuit-side solution which can filter out the phase noise. For high-resolution LiDAR, new optimal spectral estimation algorithm, which can take into account the phase noise in its signal model, is needed.

6.1 FMCW Measurement in the Presence of Phase Noise

To illustrate how the laser phase noise affects the signal at the output of the optical coherent detection frontend, we can go back to Equation 3.5 and this time keep the noise term.

$$\phi_{I_{RX}}(t;\tau) = \int_{t-\tau}^{t} 2\pi f(u)du + \phi_n(t) - \phi_n(t-\tau) = \omega_b t + \phi_0 + \phi_{n,I_{RX}}(t), \quad \omega_{I_{RX}} = 2\pi\gamma\tau \qquad (6.2)$$

$$\phi_{n,I_{RX}}(t;\tau) = \phi_n(t) - \phi_n(t-\tau). \qquad (6.3)$$

As mentioned in Section 3.2, the phase noise process of the coherent receiver output signal is the difference between the laser phase noise process and another copy of the laser phase noise but delayed by τ (i.e. $\phi_n(t)$ vs. $\phi_n(t-\tau)$). Two processes are perfectly correlated and cancel each

other when τ is zero. They become less correlated as τ, or the target distance R, increases, and eventually become completely incoherent.

We can also express the coherent receiver output phase noise in the spectral domain as follows:

$$\Phi_{n,I_{RX}}(\omega; \tau) = \Phi_n(\omega) - e^{-j\omega\tau}\Phi_n(\omega) \tag{6.4}$$

$$= \Phi_n(\omega)e^{-j\omega\tau/2}(e^{j\omega\tau/2} - e^{-j\omega\tau/2}) \tag{6.5}$$

$$= (2j)\Phi_n(\omega)e^{-j\omega\tau/2}\sin(\omega\tau/2). \tag{6.6}$$

This implies that the power spectral density of the receiver output phase noise and the laser phase noise has the simple relationship as the following:

$$S_{\phi_{n,I_{RX}}}(\omega; \tau) = S_{\phi_n}(\omega)\left|2\sin\left(\frac{\omega\tau}{2}\right)\right|^2. \tag{6.7}$$

Namely, the phase noise spectrum of the receiver output is the same as the phase noise spectrum of the laser filtered by a sine-square frequency response of period $1/\tau$. Note that following approximation can be made at either the low-frequency or the high-frequency range with respect to $\omega = 1/\tau$:

$$S_{\phi_{n,I_{RX}}}(\omega; \tau) \approx \begin{cases} |\omega\tau|^2 S_{\phi_n}(\omega), & \omega \ll 1/\tau \\ 2S_{\phi_n}(\omega), & \omega \gg 1/\tau \end{cases} \tag{6.8}$$

In other words, FMCW measurement, or self-interferometry, can be roughly interpreted as a first-order high pass filter in the phase domain with the passband gain of $\sqrt{2}$ and the cutoff frequency at $1/(\sqrt{2}\pi\tau)$.

To precisely examine the impact of the phase noise on the signal to noise ratio, we need to derive the expression for the spectral density of the coherent receiver output current signal ($S_{I_{RX}}(\omega)$) in the presence of phase noise given by Equation 6.6. Note that we can express the output current as the following:

$$I_{RX}(t; \tau) = A_{I_{RX}}\cos\left(\omega_{I_{RX}}t + \phi_{n,I_{RX}}(t; \tau)\right) = A_{I_{RX}}\cos\left(\omega_{I_{RX}}t + \phi_n(t) - \phi_n(t - \tau)\right). \tag{6.9}$$

Phase offset ϕ_0 is omitted as it does not affect the spectral density. Ultimately, we want to find out a way to transfer the noise in the phase domain to the current domain through cosine function. First, let's remind that the power spectral density (PSD) of a signal is in a Fourier relationship with the autocorrelation function (Wiener-Khinchin theorem). Therefore, we can derive the expression for the autocorrelation of the signal first and then take the Fourier transform to find the PSD. Assuming that the phase noise is ergodic and stationary, the autocorrelation of I_{RX} is given

as follows:

$$R_{I_{RX}}(u; \tau) = \langle I_{RX}(t; \tau) I_{RX}(t - u; \tau) \rangle_t \tag{6.10}$$

$$= A_{I_{RX}}^2 \langle \cos\left(\omega_{I_{RX}} t + \phi_{n,I_{RX}}(t; \tau)\right) \cos\left(\omega_{I_{RX}}(t - u) + \phi_{n,I_{RX}}(t - u; \tau)\right) \rangle_t \tag{6.11}$$

$$= \frac{A_{I_{RX}}^2}{2} \left[\langle \cos\left(\omega_{I_{RX}} u + \Delta\phi_{n,I_{RX}}(t; u, \tau)\right) \rangle_t + \langle \cos\left(\omega_{I_{RX}}(2t - u) + \phi_{n,I_{RX}}(t; \tau) + \phi_{n,I_{RX}}(t - u; \tau)\right) \rangle_t \right],$$

$$\Delta\phi_{n,I_{RX}}(t; u, \tau) = \phi_{n,I_{RX}}(t; \tau) - \phi_{n,I_{RX}}(t - u; \tau)$$

$$\tag{6.12}$$

$$= \frac{A_{I_{RX}}^2}{2} \cos\left(\omega_{I_{RX}} u\right) \langle \cos\left(\Delta\phi_{n,I_{RX}}(t; u, \tau)\right) \rangle_t - \frac{A_{I_{RX}}^2}{2} \sin\left(\omega_{I_{RX}} u\right) \langle \sin\left(\Delta\phi_{n,I_{RX}}(t; u, \tau)\right) \rangle_t. \tag{6.13}$$

Note that the second term in Equation 6.12 goes to zero as the time-domain average of cosine function is taken. Since $\phi_{n,I_{RX}}(t; \tau)$ is an accumulation of independent random variables along the time series, both $\phi_{n,I_{RX}}(t; \tau)$ and $\Delta\phi_{n,I_{RX}}(t; \tau, u)$ are zero-mean Gaussian random variables because of the central limit theorem. For a zero-mean Gaussian random variable x with standard deviation σ_x, the following identity holds:

$$E\left[\cos x\right] = E\left[\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}\right] = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} E\left[x^{2n}\right] = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \sigma_x^{2n}(2n - 1)!!$$

$$= \sum_{n=0}^{\infty} \frac{(-\sigma_x^2)^n}{(2n)!!} = \sum_{n=0}^{\infty} \frac{(-\sigma_x^2/2)^n}{n!} = \exp\left(-\frac{\sigma_x^2}{2}\right), \tag{6.14}$$

$$E\left[\sin x\right] = E\left[\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n + 1)!} x^{2n+1}\right] = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n + 1)!} E\left[x^{2n+1}\right] = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n + 1)!} 0 = 0. \tag{6.15}$$

We can now simplify Equation 6.13 as follows:

$$R_{I_{RX}}(u; \tau) = \frac{A_{I_{RX}}^2}{2} \cos\left(\omega_{I_{RX}} u\right) \exp\left(-\frac{\sigma_{\Delta\phi_{n,I_{RX}}}^2(u; \tau)}{2}\right). \tag{6.16}$$

It can be derived from Equation 6.16 that the power spectrum density of the receiver signal is expressed as follows:

$$S_{I_{RX}}(\omega; \tau) = \frac{A_{I_{RX}}^2}{2} \left(\frac{\delta(\omega - \omega_{I_{RX}}) + \delta(\omega + \omega_{I_{RX}})}{2}\right) * \mathcal{F}\left[\exp\left(-\frac{\sigma_{\Delta\phi_{n,I_{RX}}}^2(u; \tau)}{2}\right)\right] \tag{6.17}$$

$$= \frac{A_{I_{RX}}^2}{4} \left(S_{I_{RX}}^\star(\omega + \omega_{I_{RX}}; \tau) + S_{I_{RX}}^\star(\omega - \omega_{I_{RX}}; \tau)\right), \quad S_{I_{RX}}^\star(\omega; \tau) = \mathcal{F}\left[\exp\left(-\frac{\sigma_{\Delta\phi_{n,I_{RX}}}^2(u; \tau)}{2}\right)\right]. \tag{6.18}$$

We can further expand the variance of $\Delta\phi_{n,I_{RX}}$ as the following:

$$\sigma^2_{\Delta\phi_{n,I_{RX}}}(u;\tau) = \mathrm{E}\left[\left(\Delta\phi_{n,I_{RX}}(t;u,\tau)\right)^2\right] - \mathrm{E}\left[\Delta\phi_{n,I_{RX}}(t;u,\tau)\right]^2 \tag{6.19}$$

$$= \mathrm{E}\left[\left(\phi_{n,I_{RX}}(t;\tau) - \phi_{n,I_{RX}}(t-u;\tau)\right)^2\right] - 0 \tag{6.20}$$

$$= 2\sigma^2_{\phi_{n,I_{RX}}}(\tau) - 2\mathrm{E}\left[\phi_{n,I_{RX}}(t;\tau)\phi_{n,I_{RX}}(t-u;\tau)\right], \quad (\mu_{\phi_{n,I_{RX}}} = 0) \tag{6.21}$$

$$= 2\sigma^2_{\phi_{n,I_{RX}}}(\tau) - 2\mathrm{E}\left[\left(\phi_n(t) - \phi_n(t-\tau)\right)\left(\phi_n(t-u) - \phi_n(t-u-\tau)\right)\right] \tag{6.22}$$

$$= 2\sigma^2_{\phi_{n,I_{RX}}}(\tau) - 2\mathrm{E}\left[\phi_n(t)\phi_n(t-u)\right] + 2\mathrm{E}\left[\phi_n(t)\phi_n(t-u-\tau)\right]$$
$$\quad + 2\mathrm{E}\left[\phi_n(t-\tau)\phi_n(t-u)\right] - 2\mathrm{E}\left[\phi_n(t-\tau)\phi_n(t-u-\tau)\right] \tag{6.23}$$

$$= 2\sigma^2_{\phi_{n,I_{RX}}}(\tau) + 2\sigma^2_{\phi_n} - 2\mathrm{E}\left[\phi_n(t)\phi_n(t-u)\right] - 2\sigma^2_{\phi_n} + 2\mathrm{E}\left[\phi_n(t)\phi_n(t-u-\tau)\right]$$
$$\quad - 2\sigma^2_{\phi_n} + 2\mathrm{E}\left[\phi_n(t-\tau)\phi_n(t-u)\right] + 2\sigma^2_{\phi_n} - 2\mathrm{E}\left[\phi_n(t-\tau)\phi_n(t-u-\tau)\right] \tag{6.24}$$

$$= 2\sigma^2_{\phi_{n,I_{RX}}}(\tau) + \mathrm{E}\left[\left(\phi_n(t) - \phi_n(t-u)\right)^2\right] - \mathrm{E}\left[\left(\phi_n(t) - \phi_n(t-u-\tau)\right)^2\right]$$
$$\quad - \mathrm{E}\left[\left(\phi_n(t-\tau) - \phi_n(t-u)\right)^2\right] + \mathrm{E}\left[\left(\phi_n(t-\tau) - \phi_n(t-u-\tau)\right)^2\right] \tag{6.25}$$

$$= 2\sigma^2_{\phi_{n,I_{RX}}}(\tau) + 2\sigma^2_{\phi_{n,I_{RX}}}(u) - \sigma^2_{\phi_{n,I_{RX}}}(u+\tau) - \sigma^2_{\phi_{n,I_{RX}}}(u-\tau). \tag{6.26}$$

To find the variance of $\phi_{n,I_{RX}}$, we can again utilize Fourier relationship between the autocorrelation and power spectral density:

$$\sigma^2_{\phi_{n,I_{RX}}}(\tau) = R_{\phi_{n,I_{RX}}}(0;\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{\phi_{n,I_{RX}}}(\omega;\tau)d\omega \tag{6.27}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{\phi_n}(\omega)\left|2\sin\left(\frac{\omega\tau}{2}\right)\right|^2 d\omega. \tag{6.28}$$

To summarize, once you know the power spectral density of the laser phase noise (S_{ϕ_n}), one can find the autocorrelation function of the coherent receiver output ($R_{I_{RX}}$) using Equation 6.28, Equation 6.26, and Equation 6.18. There is no general closed-form expression for the integration in Equation 6.28. However, we can find one if the laser frequency noise is white, which is the case when the laser phase noise is dominated by spontaneous emission. If the 3 dB linewidth of the source laser is $\Delta\nu$, spectral density of the laser frequency and phase noise with whiteness assumption are expressed as follows:

$$S_{\omega_n}(\omega) = 2\pi\Delta\nu = \Delta\omega, \tag{6.29}$$

$$S_{\phi_n}(\omega) = \frac{\Delta\omega}{\omega^2}. \tag{6.30}$$

We can now plug Equation 6.30 in to Equation 6.28:

$$\sigma^2_{\phi_{\text{n}}, I_{\text{RX}}}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\Delta\omega}{\omega^2} \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega \tag{6.31}$$

$$= \frac{\Delta\omega\tau}{\pi} \int_{-\infty}^{\infty} \left(\frac{2}{\omega\tau}\right)^2 \left| \sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\left(\frac{\omega\tau}{2}\right) \tag{6.32}$$

$$= \Delta\omega\tau = 2\tau/\tau_{\text{coh}}. \tag{6.33}$$

Then, from Equation 6.26:

$$\sigma^2_{\Delta\phi_{\text{n}}, I_{\text{RX}}}(u; \tau) = 4\tau/\tau_{\text{coh}} + 4|u|/\tau_{\text{coh}} - 2|u + \tau|/\tau_{\text{coh}} - 2|u - \tau|/\tau_{\text{coh}} \tag{6.34}$$

$$= \begin{cases} 4\tau/\tau_{\text{coh}}, & |u| > \tau \\ 4|u|/\tau_{\text{coh}}, & |u| \leq \tau \end{cases} \tag{6.35}$$

Finally, this can be plugged into Equation 6.18 to get the expression for receiver signal PSD:

$$S^\star_{I_{\text{RX}}}(\omega; \tau) = \mathcal{F}\left[\exp\left(-\frac{\sigma^2_{\Delta\phi_{\text{n}}, I_{\text{RX}}}(u; \tau)}{2}\right) \right] = \int_{-\infty}^{\infty} \exp\left(-\frac{\sigma^2_{\Delta\phi_{\text{n}}, I_{\text{RX}}}(u; \tau)}{2}\right) e^{-j\omega u} du \tag{6.36}$$

$$= \int_{-\infty}^{\infty} e^{-\frac{2\tau}{\tau_{\text{coh}}}} e^{-j\omega u} du - \int_{-\tau}^{\tau} e^{-\frac{2\tau}{\tau_{\text{coh}}}} e^{-j\omega u} du + \int_{-\tau}^{\tau} e^{-\frac{2|u|}{\tau_{\text{coh}}}} e^{-j\omega u} du \tag{6.37}$$

$$= 2\pi\delta(\omega)e^{-\frac{2\tau}{\tau_{\text{coh}}}} + \frac{2}{\omega} e^{-\frac{2\tau}{\tau_{\text{coh}}}} \sin(\omega\tau) + 2\int_{0}^{\tau} e^{-\frac{2u}{\tau_{\text{coh}}}} \cos(\omega u) du. \tag{6.38}$$

The last term is computed using integration by parts:

$$S^\star_{I_{\text{RX}}}(\omega; \tau) = 2\pi\delta(\omega)e^{-\frac{2\tau}{\tau_{\text{coh}}}} + \frac{\tau_{\text{coh}}}{1 + (\omega/\Delta\omega)^2} \left\{ 1 - e^{-\frac{2\tau}{\tau_{\text{coh}}}} \left[\cos(\omega\tau) + \frac{\Delta\omega}{\omega} \sin(\omega\tau) \right] \right\}. \tag{6.39}$$

In reality, the length of the measurement window (T_{meas}) is always finite. Assuming $T_{\text{meas}} \gg \tau_{\text{coh}}, \tau$, we can modify Equation 6.39 as follows:

$$S^\star_{I_{\text{RX}}}(\omega; \tau, T_{\text{meas}}) = \frac{1}{2\pi} S^\star_{I_{\text{RX}}}(\omega; \tau) * T_{\text{meas}} \text{sinc}^2\left(\frac{\omega T_{\text{meas}}}{2}\right) \tag{6.40}$$

$$= T_{\text{meas}} \text{sinc}^2\left(\frac{\omega T_{\text{meas}}}{2}\right) e^{-\frac{2\tau}{\tau_{\text{coh}}}} + \frac{\tau_{\text{coh}}}{1 + (\omega/\Delta\omega)^2} \left\{ 1 - e^{-\frac{2\tau}{\tau_{\text{coh}}}} \left[\cos(\omega\tau) + \frac{\Delta\omega}{\omega} \sin(\omega\tau) \right] \right\}. \tag{6.41}$$

One can notice that the first term in Equation 6.41 represents the beating tone, and the second term resembles the Lorentzian function which is the laser lineshape function assuming white frequency noise. If the target distance is much shorter than the coherence range, or the delay τ is much smaller than τ_{coh}, the second term is negligible, and the phase noise lowers the power of
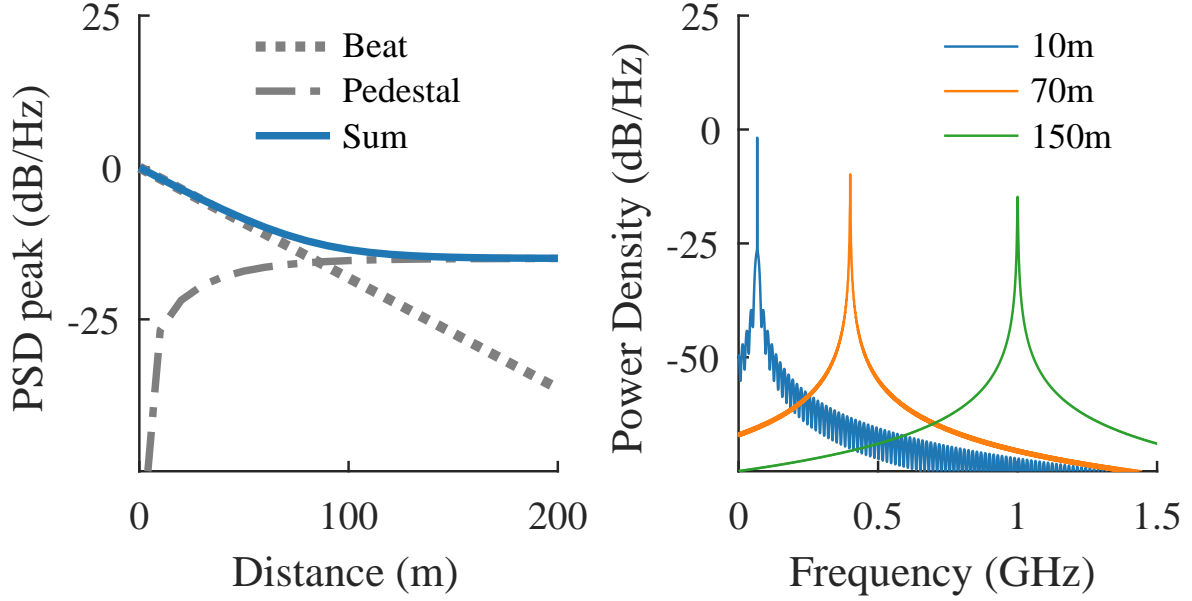
Figure 6.1: Signal peak magnitude of the FMCW LiDAR signal for different target distance (left) and corresponding PSD in the frequency domain (right). Power density was normalized to the peak at zero distance and free-space loss is not included ($T_{\mathrm{meas}} = 10\,\mu s$, $\Delta v = 1\,\mathrm{MHz}$, $R_{\mathrm{coh}} \approx 48\,\mathrm{m}$).

the beating tone through the attenuation factor $e^{-\frac{2\tau}{\tau_{\mathrm{coh}}}}$. As a result, SNR goes down exponentially as the distance, or τ, increases. Once the target range is comparable or larger than the coherence range, the second term (or phase noise *pedestal* term) starts to dominate, and it converges to the Lorentzian shape with the 3 dB linewidth of $2\Delta v$, two times bigger than the original source laser. In other words, LO and RX beams are no longer coherent. Figure 6.1 visualizes the simulated receiver signal PSD for different target distances in the presence of laser phase noise ($T_{\mathrm{meas}} = 10\,\mu s$, $\Delta v = 1\,\mathrm{MHz}$, $R_{\mathrm{coh}} \approx 48\,\mathrm{m}$). The free-space loss discussed in Equation 3.34 is not included to isolate the impact of phase noise. The PSD at the two extremes of the target distance can be summarized as follows:

$$S^{\star}_{I_{\mathrm{RX}}}(\omega; \tau, T_{\mathrm{meas}}) = \begin{cases} T_{\mathrm{meas}}\,\mathrm{sinc}^2\left(\frac{\omega T_{\mathrm{meas}}}{2}\right)e^{-\frac{2\tau}{\tau_{\mathrm{coh}}}}, & \tau \ll \tau_{\mathrm{coh}} \\ \frac{\tau_{\mathrm{coh}}}{1+(\omega/\Delta\omega)^2}, & \tau \gg \tau_{\mathrm{coh}} \end{cases} \tag{6.42}$$

Note that the ratio of the signal PSD density peak ($S^{\star}_{I_{\mathrm{RX}}}|_{\omega=0}$) between completely coherent and incoherent case is $T_{\mathrm{meas}}/\tau_{\mathrm{coh}}$, and this eventually converges to 1 for $\tau_{\mathrm{coh}} \geq T_{\mathrm{meas}}$. In the context of FMCW LiDAR, this represents the suppression factor for incoherent, interfering signals. Interestingly, this suggests that it is actually not desirable to use a laser with overly narrow linewidth
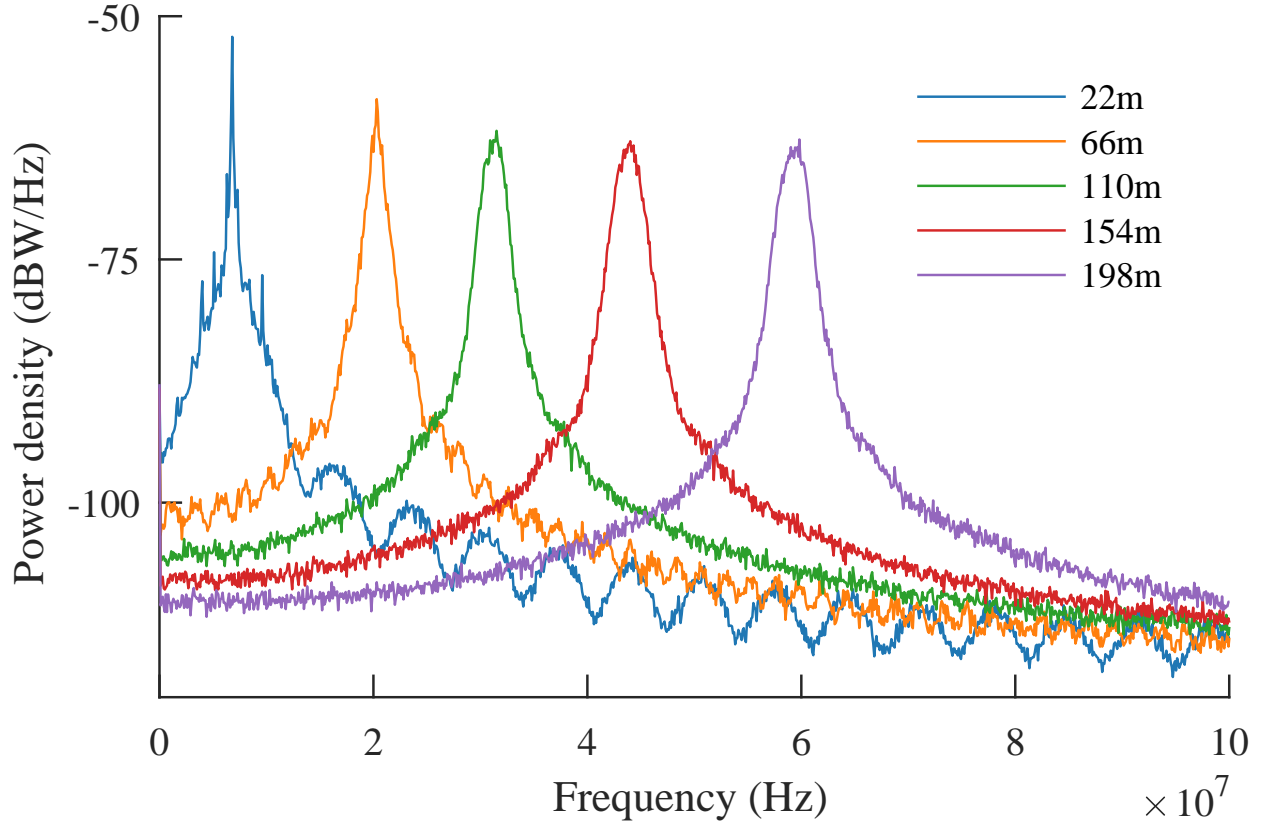
Figure 6.2: Measured FMCW LiDAR receiver signal spectrum for different target distance emulated by single-mode fibers. Similar to Figure 6.1, free-space loss is not introduced to emphasize the SNR degradation from laser phase noise.

from interference rejection perspective. Still, the worst-case absolute SNR of the LiDAR signal is degraded by $\exp\left(-2\tau_{\max}/\tau_{\mathrm{coh}}\right) = \exp\left(-4R_{\max}/(c\tau_{\mathrm{coh}})\right)$, and there is a direct trade-off between interference rejection and SNR degradation at the maximum range.

To validate the discussion so far by an experiment, simple laser self-homodyne setup shown in Figure 5.7 was again used. Different target distances were emulated by changing the length of the extra fiber on the top arm, which makes the corresponding target distance equal to one half of the fiber length times the fiber group index. Figure 6.2 shows the measured power spectral density (each spectrum was averaged from 10 measurements), which shows the pattern closely matched to what was expected from Figure 6.1.
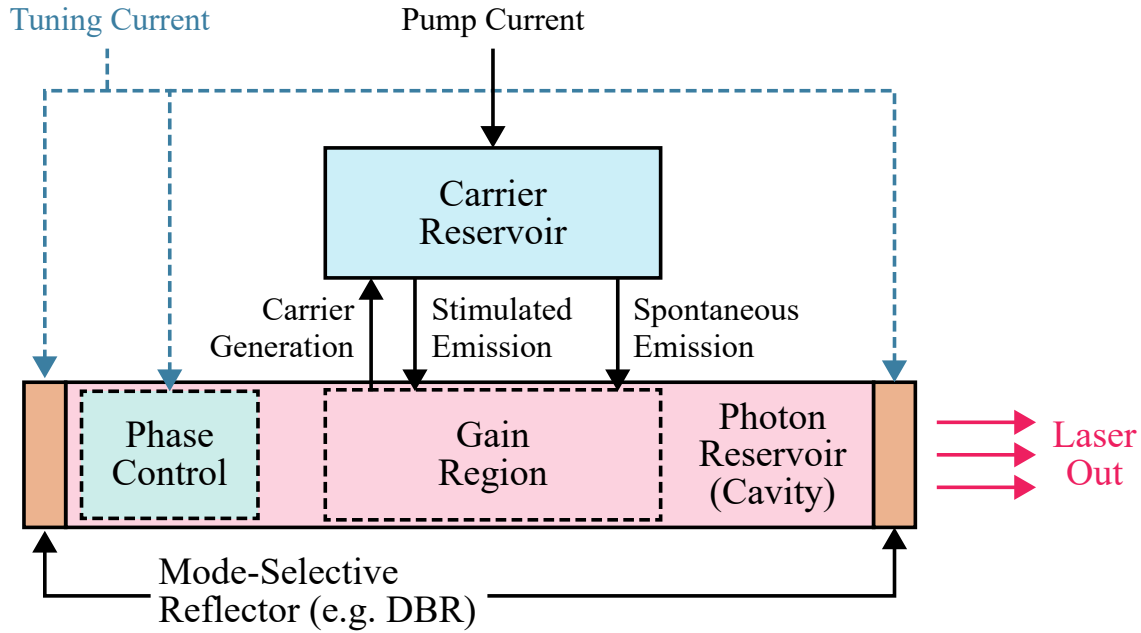
Figure 6.3: A simplified overview of tunable single-frequency laser where the lasing wavelength is changed by adjusting the cavity mode with an embedded phase shifter or changing the passband of the mode-selective reflectors.

6.2 Single-Frequency Tunable Lasers for FMCW LiDAR

Figure 6.3 shows a simplified overview of the single-frequency tunable laser often used as the FMCW source. The laser cavity is formed by a wave guiding structure capped on both sides by some type of reflectors. Wavelength-selective reflectors such as distributed Bragg reflectors (DBR) are often used in single-frequency tunable lasers to enhance the mode stability. The pumping current is injected into the active gain region placed within the laser cavity, serving as the source of carriers. Within the active region, carriers recombine and produce photons either through stimulated emission or via spontaneous emission. Note that stimulate emission results in additional photons in the lasing mode, and photons that are both coherent and incoherent to the lasing mode are added through spontaneous emission [39]. At the same time, some of the photons in the cavity lasing mode are absorbed back within the active region to generate carriers, and some exit the cavity through one of the reflectors. Ultimately, those interactions become balanced and reach the steady-state, where the density of the carrier within the active region and the density of the lasing-mode photon in the cavity does not change. Part of the photons leaking through one of the reflector form the laser output.

The lasing wavelength of the structure in Figure 6.3 is determined both by the longitudinal resonant mode of the cavity and the wavelength-selective reflectors. To tune the laser wavelength, the effective optical path length of the cavity round-trip or passband of the reflector should be changed. In semiconductor laser diodes, this is typically achieved through free-carrier plasma dispersion effect: current can be injected into a dedicated electro-optic phase shifter segment placed within the cavity to alter the local carrier density, which in turn changes the effective index as well as the cavity optical path length and resonant wavelength. The reflector can also be configured in a similar way: in the DBR case, index shift from injection current changes effective grating spacing and also the location of the passband. Note that the tuning speed of carrier injection-based tunable laser is determined by the carrier lifetime, and the 3 dB bandwidth of the wavelength tuning speed (f_{tuning}) is typically on the order of 100 MHz [86], [88], [89].

While basic tuning principle is common among different lasers, the details of the tunable laser design can vary depending on the application. For automotive LiDARs, continuous tuning range requirement is relatively small (a few GHz) as the ranging resolution goal is relaxed. Still, total tuning range must to be very wide if the grating antenna-based beam steering, illustrated in Chapter 4 and Chapter 5, is used. One of the well-known solutions for tuning range extension is utilizing sampled-grating distributed Bragg reflector (SG-DBR) [90]: leveraging Vernier effect between two SG-DBRs with different sample spacing as front and back reflector, very wide range (> 50 nm) is achieved over multiple mode hops. In addition, local, narrowband continuous tuning for FMCW modulation is performed by driving two DBRs with common injection current and/or manipulating the phase shifter dedicated for cavity longitudinal mode shift. On the other hand, high-resolution applications do need wide mode-hop-free tuning range. To achieve this, DBR lasers with very short cavity [88], [89], [91] are often used to extend the cavity mode spacing, and continuous tuning over ∼ 5 nm has been demonstrated. In such short-cavity DBR (SC-DBR) lasers, phase shifter for cavity mode tuning is often omitted [88], [89], and single input drives continuous DBRs on both front and back-side of the cavity.

Unfortunately, enhanced tuning capability often comes at a price: most widely tunable lasers have much larger linewidth than fixed-frequency lasers. The main difference is the existence of current-injection modulation. As mentioned in Section 6.1, the frequency noise in fixed-frequency laser is often assumed to be white and mostly dominated by the spontaneous emission in the optical gain region. In widely tunable lasers, the injection current applied to the tuning section produces significant shot noise, which perturbs the carrier density and eventually the laser frequency through the same tuning mechanism illustrated above [86]. Assuming that the

injection current-laser frequency gain (Hz/A) is α, following relationship holds:

$$S_{\omega_n}(\omega) = \frac{4\pi^2 \alpha^2 i_{n,\text{shot}}^2}{1 + (\omega/\omega_{\text{tuning}})^2} = \frac{4\pi^2 \alpha^2 q I_{\text{tuning}}}{1 + (\omega/\omega_{\text{tuning}})^2}. \tag{6.43}$$

Since this is no longer white, the receiver signal PSD of a FMCW measurement with a widely tunable laser will be different than Equation 6.41. We can go back to Equation 6.28 and express the variance of the coherent receiver output phase noise with new frequency noise expression in Equation 6.43.

$$\sigma_{\phi_{n,I_{\text{RX}}}}^2(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{\phi_n}(\omega) \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_{\omega_n}}{\omega^2} \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega \tag{6.44}$$

$$= \frac{1}{\pi} \int_0^{\infty} \frac{1}{\omega^2} \frac{4\pi^2 \alpha^2 q I_{\text{tuning}}}{1 + (\omega/\omega_{\text{tuning}})^2} \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega \tag{6.45}$$

$$\approx \frac{1}{\pi} \int_0^{\infty} \frac{4\pi^2 \alpha^2 q I_{\text{tuning}}}{\omega^2} \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega - \frac{1}{\pi} \int_{\omega_{\text{tuning}}}^{\infty} \frac{4\pi^2 \alpha^2 q I_{\text{tuning}}}{\omega^2} \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega$$

$$+ \frac{1}{\pi} \int_{\omega_{\text{tuning}}}^{\infty} \frac{1}{\omega^2} \frac{4\pi^2 \alpha^2 q I_{\text{tuning}}}{1 + (\omega/\omega_{\text{tuning}})^2} \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega \tag{6.46}$$

$$= 4\pi^2 \alpha^2 q I_{\text{tuning}}\tau - \frac{1}{\pi} \int_{\omega_{\text{tuning}}}^{\infty} \frac{4\pi^2 \alpha^2 q I_{\text{tuning}}}{\omega^2} \frac{(\omega/\omega_{\text{tuning}})^2}{1 + (\omega/\omega_{\text{tuning}})^2} \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega \tag{6.47}$$

$$\approx 4\pi^2 \alpha^2 q I_{\text{tuning}}\tau - \frac{1}{\pi} \int_{\omega_{\text{tuning}}}^{\infty} \frac{4\pi^2 \alpha^2 q I_{\text{tuning}}}{\omega^2} \left| 2\sin\left(\frac{\omega\tau}{2}\right) \right|^2 d\omega. \tag{6.48}$$

It is also worth remembering from Equation 6.8 that through the self-mixing process in FMCW measurement, laser phase domain noise is transferred to the coherent receiver output phase domain through a high-pass filter whose cutoff frequency is $\omega_{\text{cutoff}}(\tau) = \sqrt{2}/\tau = c/(\sqrt{2}R)$ (e.g. $\sim 30\,\text{cm}$ target corresponds to 112 MHz cutoff frequency). Therefore, if $\omega_{\text{cutoff}}(\tau) < \omega_{\text{tuning}}$, Equation 6.48 can be further approximated using Equation 6.8:

$$\sigma_{\phi_{n,I_{\text{RX}}}}^2(\tau) \approx 4\pi^2 \alpha^2 q I_{\text{tuning}}\tau - \frac{1}{\pi} \int_{\omega_{\text{tuning}}}^{\infty} \frac{8\pi^2 \alpha^2 q I_{\text{tuning}}}{\omega^2} d\omega = 4\pi^2 \alpha^2 q I_{\text{tuning}}\tau - C_0. \tag{6.49}$$

Note that the last term in Equation 6.49 is no longer dependent on τ and can be treated as a constant. For $|u| > 2\tau$ and $\omega_{\text{cutoff}}(\tau) < \omega_{\text{tuning}}$, we can express $\sigma_{\Delta\phi_{n,I_{\text{RX}}}}^2$ from Equation 6.26 as the following:

$$\sigma_{\Delta\phi_{n,I_{\text{RX}}}}^2(u;\tau) = 8\pi^2 \alpha^2 q I_{\text{tuning}}\tau - 2C_0, \qquad (u > 2\tau, \ \omega_{\text{cutoff}}(\tau) < \omega_{\text{tuning}}). \tag{6.50}$$

Namely, $\sigma_{\Delta\phi_{n,I_{\text{RX}}}}^2(u;\tau)$ with bandlimited laser frequency noise is independent of u when u is sufficiently larger than τ (similar to white frequency noise case in Equation 6.35) and if τ is large

enough that the cutoff frequency of the phase domain high pass filter from self-mixing is lower than the bandwidth of the laser frequency noise. As seen from Equation 6.38, constant $\sigma^2_{\Delta\phi_{\text{n},I_{\text{RX}}}}$ value for large u implies exponential decay of the beating tone power with respect to increasing target distance. As mentioned above, f_{tuning} is on the order of $100\,\text{MHz}$, and SNR attenuation beyond a few 10s of centimeters will be $e^{-4\pi^2\alpha^2 qI_{\text{tuning}}\tau}$ where effective linewidth is $2\pi\alpha^2 qI_{\text{tuning}}$. In typical widely tunable lasers, this is substantially higher than fixed frequency lasers and on the order of $1\,\text{MHz} \sim 10\,\text{MHz}$ (or the coherence range of $5 \sim 50\,\text{m}$).

To sumarize, we can realize compact integrated lasers with fast, wide tuning range by the combination of effective index modulation through carrier injection and reflector/cavity design techniques, which is very promising for consumer-grade FMCW LiDAR implementation. However, increased laser phase noise due to the shot noise in the tuning section can significantly limit the ranging performance. In the next two sections, I will present circuit- and algorithm-side solutions to mitigate the performance degradation in both long-range and high-precision imaging applications.

6.3 Feedforward Phase Noise Cancellation with FMCW Modulation

Laser frequency/phase stabilization techniques are commonly used in scientific experiments whenever high-precision spectroscopy or interferometry measurements are needed. For instance, it was a crucial part of the LIGO system [92] designed for detecting the gravitational wave. Although existing techniques are used mostly in the context of lab setting, there is also an active effort to miniaturize the stabilization system and ultimately bring low-noise lasers to portable devices, and this is becoming even more promising with the advancements in the integrated photonics technology [93], [94].

Although there are a number of different methods, the principle of operation is similar and summarized as follows:

1. The frequency/phase of the laser is detected and converted to an electrical signal. In most cases, the frequency of the laser is firstly converted to optical intensity using a resonant structure such as Fabry-Pérot cavity [95] or asymmetric Mach-Zehnder Interferometer (MZI), and then the intensity is detected by a photodetector what outputs electrical current. A resonator and a photodetector is often collectively referred to as an optical frequency discriminator.
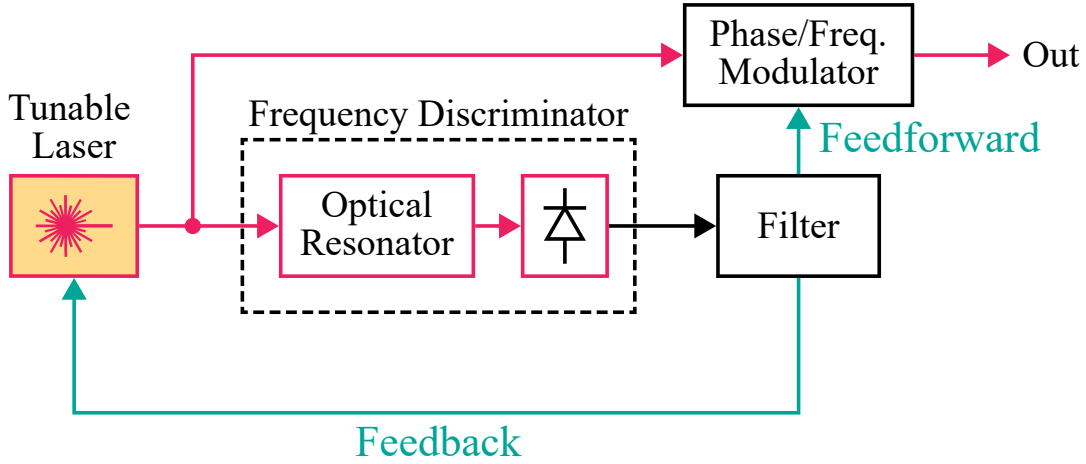
Figure 6.4: Principle of laser stabilization techniques based on a frequency discriminator and feedback/feedforward control.

2. The output of the optical frequency detector is compared against the desired *reference* value to create the error signal, and then passed through a filter to generate proper control signal.

3. Finally, generated control signal is fed back into the laser tuning input to adjust the laser frequency/phase so as to reduce the error. Alternatively, one can take a copy of the seed laser, and manipulate the laser using dedicated optical phase/frequency modulator driven by the generated control signal: this concept is referred to as feedforward laser stabilization scheme [96].

Figure 6.4 also shows an overview of the laser stabilization system. Generally, feedforward scheme has advantages in achieving wider stabilization bandwidth and also tends to be *laser agnostic*: for instance, it is not affected by laser tuning gain- and laser power-dependent loop stability issues. On the flip side, it requires precise gain and latency matching between feedforward path and discrimination path. Moreover, nonidealities associated with the standalone phase/frequency modulator (e.g. finite modulation range) can limit the effectiveness of the scheme [96].

   Although the method described above is applied mostly for fixed-frequency lasers, it is possible to apply the same principle to frequency modulated lasers. If an FMCW source laser modulated with a sawtooth or triangular wave (Equation 3.2) is used in the system shown in Figure 6.4, and if the frequency discriminator is ideal, i.e. has linear relationship between the input laser frequency and the output current, the output of the frequency discrimination should also be
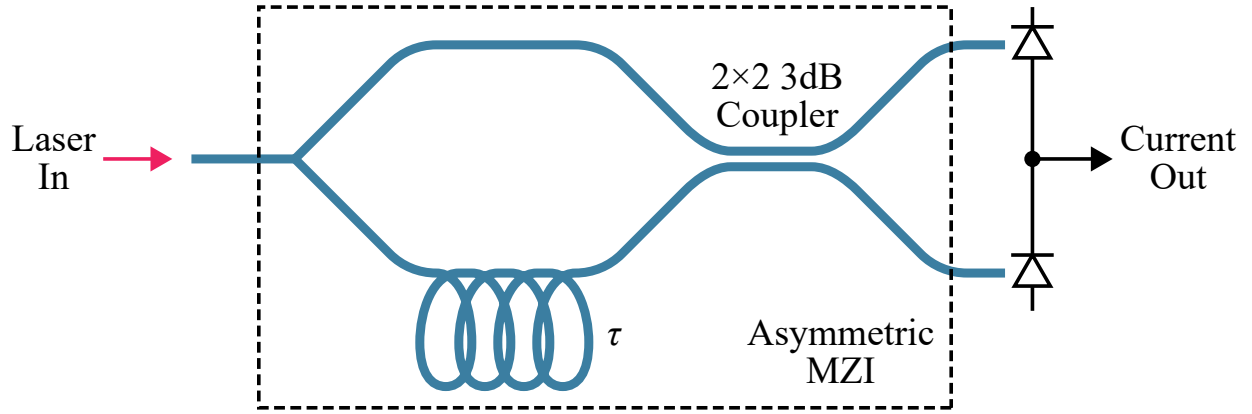
Figure 6.5: Optical frequency discrimination circuit based on an asymmetric Mach-Zehnder interferometer and a balanced detector.

sawtooth/triangular wave. In other words, one can calculate the difference between the *reference* frequency pattern and the discriminator output to correct the laser input.

However, existing frequency discriminators are not linear. As an example, Figure 6.5 shows the asymmetric MZI-based discriminator commonly used as a frequency discriminator. It includes a MZI with optical delay mismatch of τ between two arms, and then the output of the MZI is converted to current via balanced photodetector. One can notice that this is physically equivalent to the FMCW LiDAR measurement illustrated in Section 3.1 with free-space time of flight τ, and we can expressed the output of this discriminator from Equation 3.25:

$$I_{\text{disc.}}(t) = A_{\text{disc.}} \cos\left(\phi_{\text{laser}}(t) - \phi_{\text{laser}}(t - \tau)\right), \quad A_{\text{disc.}} = R_{\text{PD}} P_{\text{laser}}. \tag{6.51}$$

Note that Equation 6.51 assumes that the split ratio of the 1×2 splitter at the input of the MZI is 50:50. For small τ, we can make following simplification to Equation 6.51:

$$I_{\text{disc.}}(t) = A_{\text{disc.}} \cos\left(\tau \frac{\phi_{\text{laser}}(t) - \phi_{\text{laser}}(t - \tau)}{\tau}\right) \tag{6.52}$$

$$\approx A_{\text{disc.}} \cos\left(\tau \dot{\phi}_{\text{laser}}(t)\right). \tag{6.53}$$

Namely, the discriminator detects the laser frequency, multiplies the frequency by MZI mismatch, and maps to the output current through a cosine function. In fact, we have already seen in Equation 6.8 that the FMCW measurement is approximated to a differentiation in the phase domain with path mismatch τ as the gain, when the frequency of interest is lower than $1/\tau$.
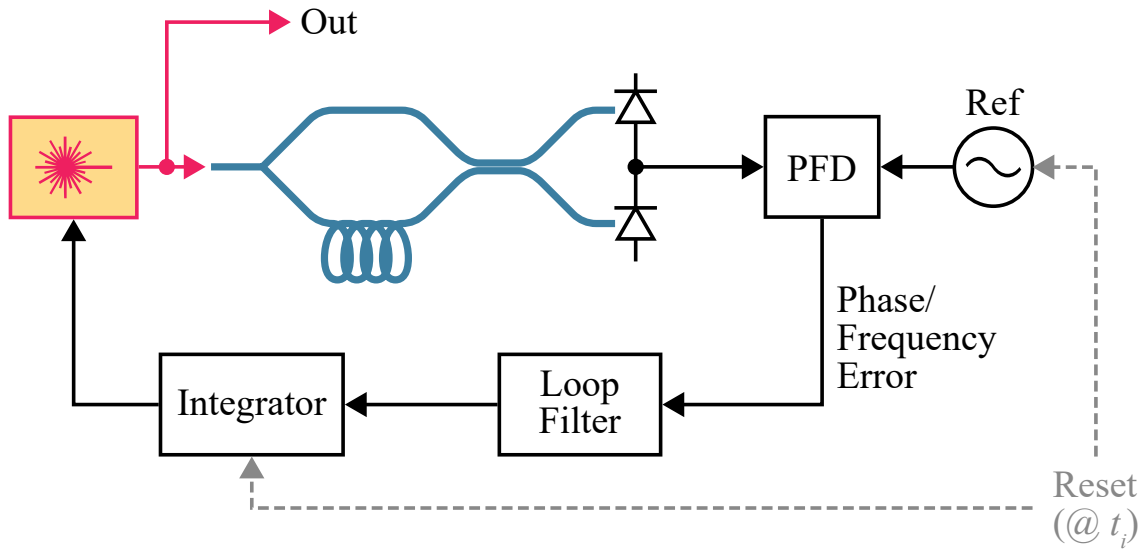
Figure 6.6: FMCW modulation stabilization system based on electro-optic phase-locked loop.

Given that, for an ideal FMCW source with sawtooth modulation (Equation 3.2), we can find the expected discriminator output:

$$I_{\text{disc.,ideal}}(t) \approx A_{\text{disc.}} \cos\left(\frac{2\pi f_{\text{BW}}}{T_{\text{mod}}}\tau(t - t_i) + 2\pi\tau f_{\text{min}}\right),\ t_i \leq t < t_{i+1},\ i \in \mathbb{Z}. \tag{6.54}$$

In other words, we can compare the frequency and the phase of the discriminator output signal to the reference sinusoid to set both the modulation slope ($\gamma = f_{\text{BW}}/T_{\text{mod}}$) and the minimum frequency (f_{min}). This can be done with a phase-locked loop, and chirp linearization schemes based on an electro-optic phase-locked loop in [86], [97] (also shown in Figure 6.6) are based on this principle. Moreover, since a phase-locked loop introduces a high-pass transfer function in the phase domain, the laser phase noise is also suppressed up to the closed-loop bandwidth of the PLL.

Unfortunately, the idea of using optical PLL for laser phase noise suppression makes little sense in the context of long-range automotive LiDARs. Note from Equation 6.54 that the reference sinusoid frequency is determined by the multiplication of the modulation slope and the delay mismatch in the asymmetric MZI ($f_{\text{ref}} = \gamma\tau$). According to the discussion in Sections 3.2 and 3.5, typical modulation slope in long-range FMCW LiDAR is around a few GHz over $\sim 100\,\mu\text{s}$, which results in the modulation slope of $\sim 10\,\text{THz/s}$. If the delay mismatch is 1 ns (which corresponds to physical delay of 30 cm), resulting reference frequency is only 10 kHz. In other words, since the target modulation slope of the FMCW sources for long-range LiDARs are rather small, reference

oscillation for the optical PLL is very slow unless $> 1\,\mathrm{m}$ scale mismatch is allowed in the MZI, which is clearly unrealistic for integrated MZIs. Low reference oscillation frequency in turn limits dynamic response of the PLL, since closed-loop bandwidth of the PLL is typically limited to around 10% of the reference frequency [98]. Although it is possible to relax this constraint by using analog phase detector [99], the speed of the loop is still determined primarily by the reference frequency.

One can then suggest to use the feedforward scheme for FMCW source because it is not limited by the loop dynamics. From Equation 6.53 and Equation 6.54, we can express the discriminator output for a FMCW source whose modulation is ideal but still corrupted by the phase noise:

$$I_{\mathrm{disc.}}(t) \approx A_{\mathrm{disc.}} \cos\left(2\pi\gamma\tau(t - t_i) + 2\pi\tau f_{\min} + \tau\dot{\phi}_{\mathrm{laser}}(t)\right),\ t_i \leq t < t_{i+1},\ i \in \mathbb{Z}. \qquad (6.55)$$

Again, this expression is only valid up to $\sim 1/\tau$, but it is still quite wide ($> 1\,\mathrm{GHz}$) if an on-chip MZI with smaller than centimeter-scale mismatch was used. Equation 6.55 implies that to realize feedforward cancellation of laser phase noise, first two terms within the cosine should somehow be removed. In fact, that was the exact role of the phase detector in electro-optic PLLs, and direct analog mixing against a reference tone with $f_{\mathrm{ref}} = \gamma\tau$ could be a way to isolate the phase noise term [99]. Unfortunately, the output of the mixer will also contain the second harmonic at $2f_{\mathrm{ref}}$, not just the desired baseband signal. This is a problem not only because it requires an additional filter to reject the second harmonics and makes the sensing bandwidth to be effectively limited again by f_{ref}, but also because the phase noise skirt of the second harmonics will be present at the baseband and cause aliasing (Figure 6.7(b)). To eliminate the second harmonics, it also possible to insert a Hilbert filter at the output of the discriminator to create both in-phase and quadrature tone (i.e. IQ generation) and then perform single-sideband downconversion [100]. However, the fact that the bandwidth of interest around the carrier ($\sim 1/\tau$) is wider than the carrier frequency (f_{ref}) still causes noise aliasing, as illustrated in Figure 6.7.

To avoid noise aliasing during single-sideband downconversion, IQ generation must be performed in the optical domain where the carrier frequency is much higher than $1/\tau$. This can be done by using optical coherent receiver frontend with full IQ generation [101] in the frequency discriminator. Figure 6.8 shows the modified frequency discriminator which outputs both in-phase and quadrature version of Equation 6.55. After single-sideband downconversion, the output of the frequency discriminator is expressed as follows:

$$I_{\mathrm{disc.}}(t) \approx A_{\mathrm{disc.}} \cos\left(\tau\dot{\phi}_{\mathrm{laser}}(t) + \phi_0 + \phi_{\mathrm{offset}}\right). \qquad (6.56)$$
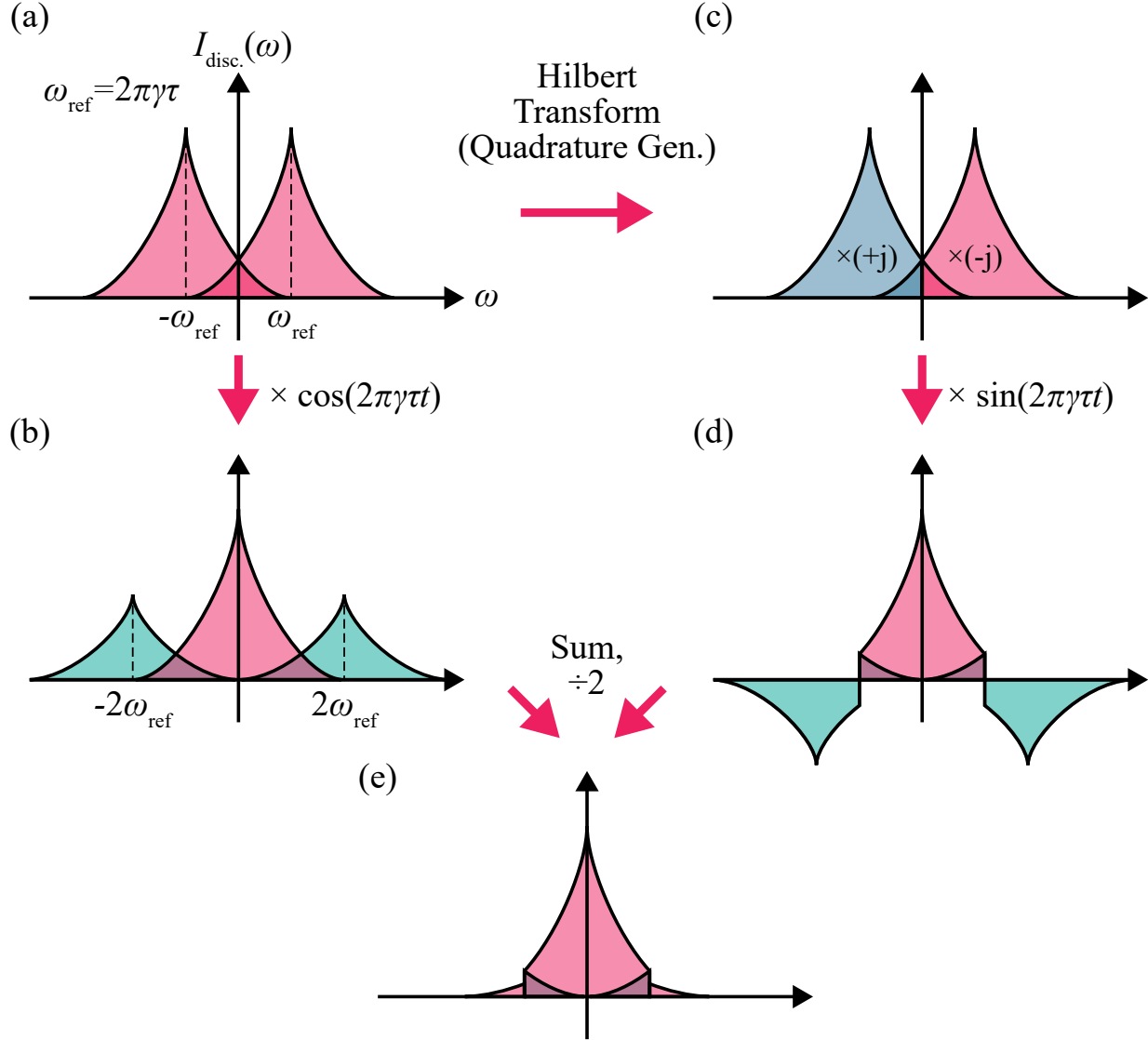
Figure 6.7: (a) Spectrum of the optical frequency discriminator output when FMCW modulated laser with low chirp rate is used as the input. (b) Spectrum of the frequency discriminator output where analog mixer is added after the optical coherent detection frontend, highlighting the noise aliasing caused by the phase noise skirt of the second-order harmonics. (c-e) Spectrum of the frequency discriminator output where single-sideband downconverting mixer is added after the optical frontend, illustrating the problem of noise aliasing when ω_{ref} is low.

Figure 6.8: Modified asymmetric MZI-based frequency discriminator with both in-phase and quadrature output.

$\phi_t ext of fset$ is the phase offset introduced to the reference oscillator, which can be configured to satisfy $\phi_0 + \phi_{\text{offset}} = \pi/2$. Resulting output of the discriminator is as follows.

$$I_{\text{disc.}}(t) \approx A_{\text{disc.}} \sin\left(\tau \dot{\phi}_{\text{laser}}(t)\right) \approx A_{\text{disc.}} \tau \dot{\phi}_{\text{laser}}(t). \tag{6.57}$$
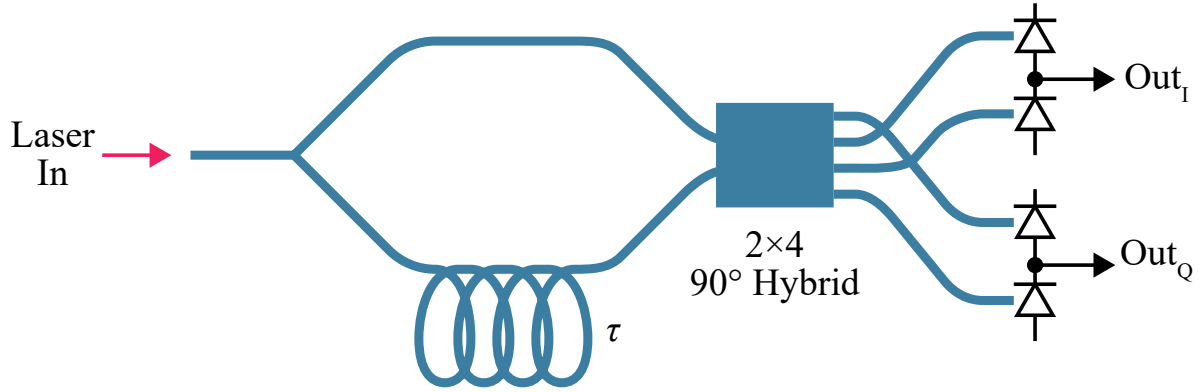
By taking integration of the discriminator output, one can directly sense the phase noise of the laser, and this can be used to realize feedforward noise cancellation shown in Figure 6.9. Note that in order to realize perfect cancellation, latency and the gain of the feedforward and noise sensing path has to be matched, and this can be fulfilled by having dedicated delay line and variable gain amplifier (VGA).

To validate the proposed scheme, behavioral simulation was done using Simulink model shown in Figure 6.10. Chirp rate was set to $2\,\text{GHz}/10\,\mu\text{s}$ and the MZI delay mismatch was set as $1\,\text{ns}$, resulting in f_{ref} of $200\,\text{kHz}$. Laser linewidth was $1\,\text{MHz}$ ($\tau_{\text{coh}} = 3.14\,\mu\text{s}$, $R_{\text{coh}} \sim 47.7\,\text{m}$). Figure 6.11(a) shows the frequency noise PSD of the laser with and without the proposed scheme, and it is clear that feedforward cancellation suppresses the laser phase noise up to $1/(2\pi\tau) \sim 160\,\text{MHz}$. Figure 6.11(b) shows the receiver output spectrum for a FMCW LiDAR measurement with $200\,\text{m}$ target distance and measurement window of $10\,\mu\text{s}$. From the baseline measurement spectrum, it is clear that the FMCW measurement is in the coherent regime and the spectral shape is close to Lorentzian. On the other hand, with the feedforward cancellation on, the peak SNR is improved by $15\,\text{dB}$, which agrees with to the coherence gain in Equation 6.42 ($T_{\text{meas}}/\tau_{\text{coh}} = 10\,\mu\text{s}/3.14\,\mu\text{s} \sim 31.4$). Figure 6.11(c,d) shows the performance degradation of the proposed scheme in he presence of delay and gain matching error. While the SNR gain is preserved even with 100% delay mismatch error, the performance was sensitive to the gain matching error. This can be a potential
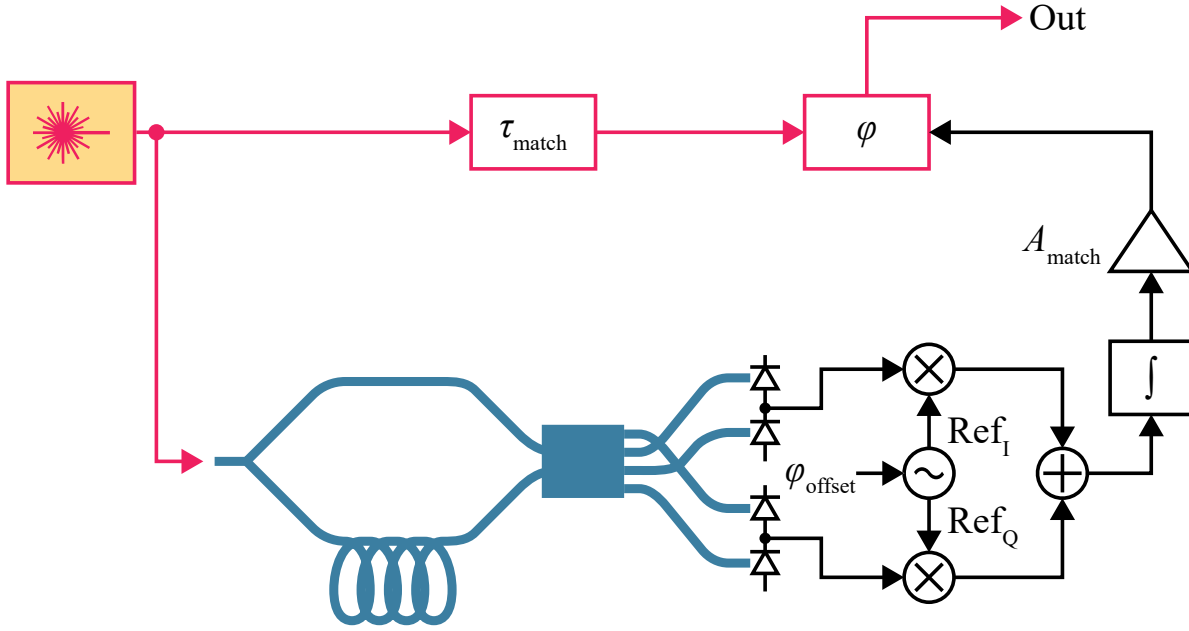
Figure 6.9: Feedforward FMCW source phase noise cancellation system based on IQ frequency discriminator from Figure 6.8.

issue considering that the gain of the frequency discriminator is proportional to the input laser power. Moreover, the power of the tunable laser is often coupled to the frequency tuning because of the free carrier absorption effect in the current injection phase shifter within the cavity. Fortunately, IQ coherent receiver enables direct sensing of input laser power measurement independent of phase measurement ($P_{\text{laser}} \propto \sqrt{A_{\text{disc.,I}}^2 + A_{\text{disc.,Q}}^2}$), and the impact of power fluctuation can be compensated by adjusting the VGA using measured power.
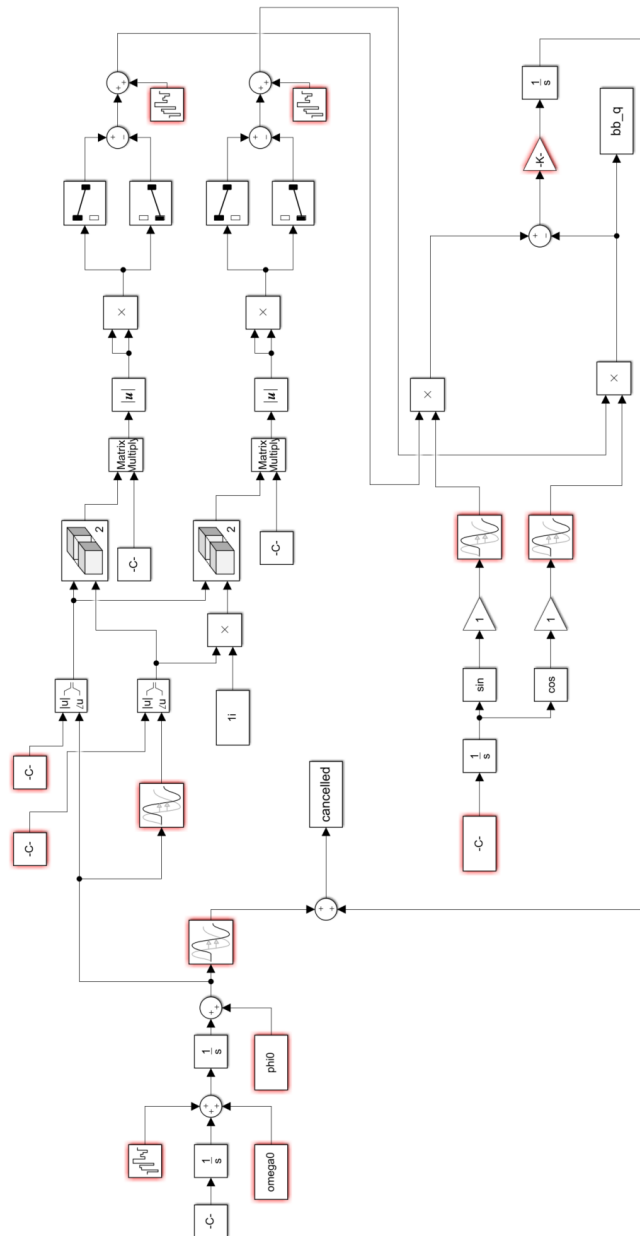
Figure 6.10: Simulink behavioral model of the feedforward FMCW source phase noise cancellation system.
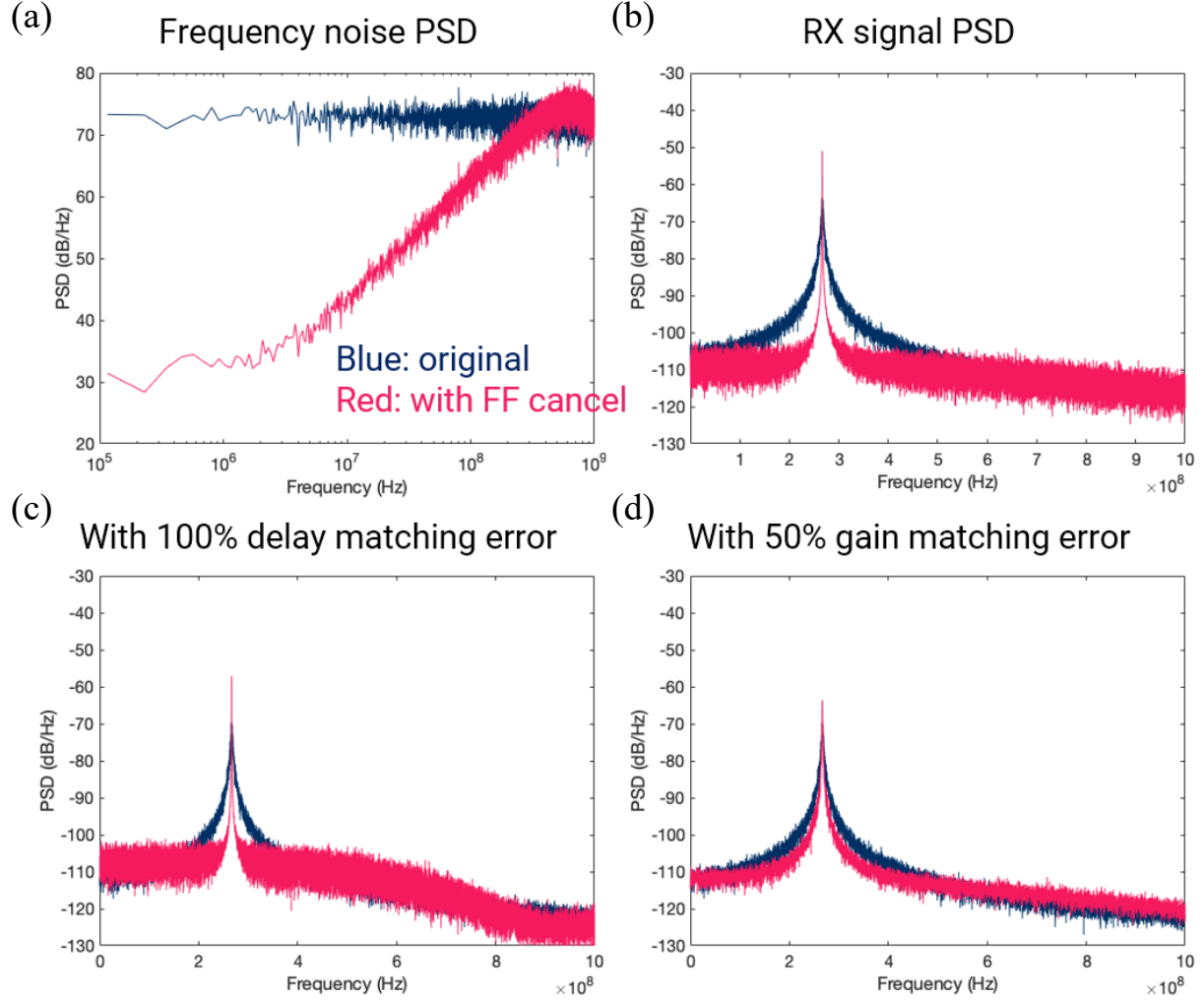
Figure 6.11: Behavioral simulation result of the feedforward FMCW source phase noise cancellation system ($\gamma = 2\,\text{GHz}/10\,\mu\text{s}$, $\tau_{\text{MZI}} = 1\,\text{ns}$, $\Delta\nu = 1\,\text{MHz}$, $R = 200\,\text{m}$, $T_{\text{meas}} = 10\,\mu\text{s}$).

6.4 Optimal Spectral Estimation Algorithm for Incoherent FMCW Measurements[1]

It has been commonly assumed that the detection range of the FMCW LiDAR is fundamentally limited by the coherence range [102]. In other words, once the target distance becomes longer than the coherence range, the receiver has been deemed phase noise-limited regardless of the SNR. This can be highly limiting for the laser design in the context of short-distance, high-precision applications such as OCT or microimagers. As explained in Equation 3.15, high ranging resolution/precision requires wide laser frequency modulation bandwidth, which means it also requires a tunable laser with wide mode hopping-free tuning range. However, as discussed in Section 6.2, there is a fundamental trade-off between phase noise and continuous tuning range. Alternatively, a chirped laser can also be generated using continuous-wave laser followed by an external I/Q modulator [103]. However, this approach requires fast drivers and chirp generator over multi-GHz bandwidth. In other words, required electrical bandwidth is directly determined by f_{BW}, and this removes the major advantage of FMCW LiDAR that it is possible to use slow electrical circuits to achieve high resolution, as explained in Section 3.1.

As mentioned in Section 3.1, the role of the receiver DSP backend is to solve a classical problem of line spectra estimation. It is one of the most well-studied topics in signal processing [87], and there are numerous algorithms one can choose depending on the nature of additive noise and affordable complexity. In general, it is possible to achieve the theoretical accuracy bound (i.e. Cramér-Rao lower bound) as long as the SNR is higher than the threshold. The signal model used commonly in standard line spectral estimation algorithm is as following [87]:

Note that the phase offset of the sinusoidal tone may be unknown but can be modeled as constant within single observation. Unfortunately, such constant phase offset assumption is invalid for the incoherent FMCW measurement. In [104], it was shown that the fundamental lower bound of the frequency estimation variance becomes a function of the amount of the phase noise, in addition to the measurement time and additive noise. Especially, it was also shown that any algorithm designed assuming constant phase offset performs poorly in presence of the phase noise.

We can actually notice that even though the signal spectrum in the incoherent regime has a different shape and is suppressed by $T_{\mathrm{meas}}/\tau_{\mathrm{coh}}$, the PSD of the photocurrent is still dependent on the target distance as it is evident in Figure 6.1 and Figure 6.2. Moreover, we have the prior information from Equation 6.18 and Equation 6.42 that the PSD corresponding to an incoherent

---

[1]The content of this chapter was derived from [17], [18].

measurement is a Lorentzian shifted by the beat frequency ($\omega_{I_{RX}}$). From this observation, we can reason that by leveraging such knowledge to design optimal spectral estimation algorithm, it may be possible to achieve decent ranging performance in the incoherent regime. From Equation 6.18, Equation 6.42, and Equation 3.26, we can define the signal model of the single-sided power spectral density of any incoherent, shot noise-limited FMCW measurement as the following:

$$\tilde{S}_{I_{RX}}(\omega; \alpha, \omega_{I_{RX}}) = \sum_{i=1}^{N} \frac{\alpha_i}{(\omega - \omega_{I_{RX},i})^2 + \Delta\omega^2} + 2qR_{PD}P_{LO}. \tag{6.58}$$

If the number of possible targets is assumed to be N, there are $2N$ parameters: $\omega_{I_{RX},i}$ and α_i are the center frequency and relative power of the ith Lorentzian, respectively. Given this model and the periodogram estimate of the PSD ($\hat{S}_{I_{RX}}$) from the measured waveform, we can simply perform nonlinear least-squares to estimate those parameters.

$$\alpha^*, \omega_{I_{RX}}^* = \underset{\alpha, \omega_{I_{RX}}}{\arg\min} \left| \hat{S}_{I_{RX}}(\omega) - \tilde{S}_{I_{RX}}(\omega; \alpha, \omega_{I_{RX}}) \right|^2 \tag{6.59}$$

Note that the Lorentzian-shape PSD itself is deterministic, but its estimate using periodogram adds uncertainty. It can also have small bias if the length of the measurement is too small [37], but this is generally negligible considering realistic sample rates and observation times. In addition, it is possible that the lineshape of the practical laser deviates from Lorentzian shape, depending on its dominant phase noise mechanism [86]. However, it is possible to measure the lineshape of the laser under test and improve the parametric model using different prior (for example such as the Voigt function [105]).

In order to test the performance of the proposed Lorentzian least squares estimation (LLSE) and compare it to the standard frequency estimation schemes with constant phase offset model, we built a behavioral model of the FMCW LiDAR using Simulink and ran transient simulations to generate realistic data. We have assumed that the E_{LO} beam is strong enough for the receiver to be shot-noise limited, and only one target exists in the measurement. Baseline laser parameters were $f_{BW} = 10\,\text{GHz}$, $T = 10\,\mu\text{s}$, $\Delta\nu = 1\,\text{MHz}$ which correspond to $\gamma = 1\,\text{GHz}/1\,\mu\text{s}$. R_{PD} was $1\,\text{A/W}$. With simulated time-domain data for target distance up to $100\,\text{m}$, we applied different algorithms including the proposed LLSE method to estimate the distance and recorded estimation variance from 100 Monte Carlo simulations per each distance. Among a number of constant-phase frequency estimation methods, Rife and Boorstyn's [38] and MUSIC [106] were used for comparison.

Figure 6.12(a) shows the high-SNR result where P_{RX} is $1\,\text{mW}$. Estimated PSD shows that the shot noise floor is almost negligible compared to the Lorentzian noise *pedestal*. In this case,
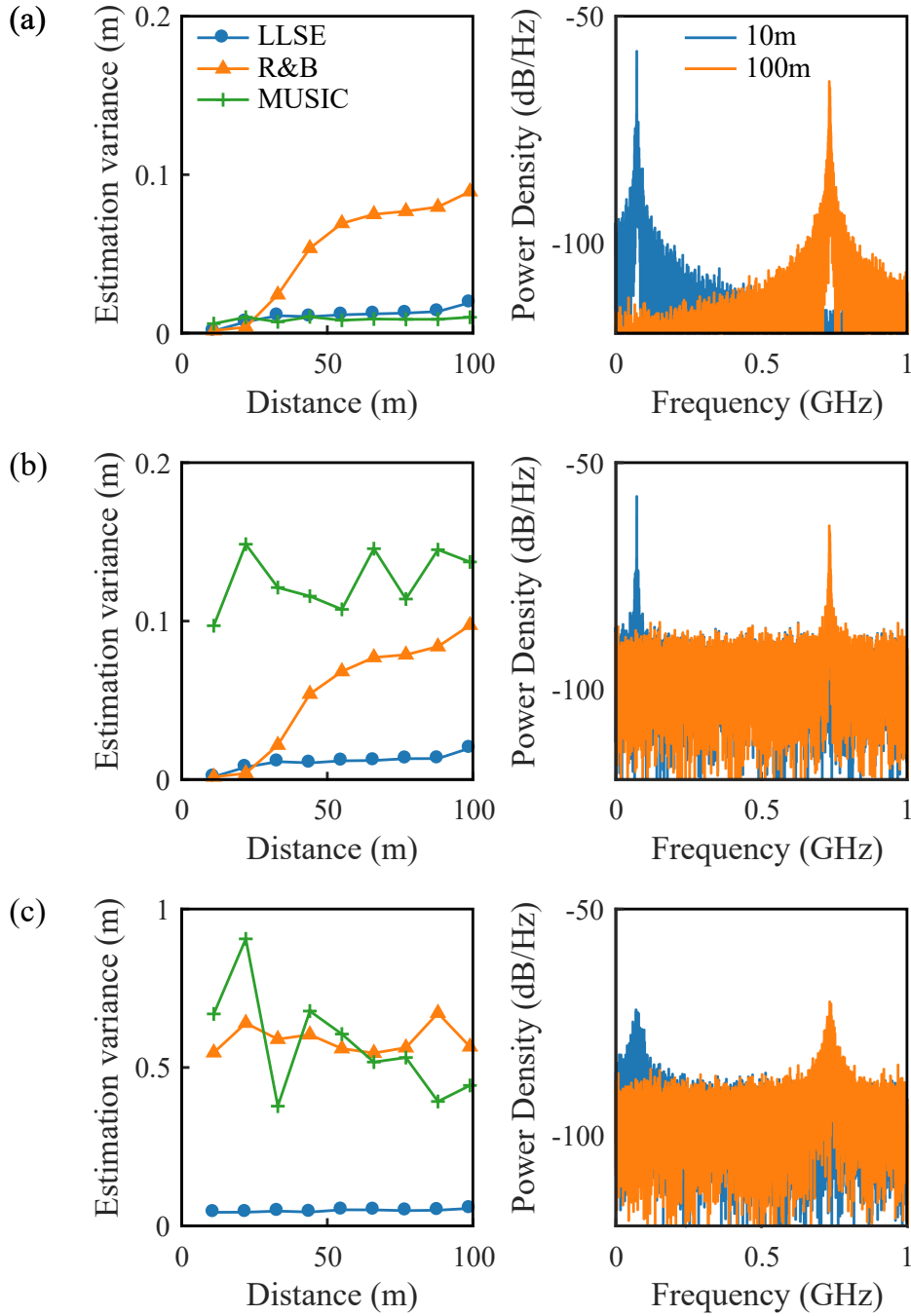
Figure 6.12: Impact of frequency estimation algorithm on FMCW measurement. Distance estimation variance for different algorithms and periodogram PSD estimates with Least squares fit are shown for (a) $(P_{RX}, \Delta v) = (1\,mW, 1\,MHz)$ (b) $(P_{RX}, \Delta v) = (1\,nW, 1\,MHz)$ (c) $(P_{RX}, \Delta v) = (1\,nW, 10\,MHz)$
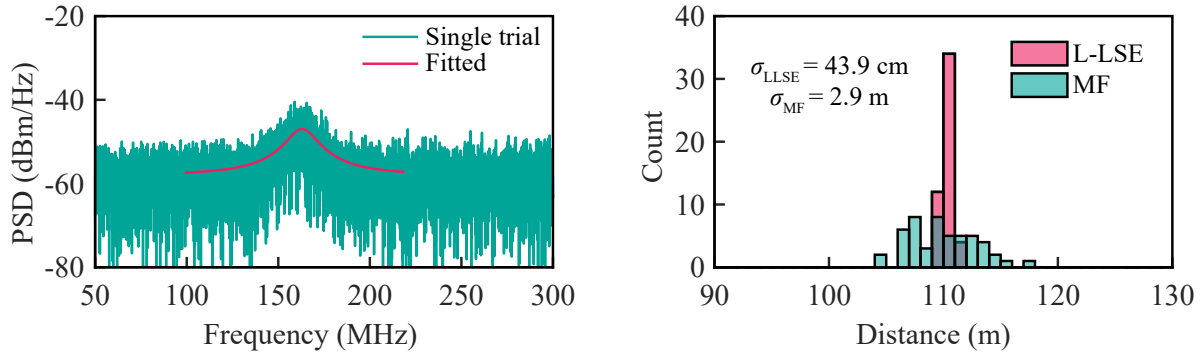
Figure 6.13: Experimental demonstration of proposed LSE-based spectral estimation (γ = 220 THz/s, $\Delta v \sim 7$ MHz).

the performance of MUSIC algorithm is the best, and proposed LLSE shows similar, but slightly worse performance. The accuracy of R&B algorithm is much worse than other algorithms except for deeply coherent regime. In Figure 6.12(b), P_{RX} is reduced to 1 nW so that the SNR level is more relevant for actual LiDAR with various sources of insertion loss (Section 3.5). Under this setting, the performance of the LLSE and R&B was almost unchanged, but the MUSIC algorithm performed very poorly compared to the high SNR case. This is not surprising since frequency estimation algorithms based on eigendecomposition of autocorrelation matrix, including MUSIC, rely heavily on the model for additive noise and are known to be unstable in the general case [107]. For both high SNR and low SNR case, the proposed LLSE algorithm showed consistently excellent performance.

Lastly, while keeping the same P_{RX} as Figure 6.12(b), linewidth of the laser is increased to 10MHz in Figure 6.12(c). The measured PSD clearly shows that E_{LO} and E_{RX} are completely incoherent for both 10 m and 100 m cases. With such noisy laser and low SNR, the proposed LLSE is the only algorithm that can yield acceptable performance. For a 100 m target, the variance of the LLSE estimator was 4.82 cm in contrast to 56 cm of the R&B estimation, showing > 10× improvement. From this result, we can clearly see that the impact of the frequency estimation algorithm choice for the system is critical, especially assuming realistic SNR and incoherent measurements.

The experimental setup in Figure 5.7 was used again to demonstrate the efficacy of the proposed scheme. In this case, a short-cavity DBR laser diode [89] was used as our tunable laser source. This laser has very wide and fast tuning capability without significant output power variation thanks to embedded gain within the DBR section, which makes it much easier to apply advanced wavelength modulation techniques with feedback control [86]. However, its linewidth is relatively wide (> 2.6 MHz) and the coherence distance is limited to sub-meter. Chirping rate was

set to 220 THz/s and negligible dispersion from the fiber was ensured. Figure 6.2(b) shows measured spectrum of the receiver photocurrent for target distance of 110 m, and estimated linewidth was 7 MHz. The effective power of the reflected beam can also be estimated from the spectrum, which was set to be 0.1 nW. Resulting distribution of the estimated distance is shown in Fig. 2(c), where 6.6× improvement in accuracy is achieved through the proposed method.

In summary, even for relatively narrow-linewidth laser sources such as injection current-modulated external cavity lasers, the detection range of the FMCW LiDAR is eventually limited by the spectral purity of the source if the coherence distance is treated as a hard constraint. However, proper pre-characterization of the laser spectrum and detection algorithm optimization, reasonable ranging accuracy for long-range applications is indeed possible from incoherent measurement. As mentioned above, utilizing laser pre-characterization, the Lorentzian model can be replaced in the proposed LSE detection, to tailor the detection algorithm to the specific noise properties of the laser in use.

6.5 Chapter Summary

Single-frequency continuous-wave laser is the essential part of an FMCW LiDAR, but it is fundamentally hard to achieve both narrow linewidth and wide tuning range. In this chapter, I have presented a quantitative study on the reason why the laser phase noise is detrimental in both SNR-limited long-range applications and short-range, high-resolution applications. To break the trade-off between linewidth and tuning range, I have introduced both circuit-level technique (feedforward phase noise cancellation with FMCW modulation leveraging optical IQ coherent frontend) and algorithm-level technique (Lorentzian least squares-based spectral estimation to take into account the laser spectral shape) to enable FMCW measurement beyond the inherent coherence range limit. Both techniques clearly highlight the importance of "circuits and systems" perspective in electro-optic heterogeneous system design: proper utilization of available mixed-signal CMOS circuits in heterogeneous integration platforms may yield simple solutions to seemingly intractable problems in photonics.

Chapter 7

Conclusion

A comprehensive study on integrated coherent LiDAR is presented so far: the scope of the thesis included identification of system level requirements in the context of self-driving car applications, construction of quantitative framework for coherent LiDAR performance analysis and feasibility study, actual siliconization of fully-integrated coherent LiDAR system equipped with optical phased array-based solid state beam steering, evaluting the impact of laser phase noise on an FMCW LiDAR system performance and proposing possible solutions from electro-optic circuits and systems perspective.

The major takeaway points from this thesis are summarized as follows:

- Solid-state implementation of beam-steering coherent LiDAR, operating in optical C-band, appears to be the most plausible system to ultimately realize main, forward-looking LiDAR for self-driving cars (with long operation range and interference rejection for outdoor operation) at reasonable price point.

- To address long detection range and reasonable frame rate, at least a few centimeter square-scale receiving aperture is required. Millimeter square-scale aperture may enable mid-range LiDARs (<100 m target range), relevant for low-speed driving cases or side/rearview installations.

- Optical phased array is a powerful technique to realize low-cost solid-state beam scanning transmitter module, achieving steering range and lateral resolution relevant to automotive LiDARs is now becoming reality. However, further studies on ways to improve maximum power handling, temperature sensitivity, and array calibration method are needed before the OPAs to be successfully deployed in practical systems.

- With modern silicon photonics technology, it is certainly possible to realize fully integrated coherent LiDAR system on a single chip. However, realizing centimeter-scale aperture on
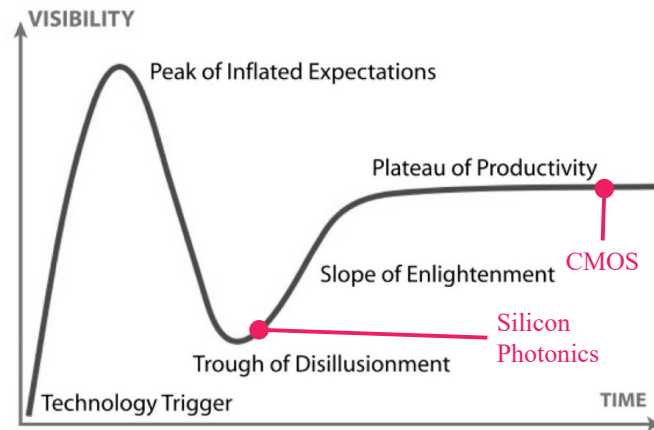
Figure 7.1: Gartner hype curve.

a chip to reach long operation range is extremely challenging. Co-design of imaging optics and on-chip optical antenna array / optical switch network may decouple the problem of on-chip complexity and effective receiving aperture.

- There is a fundamental trade-off between wavelength tuning range and linewidth in a single-frequency solid-state tunable laser for FMCW source, and it is detrimental for both SNR-limited long-range LiDARs and high-resolution depth vision applications. This trade-off can be overcome via circuit- and algorithm-level techniques, which also highlights the appeal of CMOS-assisted photonic systems.

It is certainly an exciting time for anyone working in the silicon photonics industry. As always, the main application that is driving the whole field is optical I/O: traditional electrical wireline I/Os are finally running out of steam, and having hard time meeting ever-increasing demand for bandwidth in datacenters. At the same time, we have gained much better understanding of the silicon photonics technology: beyond the initial hype, now we know exactly what it does and doesn't offer, and what it takes to actually build practical systems (Figure 7.1). Optical links are now on the verge of reaching the "last meter" [108] within the datacenter racks, and a series of silicon photonics company acquisitions by Cisco (Luxtera, Acacia) is reinforcing the impression that the *grand transition* to optical I/O, which was once promised in the early 2000s, is finally taking place. As illustrated in this thesis, the rise of autonomous vehicles has a great potential to open up another high-volume market for silicon photonics.

However, it is extremely important to always keep in mind that the silicon photonics is attrac-

tive not because it offers the best device performance or the highest reliability but because it is the only realistic platform to actually make up a "system" spanning over multiple signal domains, which can offer unique functionality as a whole. I would say that was the most important lesson from my Ph.D. study.

Bibliography

[1] "Global Semiconductor Sales Increase 13.7 Percent to $468.8 Billion in 2018," *Semiconductor Industry Association*, Feb. 2019.

[2] Y. Lee, A. Waterman, H. Cook, B. Zimmer, B. Keller, A. Puggelli, J. Kwak, R. Jevtić, S. Bailey, M. Blagojević, P. F. Chiu, R. Avizienis, B. Richards, J. Bachrach, D. Patterson, E. Alon, B. Nikolić, and K. Asanović, "An Agile Approach to Building RISC-V Microprocessors," *IEEE Micro*, vol. 36, no. 2, pp. 8–20, Mar. 2016.

[3] E. Alon, K. Asanović, J. Bachrach, and B. Nikolić, "Open-Source EDA Tools and IP, A View from the Trenches," in *2019 ACM/IEEE Design Automation Conference (DAC)*, Jun. 2019.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[5] H. Yoo, "Intelligence on Silicon: From Deep-Neural-Network Accelerators to Brain Mimicking AI-SoCs," in *2019 IEEE International Solid-State Circuits Conference (ISSCC)*, Feb. 2019, pp. 20–26.

[6] H. Qu, "CMOS MEMS Fabrication Technologies and Devices," *Micromachines*, vol. 7, no. 1, p. 14, Jan. 2016.

[7] V. Stojanović, R. J. Ram, M. Popović, S. Lin, S. Moazeni, M. Wade, C. Sun, L. Alloatti, A. Atabaki, F. Pavanello, N. Mehta, and P. Bhargava, "Monolithic Silicon-Photonic Platforms in State-Of-The-Art CMOS SOI Processes," *Optics Express*, vol. 26, no. 10, pp. 13 106–13 121, May 2018.

[8] N. P. Jouppi, C. Young, N. Patil, D. Patterson, G. Agrawal, R. Bajwa, S. Bates, S. Bhatia, N. Boden, A. Borchers, *et al.*, "In-Datacenter Performance Analysis of a Tensor Processing Unit," in *2017 ACM/IEEE Annual International Symposium on Computer Architecture (ISCA)*, 2017, pp. 1–12.

[9] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press, 2005.

[10] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, *et al.*, "Stanley: The Robot That Won the DARPA Grand Challenge," *Journal of field Robotics*, vol. 23, no. 9, pp. 661–692, 2006.

[11] Y. Wang, W.-L. Chao, D. Garg, B. Hariharan, M. Campbell, and K. Q. Weinberger, "Pseudo-LiDAR from Visual Depth Estimation: Bridging the Gap in 3D Object Detection for Autonomous Driving," in *IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2019, pp. 8445–8453.

[12] J. Lee, Y.-A. Li, M.-H. Hung, and S.-J. Huang, "A Fully-Integrated 77-GHz FMCW Radar Transceiver in 65-nm CMOS Technology," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 12, pp. 2746–2756, Dec. 2010.

[13] M. E. Warren, "Automotive LIDAR Technology," in *2019 IEEE Symposium on VLSI Circuits*, Jun. 2019, pp. 254–255.

[14] T. Kim, P. Bhargava, C. V. Poulton, J. Notaros, A. Yaacobi, E. Timurdogan, C. Baiocco, N. Fahrenkopf, S. Kruger, T. Ngai, Y. Timalsina, M. R. Watts, and V. Stojanović, "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.

[15] ——, "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.

[16] P. Bhargava, T. Kim, C. V. Poulton, J. Notaros, A. Yaacobi, E. Timurdogan, C. Baiocco, N. Fahrenkopf, S. Kruger, T. Ngai, Y. Timalsina, M. R. Watts, and V. Stojanović, "Fully Integrated Coherent LiDAR in 3D-Integrated Silicon Photonics/65nm CMOS," in *2019 IEEE Symposium on VLSI Circuits*, Jun. 2019, pp. 262–263.

[17] T. Kim, P. Bhargava, and V. Stojanović, "Optimal Spectral Estimation and System Trade-Off in Long-Distance Frequency-Modulated Continuous-Wave Lidar," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018, pp. 1583–1587.

[18] T. Kim, P. Bhargava, and V. Stojanović, "Overcoming the Coherence Distance Barrier in Long-Range FMCW LIDAR," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2018, STh3L.7.

[19] H. Sarbolandi, D. Lefloch, and A. Kolb, "Kinect Range Sensing: Structured-Light versus Time-of-Flight Kinect," *arXiv:1505.05459 [cs]*, May 2015, arXiv: 1505.05459.

[20] C. Niclass, M. Soga, H. Matsubara, M. Ogawa, and M. Kagami, "A 0.18-µm CMOS SoC for a 100-m-Range 10-Frame/s 200 × 96-Pixel Time-of-Flight Depth Sensor," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 1, pp. 315–330, Jan. 2014.

[21] P. M. Woodward, *Probability and Information Theory, with Applications to Radar*. Pergamon Press, 1953.

[22] B. Behroozpour, P. A. M. Sandborn, M. C. Wu, and B. E. Boser, "Lidar System Architectures and Circuits," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 135–142, Oct. 2017.

[23] C. Weitkamp, *Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere*. Springer Science & Business, 2006, vol. 102.

[24] A. Kadambi and R. Raskar, "Rethinking Machine Vision Time of Flight With GHz Heterodyning," *IEEE Access*, vol. 5, pp. 26 211–26 223, Nov. 2017.

[25] X. Li and Y. Liang, "Remote Measurement of Surface Roughness, Surface Reflectance, and Body Reflectance with LiDAR," *Applied Optics*, vol. 54, no. 30, p. 8904, Oct. 2015.

[26] R. Huffaker and R. Hardesty, "Remote Sensing of Atmospheric Wind Velocities Using Solid-State and CO_2 Coherent Laser Systems," *Proceedings of the IEEE*, vol. 84, no. 2, pp. 181–204, Feb. 1996.

[27] A. T. Mecherikunnel, J. A. Gatlin, and J. C. Richmond, "Data on Total and Spectral Solar Irradiance," *Applied Optics*, vol. 22, no. 9, p. 1354, May 1983.

[28] "Safety of Laser Products - Part 1: Equipment Classification and Requirements," International Electrotechnical Commission, Standard, 2014.

[29] "Failure Mechanism Based Stress Test Qualification for Integrated Circuits," Automotive Electronics Council, Standard, Sep. 2014.

[30] D. Xu, D. Anguelov, and A. Jain, "PointFusion: Deep Sensor Fusion for 3d Bounding Box Estimation," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 244–253.

[31] *Velodyne Alpha Puck™ Datasheet*, Velodyne LiDAR.

[32] *Highway Design Manual*, California Department of Transportation, Jul. 2018.

[33] B. Park, I. Park, W. Choi, and Y. Chae, "A 64 × 64 APD-Based ToF Image Sensor with Background Light Suppression up to 200 klx Using In-Pixel Auto-Zeroing and Chopping," in *2019 IEEE Symposium on VLSI Circuits*, Jun. 2019, pp. 256–257.

[34] J. R. Barry and E. A. Lee, "Performance of Coherent Optical Receivers," *Proceedings of the IEEE*, vol. 78, no. 8, pp. 1369–1394, Aug. 1990.

[35] M. Perenzoni, "Single-Photon Avalanche Diode-Based Detection and Imaging: Bringing the Photodiode Out of Its Comfort Zone," *IEEE Solid-State Circuits Magazine*, vol. 10, no. 3, pp. 26–34, 2018.

[36] S.-C. Lin, Y. Zhang, C.-H. Hsu, M. Skach, M. E. Haque, L. Tang, and J. Mars, "The Architectural Implications of Autonomous Driving: Constraints and Acceleration," in *Proceedings of the Twenty-Third International Conference on Architectural Support for Programming Languages and Operating Systems*, Williamsburg, VA, USA: ACM, Mar. 2018, pp. 751–766.

[37] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, 1st. John Wiley & Sons, Inc., 1996.

[38] D. Rife and R. Boorstyn, "Single Tone Parameter Estimation from Discrete-Time Observations," *IEEE Transactions on Information Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.

[39] L. Coldren, M. Mashanovitch, and S. Corzine, *Diode Lasers and Photonic Integrated Circuits*. John Wiley & Sons, 2012.

[40] A. Vasilyev, "The Optoelectronic Swept-Frequency Laser and Its Applications in Ranging, Three-Dimensional Imaging, and Coherent Beam Combining of Chirped-Seed Amplifiers," PhD thesis, California Institute of Technology, 2013.

[41] X. Zhang, J. Pouls, and M. C. Wu, "Laser Frequency Sweep Linearization by Iterative Learning Pre-Distortion for FMCW LiDAR," *Optics Express*, vol. 27, no. 7, p. 9965, Apr. 2019.

[42] D. Li, G. Minoia, M. Repossi, D. Baldi, E. Temporiti, A. Mazzanti, and F. Svelto, "A Low-Noise Design Technique for High-Speed CMOS Optical Receivers," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 6, pp. 1437–1447, Jun. 2014.

[43] W. Wang, C. Chan, Y. Zhu, and R. P. Martins, "A 72.6dB-SNDR 100MHz-BW 16.36mW CTDSM with Preliminary Sampling and Quantization Scheme in Backend Subranging QTZ," in *2019 IEEE International Solid-State Circuits Conference (ISSCC)*, Feb. 2019, pp. 340–342.

[44] K. Tan, X. Cheng, and X. Cheng, "Modeling Hemispherical Reflectance for Natural Surfaces Based on Terrestrial Laser Scanning Backscattered Intensity Data," *Optics Express*, vol. 24, no. 20, pp. 22 971–22 988, Oct. 2016.

[45] M. J. Byrd, C. V. Poulton, M. Khandaker, E. Timurdogan, D. Vermeulen, and M. R. Watts, "Free-space Communication Links with Transmitting and Receiving Integrated Optical Phased Arrays," in *Frontiers in Optics / Laser Science*, Sep. 2018, FTu4E.1.

[46] K. L. G. Parkin, "The Breakthrough Starshot System Model," *Acta Astronautica*, vol. 152, pp. 370–384, Nov. 2018.

[47] U. Hofmann, F. Senger, F. Soerensen, V. Stenchly, B. Jensen, and J. Janes, "Biaxial Resonant 7mm-Mems Mirror for Automotive LiDAR Application," in *2012 International Conference on Optical MEMS and Nanophotonics*, Aug. 2012, pp. 150–151.

[48] J. J. López, S. A. Skirlo, D. Kharas, J. Sloan, J. Herd, P. Juodawlkis, M. Soljačić, and C. Sorace-Agaskar, "Planar-Lens Enabled Beam Steering for Chip-scale LIDAR," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2018, SM3I.1.

[49] D. Inoue, T. Ichikawa, A. Kawasaki, and T. Yamashita, "Demonstration of a New Optical Scanner Using Silicon Photonics Integrated Circuit," *Optics Express*, vol. 27, no. 3, p. 2499, Feb. 2019.

[50] S. R. Davis, S. D. Rommel, S. Johnson, M. H. Anderson, and A. W. Yu, "Liquid Crystal Clad Waveguide Laser Scanner and Waveguide Amplifier for LADAR and Sensing Applications," in *Integrated Optics: Devices, Materials, and Technologies XIX*, Feb. 2015, 93650N.

[51] G. Takeuchi, Y. Terada, M. Takeuchi, H. Abe, H. Ito, and T. Baba, "Thermally Controlled Si Photonic Crystal Slow Light Waveguide Beam Steering Device," *Optics Express*, vol. 26, no. 9, p. 11 529, Apr. 2018.

[52] K. Van Acoleyen, W. Bogaerts, J. Jágerská, N. Le Thomas, R. Houdré, and R. Baets, "Off-Chip Beam Steering with a One-Dimensional Optical Phased Array on Silicon-On-Insulator," *Optics Letters*, vol. 34, no. 9, p. 1477, May 2009.

[53] J. Sun, E. Timurdogan, A. Yaacobi, E. S. Hosseini, and M. R. Watts, "Large-Scale Nanophotonic Phased Array," *Nature*, vol. 493, no. 7431, pp. 195–199, Jan. 2013.

[54] C. V. Poulton, M. J. Byrd, P. Russo, E. Timurdogan, M. Khandaker, D. Vermeulen, and M. R. Watts, "Long-Range LiDAR and Free-Space Data Communication With High-Performance Optical Phased Arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–8, Sep. 2019.

[55] D. N. Hutchison, J. Sun, J. K. Doylend, R. Kumar, J. Heck, W. Kim, C. T. Phare, A. Feshali, and H. Rong, "High-Resolution Aliasing-Free Optical Beam Steering," *Optica*, vol. 3, no. 8, p. 887, Aug. 2016.

[56] S. A. Miller, C. T. Phare, Y.-C. Chang, X. Ji, O. Jimenez, A. Mohanty, S. Roberts, M. C. Shin, B. Stern, M. Zadka, and M. Lipson, "512-Element Actively Steered Silicon Phased Array for Low-Power LIDAR," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2018, JTh5C.2.

[57] S. Chung, H. Abediasl, and H. Hashemi, "A Monolithically Integrated Large-Scale Optical Phased Array in Silicon-on-Insulator CMOS," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 1, pp. 275–296, Jan. 2018.

[58] R. Fatemi, A. Khachaturian, and A. Hajimiri, "A Nonuniform Sparse 2-D Large-FOV Optical Phased Array With a Low-Power PWM Drive," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 5, pp. 1200–1215, May 2019.

[59] C. V. Poulton, A. Yaacobi, D. B. Cole, M. J. Byrd, M. Raval, D. Vermeulen, and M. R. Watts, "Coherent Solid-State LiDAR with Silicon Photonic Optical Phased Arrays," *Optics Letters*, vol. 42, no. 20, p. 4091, Oct. 2017.

[60] H. Abediasl and H. Hashemi, "Monolithic Optical Phased-Array Transceiver in a Standard SOI CMOS Process," *Optics Express*, vol. 23, no. 5, p. 6509, Mar. 2015.

[61] C. A. Balanis, *Antenna Theory: Analysis and Design.* John Wiley & Sons, 2016.

[62] R. J. Mailloux, *Phased Array Antenna Handbook.* Artech House, 2017.

[63] L. Chrostowski, X. Wang, J. Flueckiger, Y. Wu, Y. Wang, and S. T. Fard, "Impact of Fabrication Non-Uniformity on Chip-Scale Silicon Photonic Integrated Circuits," in *2014 Optical Fiber Communication Conference and Exposition (OFC)*, Mar. 2014, pp. 1–3.

[64] Y. Yang, Y. Ma, H. Guan, Y. Liu, S. Danziger, S. Ocheltree, K. Bergman, T. Baehr-Jones, and M. Hochberg, "Phase Coherence Length in Silicon Photonic Platform," *Optics Express*, vol. 23, no. 13, p. 16 890, Jun. 2015.

[65] A. Yaacobi, J. Sun, M. Moresco, G. Leake, D. Coolbaugh, and M. R. Watts, "Integrated Phased Array for Wide-Angle Beam Steering," *Optics Letters*, vol. 39, no. 15, pp. 4575–4578, Aug. 2014.

[66] C. McDonough, D. La Tulipe, D. Pascual, P. Tariello, J. Mucci, M. Smalley, A. Nguyen, Tuan Vo, C. Johnson, P. Nguyen, J. Hebding, G. Leake, M. Moresco, E. Timurdogan, V. Stojanović, M. R. Watts, and D. Coolbaugh, "Heterogeneous Integration of a 300-mm Silicon Photonics-CMOS Wafer Stack by Direct Oxide Bonding and Via-Last 3-D Interconnection," *Journal of Microelectronic & Electronic Packaging*, vol. 13, no. 2, pp. 71–76, Apr. 2016.

[67] F. Boeuf, S. Crémer, N. Vulliet, T. Pinguet, A. Mekis, G. Masini, L. Verslegers, P. Sun, A. Ayazi, N.-. Hon, *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 (IEDM)*, Dec. 2013, pp. 13.3.1–13.3.4.

[68] C. T. DeRose, N. J. Martinez, R. D. Kekatpure, W. A. Zortman, A. L. Starbuck, A. Pomerene, and A. L. Lentine, "Thermal Crosstalk Limits for Silicon Photonic DWDM Interconnects," in *2014 Optical Interconnects Conference (OIC)*, May 2014, pp. 125–126.

[69] M. Raval, C. V. Poulton, and M. R. Watts, "Unidirectional Waveguide Grating Antennas with Uniform Emission for Optical Phased Arrays," *Optics Letters*, vol. 42, no. 13, p. 2563, Jul. 2017.

[70] C. Sun, M. Georgas, J. Orcutt, B. Moss, Y.-H. Chen, J. Shainline, M. Wade, K. Mehta, K. Nammari, E. Timurdogan, D. Miller, O. Tehar-Zahav, Z. Sternberg, J. Leu, J. Chong, R. Bafrali, G. Sandhu, M. Watts, R. Meade, M. Popović, R. Ram, and V. Stojanović, "A Monolithically-Integrated Chip-to-Chip Optical Link in Bulk CMOS," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 4, pp. 828–844, Apr. 2015.

[71] C. V. Poulton, M. J. Byrd, E. Timurdogan, P. Russo, D. Vermeulen, and M. R. Watts, "Optical Phased Arrays for Integrated Beam Steering," in *2018 IEEE 15th International Conference on Group IV Photonics (GFP)*, Aug. 2018, pp. 1–2.

[72] J. C. Hulme, J. K. Doylend, M. J. R. Heck, J. D. Peters, M. L. Davenport, J. T. Bovington, L. A. Coldren, and J. E. Bowers, "Fully Integrated Hybrid Silicon Two Simensional Beam Scanner," *Optics Express*, vol. 23, no. 5, p. 5861, Mar. 2015.

[73] C. V. Poulton, M. J. Byrd, M. Raval, Z. Su, N. Li, E. Timurdogan, D. Coolbaugh, D. Vermeulen, and M. R. Watts, "Large-Scale Silicon Nitride Nanophotonic Phased Arrays at Infrared and Visible Wavelengths," *Optics Letters*, vol. 42, no. 1, p. 21, Jan. 2017.

[74] J. Chen, "Advanced Architectures for Efficient mm-Wave CMOS Wireless Transmitters," PhD thesis, EECS Department, University of California, Berkeley, 2014.

[75] H. K. Tsang, C. S. Wong, T. K. Liang, I. E. Day, S. W. Roberts, A. Harpin, J. Drake, and M. Asghari, "Optical Dispersion, Two-Photon Absorption and Self-Phase Modulation in Silicon Waveguides at 1.5 µm Wavelength," *Applied Physics Letters*, vol. 80, no. 3, pp. 416–418, Jan. 2002.

[76] B. J. Isaac, B. Song, S. Pinna, L. A. Coldren, and J. Klamkin, "Indium Phosphide Photonic Integrated Circuit Transceiver for FMCW LiDAR," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 6, pp. 1–7, Apr. 2019.

[77] M. A. Tran, T. Komljenovic, D. Huang, L. Liang, M. Kennedy, and J. E. Bowers, "A Widely-Tunable High-SMSR Narrow-Linewidth Laser Heterogeneously Integrated on Silicon," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2018, AF1Q.2.

[78] D. Huang, M. A. Tran, J. Guo, J. Peters, T. Komljenovic, A. Malik, P. A. Morton, and J. E. Bowers, "High-Power Sub-kHz Linewidth Lasers Fully Integrated on Silicon," *Optica*, vol. 6, no. 6, pp. 745–752, Jun. 2019.

[79] J. Notaros, N. Li, C. V. Poulton, Z. Su, M. J. Byrd, E. S. Magden, and M. R. Watts, "CMOS-Compatible Optical Phased Arrays with Monolithically-Integrated Erbium Lasers," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2018, STu4B.2.

[80] N. Li, Purnawirman, Z. Su, E. Salih Magden, P. T. Callahan, K. Shtyrkova, M. Xin, A. Ruocco, C. Baiocco, E. P. Ippen, F. X. Kärtner, J. D. B. Bradley, D. Vermeulen, and M. R. Watts, "High-Power Thulium Lasers on a Silicon Photonics Platform," *Optics Letters*, vol. 42, no. 6, pp. 1181–1184, Mar. 2017.

[81] N. Li, D. Vermeulen, Z. Su, E. S. Magden, M. Xin, N. Singh, A. Ruocco, J. Notaros, C. V. Poulton, E. Timurdogan, C. Baiocco, and M. R. Watts, "Monolithically Integrated Erbium-Doped Tunable Laser On a CMOS-Compatible Silicon Photonics Platform," *Optics Express*, vol. 26, no. 13, pp. 16 200–16 211, Jun. 2018.

[82] R. W. Gerchberg and W. O. Saxton, "A Practical Algorithm for the Determination of the Phase from Image and Diffraction Plane Pictures," *Optik*, vol. 35, pp. 237–246, May 1972.

[83] M. J. Byrd, E. Timurdogan, Z. Su, C. V. Poulton, N. M. Fahrenkopf, G. Leake, D. D. Coolbaugh, and M. R. Watts, "Mode-Evolution-Based Coupler for High Saturation Power Ge-On-Si Photodetectors," *Optics Letters*, vol. 42, no. 4, p. 851, Feb. 2017.

[84] F. Aflatouni, B. Abiri, A. Rekhi, and A. Hajimiri, "Nanophotonic Coherent Imager," *Optics Express*, vol. 23, no. 4, pp. 5117–5125, Feb. 2015.

[85] X. Chen and H. K. Tsang, "Polarization-Independent Grating Couplers for Silicon-On-Insulator Nanophotonic Waveguides," *Optics Letters*, vol. 36, no. 6, p. 796, Mar. 2011.

[86] B. Behroozpour, P. A. M. Sandborn, N. Quack, T. J. Seok, Y. Matsui, M. C. Wu, and B. E. Boser, "Electronic-Photonic Integrated Circuit for 3D Microimaging," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 1, pp. 161–172, Jan. 2017.

[87] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Prentice Hall, 2005.

[88] N. Fujiwara, T. Kakitsuka, M. Ishikawa, F. Kano, H. Okamoto, Y. Kawaguchi, Y. Kondo, Y. Yoshikuni, and Y. Tohmori, "Inherently Mode-Hop-Free Distributed Bragg Reflector (DBR) Laser Array," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 9, no. 5, pp. 1132–1137, Sep. 2003.

[89] H. Arimoto, T. Kitatani, T. Tsuchiya, K. Shinoda, A. Takei, H. Uchiyama, M. Aoki, and S. Tsuji, "Wavelength-Tunable Short-Cavity DBR Laser Array With Active Distributed Bragg Reflector," *Journal of Lightwave Technology*, vol. 24, no. 11, pp. 4366–4371, Nov. 2006.

[90] V. Jayaraman, Z. Chuang, and L. A. Coldren, "Theory, Design, and Performance of Extended Tuning Range Semiconductor Lasers with Sampled Gratings," *IEEE Journal of Quantum Electronics*, vol. 29, no. 6, pp. 1824–1834, Jun. 1993.

[91] Y. Kotaki, M. Matsuda, M. Yano, H. Ishikawa, and H. Imai, "1.55 µm Wavelength Tunable FBH-DBR Laser," *Electronics Letters*, vol. 23, no. 7, pp. 325–327, Mar. 1987.

[92] B. P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Physical Review Letters*, vol. 116, no. 6, p. 061 102, 2016.

[93] M. H. Idjadi and F. Aflatouni, "Integrated Pound-Drever-Hall Laser Stabilization System in Silicon," *Nature Communications*, vol. 8, no. 1, p. 1209, 2017.

[94] D. T. Spencer, T. Drake, T. C. Briles, J. Stone, L. C. Sinclair, C. Fredrick, Q. Li, D. Westly, B. R. Ilic, A. Bluestone, *et al.*, "An Optical-Frequency Synthesizer Using Integrated Photonics," *Nature*, vol. 557, no. 7703, p. 81, 2018.

[95] E. D. Black, "An Introduction to Pound-Drever-Hall Laser Frequency Stabilization," *American Journal of Physics*, vol. 69, no. 1, pp. 79–87, 2001.

[96] F. Aflatouni, M. Bagheri, and H. Hashemi, "Design Methodology and Architectures to Reduce the Semiconductor Laser Phase Noise Using Electrical Feedforward Schemes," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 11, pp. 3290–3303, Nov. 2010.

[97] N. Satyan, A. Vasilyev, G. Rakuljic, V. Leyva, and A. Yariv, "Precise Control of Broadband Frequency Chirps Using Optoelectronic Feedback," *Optics express*, vol. 17, no. 18, pp. 15 991–15 999, 2009.

[98]   B. Razavi, "Design of Monolithic Phase-Locked Loops and Clock Recovery Circuits-A Tutorial," *Monolithic Phase-locked loops and Clock recovery circuits: Theory and Design*, pp. 1–39, 1996.

[99]   J. Sharma, S. Ahasan, C. T. Phare, M. Lipson, and H. Krishnaswamy, "Continuous-Time Electro-Optic PLL with Decimated Optical Delay/Loss and Spur Cancellation for LiDAR," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2018, JTu2A.81.

[100] S. Ahasan, A. Binaie, C. T. Phare, M. Lipson, and H. Krishnaswamy, "A Compact, Low Loss Integrated Continuous-Time Electro Optic-PLL with Maximum Range of > 3.3 m," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2019, SM1N.6.

[101] S. Bottacchi, A. Beling, A. Matiss, M. L. Nielsen, A. G. Steffan, G. Unterbörsch, and A. Umbach, "Advanced Photoreceivers for High-Speed Optical Fiber Transmission Systems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1099–1112, Sep. 2010.

[102] J. Zheng, "Coherence Analysis of Optical Frequency-Modulated Continuous-Wave Interference," *Applied Optics*, vol. 45, no. 16, p. 3681, 2006.

[103] P. A. M. Sandborn, N. Kaneda, Y.-K. Chen, and M. C. Wu, "Dual-Sideband Linear FMCW Lidar with Homodyne Detection for Application in 3d Imaging," in *Conference on Lasers and Electro-Optics (CLEO)*, Jun. 2016.

[104] A. Barbieri, D. Bolletta, and G. Colavolpe, "On the Cramer-Rao bound for carrier frequency estimation in the presence of phase noise," in *2005 IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 3, Nov. 2005.

[105] G. M. Stéphan, T. T. Tam, S. Blin, P. Besnard, and M. Têtu, "Laser Line Shape and Spectral Density of Frequency Noise," *Physical Review A*, vol. 71, no. 4, Apr. 2005.

[106] R. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[107] V. I. Morgenshtern and E. J. Candès, "Super-Resolution of Positive Sources: The Discrete Setup," *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, pp. 412–444, Jan. 2016.

[108] "Silicon Photonics Stumbles at the Last Meter," *IEEE Spectrum*, Aug. 2018.