

## Silicon Photonic Modulator Neuron

Alexander N. Tait,<sup>\*,†</sup> Thomas Ferreira de Lima, Mitchell A. Nahmias, Heidi B. Miller,<sup>‡</sup>  
Hsuan-Tung Peng, Bhavin J. Shastri,<sup>‡</sup> and Paul R. Prucnal

*Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA*

 (Received 9 December 2018; revised manuscript received 3 February 2019; published 18 June 2019)

There has been recent interest in neuromorphic photonics, a field with the promise to access pivotal and unexplored regimes of machine intelligence. Progress has been made on isolated neurons and analog interconnects; nevertheless, this renewal of interest has yet to produce a demonstration of a silicon photonic neuron capable of interacting with other like neurons. We report a modulator-class photonic neuron fabricated in a conventional silicon photonic process line. We demonstrate the behaviors of transfer-function configurability, fan-in, inhibition, time-resolved pulse processing, and, crucially, autaptic cascadability—a sufficient set of behaviors for a device to act as a neuron participating in a network of like neurons. The silicon photonic modulator neuron constitutes the final piece needed to make photonic neural networks fully integrated on currently available silicon photonic platforms.

DOI: [10.1103/PhysRevApplied.11.064043](https://doi.org/10.1103/PhysRevApplied.11.064043)

### I. INTRODUCTION

Renewed interest in neuromorphic photonics has been heralded by advances in photonic integration technology [1–3], roadblocks in conventional computing performance [4,5], the return of neuromorphic electronics [6–10], and the inundation of machine learning (ML) with neural models [11]. Neural networks have played some role in ML (e.g., image and voice recognition, language translation, pattern detection, and others) since the 1950s [12,13]. They fell out of favor in the 1990s because they are difficult to train.

Over the past decade, neural network models have decisively retaken the helm of ML under the alias of “deep networks” [14]. There are three main reasons: (1) major algorithmic innovations [15,16], (2) the Internet—an inexhaustible source of millions of training examples—and (3) new hardware, specifically graphical processing units (GPUs) [17]. Central processing units (CPUs) are woefully inefficient at evaluating these models because they are centralized and instruction based, whereas networks are distributed and capable of adaptation without a programmer. GPUs are more parallel but, today, even they have been pushed to their limits [18].

Today’s demand for evaluating neural network models necessitates new hardware. High-tech behemoths and

research agencies—notably IBM [6], HP [19], Intel [10], Google [20,21], the Human Brain Project [22], and DARPA SyNAPSE [23]—have invested heavily in massively parallel application-specific integrated circuits (ASICs) for evaluating neural network models more efficiently. Some of these architectures aim to be ML number crunchers [20,24] and others have enabled novel neuroscientific tools [25,26] and previously unforeseen low-power mobile applications [27].

The primary performance driver for the neuromorphic electronics community is computational power efficiency; speed is a secondary consideration. Neuromorphic electronics have largely focused on biological-time-scale neural networks: 1 kHz bandwidth (with one 10 MHz exception [25]). They universally rely on digital time and event multiplexing [[28], Sec. 3.3], which means that they cannot simply run faster by turning up the clock. Nevertheless, there are compelling applications for neural networks with gigahertz bandwidths. Some applications could be offline (i.e., number crunching), such as accelerators for deep-network training and inference; others could be online (i.e., real-time), such as pattern detectors for wide-band radio-frequency (rf) signals and feedback controllers for systems subject to short-time-constant instabilities.

Moving beyond the nanosecond will require moving beyond purely electronic physics.

Photonic physics exhibit properties distinct from those of electronics in terms of multiplexing, energy dissipation, and cross talk. These properties are favorable for dense, high-bandwidth interconnects [29] in addition to configurable analog signal processing [30–32]. Consequently, neuromorphic photonic systems could operate 6–8 orders of magnitude faster than neuromorphic electronics [33]

<sup>\*</sup>atait@ieee.org

<sup>†</sup>Now at National Institute of Standards and Technology, Boulder, CO 80305, USA.

<sup>‡</sup>Also at Department of Physics, Engineering Physics & Astronomy, Queen’s University, Kingston ON, K7L 3N6, Canada.

with potentially higher energy efficiencies [34]. Neural interconnects based on field evolution in free space [35, 36], holograms [37,38], and fiber [39] have been demonstrated but have not been widely adopted, in part because they cannot be integrated and thereby scaled robustly and manufactured cheaply. Analog interconnects integrated on a silicon photonic platform have been demonstrated [40, 41]. These interconnects are optically coherent, and a cascable photonic neuron that regenerates phase from layer to layer has not been proposed.

In this work, we fabricate and demonstrate a silicon photonic modulator neuron. The neuron consists of a balanced photodetector directly connected to a microring (MRR) modulator. We demonstrate that this device possesses the necessary capabilities of a network-compatible neuron: fan-in, high-gain optical-to-optical nonlinearity, and indefinite cascability. Furthermore, we demonstrate optional but useful capabilities of transfer-function configurability, inhibitory fan-in, pulse compression, and time-resolved processing.

Network-compatible photonic neurons are optical-in optical-out devices that must be able to (1) convert multiple independently weighted inputs into a single output (i.e., *fan-in*), (2) apply a nonlinear transfer function to the weighted sum of the inputs, and (3) produce an output capable of driving multiple other neurons, including themselves (i.e., *cascability*). There are fundamental reasons why fan-in, nonlinearity, and cascability are difficult to achieve in all-optoelectronic devices, as discussed further in Sec. IV A. The modulator neuron addresses these issues using an optical-electrical-optical (O-E-O) signal pathway with a lumped, intermediate electronic connection.

Research in neuromorphic photonics has seen an abundance of semiconductor lasers that exhibit nonlinear transfer functions, particularly those of a spiking neuron [42–50], reviewed in Ref. [51]. Nevertheless, all of the conditions of network compatibility have yet to have been conclusively demonstrated in a single device (with the exception of the fiber laser in Ref. [52]), and much of

this research overlooks fan-in and/or cascability entirely. Other works pertaining to the broadcast-and-weight architecture have shown proof of concept of an O-E-O neural pathway: in Ref. [53], a fiber detector was connected to a fiber Mach-Zehnder modulator; in Ref. [48], a III-V detector was connected to a III-V laser die. This O-E-O neural pathway has not been shown in a monolithically integrated silicon device.

The neuron shown here is integrable with the silicon photonic neural network demonstrated in Ref. [53]. Modulator neurons and the MRR weight network combine to form a neuromorphic photonic architecture called broadcast-and-weight [54,55], diagrammed in Fig. 1. In this architecture, photonic neurons output optical signals with unique wavelengths. These are wavelength-division multiplexed (WDM) and broadcast to all others, weighted, and detected. Each connection between a pair of neurons is configured independently by one MRR weight [57,58], and the WDM carriers do not mutually interfere when detected by a single photodetector. In the context of prior work, the MRR modulator neuron illustrates a path to fully integrated photonic neural networks compatible with currently available silicon technology. This compatibility is a crucial element of the scalability, cost, and feasibility.

Silicon photonics presents an opportunity for optical technologies to venture beyond pure communication links and specialized co-processing boxes. Recent advances have brought about a stable academic and industrial ecosystem surrounding silicon photonics [59,60], propelling technology road mapping [3], standardized multiproject wafer runs [1,2], broadened accessibility to academic research [61], and economies of scale previously enjoyed solely by microelectronics. Despite its origins in cutting the cost of midreach communications, silicon photonic integration could be instrumental to large-scale photonic information processing concepts—beyond what could be considered in fiber, III-V, or holographic platforms.

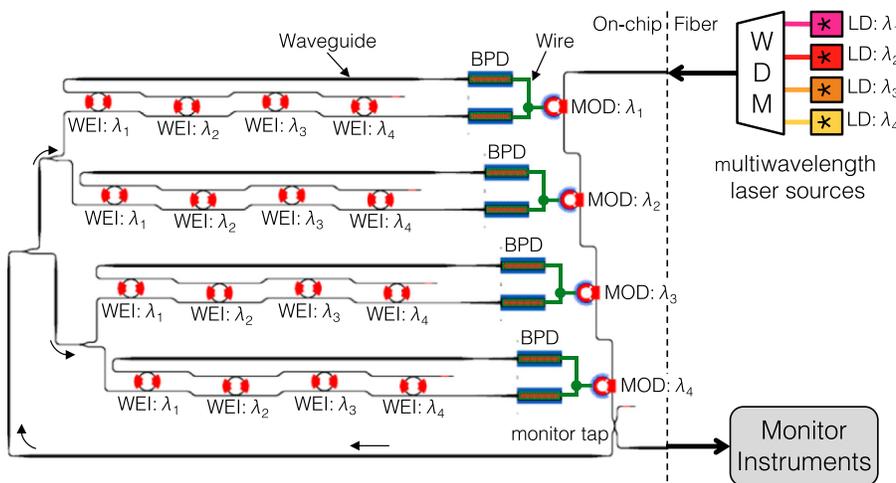


FIG. 1. A broadcast-and-weight network [55] using MRR modulator neurons and MRR weight banks (WEI) [56]. The entire network is integrated with the exception of the pump-laser diodes (LD), which lie outside of the high-bandwidth signal pathway. The red (blue) arcs are  $n$ -doped ( $p$ -doped) regions. The green lines are the wires connecting balanced photodiodes (BPD) to modulators (MOD) within each neuron. Each MRR interacts with only one wavelength,  $\lambda_i$ .

## II. METHODS

### A. Device description

#### 1. Theory of operation

The modulator neuron is an optical-to-electrical-to-optical (O-E-O) device consisting of two photodetectors (PDs) connected electrically to a microring (MRR) modulator. It takes two optical inputs, subtracts their photocurrents electronically, and remodulates a signal onto a new wavelength [Fig. 2(a)]. The output signal is a nonlinear function of inputs determined by the electro-optic transfer function of the modulator.

Figure 2(a) shows the neuron circuit diagram. The WG labeled “pump” directs a continuous-wave (cw) laser signal at  $\lambda_n = 1544.8$  nm to the MRR modulator neuron. When modulated, this signal serves as the neuron’s optical output. The PDs are reverse biased and convert impinging light at all wavelengths into photocurrents. Photocurrent from the positive PD adds to the injected current, while photocurrent on the negative PD shunts away some of the injected current coming from the bias port. These photocurrents are combined with a bias current  $I_b$  such that their sum affects the refractive index of the MRR modulator via free-carrier injection to a  $p$ - $n$  junction. This index change affects its transmission at the pump wavelength and ultimately modulates the amount of light transmitted at the output port of the neuron. The MRR resonance wavelength can also be thermally tuned by a in-ring heater with a current  $I_h$ . The heater is used for strong, slow resonance tuning, while the  $p$ - $n$  junction is used for weak, fast resonance modulation.

#### 2. Fabrication

Figure 2(b) shows a false-color optical micrograph of the fabricated neuron. Waveguides (WGs), shown in yellow, are formed in the 220-nm-thick silicon layer. They are patterned to be 500 nm wide except in long, straight segments, WGs are widened to 3  $\mu\text{m}$  to minimize side-wall scattering. The MRR modulator consists of a circular WG with a designed radius of 11.5  $\mu\text{m}$  coupled to one WG with a gap of 200 nm. The WGs containing active devices are partially etched to a 90-nm-thick pedestal (light green) that hosts the dopants.

The doping profiles of the active regions are diagrammed in Fig. 2(c). The MRR has two doped regions: an  $n^+$ - $n$ - $n^+$  junction to act as an ohmic heater and an  $n^+$ - $n$ - $p$ - $p^+$  junction to act as a high-speed modulator. The phosphorous and boron concentrations are the same as in Ref. [62].

A germanium layer is deposited and patterned (dark green) along with additional dopants that create a vertical  $p$ - $i$ - $n$  photodiode [Fig. 2(d)]. Silicon dioxide is then deposited, vias are etched, and aluminum traces are deposited (red); this sequence is repeated to form a second

wiring layer (orange). Metal traces are patterned to break out the device electrical ports to probe pads.

### B. Device characterization

The optical behavior of the MRR modulator and its response to different effects is characterized by the profile of its resonance dip. We examine the optical transmission spectrum of the MRR modulator in response to three independent quantities:

- (1) Heater current bias ( $I_h$ )
- (2) Modulator current bias ( $I_b$ )
- (3) Optical power into the IN+ and IN− ports ( $P_+$  and  $P_-$ , respectively)

The spectra are plotted in Fig. 3, taken by a transmission-spectrum analyzer (Apex AP2440A). The bare resonance is seen in Fig. 3(a) as the blue curve. The modulator’s  $p$ - $n$  junction is forward biased by currents  $I_b$  such that dynamic inputs induce an index change through free-carrier injection. When current bias is applied, the peak blue shifts to the left, seen over different rows. When heat is applied, it red shifts to the right, seen over the different columns. The orange, green, and red curves show different optical input states, described below.

Comparing the blue curves of (a)–(d) with (e)–(h), we see that thermal tuning does not significantly affect the response besides shifting the absolute wavelength with an efficiency of 0.24 nm/mW. This is desirable so that the thermal degree of freedom can be used as an independent parameter to lock the MRR onto a WDM channel of interest without affecting the electro-optic response. In contrast, the electrical-bias parameter does affect the depth and quality factor (from 14,500 to 3,500) of the peak. The electrical bias can be used to configure the response, while the combination of electrical and thermal parameters can be used to maintain the desired wavelength of interest. The electrical-tuning efficiency below the diode threshold is zero because no free carriers are injected. Above the diode threshold, the tuning efficiency is at most 0.26 nm/mA.

Two lasers are used to characterize the O-E-O response. When the optical inputs are nonzero, the effect is the same as modifying the electrical bias, because they deflect the current seen on the modulator. The orange curves represent the optical power (10 dBm at facet) on the IN+ port with the IN− laser turned off. The photocurrent adds to the bias current, causing a blue-shift. The fiber-to-fiber insertion loss through the chip is measured to be 18 dB, so we estimate that the input grating coupler represents 9.0 dB of loss. Therefore, the optical power reaching PD+ would be 1.26 mW. This causes a

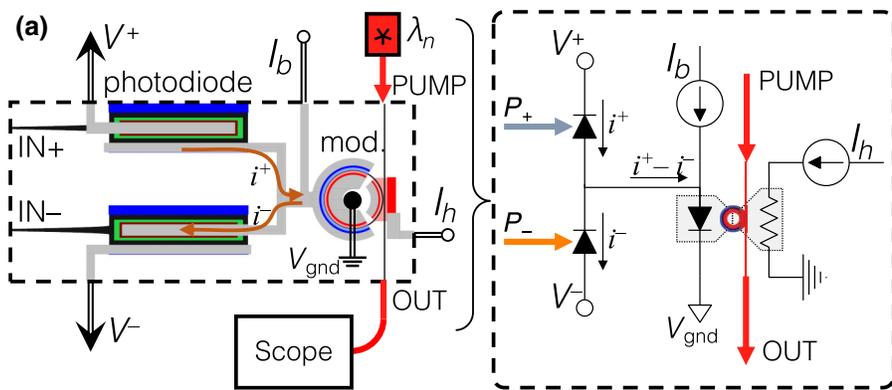
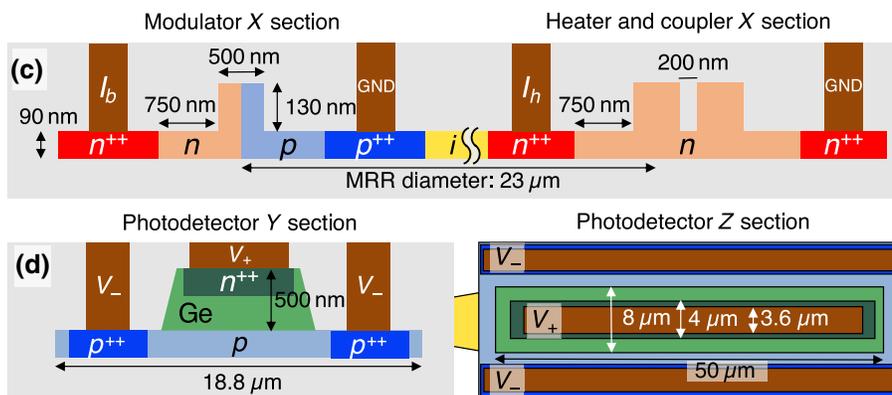
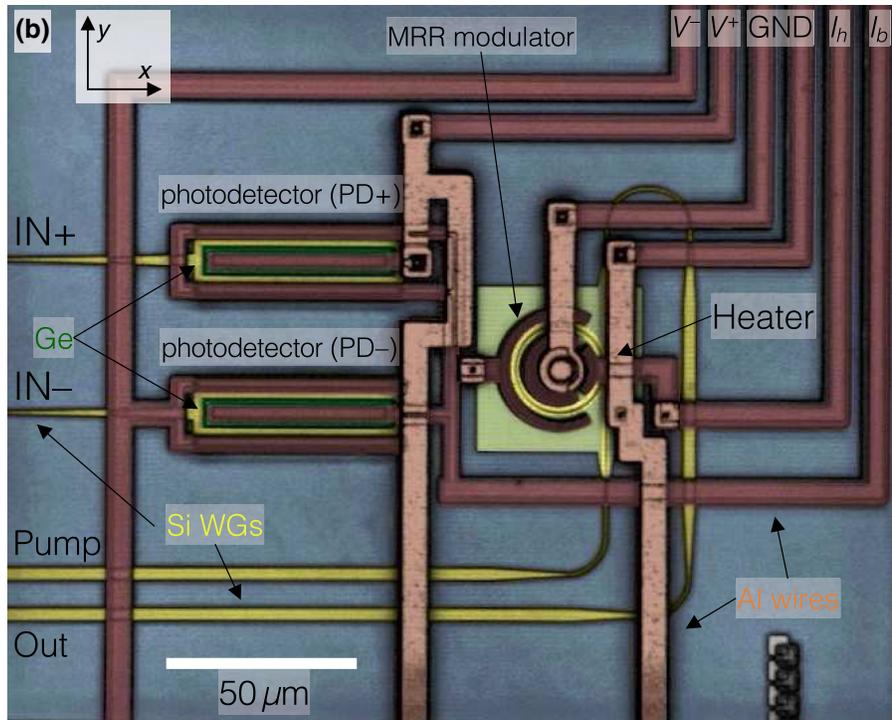


FIG. 2. (a) A simplified equivalent circuit diagram of the MRR modulator neuron. (b) A false-color confocal micrograph of the fabricated device. (c) A cross section of the MRR modulator with embedded *p-n* modulator and *n*-doped heater. (d) A cross section of the Si-Ge photodetector, showing the cathode ( $V_+$ ) and the anode ( $V_-$ ).



deflection equivalent to 0.96 mA, resulting in an estimated PD responsivity of 0.76 A/W, slightly less than in other reports [63]. The green curve is the opposite case: IN+ laser off and IN- laser at 10 dBm at facet. The red curve is with both lasers on. The fact that red curves match up with

blue (lasers off) curves is significant, because it shows that the +/- optical inputs cancel out. This means that excitatory and inhibitory effects can be realized by switching the port of impinging inputs. The microring weight banks [56,57] explored in other work were specifically designed

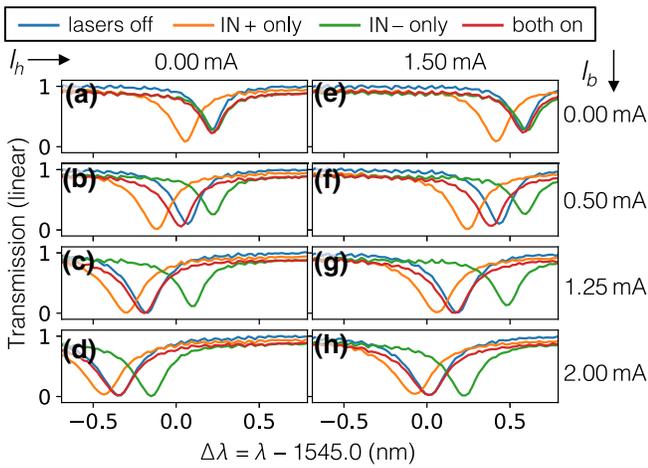


FIG. 3. Spectrum vs heat and forward-current bias under different permutations of illumination. Columns distinguish heating; rows distinguish the modulator bias current. Traces distinguish optical input conditions: blue, no light, just electrical bias; orange, green, only one optical input at a time; red, both inputs creating complementary effects. Ideally, the red traces should overlap with the blue.

to perform this task of directing multiple-wavelength channels between the ports of a balanced photodetector.

The optical-to-optical gain is measured by comparing the amplitude of the modulation swing on the  $\lambda_1$  input to the amplitude of the modulation swing on the  $\lambda_n$  output coming from the neuron. It is found to be  $2.16 \times 10^{-2}$ , dominated by fiber-to-chip insertion loss measured at 18 dB. Removing the coupling loss results in an on-chip gain estimate of 1.36. The gain scales in proportion to the pump

power, which is held artificially low during this measurement to eliminate the chance of the pump influencing the carrier concentration or temperature of the modulator. The pump power is measured as  $-5.5$  dBm at the chip facet, with an estimated  $-14.5$  dBm arriving at the neuron, less than 100 times weaker than it could be under realistic assumptions.

We use injection modulation because we find that the depletion effect is too weak to demonstrate the desired results with the nonoptimized modulators that we fabricate. Injection modulation with a forward-biased junction is slower than depletion modulation with a reverse-biased junction. The injection-modulation bandwidth can reach up to 6.25 GHz [64] compared to depletion modulators shown up to 40 GHz [65]. All of the signal-processing concepts of modulator-class neurons shown here hold over different types of modulation mechanisms.

### C. Experimental setup

The multichannel input generator used in the following experiments is shown in Fig. 4. All time-varying signals coming into and out of the chip are optical. Two input wavelengths ( $\lambda_1 = 1546.4$  nm and  $\lambda_2 = 1548.2$  nm) are wavelength-division multiplexed and power modulated by a fiber Mach-Zehnder modulator (MZM). Distinct input signals are then created by a relative delay,  $\Delta T$ , between these wavelengths, following the method of Ref. [66]. The laser pumps for the external inputs and the neuron itself come from a distributed-feedback (DFB) laser array (ILX 7900B). The  $\lambda_2$  signal can be switched between the excitatory and inhibitory ports.

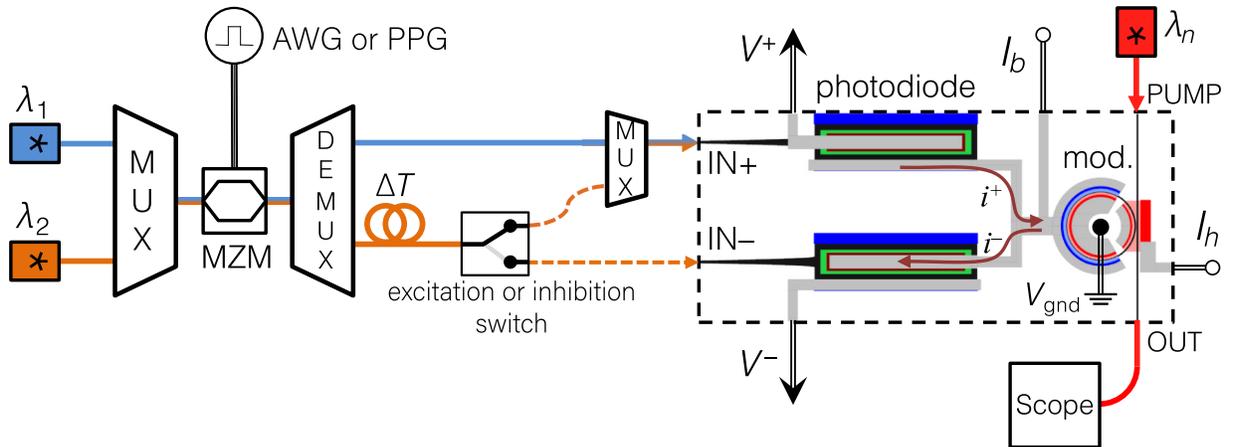


FIG. 4. The experimental setup. A Mach-Zehnder modulator (MZM), followed by a DEMUX and  $\Delta T$  delay, is used to create two distinct signals. An arbitrary-wave-form generator (AWG) or a PPG is used to drive the modulator. A fiber switch directs the second ( $\lambda_2$ ) input either to the same IN+ port as the first input ( $\lambda_1$ ) or to the complementary IN- port. Light going into IN(+/-) waveguides impinges on the photodetectors, resulting in complementary photocurrents  $i^{+/-}$ . Current injected to the MRR modulator (“mod.”) is the sum of the photocurrents and the bias current ( $I_b$ ):  $I_b + i^+ - i^-$ . The MRR resonance wavelength is also affected by a heater biased at  $I_h$ . The transmission of the modulator is probed by a cw pump (“PUMP”) at  $\lambda_n$ . Its optical output is monitored by an oscilloscope (“Scope”). The dashed box indicates the on-chip/off-chip boundary.

There are three signal generators used in the following experiments, two analog (a.k.a. synths) and one binary. A simple slow-wave-form generator (HP 8116A) is used to acquire the transfer functions (Sec. III A) and the autapse behavior (Sec. III E). The 8116A offers control of sawtooth wave forms that can be used to separate rising and falling aspects. Burst inputs are generated by a Rohde and Schwartz SMBV 100A VG (R&S), which is used in Secs. III A, B, and C. The R&S burst can also be viewed as trains of return-to-zero (RZ) pulses of varying amplitude. The binary-pulsed inputs used in Sec. III D are generated by a pulse-pattern generator (PPG) (Anritsu MP1761B). The PPG provides the highest instantaneous bandwidth but the least control over wave forms.

The neuron's output is coupled off-chip, detected, and observed in a sampling oscilloscope (Tektronix DSA8300). Between the output coupler and the oscilloscope, there is a signal-to-noise enhancement stage, not diagrammed, consisting of an erbium-doped fiber amplifier (EDFA), an optical band-pass filter at  $\lambda_n$ , a discrete photodetector (Discovery Semiconductors, Inc. DSC-R405ER), and, for low-bandwidth experiments, an electrical low-pass filter. The prechip subsystem contains two EDFAs and polarization controllers that are not diagrammed. All of the above instrumentation is controlled remotely via LIGHTLAB, a free PYTHON software package [67].

The chip containing the neuron is placed on a temperature-controlled alignment stage and kept in place by a back-side vacuum. Four fibers of a V-groove array are aligned to grating couplers leading to the IN+, IN−, PUMP, and OUT ports. DC bias signals are applied through six probe tips in an array (GGB MCW-26-8146). The power and ground signals are derived from a power supply, and the tunable biases,  $I_b$  and  $I_h$ , are derived from two Keithley 2400 current sources.

### III. RESULTS

We conduct a series of experiments on this silicon photonic device to support the claim that it exhibits the three properties necessary to act as a network-compatible neuron (nonlinearity, fan-in, and cascability), in addition to several optional but useful properties. Section III A demonstrates nonlinear optical-to-optical conversion and transfer-function configurability. Section III B demonstrates high-bandwidth operation and pulse compression. Section III C demonstrates fan-in and inhibitory fan-in. Section III D demonstrates time-resolved spike processing that is excitatory and inhibitory. Section III E demonstrates indefinite cascability through the observation of bistability in an autapse circuit.

#### A. Transfer functions

The photodetector-modulator system exhibits a variety of nonlinear optical-to-optical transfer functions that are relevant for a wide variety of neural-processing tasks. To obtain these responses, the optical input is modulated by a slow-rising sawtooth wave form at 200 kHz, derived from the HP synth. The type of response depends on the wavelength offset between the pump and the MRR resonance. To change the response, the MRR resonance is shifted by varying the heater bias.

Figure 5 shows six response shapes, each relevant in different areas of neural processing and machine learning. They are obtained under different biasing conditions. This variety indicates that the MRR modulator neuron can yield nonlinear configurability, to complement the configurability of the network's linear weights. The sigmoid shapes of Figs. 5(a) and 5(b) are commonly used in recurrent Hopfield networks for nonlinear optimization [68]. They are obtained by biasing at the maximum slope point. The rectified linear unit (ReLU) shape of Figs. 5(c) and 5(d)

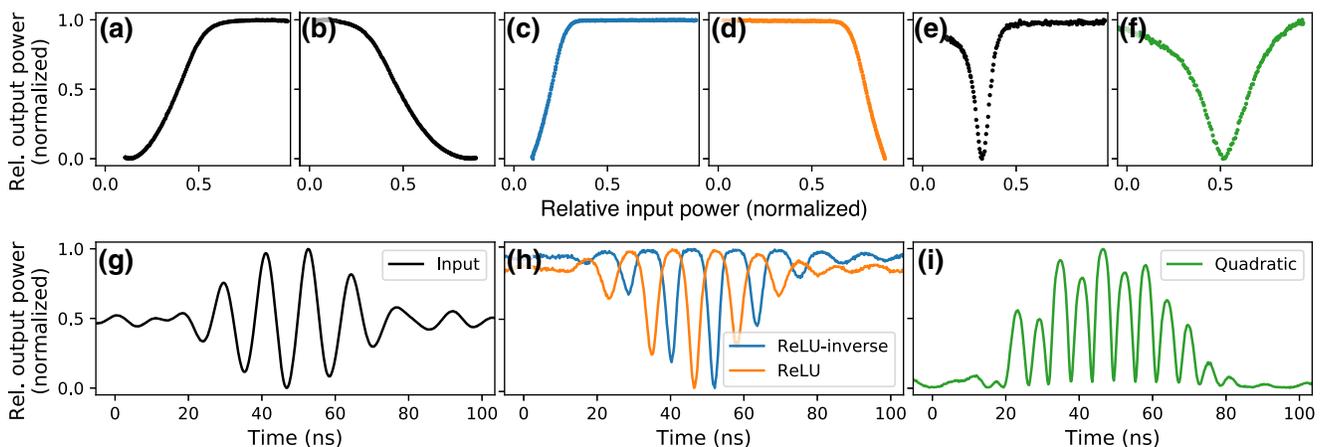


FIG. 5. (a)–(f) A variety of relevant O-E-O transfer functions seen from the PD-modulator pair, taken at different bias conditions: (a),(b) sigmoids; (c),(d) rectified linear units (“ReLU”); (e) a radial basis function (RBF); and (f) a quadratic function. (g)–(i) Time resolved pictures of these transfer functions: (g) the input is a 40-ns burst of a 100-MHz carrier; (h) both ReLUs; (i) the quadratic function. All plots display experimental data.

is widely used in feedforward-machine-learning networks today, i.e., in multilayer perceptrons (MLPs) and convolutional neural networks (CNNs) [20]. Positive and negative ReLUs are obtained by biasing slightly off resonance, either above or below the pump wavelength. A network that combines sigmoid and ReLU neurons is well suited to solving nonlinear optimization problems with constraints, some of which are reviewed in Ref. [69]. The peaked transfer functions of Figs. 5(e) and 5(f) are known as radial basis functions (RBFs). When biased on resonance, the RBF is centered at zero, resulting in a quadratic or rectifying transfer function. The off-centered RBF is obtained by setting the electrical bias to achieve the highest resonator  $Q$  factor, then thermally biasing far off resonance. RBFs are commonly used for ML based on support-vector machines [[70], Sec. 5.2].

These nonlinear responses stem from the electro-optic transfer function of the modulator. The photodetectors' O-E response is linear and the modulator's E-O response is a strongly nonlinear peaked Lorentzian function.

It is apparent that the different responses correspond to different pieces of the MRR modulator's Lorentzian peak shape. To some extent, the type of response observed depends on the amplitude of the input, which is not ideal. The ideal ReLU continues to increase indefinitely and the ideal sigmoid saturates indefinitely. We note, however, that every neuron has some limitation to its range of inputs, whether it be imposed by register-bit depth (digital) or state bleaching (analog, including biological).

## B. Response to high-bandwidth inputs

The response of the neuron to a much faster input is shown in Fig. 6. The input rf envelope is a 25.0-ns pulse of a 1.0-GHz carrier, which is modulated on an optical carrier at  $\lambda_1$ . The second input at  $\lambda_2$  is zeroed. The outputs indicate that, under different biasing conditions, the neuron can either reproduce or apply nonlinear transformations to a fast input. The linear response of Fig. 6(b)

is obtained by biasing on the maximum slope point and driving with a signal weaker than that needed to saturate the sigmoid of Fig. 5(b). This demonstration is significant, in part, because it is a faithful conversion of an rf signal from one wavelength,  $\lambda_1$ , to another,  $\lambda_n$ . A voltage gain of 0.2 is shown, although this is not necessarily representative of the fully integrated case because gain is affected by fiber-to-chip insertion loss and the fiber EDFA configuration.

In Fig. 6(c), the ReLU of Fig. 5(c) is biased below its elbow. Viewed on the burst level, it exhibits a thresholding phenomenon, resulting in a compressed burst width of 20.4 ns (19% compression). As seen in the enlarged traces of Figs. 6(d)–6(f), the ReLU also exhibits a pulse-compression effect. The 1 GHz input can be viewed as a train of 500-ps full-width-at-half-maximum (FWHM) pulses that are turned into a train of 324-ps pulses (36% compression).

Pulse compression is an important behavior for maintaining the integrity of pulses cascaded over multiple stages of neurons. It occurs because the leading and lagging pulse edges take on values that lie below the ReLU elbow. Both of these thresholding and compression behaviors are due to the positive third-order nonlinearity possessed by the ReLU. It is more common for optical devices to exhibit saturating behaviors with negative third-order nonlinearities; for example, those subject to two-photon absorption and laser diodes well above the lasing threshold. If saturating devices were used in a neural network, the pulse FWHM from stage to stage would increase (i.e., degrade) indefinitely.

## C. Response to multiple inputs

We then consider the case of two inputs at different wavelengths,  $\lambda_1$  and  $\lambda_2$ . We demonstrate fan-in behavior across wavelengths—the ability to perform a summation operation across multiple wavelengths while simultaneously converting them to a single wavelength,  $\lambda_n$ . Fan-in

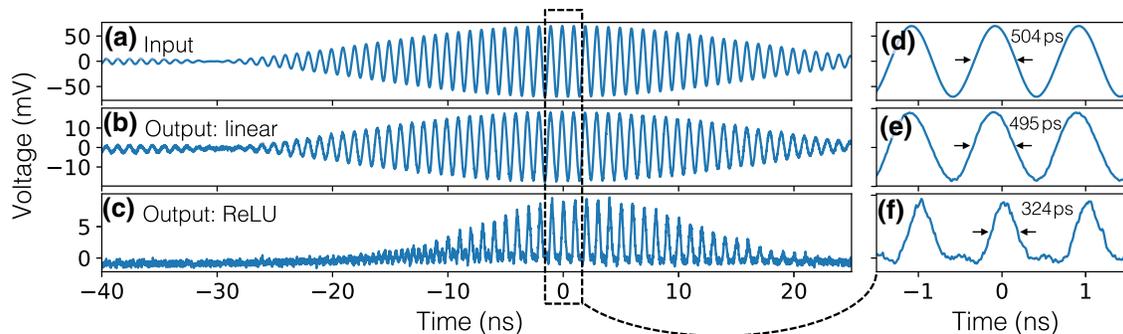


FIG. 6. Reproduction, rectification, and pulse compression. (a) The input is a 25.0-ns burst of a 1.0-GHz rf carrier modulated on a  $\lambda_1$  optical carrier. (b) In a linear regime, the modulating signal is faithfully reproduced on a different carrier wavelength,  $\lambda_n$ . (c) In the nonlinear ReLU regime, the output wave form is a rectified version of the input. (d)–(f) Enlarged windows interpreting the signals as pulsed signals: (d) input, (e) reproduction, and (f) rectification resulting in pulse-width compression.

is central to the idea of network-based processing, so it is particularly important to demonstrate this feature directly in order to claim a device as a photonic neuron.

The setup for addition and subtraction is shown in Fig. 4. The input envelopes are delayed versions of one another, i.e.,  $B(t) = A(t + \Delta T)$ . For subtraction, the two inputs are sent into complementary ports of the neuron's balanced PD. For addition, they are multiplexed and sent into the same port. This is represented by the excitatory or inhibitory switch in Fig. 4. The modulating signal, a 900-MHz rf carrier that is amplitude modulated (AM) at 50 MBaud, is chosen to demonstrate several behaviors of interest to rf signal processing, discussed below. The delay,  $\Delta T$ , is adjusted so that the rf carrier waves are in phase and the modulating bit pattern of [1, 1, 0] is delayed by one bit period.

Figures 7(a) and 7(b) show the inputs  $A(t)$  and  $B(t)$  after detection. Figure 7(c) is the optical output of the MRR neuron when these signals are multiplexed into the IN+ port. This illustrates basic optical fan-in of two WDM signals onto a single wavelength. Figure 7(d) is a more complex case, showing fan-in followed by a rectifying conversion from the electronic to the optical domain. Figures 7(e) and 7(f) illustrate inhibitory behavior where the inputs counteract one another—more than just an inversion of the excitatory case. Figure 7(e) is with a linear O-E-O transfer function and Fig. 7(f) is with a rectifying transfer.

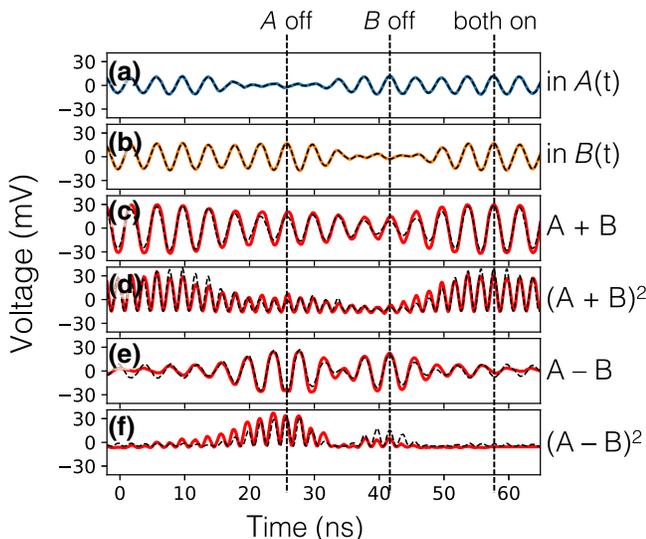


FIG. 7. Burst addition and two-channel rectification. The solid lines are experimentally measured; their colors correspond to the wavelengths in Fig. 4. (a),(b) Inputs. (c)–(f) Outputs under different biasing configurations performing the two-variable functions on the right. The dashed black curves are ideal outputs calculated from the measured inputs. The functions shown demonstrate linear and/or nonlinear capabilities as well as excitatory and/or inhibitory capabilities. Input  $B$  is slightly stronger than input  $A$ .

The fact that input optical signals effect changes in the output optical signal is significant because that output could, in principle, be fed to other neurons; furthermore, the fact that multiple signals can be “weighted” by positive and negative values, with their sum then influencing the output, is an indicator that the MRR neuron can be networked with multiple inputs and outputs. The photodetector-modulator device can also be viewed as a wavelength converter with higher gain than all-optical approaches [71]. Wavelength conversion on a conventional silicon platform has uses outside of neuromorphic photonics.

The signals themselves are chosen to show basic signal-processing tasks for typical AM radio signals. Linear fan-in is used in rf processing for dimensionality reduction and principal-component analysis [72]. Followed by a rectifying stage, this system effectively acts as a multichannel envelope detector [i.e.,  $(A + B)^2$  and  $(A - B)^2$ ]. The expansion of these squares yields a  $\pm 2AB$  term—a measure of correlation. In radio signal processing, some algorithms for channel estimation are based on calculating correlation to update estimation parameters [72,73]. Fast measurements of correlation could accelerate the convergence time and tracking rate of those algorithms.

#### D. Response to pulses

We examine the response of the neuron to two pulsed signals and show that the modulator neuron can be made to exhibit enhancing and saturating nonlinearity as well as inhibition in the pulsed domain. The significance of using pulses is the demonstration of time-resolved processing. When the magnitude of the response is dependent on the precise timing of inputs, the neuron is then capable of tasks that involve temporally coded signals.

In this experiment, the input is derived from the PPG, producing a pulse doublet with 2.0 ns width and 30 ns (15-bit periods) interpulse delay. This signal is modulated onto  $\lambda_1$  and  $\lambda_2$ , which are then delayed separately. The inter-channel delay is adjusted to correspond to the interpulse delay, such that the middle pulses coincide. At the same time, the leading and lagging pulses show the noncoincident response for comparison. The inputs post delay are shown in Figs. 8(a) and 8(b). Figure 8(c) compares the enhancing, saturating, and inhibitory cases to the linear sum of inputs (black trace). The traces are normalized to 0.5 at the first noncoincident pulse. This means that the linear sum at coincidence is 1.0.

Pulse-coincidence detection is shown in the blue trace of Fig. 8(c). The coincident pulse peaks at 1.57, meaning that coincidence is overemphasized. Enhancement is observed when the transfer function corresponds to the ReLU in Fig. 5(c), biased in the flat region. One pulse is not sufficient to reach the elbow, while two pulses are. We note

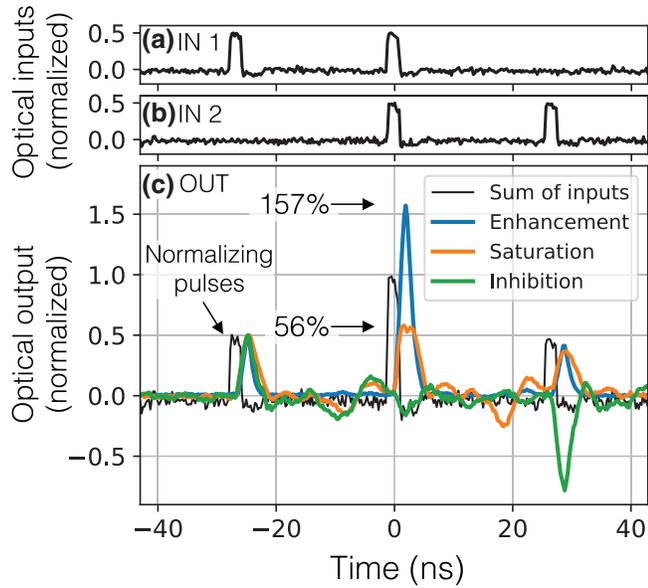


FIG. 8. Pulse-coincidence experimental results. (a),(b) The input is a 2-ns pulse doublet, which is delayed between two input wavelengths to create a pulse coincidence at  $t = 0$ . (c) Neuron outputs under different biasing conditions, compared to the linear solution (sum of inputs, black). The first pulse is used as a normalization to 0.5, meaning that the linear sum should peak at 1.0 (i.e., 100%): blue, biased for nonlinear enhancement, where coincidence is 157% of the linear solution; orange, biased for nonlinear saturation, where coincidence is 56% of the linear solution; green, inhibition; the fiber connections are switched such that IN 2 is directed to the IN– photodetector, and the neuron is biased in a linear region. The coincident pulses result in approximately zero output, and the noncoincident pulses cause complementary perturbations. The leading pulse from IN 1 results in a positive perturbation, while the lagging pulse from IN 2 results in a negative perturbation.

that the single-pulse suppression is not complete because their bandwidth is near the 3-dB bandwidth of the modulator. Coincidence detection is considered to be an essential property of nonlinear units within a pulsed neural network. A positive third-order nonlinearity resulting in pulse compression (shown in Fig. 6) is also essential for cascability in a pulsed neural network, because it counteracts pulse-width spreading. In the enhancement experiment here, the output pulses are not compressed. They have a FWHM of 3.8 ns (90% spread) due to the limited bandwidth of injection modulation. Spreading can be decreased by using longer pulses or faster modulation, such as depletion mode.

Pulse saturation is shown in the orange trace in Fig. 8(c). Again, the output is normalized to the leading pulse. In this case, the coincident pulses evoke about the same response as the noncoincident ones. The response corresponds to the inverse ReLU in Fig. 5(d), biased in the sloped region. The negative deflection around 20 ns is a spurious artifact due to carrier relaxation. Pulse inhibition

is shown in the green trace in Fig. 8(c). Signal IN 2 is switched to the IN– port of the neuron. The noncoincident pulses can be seen to produce positive (leading-pulse) and negative (trailing-pulse) effects. When the pulses coincide, the output vanishes. This result indicates more than an inversion of the response from the last experiment but the simultaneous capability for complementary excitable and inhibitory responses. In a real network, fibers cannot be switched. Instead, wavelength channels can be switched between IN+ and IN– ports by thermally configurable optical filters [57].

### E. Cascadability

In order for neurons to be cascable, signals must maintain their integrity from stage to stage indefinitely. The condition of gain cascability can be expressed as the existence of an operating point where large-signal gain is unity and small-signal differential gain,  $g$ , is greater than unity. The condition of physical cascability is expressed, for photonic neurons, as the input and output both being optical and at the same wavelength. These conditions can be observed using a self-afferent neuron—also known as an autapse. The autapse is a nonlinear feedback system the input and output of which are the same signal by definition. As derived in Appendix A, such a system becomes bistable when its differential gain is greater than unity. Therefore, an observation of a transition between mono- and bistability is equivalent to a demonstration of cascability.

Here, we construct an autapse circuit and observe this bifurcation therein, providing evidence that the MRR modulator neuron is indefinitely cascable.

The autapse setup is shown in Fig. 9(b). There is one external input that is modulated by a kilohertz-time-scale sawtooth wave form that is either rising or falling. The output of the neuron at  $\lambda_n$  is fed back via fiber and multiplexed with this external signal. Then both are fed into the positive PD. In line with the feedback pathway is a polarization controller (PC), a manually variable feedback attenuator ( $W_F$ ), an optical tap to the oscilloscope, and an EDFA to counter fiber-to-chip insertion loss.

Figure 9(c) shows the baseline input-output relation for the case where  $W_F = 0$ . The output (blue) is a regular function of the input, as seen in blue in Fig. 9(h). Figures 9(d)–9(g) show the rising and falling cases where  $W_F = 1$ . The output now depends on the direction of the input (black). Figure 9(h) shows each output plotted vs the input. The colors correspond to the left panels. In Figs. 9(d)–9(g), the bias is the same, whereas the bias is different in Fig. 9(c) because the average input optical power is less. These two different qualitative behaviors reveal the presence of a cusp bifurcation between  $W_F \in [0, 1]$ .

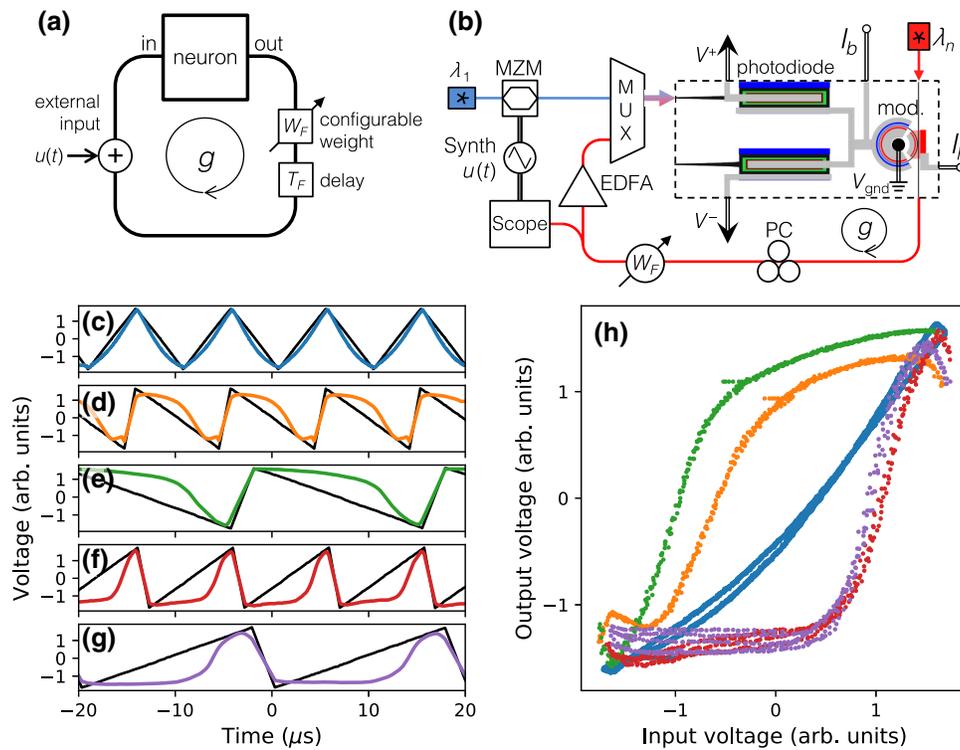


FIG. 9. The autapse concept and experiment. (a) An autapse is a neuron fed back to itself with a unitless round-trip gain of  $g$ . (b) The experimental setup. The output of the neuron on  $\lambda_n$  is wavelength multiplexed with a sawtooth-modulated external input  $u(t)$  on  $\lambda_1$ . Both are then fed into the positive-input photodetector. An oscilloscope records the output and the external input. (c) The response to a triangle-wave input (black), showing monostability. The feedback attenuator is blocking so that  $W_F = 0$ . (d)–(g) The response with feedback  $W_F = 1$  and  $g > 1$ , showing bistability. This is apparent from the difference in falling (d) and (e) and rising (f) and (g) edge responses. Changing the input from 100 kHz (d),(f) to 50 kHz (e),(g) makes little appreciable difference. (h) The output plotted against the input, illustrating the opening of a bistable regime. The colors correspond to the output traces in (c)–(g).

## IV. DISCUSSION

### A. Claiming cascability and fan-in

Many all-optical and optoelectronic elements exhibit nonlinear or neuron-like input-output transfer functions, but this is not a sufficient condition for the said devices to be capable of networking with other like devices. Optics faces special challenges in satisfying the critical requirements of cascability and fan-in, which we will refer to together as network compatibility. These fundamental challenges were expressed by Keyes and Goodman more than 30 years ago [74,75], and they remain challenges today. Due to their historical difficulty and central importance, we believe that there should be a high burden of proof placed on proposed photonic neurons to clearly demonstrate cascability and fan-in. Here, we will consider how other works have treated these concepts.

There are several physical properties that pose barriers to cascability [74]. In optical devices, the nonlinear mechanism of which is based on semiconductor carrier-mediated cross-gain modulation (XGM) or the Kerr effect, the controlled signal (probe) affects the material properties in the same way as the controller signal (pump). This

necessitates weak probes and very small pump-to-probe gains (e.g., the fiber neurons in Ref. [39,76]). Optical resonators can be used to strengthen these nonlinear effects at the cost of imposing wavelength constraints. In some approaches, the output wavelength must be different than the input (e.g., with Kerr [77] and with XGM [42,49,78]); in these cases, one gate cannot drive another.

In some recent cases, cascability has been used to refer to a single feedforward connection [44,49,79,80] but we argue that this approach does not conclusively demonstrate indefinite physical cascability. The format of the upstream input can be different than that of the downstream output such that the second might not be able to drive the first.

In this and prior works, we have argued that observing bifurcation in an autapse provides sound empirical evidence of indefinite cascability and fan-in. By definition, the output and input signals are the same, and the upstream and downstream neurons are the same. The recurrent circuit is equivalent to a neuron driving an indefinite feedforward chain of other identical neurons. To drive the autapse between bistable equilibria, an external input is needed. There must be a way for both the external input and the

feedback signal to influence the same neuron. In other words, the experiment also verifies the presence of a fan-in mechanism.

There are several works that have successfully approached cascadability by avoiding all-optoelectronic signal pathways and instead using an O-E-O chain consisting of a photodetector connected to a laser [48,81,82] or modulator [53,71]. Wavelength constraints and phase sensitivity vanish because this information is lost in the electronic domain. In addition, the E-O conversion step can offer strong nonlinearity, as employed here, and the electronic domain itself offers efficient mechanisms for nonlinearity and amplification. In Ref. [83], an O-E-O neuron based on cryogenic silicon LEDs, superconducting detectors, and superconducting amplifiers [84] was proposed. Its physical cascadability was demonstrated by the E-O-E LED-detector link shown in Ref. [85], and its gain cascadability has been addressed in more recent simulation works [81,82]. A potential downside of O-E-O is a vulnerability to electrical parasitics; however, these parasitics can remain small regardless of the network scale because O-E-O occurs entirely within a neuron, not between neurons.

When light combines, it interferes, posing a fundamental challenge to fan-in [75]. Optical fan-in results in either phase dependence, when coherent, or  $N$ -fold loss, when incoherent (e.g., 3 dB at  $N = 2$ ). In some all-optical devices where the in-out wavelengths can be the same (cascadable), these wavelengths also must be the same, meaning they cannot have more than one input [44,79]. Fan-in with coherent signals can be achieved by exerting complete control of the optical phase in the interconnect [41] but then signal-dependent phase changes in a neuron profoundly affect the behavior of the subsequent interconnect, precluding any cascadability. In Ref. [41], neuron calculations were implemented at low speed in a CPU; a neuron based on saturable absorption was contemplated, but it was not discussed how this element would regenerate a consistent optical phase.

Fan-in has also been achieved using inputs that are coherent but mutually incoherent, such as different spatial modes [86,87], different polarizations, or different wavelengths [66,88,89]. These signals do not interfere and, since they are individually coherent, can be multiplexed and routed and/or weighted independently by tunable resonators [58,90]. The total power is sensed by a photodetector (O-E), making this fan-in approach compatible with the O-E-O approach to cascadability. Multiwavelength-weighted addition was combined with O-E-O laser neurons in Ref. [48,50], wherein cascadability was also considered but not directly demonstrated. A downside of relying on multiple wavelengths is the need for a different laser source for each channel. The size of a single all-to-all subnetwork is capped by the available spectrum and the ability to distinguish adjacent channels, found in Ref. [91] to be less

than 950 if using the resonators of Ref. [92]; however, multiples of these subnetworks could be interfaced on a single chip [55].

## B. Nonspiking photonic neurons

The great majority of the work on photonic neurons has focused on lasers that implement spiking models similar to biological neurons [42–50], reviewed in Ref. [51]. To claim a nonspiking modulator as a photonic neuron represents a departure from the origins of neuromorphic photonics. Modulator-class photonic neurons are easier to fabricate and dissipate less power on chip. On the other hand, the processing repertoire of spiking neurons is theoretically greater than that of continuous-valued neurons.

The question is whether a move to continuous-valued photonic neurons would forfeit too much processing richness. Spiking networks can perform any continuous-valued task by representing values as a mean firing rate (MFR). These are converted to analog signals by low-pass filtering. A mathematical derivation of the MFR transform is found in Ref. [93]. In many applications, a valid alternative is to use physically analog signals to represent the MFR directly. There is no debate over whether continuous-valued neurons have numerous engineering applications. For example, in early artificial neural networks, the spike rate is represented by the analog voltage [12,13]. Modern machine learning almost universally makes the MFR transform, representing the MFR as a floating-point number. Another example is the Neural Engineering Framework (NEF) compiling process [94,95], which only looks at MFR-to-MFR transfer functions. By replacing the MFR transfer functions with analog optical-to-optical transfer functions, it was found that the entire NEF is still applicable to networks of modulator neurons [53].

There are at least four situations in which spiking artificial neurons have essential motivation; without these motivations, continuous-valued neurons are just as applicable. These motives include (1) studying brain function [18,26], (2) a need for increased robustness to amplitude noise, and (3) the use or study of temporal coding. Temporal coding has been proposed as one of the keys to energy efficiency in biological and perhaps artificial neural networks [96]; however, this proposition is debated and is the subject of active theoretical research (generally originating from Ref. [97]). Finally, a need for spiking can stem from (4) hardware constraints. The modern wave of neuromorphic electronics is based exclusively on spiking neural models because metal wires perform poorly when implementing dense, analog interconnects. A digital packet routing scheme (a.k.a., address event representation) is used to multiplex a small number of wires to create a large number of virtual interconnects [7,98]. In these cases, the packets are treated as spikes.

There is no fundamental reason why photonics must have one of these motivations, so we believe that continuous-valued modulator neurons do not sacrifice too much computational richness. We postulate that, due to their complementary nature, laser and modulator neurons will come to address complementary application domains. It is even possible that they could exist on the same chip to handle different types of tasks within a single neural network.

### C. Caveats of the fiber-based autapse

Both here and in Ref. [53], the feedback loop used for an autapse partially consists of fiber, resulting in substantial gain, frequency, and robustness discrepancies between the experiment and a fully integrated autapse. The insertion loss of fiber-to-chip couplers necessitates an EDFA in the feedback loop, weakening the claim that we have directly demonstrated gain cascading. Direct characterization yields a gain of 1.36, accounting for fiber-to-chip insertion loss and with an artificially weak pump (Sec. II B); device-level calculations indicate that, without fiber-to-chip insertion loss, gain can readily exceed unity.

The kilohertz frequencies of the input sawtooths in Figs. 9(c)–9(g) are chosen to isolate mono- and bistable equilibrium effects from dynamical effects. If the frequency were too high, then the long fiber feedback delay would lead to obfuscating time-delayed dynamics [99]. On the other hand, a modulation bandwidth that is too slow will also obscure equilibrium dynamics governed by carrier-injection modulation. The thermal dissipation time constant is near 1 ms, and the EDFA gain population time constant is 9 ms. The temperature and population states effectively act as integrators around 1 kHz and 100 Hz, respectively. The resulting phase lag in the output closely resembles bistability. To eliminate the possibility that we are actually observing phase lag dynamics, two different frequencies (50 kHz and 100 kHz) are compared in Figs. 9(d) vs 9(e) and Figs. 9(f) vs 9(g). The fact that they produce very similar responses provides evidence that we are observing equilibrium effects, as opposed to dynamic effects that are frequency dependent.

The fiber setup has poor environmental robustness. The couplers are polarization sensitive, and the feedback fiber can drift. This means that the polarization controller must be tuned relatively often to maintain a consistent round-trip gain. The average optical power coming out of the modulator is nonzero, thereby affecting the steady-state bias on the modulator. To further complicate that effect, the average power is weighted by the same amount as the varying signal (see Eq. (5) of Ref. [57]). To counter this effect, the prior work [53] used an inline capacitor as a dc block. That is not possible here because the key PD-modulator junction is fixed on chip, so we perform this correction manually

for the two feedback weight values. In a real network, it would be straightforward to calculate the bias deflection due to the weighted average power and then to correct the applied bias current to result in the desired net bias. We leave the automated version of that control algorithm to future work.

The fiber-based autapse has intriguing resemblances to photonic reservoir computers based on time-delayed feedback [47,100–103], although they represent distinct approaches to computing. In fiber reservoirs, the feedback fiber is used to induce complex dynamics, unlike autapse demonstrations, in which we attempt to minimize these dynamics in order to show isomorphism to a simple model. The reservoir concept complements neural-network approaches, as they are typically understood, precisely because it does not rely on maintaining a correspondence between hardware and a theoretical model. This means that reservoirs can employ continuous substrates that are easy to fabricate but difficult or impossible to make adhere to a model, such as optical phase around a fiber or a wave amplitude in a bucket of water [104]. A network of neurons can also, of course, act as a reservoir by remaining ignorant to its theoretical model [105]. Reservoir computers rely on instance-specific training to isolate dynamics desirable for processing—there is no guarantee that a particular hardware instance will exhibit the desired dynamics at all. In contrast, any artificial system isomorphic to an abstract neural network is guaranteed to be capable of all dynamics possible in that model.

### D. Further work

The results of this paper pose questions for further research. It is apparent how the modulator neuron introduced in this work could be co-integrated with the microring weight of Ref. [53], shown together in Fig. 1. An important demonstration for future work will be a fully integrated autapse. An integrated autapse would not suffer from fiber-induced latency or fiber-to-chip coupling loss. As hypothesized in Ref. [[106], Sec. 4.2], an integrated autapse could provide a means to accurately and experimentally quantify the energy consumption of photonic neurons.

Further work should look to reproduce these results at higher bandwidth. There are two bandwidth limiters: (1) the use of carrier injection modulation and (2) no isolation from the capacitance of the  $I_b$  connection. The first limitation could be addressed by optimizing the MRR modulator for stronger depletion modulation. The second could be addressed by an on-chip resistor and a series inductor. The capacitance of the modulator junction is affected by the metal between it and the dominant impedance of the biasing source. In this work, that includes the junction itself, its breakout trace to the pad, the probe contacting the pad, and the wire leading to the Keithley current source. An on-

chip resistor could constrain this capacitive region to just the junction itself.

The resonant wavelength of MRR modulators is highly sensitive to refractive index, a sensitivity shared by MRR filters used as weights. In this case, index changes result in large changes in the neural transfer function. On the one hand, this is a key advantage of resonator-based approaches since, in principle, very little additional tuning power is needed to reconfigure the response. On the other, this sensitivity results in poor environmental robustness, which can be addressed by one of two approaches. The first is precalibration, feedforward control, and, frankly, laboratory-grade isolation from the environment, as in Ref. [58]. The second is feedback control based on active monitoring of the optical intensity in the MRR, as in Ref. [56], which is feasible in nonlaboratory settings. High sensitivity also results in a poor ability to set the as-fabricated resonance to a desired wavelength, which is countered here by the  $n^+-n-n^+$  heater applying a relatively large baseline heating power. Eventually, this baseline power could come to dominate the system power [[106], Sec. 6.3.2]. This power contribution is viewed as a crucial, nontrivial, yet addressable engineering problem to explore in further work. Candidate approaches include postfabrication trimming [107], stronger athermal tuning mechanisms [108], and fabrication-tolerant MRR designs [109].

Further work could also explore computational behavior by combining the experimentally validated transfer functions of MRR modulator neurons with existing neural algorithms and compilers such as the NEF [94]. These compilers take a neuron transfer function and a high-level task specification, then returning a weight matrix. By mapping these weights back to hardware, the physical neural network will perform the task. This is analogous to a program compiler that takes a computer architecture and a program written in a high-level language, then returning machine code. Prior work [53,110] has demonstrated this strategy in photonics using the NEF. Likewise, the validated transfer functions could be injected into algorithms for machine learning [111].

## V. CONCLUSION

We fabricate a silicon photonic circuit consisting of two photodetectors electrically connected to a microring modulator. We demonstrate its ability to act as a network-compatible photonic neuron; namely, the properties of optical-to-optical nonlinearity, fan-in, and indefinite cascability. These properties have not been shown together in an integrated optoelectronic device.

Combined with the microring weight banks demonstrated in Ref. [53,56], the modulator neuron constitutes the final component needed to implement broadcast-and-weight [55] neural networks: to date, the only architecture to propose a way to unite well-defined photonic

neurons and well-defined neural interconnects upon a mainstream silicon photonic platform. Compatibility with silicon photonics provides crucial aspects of feasibility, scalability, and economies of scale to modern-day photonic neural networks. Due to their unprecedented potential speed, neuromorphic photonic systems can come to bear on society's growing demand for data-center machine learning and could, additionally, open up unexplored regimes of ultrafast machine intelligence.

## APPENDIX: EQUIVALENCE OF CASCADABILITY AND BIFURCATION

The assertion that a particular piece of semiconductor exhibits a mathematical isomorphism to a desired model should be supported by experiment. This evidence is important because if and only if a device is isomorphic to a neuron model can that neuron model be used to predict behaviors of larger networks of those devices.

Our approach to a rigorous photonic neural network demonstration is to derive qualitative predictions of a characteristic model, and then to observe them in experiment. In Sec. III E, we show a monostable-to-bistable bifurcation in an autapse in attempts to demonstrate cascability.

Here, we derive the equivalence between the cascability condition and a bifurcation threshold in a continuous-time neuron model, only the latter of which is observable as a qualitative phenomenon.

A simplified modulator neuron autapse is shown in Fig. 10. The calculations below assume that a voltage-mode depletion modulation effect is used instead of current-mode injection modulation because it will eventually be the preferred approach. The cascability condition

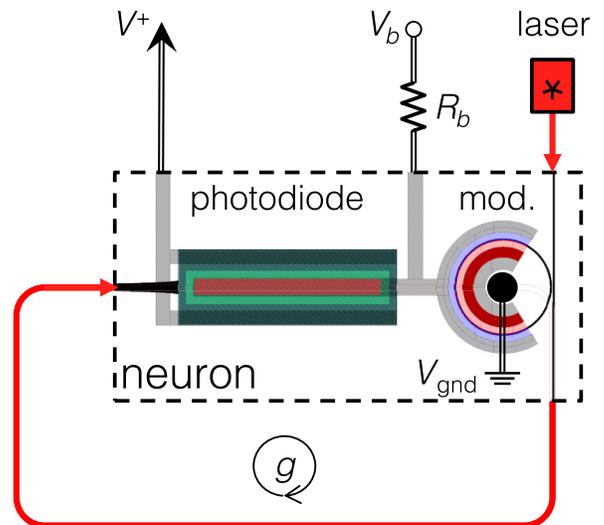


FIG. 10. A photonic autapse implemented by a microring modulator neuron. The modulator junction is reverse biased by a voltage source with impedance  $R_b$ .

is met when the neuron's differential optical-to-optical gain,  $g$ , exceeds unity.

### 1. Theory

The gain  $g$  can be derived from the device properties. It is defined as follows:

$$g = \frac{dP_{\text{out}}}{dP_{\text{in}}} \quad (\text{A1})$$

$$= \frac{dP_{\text{out}}}{dT_{\text{mod}}} \frac{dT_{\text{mod}}}{dV} \frac{dV}{dP_{\text{in}}}, \quad (\text{A2})$$

where  $T_{\text{mod}}$  is modulator transmission, and  $V$  is the junction voltage. The power leaving the modulator is as follows:

$$P_{\text{out}} = P_{\text{pump}} T_{\text{mod}}(V), \quad (\text{A3})$$

$$\frac{dP_{\text{out}}}{dT_{\text{mod}}} = P_{\text{pump}}, \quad (\text{A4})$$

where  $P_{\text{pump}}$  is the optical power entering the modulator. We assume that the modulator is biased at its maximum slope point for optimal modulation-slope efficiency. When there are no inputs,  $V = V_b$ , where  $V_b$  is the bias voltage. The optimal bias voltage is  $V_b^*$ :

$$V_b^* = \operatorname{argmax}_V \frac{dT_{\text{mod}}(V)}{dV}, \quad (\text{A5})$$

$$\left. \frac{dT_{\text{mod}}}{dV} \right|_{V_b=V_b^*} = \frac{\pi}{2V_\pi}, \quad (\text{A6})$$

where  $V_\pi$  is called the  $\pi$  voltage, used for notational convenience.  $V_\pi$  is inversely proportional to  $dT_{\text{mod}}/dV$ . It means the voltage that induces a  $\pi$  shift in the Mach-Zehnder modulator with equivalent modulation slope efficiency, not the voltage that induces a  $\pi$  shift around the MRR.

The detector and receiver component is as follows:

$$V = V_b + R_{\text{pd}} R_b P_{\text{in}}, \quad (\text{A7})$$

$$\frac{dV}{dP_{\text{in}}} = R_{\text{pd}} R_b, \quad (\text{A8})$$

where  $R_{\text{pd}}$  is the photodetector responsivity, and  $R_b$  is the biasing impedance. Combining the above terms,

$$g = P_{\text{pump}} \frac{\pi}{2V_\pi} R_{\text{pd}} R_b, \quad (\text{A9})$$

$$P_{\text{pump}} \Big|_{g=1} = \frac{2V_\pi}{\pi R_{\text{pd}} R_b}. \quad (\text{A10})$$

This pump power is the minimum needed for cascability.

### 2. Observation

The physical autapse circuit can be modeled as a dynamical system the state variable of which is the modulator voltage. A bifurcation is a qualitative change that occurs when the number or stability of fixed points changes. In the Supplementary Information of Ref. [53], the fixed points were derived with a simplified model. Here, we examine the fixed-point stability in a physical model and find the bifurcation point to be identical to the expression in Eq. A10.

The modulator junction has the following dynamics:

$$\frac{dV}{dt} = -\frac{V - V_b}{R_b C_{\text{mod}}} + \frac{i_{\text{pd}}}{C_{\text{mod}}}, \quad (\text{A11})$$

where  $C_{\text{mod}}$  is the capacitance of the modulator junction, and  $i_{\text{pd}}$  is the photocurrent. The photocurrent is induced by either the external signal,  $P_{\text{ext}}$ , or the self-feedback signal,  $P_{\text{in}}$ . Due to the feedback,  $P_{\text{in}} = P_{\text{out}}$ :

$$i_{\text{pd}} = R_{\text{pd}} [P_{\text{out}}(V) + P_{\text{ext}}]. \quad (\text{A12})$$

Zeroing the external input, this results in the following:

$$\frac{dV}{dt} = -\frac{V - V_b}{R_b C_{\text{mod}}} + \frac{R_{\text{pd}}}{C_{\text{mod}}} P_{\text{pump}} T_{\text{mod}}(V). \quad (\text{A13})$$

We do not seek the solutions for the steady-state values because they have an inelegant dependence on the parameters. Instead, we assess the stability using the Jacobian linearization. In general, dynamical systems are described by  $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ , where  $\mathbf{F}$  is a matrix of functions. They can be linearized around fixed points,  $\mathbf{x}_{\text{fp}}$ , as follows:

$$\lim_{\Delta \mathbf{x} \rightarrow 0} \Delta \dot{\mathbf{x}} = \mathbf{J} \Delta \mathbf{x}, \quad (\text{A14})$$

where  $\mathbf{J} \equiv \left. \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right|_{\mathbf{x}_{\text{fp}}}$ ,

in which  $\Delta \mathbf{x} = \mathbf{x} - \mathbf{x}_{\text{fp}}$ , and  $\mathbf{J}$  is called the Jacobian matrix. When an eigenvalue of  $\mathbf{J}$  is positive,  $\Delta \mathbf{x}$  will grow exponentially, meaning that the fixed point is unstable. In our one-dimensional case, the Jacobian is just a scalar,  $J$ . A bifurcation occurs when stability changes, meaning when  $J = 0$ . Of course, the exponential growth saturates at one of two other fixed points. In other words, in a one-dimensional system, this bifurcation is a transition from monostability to bistability.

Using the physical description of the autapse dynamical system from Eq. (A13),

$$J = -\frac{1}{R_b C_{\text{mod}}} + \frac{R_{\text{pd}}}{C_{\text{mod}}} P_{\text{pump}} \left. \frac{dT}{dV} \right|_{V_b^*}, \quad (\text{A15})$$

which crosses zero when

$$P_{\text{pump}}|_{J=0} = \frac{2V_{\pi}}{\pi R_{\text{pd}} R_b}. \quad (\text{A16})$$

Thus, the expression for the pump power where the autapse loses monostability corresponds exactly with that where the cascability condition is met in Eq. (A10).

### ACKNOWLEDGMENTS

The devices were fabricated at the IME A\*STAR foundry in Singapore. Fabrication support was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) Silicon Electronic-Photonic Integrated Circuits (SiEPIC) Program and the Canadian Microelectronics Corporation (CMC). National Science Foundation (NSF) (Grants No. ECCS 1642962 and No. DGE 1148900).

- [1] Jason S. Orcutt, Benjamin Moss, Chen Sun, Jonathan Leu, Michael Georgas, Jeffrey Shainline, Eugen Zraggen, Hanqing Li, Jie Sun, Matthew Weaver, Stevan Urošević, Miloš Popović, Rajeev J. Ram, and Vladimir Stojanović, Open foundry platform for high-performance electronic-photonic integration, *Opt. Express* **20**, 12222 (2012).
- [2] A. E.-J. Lim, Junfeng Song, Qing Fang, Chao Li, Xiaoguang Tu, Ning Duan, Kok Kiong Chen, R. P.-C. Tern, and Tsung-Yang Liow, Review of silicon photonics foundry efforts, *IEEE J. Sel. Top. Quantum Electron.* **20**, 405 (2014).
- [3] David Thomson, Aaron Zilkie, John E. Bowers, Tin Komljenovic, Graham T. Reed, Laurent Vivien, Delphine Marris-Morini, Eric Cassan, Léopold Virot, Jean-Marc Fédéli, Jean-Michel Hartmann, Jens H. Schmid, Dan-Xia Xu, Frédéric Boeuf, Peter O'Brien, Goran Z. Mashanovich, and M. Nedeljkovic, Roadmap on silicon photonics, *J. Opt.* **18**, 073003 (2016).
- [4] B. Marr, B. Degnan, P. Hasler, and D. Anderson, Scaling energy per operation via an asynchronous pipeline, *IEEE Trans. Very Large Scale Integration (VLSI) Syst.* **21**, 147 (2013).
- [5] Jennifer Hasler and Harry Bo Marr, Finding a roadmap to achieve large neuromorphic hardware systems, *Front. Neurosci.* **7**, 00 (2013).
- [6] Paul A. Merolla *et al.*, A million spiking-neuron integrated circuit with a scalable communication network and interface, *Science* **345**, 668 (2014).
- [7] S. B. Furber, F. Galluppi, S. Temple, and L. A. Plana, The SpiNNaker project, *Proc. IEEE* **102**, 652 (2014).
- [8] B. V. Benjamin, Peiran Gao, E. McQuinn, S. Choudhary, A. R. Chandrasekaran, J.-M. Bussat, R. Alvarez-Icaza, J. V. Arthur, P. A. Merolla, and K. Boahen, Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations, *Proc. IEEE* **102**, 699 (2014).
- [9] K. Meier, in *2015 IEEE International Electron Devices Meeting (IEDM)* (IEEE, Washington, DC, USA, 2015), p. 4.6.1.
- [10] M. Davies *et al.*, Loihi: A neuromorphic manycore processor with on-chip learning, *IEEE Micro* **38**, 82 (2018).
- [11] Jürgen Schmidhuber, Deep learning in neural networks: An overview, *Neural Netw.* **61**, 85 (2015).
- [12] John Von Neumann, Probabilistic logics and the synthesis of reliable organisms from unreliable components, *Automata Stud.* **34**, 43 (1956).
- [13] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain, *Psychol. Rev.* **65**, 386 (1958).
- [14] Maryam M. Najafabadi, Flavio Villanustre, Taghi M. Khoshgoftaar, Naeem Seliya, Randall Wald, and Edin Muharemagic, Deep learning applications and challenges in big data analytics, *J. Big Data* **2**, 1 (2015).
- [15] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE* **86**, 2278 (1998).
- [16] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, Deep learning, *Nature* **521**, 436 (2015).
- [17] Vivek K. Pallipuram, Mohammad Bhuiyan, and Melissa C. Smith, A comparative study of GPU programming models and architectures using neural networks, *J. Supercomput.* **61**, 673 (2012).
- [18] Alan Diamond, Thomas Nowotny, and Michael Schmuker, Comparing neuromorphic solutions in action: Implementing a bio-inspired solution to a benchmark classification task on three parallel-computing platforms, *Front. Neurosci.* **9**, 00 (2016).
- [19] Matthew D. Pickett, Gilberto Medeiros-Ribeiro, and R. Stanley Williams, A scalable neuristor built with Mott memristors, *Nat. Mater.* **12**, 114 (2013).
- [20] Norman P. Jouppi *et al.*, In-Datcenter Performance Analysis of a Tensor Processing Unit, arXiv:1704.04760 (2017).
- [21] Alex Graves *et al.*, Hybrid computing using a neural network with dynamic external memory, *Nature* **538**, 471 (2016).
- [22] Henry Markram, The human brain project, *Sci. Am.* **306**, 50 (2012).
- [23] A. S. Cassidy, P. Merolla, J. V. Arthur, S. K. Esser, B. Jackson, R. Alvarez-Icaza, P. Datta, J. Sawada, T. M. Wong, V. Feldman, A. Amir, D. B. D. Rubin, F. Akopyan, E. McQuinn, W. P. Risk, and D. S. Modha, in *The 2013 International Joint Conference on Neural Networks (IJCNN)* (IEEE, Dallas, TX, USA, 2013), p. 1.
- [24] D. Miyashita, S. Kousai, T. Suzuki, and J. Deguchi, A neuromorphic chip optimized for deep learning and CMOS technology with time-domain analog and digital mixed-signal processing, *IEEE J. Solid State Circuits* **52**, 2679 (2017).
- [25] Thomas Pfeil, Andreas Grübl, Sebastian Jeltsch, Eric Müller, Paul Müller, Mihai A. Petrovici, Michael Schmuker, Daniel Brüderle, Johannes Schemmel, and Karlheinz Meier, Six networks on a universal neuromorphic computing substrate, *Front. Neurosci.* **7**, 11 (2013).
- [26] Simon Friedmann, Nicolas Frémaux, Johannes Schemmel, Wulfram Gerstner, and Karlheinz Meier, Reward-based

- learning under hardware constraints—using a RISC processor embedded in a neuromorphic substrate, *Front. Neurosci.* **7**, 00 (2013).
- [27] W. Y. Tsai, D. Barch, A. Cassidy, M. Debole, A. Andreopoulos, B. Jackson, M. Flickner, J. Arthur, D. Modha, J. Sampson, and V. Narayanan, Always-on speech recognition using TrueNorth, a reconfigurable, neurosynaptic processor, *IEEE Trans. Comput.* **66**, 996 (2016).
- [28] Andrew Mundy, Ph.D. thesis, School University of Manchester, School of Computer Science, 2016.
- [29] S. Rakheja and V. Kumar, in *2012 13th International Symposium on Quality Electronic Design (ISQED)* (IEEE, Santa Clara, CA, USA, 2012), p. 732.
- [30] Andrew M. Weiner, Ultrafast optical pulse shaping: A tutorial review, *Opt. Commun.* **284**, 3669 (2011).
- [31] Daniel Pérez, Ivana Gasulla, Lee Crudgington, David J. Thomson, Ali Z. Khokhar, Ke Li, Wei Cao, Goran Z. Mashanovich, and José Capmany, Multipurpose silicon photonics signal processor core, *Nat. Commun.* **8**, 636 (2017).
- [32] Weilin Liu, Ming Li, Robert S. Guzzon, Erik J. Norberg, John S. Parker, Mingzhi Lu, Larry A. Coldren, and Jianping Yao, A fully reconfigurable photonic integrated signal processor, *Nat. Photon.* **10**, 190 (2016).
- [33] Bhavin J. Shastri, Alexander N. Tait, Thomas Ferreira de Lima, Mitchell A. Nahmias, Hsuan-Tung Peng, and Paul R. Prucnal, Principles of Neuromorphic Photonics, arXiv:1801.00016 (2018).
- [34] Thomas Ferreira de Lima, Bhavin J. Shastri, Alexander N. Tait, Mitchell A. Nahmias, and Paul R. Prucnal, Progress in neuromorphic photonics, *Nanophotonics* **6**, 00 (2017).
- [35] Daniel Brunner and Ingo Fischer, Reconfigurable semiconductor laser networks based on diffractive coupling, *Opt. Lett.* **40**, 3854 (2015).
- [36] Demitri Psaltis and Yong Quio, Optical neural networks, *Opt. Photon. News* **1**, 17 (1990).
- [37] J. W. Goodman, A. R. Dias, and L. M. Woody, Fully parallel, high-speed incoherent optical method for performing discrete fourier transforms, *Opt. Lett.* **2**, 1 (1978).
- [38] Praveen Asthana, Gregory P. Nordin, Armand R. Tanguay, Jr., and B. Keith Jenkins, Analysis of weighted fan-out/fan-in volume holographic optical interconnections, *Appl. Opt.* **32**, 1441 (1993).
- [39] M. T. Hill, E. E. E. Frietman, H. de Waardt, G.-D. Khoe, and H. J. S. Dorren, All fiber-optic neural network using coupled SOA based ring lasers, *IEEE Trans. Neural Netw.* **13**, 1504 (2002).
- [40] Jacques Carolan, Christopher Harrold, Chris Sparrow, Enrique Martí-López, Nicholas J. Russell, Joshua W. Silverstone, Peter J. Shadbolt, Nobuyuki Matsuda, Manabu Oguma, Mikitaka Itoh, Graham D. Marshall, Mark G. Thompson, Jonathan C. F. Matthews, Toshikazu Hashimoto, Jeremy L. O'Brien, and Anthony Laing, Universal linear optics, *Science* **349**, 711 (2015).
- [41] Yichen Shen, Nicholas C. Harris, Scott Skirlo, Mihika Prabhu, Tom Baehr-Jones, Michael Hochberg, Xin Sun, Shijie Zhao, Hugo Larochelle, Dirk Englund, and Marin Soljačić, Deep learning with coherent nanophotonic circuits, *Nat. Photon.* **11**, 441 (2017).
- [42] M. A. Nahmias, B. J. Shastri, A. N. Tait, and P. R. Prucnal, A leaky integrate-and-fire laser neuron for ultrafast cognitive computing, *IEEE J. Sel. Top. Quantum Electron.* **19**, 1 (2013).
- [43] W. Coomans, L. Gelens, S. Beri, J. Danckaert, and G. Van der Sande, Solitary and coupled semiconductor ring lasers as optical spiking neurons, *Phys. Rev. E* **84**, 036209 (2011).
- [44] Thomas Van Vaerenbergh, Koen Alexander, Joni Dambre, and Peter Bienstman, Excitation transfer between optically injected microdisk lasers, *Opt. Express* **21**, 28922 (2013).
- [45] Maia Brunstein, Alejandro M. Yacomotti, Isabel Sagnes, Fabrice Raineri, Laurent Bigot, and Ariel Levenson, Excitability and self-pulsing in a photonic crystal nanocavity, *Phys. Rev. A* **85**, 031803 (2012).
- [46] F. Selmi, R. Braive, G. Beaudoin, I. Sagnes, R. Kuszelewicz, and S. Barbay, Relative Refractory Period in an Excitable Semiconductor Laser, *Phys. Rev. Lett.* **112**, 183902 (2014).
- [47] B. Romeira, R. Avó, José M. L. Figueiredo, S. Barland, and J. Javaloyes, Regenerative memory in time-delayed neuromorphic photonic resonators, *Sci. Rep.* **6**, 19510 (2016).
- [48] Mitchell A. Nahmias, Alexander N. Tait, Leonidas Tolia, Matthew P. Chang, Thomas Ferreira de Lima, Bhavin J. Shastri, and Paul R. Prucnal, An integrated analog O/E/O link for multi-channel laser neurons, *Appl. Phys. Lett.* **108**, 151106 (2016).
- [49] T. Deng, J. Robertson, and A. Hurtado, Controlled propagation of spiking dynamics in vertical-cavity surface-emitting lasers: Towards neuromorphic photonic networks, *IEEE J. Sel. Top. Quantum Electron.* **23**, 1 (2017).
- [50] H. T. Peng, M. A. Nahmias, T. Ferreira de Lima, A. N. Tait, B. J. Shastri, and P. Prucnal, Neuromorphic photonic integrated circuits, *IEEE J. Sel. Top. Quantum Electron.* **24**, 00 (2018).
- [51] Paul R. Prucnal, Bhavin J. Shastri, Thomas Ferreira de Lima, Mitchell A. Nahmias, and Alexander N. Tait, Recent progress in semiconductor excitable lasers for photonic spike processing, *Adv. Opt. Photon.* **8**, 228 (2016).
- [52] Bhavin J. Shastri, Mitchell A. Nahmias, Alexander N. Tait, Alejandro W. Rodriguez, Ben Wu, and Paul R. Prucnal, Spike processing with a graphene excitable laser, *Sci. Rep.* **6**, 19126 (2016).
- [53] Alexander N. Tait, Thomas Ferreira de Lima, Ellen Zhou, Allie X. Wu, Mitchell A. Nahmias, Bhavin J. Shastri, and Paul R. Prucnal, Neuromorphic photonic networks using silicon photonic weight banks, *Sci. Rep.* **7**, 7430 (2017).
- [54] Paul R. Prucnal and Bhavin J. Shastri, *Neuromorphic Photonics* (CRC Press, CRC Press, Boca Raton, FL, USA, 2017).
- [55] A. N. Tait, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, Broadcast and weight: An integrated network for scalable photonic spike processing, *J. Lightwave Technol.* **32**, 4029 (2014).
- [56] Alexander N. Tait, Hasitha Jayatilleka, Thomas Ferreira De Lima, Philip Y. Ma, Mitchell A. Nahmias, Bhavin J. Shastri, Sudip Shekhar, Lukas Chrostowski, and Paul R. Prucnal, Feedback control for microring weight banks, *Opt. Express* **26**, 26422 (2018).
- [57] A. N. Tait, A. X. Wu, T. Ferreira de Lima, E. Zhou, B. J. Shastri, M. A. Nahmias, and P. R. Prucnal, Microring

- weight banks, *IEEE J. Sel. Top. Quantum Electron.* **22**, 00 (2016).
- [58] Alexander N. Tait, Thomas Ferreira de Lima, Mitchell A. Nahmias, Bhavin J. Shastri, and Paul R. Prucnal, Multi-channel control for microring weight banks, *Opt. Express* **24**, 8895 (2016).
- [59] R. Soref and B. Bennett, Electrooptical effects in silicon, *IEEE J. Quantum Electron.* **23**, 123 (1987).
- [60] Richard Soref, Silicon photonics: A review of recent literature, *Silicon* **2**, 1 (2010).
- [61] Lukas Chrostowski and Michael Hochberg, *Silicon Photonics Design: From Devices to Systems* (Cambridge University Press, Cambridge, United Kingdom, 2015).
- [62] Tom Baehr-Jones, Ran Ding, Ali Ayazi, Thierry Pinguet, Matt Streshinsky, Nick Harris, Jing Li, Li He, Mike Gould, Yi Zhang, Andy Eu-Jin Lim, Tsung-Yang Liow, Selin Hwee-Gee Teo, Guo-Qiang Lo, and Michael Hochberg, A 25 Gb/s Silicon Photonics Platform, arXiv:1203.0767 (2012).
- [63] Monireh Moayed Pour Fard, Glenn Cowan, and Odile Liboiron-Ladouceur, Responsivity optimization of a high-speed germanium-on-silicon photodetector, *Opt. Express* **24**, 27738 (2016).
- [64] Qianfan Xu, Sasikanth Manipatruni, Brad Schmidt, Jagat Shakya, and Michal Lipson, 12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators, *Opt. Express* **15**, 430 (2007).
- [65] A. Novack, Y. Liu, R. Ding, M. Gould, T. Baehr-Jones, Q. Li, Y. Yang, Y. Ma, Y. Zhang, K. Padmaraju, K. Bergmen, A. E. J. Lim, G. Q. Lo, and M. Hochberg, in *10th International Conference on Group IV Photonics* (IEEE, Seoul, South Korea, 2013), p. 7.
- [66] Alexander N. Tait, John Chang, Bhavin J. Shastri, Mitchell A. Nahmias, and Paul R. Prucnal, Demonstration of WDM weighted addition for principal component analysis, *Opt. Express* **23**, 12758 (2015).
- [67] Alex Tait, Thomas Ferreira de Lima, Philip Y. Ma, Aashu Jha, Hsuan-Tung Peng, Heidi Miller, and Paul R. Prucnal, *lightwave-lab/LIGHTLAB: Version 1.0.5* (2018).
- [68] J. J. Hopfield and D. W. Tank, "Neural" computation of decisions in optimization problems, *Biol. Cybern.* **52**, 141 (1985).
- [69] Ue-Pyng Wen, Kuen-Ming Lan, and Hsu-Shih Shih, A review of Hopfield neural networks for solving mathematical programming problems, *Eur. J. Oper. Res.* **198**, 675 (2009).
- [70] Corinna Cortes and Vladimir Vapnik, Support-vector networks, *Mach. Learn.* **20**, 273 (1995).
- [71] Kengo Nozaki, Shinji Matsuo, Takuro Fujii, Koji Takeda, Eiichi Kuramochi, Akihiko Shinya, and Masaya Notomi, in *Conference on Lasers and Electro-Optics* (Optical Society of America, San Jose, CA, USA, 2018), p. SF3A.3.
- [72] A. N. Tait, T. F. de Lima, P. Y. Ma, M. P. Chang, M. A. Nahmias, B. J. Shastri, P. Mittal, and P. R. Prucnal, in *52nd Annual Conference on Information Sciences and Systems (CISS)* (IEEE, Princeton, NJ, USA, 2018), p. 1.
- [73] S. Kutty and D. Sen, Beamforming for millimeter wave communications: An inclusive survey, *IEEE Commun. Surv. Tutorials* **18**, 949 (2016).
- [74] Robert W. Keyes, Optical logic—in the light of computer technology, *Opt. Acta: Int. J. Opt.* **32**, 525 (1985).
- [75] Joseph W. Goodman, Fan-in and fan-out with optical interconnections, *Opt. Acta: Int. J. Opt.* **32**, 1489 (1985).
- [76] David Rosenbluth, Konstantin Kravtsov, Mable P. Fok, and Paul R. Prucnal, A high performance photonic pulse processing device, *Opt. Express* **17**, 22767 (2009).
- [77] Qianfan Xu and Michal Lipson, All-optical logic based on silicon micro-ring resonators, *Opt. Express* **15**, 924 (2007).
- [78] S. Barbay, I. Sagnes, R. Kuszelewicz, and A. M. Yacomotti, in *2011 Fifth Rio De La Plata Workshop on Laser Dynamics and Nonlinear Photonics* (IEEE, Colonia del Sacramento, Uruguay, 2011), p. 1.
- [79] Thomas Van Vaerenbergh, Martin Fiers, Pauline Mechet, Thijs Spuesens, Rajesh Kumar, Geert Morthier, Benjamin Schrauwen, Joni Dambre, and Peter Bienstman, Cascadable excitability in microrings, *Opt. Express* **20**, 20292 (2012).
- [80] S. Y. Xiang, H. Zhang, X. X. Guo, J. F. Li, A. J. Wen, W. Pan, and Y. Hao, Cascadable neuron-like spiking dynamics in coupled vesels subject to orthogonally polarized optical pulse injection, *IEEE J. Sel. Top. Quantum Electron.* **23**, 1 (2017).
- [81] Jeffrey M. Shainline, Sonia M. Buckley, Adam N. McCaughan, Manuel Castellanos-Beltran, Christine A. Donnelly, Michael L. Schneider, Richard P. Mirin, and Sae Woo Nam, Superconducting Optoelectronic Neurons II: Receiver Circuits, arXiv:1805.02599 (2018).
- [82] Jeffrey M. Shainline, Adam N. McCaughan, Sonia M. Buckley, Richard P. Mirin, and Sae Woo Nam, Superconducting Optoelectronic Neurons IV: Transmitter Circuits, arXiv:1805.01941 (2018).
- [83] Jeffrey M. Shainline, Sonia M. Buckley, Richard P. Mirin, and Sae Woo Nam, Superconducting Optoelectronic Circuits for Neuromorphic Computing, *Phys. Rev. Appl.* **7**, 034013 (2017).
- [84] Adam N. McCaughan and Karl K. Berggren, A superconducting-nanowire three-terminal electrothermal device, *Nano Lett.* **14**, 5748 (2014).
- [85] Sonia Buckley, Jeffrey Chiles, Adam N. McCaughan, Galan Moody, Kevin L. Silverman, Martin J. Stevens, Richard P. Mirin, Sae Woo Nam, and Jeffrey M. Shainline, All-silicon light-emitting diodes waveguide-integrated with superconducting single-photon detectors, *Appl. Phys. Lett.* **111**, 141101 (2017).
- [86] F. B. McCormick, T. J. Cloonan, F. A. P. Tooley, A. L. Lentine, J. M. Sasian, J. L. Brubaker, R. L. Morrison, S. L. Walker, R. J. Crisci, R. A. Novotny, S. J. Hinterlong, H. S. Hinton, and E. Kerbis, Six-stage digital free-space optical switching network using symmetric self-electro-optic-effect devices, *Appl. Opt.* **32**, 5153 (1993).
- [87] Mable P. Fok, Yanhua Deng, Konstantin Kravtsov, and Paul R. Prucnal, Signal beating elimination using single-mode fiber to multimode fiber coupling, *Opt. Lett.* **36**, 4578 (2011).
- [88] Maroof H. Khan, Hao Shen, Yi Xuan, Lin Zhao, Shijun Xiao, Daniel E. Leaird, Andrew M. Weiner, and Minghao Qi, Ultrabroad-bandwidth arbitrary radiofrequency

- waveform generation with a silicon photonic chip-based spectral shaper, *Nat. Photonics* **4**, 117 (2010).
- [89] J. Chang, J. Meister, and P. R. Prucnal, Implementing a novel highly scalable adaptive photonic beamformer using “blind” guided accelerated random search, *J. Lightwave Technol.* **32**, 3623 (2014).
- [90] Qianfan Xu and Richard Soref, Reconfigurable optical directed-logic circuits using microresonator-based optical switches, *Opt. Express* **19**, 5244 (2011).
- [91] Alexander N. Tait, Allie X. Wu, Thomas Ferreira de Lima, Mitchell A. Nahmias, Bhavin J. Shastri, and Paul R. Prucnal, Two-pole microring weight banks, *Opt. Lett.* **43**, 2276 (2018).
- [92] Mohammad Soltani, Qing Li, Siva Yegnanarayanan, and Ali Adibi, Toward ultimate miniaturization of high  $Q$  silicon traveling-wave microresonators, *Opt. Express* **18**, 19541 (2010).
- [93] Ping Tak Peter Tang, Tsung-Han Lin, and Mike Davies, Sparse Coding by Spiking Neural Networks: Convergence Theory and Computational Results, arXiv:1705.05475 (2017).
- [94] Terrence C. Stewart and Chris Eliasmith, Large-scale synthesis of functional spiking neural circuits, *Proc. IEEE* **102**, 881 (2014).
- [95] Chris Eliasmith and Charles H. Anderson, *Neural engineering: Computation, representation, and dynamics in neurobiological systems* (MIT Press, Cambridge, MA, USA, 2004).
- [96] Rahul Sarpeshkar, Analog versus digital: Extrapolating from electronics to neurobiology, *Neural Comput.* **10**, 1601 (1998).
- [97] Wolfgang Maass, Networks of spiking neurons: The third generation of neural network models, *Neural Netw.* **10**, 1659 (1997).
- [98] F. Akopyan, J. Sawada, A. Cassidy, R. Alvarez-Icaza, J. Arthur, P. Merolla, N. Imam, Y. Nakamura, P. Datta, Gi-Joon Nam, B. Taba, M. Beakes, B. Brezzo, J. B. Kuang, R. Manohar, W. P. Risk, B. Jackson, and D. S. Modha, TrueNorth: Design and tool flow of a 65 mW 1 million neuron programmable neurosynaptic chip, *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **34**, 1537 (2015).
- [99] Bruno Romeira, Fanqi Kong, Wangzhe Li, José M. L. Figueiredo, Julien Javaloyes, and Jianping Yao, Broad-band chaotic signals and breather oscillations in an optoelectronic oscillator incorporating a microwave photonic filter, *J. Lightwave Technol.* **32**, 3933 (2014).
- [100] L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutiérrez, L. Pesquera, C. R. Mirasso, and I. Fischer, Photonic information processing beyond Turing: An optoelectronic implementation of reservoir computing, *Opt. Express* **20**, 3241 (2012).
- [101] M. Hermans, J. Dambre, and P. Bienstman, Optoelectronic systems trained with backpropagation through time, *IEEE Trans. Neural. Netw. Learn. Syst.* **26**, 1545 (2015).
- [102] S. Ortú, M. C. Soriano, L. Pesquera, D. Brunner, D. San Martín, I. Fischer, C. R. Mirasso, and J. M. Gutiérrez, A unified framework for reservoir computing and extreme learning machines based on a single time-delayed neuron, *Sci. Rep.* **5**, 14945 (2015).
- [103] Miguel C. Soriano, Daniel Brunner, Miguel Escalona-Morán, Claudio R. Mirasso, and Ingo Fischer, Minimal approach to neuro-inspired information processing, *Front. Comput. Neurosci.* **9**, 68 (2015).
- [104] Chrisantha Fernando and Sampsa Sojakka, in *Advances in Artificial Life*, edited by Wolfgang Banzhaf, Jens Ziegler, Thomas Christaller, Peter Dittrich, and Jan T. Kim (Springer, Berlin, Heidelberg, 2003), p. 588.
- [105] Wolfgang Maass, Thomas Natschläger, and Henry Markram, Real-time computing without stable states: A new framework for neural computation based on perturbations, *Neural Comput.* **14**, 2531 (2002).
- [106] Alexander N. Tait, Ph.D. thesis, School Princeton University, School of Engineering and Applied Science, 2018.
- [107] Payam Alipour, Amir H. Atabaki, Murtaza Askari, Ali Adibi, and Ali A. Eftekhar, Robust postfabrication trimming of ultracompact resonators on silicon on insulator with relaxed requirements on resolution and alignment, *Opt. Lett.* **40**, 4476 (2015).
- [108] L. Midolo, S. N. Yoon, F. Pagliano, T. Xia, F. W. M. van Otten, M. Lerner, S. Höfling, and A. Fiore, Electromechanical tuning of vertically-coupled photonic crystal nanobeams, *Opt. Express* **20**, 19255 (2012).
- [109] Timo Lipka, Lennart Moldenhauer, Jörg Müller, and Hoc Khiem Trieu, Athermal and wavelength-trimmable photonic filters based on TiO<sub>2</sub>-cladded amorphous-SOI, *Opt. Express* **23**, 20075 (2015).
- [110] T. F. de Lima, H. Peng, A. N. Tait, M. A. Nahmias, H. B. Miller, B. J. Shastri, and P. R. Prucnal, Machine learning with neuromorphic photonics, *J. Lightwave Technol.* **37**, 1515 (2019).
- [111] Jonathan K. George, Armin Mehrabian, Rubab Amin, Jiawei Meng, Thomas Ferreira de Lima, Alexander N. Tait, Bhavin J. Shastri, Tarek El-Ghazawi, Paul R. Prucnal, and Volker J. Sorger, Neuromorphic photonics with electro-absorption modulators, *Opt. Express* **27**, 5181 (2019).