

Foundry Manufactured 6-bit Resolution, 150 μm Long Slow-Light Electro-Optic Modulator for On-Chip Photonic Tensor Computing

Meng Zhang¹, Amir Begović¹, Dennis Yin², Nicholas Gangi¹, Jiaqi Gu², Rena Huang¹

¹ Rensselaer Polytechnic Institute, Department of Electrical, Computer and Systems Engineering, 110 8th Street, Troy, NY 12180, USA
Author e-mail address: {zhangm19, begova, gangin2, huangz3}@rpi.edu

² Arizona State University, School of Electrical, Computer and Energy Engineering, Tempe, AZ 85281, USA
Author e-mail address: {ziangyin, jiaqigu}@asu.edu

Abstract: We demonstrate 6-bit DAC resolution with an ultra-compact slow-light electro-optic modulator. The 10 \times modulation length reduction enables 31 \times compute density, 1.17 \times energy efficiency, and 36.1 \times energy efficiency per unit area for photonic tensor computing.

1. Introduction

On-chip electro-optic Mach-Zehnder Modulators (MZMs) are an essential element in converting electrical signals to the optical domain for on-chip optical tensor core computing. For an $N \times N$ dynamic photonic tensor core (PTC) [1], $2N$ MZM devices are required to modulate both input matrices. While the standard Si MZM processing design kit (PDK) element offered at Si photonics foundries has a typical phase shifter length of 1 mm to 3 mm, it severely impacts the computing unit area efficiency due to high cost arising from Si chip surface area consumption and ultimately will limit the scalability of photonic accelerators due to the maximum manufacturable wafer size.

Up to date, up to 5-bit resolution are reported for PTC applications [2,3]. In this work, we report the demonstration of a 6-bit resolution, ultra-compact slow-light enabled MZM (SL-MZM) with 150 μm modulation length for photonic analog computing in a dynamically operated PTC [1]. The SL-MZM has demonstrated a $\sim 10\times$ modulation length reduction compared to the MZMs in the standard PDK, which greatly reduces the system footprint and power consumption.

2. SL-MZM design and characterization

A 1D dielectric photonic crystal waveguide, namely a rectangular-shaped Bragg grating, is employed to provide a slow-light effect for modulator size reduction. A schematic of the SL-MZM is shown in Fig. 1(a), and it has a design similar to the previously reported SL-MZM device in [4]. The center waveguide width (W_{min}) is 400 nm, and the maximum grating width (W_{max}) is 800 nm. The grating is constructed with a period (Λ) of 290 nm and follows super Gaussian apodization for side mode suppression. Both arms have identical doping profiles with concentrations of $\sim 10^{19} \text{ cm}^{-3}$, while single end/single side driving scheme is utilized in the grating arm. The high doping concentration allows the phase shifter to be very compact. Hence, the phase shifter length (L_{PS}) is only 150 μm . Heaters were also included in the MZM but were unused for this testing. The SL-MZM was fabricated at American Institute for Manufacturing (AIM) Photonics under multi-project wafer (MPW) run.

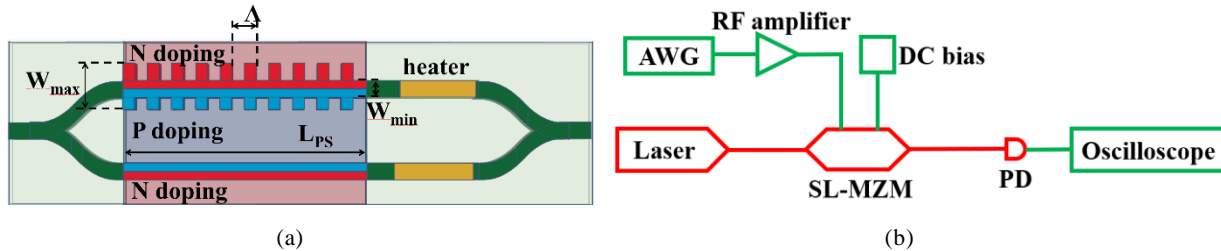


Figure 1. SL-MZM for bit resolution testing. (a) Schematic of the SL-MZM. (b) Schematic of the testing setup.

During device characterization, an arbitrary waveform generator (AWG) was used to generate 3-bit, 4-bit, 5-bit, and 6-bit signal trains. The AWG signal was amplified by an RF amplifier to reach a maximum V_{pp} of 3.5V. The modulator is reverse-biased at -3.8V, so the input data spread over the range of $-3.8\text{V} \pm 1.75\text{V}$. The modulator output is routed to an on-chip Ge photodetector (PD) that is a standard AIM PDK component. A schematic of the testing setup is shown in Fig. 1(b). The readout signals from the PD are displaced on a real-time oscilloscope shown in Fig. 2.

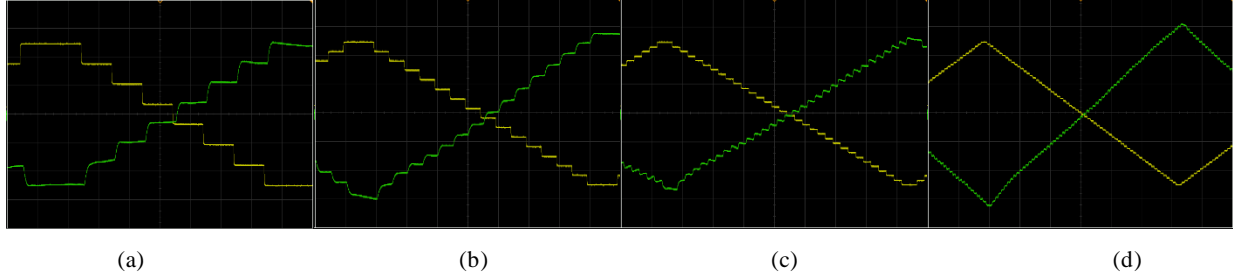


Figure 2. Bit resolution testing results of SL-MZM at (a) 3 bits (Clock rate: 10MHz), (b) 4 bits (Clock rate: 20MHz), (c) 5 bits (Clock rate: 40MHz), and (d) 6 bits (Clock rate: 100MHz). Yellow curves show direct driving signals from the AWG, while the green curves represent the SL-MZM response readout by the on-chip PD.

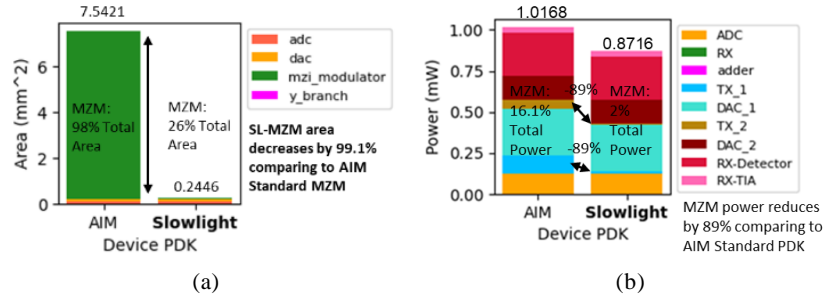


Figure 3. (a) Area and (b) power breakdown, simulated with four 5×5 PTCs @ 10GHz, 1550nm.

Table I: Simulated computation density and energy efficiency for four 5×5 PTCs @ 10GHz, 1550 nm.

Input Mod.	TOPS	TOPS /mm ²	TOPS /W	TOPS /W/mm ²
AIM Foundry MZM	50	6.6	49.2	6.5
Slow-Light MZM	50	204.4	57.4	234.5

3. System-level simulation results on dynamic photonic tensor cores

To evaluate the advantage of our demonstrated slow-light MZM at the system level, we construct four single-wavelength 5×5 dynamic tensor cores [1] with AIM foundry MZM and our slow-light MZM as the input modulators and simulate the area, power breakdown in Fig. 3. With the AIM foundry MZM, the PTCs take 7.5 mm^2 area, where MZMs contribute 7.36 mm^2 (97.6%). In contrast, our slow-light MZM is very compact, which only takes 0.0625 mm^2 , accounting for 25.6% of the total area (0.245 mm^2). Regarding power consumption, we estimated the power according to $1/4CV^2$ and a $\sim 10\times$ reduction in capacitance for the SL-MZM, leading to an 89% reduction in input modulation power and a 14.28% reduction in system-level power consumption. At 10 GHz input rate and a single wavelength (1550 nm), we compare the compute density and energy efficiency in Table I. Our compact and energy-efficient SL MZM can enable 204.4 TOPS/mm^2 compute density, 57.4 TOPS/W energy efficiency, and $234.5 \text{ TOPS/W/mm}^2$ energy efficiency per unit area, outperforming AIM foundry MZM by $31\times$, $1.17\times$, and $36.1\times$, respectively.

4. Conclusion

A miniaturized Bragg grating slow light enabled Si MZM was employed for on-chip photonic tensor computing. A $10\times$ device length reduction has demonstrated. The device-level advantage of small footprint MZMs can directly translate to significant system-level efficiency boosts, pushing forward the scalability and efficiency frontier of photonic tensor cores for machine learning applications and scientific computing.

5. References

- [1] Zhu, H., et al., "DOTA: A Dynamically-Operated Photonic Tensor Core for Energy-Efficient Transformer Accelerator," in IEEE International Symposium on High-Performance Computer Architecture (HPCA), Mar 2024, *arXiv preprint arXiv:2305.19533*.
- [2] Ma, Xiaoxuan, et al. "High-density integrated photonic tensor processing unit with a matrix multiply compiler." (2022).
- [3] Feldmann, Johannes, et al. "Parallel convolutional processing using an integrated photonic tensor core." *Nature* 589.7840 (2021): 52-58.
- [4] S. R. Anderson, A. Begović and Z. R. Huang, "Integrated Slow-Light Enhanced Silicon Photonic Modulators for RF Photonic Links," in IEEE Photonics Journal, vol. 14, no. 4, pp. 1-6, Aug. 2022, Art no. 6647106, doi: 10.1109/JPHOT.2022.3185888.