Inversely designed photonic integrated vector dot-product core with mode-division multiplexing

Zheyuan Zhu[1], Raktim Sarma[2], Seth Smith-Dryden[1], Guifang Li[1], and Shuo S. Pang[1]
[1]*CREOL, The College of Optics and Photonics, University of Central Florida, 4304 Scorpius St., Orlando, FL 32816, USA.*
[2]*Center for Integrated Nanotechnologies, Sandia National Laboratories, 1515 Eubank Blvd SE, Albuquerque, NM 87185, USA.*
Author e-mail address: zyzhu@knights.ucf.edu

Abstract: We present an inversely designed integrated photonic dot-product core based on mode-division multiplexing. The core features a 5μm×3μm footprint for scalability and can perform general-purpose vector dot-products with easily reconfigurable inputs for various computing applications. © 2024 The Author(s)

1. Introduction

Photonic computing presents a promising alternative to electronics in handling high-throughput, data-intensive computing applications, as photonic signals can pack a multitude of degrees of freedom, including wavelength, spatial mode, polarization, and phase quadratures in a single transmission line at any given time. Driven by the computing demand in artificial neural networks, analog computing platforms based on integrated photonic devices [1] have demonstrated the potential of higher efficiency than the electronic counterparts, due to the intrinsically passive photonic multiply-accumulate (MAC) process [2]. The existing photonic computing platforms mostly rely on wavelength-division multiplexing or space-multiplexing to achieve parallelism. Using spatial mode as an additional degree of freedom only begins to emerge as a viable approach in high-bandwidth optical communication [3], but its application in photonic computing has not been exploited. In this work, we present an integrated mode-division multiplexed (MDM) photonic computing core using inverse design. The core performs dot-products with easily reconfigurable inputs and supports the deployment of various computing applications. As a specific application, we have demonstrated a complex number multiplier in Sec. 3.

2. Photonic computing core design and fabrication

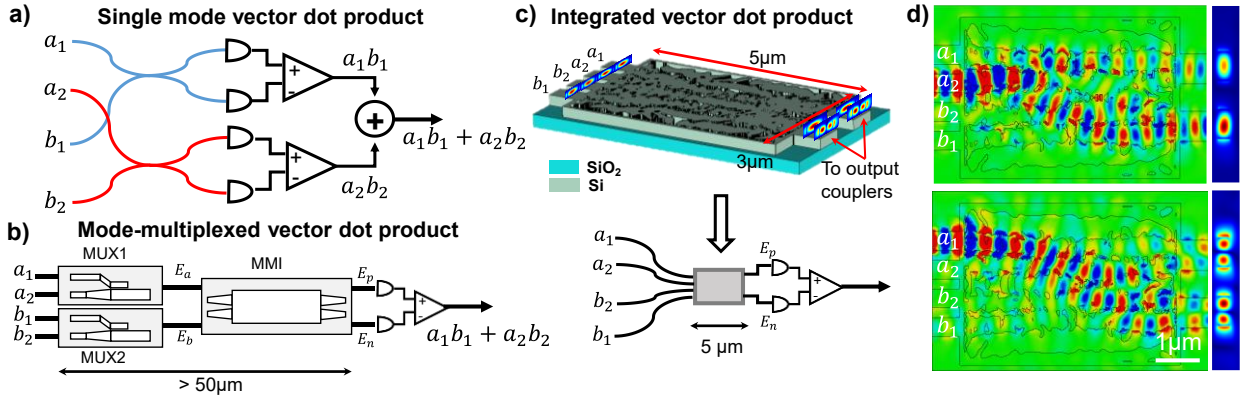

Fig. 1. Operating principle and design of the mode-multiplexed dot-product core.

A photonic MDM-based dot-product core supporting two input vectors, $\vec{a} = (a_1, a_2)^T$ and $\vec{b} = (b_1, b_2)^T$, consists of four input ports and two output ports. Fig. 1 schematically illustrates the operating principle of performing dot-product on 2-element vectors. Vector dot-product based on coherent mixing using single mode components requires 2 balanced detections for multiplications and separated accumulation in electronic domain that consumes energy (Fig. 1(a)). In the mode-multiplexed setup, two elements in the vector, a_1 (b_1) and a_2 (b_2), are mapped to the fundamental (ψ_{I}) and the second order (ψ_{II}) TE modes of a few-mode waveguide, respectively. The mode-multiplexed photonic signals, $E_a = a_1\psi_{\mathrm{I}} + a_2\psi_{\mathrm{II}}$ and $E_b = b_1\psi_{\mathrm{I}} + b_2\psi_{\mathrm{II}}$, undergo coherent mixing, producing the electrical fields on the upper and lower arms $E_p = \frac{1}{\sqrt{2}}(E_a + iE_b)$ and $E_n = \frac{1}{\sqrt{2}}(iE_a + E_b)$. Based on the orthogonality between ψ_{I} and ψ_{II}, the difference between the overall intensity of the upper and lower outputs $I_{diff} = |E_p|^2 - |E_n|^2$ naturally produces the dot-product between vectors $\vec{a}$ and $\vec{b}$. Since accumulation is carried out in optical domain, and only requires one balanced detection, the power consumption could be 2-3 folds lower than the single mode implementation.

Using conventional MDM devices, the dot-product core requires 2 mode multiplexers (MUXs, 20μm×4μm each [4]) for encoding $\vec{a}$ and $\vec{b}$, and 1 2×2 mixer based on multi-mode interference (MMI, 40μm×6μm [5]), giving an

overall footprint of at least 60μm×10μm (Fig. 1(b)). Our dot-product core packs the functionality of 2 MUXs and 1 MMI into a single symmetric layout (Fig. 1(c)) with a footprint of 5μm×3μm. The core is designed following a gradient-based optimization framework [6] using an in-house 3D vector-field finite-difference frequency-domain (FDFD) Maxwell equation solver with perfectly matched layers [7]. The upper and lower output few-mode waveguides are tapered to photonic crystal structures, which vertically couple out the light for imaging by a camera.

3. Results and discussions

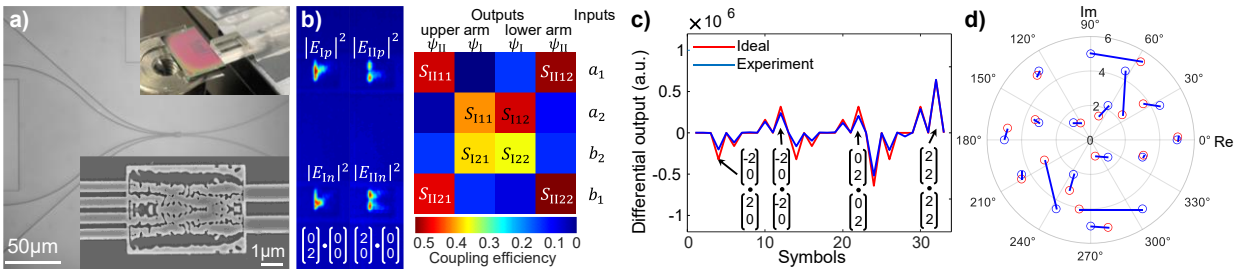

Fig. 2. Testing of the fabricated MDM dot-product core. (a) Optical, SEM, and setup images of the fabricated device under testing. (b) Spatial mode profiles and S-matrix of the fabricated device. (c) Pre- and post-compensation waveforms representing a TDM dot-product sequence. (d) Complex mulplier deployed as TDM symbol sequence on the core.

The intensity profiles on the vertical output couplers of the fabricated device (Fig. 2(a)), when input 1 and 2 are individually excited, are shown in Fig. 2(b). The profiles match the target spatial modes. Fig. 2(b) also plots the S-matrix obtained from Lumerical FDTD simulation of the fabricated device. Although the splitting ratio between the upper and lower arm is not strictly 50/50 due to fabrication imperfections, the crosstalk between two spatial modes is low. Therefore, we can treat two spatial modes independently in the S-matrix of the fabricated device.

$$\begin{bmatrix} E_{\mathrm{II}p} \\ E_{\mathrm{I}p} \\ E_{\mathrm{I}n} \\ E_{\mathrm{II}n} \end{bmatrix} = \begin{bmatrix} S_{\mathrm{II}11} & 0 & 0 & S_{\mathrm{II}12} \\ 0 & S_{\mathrm{I}11} & S_{\mathrm{I}12} & 0 \\ 0 & S_{\mathrm{I}21} & S_{\mathrm{I}22} & 0 \\ S_{\mathrm{II}21} & 0 & 0 & S_{\mathrm{II}22} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ b_2 \\ b_1 \end{bmatrix}. \quad (1)$$

Following Eq. (1), the differential output in mode I is $I_{Idiff} = |a_2|^2(|S_{\mathrm{I}11}|^2 - |S_{\mathrm{I}21}|^2) + |b_2|^2(|S_{\mathrm{I}12}|^2 - |S_{\mathrm{I}22}|^2) + 2Re[a_2 b_2^*(S_{\mathrm{I}11}S_{\mathrm{I}12}^* - S_{\mathrm{I}21}S_{\mathrm{I}22}^*)]$, and similar output can be derived for mode II. This indicates that to produce the result proportional to the dot-product, the uneven splitting can be corrected by a calibration process that removes the differential outputs when only one of the four inputs is excited.

We generated two 2-element input vectors containing arbitrary, signed elements from four off-chip fiber-based Mach-Zehnder modulators (JDSU OC-192) and edge-coupled the modulated signals into the core. Fig. 2(c) plots the dot-product results after correcting for the uneven splitting between the upper and lower arms. After calibration, the core natively supports signed 3-bit inputs (integers from -2 to 2) with an output dynamic range of signed 4-bits (integers from -8 to 8).

Fig. 2(d) demonstrates a general-purpose complex number multiplier ($(a + bi) \times (c + di)$) on the core. The real and imaginary parts of the result ($(ac - bd) + (ad + bc)i$) are split into two equivalent dot-products deployed in a time-division multiplexing (TDM) symbol sequence. 16 complex numbers represented by a sequence of 32 symbols were multiplied on the core. The normalized root mean square error of all complex multiplications was 15.9%, representing signed 6.27 effective levels in the results. The computing error is primarily attributed to the time-varying phase of the four off-chip fiber inputs, which were not always constructively aligned within the symbol sequence. This error can be greatly reduced with on-chip modulators.

In summary, we have designed and fabricated a compact, integrated photonic dot-product core using inverse design. The core utilizes spatial mode as the multiplexing dimension to perform arbitrary 2-element vector dot-products. We have demonstrated the deployment of a general-purpose complex number multiplier on the core. The small (5μm×3μm) footprint enables high-density integration of the core in a parallel computing array.

References

[1] Y. Shen *et al.*, Deep learning with coherent nanophotonic circuits, Nat. Photonics, 11, 441–446 (2017)
[2] B. J. Shastri *et al.*, Photonics for artificial intelligence and neuromorphic computing, Nat. Photonics, 15, 102–114 (2021)
[3] X. Wu *et al.*, Mode-Division Multiplexing for Silicon Photonic Network-on-Chip, J. Light. Technol., 35, 3223–3228 (2017)
[4] D. Dai *et al.*, Silicon mode (de)multiplexer enabling high capacity photonic networks-on-chip with a single-wavelength-carrier light, Opt. Lett., 38, 1422 (2013)
[5] H. Shiran *et al.*, Dual-mode broadband compact 2 × 2 optical power splitter using sub-wavelength metamaterial structures, Opt. Express, 29, 23864 (2021)
[6] L. Su *et al.*, Nanophotonic inverse design with SPINS: Software architecture and practical considerations, Appl. Phys. Rev., 7, 011407 (2020)
[7] W. Shin & S. Fan, Choice of the perfectly matched layer boundary condition for frequency-domain Maxwell's equations solvers, J. Comput. Phys., 231, 3406–3431 (2012)