

Design, fabrication, and metrology of 10×100 multi-planar integrated photonic routing manifolds for neural networks

Jeff Chiles,^a Sonia M. Buckley, Sae Woo Nam, Richard P. Mirin,
and Jeffrey M. Shainline

*National Institute of Standards and Technology, Applied Physics Division, Boulder,
Colorado 80305, USA*

(Received 9 May 2018; accepted 9 July 2018; published online 26 July 2018)

We design, fabricate, and characterize integrated photonic routing manifolds with 10 inputs and 100 outputs using two vertically integrated planes of silicon nitride waveguides. We analyze manifolds via top-view camera imaging. This measurement technique allows the rapid acquisition of hundreds of precise transmission measurements. We demonstrate manifolds with uniform and Gaussian power distribution patterns with mean power output errors (averaged over 10 sets of 10 inputs) of 0.7 and 0.9 dB, respectively, establishing this as a viable architecture for precision light distribution on-chip. We also assess the performance of the passive photonic elements comprising the system via self-referenced test structures, including high-dynamic-range beam taps, waveguide cutback structures, and waveguide crossing arrays. © 2018 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/1.5039641>

I. INTRODUCTION

A. Background

The development of highly compact and energy-efficient optical interconnects¹ has been a major research objective for integrated photonics. Applications of optical interconnects range from telecommunications² to energy-efficient and high-bandwidth cross-chip communications in CMOS systems.^{3,4} The reason for photonic communication to replace electrical communication is that light experiences no charge-based parasitics and therefore can achieve higher fan-out as well as long-range communication with lower power and higher speed. However, the relatively large size of photonic components presents challenges to their integration. In a system with both photonic and electronic components, the chip area consumed by photonics grows rapidly as the number of communicating nodes, and their degree of connectivity, is increased. For densely connected systems, the requisite number of waveguides can increase to the point where they cannot fit on one plane. Wavelength-division-multiplexing (WDM)^{4,5} or mode-division multiplexing⁶ can partially alleviate this problem. Only one or a small number of master communication buses is then required to satisfy the information bandwidth requirements. When the number of nodes and their degree of connectivity are small, this provides an elegant and cost-effective solution to mitigating the von Neumann bottleneck.⁷

However, neural computing departs significantly from the von Neumann architecture. In a neural system, each processing node (neuron) contains local memory and communicates to many other nodes of the network across local and global spatial scales.^{8,9} The information processing of biological neural systems is approximated in feed-forward neural networks, which have proven technologically useful.¹⁰ A feed-forward neural network consists of multiple layers of neurons, which each integrate several inputs and transmit a signal when a threshold condition is reached. Each layer consists of some number of neurons, which have directed connections to the next downstream layer of neurons.

^aElectronic mail: jeffrey.chiles@nist.gov

Computation and memory are distributed, largely eliminating the bottleneck of processor-memory communication but necessitating significant communication to and from each neuron.

Light is naturally suited to perform this communication. Because photons are uncharged and massless, photons avoid charge-based wiring parasitics. Using light for communication in neural systems is very promising,^{11–16} but constructing a network of nodes each with thousands of connections presents a formidable routing challenge. Using WDM alone is untenable, as it would require an extremely fine and precise wavelength spacing to be constantly maintained. The ability to scale to greater connectivities thus depends on the number of waveguides that can be integrated on a substrate. A suitable solution is the use of multiple planes of photonic waveguides, a field which has seen significant progress over the last decade.^{17–22} The stacking of waveguides allows for dense integration with low-loss and low-cross talk waveguide crossings. In the present work, we present the design and implementation of a two-plane signal distribution network routing 10 input nodes in one network layer to 100 connections on 10 output nodes. This routing manifold accomplishes the routing between two layers of a feed-forward neural network with 10 neurons per layer and all-to-all connectivity. We recently reported a theoretical analysis of the performance and scaling of multi-planar routing strategies for neural computing.²³

B. Design

Feed-forward neural networks commonly leverage topologies where a given layer has order N^2 synaptic connections, where N is the number of neurons in a layer. In this work, we design, fabricate, and experimentally characterize a distributed passive photonic routing manifold capable of realizing connectivities of order N^2 . The routing network can be pruned to achieve any subset of connections. Communication with this manifold requires neither wavelength nor time multiplexing yet can be straightforwardly extended to utilize either. The design of the proposed manifold (Fig. 1) is based on two vertically integrated planes of waveguides. The lower plane (P_1) predominantly runs east, while the second plane (P_2) runs south, thus avoiding in-plane crossings. The light in P_1 bus waveguides originating from each input node is tapped sequentially into P_2 waveguides as the light propagates eastward. This Manhattan-like routing architecture reduces the number of waveguides relative to a scheme where each input is immediately fanned-out with a star coupler.

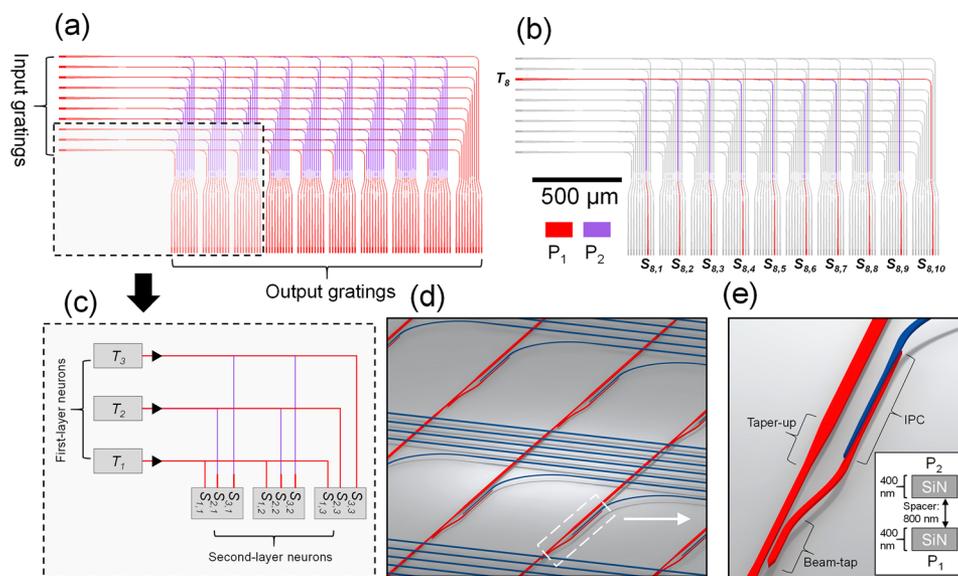


FIG. 1. Proposed photonic routing manifold design. (a) Top-view of the schematic layout; (b) the same view, with one transmission bus and associate output paths highlighted; (c) the representation of the boxed area in (a) as a neural network, showing the notation scheme used throughout this work; (d) 3D perspective of the multi-planar system; (e) zoom view of the tap-and-transition device. Inset: cross-sectional view of the designed layer stack, showing silicon nitride waveguide cores surrounded by SiO_2 . IPC: inter-planar coupler; P_1 and P_2 refer to the bottom and top waveguide planes, respectively.

The manifold implements two layers of a feed-forward neural network with 10 upstream neurons (first layer), 10 downstream neurons with 100 synapses (second layer), and all-to-all connectivity. Figure 1(c) provides this perspective for a reduced section of the manifold. Throughout the remainder of this paper, we will use the labeling scheme shown in Figs. 1(b) and 1(c): inputs (transmitters of the first layer of neurons) are denoted as T_x and the receivers or outputs (synapses of the second layer neurons) are denoted as $S_{x,y}$. For example, $S_{8,5}$ refers to the synapse on the fifth output neuron receiving input from the eighth input neuron, T_8 . The crossbar-like network allows each input node to be routed into a group of 10 outputs representing the whole input array [see the single-input case shown in Fig. 1(b)]. Each output group acts as the synapses (receivers) for that downstream neuron.

The goal of the manifold is to route each input to one synapse on each output, following a pre-determined power distribution pattern. Here, we pursue two schemes to demonstrate control of the output intensity: uniform (each output synapse receives the same power) and Gaussian (the synapses from middle neurons of the upstream layer receive the most power, and the synapses from peripheral neurons receive much less). A script was developed to automatically generate the layouts for the manifolds in both cases: variables in the script set neuron numbers and intensity distribution profiles. The core element of the manifold is the tap-and-transition device shown in Fig. 1(e). It comprises a beam-tap and an inter-planar coupler (IPC) in close proximity. Its function is to divert a certain fraction of the bus power into a perpendicular waveguide on the upper plane. Between bends, gratings, and tap-and-transition devices, the P_1 and P_2 waveguides are adiabatically tapered to and from a larger width ($1.5 \mu\text{m}$) to minimize scattering loss over most of their length. The IPC is a similar design to the one presented in Ref. 17. In the present work, the input waveguide (on P_1) is tapered down to a width of 400 nm over a distance of $12 \mu\text{m}$ and is then routed at a constant width for a distance of $18 \mu\text{m}$. It is finally tapered down to a minimum width of 200 nm over $12 \mu\text{m}$. The other waveguide (on P_2 receiving from P_1) follows the same pattern in reverse over the same length. The total IPC length is $42 \mu\text{m}$.

In a network of this size, a significant dynamic range is required in the power-tap coefficients to achieve either uniform or Gaussian distributions. If only a single coupling gap is utilized, two limits are encountered: (1) the finite size of the sine bend in the tap waveguide results in a certain minimum coupling coefficient and (2) an excessively long interaction length is required to achieve a high coupling coefficient. To address these issues, the manifold makes use of three coupling gaps and variable coupling lengths to improve the dynamic range of the power distribution network. The layout script selects the coupling gap from a look-up table generated from prior measurements of the tap coefficients. The three gap values are 300 nm , 400 nm , and 500 nm . Coupling lengths range from $2.7 \mu\text{m}$ to $19 \mu\text{m}$.

II. EXPERIMENT

A. Fabrication

We fabricated the photonic routing manifolds at the NIST Boulder Microfabrication Facility on 76 mm silicon wafers. Several images of the fabricated structures are shown in Fig. 2. The fabrication process is similar to Ref. 17. In the present work, the two waveguiding planes consist of 400 nm -thick silicon nitride (SiN), with an inter-planar pitch of approximately $1.2 \mu\text{m}$ and a nominal width of 800 nm (on-mask). The waveguides are cladded in plasma-deposited silicon dioxide (SiO_2) on all sides. The SiN material was deposited at a low temperature^{24,25} of 40°C to minimize stress from intrinsic sources and thermal expansion mismatch. As such, the film was not optimized for propagation loss in this experiment, which was of little consequence over the short propagation lengths of the structures under test. The SiN film exhibited a refractive index of 1.96 and a slab propagation loss of $\sim 5 \text{ dB/cm}$ at $\lambda = 1310 \text{ nm}$, measured via prism-coupling.

B. Characterization

Each manifold under consideration each have 10 input ports and 100 output ports. While it is possible to measure these devices with the common approach of aligning optical fibers to grating couplers or facet-terminated waveguides, this measurement technique has significant limitations. First, repeatability strongly depends on the operator's ability to consistently optimize the fiber position

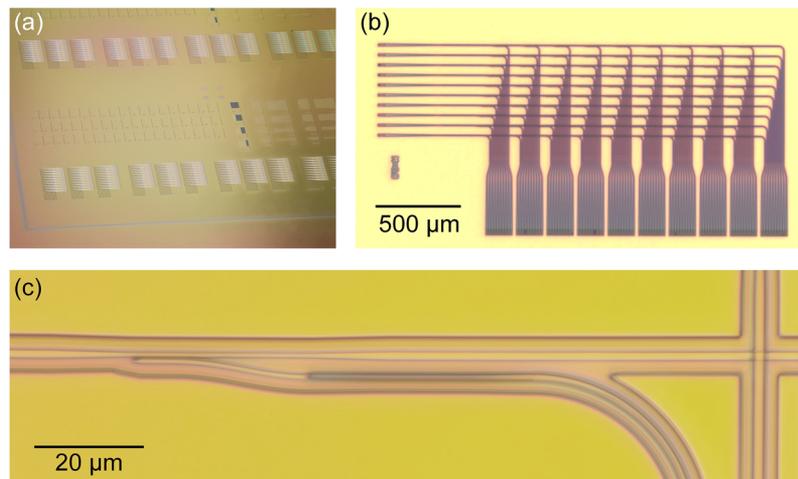


FIG. 2. Optical images of the fabricated devices: (a) focus-stacked view of several test devices on the wafer; (b) photonic routing manifold; and (c) zoom view of a beam-tap and an interplanar coupler to move light from the P_1 to the P_2 plane.

on both ends using micro-positioning stages. Second, the sample and fiber position drift are likely to disturb any power normalization by the time all the output ports are measured. V-groove arrays of fibers may alleviate the problem, but cannot accommodate densely packed structures, nor can the inter-fiber spacing be readily adjusted for different device configurations. Realizing precise fiber array alignment to the sub-dB level is challenging.

Here we pursue an alternative method of transmission measurements for this experiment: top-view imaging with a microscope and a camera. We couple transverse-electric (TE) polarized laser light near $\lambda = 1320$ nm onto the chip through a fiber-to-waveguide grating coupler, and light is coupled out through one or more grating couplers designed for vertical emission (width of $6 \mu\text{m}$). Instead of collecting the light with fibers, we focus it onto a 640×512 pixel, 12-bit-depth (approximately up to 36 dB dynamic range in a single exposure) indium gallium arsenide image sensor array through a microscope objective. The light from each output port is integrated over a small window and normalized to the brightest port in the frame, allowing simultaneous acquisition of many outputs. For most of the devices, a reference port is included near the input to allow straightforward normalization of the input power. An *in situ* image of this arrangement is shown in Fig. 3(a). To obtain low-noise and repeatable measurements, we take care to meet several conditions during all measurements: (1) the camera's gamma (intensity curve) is always fixed at 1.0 to ensure linear power dependence and no gain is applied, (2) a pixel correction mask is applied to remove bright pixels and nonuniformities, (3) background light is filtered out via an 1150 nm long-pass filter inserted in the microscope tube, and (4) all output ports utilize an identical grating design and orientation. Proximity effect correction is applied during lithography to prevent distortion of the gratings in densely loaded areas. Any measurements with saturated pixels are rejected and repeated at a lower exposure time. Likewise, measurements that are too close to the noise floor are repeated at a higher exposure time. The grating coupler efficiency was not characterized, but it is more than sufficient to conduct the measurements with a high signal-to-noise ratio (SNR). Most measurements were conducted with a laser power of only a few hundred microwatts exiting the input fiber.

Images from the camera are analyzed with in-house software, which locates the optical modes of the output ports and extracts a relative power measurement from the set. The data analysis proceeds as follows: (1) Pixels corresponding to scattered light from the input fiber are set to zero; (2) the mean background intensity (calculated from an area away from the ports) is subtracted from all pixels; (3) pixels with negative values are set to zero; (4) a convolution filter with a six-pixel by six-pixel window is applied to locate bright spots; and (5) power is integrated near each port, and the rectangular integration window is expanded in both dimensions until convergence to a specified residual is achieved or a maximum width is reached (typically 20 pixels, or up to half the spacing between ports if they are close together). However, the measurements are generally insensitive to the

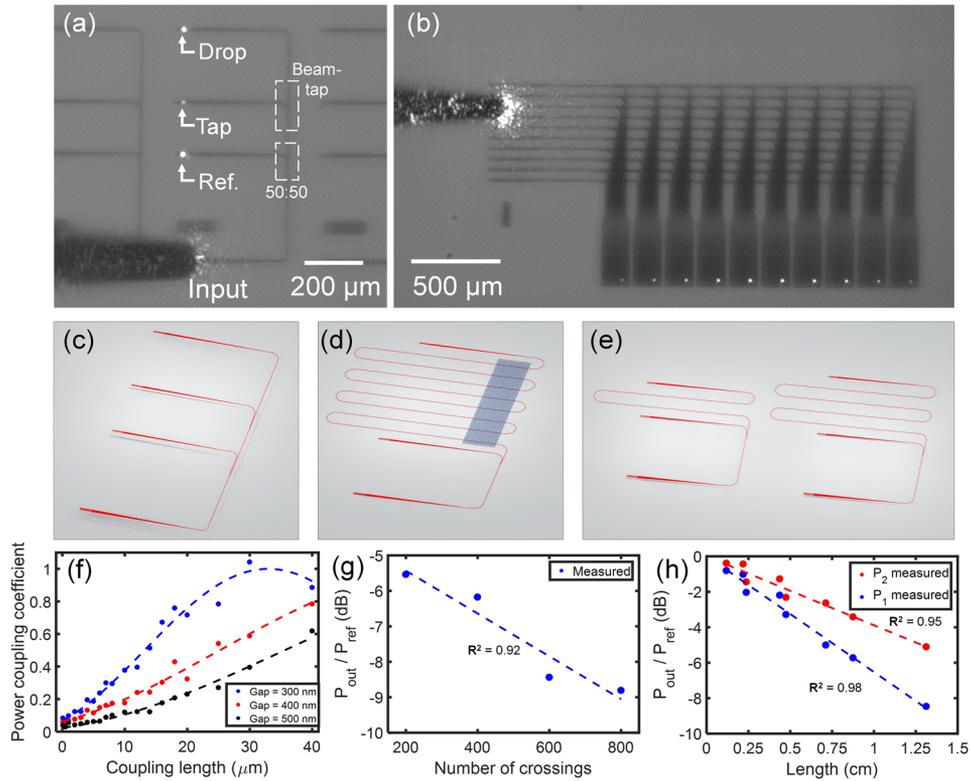


FIG. 3. (a) Infrared top-view image of a test structure for measuring a beam-tap. The reference port allows immediate normalization of the data in the other two ports. (b) Infrared top-view image of a Gaussian-distribution photonic routing manifold. The bright spots are light emitted from the output grating couplers on the connected synapses. [(c)–(e)] 3D perspective views of the beam-tap, waveguide crossing, and cut-back test structures, respectively. (f) Measured beam-tap coefficients for three different coupling gaps used in the high-dynamic range power distribution system for the manifolds. (g) Measured data for the P_1/P_2 waveguide crossings, showing 6 ± 1 m dB loss per crossing. (h) Measured waveguide cutback data, showing 6.5 ± 0.4 and 3.9 ± 0.4 dB/cm propagation loss for the P_1 and P_2 waveguides, respectively. In (f)–(h), circles represent the measured data points and dashed lines indicate fits to the data.

size of the window; for example, changing the width from 12 to 20 pixels usually resulted in a relative power error of 1% or less. We also note that the ports were always separated from each other far enough that there is no detectable cross talk. The output of the script is an array of power values, normalized to the largest value in the set.

This measurement technique allows many photonic devices to be analyzed in parallel with high precision. In this work, we investigate up to 10 ports at once, but many more ports can be analyzed, limited mainly by the imaging performance of the optics and camera which dictate some minimum spacing between ports. Consider the test device in Fig. 3(a), which starts with an input grating coupler. Light is then split into two paths in a 50:50 power splitter (based on a Y-junction). The path on the left leads to a reference output grating coupler. On the right, the path leads to the device under test, in this case a beam-tap. The coupling coefficient of the beam-tap is simply the ratio of the tap output power divided by the reference port's power. The loss of the grating couplers, input waveguide section, and 50:50 splitter are normalized out. Consequently, the measurement has high throughput (fully parallel measurement of many ports) and is robust to alignment errors. Most structures reported in this work, with the exception of the manifolds, were designed with this configuration. In the case of the manifolds, the output ports (synapses) for a given input are measured relative to each other.

1. Passive components

First, we characterized the performance of the different passive components that are used in the manifold. The most critical feature is the high-dynamic-range power distribution system. To analyze

the constituent components, we measure an array of beam-tap test devices. Across the array, the three coupling gaps of 300 nm, 400 nm, and 500 nm are implemented with a variety of coupling lengths. Each test device comprises a 50:50 splitter and a reference port followed by two device output ports: the tap output and the drop output (indicating the untapped power). The measured data are plotted in Fig. 3(f), along with a sine-squared fit of the coupling coefficient to the coupling length. A tight fit is observed for all three coupling gaps, providing a reliable model for future routing manifold designs based on the same platform.

Next, we analyzed the performance of the P_1/P_2 waveguide crossings. The distribution of these crossings is not uniform in the manifold design presented here, so some waveguides experience more crossing loss than others. The T_1 bus waveguide [Figs. 1(a) and 1(c)] encounters 81 crossings, the maximum in this design. A test structure for waveguide crossings is shown in Fig. 3(d). It consists of a meandered P_1 waveguide passing under a cluster of P_2 waveguides above. It crosses the P_2 waveguide cluster a total of 8 times. Test structures with a total of 200, 400, 600, and 800 crossings were measured [Fig. 3(g)]. The P_2 waveguides are 800 nm wide (the same width as the P_1 waveguides) and are spaced by a nominal period of 4 μm , with a random variation between ± 400 nm to ensure that no grating effects are introduced. The data are fit with linear regression to a loss of 6 ± 1 mdB per crossing. Considering the worst case of 81 crossings (path $S_{1,10}$), this constitutes a maximum link loss contribution of 0.49 dB. In the manifolds presented later in this work, waveguide crossings occur between 1500 nm-wide waveguides, which may have slightly lower crossing losses due to tighter optical confinement; nevertheless, this measurement places a conservative bound on the loss value.

Waveguide propagation loss is also important to consider when trying to fabricate precision routing manifolds. Cutback test structures are shown in Fig. 3(e). Eight different path lengths between 1.2 and 13.0 mm were tested and identical structures were fabricated for both the P_1 and P_2 planes. The data are shown in Fig. 3(h). A good fit via linear regression is again observed, indicating propagation losses of 6.5 ± 0.4 and 3.9 ± 0.4 dB/cm, for the P_1 and P_2 waveguides, respectively. The higher P_1 loss could be from mechanical degradation of its top oxide cladding in successive processing steps, which can be addressed with dense and robust sputtered oxide films. Notably, the P_2 waveguide propagation loss is lower than the previously measured slab loss. This may be from two effects: (1) run-to-run variability in the material loss (observed to be on the order of 1 dB/cm) or (2) a change in material properties induced by subsequent processing steps. Future studies will include co-optimization of the optical and material properties of the SiN film to enable scaling to larger numbers of waveguiding planes.

Finally, we discuss the characterization of the IPCs. On this mask, the IPC test structures were placed too far from the optimal zone in the middle of the wafer (where the planarization was on-target), increasing the inter-planar pitch and possibly increasing the losses. Due to the non-ideal planarization, the SiO_2 spacer layer thickness increased by roughly 80 nm over a radial distance of 1 cm from the center of the wafer, with some local variation occurring nearby dense features. Since there were 64 IPCs back-to-back, the total loss exceeded the dynamic range possible in the measurement. Fortunately, the IPC performance could still be straightforwardly characterized by comparing power transmission through two particular synapses on the manifolds: $S_{1,2}$ and $S_{2,2}$. The only difference between them is that $S_{2,2}$ has two IPCs and 180 μm extra P_1 propagation length. We carefully aligned the fiber to each of the two inputs and recorded the power transmitted through the respective synapse. At $\lambda = 1320$ nm (the nominal wavelength for most tests in this work), a value of 0.6 dB per IPC is measured (after subtracting the 0.1 dB loss acquired from the extra propagation length). This is sufficiently low loss to enable good power uniformity, since any two synapses may differ only by up to two IPCs in their routed paths. Still, the loss is higher than anticipated, probably due to a deviation in the fabricated dimensions from the design. In future work, we expect pre-compensation information to improve this to levels similar to our previous work on amorphous silicon.¹⁷ At this point, we can also make an informed estimate of the total link loss experienced in two representative paths through the manifold. First, we consider the path $S_{2,9}$ (Fig. 1), which encounters a relatively large loss compared to the other connections. It has a long propagation length (2.9 mm, all on P_1), 72 waveguide crossings, and 2 IPCs. Utilizing the information collected from the passive measurements earlier in this section, we estimate the $S_{2,9}$ link loss to be 3.5 dB. The smallest link loss occurs

on $S_{1,1}$, which consists of 1.1 mm of P_1 propagation length, leading to 0.7 dB loss. However, it should be noted that these losses are probably larger than the actual values because we have used the propagation loss value from an 800 nm-wide waveguide. Because of the tapered straight propagation sections, the manifolds employ 1500 nm-wide waveguides over most of the propagation length, which could reduce the link loss in the $S_{2,9}$ case.

2. Uniform-distribution manifold

The first type of routing manifold we analyze is the uniform distribution pattern. For any given input, the power delivered to each connected output synapse should be equal; for example, after applying input light to port T_x , we should observe a power distribution of $S_{x,1} = S_{x,2} = S_{x,3} \cdots = S_{x,10}$. To satisfy this requirement, the tap coefficients range from 0.1 to 0.5. An infrared image of the manifold under test is shown in Fig. 3(b), showing light emerging from the output ports. The measured intensities (normalized for each input case) are plotted together in Fig. 4(a), as well as the errors are plotted in Fig. 4(b). While there are a few outliers, the vast majority of synapses exhibit good uniformity. The measured power uniformity of the outputs for input T_8 is shown in Fig. 4(c) as a representative case. Error is calculated as the deviation of each point from the mean of that set. In Fig. 4(d), the mean is calculated for the absolute value of the errors in each row in Fig. 4(b). The grand mean of this data results in an overall average error of 0.7 dB.

Next, we consider the spectral dependence of the uniform routing manifold. For this study, we couple into a single input node T_8 and observe the changes to output uniformity, while scanning the wavelength. The power dependence on the wavelength is plotted in Fig. 5(a) and the error is plotted in Fig. 5(b). The lowest mean error of 0.46 dB is observed at a wavelength of 1320 nm [Fig. 5(c)], and the value remains below 1 dB over a bandwidth of at least 50 nm, providing sufficient tolerance for many applications. We note that the mean error value only differs by 0.1 dB with the measurement of that same node, T_8 , in the earlier series of measurements [see Fig. 4(d), input number 8]. This indicates that the measurement approach is highly repeatable.

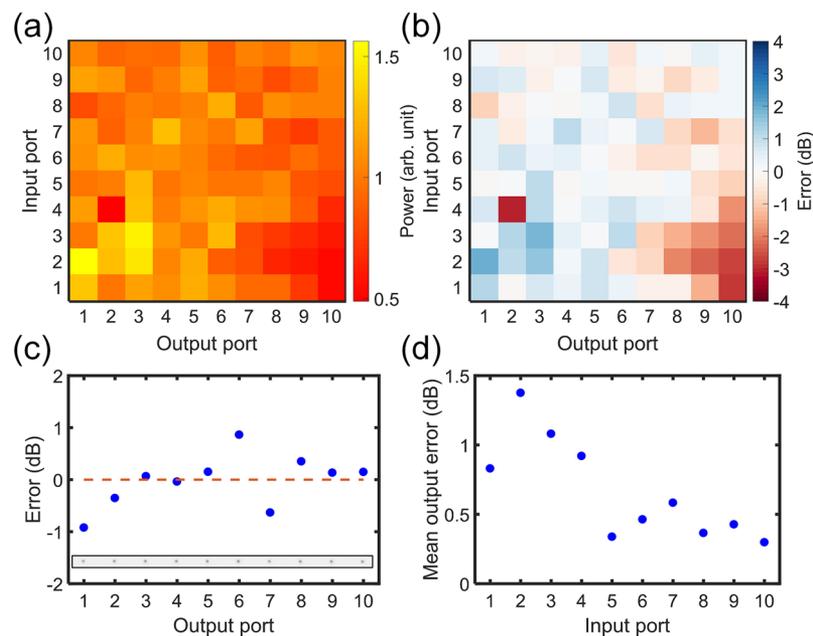


FIG. 4. Measured data for the uniform-distribution manifold measured at 1320 nm. (a) Power distribution for all 100 synapses (each row is normalized to the data in its source microscope image) and (b) error plot for the same data. Error is calculated as the deviation of a synapse's value from the mean for that input. (c) Error plotted for input T_8 , showing that the power on all synapses is within 1 dB of the mean, giving <2 dB spread. The plot in (c) can be viewed as a horizontal slice of (a) for input T_8 . Inset: an infrared image of light from the ports on each synapse, with contrast enhanced for visibility. (d) The mean output error for each input case to observe trends in nonuniformity. Higher error in the early inputs likely results from the higher density of crossings encountered by those waveguides.

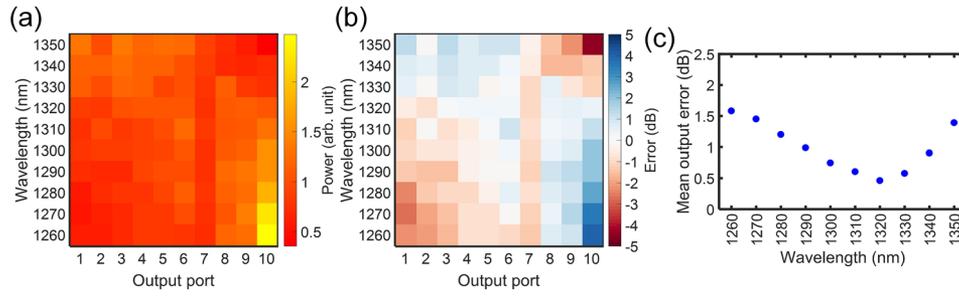


FIG. 5. Spectral performance for the uniform-distribution manifold. (a) Output power for one input case (T_8) as the wavelength is varied; (b) resulting error plot calculated as the deviation from the mean; and (c) mean error for each wavelength, showing optimal performance at $\lambda = 1320$ nm.

3. Gaussian-distribution manifold

We continue the analysis with the Gaussian-distribution routing manifold. This manifold is designed such that the synapses receive power following a Gaussian envelope. The designed envelope is plotted on top of the experimentally measured synaptic power distribution for input node 8 in Fig. 6(c), showing good agreement. The rest of the analysis follows the same pattern as for the uniform case. Measured intensities are plotted together in Fig. 6(a), as well as the errors in Fig. 6(b). For this manifold, the normalization for each input is done by least-squares fitting of the amplitude a of the Gaussian power envelope $P(k)$ according to

$$P(k) = ae^{-\frac{(4 \ln(2)(k-b)^2)}{w^2}}, \quad (1)$$

where k is the index of the output synapse, b is the index of the peak value, and w is the FWHM of the Gaussian envelope (both b and w are equal to 6 and are not fitted in the analysis). Once a is

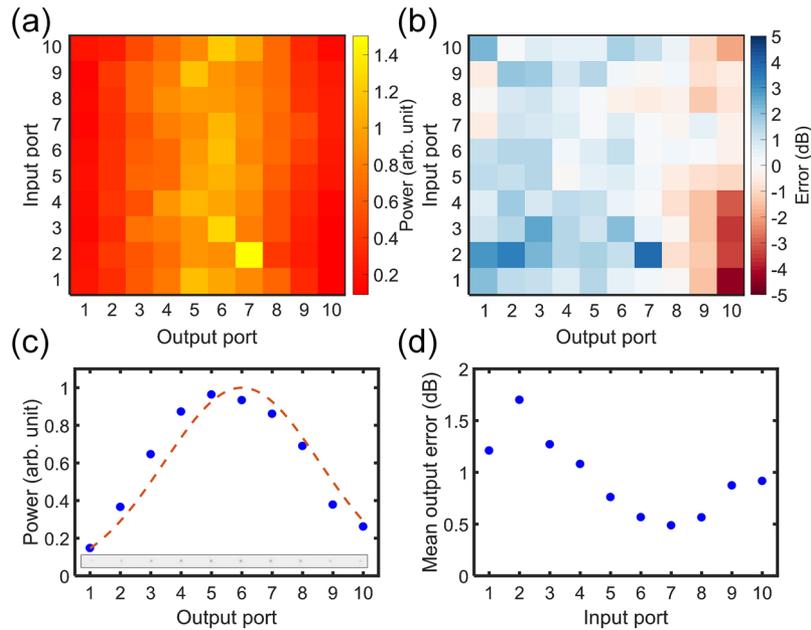


FIG. 6. Measured data for the Gaussian-distribution manifold measured at 1320 nm. (a) Power distribution for all 100 synapses (each row is normalized to the data in that image) and (b) error plot for the same data. Error is calculated as the deviation of a synapse's value from the ideal Gaussian envelope that it was designed to follow (after fitting of the Gaussian amplitude and normalizing). (c) Output powers plotted for the case of input T_8 . Circles are measured data, and the dashed line represents the designed Gaussian. Inset: infrared image of light from each synapse, with contrast enhanced for visibility. (d) Mean error for each input case.

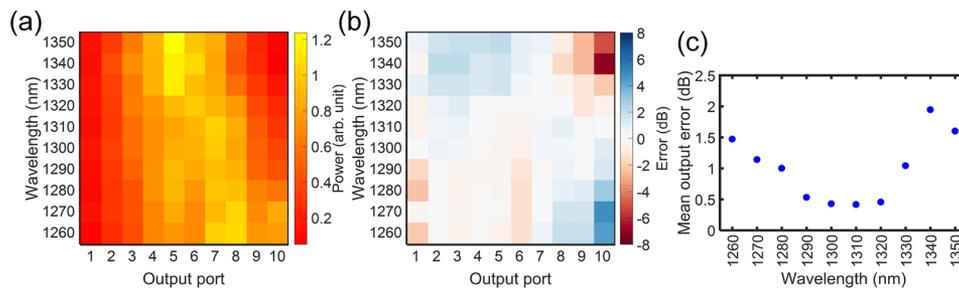


FIG. 7. Spectral performance for the Gaussian-distribution manifold; (a) output power for one input case (T_8) as the wavelength is varied; (b) resulting error plot calculated as the deviation from the ideal Gaussian envelope; (c) the mean error for each wavelength.

fitted, the output powers are normalized to that amplitude, so the envelopes remain in-line despite the occasional bright or dark synapse. In Fig. 6(d), the mean is calculated for the absolute value of the errors in each row in Fig. 6(b). The grand mean of the errors results in an overall average error of 0.9 dB.

The spectral dependence of the Gaussian routing manifold is analyzed with a similar method to the uniform manifold. As before, light is coupled solely into T_8 . The power dependence on the wavelength is plotted in Fig. 7(a) and the error in Fig. 7(b). A trend in the movement of the envelope's centroid toward lower-numbered synapses is seen in Fig. 7(a) as the wavelength is increased. This is consistent with the expectation that the coupling coefficients of the beam-taps will generally increase in the same direction. The lowest error of 0.42 dB is observed at a wavelength of 1310 nm [Fig. 7(c)]. Since the uniform and Gaussian manifolds have the lowest error near a similar wavelength, we can conclude that the beam-tap coefficients are well-calibrated at 1310-1320 nm.

4. Discussion of manifold measurements

Now, we briefly discuss some observations made from the characterization of the photonic routing manifolds in Subsections II B 2 and II B 3. Comparing the error maps of Figs. 4(b) and 6(b), we see a tendency for the higher-numbered output ports to exhibit a power deficiency, especially for those attached to lower-numbered input ports. A likely explanation for this error is that these particular waveguide paths have a higher number of crossings, increasing the total loss relative to other paths. Additionally, propagation loss is accrued (on the order of 1 dB) for the longest paths, which also affects the power uniformity. These loss mechanisms were not accounted for in the design of the beam tap network.

Another type of experimental error observed is the presence of several bright and dark synapses, visible as an abrupt change in the error distribution. One example is synapse $S_{2,7}$ in the Gaussian manifold case [Fig. 6(b)]. These likely originate from mechanical damage to a small number of the output gratings which occurred during the planarization step.

It should be noted that with the data acquired in this study, the errors from crossings and propagation loss can be readily compensated for in the manifold design, by modifying the beam-tap coefficients to distribute more light to the outer synapses as needed. This could greatly improve the power distribution fidelity of the manifolds. At this point, the main source of error would likely be random defects in the waveguides that add loss or a systematic error in the beam-tap coefficients from dimensional variations.

Depending on the application, intensity errors may affect the energy efficiency of the system without impacting information processing. For example, in a neuromorphic application, each synapse will require a certain minimum number of photons to trigger a response. If the optical power distribution network from a neuron to its synaptic connections has nodes which inadvertently receive anomalously low photon numbers, the total amount of light produced by the neuron will need to be increased to ensure that the dimmest connections receive the necessary optical signal. However, in some optical neuron designs,²³ the synaptic weight (which determines information transfer) is set in the electronic domain, so as long as the synapse receives a detectable optical signal, information

processing can occur without error. The choice to instantiate flat and Gaussian distribution manifolds was intended to demonstrate versatile control over the power distribution network. These and other routing patterns may be useful in practice depending on the network architecture.

An important consideration for this architecture is how it scales with the number of input and output nodes. With regards to optical loss, the limiting factor of this design is mainly the loss from waveguide crossings. In the all-to-all connected scheme explored in this work, the maximum number of crossings in any path is proportional to the square of the number of nodes. At a loss-per-crossing of 6 mdB (the value measured experimentally here), an ~ 3 dB loss penalty is incurred in the case of a manifold with 22 input and 22 output nodes. However, there are straightforward methods to mitigate the crossing loss; for example, the interplanar gap can be increased. In doing so, the only compromise will be a somewhat longer interplanar coupler and increased chip area consumption. Further analysis of this scaling has been investigated in Ref. 26.

III. SUMMARY

In this work, we propose, fabricate, and characterize an integrated photonic routing manifold capable of distributing light with high precision across a 10×100 network. The approach utilizes multiple planes of waveguides and a distributed routing scheme to make efficient use of area. The manifold can instantiate custom power distribution patterns, such as uniform or Gaussian, based on the values of beam tap coefficients. This design is topologically equivalent to a feed-forward, 10×10 , all-to-all-connected neural network. In analyzing the network and its sub-components, we employ a method for rapidly acquiring insertion loss measurements. Using fiber-based input coupling and vertical grating emission onto an InGaAs imaging sensor, photonic routing manifolds with 100 output ports are fully characterized in less than 4 min. At a wavelength of 1320 nm, the uniform and Gaussian manifolds were found to have mean output power errors (averaged over 10 rows of 10 inputs) of 0.7 and 0.9 dB, respectively. These routing and measurement techniques offer new opportunities for complex integrated photonic systems in computing, telecommunications, and other applications.

ACKNOWLEDGMENTS

Official contribution of the National Institute of Standards and Technology, not subject to copyright in the United States.

- ¹ D. A. B. Miller, *Proc. IEEE* **88**, 728 (2000).
- ² C. Zhang and J. E. Bowers, *Opt. Fiber Technol.* **44**, 2 (2017).
- ³ C. Sun, M. Wade, Y. Lee, J. Orcutt, L. Alloatti, M. Georgas, A. Waterman, J. Shainline, R. Avizienis, S. Lin, B. Moss, R. Kumar, F. Pavanello, A. Atabaki, H. Cook, A. Ou, J. Leu, Y.-H. Chen, K. Asanović, R. Ram, M. Popović, and V. Stojanović, *Nature* **528**, 534 (2015).
- ⁴ C. Sun, M. T. Wade, Y. Lee, J. S. Orcutt, L. Alloatti, M. S. Georgas, A. S. Waterman, J. M. Shainline, R. R. Avizienis, S. Lin, B. R. Moss, R. Kumar, F. Pavanello, A. H. Atabaki, H. M. Cook, A. J. Ou, J. C. Leu, Y. H. Chen, K. Asanovic, R. J. Ram, M. A. Popović, and V. M. Stojanović, in *Optical Fiber Communications Conference and Exhibition (OFC)* (OSA, 2017), pp. 1–3.
- ⁵ L. Zhou, S. S. Djordjevic, R. Proietti, D. Ding, S. J. B. Yoo, R. Amirtharajah, and V. Akella, *Appl. Phys. A* **95**, 1111 (2009).
- ⁶ H. Jia, T. Zhou, X. Fu, J. Ding, L. Zhang, and L. Yang, *Opt. Express* **26**, 9740 (2018).
- ⁷ J. Backus, *Commun. ACM* **21**, 613 (1978).
- ⁸ O. Sporns, *Biosystems* **85**, 55 (2006).
- ⁹ E. Bullmore and O. Sporns, *Nat. Rev. Neurosci.* **10**, 186 (2009).
- ¹⁰ D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, *Nature* **529**, 484 (2016).
- ¹¹ A. Tait, M. Nahmias, B. Shastri, and P. Prucnal, *J. Lightwave Technol.* **32**, 4029 (2014).
- ¹² Y. Shen, N. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljacic, *Nat. Photonics* **11**, 441 (2016).
- ¹³ Z. Cheng, C. Rios, W. Pernice, C. Wright, and H. Bhaskaran, *Sci. Adv.* **3**, e1700160 (2017).
- ¹⁴ A. Tait, T. F. de Lima, E. Zhou, A. Wu, M. Nahmias, B. Shastri, and P. Prucnal, *Sci. Rep.* **7**, 7430 (2017).
- ¹⁵ J. M. Shainline, S. M. Buckley, R. P. Mirin, and S. W. Nam, *Phys. Rev. Appl.* **7**, 034013 (2017).
- ¹⁶ J. M. Shainline, S. M. Buckley, A. N. McCaughan, J. Chiles, R. P. Mirin, and S. W. Nam, e-print [arXiv:1805.01929](https://arxiv.org/abs/1805.01929) (2018).
- ¹⁷ J. Chiles, S. Buckley, N. Nader, S. W. Nam, R. P. Mirin, and J. M. Shainline, *APL Photonics* **2**, 116101 (2017).
- ¹⁸ W. D. Sacher, J. C. Mikkelsen, P. Dumais, J. Jiang, D. Goodwill, X. Luo, Y. Huang, Y. Yang, A. Bois, P. G.-Q. Lo, E. Bernier, and J. K. S. Poon, *Opt. Express* **25**, 30862 (2017).

- ¹⁹ K. Shang, S. Pathak, B. Guan, G. Liu, and S. J. B. Yoo, *Opt. Express* **23**, 21334 (2015).
- ²⁰ A. H. Hosseinnia, A. H. Atabaki, A. A. Eftekhar, and A. Adibi, *Opt. Express* **23**, 30297 (2015).
- ²¹ J. Kang, N. Nishiyama, Y. Atsumi, T. Amemiya, and S. Arai, in *Optoelectronic Interconnects XIII* (SPIE, 2013), Vol. 8630, pp. 863008–863012.
- ²² T. J. Seok, N. Quack, S. Han, R. S. Muller, and M. C. Wu, *Optica* **3**, 64 (2016).
- ²³ J. M. Shainline, J. Chiles, S. M. Buckley, A. N. McCaughan, R. P. Mirin, and S. W. Nam, e-print [arxiv:1805.0193](https://arxiv.org/abs/1805.0193) (2018).
- ²⁴ Z. Shao, Y. Chen, H. Chen, Y. Zhang, F. Zhang, J. Jian, Z. Fan, L. Liu, C. Yang, L. Zhou, and S. Yu, *Opt. Express* **24**, 1865 (2016).
- ²⁵ J. M. Shainline, S. M. Buckley, N. Nader, C. M. Gentry, K. C. Cossel, J. W. Cleary, M. Popović, N. R. Newbury, S. W. Nam, and R. P. Mirin, *Opt. Express* **25**, 10322 (2017).
- ²⁶ J. M. Shainline, J. Chiles, S. M. Buckley, A. N. McCaughan, R. P. Mirin, and S. W. Nam, e-print [arXiv:1805.01942](https://arxiv.org/abs/1805.01942) (2018).