An All-Optical Neuron for Scaling Integrated Photonic Neural Networks

Md Saiful Islam Sumon[1], Mihai Crisan[1], Weicheng You[1], Shrivatch Sankar[1], Imad I. Faruque[2], Sarvagya Dwivedi[3], and Shamsul Arafin[1*]
1. Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43201, USA
2. Quantum Engineering Technology Laboratories (QET Labs), University of Bristol, BS8 1TL Bristol, UK
3. Rockley Photonics, Pasadena, CA 91101, USA
arafin.1@osu.edu (Email address of corresponding author)

2023 IEEE Photonics Conference (IPC) | 979-8-3503-4722-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/IPC57732.2023.10360538

***Abstract*—We propose an all-optical neural network where signal processing requires no optical-to-electrical conversion. Weight multiplication, addition, and nonlinear activation of artificial neurons are performed in the photonic domain. Successful implementation will advance photonic neuromorphic computing, enabling practical solutions in artificial intelligence-driven tasks.**

***Keywords*—neuromorphic computing, optical neural networks (ONN), photonic neuron**

I. INTRODUCTION

Artificial neural networks (ANNs) have demonstrated an immense capacity to learn patterns while providing state-of-the-art performance in a variety of artificial intelligence tasks, particularly in voice and image recognition. However, scaling contemporary von Neumann computing architectures to meet the unsatiated computational demands of ANNs has hindered the development of a reliable and energy-efficient computing platform. As a result, there has been immense research in exploring alternative hardware platforms, such as neuromorphic computers [1,2] in an effort to both meet time- and energy-efficiency demands during ANN inferencing. While electronic neuromorphic computers such as Google's TPU and Intel's Loihi have shown drastic improvements in the reduction of ANN training times, such systems, however, are inherently prohibited by the hardware interconnectivity and the stringent cooling energy requirements [3]. This, in turn, has sparked interest in optical neural networks (ONNs), which are a leading form of analog-based neural networks that can be accelerated in the photonics domain.

Extensive research in the past decade on ONNs and photonics has demonstrated their potential as a promising alternative to digital neuromorphic processing. While photonic neuron-level implementations have shown moderate success, optical-electrical conversions at each neuron using optical-electronic-optical (O/E/O) links present challenges for large-scale photonic neural networks [4,5]. To overcome this challenge, we propose and design a programmable nanophotonic neuron (PNN) that eliminates the requirement of O/E/O conversion, leading to enhanced power efficiency and scalability benefits.

II. PHOTONIC NEURON

To implement a full-ONN purely within the optical domain (without signal conversion), our approach requires several key components as follows: (1) weight multiplication is achieved through a Mach-Zehnder interferometer (MZI) with an unused waveguide arm; (2) optical weighted addition is achieved through another MZI loaded with an external and internal phase shifter, performing 2×2 unitary operation, which is then controlled actively with feedback from an external photodetector [6]; (3) nonlinear activation function is achieved passively through optical bistability in a silicon microring resonator (MRR) [7].

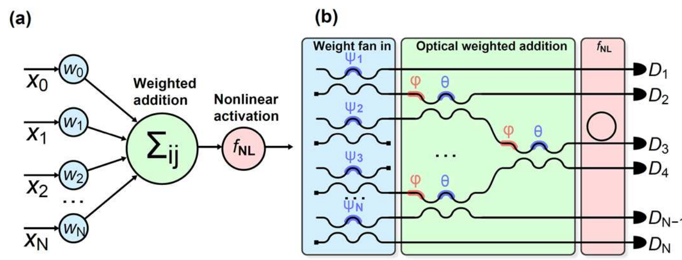

Fig. 1 (a) An artificial neuron with a fan in of N inputs with assigned weights. Each neuron sums the inputs which are then fed to a nonlinear activation function to determine the strength of the output signal, (b) the corresponding photonic version showing different associated operations.

Figures 1(a) and 1(b) illustrates the generalized artificial neuron for the hidden layer with N inputs and the corresponding photonic implementation of a neuron. In the photonic implementation, N attenuators are employed for weighted multiplication, represented by the blue-shaded region in Figure 1(b), and the green-shaded region highlighted the optical weighted addition implemented through a binary-tree of tunable MZIs [6]. After summation of the optical inputs, our neuron then performs a nonlinear activation function through optical bistability in a silicon MRR as shown highlighted in red in Figure 1(b). This comprehensive approach enables the realization of fast, all-optical neural network, eliminating the need for signal conversion.

III. PHOTONIC NEURAL NETWORK

We assessed the feasibility of our proposed photonic neuron implementation by training an ANN to predict the behavior a nonlinear curve derived from the equation $y(x) = 2x^2 + 1$. This choice was made twofold: (1) the simplicity of the defined problem; (2) there are easily discernable outcomes.

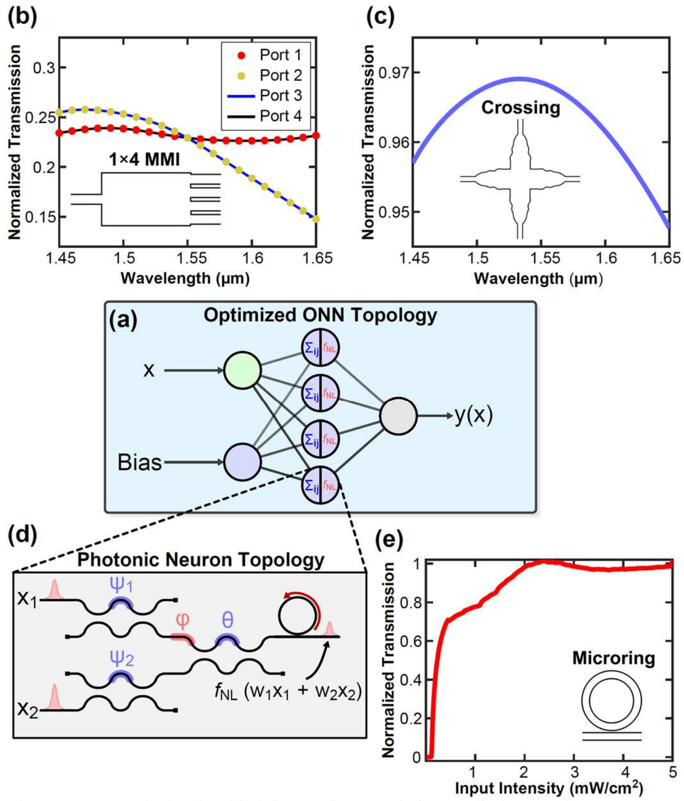

Fig. 2 (a) An optimized artificial neural network for predicting nonlinear curves, (b) 1×4 multi-mode interferometer simulated in FDTD achieving ~25% power splitting, (c) waveguide crossing broadband response while achieving loss of 0.14dB (d) displays the two-input photonic neuron topology, (e) shows the nonlinear activation function performed by the silicon MRR.

Through optimization, we fine-tuned the neural network to incorporate two input layer neurons, four hidden layer neurons, and one output layer neuron, enabling the accurate representation of the given nonlinear function within the range of $-1 < x < 1$, as shown in Figure 2(a). The input layer neuron plays a simple yet crucial role of splitting the input into four equal parts, a task seamlessly accomplished through a 1×4 multimode interferometer (MMI).

The transmission vs. wavelength curve is presented in Figure 2(b) and was obtained using an eigenmode expansion (EME) solver. To implement any desired ONN topology, we designed low-loss waveguide crossings to facilitate the connection between the input and hidden layers. To address this, we employed a particle swarm algorithm (PSO) in conjunction with the finite-difference time-domain (FDTD) method to design a low-loss waveguide crossing. Figure 2(c) displays the transmission vs. wavelength curve for the crossing.

The hidden layer consists of standard neurons with two inputs, incorporating functionalities such as weighted addition and nonlinear activation. Our proposed photonic neuron realizes this capability by utilizing an adaptive beam coupler with feedback control and photodetectors [6], as depicted in Figure 2(d). Figure 2(e) shows the input-output intensity

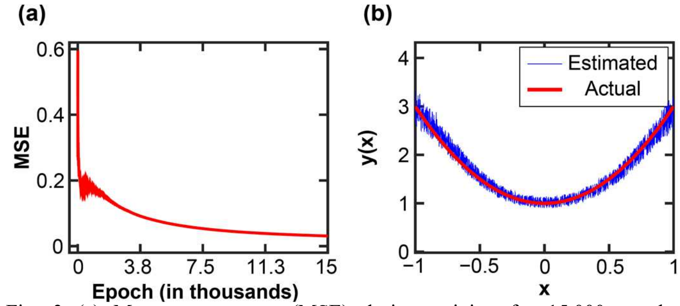

Fig. 3 (a) Mean-square error (MSE) during training for 15,000 epochs considering a uniform distribution of nonidealities attributed to the physical implementation of the ONN, (b) resulting estimated and actual curve for a given nonlinear function of $y(x) = 2x^2 + 1$.

relationship of our silicon MRR bistability nonlinear activation function. Finally, the output layer consists of a single neuron, which simply sums the four inputs using the adaptive beam. coupler. This configuration enables efficient and accurate computation, representing an addition of four inputs.

To assess the viability of implementing a photonic neural network, we trained a neural network to model the nonlinear curve $y(x) = 2x^2 + 1$. We introduced uniformly sampled random variations ranging from 0 to 20% in attenuation, addition, and nonlinear activation. These variations were incorporated to account for nonidealities that may arise from physical implementations. As depicted in Figure 3(a), utilizing 15,000 training samples, we successfully minimized the mean square error (MSE) to ~0.03. Notably, during testing, we considered the presence of nonidealities, and Figure 3(b) demonstrates that the trained neural network is capable of capturing the overall trend of the actual curve. This outcome shows the potential of our proposed method to effectively implement an optical neural network, provided that meticulous design and fabrication are achieved.

By cascading individual neurons according to the required number of inputs and layers, a complete neural network can be constructed in a seamless manner, eliminating the need for O/E/O conversion. This approach promises to enhance the overall efficiency, practicality, and complexity of implementing large-scale, photonic neural networks.

REFERENCES

[1] C. D. Schuman *et al.*, "A survey of neuromorphic computing and neural networks in hardware," arXiv preprint arXiv:1705.06963, 2017.

[2] A. Graves *et al.*, "Hybrid computing using a neural network with dynamic external memory," *Nature*, vol. 538, no. 7626, pp. 471-476, 2016.

[3] M. Davies *et al.*, "Loihi: A neuromorphic manycore processor with on-chip learning," *IEEE Micro*, vol. 38, no. 1, pp. 82-99, 2018.

[4] A. N. Tait *et al.*, "Silicon photonic modulator neuron," *Phys. Rev. Appl.*, vol. 11, no. 6, p. 064043, 2019.

[5] A. N. Tait *et al.*, "Broadcast and weight: an integrated network for scalable photonic spike processing," *J. Lightw. Technol.*, vol. 32, no. 21, pp. 3427-3439, 2014.

[6] D. A. Miller, "Self-aligning universal beam coupler," *Opt. Express*, vol. 21, no. 5, pp. 6360-6370, 2013.

[7] I. D. Rukhlenko *et al.*, "Analytical study of optical bistability in silicon ring resonators," *Opt. Lett.*, vol. 35, no. 1, pp. 55-57, 2010.