

A Silicon Nitride Reconfigurable Linear Optical Processor

L. De Marinis¹, G. Contestabile¹, P. Castoldi¹, and N. Andriolli²

⁽¹⁾ Scuola Superiore Sant'Anna, Pisa, Italy

⁽²⁾ National Research Council of Italy (CNR-IEIT), Pisa, Italy

lorenzo.demarinis@santannapisa.it

Abstract: The characterization of a broadband Si_3N_4 integrated linear optical processor operating in the C-band is reported. The impact of losses on the processor accuracy is discussed towards the photonic implementation of state-of-the-art neural networks. © 2021 The Author(s)

1. Introduction

In the field of artificial intelligence, deep learning (DL) has become the most powerful and versatile tool, exploited in an ever-growing number of applications [1]. This resulted in the exponential growth of DL-related computational needs. In this scenario, reconfigurable linear optical processors are promising to answer this computational challenge, bringing the high speed and inherent parallelism of optics in neuromorphic applications. Photonic processors based on Mach-Zehnder interferometer (MZI) meshes are particularly suitable to be developed within integrated chips [2]. Several works discuss the implementation of such MZI-based processors in Silicon Photonics [3–5]. In this paper we report the characterization and resolution analysis of a C-Band 4×4 reconfigurable linear optical processor designed for the low-loss Si_3N_4 TriPleX platform [6]. The device has been fabricated and packaged within a multi-project wafer run by LioniX International. Differently from [7], we focused on reducing footprint and losses by exploiting a more balanced photonic integrated circuit topology [2]. In the last part of this manuscript, we discuss the scalability of passive optical processors towards the implementation of state-of-the-art neural networks in optics.

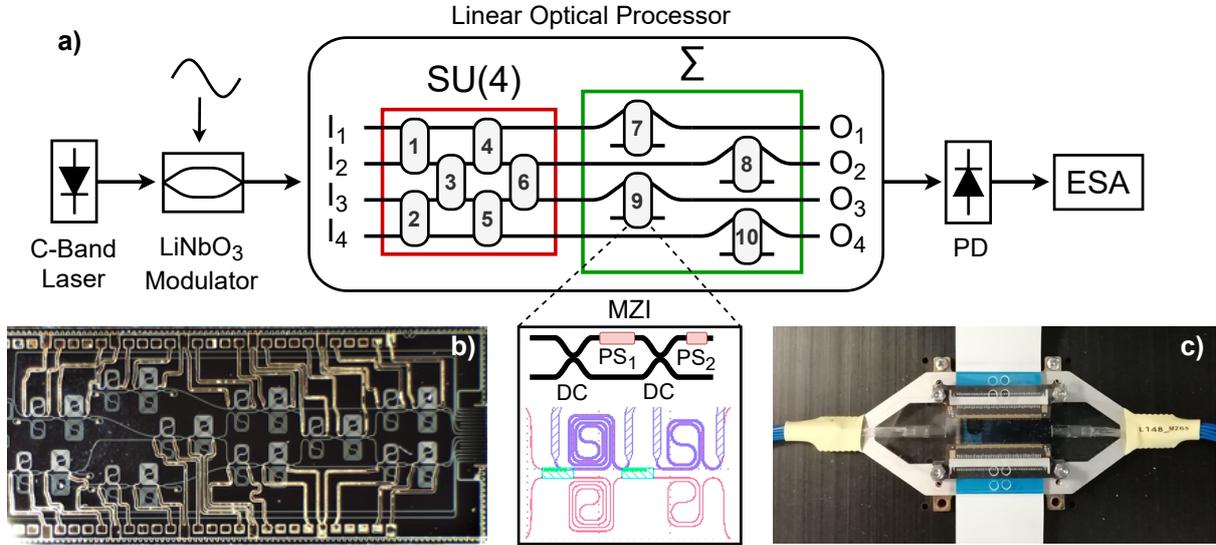


Fig. 1: (a) Scheme of the linear optical processor and measurement setup (PD: photodetector; ESA: electrical spectrum analyzer); (b) Si_3N_4 chip; (c) Packaged device with optical and electrical I/O.

2. Silicon Nitride Linear Optical Processor

The 4-port reconfigurable integrated optical processor presented in this paper can implement any linear transformation matrix and enable time-of-flight all-optical matrix-vector multiplications required in each layer of a deep neural network. Fig. 1(a) reports the processor architecture. The device has been fabricated and packaged by LioniX International in a multi-project wafer run exploiting the low-loss Si_3N_4 TriPleX platform that uses asymmetric double-strip waveguides [6]. It is made by ten thermally-tuned balanced MZI elements [4], each of them realized using two 3-dB couplers and two phase shifters: one inserted in one internal arm (PS_1) and the other in one of the outputs waveguides (PS_2), as shown in the inset of Fig. 1(a). The first one is used to adjust the coupling ratio at the MZI output, while the second one controls the relative phase of the outputs. The first six MZIs implement the 4×4 unitary transformation matrix (i.e., SU(4)), while the last four MZIs implement the diagonal matrix

multiplication section (i.e., Σ) used to control the optical power at the four outputs. The chip implements the so called ‘‘Clements’’ architecture, where the MZIs are arranged in a rectangular shape [2]. This is opposed to the triangular (‘‘Reck’’ architecture [2]) or diamond shape [7]. In this way, balanced optical paths and robustness to unbalanced optical losses are obtained, while also reducing the depth (defined as the maximum number of MZIs in the longest path) by one half.

The four inputs and the four outputs of the optical linear processor are terminated at the chip edge with spot size converters that guarantee low-loss coupling to SMF. To avoid the use of large driving currents, we designed relatively long low-loss phase shifters: 8 mm-long for the inner shifter, 4 mm-long for the external one. Also, spirals with a minimum radius of 105 μm have been used for a compact design. The resulting footprint of the the optical linear processor (shown in Fig. 1(b)) is about $16 \times 8 \text{ mm}^2$. For an easier chip operation, it has been placed on a submount (acting as a heat sink and supporting the fiber arrays) and then packaged, as depicted in Fig. 1(c): the right-hand side 16-element fiber array is dedicated to the optical linear processor and used to inject and retrieve inputs and outputs, while heater pads located on the top and bottom sides of the chip have been wire-bonded to the host PCB and connected to two FFC connectors for electrical driving.

3. Measurements and Results

3.1. Characterization of the MZI basic elements

The device has been characterized by finding the phase-current relation of every phase shifter based on the procedure described in [4]. Fig. 2(a) reports the normalized transmission spectrum of the worst-case transmission path, exhibiting a 3-dB bandwidth of about 40 nm, encompassing the whole C-band. Fig. 2(b) and 2(c) show the normalized transmission at 1550 nm of MZI 4 in cross and MZI 7 in bar, i.e., the best and worst one from the extinction ratio (ER) standpoint (16.1 dB and 7.5 dB, respectively). The average ER has been found to be 12.8 dB, mainly limited by the directional coupler unbalance from the nominal 50/50 splitting ratio.

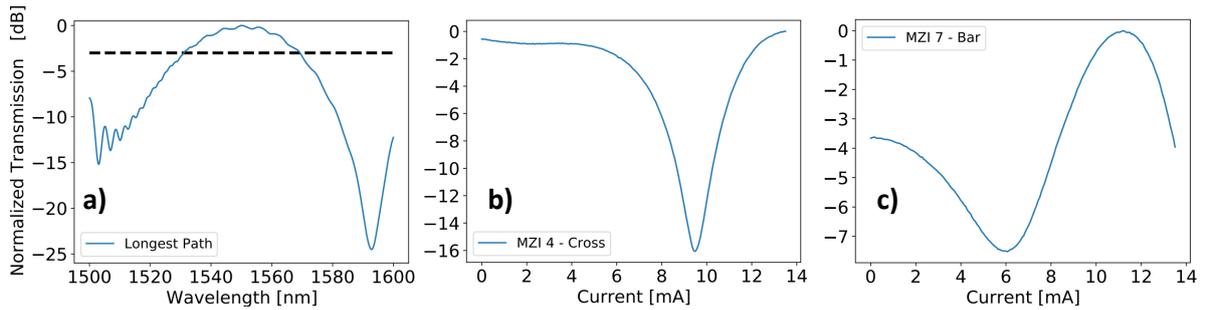


Fig. 2: (a) Optical processor bandwidth on the longest path; (b) MZI 4 normalized transmission in Cross; (c) MZI 7 normalized transmission in Bar.

To evaluate the insertion loss (IL) of the spot size converters and of the MZIs, transmission measurements were made on two waveguide loops used for the alignment of the fiber array and on the optical processor in different configurations. The average IL was 1.5 dB per spot-size converter and 1.5 dB per MZI at 1550 nm. MZI losses are mainly due to unexpected excess loss of the directional couplers.

3.2. Resolution Analysis

The presented photonic integrated circuit implements analog matrix-vector multiplications handling continuous signals. Nonetheless, in analog systems the ability to distinguish between two different values is limited by noise and distortions. This limitation is characterized by the SINAD, i.e., the signal over noise and distortion ratio, which directly translates into a finite resolution, measured by means of the effective number of bits (ENOB) [8]. This metric is useful when the analog optical processor acts as accelerator of a classical digital electronic processor. Indeed, although some recent interesting works on photonic memories [9], up to now there is no established photonic data storage system. For this reason, in order to store the output signals from the photonic analog engines, they must be converted back to the electronic domain with a following analog-to-digital conversion.

Fig. 1(a) shows the measurement setup used to evaluate the resolution of the photonic engine. A Lithium Niobate (LiNbO_3) intensity modulator operated in its linear region is used to impress a 10 GHz sinusoidal signal on continuous wave (CW) emitted by a laser at 1550 nm. The modulated signal passes through the photonic chip, is photodetected and analyzed by an electrical spectrum analyzer (ESA). As a reference, the electrical spectrum right after the LiNbO_3 modulator is also acquired to be compared to the spectra after crossing the Si_3N_4 chip.

Fig. 3(a) reports the ENOB reduction due to the device, as a function of the laser power, with respect to the reference signal (CW laser power range 1 mW-10 mW). Along with the reference signal, the graph reports the

ENOB reduction in the best and worst device path, i.e., $I_3 \rightarrow O_4$ and $I_4 \rightarrow O_1$ respectively. The acquired spectra do not show any significant distortion (< 1 dB total harmonic distortion degradation), therefore the ENOB reduction is mainly caused by chip losses impacting the signal to noise level. For large input power, the ENOB degradation with respect to reference is below 2, while it grows to 3 at lower input power. The ENOB reduction spread among different paths is below one for any laser power.

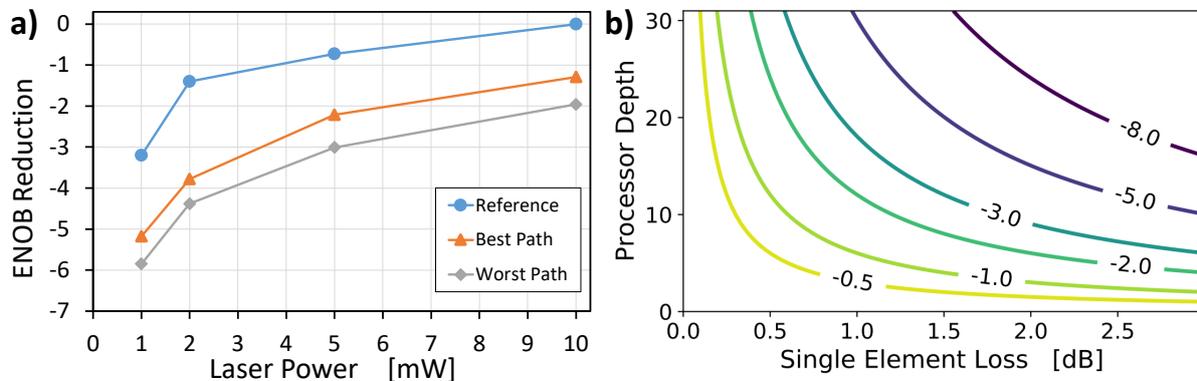


Fig. 3: (a) ENOB reduction as a function of laser power for a reference signal, best and worst path in the photonic device; (b) ENOB reduction as a function of processor depth and single MZI element loss.

As PIC losses are the main limiting factor for the scalability of these MZI-based optical processors, in Fig. 3(b) we also report the ENOB reduction as a function of the loss of a single MZI element and of the processor depth (i.e., the maximum number of crossed elements). Using the ‘‘Clements’’ architecture, $SU(N)$ matrices can be implemented with a processor depth N equal to the number of input ports. In order to implement any matrix using singular value decomposition, a diagonal matrix Σ and an additional $SU(N)$ section are required for a total processor depth of $2N + 1$ [2]. As the realized chip is characterized by a single element (MZI) loss of 1.5 dB, $N = 4$ inputs can be supported by admitting an ENOB reduction ≤ 2 , as confirmed by Fig. 3(a).

To put our results into the perspective of a real application, we consider state-of-the-art convolutional neural network (CNN), like the Google Inception-ResNet-v2, which exhibits super-human abilities in computer vision [10]. It uses convolutional filters implementing matrix-vector multiplications with up to $N=9$ inputs. Therefore, to implement Inception-ResNet-v2 with a linear optical processor, a depth of 19 must be guaranteed, requiring a single element loss < 0.7 dB (to keep ENOB reduction ≤ 2).

4. Conclusions

In this paper we presented a promising MZI-based linear optical processor manufactured and packaged in a commercial silicon nitride platform, implementing a 4×4 unitary transformation matrix (i.e., $SU(4)$) followed by a diagonal matrix multiplication section. Beyond the optical characterization of the PIC, we focused our analysis on the study of the impact of the device on the signal resolution. We found a worst-case 2 to 3 ENOB reduction on the tested power range (1 mW–10 mW), mainly due to chip losses. Despite the promising results, a further reduction of the single MZI losses is required to enable the photonic implementation of 9-input convolutional filters for state-of-the-art CNNs.

References

1. V. Sze *et al.* ‘‘Efficient processing of deep neural networks: A tutorial and survey,’’ *Proc. IEEE* 105.12 (2017): 2295–2329.
2. W. R. Clements *et al.* ‘‘Optimal design for universal multiport interferometers,’’ *Optica* 3.12 (2016): 1460–1466.
3. N. Harris *et al.* ‘‘Linear programmable nanophotonic processors,’’ *Optica* 5.12 (2018) 1623–1631.
4. F. Shokraneh *et al.* ‘‘Theoretical and Experimental Analysis of a 4×4 Reconfigurable MZI-Based Linear Optical Processor,’’ *J. Lightw. Technol.* 38.6 (2020): 1258–1267.
5. L. De Marinis *et al.* ‘‘Characterization and ENOB Analysis of a Reconfigurable Linear Optical Processor,’’ in *Proc. Photonics in Switching and Computing*, OSA (2020): PsW1F.4.
6. C.G.H. Roeloffzen *et al.* ‘‘Low-loss Si₃N₄ TriPleX optical waveguides: Technology and applications overview,’’ *IEEE J. Sel. Topics Quantum Electron.* 24.4 (2018): 4400321.
7. C. Taballione *et al.* ‘‘ 8×8 reconfigurable quantum photonic processor based on silicon nitride waveguides,’’ *Opt. Express* 27.19 (2019): 26842–26857.
8. A. Catania *et al.* ‘‘Analysis and Simulation of Chopper Stabilization Techniques Applied to Delta-Sigma Converters,’’ in *Proc. SMACD*, IEEE (2018): 253–256.
9. T. Alexoudi *et al.* ‘‘Optical RAM and integrated optical memories: a survey,’’ *Light: Science & Applications* 9.1 (2020).
10. C. Szegedy *et al.* ‘‘Inception-v4, Inception-ResNet and the impact of residual connections on learning,’’ in *Proc. AAAI Conference on Artificial Intelligence*, (2017) 4278–4284.