Supplementary Material

Serpentine optical phased arrays for scalable integrated photonic lidar beam steering: supplementary material

NATHAN DOSTART1,‡,*, BOHAN ZHANG2,‡, ANATOL KHILO2,3, MICHAEL BRAND1, KENAISH AL QUBAISHI2, DENIZ ONURAL2, DANIEL FELDKHUN1, KEVIN WAGNER1, AND MILOŠ POPOVIĆ2

1Department of Electrical, Computer, and Energy Engineering, University of Colorado, Boulder, CO, 80309, USA
2Department of Electrical and Computer Engineering, Boston University, Boston, MA, 02215, USA
3Currently with Ayar Labs, 6460 Hollis St, Ste A, Emeryville, CA, 94608, USA
*Corresponding author: nathan.dostart@colorado.edu

Published 17 June 2020

This document provides supplementary information to “Serpentine optical phased arrays for scalable integrated photonic LIDAR beam steering,” https://doi.org/10.1364/OPTICA.389006. The supplementary material includes measurement details for SOPA components and OPA performance. Additionally this document discusses the limitations on OPA performance imposed by grating and tiling lobes in the form of SNR and ambiguity as well as lobe suppression techniques. The definition of fill-factor used in the main text is derived in full here. Potential detection techniques for tiled array transceivers are also presented.

1. COMPONENT LOSS MEASUREMENTS

The SOPA consists of four main optical components - waveguides, gratings, tapers and bends - that were serially connected together, forming a long optical path length of 6.4 cm in the design demonstrated here. Because the optical path is so long (∼40,000 times the wavelength), the components needed to be designed for very low loss such that the light was able to traverse the entire SOPA path and sufficiently illuminate the aperture.

The four components and their measured losses are summarized in Fig. S1. For low-loss routing in the grating-waveguides and the flybacks, the waveguides were made to be 6.5 µm wide so that the fundamental mode of 1/λ² width 3 µm was highly confined to the center of the waveguide. By confining the mode to the center of the wide waveguide, the mode had less interaction with the sidewalls and less surface roughness-induced scattering loss [1]. We measured propagation loss of approximately 0.06 dB/cm for 2 µm wide waveguides, which to our knowledge is a record for the AIM process. Propagation loss for 6.5 µm wide waveguides was too low to accurately measure. Low-loss 180° bends with adiabatically changing curvature [2, 3] are designed to connect rows together. The waveguide width for the bends was chosen to be sufficiently small to only support the fundamental mode: 0.5 µm. The input to output pitch of the bends was 8 µm, setting the row-to-row pitch as 16 µm. We measured an insertion loss of 0.003 dB per bend. Adiabatic tapers were designed to connect the 0.5 µm wide bends to the 6.5 µm wide waveguides. We measured a taper insertion loss of 0.07 dB per taper in our first implementation, resulting in the tapers contributing the large majority of loss across the full SOPA aperture (9 dB).

For low-loss gratings, various grating teeth were used. Sidewall perturbations ranging from 10 nm to 200 nm were used as a silicon-only grating design with grating strength up to approximately 1 dB/cm. A sidewall perturbation of 100 nm was used for the two tile results in the main text (20 dB insertion loss into the main lobe). Other grating designs used bars in one of the two nitride layers as grating teeth. The nitride layer which was closer to the silicon waveguide had a grating strength upwards of 30 dB/cm which radiated all of the light in the first several rows (10 dB insertion loss into the main lobe). The nitride layer which was further from the waveguide had a lower perturbation and
| Component            | Design | Measured Loss | Units/Tile | Total Loss/Tile | Design Loss (Next Gen) |
|----------------------|--------|---------------|------------|-----------------|------------------------|
| Euler                |        | 0.003 dB      | 64         | 0.2 dB          | 0.001 dB               |
| Bend                 |        | 0.07 dB       | 128        | 9 dB            | 0.004 dB               |
| Adiabatic            |        | <0.06 dB/cm   | 6.4 cm     | <0.4 dB         | <0.06 dB/cm            |
| Taper                |        | 0.15 dB/cm    | 3.2 cm     | 1 dB            | 0.15 dB/cm             |
| Low-Loss Waveguide   | 6.5 µm |               |            |                 |                        |
| Weak Nitride Grating |        | 0.15 dB/cm    | 3.2 cm     | 1 dB            | 0.15 dB/cm             |

**Fig. S1. Component Loss Measurements.**

has an estimated grating strength of 0.15 dB/cm, allowing approximately uniform emission from the SOPA in the absence of component losses (30 dB insertion loss into the main lobe). This second, weaker nitride grating was used for the single tile results. Future designs will be optimized to achieve 0.5 – 1 dB/cm unidirectional grating strength by utilizing multiple layers for the grating teeth.

2. GRATING LOBES

For element pitches greater than \( \lambda/2 \), which cannot be achieved in the row-to-row spacing of the SOPA, multiple lobes are emitted into the far-field, commonly referred to as grating lobes. All 2D integrated photonic OPAs demonstrated to-date have been unable to meet this spacing requirement, and therefore grating lobes are common to all demonstrations so far. The grating lobe spacing along the fast axis of the SOPA limits the FOV and is related to the row-to-row pitch by \( \Delta \theta = \sin^{-1}(\lambda/\Lambda_y) \). While the grating lobe spacing is set by the row-to-row pitch, the distribution of optical power emitted into the grating lobes is determined by the element pattern and spacing of the gratings. For a uniform grating-waveguide of width \( w \) the proportion of radiated power which is located in grating lobes is approximately \( P_G = P_{rad}(1 - w/\Lambda_y) \). Notably this ratio is the fill-factor \( FF_T \) of the of the OPA, see Sec. 5A for additional discussion.

To minimize the power radiated to grating lobes we therefore require to maximize the ratio of the grating-waveguide width to the grating-waveguide spacing.

The SOPA grating lobes, due to the relatively large element spacing of 16 µm, significantly limit the FOV to 5.5° along \( \theta_y \). With a grating width of 6.5 µm, the SOPA fill-factor is 40.6% corresponding to just under 60% of the radiated power being emitted into grating lobes. Grating lobes can be suppressed in order to avoid potential ambiguity and to recover the full FOV using a variety of approaches, discussed further in Sec. 6.

Without a suppression approach, grating lobes introduce measurement ambiguity directly analogous to the ambiguity associated with tiling lobes, discussed further in Sec. 5D. There is also power loss due to grating lobes which will decrease SNR in a LIDAR system, again analogous to the power loss due to tiling lobes discussed further in Sec. 5C. In particular, we can perform the exact same derivations from Secs. 5C-5D simply substituting the fill-factor of the tile \( FF_T \) in Sec. 5A) for TFF and arrive at identical conclusions with substituted fill-factor.

3. METHODS

Our devices were fabricated in the AIM Photonics multi project wafer (MPW) process (run Feb. 2018), which was chosen for its low waveguide loss and flexibility to use multiple independently pattern-able device layers [4].

We tested the performance of individual tiles and pairs of tiles by imaging the far-field distribution with the setups shown in Fig. S2. We used two different optical setups to image the far-field: a matte planar surface at approximately the Rayleigh range, and a Fourier transform lens (IR, 100 mm focal length, 0.45 NA) focusing onto the InGaAs SWIR detector (Photon Focus MV3-D640I-MO1-144-G2). In the first case, the surface was placed beyond the Rayleigh range and imaged using the InGaAs detector with a camera objective (Fig. S2a1). In the second case, the lens was placed between the tile and the focal plane, with one focal length on each side (Fig. S2a2,b). The first setup allows for a larger FOV in order to capture scanning but is complicated by trigonometric projection factors resulting in curved beam steering loci, while the second gives a higher resolution image of the spot and spot overlap, as well as more rectilinear scanning.

In both cases, the tile is placed on a positioning stage and light is coupled via an optical fiber into AIM PDK component library standard fiber edge couplers [4]. A mirror is used to reflect the beam into the plane of the optical table and onto the far-field surface or into the lens. A tunable laser (Keysight 8164B Lightwave Measurement System with 81608A Laser Module) is used to sweep the wavelength and thereby the angle in 2D; for the two-tile case, the laser passes through an optical amplifier (Amonics AEDFA-HP) and 50:50 splitter before entering the chip through a fiber v-groove array (Meisu 16 channel PM fiber array, 127 µm pitch, Corning PM15-U25D fiber). In this initial demonstration, relative phase shift between the tiles is introduced by straining one of the fibers following the splitter.

All presented images have had a measured background image subtracted from the raw data to remove fixed extraneous scattering and camera hot pixels. For the 200nm sweep figure, we computationally equalize the power in the main lobe to correct for variations in laser power and radiation efficiency across the wavelength range. The scan image is generated by taking 1,500 individual images at a constant 17 GHz frequency interval (and subtracting the background), pseudo-color-encoding each image according to the wavelength, and creating a composite image by adding all the individual images. Visualization 1 comprises the individual images, demonstrating the sequential raster scan.
In order to determine spot size, the number of addressable spots attainable by the SOPA implementation, the frequency steps required for scanning along the fast and slow directions, the FOV, and dual tile operation several versions of experimental setups were used, discussed in Sec. 3. For both single- and dual-tile measurements where a detailed image of the spot is needed, a long focal length lens is used to measure the far-field by placing the lens one focal length away from the CCD. For a rectangular grid of rectangular apertures, it can be found in Fig. S3 for a fast scan at 1550 nm. The SOPA operates over an even wider wavelength range, with normal emission at $\lambda = 1300$ nm, but our laser is limited to a 200 nm wavelength scan range from 1450 – 1650 nm. Notably, in Fig. S3a where every other resolvable spot is shown it can be seen that a fast scan across 41.3 GHz results in a spot which is aligned along the fast axis with the initial spot but shifted along the slow axis to the next addressable spot. This demonstrates the fast/slow raster scanning pattern desired. Taking a projection of each spot image along $\theta_y$ allows for showing the spot cross-section with normalized spot powers in Fig. S3b, and with every resolvable spot in Fig. S3c. This 27 spot fast scan (where the 28th spot is the next slow scan spot) demonstrates the FWHM resolvability of the 27 spots along the fast direction. A key point is that the spot width and grating lobe spacing scale together at different wavelengths, and the number of resolvable spots is invariant across the 200 nm sweep. For an ideal realization there will be as many resolvable spots as grating-waveguide rows (here, 32).

Slow scan measurements are shown in Fig. S4 across the full 200 nm sweep. A slow scan where only every 10th spot (62 spots total) is sampled is shown in cross-sectional view, where the spot powers are normalized, in Fig. S4a and 2D image of the scan is shown in Fig. S4d. The full slow scan, with all 610 spots, is shown in cross-sectional view in Fig. S4b. Notably the frequency spacing required to move by one slow scan spot varies across the 200 nm wavelength range due to group velocity dispersion and the wavelength increment is inherently nonlinear across such a wide wavelength range. We measured 40.8 GHz frequency steps at 1450 nm, 41.3 GHz frequency steps at 1550 nm, and 41.7 GHz frequency steps at 1650 nm to obtain the approximately ‘straight’ slow scan pattern depicted in Fig. S4d. Because these frequency spacings do not vary linearly, both 1st and 2nd order group velocity dispersion terms must be accounted for in this large wavelength range.

To demonstrate the number of addressable slow scan spots, two slow scan spot cross-sections are shown in Fig. S4c using the Fourier plane setup. The two spots cross at approximately 70% of peak power, and are nearly Sparrow resolvable (crossing at 50% of peak power). The slow scan spot spacing is controlled by the relative amount of delay in the flybacks as compared to the grating waveguides, and later designs will reduce this ratio to increase spot resolvability along the slow axis such that all addressable spots are fully resolvable.

5. TILING LOBES

When the separation of phased array elements (in this case, tiles) is larger than $\lambda/2$, the array will produce radiation lobes: additional beams that radiate to undesired angles, thereby reducing power in the desired beam and increasing signal ambiguity [5]. These ‘tiling lobes’ are directly analogous to the grating lobes formed by element spacing larger than $\lambda/2$ within a single OPA tile. We quantify the tiling density using the tiling fill-factor (TFF), defined in the main text and discussed in detail in Sec. 5A.

For a rectangular grid of rectangular apertures, it can be seen that there are $N = 1/TFF$ beams radiated within the sinc envelope of a single aperture, and the main lobe carries $1/N$ of the radiated power (see also Eq. (S5) in Sec. 5C). This can also be seen from the directivity of the tiled array into the hemisphere [5] which is approximately $D = (TFF)2\pi\lambda^2/A$ for a spot at a distance $z$ with area $A$. Since the TFF is a metric which captures the number of tiling lobes, and correspondingly power loss to tiling lobes, it provides a useful metric of array performance.
Supplementary Material

Fig. S3. Fast scan measurements. a False color coded fast scan with 14 spots at 1550 nm with a 41.3 GHz sweep where the 14th spot is the next slow scan spot. b Projection along $\theta_x$ to acquire spot cross-sections for 13 of the spots in part a with normalized power. c The same cross-section with 27 spots, spaced at 1.53 GHz intervals, demonstrating our claim of 27 resolvable spots along the fast axis.

Two main issues arise due to tiling lobes: signal power loss and ambiguous measurements due to the presence of erroneous return signals from tiling lobes. The first effect is a power loss penalty (decreased SNR) due to tiling lobes and discussed in Sec. 5C. The second effect is detrimental in that erroneous signal (due to tiling lobes) can hide the desired signal (due to the main lobe), obfuscating the measurement and discussed in Sec. 5D.

A. Fill-Factor Definitions

In this section we wish to define the fill-factors introduced in the text in a rigorous manner. We additionally desire to design a fill-factor expression which applies equally well to a single OPA tile, or a tiled-array of OPAs. For a single OPA tile we will refer to an example design which is composed of $M$ grating rows of length $L$, width $w_T$, and periodicity $\Lambda_T > w_T$. For a tiled aperture we will use an array of identical tiles in a rectangular pattern with $N_x \times N_y$ individual tiles at periods of $D_x, D_y$ with aperture widths $W_x, W_y$ with each tile emitting the same field pattern $U_T$.

For an arbitrary OPA design, we wish to define a fill-factor which defines the ratio of the emitting area (effective aperture area) of the OPA to the total area taken up by the OPA. We can define this fill-factor as

$$\text{FF} = \frac{A_{\text{eff}}}{A_{\text{tot}}}.$$  (S1)

In order to clarify this definition, we require definitions of both the area terms.

We begin by finding a mathematical definition of effective radiating (aperture) area $A_{\text{eff}}$, the emitting area of the OPA, to use for the numerator of the fill-factor. Henceforth we refer to this area as simply the 'effective area' for short. For OPAs which emit at a uniform amplitude, such as apertures which can be represented in terms of rect and circ functions, this effective area can be calculated by inspection by summing up the area of these window functions. For the single OPA tile example the effective area is $A_{T_{\text{eff}}} = Mw_T L$, and for the tiled-array the effective area is $A_{A_{\text{eff}}} = N_x N_y A_{T_{\text{eff}}}$. For non-uniform emission, however, the effective area is not clear upon inspection and we therefore require a mathematical expression.

We define the effective area of a non-uniform field to align with the uniform case. For a uniform field with unknown, constant amplitude $U$ and unknown area $A_{\text{eff}}$ we can find the effective area by considering two integrals: the field power...
Supplementary Material 5

Power

\( \theta_x (\degree) \)

Power

\( b \)

\( a \)

Power
d

\( \theta_x (\degree) \) \( \theta_x (\degree) \)

\( \theta_y (\degree) \)

Fig. S4. Slow scan measurements. a Projection along \( \theta_x \) to acquire spot cross-sections for every 10\(^{th} \) spot along a slow scan (62 of 610 spots) with normalized spot power. The spot widths are exaggerated due to detector saturation and a measurement plane not in the far-field. b The same cross-section with all 610 spots, spaced at intervals ranging from 41.7 GHz to 40.8 GHz from 1450 nm to 1650 nm. c Two adjacent, power-normalized slow scan spots at 1550 nm measured in the far-field (Fourier plane) demonstrating the resolvability of the addressable spots. d False color coded slow scan image with every 10\(^{th} \) spot.

\[
\int \int |U(\vec{r})|^2 d\vec{r} = \|\overrightarrow{A}\|^2 \text{A}_{\text{eff}} \quad \text{and the integral over square intensity} \\
\int \int |U(\vec{r})|^4 d\vec{r} = \|\overrightarrow{A}\|^4 \text{A}_{\text{eff}}. \]

While the first integral is chosen as it is directly proportional to the optical power, the second integral we choose out of convenience as this term will appear in our calculations of detected signal power. We can then find the effective area as the square of the first integral divided by the second \( \text{A}_{\text{eff}} = (\|\overrightarrow{A}\|^2 \text{A}_{\text{eff}})^2 / \|\overrightarrow{A}\|^4 \text{A}_{\text{eff}}. \) The full integral from is then

\[
\text{A}_{\text{eff}} = \frac{\int \int |U(\vec{r})|^2 d\vec{r}^2}{\int \int |U(\vec{r})|^4 d\vec{r}}.
\] (S2)

If \( U \) is normalized such that \( \int \int |U|^2 d\vec{r} = 1 \), then the effective area is simply \( \text{A}_{\text{eff}} = 1 / \|\overrightarrow{A}\|^4 \text{A}_{\text{eff}}. \) For a non-uniform field, this definition finds the area of an equivalent, uniform aperture while maintaining the power of the field. As an example, this definition of effective area gives \( \text{A}_{\text{eff}} = \pi w^2 \) for a circular Gaussian of the form \( U = U_0 \exp[-r^2/(w^2)]. \)

With a definition of the effective area, we now require to similarly define the total area of the OPA, the total footprint on-chip required for routing, modulation, and the aperture (effective area) itself. However, we will find that there is no obvious way to explicitly, mathematically define the total area for the generic case of an array of non-uniform tiles. We are able to define the total area in general terms, and for a specific OPA emitted field the total area can be found.

There are two relevant choices for defining the total area of an OPA: the minimum-area shape which encloses the radiating aperture, and the minimum-area shape which encloses both the radiating aperture and all associated components. The first area we call the ‘aperture area’, and we define it as the area of the smallest convex polygon fully enclosing the space in which \( |U(\vec{r})| > 0 \). The second area we call the ‘tile area’, and we define it as the area of the smallest convex polygon fully enclosing all components associated with the radiating aperture (phase-shifters, splitter tree, etc.). The ‘aperture area’ is the relevant metric for calculating aperture fill-factors, whereas the ‘tile area’ is relevant for fabricating the tiled-array and the area needed to calculating TFF. For simplicity, here and in the main text we use the smallest rectangle enclosing the relevant space for calculating the aperture and tile areas.

To clarify the difference between the ‘aperture area’ and ‘tile area’, we consider using the single tile example case as the unit cell in the tiled-array example case. The ‘aperture area’ we denote with the \( \text{tot} \) subscript and for the single tile can be seen by inspection to be \( A_{\text{A,tot}} = M \Lambda_y L_y \) the size of the rectangle enclosing all grating rows. The tiled aperture’s ‘aperture area’ is \( A_{\text{A,tot}} = D_x N_x D_y N_y \) the rectangle enclosing all tiles. In this case the ‘tile area’ of the single OPA tile is \( D_x D_y \), the rectangle enclosing the tile, routing waveguides, and control elements.
The tiled array also has a ‘tile area’ which involves e.g. the splitter tree and control elements to distribute and control the individual tiles, which is not of interest here. If we now apply the fill-factor equation to the examples cases, we can calculate the single tile fill-factor as $FF_T = A_{T_{eff}} / A_{T_{tot}} = MwT_L / MA_T = w_T / A_T$ and the tiled-array fill-factor as $FF_A = A_{A_{eff}} / A_{A_{tot}} = N_x N_y A_{A_{eff}} / N_x N_y D_x D_y = A_{T_{eff}} / D_x D_y$. We use the fill-factor definition to identify $A_{T_{eff}} = A_{T_{tot}} (FF_T)$ and therefore $FF_A = A_{T_{tot}} / D_x D_y (FF_T)$. This substitution clarifies that the array fill-factor includes both the fill-factor of the tile $FF_T$ and the ratio of the tile’s ‘aperture area’ to its ‘tile area’, $A_{T_{tot}} / D_x D_y$. The TFF defined in the text is exactly this ratio and is therefore antigenic to the tile design (fill-factor of the tile, FF_T). It can then be seen that the total fill-factor of the tiled-array is the TFF multiplied by the fill-factor of the tile $FF_A = (TFF) (FF_A)$, as expected.

We then identify that, whereas TFF is a key metric in determining the effects of tiling lobes, the effects of grating lobes are related to the single tile fill-factor FF_T and the effects of both tiling and grating lobes are related to the total aperture fill-factor $FF_A$. It is therefore this total aperture fill-factor which is a good overall metric for determining the degradation in system performance due to unwanted lobes (grating and tiling) by fully accounting for both.

B. Tiling Lobe Suppression

The detrimental effects of tiling lobes, signal ambiguity and power loss, can be partially compensated with appropriate design. While the power lost to these lobes cannot be recovered without improving the fill-factor, it is possible to address the issue of signal ambiguity without changing the fill-factor. The approaches developed for grating lobe suppression (see Sec. 6) such as aperiodic arrays [6–8] and Vernier lobe spacing [9, 10], are equally applicable to tiling lobe suppression. For example, varying the tile periodicity between transmit and receive tiled-apertures can suppress tiling lobes, or even allow for sparse apertures which could be interleaved into a single, composite transceiver aperture.

C. SNR of a Tiled-Aperture

The effects of tiling on the SNR of a LIDAR system are discussed briefly here. In particular, we are interested in how the TFF affects the SNR. Notably the following derivation is exact for apertures composed of separable rect functions, but only a useful approximation (which still accurately captures the scaling relations) for realistic aperture functions. The detector current due to the return signal is proportional to the received power from the main lobe $P_{rec}$ for direct detection, $i_{DD}^2 \propto P_{rec}$, and the geometric mean of the received power and the local oscillator (LO) power $P_{LO}$ for heterodyne detection, $i_{DD}^H \propto \sqrt{P_{rec} P_{LO}}$ [11].

The SNR is the temporally averaged ratio of electrical signal power $i_{DD}^2$ to electrical noise power $i_{N}^2$

$$SNR = 10 \log_{10} \left( \frac{i_{DD}^2}{i_{N}^2} \right) \tag{S3}$$

where the temporal averaging is denoted by $\langle \rangle$. Therefore the SNR is proportional to the square of the received power for direct detection $SNR^D \propto 20 \log_{10}(P_{rec})$ or linear in the received power for heterodyne detection $SNR^H \propto 10 \log_{10}(P_{rec} P_{LO})$ [11].

In order to calculate $P_{rec}$, we require a ‘capture efficiency’ $\eta_C$ which relates the radiated power $P_{rad}$ to the received power such that $P_{rec} = \eta_C P_{rad}$. This capture efficiency can be found by overlapping the field scattered off the target and receiver’s back-projected field at a plane of our choosing, such as the target. If both fields are normalized ($\int |U|^2dA = 1$), we can write

$$\eta_C = \left( \int \langle \text{Main Lobe} \rangle U_{TX}(x,y) U_{RX}(x,y) dxdy \right)^2 \tag{S4}$$

for a field reflectivity profile $R$, transmitter field on the target plane $U_{TX}$, and receiver back-projected field on the target plane $U_{RX}$. Note that the receiver field is not conjugated as would be expected in an overlap integral (inner product) because the back-propagation operation (which is a time-reversal operation) applies a second conjugation cancelling out the conjugation from the overlap integral. With the notation here, $U_{RX}$ should be understood as the field on the target plane which would be generated if the receiver array was used as a transmitter.

We consider the case of both transmitter and receiver using identical OPAs. For this case it is easy to distinguish that the received power due to the signal (main lobe) is simply the overlap of the (identical) main lobes of both the transmitter and receiver and the integration term is $U_{TX} U_{RX} \rightarrow U_{TX}^2$. For an example array of $N_x \times N_y$ rectangular tiles with aperture size $w_x \times w_y$ at periodicity $D_x \times D_y$ the TFF is $w_x w_y / D_x D_y$ and the normalized field pattern on a target in the far-field at distance $z$ is

$$U_{TX} = \sqrt{\frac{N_x N_y D_x D_y}{\lambda z^2}} \sin \left( \frac{N_x D_x x}{\lambda z} \right) * \sin \left( \frac{w_x x}{\lambda z} \right) \text{comb} \left( \frac{D_x x}{\lambda z} \right) \tag{S5}$$

$$\text{comb} \left( \frac{D_x x}{\lambda z} \right)$$

where we have pulled the quadratic phase factor into the reflection phase and $\ast$ denote 1D spatial convolutions. For a diffuse reflector the capture efficiency is proportional to the field overlap $\eta_C \propto (\int |U_x|^2dA)^2$ (as well as aperture size and range) which, integrating over only the main lobe centered at $(x,y) = (0,0)$, can be found as $\eta_C \propto TFF^2$. For the case where only the transmitter is an OPA, it can be seen that $\eta_C$ includes only a single factor of TFF.

For identical transmit and receive OPAs we can therefore see that the detected signal power is proportional to the fourth power of TFF for direct detection and the square of TFF for heterodyne detection:

$$\text{SNR}^D \propto TFF^4 P_{rad}^2 \tag{S6}$$

$$\text{SNR}^H \propto TFF^2 P_{rad} P_{LO} \tag{S7}$$

The ‘tiling penalty’ discussed in the main text is then the factor by which the SNR is degraded due to power loss, such that large tiling penalties correspond to low SNRs. For identical transmit and receive OPAs the tiling penalty TP is

$$\text{TP}^D \propto \frac{1}{TFF^4} \tag{S8}$$

$$\text{TP}^H \propto \frac{1}{TFF^2} \tag{S9}$$

For the case where only the transmitter uses an OPA and the receiver is simply a large area, wideband detector behind a focusing telescope, these correspond to factors of $1/TFF^2$ and $1/TFF$ respectively.
D. Ambiguity
In our SNR calculation we did not count the effects of tiling lobs as noise because erroneous signal due to tiling lobs is, in a sense, worse than noise. There are many signal processing methods for increasing SNR by boosting the signal power relative to the noise power. As one example, coherently integrating over multiple return pulses, which will be correlated in their arrival time, will suppress the noise relative to the signal proportional to the number of pulses. However, the tiling lobs are indistinguishable in all aspects from the main lobe post-detection and therefore these methods cannot be applied to improve the ratio of signal to erroneous signal. Because this erroneous signal is not noise but must still be accounted for, we introduce the ambiguity metric (Amb) which is the ratio of the erroneous signal power to the signal power. We define this metric along the lines of SNR in terms of electrical signal power. The ambiguity is then

$$\text{Amb} = \frac{\sum_i (P_i^2)}{(P_L^2)}$$

(S10)

the sum over the signal powers of all tiling lobs divided by the desired signal power. For this definition, Amb $\gg 1$ corresponds to high ambiguity where the desired signal is indistinguishable from the erroneous signal, whereas Amb $\ll 1$ denotes no ambiguity and the desired signal is significantly stronger than the erroneous signal.

The return signal due to a tiling lobe can be found identically with the main lobe case, and we therefore only require to calculate the sum over the squared of tiling lobe powers (for direct detection) or sum over the tiling lobe powers (for heterodyne detection). The sum over powers, in both cases, is not analytically tractable for our example array but can be approximated for both low and high TFF cases as

$$\sum_{i \neq 0} P_{i,L}^2 \propto \begin{cases} \text{TFF}^4 \left(1 - \text{TFF}\right)^{10} & \text{TFF} > 0.5 \\ \text{TFF}^4 \left(\frac{0.2298}{\text{TFF}} - 1\right) & \text{TFF} < 0.15 \end{cases}$$

(S11)

for direct detection and

$$\sum_{i \neq 0} P_{i,L}^2 \propto \begin{cases} \text{TFF}^2 \left(1 - \text{TFF}\right)^4 & \text{TFF} > 0.6 \\ \text{TFF}^2 \left(\frac{4}{\text{TFF}} - 1\right) & \text{TFF} < 0.3 \end{cases}$$

(S12)

for heterodyne detection.

The ambiguities in the direct detection and heterodyne cases (where the other scaling factors cancel and therefore we avoid writing proportionality) are then

$$\text{Amb}^D = \begin{cases} 
(1 - \text{TFF})^{10} & \text{TFF} > 0.5 \\
0.2298 - 1 & \text{TFF} < 0.15
\end{cases}$$

(S13)

$$\text{Amb}^H = \begin{cases} 
(1 - \text{TFF})^4 & \text{TFF} > 0.6 \\
\frac{4}{\text{TFF}} - 1 & \text{TFF} < 0.3
\end{cases}$$

(S14)

where in both cases ambiguity grows inversely with TFF for low TFF, but approaches 0 as TFF approaches unity. The exact solution, as well as high TFF and low TFF approximations, are plotted for both direct detection and heterodyne detection systems in Fig. 55. Notably, while SNR was reduced by TFF more strongly for direct detection than heterodyne detection, the ambiguity of direct detection is more immune to TFF than heterodyne detection for TFF near unity by virtue of the stronger signal dependence on received power.

An ambiguity greater than 1 makes the system effectively useless except for specific LIDAR applications where there is no potential for reflection of grating lobs to create ambiguity: for example, observing spacecraft against a background of empty space. While SNR $\ll 1$ is common, post-detection signal processing recover this low SNR, resulting in an SNR $> 1$ after signal processing that allows these systems to form useful LIDAR returns. Because there is not ‘signal processing’ which allows for distinguishing the main lobe returns from tiling (or grating) lobe returns, we require an ambiguity less than 1 to allow the signal to be distinguished from other lobs. For direct and heterodyne detection, this limit corresponds to TFFs of 12% and 22%, respectively. Tiled apertures with TFFs below these limits will, on average, not be able to distinguish the main lobe signal from tiling lobe signals unless an ambiguity suppression approach is used, as discussed below. Only very few OPA designs, such as the SOPA, currently meet this requirement as seen in Table 1 in the main text. The presence of grating lobes will only worsen the issue, increasing the TFF required for unambiguous returns.

6. GRATING LOBE SUPPRESSION

There are two main approaches to suppressing the effects of grating lobes: removing the presence of such lobes entirely by aperiodic spacing of emitters or misaligning the received grating lobs from the transmitted lobs. The first approach is generally referred to as aperiodic or random arrays, while the second (which uses periodic spacings) is referred to as Verniered arrays.

Aperiodic (or random) arrays [6–8] can be applied independently to transmitters and receivers and this approach spreads the unwanted lobe power out uniformly for a constant ‘background’ while maintaining a diffraction-limited main lobe. The same amount of power is lost to the additional lobes, but the erroneous signal due to these lobes is suppressed and the ambiguity is removed. This approach necessitates spacing the gratings in an aperiodic manner, and will therefore slightly lower the fill-factor.

The Vernier approach [9, 10] requires designing the transmitter and receiver OPAs together such that they have different grating spacing periods. The different periods cause the additional lobes radiated by the transmitter to be misaligned with the equivalent lobes captured by the receiver, thereby supressing the return signal. Just as with aperiodic approaches, the ambiguity is reduced at the cost of a slight decrease in fill-factor.

7. SPECKLE SUPPRESSION THROUGH TILE CURRENT SUMMING

When using OPAs as receivers in a LIDAR system it should be noted that OPAs suffer from speckle more than a standard incoherent detector due to the coherent capture of the backscattered field, analogous to the heterodyne/mixing efficiency of a heterodyne bulk optical LIDAR even if used for a direct detection LIDAR. Specifically, for an aperture which contains $M$ speckles of the back-scattered field all the single OPA aperture, a heterodyne system or OPA will only convert one speckle’s power to electrical signal on average due to the integral over the random phases of the speckle field. In contrast, an incoherent detector of the same size integrates over intensity and therefore converts a factor of $M$ more incident optical power to signal power (with variance of this speckled signal sensitivity increasing as $\sqrt{M}$) [11].

For a transmitter of a given size projecting a beam onto a target at least as large as the beam spot, the correlation function
of the speckle is given by the autocorrelation of the transmit aperture. Therefore, the speckle size at the receiver (in the same plane as the transmitter) is approximately twice the size of the transmitter. Therefore we optimally should build a receiver OPA half as large as the transmitter OPA, but no larger than the transmitter OPA, because there will be no benefit in terms of signal power from larger receivers. This motivation for a matched receiver can be inferred from the far-field overlap integral of the transmitted field and back-propagated capture field, which is maximized for identical transmit and capture fields.

It is possible to exceed this limitation by using a tile current summing approach, where each tile is coupled to its own detector and the detector currents are summed using a delay-matched tree to form the accumulated signal from the array of tiles. This contrasts with the more obvious approach of making the receive array identical to a transmit array: the laser feeding an array of tiles via a distribution network becomes an array of tiles which coherently combine their captured optical fields via a distribution network which feeds into a single detector. The benefit of the tile current summing approach is that, because each tile is smaller than a speckle, we effectively perform an incoherent integration at the tiled-aperture level and avoid the loss associated with coherent integration. The main drawback of this approach is that we cannot obtain any resolution enhancement beyond a single tile spot on the receiver end: we rely entirely on the tiled-aperture transmitter to provide the increased resolution. This effectively decouples the relative sizes of receiver and transmitter in terms of speckled return by making the receiver an incoherent detector on the scale of the speckle.

8. LIMITATIONS ON RANGING WITH WAVELENGTH-STEERED OPAS

LIDAR is the main application intended for the SOPA demonstrated here, so some additional ranging restrictions imposed by wavelength steering should be noted. Consider a LIDAR system consisting of a continuous wave (CW) laser and a modulator with RF bandwidth $B$ which will be used for pulse-shaping for time-of-flight ranging. The shortest pulse which can be created by this modulator has width of $\tau \approx 1/B$, such that higher bandwidth modulators can create shorter pulses. The temporal pulse width can be directly equated with a free-space pulse length as $l = \tau c$, the spatial extent of the pulse along the propagation axis.

The ranging operation itself involves detecting a pulse or coded waveform ’echo’ reflected off the target, measuring the delay $\Delta t$ between the emitted and detected waveform, and estimating the target range as $R = c\Delta t/2$. In the ‘many photon’ regime, a large number of photons are detected each with their own measured time delay (or equivalently a classical, high speed detector) and the pulse shape can be reconstructed accurately. In this case the center of the pulse can be located well within the pulse width $\tau$, corresponding to a range resolution well below the pulse length $\tau/2$. However, LIDAR generally operates in a photon-starved (few photon) regime, where only several photons are detected per return pulse. In this case, it is common that the first photon is detected, but not later photons from the same pulse, as single photon detectors can have reset times much longer than the pulse width. Even if multiple photons are detected, noise in timing of the return pulse leads to a measurement uncertainty on the order of the pulse width. Thus the range resolution $\Delta R$ is given by $c\tau/2$, or in terms of bandwidth $\Delta R \approx c/2B$. While this relation is derived using a pulsed LIDAR approach, the same relation between RF bandwidth and range resolution can be applied to CW ranging schemes such as frequency-modulated continuous wave (FMCW) LIDAR [12]. Both LIDAR ranging schemes are compatible with SOPAs.

The SOPA, and other OPA approaches which utilize wavelength steering, map the optical frequency axis to a beam emission angle in 1D or 2D. Ranging uses the available bandwidth to determine the target range. The question therefore arises as to how to reconcile these two uses of the frequency domain and perform ranging in a wavelength-steered OPA. The answer lies in the disparity between the frequency ranges needed for wavelength-steering and ranging. For the predominant 1D wavelength-steered OPAs, 40-100 GHz frequency steps are used to steer a beam by one spot width, and only a few GHz of modulation is used for ranging, so the ranging signal does not cause noticeable beam deflection and these operations are effectively independent. For a ranging bandwidth greater than the beam steering frequency step (for example using picosecond laser pulses) the frequency content of the pulse will cause beam deviations greater than the spot width, resulting in each frequency

---

**Fig. S5. Signal ambiguity due to tiling lobes.**

a) Signal ambiguity due to tiling lobes for a direct detection system. The exact (numerical) solution is plotted against the high and low TFF approximations. b) Signal ambiguity due to tiling lobes for a heterodyne detection system with exact, high TFF, and low TFF solutions.
line being emitted at a slightly different angle creating a ‘blurred’ spot which is wider than that achieved at a single excitation frequency. By contrast, if the bandwidth used for ranging is less than the beam steering frequency step, the spot centroid associated with each projected frequency will remain within a spot width. It can then be seen that a wavelength-steered OPA is only capable of projecting a bandwidth onto a single target spot at most equal to the spot steering frequency step, limiting the ranging resolution of a wavelength-steered OPA for a single resolvable spot.

In the case of the SOPA, as discussed in the main text, a 1.5 GHz step is required to steer by one resolvable spot along the fast axis. By design, this frequency step corresponds to a range resolution \( \Delta R = 10 \text{ cm} (B = c/2\Delta R = 1.5 \text{ GHz}) \), approximately the resolution desired for automotive applications. As described briefly in the main text, the minimum resolvable steering frequency step (corresponding to maximum ranging bandwidth) is determined by the total time delay across the aperture \( \Delta R = c/(2\Delta f_y) = cMC\tau/2 = L\Delta g \). The numerator is simply the distance traveled by light when traversing the entire delay length accumulated across the aperture, which we might call the effective length of the SOPA. In an interesting equivalence, it can therefore be seen that the on-chip delay line optical path length \( L\Delta g \) is the limit of the achievable range resolution.

It should be noted that in the previous discussion we made no mention of the effects of the wavelength sweep range on ranging. This is because they are independent; the ranging bandwidth is entirely determined by the geometry (as is the spot width). Increasing the wavelength sweep range increases the FOV of a wavelength-steered OPA, and number of spots, but does not affect the spot size itself or the ranging bandwidth.

†These authors contributed equally to this work.

REFERENCES

1. H. Gevorgyan, K. A. Qubaisi, M. S. Dahlem, and A. Khilo, “Silicon photonics time-wavelength pulse interleaver for photonic analog-to-digital converters,” Opt. Express 24, 13489–13499 (2016).
2. M. Cherchi, S. Ylinen, M. Harjanne, M. Kapulainen, and T. Aalto, “Dramatic size reduction of waveguide bends on a micron-scale silicon photonic platform,” Opt. Express 21, 17814–17823 (2013).
3. H. Zafar, P. Moreira, A. M. Taha, B. Paredes, M. S. Dahlem, and A. Khilo, “Compact silicon TE-pass polarizer using adiabatically-bent fully-etched waveguides,” Opt. Express 26, 31850–31860 (2018).
4. AIM Photonics, “AIM silicon photonics PDK v2.5a,” (2018).
5. R. C. Johnson and H. Jasik, “Antenna engineering handbook,” New York, McGraw-Hill Book Co. (1984).
6. 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 3, 887–890 (2016).
7. T. Komljenovic, R. Helkey, L. Coldren, and J. E. Bowers, “Sparse aperiodic arrays for optical beam forming and LiDAR,” Opt. Express 25, 2511–2528 (2017).
8. K. Sayyah, O. Efimov, P. Patterson, J. Schaffner, C. White, J.-F. Seurin, G. Xu, and A. Miglo, “Two-dimensional pseudorandom optical phased array based on tandem optical injection locking of vertical cavity surface emitting lasers,” Opt. Express 23, 19405–19416 (2015).
9. S. Pinna, B. Song, L. A. Coldren, and J. Klamkin, “Vernier transceiver architecture for side-lobe-free and high-extendue LiDAR,” in CLEO: Applications and Technology, (Optical Society of America, 2018), pp. ATu3R–3.
10. N. Dostart, M. Brand, B. Zhang, D. Feldkhn, K. Wagner, and M. A. Popović, “Vernier Si-photonic phased array transceiver for grating lobe suppression and extended field-of-view,” in CLEO: Applications and Technology, (Optical Society of America, 2019), pp. AW3K–2.
11. G. R. Osche, “Optical detection theory for laser applications,” Opt. Detect. Theory for Laser Appl. by Gregory R. Osche, pp. 424. ISBN 0-471-22411-1. Wiley-VCH, July 2002. p. 424 (2002).
12. D. J. Lum, S. H. Knarr, and J. C. Howell, “Frequency-modulated continuous-wave lidar compressive depth-mapping,” Opt. Express 26, 15420–15435 (2018).