Adaptive all-optical sigmoid activation functions for Photonic Neural Networks using Fabry-Perot laser diodes under optical injection

Petar Atanasijević[1], Christos Pappas[2], Mladen Banović[1], Jasna Crnjanski[1], Apostolos Tsakyridis[2], Miltiadis Moralis-Pegios[2], Konstantinos Vyrsokinos[2], Marko Krstić[1], Peđa Mihailović[1], Slobodan Petričević[1], Nikos Pleros[2], Dejan Gvozdić[1]

[1]School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia
[2]Centre for Interdisciplinary Research and Innovation, Informatics Dept. Aristotle University of Thessaloniki, Greece
Author e-mail address: petarat@etf.bg.ac.rs

Abstract: We experimentally validate the all-optical activation functions in Fabry-Perot lasers under optical injection for random and non-random inputs. Sigmoid-like activations for 100 ps pulses are reconfigured using injection parameters, consuming 1.4 pJ per nonlinear operation. © 2024 The Authors

1. Introduction

Interest in neuromorphic computing grows ever stronger, as the bottlenecks of traditional computing hardware become eminent. In the neuromorphic race, photonics brings unique promises of higher computing speed and energy efficiency over its electronic counterpart [1,2]. However, the complete realization of the photonic neuromorphic hardware requires robust physical implementations of both the linear and the nonlinear parts of the artificial neurons. The latter proves to be especially difficult to achieve, usually requiring optoelectronic, or opto-electro-optic conversions that utilize transimpedance amplifiers (TIA) and electro-optic modulators [3]. Up to this point, different reconfigurable optoelectronic nonlinear units were demonstrated, usually exhibiting sigmoid-like and ReLU-like profiles [3,4]. More recently, a Tanh-like activation function was experimentally demonstrated, utilizing a programmable TIA + balanced photodetector on a chip, covering a range of different activations [5]. On the other hand, the speed and/or energy limitations associated with the optoelectronic nonlinearities have inspired research towards all-optical activation functions [6]. However, most of the all-optical nonlinear activation functions reported so far have hardly managed to perform at clock rates higher than 1 GHz [7,8], with all-optical activation functions reaching a 10 GHz operational rate having been reported so far only through the use of power-hungry semiconductor optical amplifier (SOA)-based circuitry [9,10].

In this paper, we build upon the theoretical foundation for a reconfigurable all-optical perceptron [11], and present for the first time, to the best of our knowledge, an experimental characterization of an all-optical nonlinear activation unit based on a single Fabry-Perot laser diode (FP-LD) under optical injection. The unit's all-optical activation functions are shown to be adaptive while exhibiting sigmoid-like trends. The nonlinearity was evaluated for pulses with 100 ps widths using linearly increasing and randomly distributed input pulse powers. The energy consumption of only 1.4 pJ per nonlinear operation is estimated for 100 ps pulses.

2. Experiment design, bistability, and pulsed injection

The experimental setup for optical injection into a FP-LD is presented in Fig. 1(a). The continuous wave (CW) tunable master laser (TL) is modulated using a Mach-Zehnder modulator (MZM), preceding a variable optical attenuator (VOA) used for signal power adjustment. A 50:50 fiber coupler is added to simultaneously monitor the modulated TL signal using a photodiode (PD, 20 GHz bandwidth) and inject it in the slave FP-LD. A polarization controller (PC) is placed before the optical circulator (OC), used for injection and readout signal separation, to adjust the polarization of TL's light to match the FP-LD's polarization state. A 1% tap is placed at the output of the OC, allowing for FP-LD's output spectrum monitoring (using an optical spectrum analyzer, OSA) prior to filtering of only the TL's wavelength using a 0.35 nm bandpass tunable optical filter (TF). The filtered signal is amplified using a gain block consisting of an erbium-doped fiber amplifier (EDFA), a 1 nm tunable optical filter, and an attenuator for precise power control. The amplified output is monitored using a PD (70 GHz bandwidth). Both the input and output monitoring PDs are connected to the oscilloscope (OSC), used for data acquisition. During the experiments, the FP-LD's current and temperature were regulated to 14 mA and 20°C, respectively, keeping the diode near the 10 mA threshold.

The output spectrum of a free-running FP-LD with the CW TL signal injected near the mode m = +2 (lower wavelength modes being positive) is shown in Fig. 1(b). The wavelength detuning Δλ between the injected mode and the TL's signal is recognized as one of the key parameters affecting the injection locking stability and static bistability [12]. The intermodal spacing of the FP-LD is 1.62 nm, setting a theoretical limit to $\Delta\lambda$ values. The experimentally recorded dependencies of output vs. injected CW FP-LD optical power at the wavelength of the TL, for three different detunings are provided in Fig. 1(c). The hysteresis loops, showing bistability, are shifted towards higher powers as the wavelength of the TL moves further away from the injected mode, i.e., the detuning is increased. The arrows in the branches of the hysteresis loop indicate the direction of the loop. However, going from the CW regime towards faster changes in the injected optical power, i.e., optical pulses with widths under 1 ns, results in a loss of a two-branched static hysteresis loop between the stable states. Instead, a single nonlinear function is predicted to relate the peak optical powers of the pulses injected and emitted from the FP-LD [11]. To experimentally evaluate the all-optical nonlinearity, the MZM was biased near a null point, and modulated using a linearly increasing Gaussian pulse train with 100 ps bit slots (10 Gb/s), separated by 3 ns pauses (30 zeros). The experimental results showing the input and output pulse train powers in time are shown in Fig. 1(d).

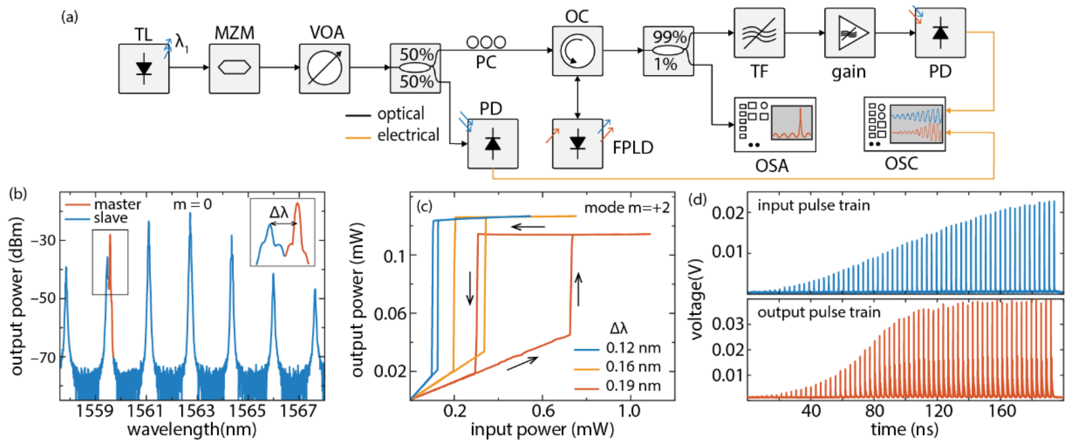

Fig. 1: a) A schematic of the experimental setup, used for optical injection into a FP-LD, b) spectrum of a free-running FP-LD with the injected master wavelength, detuned by $\Delta\lambda$ with respect to the injected mode $m = +2$, c) static hysteresis loops for different detunings, and d) a linearly increasing 100 ps pulse train, injected in the FP-LD, with the corresponding output response.

2. The adaptive all-optical sigmoid-like nonlinearity

The dependence of the output FP-LD peak pulse powers (after the TF) on the input peak pulse powers, injected near the $m = +2$ mode of the FP-LD, is evaluated in the case of four different detunings, 0.12, 0.14, 0.16, and 0.19 nm. Measured all-optical activation functions, together with the fitted sigmoid-like curves are presented in Fig. 2(a). A high degree of adaptivity of both the steepness and threshold of the activation function is demonstrated by frequency variation of under 10 GHz, for 10 Gb/s pulses and 0.32 Gb/s repetition rates. The results agree with the theoretically predicted trends, shown in Fig. 2(b). The repetition rate was gradually increased up to ten times (3.33 Gb/s - 2 zeros after a 100 ps pulse). The observed changes in the activation function's saturation levels for $\Delta\lambda = 0.12$ nm and the injected FP-LD mode $m = +6$ are presented in Fig. 2(c). Lowering of the saturation with the increase in repetition rate demonstrates the inter influence of time-adjacent input pulses. However, the nonlinear behavior of the activation function is preserved.

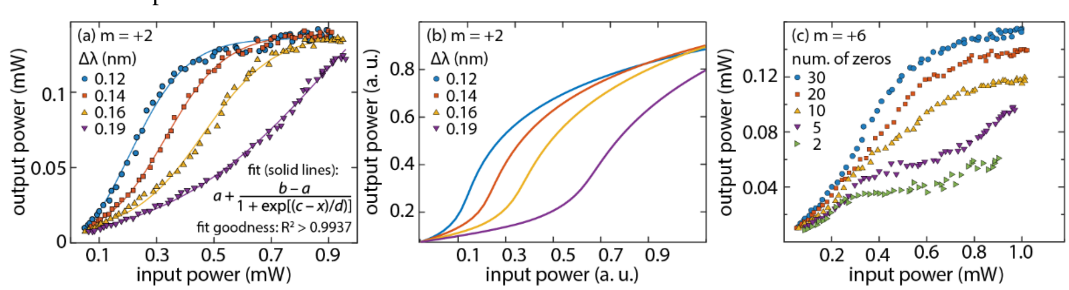

Fig. 2: a) Experimentally measured, and b) theoretically predicted sigmoid-like activation functions, showing adaptivity due to a variation of $\Delta\lambda$. c) Changes in the nonlinearity due to a different number of zeros after a 100 ps pulse.

This adaptive sigmoid-like activation function was further tested using a train of 100 ps-long input pulses with randomly distributed peak power levels at a 0.91 Gb/s repetition rate, resembling in this way the randomness of the pattern that is typically provided as the summed weighted input in a real-world perceptron operation. An example of a random input pulse train, followed by a matching output pulse train is presented in Fig. 3(a). A distinctive pulse pattern was applied at the beginning of every randomized sequence, enabling the correct matching of the input-output pulses. A set of ten different random pulse train measurements per detuning, per injected FP-LD mode, was evaluated for detunings of 0.12 and 0.16 nm, and mode numbers $m = +2$ and $m = +4$. The obtained adaptive all-optical activation functions are shown in Figs. 3(b) and (c), respectively.

Experimental results are plotted with points, while lines show an average trend, obtained using a 50-point moving average digital filter on the data. The results exhibit dispersion around the sigmoid-like curve, shown to decrease for a higher mode number. This "noisy" activation response can be compensated through noise-aware Neural Network training approaches and Optics-informed NN training models that have been shown to effectively support high-quality NN performance over "noisy" photonic hardware [13-15]. Furthermore, it is shown that the mode number influences not only the level of noise but also the saturation and threshold values of the achievable all-optical activation functions, providing another degree of freedom in the adaptivity of the demonstrated solution.

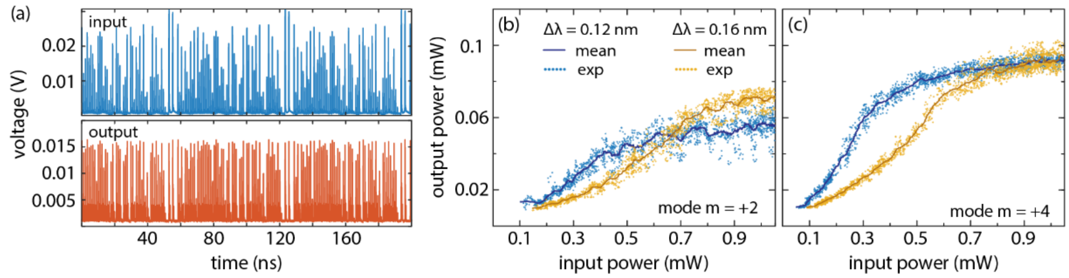

Fig. 3 a) A randomized input pulse train, injected in the FP-LD, with the corresponding output; Plots b) and c) show the activation functions with the measurement point dispersion and an average trend, for injection in the $m = +2$ and $m = +4$ mode, respectively.

At 14 mA and 0.93 V, the FP-LD based adaptive all-optical nonlinear element consumes under 14 mW, including the injected optical pulse power, yielding an energy efficiency of only 1.4 pJ per nonlinear operation that comprises an almost 10x improvement compared to other 10 GHz photonic activation function demonstrations.

3. Acknowledgements

This research was supported by Science Fund of the Republic of Serbia, Grant no #7750121, All-optical Reservoir Computer Architecture based on Laser Bistability – ORCA-LAB, and by Serbian Ministry of Education, Science and Technological Development. AUTH also acknowledges the support by the EC via H2020 Projects PLASMONIAC (871391), SIPHO-G (101017194) and GATEPOST (101120938).

4. References

[1] B. J. Shastri *et al.*, "Photonics for artificial intelligence and neuromorphic computing," *Nature Photonics*, vol. 15, no. 2. 2021.
[2] M. Moralis-Pegios *et al.*, "Neuromorphic Silicon Photonics and Hardware-Aware Deep Learning for High-Speed Inference," *J. of Light. Technol.*, vol. 40, no. 10, 2022.
[3] J. K. George *et al.*, "Neuromorphic photonics with electro-absorption modulators," *Opt Express*, vol. 27, no. 4, 2019.
[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 J of Sel.Top. Quantum Electron.*, 2019.
[5] C. Pappas *et al.*, "Programmable Tanh-, ELU-, Sigmoid-, and Sin-Based Nonlinear Activation Functions for Neuromorphic Photonics," *IEEE J of Sel.Top. Quantum Electron.*, vol. 29, no. 6: Photonic Signal Processing, pp. 1–10, Nov. 2023.
[6] O. Destras, S. Le Beux, F. G. De Magalhães, and G. Nicolescu, "Survey on Activation Functions for Optical Neural Networks," *ACM Comput Surv*, vol. 56, no. 2, 2024.
[7] A. Jha, C. Huang, and P. R. Prucnal, "Reconfigurable all-optical nonlinear activation functions for neuromorphic photonics," *Opt. Lett*, vol. 45, no. 17, 2020.
[8] G. H. Y. Li *et al.*, "All-optical ultrafast ReLU function for energy-efficient nanophotonic deep learning," *Nanophotonics*, vol. 12, no. 5, 2023.
[9] G. Mourgias-Alexandris, A. Tsakyridis, N. Passalis, A. Tefas, K. Vyrsokinos, and N. Pleros, "An all-optical neuron with sigmoid activation function," *Opt. Express*, vol. 27, no. 7, 2019.
[10] B. Shi, N. Calabretta and R. Stabile "InP photonic integrated multi-layer neural networks: Architecture and performance analysis," APL Photon., vol. 7, no. 1, 2022, Art. no. 010801
[11] J. Crnjanski, M. Krstić, A. Totović, N. Pleros, and D. Gvozdić, "Adaptive sigmoid-like and PReLU activation functions for all-optical perceptron," *Opt Lett*, vol. 46, no. 9, pp. 2003–2006, 2021.
[12] M. M. Krstić, J. V. Crnjanski, and D. M. Gvozdić, "Injection power and detuning-dependent bistability in Fabry-Perot laser diodes," *IEEE J of Sel.Top. Quantum Electron.*, vol. 18, no. 2, 2012.
[13] G. Mourgias-Alexandris *et al.*, "Noise-resilient and high-speed deep learning with coherent silicon photonics," *Nat. Commun.*, vol. 13, 2022.
[14] N. Passalis, M. Kirtas, G. Mourgias-Alexandris, G. Dabos, N. Pleros, and A. Tefas, "Training noise-resilient recurrent photonic networks for financial time series analysis," in *European Signal Processing Conference*, 2021.
[15] S. Kovaios et al, "Programmable Tanh- and ELU-based Photonic Neurons in Optics-Informed Neural Networks", submitted to *J. Sel..Top. Quantum Electron.*