EUROPEAN CONFERENCE ON INTEGRATED OPTICS

Programmable photorefractive optical synapses in integrated photonics for neuromorphic computing

(Student paper)

Elger A. Vlieg[1*], Folkert Horst[1], Roger Dangel[1†] and Bert J. Offrein[1]

[1] IBM Research—Zurich, Säumerstrasse 4, Rüschlikon, 8803, Switzerland
*Corresponding author e-mail (vli@zurich.ibm.com)
†Previously employed at [1]

We experimentally demonstrate programmable optical synaptic connections for neuromorphic computing, based on the photorefractive effect, in an integrated photonic interconnect circuit. This provides the core components of an analog crossbar array for fast and efficient matrix-vector multiplications in artificial neural networks.
Keywords: *integrated optics, neuromorphic computing, synaptic connections, analog crossbar array, photorefractive effect, holography*

INTRODUCTION

For the past decade, the progress in artificial intelligence (AI) has been accompanied by an artificial neural network (ANN) size increase at an exponential rate close to 10x that of Moore's law [1]. This trend is unsustainable for general-purpose digital processors due to the increasing energy and hardware demands [1,2].

Analog neuromorphic computing is an emerging paradigm that aims to meet the processing requirements in AI [3,4]. A core building block of the neuromorphic computer is the synapse, which connects signal paths with a programmable coupling strength, and can be optimized by training for a desired AI task.

Analog synapses configured in a crossbar array network can accelerate matrix-vector multiplications (MVM) for the signal processing in ANNs (Fig. 1) [3-5]. Importantly, the signal propagation between densely connected layers yields a multiply-accumulate (MAC) workload that scales quadratically with network width, while the analog crossbar array can perform it with linear scaling on latency and energy [5]. Thus, the analog crossbar array is a promising architecture to enable ANNs with improved power-efficiency.

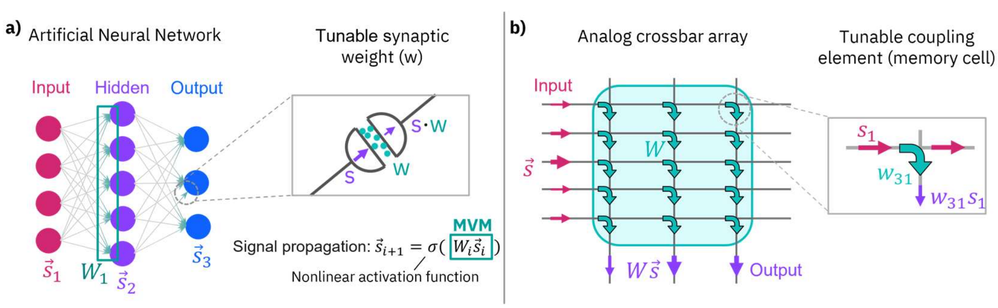

*Fig. 1. **a)** The signal transfer between fully-connected layers in artificial neural networks is given by the matrix-vector multiplication (MVM) of the synaptic weight matrix and the input neural signal vector. **b)** The analog crossbar array processor performs this synaptic signal transfer by physically implementing the synaptic connections.*

The optimal hardware implementation of the analog crossbar array is a topic of scientific investigation [3-5]. In this paper, we discuss the development of an integrated photonic implementation based on the photorefractive effect (Fig. 2) [5-7]. This architecture is particularly powerful since it supports all matrix operations for ANN backpropagation training by single-cycle execution steps [5]. The transposed matrix for back propagation is addressed by simply swapping the roles of the source and destination channels. The outer product for the synaptic weight update is enabled by a linear grating write response in both negative and positive directions (Fig 2. **b)**). Lastly, the photorefractive grating formation effectively occurs continuously, thereby offering a rich number of analog weight states required for training. On top of supporting efficient ANN training, this architecture helps to alleviate the scaling challenges in photonics by spatially superimposing the synaptic weights.

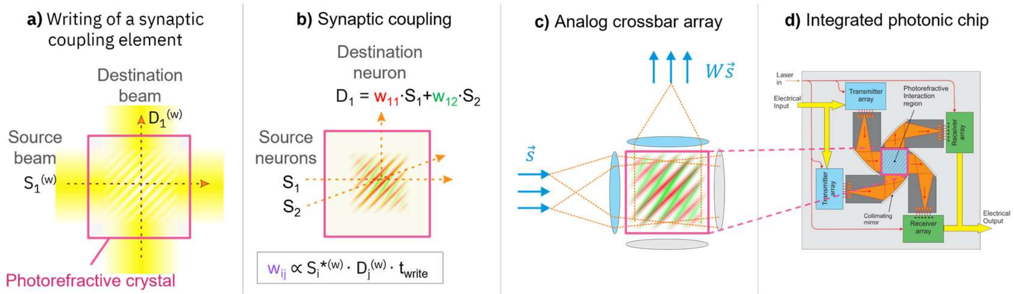

*Fig. 2. The proposed photorefractive crossbar array architecture in integrated photonics [5-7]. **a)** A source and destination laser beam yield a sinusoidal interference pattern, which is converted into a non-volatile refractive index grating by the photorefractive effect. **b)** The refractive index grating acts as a Bragg-mirror and thereby establishes a synaptic connection between the source and destination channel. Since Bragg-mirrors are angle-selective, synapses can be superimposed for angularly separated channels. **c)** Collimating optics are deployed to convert spatially separated I/O waveguides to the angular domain and back. **d)** All functional components are included on a single photonic integrated circuit.*

RESULTS

The realization of photorefractive photonic integrated circuits requires photorefractive thin-film substrates. These were achieved by thinning down photorefractive semi-insulating GaAs wafers, after wafer bonding them to Si substrates with a SiO_2 layer (Fig. 3).

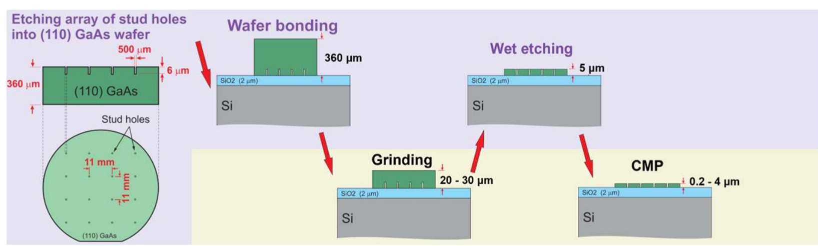

Fig. 3. Photorefractive GaAs wafers are thinned down to thicknesses suitable for photonic waveguides.

As a proof-of-concept, we first focused on demonstrating individual synapses in a prototype integrated crossbar circuit by photorefractive two-wave mixing (Fig. 4) [7]. A source and destination beam intersect inside the processor and form a photorefractive index grating. An electro-optic phase modulator (EOPM) is included on the source beam path to experimentally trigger grating write sequences by periodically shifting the phase of the optical interference pattern [5]. The EOPM is located off chip, and the source and destination signals are coupled into the photonic chip by a lensed fiber array.

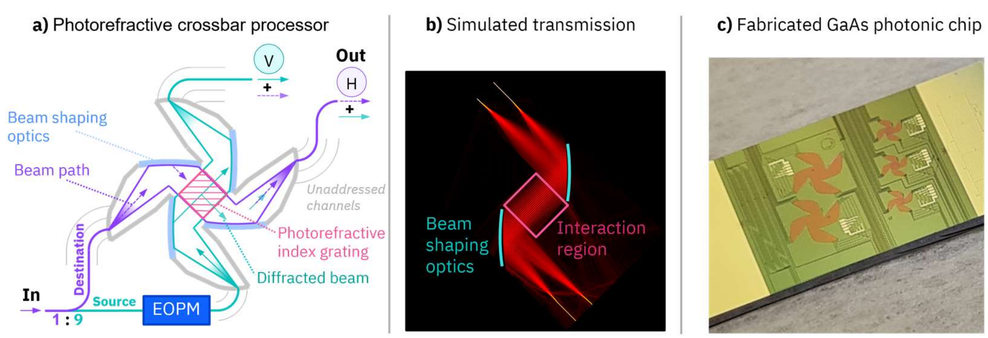

*Fig. 4. The experimental design for demonstrating photorefractive synapses by two-wave mixing [7]. **a)** A schematic of the prototype photonic crossbar circuit. **b)** The simulated transmission through the processor. **c)** The circuit realized in a photorefractive thin-film GaAs layer (Fig. 3).*

The results of the experiment in Fig. 4 are shown in Fig. 5. When the EOPM (gray) π phase-shifts the source beam, an intensity jump is visible due to the mirrored interference behavior between the output beams, followed by the rewriting of the synapse via an exponential saturation process. The analog formation of the coupling strength illustrates that the synaptic weight can in principle be programmed by write time (Fig. 2 **b)**).

Three different experiments were performed to confirm that the exponential trend corresponds to photorefractive two-wave mixing. Fig. 5 **a)** shows that based on the wavelength, the optical intensity either predominantly removes or adds electrons in the GaAs deep charge traps underlying the photorefractive grating, yielding a π phase shift thereof [8]. Fig. 5 **b)** confirms that intensity is transferred between the horizontal and vertical processor branches (Fig 4. **a)**). Fig 5. **c)** demonstrates that the grating write speed is affected by the input optical power, as indicated by the initial slope steepness, corresponding to an accelerated charge excitation rate [7]. In Fig 5. **c)**, the 1.5mW trace has insufficient time to saturate the grating, yielding a near-linear write behavior at the expense of dynamic range.

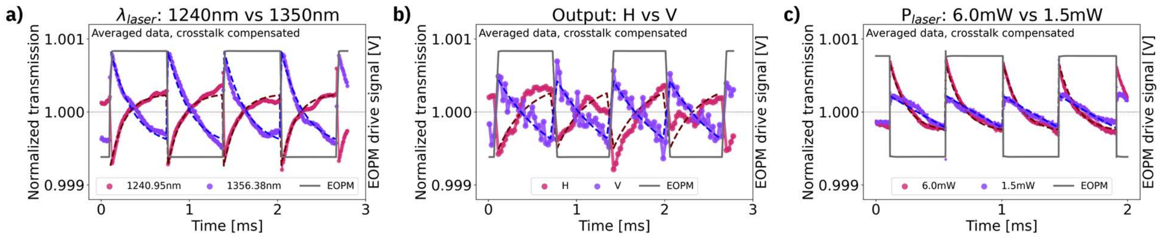

*Fig. 5. The experimental periodic writing of photorefractive synapses with threefold verification. **a)** A sign flip based on optical wavelength. **b)** A sign flip upon swapping the source and destination channels (Fig. 4 **a)**). **c)** A response time modification based on optical power.*

As an extension to Fig. 5 **c)**, the fitted photorefractive response times for different laser powers are shown in Fig. 6. The datapoint at 6.0 mW is not captured by the fit. We expect that this is due to an overestimated effective optical power for all data points due to drifting lensed fibers yielding a reduced coupling efficiency. Importantly, the linear relationship in Fig 2. **b)** appears to hold.

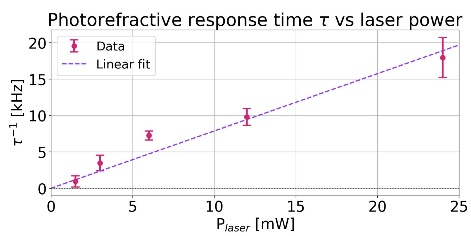

Fig. 6. The photorefractive response time τ for 5 different laser powers. The vertical error bars only indicate fit robustness, and exclude experimental reproducibility.

DISCUSSION

We report the realization of photorefractive optical synapses in integrated photonics, and thereby verified the basic operating principles of our crossbar array design. To extend this work to a fully functional crossbar array, programmable I/O circuitry needs to be included, so that multiple crossbar channels can be addressed in parallel and individually (Fig. 4 **a)**) [5]. Future work consists of fabricating a full crossbar array and demonstrating all relevant signal processing.

Acknowledgements: This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement number 828841 (ChipAI) and 860360 (Post-Digital). We thank the Binnig and Rohrer Nanotechnology Center (BRNC).

References

[1] D. Narayanan et al., *Efficient Large-Scale Language Model Training on GPU Clusters Using Megatron-LM*, SC '21 58, pp. 1-15, 2021
[2] E. Strubell et al., *Energy and Policy Considerations for Modern Deep Learning Research*, AAAI 34(09), 13693-13696, 2020
[3] D. Marković et al., *Physics for neuromorphic computing*, Nat Rev Phys 2, pp. 499–510, 2020
[4] A. Sebastian et al., *Memory devices and applications for in-memory computing*, Nat. Nanotechnol. 15, pp. 529–544, 2020
[5] E. A. Vlieg et al., *An Integrated Photorefractive Analog Matrix-Vector Multiplier for Machine Learning*, Appl. Sci. 12(09), pp. 4226, 2022
[6] Y. Owechko, B.H. Soffer, *Holographic Neurocomputer Utilizing Laser Diode Light Source*, Proc. SPIE 2565, pp. 12–19, 1995
[7] P. Yeh, *Two-Wave Mixing in Nonlinear Media*, IEEE J. Quantum Electron. 25(3), pp. 484–519, 1989
[8] F. P. Strohkendl et al., *Hole–electron competition in photorefractive gratings*, Opt. Lett. 11, pp. 312-314, 1986