Si Photonics for Practical LiDAR Solutions

Xiaochen Sun *, Lingxuan Zhang, Qihao Zhang and Wenfu Zhang

Chinese Academy of Sciences, Xi’an Institute of Optics and Precision Mechanics, Xi’an 710119, China; zhanglingxuan@opt.cn (L.Z.); zhangqihao2016@opt.cn (Q.Z.); wfuzhang@opt.ac.cn (W.Z.)

Received: 30 June 2019; Accepted: 5 September 2019; Published: 10 October 2019

Abstract: In the article the authors discuss light detection and ranging (LiDAR) for automotive applications and the potential roles Si photonics can play in practice. The authors review published research work on Si photonics optical phased array (OPA) and other relevant devices in the past decade with in-depth technical analysis with respect to practical system design considerations. The commercialization status of certain LiDAR technologies is briefly introduced.

Keywords: Si photonics; OPA; LiDAR; ToF; FMCW

1. Introduction

Silicon (Si) photonics have emerged as one of the so-called “more-than-Moore” technologies that are based upon or derived from Si complementary metal-oxide-semiconductor (CMOS) and are focused on application and diversity rather than transistor scaling. Many integrated circuit (IC) and communication companies have been heavily invested in Si photonics for over a decade, not long after the academic world put it under the spotlight. Thanks to extensive research and development on SiGe epitaxy for bipolar CMOS (BiCMOS) and silicon-on-insulator (SOI) wafers for radio frequency IC (RFIC) and partially or fully depleted CMOS, most Si photonics components can be fabricated using slightly modified SOI processes with the addition of pure Ge epitaxy and associated metallization processes. High compatibility with CMOS helps Si photonics to create a fast growing fabless-foundry ecosystem that has been repeatedly proven a key to success by the cases like image sensors. Si photonics has achieved its early success in commercialization by enabling high-volume optical transceiver modules for both data center and long-haul optical networks.

While people realize that Si photonics has its limitations and may co-exist with its fierce InP competitors for many years to come, many IC, optics and communication companies are embracing this technology and many released products in recent years. On the foundry side, starting with a couple of research or pilot fabrication lines in imec (Belgium), Institute of Microelectronics (IME) (Singapore), American Institute of Manufacturing (AIM) Photonics (US), Chinese Academy of Sciences—Institute of Microelectronics (CAS-IME) (China), etc., many semiconductor foundries have released their Si photonics commercial offerings, including GlobalFoundries (Santa Clara, CA, USA), STMicroelectronics (Geneva, Switzerland), TowerJazz (Migdal HaEmek, Israel), Taiwan Semiconductor Manufacturing Company (TSMC) (Hsinchu, Taiwan), Advanced Micro Foundry (AMF) (Singapore), etc.

As Si photonics has made its entry into the optical communications market and grasps sizable shares, the designs and processes are maturing and standardizing and the optical transceiver products it enables are increasingly competing on manufacturability, cost, yield, and supply chain instead of technological novelty. Many Si photonics researchers and start-ups are looking into other applications such as optical sensing, among which, light detection and ranging (LiDAR) sensors have enjoyed a lot attention in recent years thanks to the burgeoning autonomous driving market. While there have been many review articles on Si photonics for optical communication applications [1–5], there are fewer for
LiDAR [6], in particular, with the focus on practical implementation. The authors attempt to capture the published results of this rapidly developing field in last decade with an emphasis on practical aspects for commercialization than scientific novelty. The authors understand that it is very challenging, if not impossible, to accomplish a complete survey of such an actively developing field and we apologize in advance to those researchers and institutions whose work we have inadvertently overlooked.

2. An Overview of LiDAR Technologies

Before diving into the details of Si photonics light detection and ranging (LiDAR), we’d like to clarify some terminologies used in this article. The general application we discuss here is 3D sensing, which is about acquiring the three-dimensional information of a scene often with point cloud data constructed from a depth map of a two-dimensional image. Presently, 3D sensing can be realized by a variety of approaches including, but not limited to, interferometry, triangulation, stereo vision, structured light reconstruction, and LiDAR. LiDAR emphasizes the use of a light wave, laser light in many cases, to illuminate a scene in order to calculate depth information from the reflection. Many LiDAR solutions involve using direct or indirect time-of-flight (ToF) for distance calculation so these two terms sometimes are used interchangeably. The LiDAR term is coined from radar and its adoption began with long-range applications such as terrestrial mapping. The LiDAR term tends to be associated with relatively long-range applications, such as autonomous driving, while the ToF camera or the ToF sensor term is more commonly used for short-range applications such as facial recognition and augmented reality. The distinction between these terms may be vague and arbitrary. In this article, we refer LiDAR to a group of 3D sensing methods using light wave to generate depth maps while restrict ToF to a group of distance calculation methods.

We coarsely categorize a LiDAR solution by two main design dimensions: the distance calculation method and the scene mapping method. Table 1 summarizes our knowledge on the maturity of some popular LiDAR solutions for autonomous driving, advanced driver-assistance systems (ADAS) and consumer 3D sensing. These distance calculation methods generally leverage decades of radar development with adjustment for optical wavelength operation.

<table>
<thead>
<tr>
<th>Distance Calculation Method</th>
<th>Time-of-Flight</th>
<th>FM</th>
<th>FMCW (Optical Freq. Chirp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>Pulsed</td>
<td>AM (RF Phase Shift)</td>
<td>FM (RF Freq. Chirp)</td>
</tr>
<tr>
<td>Indirect (Modulated CW)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.** Status of popular LiDAR solutions for autonomous driving, ADAS and consumer 3D sensing with respective to the design choices in distance calculation and scene mapping.

2.1. Distance Calculation Methods

Direct ToF calculates distance from the travel time of short light pulses reflected by targets. It is a straightforward approach and is by far the most widely used approach for automotive LiDAR. It can measure a long range from hundreds of meters for automotive applications to tens of kilometers for aerial or terrestrial mapping applications. It has a high ambient light rejection thanks to the short pulse duration, e.g. several to tens of nanoseconds, and high peak optical power. It is usually paired with a high sensitivity receiver therefore gains significant momentum as Si CMOS based single-photon avalanche detector (SPAD) matures in recent years. The light pulse travel time can be calculated by a
time-correlated Geier mode detection method [7] or by a coincidence detection method which better tackles a premature sampling issue [8].

Indirect ToF refers to a group of approaches using continuous optical wave with certain modulation and sampling schemes. In an amplitude modulation (AM) scheme, the amplitude of an optical wave is modulated at a frequency so that the corresponding modulation period is longer than the round-trip travel time and the distance is calculated from the phase shift of reflected light. This approach is sometimes called AMCW, phase-shift ToF, or phase-comparison ToF. The phase shift of reflected light is determined by either directly sampling the received photocurrent or, more commonly, by integrating the photocurrent within gated period synchronized and shifted by every $\pi$ (or $\pi/2$), from the source signal. The latter method is essentially a demodulation process therefore it is also called homodyne indirect ToF. In general, the AM scheme has limited unambiguous range corresponding to the modulation period though the range can be extended by using more than one modulation frequencies to unwrap ambiguous phases. The modulation frequency can also be chirped in more steps so that a beat frequency is generated with better signal-to-noise ratio (SNR). It is called stepped frequency (SF) CW [9] or, less often, frequency modulation (FM) CW which must not be confused with a different FMCW approach described below.

FMCW is often referred to a scheme using optical wave frequency chirp, i.e., wavelength varying over time, to calculate the distance from the beat frequency generated from mixing the emitting and the reflected light. It essentially uses a local optical oscillator, often from the same light source, as a low noise optical amplifier to achieve a better SNR than what can be obtained in RF domain. Therefore, it can operate at much lower optical power for a given range than most other approaches. On the other hand, it requires a narrow linewidth laser for good temporal coherence within a targeted range, a good linear chirping and receiving system and demanding compute power for signal processing. These requirements currently make a low-cost solution challenging.

2.2. Mapping Methods

In terms of scene mapping, mechanical scanning and microelectromechanical system (MEMS) mirror scanning share a common feature that using macro or micro moving parts to steer an optical beam to different directions. Mechanical scanning spins the whole or the part of an emitter subsystem along one or two axes for up to $360^\circ$ to scan a scene in one or two dimensions. It may use multiple pairs of emitters and receivers for mapping the other dimension when the spinning scans a single dimension. In MEMS mirror scanning, light is reflected by a Si based micro mirror vibrating in one or two directions, often with limited deflection angles.

Mechanical scanning is the first commercialized and currently dominant LiDAR technology for automotive applications though it is often criticized for poor scalability, high cost and frequent re-calibration. MEMS scanning presents a potentially lower cost alternative as the MEMS mirror could follow a cost reduction path like other widely used MEMS chips. MEMS mirrors have been used for optical projection and switching applications for a long time, however, the MEMS mirror for LiDAR applications present many new challenges and its design must carefully balance the intrinsic trade-off among mirror size (for beam divergence), deflection angle (for steering range) and mirror weight (for frame rate). Many companies are diligently working toward bringing MEMS mirror-based LiDAR products to automotive market. The automotive grade reliability of MEMS mirrors, a great concern for many, is being field-tested at the time of writing. Both mechanical and MEMS scanning can handle high optical power so they both are popular choices for pulsed ToF with tens to hundreds of watts peak power for ranging hundreds of meters. Mechanically scanned multi-channel direct ToF products have already been adopted in the ADAS of luxury passenger vehicles since 2017 [10], in many robotic vehicles, and in autonomous test vehicle fleets [11].

Another way of mapping is to image a scene with a camera lens system and an array of SPAD or photodiode (PD)-based depth sensor pixels. It is a true solid-state approach with a straightforward concept as well as support from a mature camera ecosystem. ToF cameras combining this approach
with indirect phase-shift ToF are recently penetrating smartphone market for artificial reality (AR) and other applications. Most applications in automotive or consumer 3D sensing require hundreds or more pixels in at least one dimension. The requirements of large pixel area for good SNR and/or pixel-wise ToF read-out circuits currently restrict the pixel size to about 10–20 μm which is an order of magnitude larger than today’s color camera sensor pixel sizes. Nevertheless, the industry is continuously improving the sensor technology to enable automotive LiDAR companies to deploy such so-called flash LiDAR. For automotive applications, such imaging approach requires high peak optical power as it must illuminate a whole scene at far distances with sufficient reflected light at each pixel. Ultra-short pulses, on the order of nanosecond, must be used in such cases to comply with eye-safety requirement.

The optical phased array (OPA) approach is another true solid-state method for scene mapping. It generally resembles a phased array radar that uses an array of phase tunable antennae to form the shape of an output beam and steer its direction, while operating at optical frequency instead. In theory, with a large number of antennae, an OPA can generate arbitrary beam shapes and steer a beam in 2D without a lens system. In practice, however, design complexity, power consumption and cost often limit the number of phase shifters and antennae thus the optical performance. As an OPA comprises a large number of optical elements so it only returns to people’s attention as the rise of Si photonics which promises manufacturing such complex optical system on a chip at low cost.

3. Si Photonics for LiDAR

Si photonics is one type of integrated photonics which often implies monolithic integration of more than one optical element or function on a single chip. The first commercially successful integrated photonics product may be the InP based EML which stands for externally modulated laser or electro-absorption modulated laser. It was developed in the late 1980s [12], commercialized after about 10 years, and is now widely used in many telecom applications. The rise of Si photonics greatly accelerates the pace of photonic integration in terms of integration scale as well as applications.

Most Si photonics LiDAR solutions focus on automotive applications by providing certain unique value propositions. Therefore, our following analysis is mostly based on the requirement of autonomous driving and ADAS applications. Table 2 summarizes some essential specifications, proposed by a LiDAR company [13], desired by Level 4 and 5 autonomy (i.e., no human intervention) which is defined by the international Society of Automotive Engineers (SAE) J3016.

<table>
<thead>
<tr>
<th>Items</th>
<th>Specs</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>&gt;200 m</td>
<td>10% reflectivity</td>
</tr>
<tr>
<td>Horizontal FOV</td>
<td>&gt;120°</td>
<td>/</td>
</tr>
<tr>
<td>Vertical FOV</td>
<td>&gt;25°</td>
<td>/</td>
</tr>
<tr>
<td>Angular resolution</td>
<td>&lt;0.1°</td>
<td>/</td>
</tr>
<tr>
<td>Frame rate</td>
<td>&gt;15 Hz</td>
<td>/</td>
</tr>
</tbody>
</table>

The 200 m range is estimated from the braking distance of about 100 m at 70 miles per hour (MPH) on wet pavement road, computer reaction distance and some margins. Some argue 300 m may be needed for heavily loaded tracks or slippery road while some adopt less than 200 m for low speed vehicles, e.g., shuttle buses, or certain ADAS applications. The field of view (FOV) specifications listed in Table 2 consider the coverage of a wide range of distances. Some believe narrower FOVs can be tolerated if the LiDAR is used for long range only while some insist a full 360° horizontal coverage. The 0.1° angular resolution results from the requirement of resolving a human body size at 200 m. The 15 Hz frame rate is an average of required frame rate range to track moving objects in certain road conditions [14]. The industry is at its very early stage of developing and deploying LiDAR in
automotive and is far from reaching agreement on these specifications. Nevertheless, we use the values in Table 2 as a general guideline in following discussions.

3.1. Implementations of Si Photonics in LiDAR

The simple idea behind Si photonics is to replace a complex free space optical system made of many discrete optical components with a chip scale solution on Si platform. A complex optical system for LiDAR is apparently a perfect target to tackle. Among all the mapping approaches described earlier, the OPA may be the mostly suited to leverage of the benefit of Si photonics. The concept of an OPA is not new and chip-scale OPA research activities started with electro-optic medium such as liquid crystal in early 1990s [15,16] but lead to no commercial success. With the help of Si photonics researchers and entrepreneurs resurge the adoption of the OPA which is analyzed in details in Section 4.

Besides OPA, Si photonics might find other uses in a LiDAR system. The FMCW method requires a coherent receiver for mixing a local optical oscillator with the optical signal reflected from objects. It is nearly identical to a phase-shift keying (PSK) coherent receiver that is widely used for long-haul optical networks. Si photonics has been adopted by both new and incumbent optical companies to make such PSK coherent receivers with increasing market shares. Adopting Si photonics for such receivers is naturally considered for FMCW LiDAR and is discussed in Section 5.1. Si photonics can also be used for making narrow linewidth tunable lasers as the light source for FMCW LiDAR which is discussed in Section 5.2.

3.2. LiDAR System Design Considerations

Designing a Si photonics LiDAR for automotive application requires the careful consideration of many aspects of the system. We discuss some of them and analyze the pros and cons of different materials, wavelengths and components and their impact on the system performance.

3.2.1. Eye Safety

Eye safety must be kept in mind when choosing a wavelength, optical power and beam shape. Eye safety ultimately limits the range of most civilian LiDAR systems can operate. International standard International Electrotechnical Commission (IEC) 60825 exists for regulating laser safety and is generally in consensus with other related standards such as American National Standards Institute (ANSI) Z136. Most LiDAR systems for automotive and consumer applications adopt an infrared laser as the light source complying with IEC 60825 Class 1 as they are used in an environment where human eyes are exposed. Infrared lasers are particularly hazardous, since body’s protective glare aversion response, or “blink reflex,” is triggered only by visible light.

The two groups of wavelengths that are commonly used for automotive or consumer LiDAR are the 800–1000 nm range and the 1300–1600 nm range, within the range of near infrared (NIR) and short-wave infrared (SWIR). The former group takes advantage of Si photo sensors and camera optics while the latter group leverages components and technologies of fiber optic telecom industry. Light in the former group can cause retinal damage at relatively low energy as an eye is transparent and well focuses light at these wavelengths. Light in the latter group of wavelengths is largely absorbed in the crystalline lens (1300–1400 nm) or the vitreous body (1400–1600 nm) of an eye so more energy is allowed. IEC 60825 defines Class 1 lasers as optical energy below the maximum permissible exposure (MPE) for a certain type of a laser operating at a certain wavelength with certain exposure time. The MPE of some commonly used laser wavelengths at different exposure time is calculated and shown in Figure 1.

The MPE is a few orders of magnitude higher at 1550 nm than at 800–1000 nm range therefore 1550 nm is sometimes called “eye-safe” wavelength. The MPE at 1300 nm range generally falls between above two groups below 0.1 s exposure and surpasses other wavelengths when approaching continuous wave. Long range LiDAR systems using continuous wave, e.g., FMCW, tend to adopt “eye-safe” long wavelength like 1550 nm while those using pulsed light, e.g., direct ToF, mostly adopt short wavelengths for other merits such as availability of light sources and detectors.
The implementation of most Si photonics products is fabricated on silicon-on-insulator (SOI) wafers, which was originally invented as a better substrate for low power CMOS and RF applications. The type of SOI wafers used in Si photonics is with relatively thick device Si and buried oxide (BOX) layers which is closer to that for RF-CMOS than that for partially or fully depleted transistor CMOS. The device Si can act as a waveguiding core as it has larger refractive index than its surrounding silicon dioxide (SiO$_2$) or silicon nitride (SiN) materials. The material platform, most notably specified by the thickness of the device silicon, has gradually come to a consensus of around 220 nm [17–21] though some variations exist for better integration with transistors [22] or for more tolerable optical performance [23,24]. SiN, commonly found in many CMOS processes, can also act as a waveguiding material in Si photonics. The thickness of SiN layers varies a lot, from a few hundred nanometers to few microns [25–27]. In this article we exclude sub-100 nm thick SiN (for ultra-low loss applications [28]) or few µm thick Si (for better polarization and process tolerance management [29]) because they are currently lack of a comparable manufacturing maturity or with vast different design rules that complicates following quantitative analysis. We summarize and compare both waveguiding materials in Table 3 with some typical designs and properties relevant to our discussions. The SiN used for waveguides is generally not crystalline materials and its properties vary with deposition processes so some data exhibit a range.

### Table 3. A comparison of typical designs and properties between Si and SiN waveguide with SiO$_2$ claddings.

<table>
<thead>
<tr>
<th>Transparency µm</th>
<th>Thickness µm</th>
<th>Loss dB/cm</th>
<th>Bending µm</th>
<th>$\beta_{TO}$/K</th>
<th>$n2$ cm$^2$/GW</th>
<th>$\beta_{TPA}$ cm/GW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>1.1–3.7</td>
<td>0.2–0.5$^1$</td>
<td>1–3$^1$</td>
<td>5–50$^1$</td>
<td>$1.8 \times 10^{-42}$</td>
<td>$5 \times 10^{-52}$</td>
</tr>
<tr>
<td>SiN</td>
<td>0.5–3.7</td>
<td>0.2–2$^1$</td>
<td>0.2–2$^1$</td>
<td>20–200$^1$</td>
<td>$2–4 \times 10^{-52}$</td>
<td>$3–7 \times 10^{-62}$</td>
</tr>
</tbody>
</table>

$^1$ These ranges are coarsely summarized from some but not all data in Ref. [17–21,25–27].
$^2$ These data are obtained from Ref. [30–37] at around 1550 nm wavelength.

Si waveguide transmits light with wavelength above its optical bandgap at around 1.1 µm while SiN is transparent in both visible and near infrared. In most manufacturable Si photonics processes, Si and SiN waveguides are surrounded by SiO$_2$ as cladding which exhibits absorption starting from about 3.7 µm and effectively introduces the upper wavelength limit for both types of waveguides. LiDAR chips using Si waveguides often adopts a telecom wavelength band such as...
C-band (1530–1565 nm) or O-band (1260–1360 nm) to leverage the existing process design kit (PDK) of foundries, a mature supply chain, and product development experience.

The transparency of SiN in the 800–1100 nm allows possible LiDAR designs in this particularly interesting range leveraging mature light sources, Si sensors and imaging optics. On the other hand, SiN is lack of an efficient phase tuning mechanism. Thermal optics tuning is currently the dominant method for both Si and SiN whereas SiN’s thermo-optic coefficient ($\beta_{TO}$) is nearly an order of magnitude smaller than Si’s which means greater power consumption and a larger phase shifter.

The propagation loss and the bending radius data of Si and SiN waveguides in Table 3 are coarse ranges nevertheless they are good measures to evaluate the scalability of a design, i.e., how many optical elements can be packed on chip for a given size. Propagation loss is related to fabrication processes and can be moderately improved by using advanced lithography with better line edge roughness (LER). Thanks to a maturing ecosystem, these and many other data, provided by foundries’ PDK [17,20], become increasingly rich and statistically predictable. It allows a robust Si photonics product, e.g., a LiDAR chip, design which must consider tolerance and margins based on these data.

The waveguides for LiDAR often transmit optical waves at very high power at which non-linear effects cannot be neglected. The Kerr effect, or quadratic electro-optic (QEO) effect, corresponding to the real part or the imaginary part of the nonzero third order nonlinear susceptibility, is responsible for many non-linear behaviors in waveguides such as self-phase modulation (SPM), cross-phase modulation (XPM), four-wave mixing (FWM) and two-photon absorption (TPA). Other nonlinear effects also exist such as stimulated Brillouin scattering (SBS) and stimulated Raman scattering (SRS) due to photon-phonon interactions and molecular vibrational or rotational transition respectively.

TPA may be the first one come to mind when people worries the impact of non-linear effects on their Si photonics LiDAR chips. Si has a relatively large TPA coefficient i.e., $\beta_{TPA}$ as listed in Table 3 and the corresponding propagation loss can be described as [32]

$$10 \log_{10}(e) \frac{\beta_{TPA}}{A_{eff}},$$

where $A_{eff}$ is the effective mode size. Taking a $220 \times 500$ nm$^2$ Si waveguide at 1550 nm for example, the loss is about 2 dB/cm/W. Therefore, TPA introduces non-negligible optical loss when transmitting light above hundreds of mW and may limit the implementation of direct ToF requiring very high peak optical power. TPA may be greatly mitigated by sweeping the induced free carriers away from waveguide with electric field, e.g., generated by a p-i-n diode [38], therefore orders of magnitude higher optical power may be tolerated. TPA in SiN has so far not been observed.

SPM introduces optical intensity dependent phase shift which may be an issue of OPA designs relying on specific phase relations among antennae. For pulsed light operation, the induced phase shift varies with time leads to chirping which may cause pulse broadening at the presence of dispersion in waveguides. The instability of such phase modulation also introduces noise to the detected signals. A semi-quantitative way to evaluate this effect is by examining a nonlinear coefficient representing the phase shift per unit propagation length per unit power [39]

$$\gamma = \frac{2\pi n_2}{\lambda A_{eff}},$$

where $\lambda$ is wavelength and $n_2$ is Kerr coefficient of refractive index listed in Table 3. Taking the same Si waveguide for example $\gamma$ is on the order of 1 cm$^{-1}$W$^{-1}$. Therefore, this phase shift is insignificant for a mm size chip with continuous wave operation which seldomly goes beyond 1 W but is possibly an issue for pulsed light with high peak power. The $n_2$ of SiN is an order of magnitude smaller and SiN waveguides usually have larger mode size so the power constraint due to this effect may be relaxed by 10–100 times.
SRS may also be taken into consideration in Si waveguides. The effect of SRS can be analyzed by the Raman pumping threshold, which, under Lorentzian gain spectrum assumption, is estimated to be [36]

\[
\frac{16A_{eff}}{\delta g_{R}} \sim \frac{16\alpha A_{eff}}{\delta g_{R}},
\]

where \(g_R\) is the Raman gain coefficient and \(\alpha\) is propagation loss. Taking the same Si waveguide with 1 dB/cm loss for example, with a measured \(g_R \sim 7.5 \text{ cm GW}^{-1}\) [38], the threshold is about 1 W.

The effect of SBS for a commonly used non-suspended waveguide is generally minimal compared to other nonlinear effects described above.

In conclusion, SiN waveguides present better flexibility in wavelength choice and allows much higher optical power without significant nonlinear effect but suffers from an order of magnitude poorer thermo-optic phase tuning power efficiency. In comparison, Si waveguides allow efficient thermo-optic phase tuning (and other tuning mechanisms discussed in Section 4.2.3) but do not work in NIR wavelength or at optical power larger than a few hundred mW (for typical 220 nm SOI waveguides). The scalability, in terms of size and loss budget, of Si waveguides is generally better than that of SiN waveguides due to smaller bending radius while similar propagation loss.

3.2.3. Lasers and Photodetectors

The availability of high-power laser sources and high sensitivity photodetectors at reasonable cost is also a key consideration when designing a LiDAR product. Table 4 summarizes commonly used laser sources and photodetectors for different wavelengths.

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>Lasers</th>
<th>Photodetectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>850, 905, 940</td>
<td>VCSEL, EEL</td>
<td>Si PD/APD/SPAD</td>
</tr>
<tr>
<td>1064</td>
<td>DPSSL</td>
<td>Si SPAD</td>
</tr>
<tr>
<td>1310</td>
<td>EEL</td>
<td>InGaAs PD/APD</td>
</tr>
<tr>
<td>1550</td>
<td>EEL (+EDFA), fiber laser</td>
<td>InGaAs PD/APD</td>
</tr>
</tbody>
</table>

Most mechanical or MEMS-based LiDAR systems adopt 800–1000 nm wavelength band mainly because of the exceptional signal-to-noise ratio of Si avalanche photodiode (APD) or single photon avalanche detector (SPAD) [40]. With the consideration of eye-safety, pulsed (direct) ToF is often used for long distance ranging. For a Si photonics implementation in this wavelength range, SiN waveguides may be used with careful designs for coping with phase nonlinearity at high peak optical power (e.g., tens of watts) and high phase tuning power which complicates thermal management and crosstalk. At least one well-funded Si photonics LiDAR startup [41] is targeting this wavelength range using an OPA approach [42].

GaAs based semiconductor diode lasers, including both edge emitting lasers (EELs) and vertical cavity surface emitting lasers (VCSELS), are widely available at these wavelengths. Both EELs and VCSELS can achieve hundreds of watts peak power, by introducing lateral multi-mode and multi-junction stacking for the former and by arranging a 2D array of emitters for the latter, respectively. Both types of lasers present challenges for use in LiDAR to form a high-quality collimated beam due to highly asymmetric divergence and nonuniform wave front, respectively. This problem becomes more severe when they pair with a Si photonics chip because a waveguide generally requires single mode coupling. In practice, such laser sources are often packaged with a single mode fiber which feeds to a waveguide chip. A significant amount of loss is expected when coupling between a multi-mode laser and a single mode fiber.

At 1064 nm diode-pumped solid-state lasers exist and help high efficiency coupling with a single mode fiber or a waveguide. One the other hand, the sensitivity of Si APD or SPAD sensors reduces by
an order of magnitude compared to its peak value around 800 nm. The authors are not aware of any companies pairing Si photonics with this wavelength at the time of writing.

At 1310 nm or 1550 nm, diode EELs and photodetectors are very mature thanks to the telecom industry though such lasers are single mode and usually emit much lower power than what direct ToF demands. High power multi-mode lasers also exist but the same coupling issue applies. At 1550 nm, mature erbium-doped fiber amplifier (EDFA) technology, which was commercialized in the 90s and dominate optical amplification in long-haul networks, enables high output power by either amplifying an EEL or enabling a fiber laser. These lasers, however, are bulky and expensive due to the use of long fiber coils and pump laser sources. Nevertheless, some LiDAR companies are using such high power 1550 nm lasers though combining it with Si photonics has yet to be known.

When FMCW is used, the requirement on output optical power may be greatly reduced to tens of mW or less [43] thus a single diode EEL may provide sufficient optical power. Many LiDAR companies combining Si photonics with FMCW choose 1550 nm diode lasers [44,45].

3.2.4. Ambient Light Suppression

Many LiDAR systems are used in outdoor environment and the effect of sunlight on SNR must be carefully dealt with. The choice of 940 nm in consumer 3D sensors results from its superior outdoor performance. The solar irradiance at 940 nm is about a third of that at 850 nm or 905 nm and also lower than that at 1550 nm as shown in Figure 2. An extremely low solar irradiance window from 1350 to 1400 nm exists, though this wavelength band is rarely explored due to the poor availability of high-power laser sources. To suppress ambient light at the receiver, sensor read-out frontend circuit must be carefully designed to minimize effect of ambient light photocurrent. Using an ambient photocurrent cancellation circuit at the expense of sensitivity loss, some indirect ToF frontend circuits can enable a receiver operating under direct sunlight at around 100 klx illuminance. Direct ToF or FMCW is generally more immune to ambient light thanks to very short pulse duration or coherent detection, respectively.

Figure 2. (a) Solar spectral irradiance (direct normal and circumsolar at AM 1.5) plotted, with permission, from G173-03 Standard Tables for Reference Solar Spectral Irradiances: Direct Normal and Hemispherical on 37° Tilted Surface, © ASTM International [46]; (b) Solar irradiance with 20 nm integration window.

4. Si Photonics Optical Phased Array

Chip scale OPA has always been thought to be an elegant solution for many beam steering applications. Researchers explored methods using liquid crystal [15,16] and AlGaAs waveguides [47] since early 90s. Modern Si photonics OPA emerged around ten years ago [48] and quickly find its way into many scientific publications and startup business plans. In Table 5, we summarize a non-exhaustive list of published integrated photonics OPA research works.
Table 5: A comparison of selected published integrated photonics OPA research work.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Yr</th>
<th>Group</th>
<th>Ant.</th>
<th>N\textsubscript{an}</th>
<th>D\textsubscript{an} [\mu m]</th>
<th>Beam deg</th>
<th>Steer. deg</th>
<th>SLS [dB]</th>
<th>PS/St.</th>
<th>P\textpi [mW]</th>
<th>BW [kHz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[48]</td>
<td>09</td>
<td>UGent</td>
<td>LG</td>
<td>16</td>
<td>2</td>
<td>2.7, 2.5</td>
<td>50, 0.14</td>
<td>1</td>
<td>TO + \lambda</td>
<td>8.2</td>
<td></td>
</tr>
<tr>
<td>[49]</td>
<td>10</td>
<td>UGent</td>
<td>FG</td>
<td>4 x 4</td>
<td>60</td>
<td>0.4, 0.4</td>
<td>0.2 1</td>
<td>\lambda</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[50,51]</td>
<td>11</td>
<td>UCSB</td>
<td>LG</td>
<td>16</td>
<td>3.5</td>
<td>0.6, 1.6</td>
<td>20, 0.14</td>
<td>1</td>
<td>10</td>
<td>TO + \lambda</td>
<td>215</td>
</tr>
<tr>
<td>[52]</td>
<td>11</td>
<td>UTA</td>
<td>EE</td>
<td>12</td>
<td>3.1 2</td>
<td>32</td>
<td>2-7</td>
<td>TO</td>
<td></td>
<td>12</td>
<td>100</td>
</tr>
<tr>
<td>[53]</td>
<td>12</td>
<td>UCSB</td>
<td>LG</td>
<td>8</td>
<td>5.5</td>
<td>1.8, 0.6</td>
<td>12</td>
<td>7</td>
<td>TO</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>[54]</td>
<td>13</td>
<td>MIT</td>
<td>FG</td>
<td>8 x 8</td>
<td>9</td>
<td>15</td>
<td>TO</td>
<td>8.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[55]</td>
<td>13</td>
<td>UCSB</td>
<td>LG</td>
<td>32</td>
<td>0.3, 1.2</td>
<td>35, 0.14</td>
<td>15</td>
<td>CI +\lambda</td>
<td>1 x 10\textsuperscript{4}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[56]</td>
<td>14</td>
<td>Caltech</td>
<td>FG</td>
<td>4 x 4</td>
<td>50</td>
<td>0.5, 0.5</td>
<td>1, 1.8</td>
<td>1</td>
<td>CI</td>
<td>&gt;10</td>
<td></td>
</tr>
<tr>
<td>[57,58]</td>
<td>14</td>
<td>MIT</td>
<td>LG</td>
<td>16</td>
<td>2</td>
<td></td>
<td>51</td>
<td>TO</td>
<td></td>
<td>13</td>
<td>100</td>
</tr>
<tr>
<td>[59]</td>
<td>14</td>
<td>UTA</td>
<td>EE</td>
<td>16</td>
<td>4</td>
<td>1.2, 0.5</td>
<td>20, 0.15</td>
<td>6-10</td>
<td>TO + \lambda</td>
<td>20</td>
<td>7.3</td>
</tr>
<tr>
<td>[60]</td>
<td>14</td>
<td>UCSB</td>
<td>LG</td>
<td>32</td>
<td>0.6, 0.6</td>
<td>15, 0.14</td>
<td>15</td>
<td>CI +\lambda</td>
<td>14 x 10\textsuperscript{3}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[61]</td>
<td>15</td>
<td>UCSB</td>
<td>LG</td>
<td>32</td>
<td>3.4 2</td>
<td>1, 0.6</td>
<td>23, 0.13</td>
<td>8-14</td>
<td>CI +\lambda</td>
<td>14 x 5 x 10\textsuperscript{4}</td>
<td></td>
</tr>
<tr>
<td>[62]</td>
<td>15</td>
<td>USC</td>
<td>FG</td>
<td>8 x 8</td>
<td>32</td>
<td>1.6</td>
<td>11</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[63,64]</td>
<td>16</td>
<td>MIT</td>
<td>LGS</td>
<td>50</td>
<td>2</td>
<td>0.9, 0.2</td>
<td>46, 0.19</td>
<td>8</td>
<td>TO + \lambda</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>[65]</td>
<td>16</td>
<td>Intel</td>
<td>LG</td>
<td>128</td>
<td>–2 2</td>
<td>0.14, 0.1</td>
<td>80</td>
<td>8-10</td>
<td>TO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[66]</td>
<td>17</td>
<td>MIT</td>
<td>5</td>
<td>1024</td>
<td>0.02, 0.02</td>
<td>46</td>
<td>10</td>
<td>n.a.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[67]</td>
<td>17</td>
<td>PKU</td>
<td>LG</td>
<td>6</td>
<td>8</td>
<td>1.5</td>
<td>11</td>
<td>EO</td>
<td>n.a.</td>
<td>2 x 10\textsuperscript{5}</td>
<td></td>
</tr>
<tr>
<td>[68]</td>
<td>18</td>
<td>UCol</td>
<td>LGS</td>
<td>512</td>
<td>1.3</td>
<td>0.15</td>
<td>70, 0.3</td>
<td>8</td>
<td>TO + \lambda</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>[69]</td>
<td>18</td>
<td>UCol</td>
<td>LG</td>
<td>64</td>
<td>0.8</td>
<td>1.2</td>
<td>120</td>
<td>11</td>
<td>TO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[70]</td>
<td>18</td>
<td>USC</td>
<td>FG</td>
<td>1024</td>
<td>2</td>
<td>0.03</td>
<td>45</td>
<td>9</td>
<td>TO</td>
<td>54</td>
<td>5</td>
</tr>
<tr>
<td>[71]</td>
<td>19</td>
<td>Leti</td>
<td>5</td>
<td>4, 4</td>
<td>3</td>
<td>4.3, 0.7</td>
<td>17</td>
<td>4</td>
<td>TO</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>[72]</td>
<td>19</td>
<td>UCD</td>
<td>LGN</td>
<td>24</td>
<td>1.3</td>
<td>40, 0.16</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Ant., antenna design; N\textsubscript{an}, number of antennae; D\textsubscript{an}, antenna pitch; Beam, beam linewidth; Steer., beam steering range; SLS, sidelobe suppression; PS/St., phase shifter and/or other steering approach; P\textpi, \pi-shift power; BW, tuning speed or bandwidth. LG, 1D linear grating; LGS, sidewall LG; LGN SiN assisted LG; FG, focusing grating; EE, edge emitting; \lambda, wavelength sweeping; CI, current injection. 1 Data for wavelength steering with unit deg/nm. 2 Minimum pitch for nonuniform antenna pitch design. 3 A single value is SLS at 0 deg while a range represents whole steering angle. 4 Based on InP integrated photonics. 5 Based on SiN waveguide.

Table 5 includes some key OPA design aspects including antenna type, the number of antennae, array spacing and phase tuning method. It also includes experimental results including beamwidth, steering range, side lope suppression (SLS) ratio, phase tuning efficiency and tuning speed, which are further analyzed in following discussions.

4.1. OPA System Design

While great distinctions exist between doing academic research and building a Si photonics LiDAR product in practice, research results can often help understand the constraints of a technology when making a system design for a given application. In this section, we analyze a few key parameters of automotive LiDAR and discuss whether and how a Si photonics OPA can meet these targets.

4.1.1. Beam Divergence

Long-range, e.g., 100–300 m, automotive LiDAR demands a horizontal (azimuth) angular resolution of around 0.1° for distinguishing human body size objects at maximum range. The vertical angular resolution may be relaxed to a few times larger as the height information of an object usually has less impact on recognition and decision making. Some emerging short-range automotive applications may tolerate coarser angular resolution whereas we consider the mainstream 0.1° requirement for following discussions. It generally requires a well collimated optical beam with full divergence angle less than the angular resolution. Beamwidth or linewidth, measured as full (angular) width at maximum power (FWHM) is also often used to describe beam quality. The beam divergence angle has a strict relationship to beamwidth for an ideal Guassian beam. Despite the difference in rigorous sense, we may use these terms interchangeably when qualitatively describing a beam.

While mechanical or MEMS scanning LiDAR products often rely on a complex and expensive aspheric lens system to achieve good collimation within the whole steering range, an OPA can in
principle form a well collimated beam without any lenses. Diffraction limit dictates the relationship of the full divergence $\Theta$ of a beam emitted from an OPA with its effective aperture approximately described by a Gaussian beam model [73]

$$\Theta \approx 2\frac{\lambda}{\pi w_0} \approx \frac{4\lambda}{\pi D},$$

(4)

where $\lambda$ is wavelength in transmitting medium, $w_0$ is Gaussian beam waist and $D$ is the total size of an antenna array in the corresponding FOV direction. Taking the 0.1° full divergence requirement at 1550 nm for example, the minimal array size, i.e., aperture $D$, is 1.13 mm and in practical a larger aperture may be needed considering the deviation of the output beam from ideal Gaussian model. Although an mm aperture in an OPA chip does not sound unreasonable, when this requirement couples with antenna arrangement and phase shifter controlling complexity, it becomes a real challenge. The details are discussed in Section 4.2. This simple relation in Equation (4) also explains why millimeter wave radar has great challenges in achieving image grade resolution with reasonable size and cost.

4.1.2. Beam Steering Range

The requirement on the range of beam steering or FOV is comparably less stringent. Axial rotation-based mechanical beam steering can provide 360° horizontal coverage, dwarfing any other mapping methods. However, 360° steering is not always needed in many cases and the OPA solution is mostly in direct competition with MEMS mirror scanning (or other micro motion technologies [74]) or imaging (i.e., flash LiDAR) approaches on cost and manufacturing scalability. A typical commercially available MEMS mirror-based LiDAR (with a single emitter) supports less than 60° horizontal FOV and 20–30° vertical FOV. The proposed horizontal FOV requirement for autonomous driving listed in Table 2 is 120°. To accomplish such wide angles, one can in principle combine two or more emitter and receiver modules and this method is also implemented in many MEMS LiDAR product development.

In a typical OPA design, beam steering is enabled by the optical interference of an antenna array in at least one FOV direction (refer to Table 5). The steering range is theoretically limited by at least two factors: aliasing, i.e., high order constructive interference at angles other than the main direction (usually 0°) and diffraction envelope, i.e., the far field divergence of a single antenna. The aliasing angle $\theta_a$ is the angle between two adjacent constructive interference. For a uniformly distributed antenna array, the Nth order $\theta_a(N)$ is directly related to antenna spacing $d$ by [75]

$$\theta_a(N) = \sin^{-1}\left(\frac{N\lambda}{d}\right) - \sin^{-1}\left(\frac{(N - 1)\lambda}{d}\right), \quad N = 1, 2, \ldots M, \quad M\lambda \leq d.$$

(5)

The 1st order aliasing practically represents the full steering range (between $-\theta_a/2$ and $\theta_a/2$) if aliasing is the sole limiting factor as shown by the example in Figure 3a. A steering angle larger than 60° requires antenna spacing at most 1.15$\lambda$ (e.g., 1.8 $\mu$m at 1550 nm) which may lead to unwanted optical coupling between Si waveguides. In addition, such small spacing requires hundreds of antennae for the mm size emission aperture and results in design and control complexity and high power consumption.

The diffraction envelope originates from the emission profile of a single antenna. It can also be approximated by a Gaussian beam model with Equation (4). The peak beam power follows this envelope when beam is steered across its FOV. When a uniformly spaced antenna array is adopted, the divergence angle of a single antenna is always larger than the aliasing angle because the width of the antenna is less than of the spacing. For antenna arrays with large and unequal spacing to suppress aliasing peaks [52], the diffraction envelope may instead become the limiting factor because the beam power may become too little to maintain a good SNR while steering toward edge angles, as shown by the example in Figure 3b.
while most are around 0.15 deg/nm \([48,50,51,55,59–61,72]\) or slightly higher \([63,64]\). It means for the second FOV dimension, the highest efficiency of wavelength steering is about 0.3 deg/nm \([68]\) while most are around 0.15 deg/nm \([48,50,51,55,59–61,72]\) or slightly higher \([63,64]\). It means for the requirement of 20–30° vertical FOV mentioned earlier, it demands a laser source with tuning range of 100–200 nm. A tunable laser with reliable and fast operation at such wide range is generally very expensive. For example, a high-volume mature product called integrated tunable laser assembly (iTLA) is widely used in telecom and supports 100 nm tuning range. However, it may be lack of sufficient tuning speed for high scanning rate required by LiDAR, and it costs more than what a mass market LiDAR desires. In addition, a grating coupler used as an antenna in wavelength steering may not provide good upward diffraction efficiency over full scanning range especially at angles near 0° due to high back reflection. The unavailability of low-cost automotive qualified tunable lasers and the nonoptimal diffraction efficiency may pose great challenges for OPA based LiDAR products despite research works show impressive results. The efficiency of wavelength steering is related to antenna design and more details are discussed in Section 4.2.1.

4.2. Key Components

With the system design and framework analysis in Section 4.1 combing published research results summarized in Table 5, we further examine certain key components in a Si photonics OPA for design trade-offs and practicality.

4.2.1. Antenna

An antenna in an OPA is a waveguide-to-free space coupler which emits light to far field for beam forming. A waveguide based grating coupler is the most frequently used antenna in Si photonics OPA designs because it diffracts output beam nearly perpendicularly out of chip plane which eases testing and packaging. Grating couplers have been widely used for coupling between waveguides and fibers in many Si photonics products. A grating coupler has unparalleled benefit in wafer-level testing and densely arrayed coupling whereas it suffers from high coupling loss (at least 2–3 dB in practice) and
strong wavelength and polarization dependence. A waveguide edge coupler which emits light right through a waveguide facet can also be used as an antenna. It in principle presents much lower loss and wavelength dependence while limits antenna arrangement in 1D.

The grating coupler antennae used in a 2D and a 1D OPA are often quite different in design. For a 2D OPA, the spacing between adjacent antenna is usually as large as 10–30 µm [54,62] due to the large footprint of beam splitters and phase shifters. If a large steering range is demanded, with careful unequally spaced array design, the ultimate limiting factor tends to be the antenna emission diffraction envelope as shown in Figure 3b. Therefore, a small focusing grating with size of 2–3 µm [54,62] with 40–50° FOV in both directions is often adopted and an example is shown in Figure 4a. In order to diffract most light out of plane within 2–3 µm propagation, unlike the shallow-etched design for fiber couplers, the waveguide core is often completely etched for most grating teeth. These focusing gratings may achieve about 50% upward diffraction efficiency in theory and usually less in practice.

**Figure 4.** An example of (a) a deep-etched focusing grating antenna for 2D OPA (© 2019 IEEE. Reprinted, with permission, from [54]) and (b) a weak coupling strength linear grating antenna for 1D OPA (adapted with permission from [68] © The Optical Society).

For a 1D OPA, waveguide antennae may be arranged very tightly such that a very large aliasing angle can be achieved. In such cases, a long and shallowly etched linear grating is generally more preferred than a short grating as the former results in less divergence angle in the FOV plane parallel to the waveguides (the non-steering FOV). In principle a wide-angle optical collimation system can narrow the divergence for the full steering range, but it diminishes the benefit of OPA beam forming and adds cost, size and complexity. In addition, when wavelength tuning is used for beam steering, as mentioned in Section 4.1.3, the chromatic aberration of any additional optical system must also be dealt with. Therefore, a long grating, i.e., a large aperture size, is often more desirable. Such gratings generally require very weak coupling strength, i.e., small refractive index change along propagation, which may not be easily fabricated.

Different designs and fabrication methods are used to achieve such gratings (refer to Table 5): the Massachusetts Institute of Technology (MIT) and the Columbia University groups use side modulated grating [63,64,68] with width variation as small as 10 nm [68], as shown in Figure 4b, to get around 0.2° divergence; the Intel group uses very shallow grating etch (16 nm out of 400 nm thick waveguide) to get down to 0.1° [65]; and the MIT group uses SiN waveguide, with weaker mode confinement, to achieve an outstanding 0.02° record [66] (although no steering capability presents). Both shallowly etched grating and side modulated grating present processing challenges and do not comply with current commercially available Si photonics processes. The performance of a shallowly etched grating is sensitive to etch accuracy as shown in Figure 5a. Without the help of a stopping layer, it’s challenging to achieve a few nm etch accuracy with in-wafer or wafer-to-wafer consistency. For a side modulated grating, the small width difference, e.g., 10 nm, brings challenges in accuracy control, if not violates design rules, for most commercial Si photonics processes. In addition, it may be difficult to obtain good upward diffraction efficiency compared with its top side etched counterparts. Another approach is to over clad a Si waveguide with a SiN grating which has shown sub 0.1° results [76]. Such a process may not be readily available in a commercial Si photonics process and it may present
some challenges with respect to strain control and planarization. Nevertheless, in the authors’ opinion it may be a more promising approach.

\[ n_{eff} = n_{cl} \sin \theta_c + N \frac{\Lambda}{\Lambda} N = \ldots, -1, 0, 1, 2, \ldots \]  

where \( n_{eff} \) is the effective index of a waveguide, \( n_{cl} \) is the index of a cladding material, \( \theta_c \) is out-coupling diffraction angle, \( \Lambda \) is grating pitch. Therefore,

\[ \frac{d\theta_c}{d\Lambda} = -\frac{1}{n_{cl} \cos \theta_c} \left( N - \frac{dn_{eff}}{d\Lambda} \right) \]  

Most people use a detuned second order grating (\( N = 1 \)) for better upward out-of-plane diffraction efficiency and less back reflection thus the choice of grating pitch does not differ very much for common diffraction angles centered in the range of 5–15° (in SiO\(_2\) cladding). Therefore, the efficiency of wavelength steering is roughly the same (~0.15 deg/nm) for most results listed in Table 5. The results from [63,64] are slightly better and those from [68] are about two times higher. In those works, a narrow side modulated waveguide grating is used as an antenna, and the higher steering efficiency may result from higher wavelength-dependent effective index change for narrower waveguides, as shown in Figure 5b, according to Equation (6).

With the benefit of low loss and wavelength dependence, a waveguide edge coupler may be used as an antenna for 1D array cases [32]. Although an edge coupler may be considered less complex to implement, it also presents some challenges. Commonly used waveguide edge couplers, e.g., an inverse taper-based spot size convertor (SSC), are usually engineered to have a large mode size for efficient coupling to a fiber or another waveguide. However, such end couplers result in small divergence angle, i.e., far field envelope, which narrows steering range. An end coupler can be designed for large divergence, but it then may face the issue of high reflectance from its facet due to large refractive index change between a waveguide end and its adjacent medium. A simple anti-reflection (AR) coating may not be the solution. Si cannot be cleaved in straight line for die singulation, therefore deep etch is used in most Si photonics processes to make edge facet exposed. The process design rules often require leaving a short (few µm) gap between the waveguide end and the etched sidewall. This gap made of SiO\(_2\) is unfortunately inaccurate due to the MEMS type deep etch process and disturbs the performance of any AR coating in an uncontrollable fashion. This issue is not as important as in the
inverse taper SSC case as the effective index at the waveguide end, i.e., the tip, is very close to the refractive index of SiO$_2$ and the reflection from either SiO$_2$ or air is minimal. A possible solution is, instead of deep etch, to polish the edge of a chip to expose waveguide facets. However, an OPA usually includes many antennae, therefore the yield of sidewall polishing must be very high to make sure all antennae facets are of good quality. Edge polishing is generally a challenging and expensive process and does not scale well. This issue may be among the reasons that discourage people from using edge emitting antennae in a Si OPA even if a 1D array design is used.

4.2.2. Antenna Array

With a proper antenna design in place, the next important thing is to arrange them for high quality beam forming. For 1D OPA, the arrangement is relatively straightforward with all antennae, which are often long linear gratings, lined up in parallel as the example shown in Figure 6a.

As analyzed in Section 4.1.2, to achieve a wide steering angle, narrow spacing between antennae is preferred to increase aliasing angle. The spacing is ultimately constrained by unwanted optical coupling when two waveguides are placed very close to each other. The coupling becomes significantly when the spacing is below around 1.5 µm at which 60° FOV can be achieved. Smaller spacing has been demonstrated [69] by reducing waveguide coupling with uneven waveguide width to disturb coupling symmetry. This approach produces a very good FOV but may be more sensitive to fabrication errors and temperature variations in practice.

The use of narrow antenna spacing inevitably results in a large number of antennae, as well as corresponding phase shifters when the aperture size is required to be sufficiently large for narrow beamwidth. The mm size aperture easily demands hundreds of antennae and phase shifters that greatly complicates control electronics and interconnection.

For a 2D OPA array, waveguide and metal routing becomes a critical issue with limited chip real estate. A tree-branch waveguide routing architecture is a straightforward choice with a group of first level waveguide branches split from an input bus waveguide via directional couplers or multimode interferometers (MMIs), and a group of second level waveguide branches split from each first level branch. A typical 2D antenna arrangement using directional couplers is shown in Figure 6b. Such arrangement demands a phase shifter in proximity to each antenna and forms a phase shifter and antenna unit cell. The unit cell usually takes a significant footprint, from nearly 10 µm to over 30 µm [54,62] in each dimension, much larger than the pitch required for a large aliasing angle.

In principle, an unequally spaced antenna array design may be used to reduce the number of antennae and allows larger antenna spacings while maintaining steering range [52,78]. The principle behind this method is to avoid constructive interference among the higher order grating lobes. However, it actually redistributes the power originally in these grating lobes over many more angles with smaller average values rather than absorbs it in the desired main beam. Therefore, while it does help with
effective steering angle, it wastes optical energy and increases noise floor. The applicability of this method depends on the system SNR design and applications.

It is also possible to place all the phase shifters at a region away from the antenna array and use waveguide to arrange the antennae in 2D [49]. However, the density is ultimately limited by how closely two waveguides can be placed, and it quickly becomes comparable to that of the unit cell tree-branch architecture when more than 10 antennae is used in either dimension. There are other attempts, such as multi-layer waveguides, multi-antenna in a single waveguide [79], etc., to tackle this issue while their fabrication complexity and practicality may be up for debate.

4.2.3. Phase Shifter

An OPA generally needs a phase shifter for each antenna to steer output beam via optical inference. In a Si photonics OPA, thermo-optic effect is a widely used method to tune phase and it works for both Si and SiN thought the tuning power efficiency is much lower for the latter which has smaller thermo-optic coefficient (refer to $\beta_{\text{TO}}$ in Table 3). The on-chip heater to induce temperature change in a waveguide-based phase shifter can be realized by TiN or other metal resistors [48,50,68], doped polysilicon resistors [62,70], or a section of doped Si waveguide as resistors [54,57,64,65]. The former two approaches generally consume more power than the latter which directly heats waveguides at the expense of possible optical loss introduced by the dopants via free-carrier absorption. The tuning efficiency can be improved by adding thermal isolation, e.g., deep-etched trench, around phase shifters whereas array density or tuning speed may be negatively impacted. An OPA generally requires $2\pi$ phase shifting for each phase shifter. The typical power efficiency of a thermo-optic phase shifter is about 10–20 mW/$\pi$ (refer to Table 5) though higher efficiency is possible with resonance design at the expense of optical loss and wavelength sensitivity [68]. Thermo-optic tuning is not a fast mechanism, with the fastest bandwidth at around 100 kHz [52,57] due to its passive cooling process. Thermal isolation may reduce the speed by at least several times. Low tuning speed limits the frame rate of a LiDAR demanding a large number of scanned points. The tuning efficiency and speed must be carefully balanced according to the application.

The refractive index of Si can also be changed with free carriers through free carrier absorption induced index change governed by Kramers–Kronig relations [80,81]. A Si diode is usually used in Si photonics to apply the free carrier effect to shift phase. A PIN diode is generally preferred for carrier injection scheme for more injected carrier volume while a PN diode is preferred for carrier depletion scheme for larger overlap between depletion width and optical mode. This method introduces inevitable free carrier-induced optical loss, and the design trade-offs are calculated and shown in Figure 7a,b. In the calculation it assumes that carrier-changing volume completely overlaps waveguide optical field which often inapplicable in practice due to limited optical mode confinement and/or limited depletion field volume. While telecom applications prefer a carrier depletion scheme capable of fast modulation, the carrier injection scheme is more often used for OPA designs as it has smaller footprint and less optical loss. However, even a carrier injection-based phase shifter is still an order of magnitude longer than its thermo-optic counterpart. On the other hand, it supports several orders of magnitude faster tuning speeds at the expense of often larger power consumption [56,61].
1D antenna with a uniform spacing of 1.15\(\lambda\) and power. Such higher power requires good heat dissipation management. As a comparison, a modern smartphone applications processor generally has a thermal design power around 3–5 W with a die size of 50–100 mm\(^2\) and it can be passively cooled by a good thermal design. Such IC chips are often flip-chip bonded to a substrate where heat dissipates through its dense bonding bumps into a substrate.

Grouping of phase shifter electrodes reduces the number of driving circuits but consumes the same phase shifter efficiency, the average power (assuming an 18° beam divergence, the total number of antennae or phase shifters is over 600. With an approximate 10 mW/\(\pi\) phase shifter efficiency, the average power (assuming an average of \(\pi\) shift out of all phase shifters) is about 6 W and peak power may be even higher.

Figure 7. (a) Calculated refractive index change and extinction ratio change in Si as a function of free electron or hole density, respectively; (b) Calculated \(\pi\)-shift length and free carrier absorption loss at such length in Si as a function of free electron or hole density, respectively. The calculation is based on formula and coefficients found in Ref. [81].

The circuits to drive phase shifters are usually off-chip. As analyzed in earlier, a practical OPA may be not able to avoid a large number of antennae and phase shifters even for 1D design. When hundreds of phase shifters are used, the interconnection between an OPA chip and an external driver chip may not be easily accomplished with wire bonding due to the limited number of bond pads a chip can fit at its edges. Flip-chip bonding or 2.5D type of advance packaging [68] may be required which may pose challenges in fiber coupling and antenna emission.

Monolithically integrating IC with photonics on a chip solves this interconnection issue but the availability of fabrication processes is currently limited. Standard RF-SOI CMOS processes have been used to demonstrate impressive IC-integrated OPAs [62,70]. However, the optical performance, e.g., waveguide loss, grating coupler efficiency, etc. of such processes are much poorer than that of dedicated Si photonics processes and an OPA made of such processes may not be able to achieve performance goals a LiDAR product requires. As more foundries are entering the Si photonics space, and some have brought monolithic photonics and electronics processes to market in recent years, we expect to see more OPA designs with monolithic controlling IC on chip.

Another way to tackle this interconnection issue is to group some phase shifters and use a single driver circuit to control each group simultaneously [64]. It works well in theory because the relative phase change between adjacent antennae during beam steering is the same for a uniformly spaced array. However, in practice, the accumulated phase of a waveguide is highly susceptible to unavoidable and random waveguide dimension errors from fabrication non-uniformity. The initial phases among all the antennae are generally indeterministic and may vary from chip to chip. Therefore, the capability of individually controlling each antenna guarantees correct phase condition for beam steering. Such random waveguide dimension error induced phase uncertainty grows with propagation length. Nevertheless, with careful designs, it is possible to group certain number of waveguides with equal and short length and control them by a single driver with minimal impact on beam steering [70].

The phase tuning power of a practical OPA can easily approach a watt or more. Consider a 1D antenna with a uniform spacing of 1.15\(\lambda\), corresponding to aliasing angle of 60°, and 1.17 mm aperture, corresponding to 0.1° beam divergence, the total number of antennae or phase shifters is over 600. With an approximate 10 mW/\(\pi\) phase shifter efficiency, the average power (assuming an average of \(\pi\) shift out of all phase shifters) is about 6 W and peak power may be even higher. Grouping of phase shifter electrodes reduces the number of driving circuits but consumes the same power. Such higher power requires good heat dissipation management. As a comparison, a modern smartphone applications processor generally has a thermal design power around 3–5 W with a die size of 50–100 mm\(^2\) and it can be passively cooled by a good thermal design. Such IC chips are often flip-chip bonded to a substrate where heat dissipates through its dense bonding bumps into a substrate.
as well as through the entire backside to a heat sink. The OPA chip we consider above draws a similar level of power but may only be cooled from its backside due to light emission from its top surface and desired minimal thermal crosstalk between phase shifters. Therefore, a large chip area and a better thermal conduction design may be needed as active or air cooling is not desired in most cases.

An OPA using SiN waveguides faces a more severe power management issue because of poorer thermo-optic phase tuning efficiency. The required power is about an order of magnitude higher [71], i.e., tens of watts, and it may demand active cooling which draws even more power from the system. It effectively makes the solution unattractive or even inapplicable given most other LiDAR products consume 10–20 W. The authors believe that if a SiN OPA is used for its benefit in transparency of 800–1000 nm wavelength range, certain performance may need to be traded off in order to simplify the power management.

The monolithic integration of III-V materials on Si is slowly progressing toward practicality as it is called for both advanced CMOS and photonics applications. It allows the phase tuning in III-V materials via the same carrier injection mechanism therefore it doesn’t solve the power and footprint challenges. Other exotic electro-optic materials such as lithium niobite may be added to Si platform for much lower tuning power but it takes large footprint and more importantly lack of an ecosystem maturity to make products.

Another challenge is the possible thermal crosstalk among adjacent phase shifters. The thermal crosstalk disturbs optical interference phase conditions and varies at different steering conditions, so it cannot simply be calibrated out. Longer spacing between phase shifters or thermal isolation may be needed to minimize this effect for robust beam steering at the expense of footprint.

For making an OPA product, many other challenges are expected in fabrication non-uniformity, packaging yield, phase calibration, testing complexity, qualification, reliability, etc. These are closely related to production therefore are rarely discussed in publications. However, they may greatly impact on the success of commercialization of Si photonics OPA.

5. Other Si Photonics Use Cases for LiDAR

While OPA is currently at the center of the stage for Si photonics in LiDAR space, other use cases may rise as well. The authors list two potentially useful cases for possible deployment in LiDAR products in near future.

5.1. FMCW Coherent Receiver

The concept of a FMCW LiDAR is introduced in Section 2.1 and is graphically described in Figure 8a. FMCW requires coherent detection which mixes a local oscillator, usually from the same laser source, with reflected light signal in optical domain. A balanced photodiode read by differential amplifiers is often used for 3 dB SNR gain and the prevention of amplifier saturation from ambient photocurrent. A simplified system diagram is shown in Figure 8b. Similar optical coherent detection has been adopted for long haul fiber optic communications for nearly 20 years and a Si photonics solution, which integrates an optical mixer, a balanced photodiode and a local oscillator beam splitter, emerged about 10 years ago [82]. By integrating more elements such as IQ modulators, polarization splitters and combiners, fully integrated Si photonics dual polarization QPSK transceivers were demonstrated [83] and successfully commercialized by startups and incumbent system companies. Such Si photonics coherence receivers can be readily tuned for use in FMCW LiDAR.
The wavelength tuning range is as wide as 100 nm, mainly constrained by a wide band tunable laser to steer the optical beam in the FOV direction that a 1D phased array cannot steer. One way to tackle the spatial coherency issue is to use another OPA as a receiver to selectively receive light from the same angle of the steered output beam and mix them in the waveguide [66]. However, this approach limits the receiving optical aperture to the size of the OPA antenna array. This aperture size is often much smaller than that in a traditional receiver optical system using an optical lens therefore the received signal is proportionally smaller leading to a poorer system sensitivity.

5.2. Narrow Linewidth Tunable Laser

A FMCW LiDAR demands a narrow linewidth laser because the temporal coherence length $L_{coh}$ of a laser is inverse proportionally to its linewidth $\Delta \nu$. For a laser with a Lorenztian optical spectrum, this relation can be described by [84]

$$L_{coh} = \frac{2 \ln 2}{\pi} \frac{c}{\Delta \nu},$$

where $c$ is the speed of light. The maximum linewidth for a 400 m coherent length, the round trip of 200 m range, is about 300 kHz which is narrower than that of a typical DFB diode laser (several to tens of MHz). External cavity lasers or fiber lasers can achieve very narrow linewidth but their cost is usually quite high. Ranging beyond coherent length is possible through advanced digital signal processing [85] or using a referencing calibration system [86] at the expense of compute power or system complexity respectively. In addition, as shown earlier, Si photonics LiDAR often desire a wide band tunable laser to steer the optical beam in the FOV direction that a 1D phased array cannot steer. The adoption of such narrow linewidth and tunable lasers is greatly hampered by high cost as discussed in Section 4.1.3.

A narrow linewidth laser and wide band tunable laser may potentially be addressed by using a Si photonics high Q and tunable cavity chip. There has been ongoing research on this topic in the past decade [87–89] and some companies are currently bringing it to telecom applications. The schematic design and the experimental results of an early work of such Si photonics external cavity tunable laser are shown in Figure 9. Such lasers can be made to achieve 100 kHz linewidth corresponding to around 1 km coherent length. The wavelength tuning range is as wide as 100 nm, mainly constrained by the gain spectrum of a gain chip. The 100 nm wavelength range, corresponding to 15–30° steering range, may be sufficient for certain applications. Many challenges still exist to bring this solution to automotive applications due to the stringent requirement on performance and reliability. Nevertheless, it presents the potential to bring down the cost of future wavelength scanning FMCW LiDAR.

![Figure 8](https://example.com/figure8.png)  
**Figure 8.** (a) Concept of FMCW with triangle chirped optical frequency modulation; (b) A simplified typical FMCW LiDAR system diagram. Adapted with permission from [64] © The Optical Society.
The authors categorize a LiDAR by two key design dimensions, e.g., distance calculation method and scene mapping method, and analyze the pros and cons of the design choices commonly adopted in the industry. The authors analyze the system design considerations of automotive LiDAR relevant to Si photonics implementations. The authors review the published research works on Si photonics OPA in the past decade and discuss the system and component designs in details. The discussions attempt to emphasis on design tradeoffs and practicality instead of scientific novelty or exceptional performance of a single aspect. The commercialization status of certain technologies or components is briefly mentioned. The potential uses of Si photonics in coherent receivers and narrow linewidth tunable lasers for certain types of LiDAR solutions are briefly discussed.

There are many challenges to make a Si photonics LiDAR product, particularly an OPA based product. Some challenges relate to the fundamental limitations in the mechanism or the materials and some are engineering difficulties. The large non-linearity effects of Si significantly limit the measurable range in direct ToF or phase shift indirect ToF. The use of SiN allows more power, but the poor thermo-optic phase tuning efficiency makes it less competitive in power consumption, if not a thermal management nightmare. The phase tuning in Si relies on either thermo-optic or free carrier effect. The former requires a careful balance between speed and power consumption, and the latter demands large footprint. Both methods present scaling challenges while large number of antennae and phase shifters is generally desired in an OPA for narrow beamwidth and wide steering angle. Nevertheless, among integrated photonics or other novel technologies eyeing for LiDAR applications, Si photonics has the best ecosystem, including materials, design software, foundry processes, packaging tools and testing automation, as well as talents and product experience accumulated from the past two decades. These are key factors to successfully form a business and develop products. Si photonics OPA, like any other technologies, may exhibit many trade-offs in designs and its commercial success depends on careful choice of designs and associated applications.

To the best of the authors’ knowledge, Si photonics based LiDAR have yet been mass-produced or field-deployed in meaningful quantity at the time of writing. The authors expect that some upcoming LiDAR products labeled with Si photonics may adopt a Si photonics coherent receiver to realize FMCW demodulation paired with a more traditional beam steering method. It may not be a true Si photonics LiDAR in many people’s opinion, but certainly constitutes good progress. A LiDAR product using one or more Si photonics OPAs for 1D steering, and a more traditional steering method for other FOV dimension, may also enter the market soon for properly picked applications. The products using a 2D antenna array or wavelength tuning for 2D steering may take longer to develop and commercialize.

6. Summary

In this article, the authors discuss Si photonics LiDAR focusing on practical automotive applications. The authors categorize a LiDAR by two key design dimensions, e.g., distance calculation method and scene mapping method, and analyze the pros and cons of the design choices commonly adopted in the industry. The authors analyze the system design considerations of automotive LiDAR relevant to Si photonics implementations. The authors review the published research works on Si photonics OPA in the past decade and discuss the system and component designs in details. The discussions attempt to emphasis on design tradeoffs and practicality instead of scientific novelty or exceptional performance of a single aspect. The commercialization status of certain technologies or components is briefly mentioned. The potential uses of Si photonics in coherent receivers and narrow linewidth tunable lasers for certain types of LiDAR solutions are briefly discussed.

There are many challenges to make a Si photonics LiDAR product, particularly an OPA based product. Some challenges relate to the fundamental limitations in the mechanism or the materials and some are engineering difficulties. The large non-linearity effects of Si significantly limit the measurable range in direct ToF or phase shift indirect ToF. The use of SiN allows more power, but the poor thermo-optic phase tuning efficiency makes it less competitive in power consumption, if not a thermal management nightmare. The phase tuning in Si relies on either thermo-optic or free carrier effect. The former requires a careful balance between speed and power consumption, and the latter demands large footprint. Both methods present scaling challenges while large number of antennae and phase shifters is generally desired in an OPA for narrow beamwidth and wide steering angle. Nevertheless, among integrated photonics or other novel technologies eyeing for LiDAR applications, Si photonics has the best ecosystem, including materials, design software, foundry processes, packaging tools and testing automation, as well as talents and product experience accumulated from the past two decades. These are key factors to successfully form a business and develop products. Si photonics OPA, like any other technologies, may exhibit many trade-offs in designs and its commercial success depends on careful choice of designs and associated applications.

To the best of the authors’ knowledge, Si photonics based LiDAR have yet been mass-produced or field-deployed in meaningful quantity at the time of writing. The authors expect that some upcoming LiDAR products labeled with Si photonics may adopt a Si photonics coherent receiver to realize FMCW demodulation paired with a more traditional beam steering method. It may not be a true Si photonics LiDAR in many people’s opinion, but certainly constitutes good progress. A LiDAR product using one or more Si photonics OPAs for 1D steering, and a more traditional steering method for other FOV dimension, may also enter the market soon for properly picked applications. The products using a 2D antenna array or wavelength tuning for 2D steering may take longer to develop and commercialize.
Author Contributions: Analysis and writing, X.S.; data preparation, L.Z. and Q.Z.; funding acquisition, W.Z.

Funding: This work was supported by the National Natural Science Foundation of China (NSFC) (Grant No. 61635013, 61675231, 61705257, and 61805277), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB24030600) and the Youth Innovation Promotion Association, CAS (2016353).

Conflicts of Interest: The authors declare no conflict of interest.

References
6. Heck, M. Highly integrated optical phased arrays: Photonic integrated circuits for optical beam shaping and beam steering. Nanophotonics 2016, 6, 93. [CrossRef]


