Ultra-Compact Leaky ReLU Nonlinear Function on IMOS

Ultra-Compact Leaky ReLU Nonlinear Function on IMOS

Antonio Lechiara[(1)], Andrea Marchisio[(2)], Bin Shi[(1)], Yi Wang[(1)],
Andrea Carena[(2)], Yuqing Jiao[(1)], Ripalta Stabile[(1)]

[(1)] Eindhoven Hendrik Casimir Institute, Eindhoven University of Technology, Eindhoven 5612AP, The Netherlands a.lechiara@tue.nl
[(2)] Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy, andrea_marchisio@polito.it

Abstract *We demonstrate the first compact MRR-nested-MZI nonlinear function on the IMOS platform, replicating a Leaky ReLU function. Exploiting ultra-compact phase shifters, the device achieves sub-5dBm triggering power and a footprint of 0.22 mm^2, without requiring transfer function optimization.*

Introduction

Neuromorphic photonics is an emerging field[1] that leverages the high bandwidth and energy efficiency of Photonic Integrated Circuits (PICs) to implement neuromorphic computing architectures. Among the most promising approaches for photonic inter-layer connectivity is the coherent architecture, which employs interferometric structures operating with a single-wavelength input signal[2]. A key component enabling computational capability is the nonlinear activation function, which allows neuromorphic networks to address complex, nonlinearly separable problems. When co-integrated with linear optical processors, all-optical nonlinearities can surpass electronic alternatives in terms of speed, making them essential for high-speed signal processing applications.

Various photonic activation functions have been proposed for optical neural networks, including both electro-optical[3],[4] and all-optical nonlinear implementations[5]–[7]. Electro-optical designs are often limited by the bandwidth constraints of electrical interconnects and by the challenges of impedance matching between photodetectors and modulators, which also increase footprint due to additional routing requirements. In contrast, all-optical nonlinear functions are more suitable for monolithic integration with matrix multiplication units on the same photonic chip, as they eliminate the need for electro-optic conversion.

Architectures for nonlinear functions based on Micro-Ring Resonators (MRRs) coupled with Mach–Zehnder Interferometers (MZIs) are highly compatible with MZI mesh-based coherent processing units. An optical thresholder exploiting Fano resonances has recently been demonstrated on a Silicon-on-Insulator (SOI) platform[6], combining MRR resonances with MZI interference, achieving a triggering power of $5\,\mathrm{dBm}$[5]. However, many of the activation functions demonstrated to date approximate Heaviside step responses[3],[8] or exhibit sinusoidal behavior[9], and are often limited to bandwidths around $1\,\mathrm{GHz}$, while deeper networks require ReLU, PReLU, or sigmoid-type functions (e.g., hyperbolic tangent, logistic sigmoid).

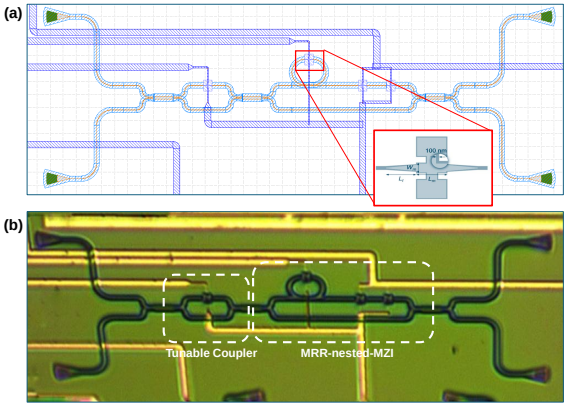

Fig. 1: (a) Mask of the MRR-nested-MZI structure. In the inset, a detailed representation of the ultra-compact PS. (b) Micrograph of the device on the IMOS chip.

The use of the Indium Phosphide (InP) material platform has enabled the demonstration of true sigmoid activation functions with bandwidths up to $20\,\mathrm{GHz}$[7], owing to the lower free-carrier lifetime in InP waveguides. Moreover, the two-photon absorption (TPA) efficiency over 20 times higher that of Silicon lowers the triggering power of the nonlinearity. However, InP-based SOA and MZI devices tend to occupy a relatively large footprint.

The Indium Phosphide Membrane on Silicon (IMOS) platform offers a compelling alternative thanks to its high refractive index contrast and the ability to monolithically integrate active and passive components. The IMOS platform exhibits values of the Kerr coefficient about 5–10 times higher than that of Si, enabling stronger nonlinear processes at lower optical power[10].

In this paper, we present for the first time a compact MRR–MZI-based nonlinear activation function implemented on the IMOS platform, capable of performing a Leaky ReLU-like response. The use of the IMOS platform, along with ultra-compact Phase Shifters (PSs), enables a significantly reduced footprint of $0.22\,\mathrm{mm}^2$. The trig-

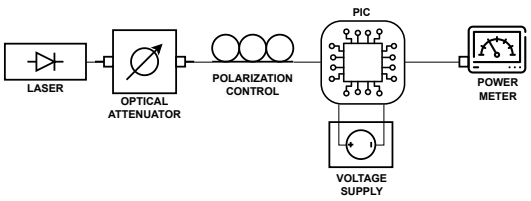

Fig. 2: Block schematic of the experimental setup.

gering power is conservatively estimated to be $\sim 3.5\,\mathrm{dBm}$, which is $30\,\%$ lower than on Si.

IMOS nonlinear function device

The IMOS integration platform enables passive-active co-integration of ultra-compact photonic components via enhanced optical confinement, obtained by inserting a low refractive index buffer layer between the InP membrane and a Si substrate. When realizing a nonlinear function compatible with the well-known MZI-mesh-based matrix multiplication unit, MZI devices turn out to be a great choice, also in view of optimal circuitry yield. However, the size of thermo-optic shifters on SOI is reported to be on the order of $\sim 250 \times 100\,\mu\mathrm{m}^2$[2], mainly attributable to the $150\,\mu\mathrm{m}$ length of the PSs. On IMOS, conventional PSs as long as $200\,\mu\mathrm{m}$ are reported. However, recently, a compact thermo-optical PS has been demonstrated based on the use of an additional doped layer on the InP membrane waveguide and designed as a crossing, for a total size of $15 \times 17\,\mu\mathrm{m}^2$, for up to 1 order of magnitude reduction in size[11].

In Fig. 1(a) and 1(b) the mask details and a micrograph of the nonlinear function device based on a MRR-nested-MZI structure are shown. This includes two main parts: a tunable coupler and an MRR-nested-MZI. The transfer function of the interferometer is reconfigured by tuning the coupling ratio of the tunable coupler and the detuning of the MRR[12]. The compact PS are used to tune the phase of the MZI and the resonance of the MRR. With a slightly detuned signal to the initial dip of the MRR and an increase in input power, which changes the phase of the arm with MRR, the transmission is changed so that a ReLU nonlinear function is achievable.

Measurements Setup

The experimental setup employed for device characterization enables both optical and electrical access to the PIC, as schematically illustrated in Fig. 2. Light from a Keysight N7711A tunable laser is vertically coupled into the PIC via surface grating couplers. Due to the angle and polarization sensitivity of the grating couplers, optimal coupling is achieved at an incident angle of 12° and with the use of a polarization controller. Electrical biasing of thermo-optic or active components on the PIC is provided via voltage probes, which are connected to a Keysight EDU36311A power supply and contact the on-chip metal lines. The output optical power, P_{out}, is measured using a Santec MPM-210H power meter for subsequent data analysis. During measurements involving the Nonlinear Activation Function (NLAF), a HP 8156A optical attenuator is used to precisely control the input power delivered from the tunable laser to the PIC.

Results

In this section, we will present the main results obtained from the measurements of the test structures realized with the InP-based ultra-compact PS[11]. Following the characterization of individual MZI and MRR components, we proceeded to measure the NLAF test structure, which integrates these subcomponents.

First, we characterized the MZI, implemented using two 2×2 Multi-Mode Interferometers (MMIs) at the input and output, connected by two waveguides. Thermal tuning is enabled by two ultra-compact PSs: one in the interferometric section ($\theta \in [0, \pi]$) and a double PS on the upper input arm ($\phi \in [0, 2\pi]$). Given their critical role in training photonic neurons, the MZI was thermo-optically characterized to assess its tunability. Figure 3(a) shows the optical output power for all four input/output port combinations as a function of the linearly increasing voltage from $0\,\mathrm{V}$ to $2\,\mathrm{V}$. These results

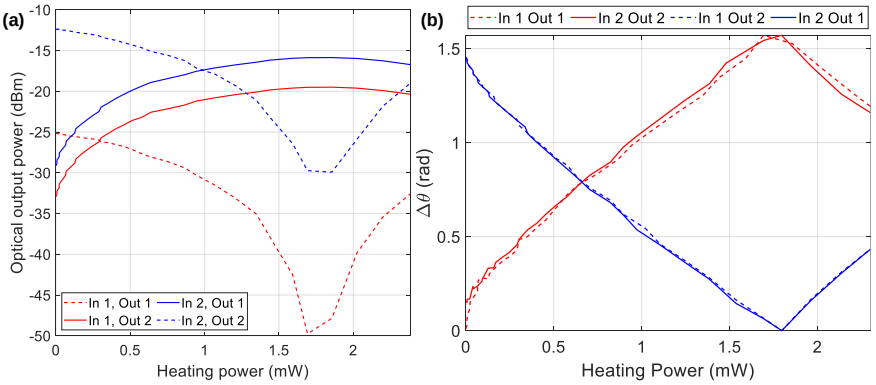

Fig. 3: MZI measurements (a) Optical power measurements of the 4 input/output combination, while applying a linearly increasing voltage between $0\,\mathrm{V}$ and $2\,\mathrm{V}$ to the θ PS; (b) Measured phase difference between the MZI arms with respect to the heating power.

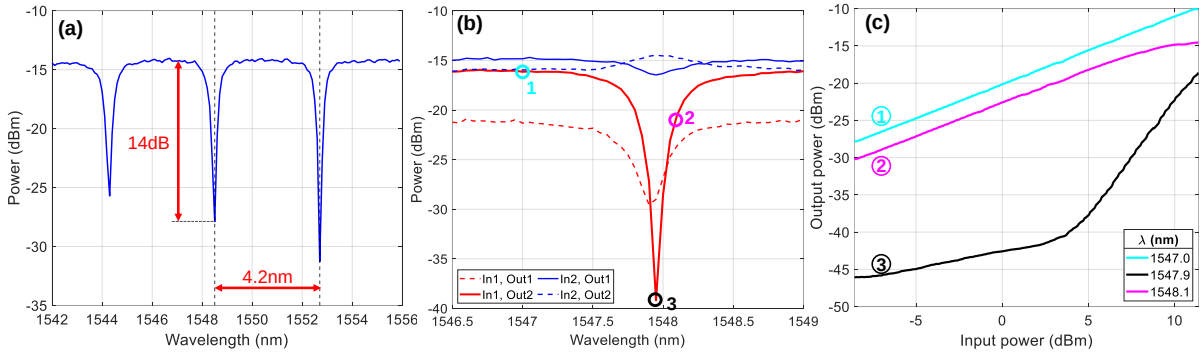

Fig. 4: (a) Transmission spectrum of a test MRR with $R = 24\,\mu\mathrm{m}$; (b) Transmission spectrum of the test NLAF, around the resonance wavelength $\lambda_{\mathrm{res}} = 1547.95\,\mathrm{nm}$; (c) Curves generated by the NLAF unit at the resonance wavelength $\lambda_{\mathrm{res}} = 1547.95\,\mathrm{nm}$ and additional wavelengths in its vicinity. A Leaky ReLU-like curve at λ_{res} is obtained.

highlight unbalanced losses for the different combinations, but it is possible to trace the trends back to the cos^2 behavior expected from theory, when converting the curves from dBm to mW. Based on this behavior, we extracted the phase difference $\Delta\theta$ between the MZI arms, plotted against heating power in Fig. 3(b), from which the π-phase shift power was determined as $P_\pi = 1.8\,\mathrm{mW}$.

In the subsequent step, we measured the transmission spectrum of an isolated MRR with radius $R = 24\,\mu\mathrm{m}$, shown in Fig. 4(a). The Free Spectral Range (FSR) is $4.2\,\mathrm{nm}$, with an Extinction Ratio (ER) of $\sim 14\,\mathrm{dB}$ at $1548.5\,\mathrm{nm}$.

Finally, we present here the results of the experimental characterization of a test structure specifically designed to implement an all-optical NLAF. The device consists of a MRR-assisted MZI, preceded by a tunable coupler, enabling nonlinear optical signal processing[13].

The MRR is equipped with the same compact micro-heater used in the MZIs, allowing precise tuning of its transmission spectrum. The adoption of compact PS, rather than conventional metallic ones, significantly reduces the overall footprint of the device. Due to its programmability, this architecture supports a wide range of photonic applications, including optical thresholding and signal processing[6].

The output power P_{out} as a function of the input wavelength λ_{in}, at a fixed input power of $P_{\mathrm{in}} = 6\,\mathrm{dBm}$, has been measured and reported in Fig. 4(b). Particular attention was paid to the resonance region near $\lambda_{\mathrm{res}} = 1547.95\,\mathrm{nm}$, with transmission spectra acquired for all four input-output port combinations, as shown in Fig. 4(b).

The device is designed to enable a variety of nonlinear activation functions, tunable via the applied voltages to the compact PSs and the MRR. Away from resonance the transmission remains approximately linear while, around resonance, nonlinear effects are observed. To evaluate the passive nonlinear behavior of the NLAF, we measured P_{out} as a function of P_{in} across several wavelengths in the vicinity of λ_{res}, without applying any external voltages to the MRR or the MZI. The optical attenuator was employed to finely sweep the input power levels. The results are depicted in Fig. 4(c).

At $\lambda_{\mathrm{res}} = 1547.95\,\mathrm{nm}$, the output response closely resembles a Leaky ReLU activation function (black curve in Fig. 4(c).). In the vicinity of the resonance, residual nonlinearities persist, as it happens at $\lambda = 1548.10\,\mathrm{nm}$ (magenta curve). Finally, at wavelengths sufficiently far from resonance (e.g., $\lambda = 1547.00\,\mathrm{nm}$, cyan curve), the behavior returns linear. These results demonstrate the NLAF's ability to perform passive nonlinear activation, without the need for electrical tuning. Further investigations into the dynamic tuning of the activation function via biasing of the PSs and MRR are planned for future work.

Conclusions and Outlook

In this work we presented and characterized the nonlinear activation function realized on IMOS and based on an MRR-assisted MZI, both spectrally and as a function of input power. Notably, a Leaky ReLU-like response was observed without applying any electrical bias, demonstrating the feasibility of passive nonlinear activation.

The experimental characterization of its main building blocks, i.e. the MRR and MZI, offered insight into their performances and losses contributions. Our preliminary measurements on straight waveguides, showing that the grating coupler as the main source of optical power losses (at about $-15\,\mathrm{dB}$), provide encouraging result for the scalability of the system, given the low propagation losses on-chip.

Future work will focus on three main directions. We plan to explore the tunability of the NLAF by biasing the PSs and MRR, enabling greater control over the activation function and expanding the device's functionality.

Secondly, the MZIs and NLAFs will be exploited as building blocks of an all-optical, ultra-compact neuron for neuromorphic computing.

Acknowledgements

This publication is funded in part by the project NL-ECO: Netherlands Initiative for Energy-Efficient Computing (with project number NWA. 1389.20.140) of the NWA research programme Research Along Routes by Consortia which is financed by the Dutch Research Council (NWO).

A.L. Ph.D. scholarship is funded by the NL-ECO program, through the NWA-ORC programme of the Dutch Research Council (NWO; project number NWA.1389.20.140).

A.M. Ph.D. scholarship is funded by the European Union Next-GenerationEU and by the Italian National Recovery and Resilience Plan (PNRR) through the Italian Ministry of University and Research (MUR) under grant D.M.352/2022.

References

[1] P. R. Prucnal and B. J. Shastri, *Neuromorphic photonics*. CRC press, 2017.

[2] Y. Shen, N. C. Harris, S. Skirlo, *et al.*, "Deep learning with coherent nanophotonic circuits", *Nature Photonics*, vol. 11, pp. 441–446, 7 Jun. 2017. DOI: 10.1038/nphoton.2017.93.

[3] A. N. Tait, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Broadcast and weight: An integrated network for scalable photonic spike processing", *Journal of Lightwave Technology*, vol. 32, no. 21, pp. 4029–4041, 2014. DOI: 10.1109/JLT.2014.2345652.

[4] I. Williamson, T. W. Hughes, M. Minkov, B. Bartlett, S. Pai, and S. Fan, "Reprogrammable electro-optic nonlinear activation functions for optical neural networks", *IEEE Journal of Selected Topics in Quantum Electronics*, pp. 1–14, 2019. DOI: 10.1109/JSTQE.2019.2930455.

[5] B. Shi, "Semiconductor optical amplifier-based photonic integrated deep neural networks", Ph.D. dissertation, Eindhoven University of Technology, 2022.

[6] C. Huang, T. Ferreira De Lima, A. Jha, *et al.*, "Programmable silicon photonic optical thresholder", *IEEE Photonics Technology Letters*, vol. 31, no. 22, pp. 1834–1837, 2019. DOI: 10.1109/LPT.2019.2948903.

[7] G. Mourgias-Alexandris, G. Dabos, N. Passalis, A. Totović, A. Tefas, and N. Pleros, "All-optical WDM recurrent neural networks with gating", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 5, pp. 1–7, 2020. DOI: 10.1109/JSTQE.2020.2995830.

[8] M. A. Nahmias, A. N. Tait, L. Tolias, *et al.*, "An integrated analog O/E/O link for multi-channel laser neurons", *Applied Physics Letters*, vol. 108, no. 15, p. 151 106, Apr. 2016. DOI: 10.1063/1.4945368.

[9] P. Ma, Y. Salamin, B. Baeuerle, *et al.*, "Plasmonically enhanced graphene photodetector featuring 100 Gbit/s Data Reception, high responsivity, and compact size", *ACS Photonics*, vol. 6, no. 1, pp. 154–161, 2019. DOI: 10.1021/acsphotonics.8b01234.

[10] Y. Jiao, J. van der Tol, V. Pogoretskii, *et al.*, "Indium phosphide membrane nanophotonic integrated circuits on silicon", *Physica Status Solidi (A) Applications and Materials Science*, vol. 217, pp. 1–12, 3 2020, ISSN: 18626319. DOI: 10.1002/pssa.201900606.

[11] Y. Wang, V. Dolores-Calzadilla, K. A. Williams, M. K. Smit, and Y. Jiao, "Ultra-compact and efficient microheaters on a submicron-thick InP membrane", *Journal of Lightwave Technology*, vol. 41, no. 6, pp. 1790–1800, 2023. DOI: 10.1109/JLT.2022.3225110.

[12] A. Jha, C. Huang, and P. R. Prucnal, "Reconfigurable all-optical nonlinear activation functions for neuromorphic photonics", *Optics Letters, Vol. 45, Issue 17, pp. 4819-4822*, vol. 45, pp. 4819–4822, 17 Sep. 2020, ISSN: 1539-4794. DOI: 10.1364/OL.398234.

[13] C. Huang, A. Jha, T. F. De Lima, A. N. Tait, B. J. Shastri, and P. R. Prucnal, "On-chip programmable nonlinear optical signal processor and its applications", *IEEE Journal of selected topics in quantum electronics*, vol. 27, no. 2, pp. 1–11, 2020. DOI: 10.1109/JSTQE.2020.2998073.