Real-time Operation of Silicon Photonic Neurons

Thomas Ferreira de Lima[1,*], Chaoran Huang[1], Simon Bilodeau[1], Alexander N. Tait[1,2], Hsuan-Tung Peng[1], Philip Y. Ma[1], Eric C. Blow[1], Bhavin J. Shastri[2], Paul Prucnal[1]

1. Lightwave Communications Research Laboratory, Department of Electrical Engineering Princeton University, Princeton, NJ, 08544 USA

2. Physical Measurement Laboratory, National Institute of Standards and Technology, Boulder, CO 80305, USA

3. Department of Physics, Engineering Physics & Astronomy, Queen's University, Kingston ON, K7L 3N6, Canada

[*]Corresponding Author: tlima@princeton.edu

Abstract: In this paper, we use standard silicon-photonic components in order to implement a neuromorphic circuit with two neurons. The network exhibits reconfigurable weights and nonlinear transfer functions, enabling high-bandwidth analog signal processing tasks.

OCIS codes: 200.4700 Optical neural systems, 200.4740 Optical processing.

1. Introduction

Neuromorphic photonics has recently attracted attention as a promising avenue for specialized, "more-than-Moore" computing hardware. It is a photonic technology that is uniquely suited to process analog, multi-variate, high-bandwidth signals by using nonlinear units emulating neurons in artificial neural networks. Neuromorphic photonic circuits can be manufactured using silicon photonics standard fabrication processes, by use of active optoelectronic devices and homogeneous- or heterogeneously-integrated infrared light sources. Such processes are now accessible via low-cost fabrication prototyping services and open-source design kits [1].

In this paper, we demonstrate real-time functionality of a small set of neurons applying a nonlinear transformation on a pair of analog signals. As we will show, these neurons can be reconfigured to provide different transfer functions or apply different synaptic weights. Mastering these operations is crucial for using a neural network as an analog computer performing low-latency classification tasks and quadratic programs [2].

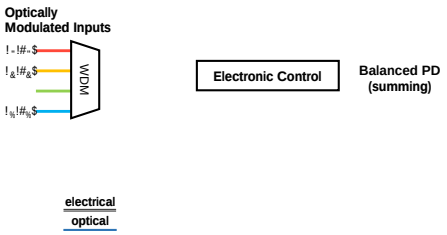

Fig. 1. (a) Description of the silicon photonic neural network under test in this experiment. This network is a subset of the larger network on chip, which consists of 4 neurons with recurrent connectivity (the output is partially coupled back to the input). (b) Experimental detail showing a V-groove fiber holder coupling light into the chip. The chip is mounted onto a plastic leadless chip carrier with 84 pins.

2. Device Description

The neural network was manufactured on a silicon photonic integrated circuit with high-speed optical ports coupled to optical fibers, and low-speed electric ports for biasing and configuration. Fig. 1(a) describes the integrated

circuit and its inputs and outputs. It is important to note, that the neuromorphic functionality of the circuit is entirely contained within the chip – the instruments outside the chip are used for generating and collecting signals, and biasing optoelectronic devices on chip, such as microring weight banks, microring modulators and photodetectors. Parts of this system were described and individually demonstrated using the same silicon photonics platform [3,4] (refer to these references for a more detailed experimental setup).

Here, we demonstrate simultaneous operation and control of two neurons networked together. There are two types of electrical traces in the circuit (Fig. 1(a)): a DC-type control traces used to bias the microring weight bank and the microring modulator (ports N, H, P, GND); and a high-speed link between the balanced photodiode, which is kept short to prevent parasitic capacitance and transmission-line effects. Only DC signals are connected outside the chip and into the printed-circuit board (PCB) (Fig. 1(b)). However, without proper biasing, AC signals leak to the PCB via the N-port. For that reason bypass capacitors were added on-chip and off-chip between ports V+/V– and P, as well as a ferrite bead in series before each port N. Without them, the PCB traces together with connected DC cables form a L-C resonator across V+/V– and N, which get excited by the balanced photodetector at around 50 MHz, which deforms the waveforms we are interested in investigating.

3. Single-variable Transfer Function

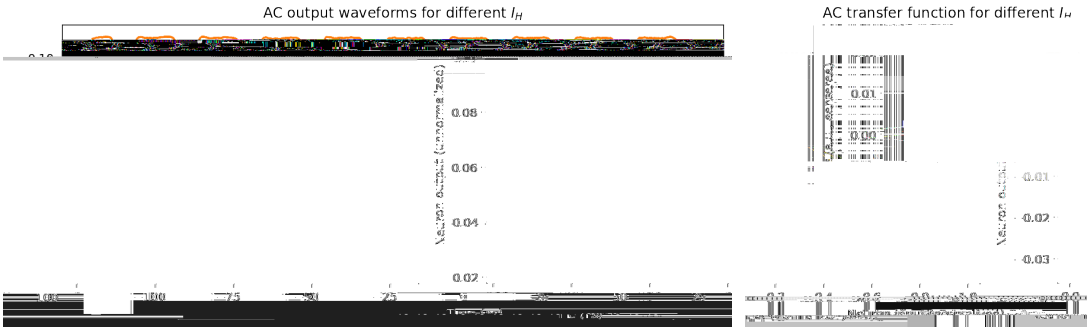

Fig. 2. (a) Neuron 1 output waveforms when excited by a 50 MHz sinusoidal waveform under different heater biasing currents applied to port H. (b) X-Y scatter plot corresponding to each biasing conditions in (a). These indicate an experimentally-measured AC transfer function. The x-axis was normalized for convenience, whereas the y-axis was centered around the lowest point of the sinusoid, in order to help visualize the transfer function shapes. The hysteresis effect is caused by limited bandwidth of the circuit – if bandwidth was infinite, the X-Y lines would not have holes inside them.

Biasing procedure First, we applied current onto the on-chip heater to bring each MRR modulator to resonance. Second, we apply a forward-bias current on the PN-junction of the modulator until the onset of the diode (requires a non-zero current because of the parallel R_{TIA} present in the circuit – see Fig. 1(a)). Then, we reverse bias each photodetector via ports V+/V– and N by 2 V. For simplicity, we biased the weight banks off-resonance so that all the optical power would impinge on the THRU photodetector, corresponding to a "–1" weight on the neuron.

Operating principle We send amplitude-modulated waveforms multiplexed into the resonant wavelength l_1, causing the generation of photocurrents proportional to the optical power (0:8 A=W responsivity). Because of the high-impedance of port N, all the current is directed across the modulator, which is the path of lowest impedance. This modulates the resonance wavelength of the modulator, affecting the amplitude of the output at that wavelength. We measured this output off-chip with a high-speed sampling scope (Fig. 2 (a)).

Individual neurons has Lorentzian-tail-shaped transfer functions (TFs) because of the microring modulators resonance optical spectra. These Lorentzian TFs can be manipulated by changing the biasing on each microring modulator (Fig. 2 (b)). Interestingly, we can create a range of TFs based on which side of the Lorentzian bell-shaped curve we bias the MRR on.

4. Multi-variable Transfer Function

Operating principle To demonstrate the functionality of a network, we explore real-time nonlinear transformation of a 2-dimensional time series. The biasing procedure and the operating principle are similar to the ones in the previous section. Here, however, we modulate a 32-bit PRBS waveform clocked at 294 MHz onto both l_1 (A)

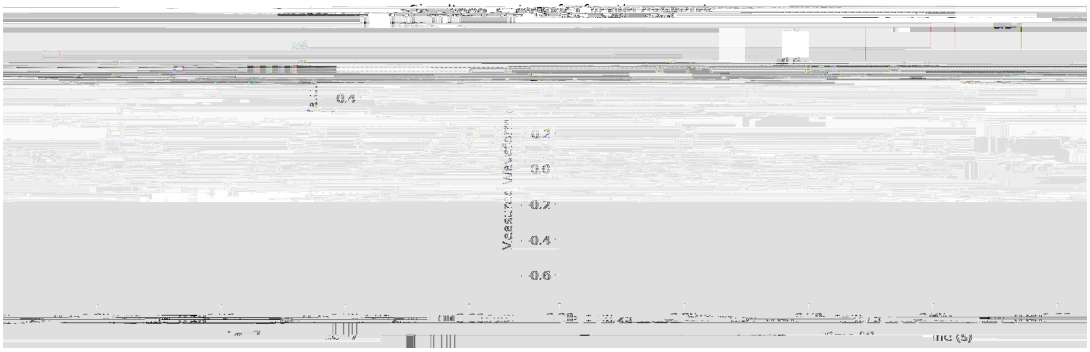

Fig. 3. Simultaneous transfer function of neurons 1 and 2. Input (black), Measured outputs of neuron 1 (red) and neuron 2 blue) after wavelength-demultiplexing (not shown in Fig. 1).

and l_2 (B). However, the one in l_2 is offset by exactly two bit periods. Since the weights are set to –1 each, both neurons are modulated with an inverted *A* + *B* waveform. Figure 3 shows the output of both neurons in this condition. We biased neuron 1 at 1.66 mA and neuron 2 at 1.68 mA so that they exhibited different transfer functions (Neuron 1 ReLU-like and Neuron 2 sigmoid-like). Neuron 1 is biased so that it suppresses the positive side of the input pattern whereas Neuron 2 amplifies it slightly. This behavior result matches the transfer functions shown in Fig. 2(b).

5. Conclusion and Outlook

There are two main advantages of architecting a photonic neuron with optical-electrical-optical (O/E/O) stages such as the one in this paper: first, it allows us to have high fan-in with high signal bandwidths thanks to wavelength-division multiplexing; second, it makes the neuron cascadable, since it generates an output spatially and spectrally isolated from its inputs. These two features allows us to network multiple layers of neurons, where one drives the next, to demonstrate low-latency nonlinear computations. They also allow us to connect the output of the neuron back to its input (not shown here), allowing for a recurrent neural network, enabling quadratic programming applications [2].

Acknowledgments

This research is supported by the Office of Naval Research (ONR) (N00014-18-1-2297), Defense Advanced Research Projects Agency (HR00111990049), and by the National Science Foundation (NSF) Enhancing Access to the Radio Spectrum (EARS) program, Grant No. (ECCS 1642962). Fabrication support was provided via the Natural Sciences and Engineering Research Council of Canada (NSERC) Silicon Electronic-Photonic Integrated Circuits (SiEPIC) Program and the Canadian Microelectronics Corporation (CMC).

References

1. L. Chrostowski, H. Shoman, M. Hammood, H. Yun, J. Jhoja, E. Luan, S. Lin, A. Mistry, D. Witt, N. A. Jaeger *et al.*, “Silicon photonic circuit design using rapid prototyping foundry process design kits,” IEEE J. Sel. Top. Quantum Electron. (2019).
2. T. Ferreira de Lima, H.-T. Peng, A. N. Tait, M. A. Nahmias, H. B. Miller, B. J. Shastri, and P. R. Prucnal, “Machine learning with neuromorphic photonics,” J. Light. Technol. **37**, 1515–1534 (2019).
3. A. N. Tait, H. Jayatilleka, T. Ferreira de Lima, P. Y. Ma, M. A. Nahmias, B. J. Shastri, S. Shekhar, L. Chrostowski, and P. R. Prucnal, “Feedback control for microring weight banks,” Opt. express **26**, 26422–26443 (2018).
4. A. N. Tait, T. Ferreira de Lima, M. A. Nahmias, H. B. Miller, H.-T. Peng, B. J. Shastri, and P. R. Prucnal, “Silicon photonic modulator neuron,” Phys. Rev. Appl. **11**, 064043 (2019).