Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Programmable Zero-Power Silicon Photonic Read-Only Memory and Compute In-Memory Circuits

Stylianos Simos

simostyl@csd.auth.gr

Aristotle University of Thessaloniki

Isaac Johnson

ij1g20@soton.ac.uk

University of Southampton

Research Article

Keywords:

Posted Date: April 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9235732/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Programmable Zero-Power Silicon Photonic Read-Only Memory and Compute In-Memory Circuits

S. Simos[*,†,1,2], I. Johnson[†,3], T. Moschos[1,2], S. Kovaios[1,2], A. Tsakyridis[1,2], M. Moralis-Pegios[1,2], T. Rangarajan[3], T. D. Bucio[3], J. Gates[3], F. Gardes[3] and N. Pleros[1,2]

[1]*Department of Informatics, Aristotle University of Thessaloniki, 54124, Thessaloniki, Greece*

[2]*Center for Interdisciplinary Research and Innovation, 57001, Thessaloniki, Greece*

[3]*Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, United Kingdom*

*simostyl@csd.auth.gr

† *These authors contributed equally to this work*

Abstract

Wide adoption of optical computing systems necessitates the use of integrated photonic platforms that can synergize non-volatile long-term information storage with efficient linear operation architectures. Non-volatile optical memory operation is emerging as a next generation compute system enabler, as this can support high-speed yet zero electrical power for traditional Read-Only-Memory (ROM) deployments required by firmware, boot procedures and lookup tables. At the same time, this enables a diverse range of compute-in-memory (CIM) edge inference tasks that rely on fixed or rarely updated model parameters, provided that non-volatile memory allows also for analog operation and gets combined with optimized linear optical circuit architectures. In this work, we report the first fully CMOS-compatible silicon photonic programmable read-only memory (SiPROM) and combine this with different architectural schemes to demonstrate experimentally a 36 bit zero-power digital SiPROM bank, a coherent analog 4x9 CIM unit, and a 2-bit 9-word zero-power 50 Gb/s content-addressable memory (CAM) bank. The SiPROM memory cell comprises an n-rich SiNx-based Mach–Zehnder interferometer (MZI) with its memory state being programmed through controlled waveguide refractive index modifications caused by ultraviolet (UV) beam exposure, enabling in this way zero-power non-volatile storage without electrical biasing, thermal tuning, or the use of exotic materials. We demonstrate a zero-power digital SiPROM bank by utilizing this memory cell technology within a matrix arrangement, with its functionality being elevated to an analog coherent CIM circuit and a high-speed CAM bank when the matrix arrangement gets combined with linear optical circuit architectures. We believe that the fully CMOS compatible nature of this SiPROM technology together with its successful demonstration across a range of scaled-up digital and analog system-scale functions can pave the way towards high-speed and zero-power memory and compute operations required by critical communications and edge AI segments.

Introduction

The on-going artificial intelligence boom is rapidly transforming computing hardware architectures from conventional von Neuman systems, towards memory-centric processing deployments, where memories are envisioned to actively participate in processing pathways [1]. This approach, commonly defined as in-memory computing, minimizes data-

movement overheads by executing matrix-vector multiplication (MVM), convolution and other linear transformations directly within the memory cells[2]-[5]. In this context, read-only memories (ROMs) are emerging as a next generation computing system enabler [6],[7], as they can both support the high-speed and low-power requirements of traditional ROM deployments, e.g. firmware, boot procedures, lookup tables, and control logic, while also enabling a diverse range of edge inference tasks that rely on parameters that remain fixed or are rarely updated, making read-only compute-in-memory (CIM) and content-addressable memory (CAM) architectures particularly attractive. Combining, this approach with a photonic platform that can support both digital and analog storage and compute operations can enable zero-power, ultra-fast, and scalable optical memory and computing for next-generation edge and AI applications [8],[9].

A diverse range of physical mechanisms has been explored to implement optical memories on integrated platforms, each offering distinct trade-offs in terms of speed, retention time, programmability, and fabrication complexity. These approaches include volatile memories based on carrier dynamics [10],[11] and non-volatile memories relying on persistent material phase transitions or magnetic domain configurations. However, existing memory schemes introduce practical constraints that impose critical limitations; volatile mechanisms rely typically on carrier dynamics and require the information to be encoded through carrier-induced refractive index or absorption changes [10], achieving in this way fast reconfiguration times but inherently consuming significant power due to their need for an electrical bias. On the other hand, non-volatile solutions, such as phase-change materials (PCMs) [12], metal-oxide-semiconductor (MOS) capacitors [13], magneto-optic elements [14], and ferroelectric devices [15], offer the advantage of persistent state retention without continuous biasing, minimizing the power consumption. However, they often rely on specialized materials or complex processing steps that are not fully compatible with standard silicon photonics and conventional CMOS fabrication, complicating large-scale integration. On top of that, many of these platforms face challenges related to long-term stability and reproducibility, which can limit their practical deployment in scalable photonic memory architectures [16].

In this paper, we experimentally present the first fully CMOS-compatible analog silicon photonic programmable read-only memory (SiPROM) technology and report on the successful demonstration of a zero-power digital SiPROM bank, a zero-power analog CiM circuit and a zero-power 50 Gb/s Content Addressable Memory (CAM) bank at a record-high memory capacity of 4x9 memory cells. The proposed prototype comprises an n-rich SiNx 4×9 memory matrix layout, where each memory cell comprises a Mach-Zehnder Interferometer (MZI) that can be programmed through ultraviolet (UV) exposure at wafer scale. This exposure induces a semi-permanent modification in the refractive index of the SiNx material, generating a controlled phase shift between the two MZI arms, which in turn produces a distinct and stable optical attenuation state with a bit resolution of 5.5 bits. The experimental characterization of the 4x9 matrix architecture reveals a successful 36-bit SiPROM bank with clearly distinguishable optical attenuation levels that correspond to stored “1” and “0” logical values. Combining this memory bank with a linear optical circuit architecture turns the SiPROM bank into a 4x9 zero-power CIM circuit that performs high accuracy matrix vector multiplication (MVM) operations with a mean error of ~2.3 %. Finally, the SiPROM bank configured within the linear circuit architecture is used for demonstrating experimentally a 50 GHz ternary content-addressable memory (TCAM) bank that delivers precise and reliable comparison and look-up operations for 50 Gb/s optical words, validating the technology credentials to perform high-speed in-memory photonic processing.

Results

Silicon photonic PROM: Basic memory cell and applications

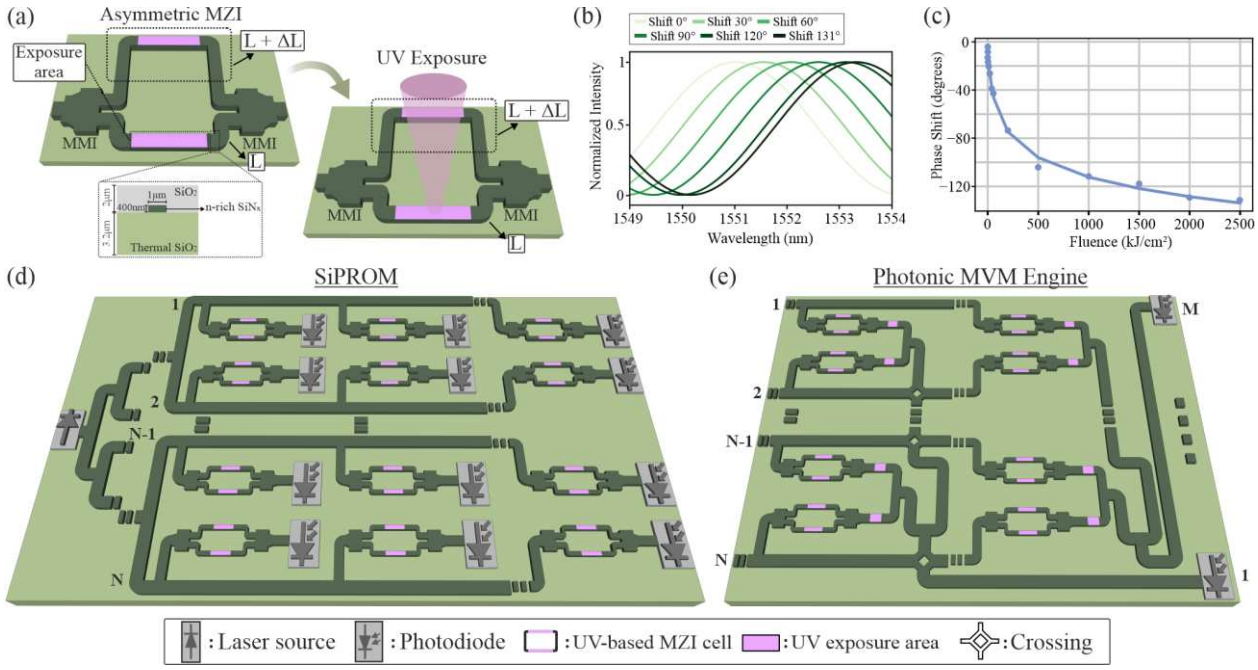


Fig. 1. a UV exposure process of an n-rich SiN AMZI. **b** Evolution of the AMZI transfer function under different UV fluence. Each irradiation step induces a controlled modification of the refractive index in the targeted arm, resulting in progressive phase shift in the transfer function **c** Achieved phase shift with respect to UV fluence. **d** N×M Xbar-based SiPROM architecture to realize ROM operation. **e** N×M Xbar-based photonic MVM engine to implement in-memory computing and CAM operation.

Figure 1 (a)-(c) summarize the process employed for the formulation of the UV-based SiPROM cells, leveraging the interaction between n-rich SiNx and UV-beams. In order to select an appropriate fluence for the UV spot, a set of test structures consisting of multiple asymmetric MZI (AMZI) memory cells are exposed, with more information about the fabrication process of the device found in Methods section: *Photonic platform and waveguide specifications*. By focusing a 244 nm continuous-wave UV beam to a ~14 μm spot on an AMZI arm, a permanent decrease in the waveguide's refractive index is induced [17]. Figure 1(b) and (c) illustrate the effect of the refractive index change on the AMZI memory cell when a single 14 um spot is UV-exposed with increasing fluence levels from 0.5 up to 2500 kJ/cm^2. Specifically, Fig. 1(b) monitors the shift of the AMZI transfer function, while Fig.1(c) quantifies the achieved phase shift. With the minimum fluence of 0.5 kJ/cm^2 level achieving a 4° phase shift, the programmable UV-based MZI memory cell can reach a total number of 45 states (180°/4°=45). This translates to a record-high bit resolution of 5.5 for non-volatile memories, significantly outperforming traditional PCM approaches and magneto-optic implementations [14]. More information on the UV-exposure process is provided in the Methods section: *UV-process: Setup and basic cell characterization* and in Supplementary section 1: *Post fabrication UV-exposure process of MZI cells*.

Building upon the UV-based memory cell, an N×M SiPROM Memory Bank architecture can be implemented by utilizing a matrix-based memory cell arrangement in rows and columns, as shown in Fig. 1(d). A continuous wave (CW) beam is uniformly divided into N rows, and each row is further equally split into M individual columns to address the full SiPROM matrix. The resulting beams are coupled into the UV-based MZI memory cells, where they interact with the stored content. The output is finally captured by a photodetector (PD) to complete the memory readout. The same UV-based MZI memory cells can also function as computing elements for CIM and CAM operations if they get incorporated in a linear optical circuit architecture, as shown in Fig. 1(e). Specifically, Fig. 1 (e), illustrates a MVM engine that synergizes several UV-based MZI memory cells with the loss-optimized Xbar-based linear operator [18] to deliver high-fidelity in memory computation. As extensively analyzed in [19], an N×M Xbar can store a N×M matrix W by mapping every matrix element to a respective memory cell and then perform MVM operation between the matrix W and an incoming N-dimensional input vector X. Specifically, N optically modulated signals representing the N elements of X are inserted into the Xbar, with each signal being injected at a different Xbar input. As such, each input signal gets then

distributed across the M columns of the Xbar. In each column, every MZI memory cell multiplies the respective input row signal with the value stored as the memory cell content and the N different products are then coherently combined through an N:1 combiner, completing in this way the dot product between the input vector and the corresponding W column vector. Coherent addition is ensured in the case of single-wavelength X-vector encoding by employing a UV-based phase shifter (PS) at every row to guarantee constructive interference among the individual products computed in each column. The same layout can, however, also support multi-wavelength X-vector encoding where every input vector element gets imprinted onto a different wavelength; in this case, all different wavelengths emerge at a single column output and have their power levels summed-up at a respective photodiode after propagating through the N:1 beam combiner. An optimized setup for incoherent addition can be supported by replacing the N:1 combiner with a N-wavelength multiplexing stage in order to avoid the 3dB losses enforced at every stage of the N:1 combiner.

36-bit SiPROM: Layout and experimental validation

Fig. 2(a) illustrates the photonic integrated circuit (PIC) prototype that was designed and fabricated on a N-rich SiNx platform and employs a total number of 36 SiPROM memory cells configured within a 4×9 Xbar architecture. The area within the white rectangle indicates the constituent 36–UV-based symmetric MZI memory cells, while the inset presents a close-up image of the basic MZI memory cell and highlights the raster pattern employed for UV beam alignment at the MZI arms. Figure 2 (b) depicts the experimental process followed for validating the 36-bit SiPROM memory bank. A laser source generates a CW signal that is inserted into the SiPROM PIC, gets then split into the 4 rows and broadcasted across all 9 columns. The stored information of each cell hosted at a certain row is accessed by blocking the propagation of the CW signal through the remaining 3 rows, so that only a single row is illuminated. As such, the power level of the signal emerging at every Xbar column output will originate from the attenuation factor enforced by a single memory cell. A logical value of "1" corresponds to the minimum optical attenuation (~0.5 dB), while a high optical attenuation (~14-21 dB) corresponds to a stored logical "0" value. The optical signals emerging at every Xbar column output are captured by a PD and corresponds to the equivalent setup of having a PD placed directly at the output of a

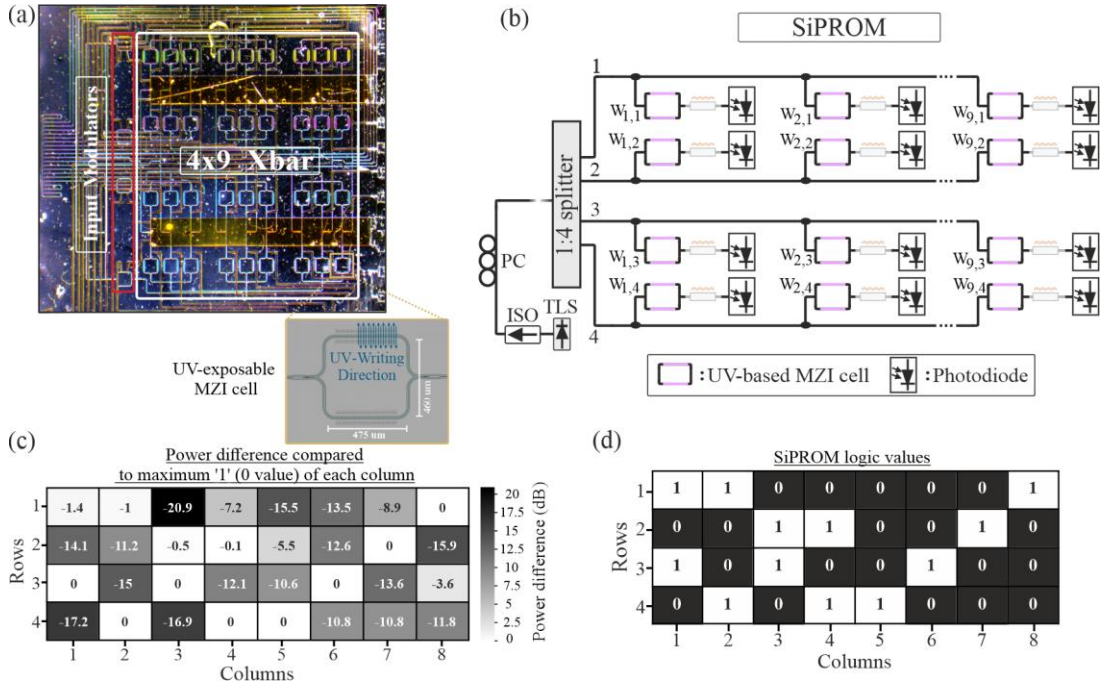


Fig. 2. a. Microscope image of the silicon photonic integrated 4×9 crossbar array, shown in the green highlighted area, along with the input modulation stage (yellow highlighted area). Inset: UV-based MZI cell, with dimensions of: 475×460 μm. **b.** Experimental setup for the validation of the SiPROM operation for the 4×9 Xbar array. **c.** Heatmap of power differences of every cell compared to the maximum "1" (0 value) of each column. **d.** SiPROM logic values assignment for all respective crossbar rows and columns.

single memory cell, as shown in Fig. 2(b), given that the simultaneous propagation of the CW light through more than one rows is not allowed.

Figure 2 (c) juxtaposes the experimentally measured attenuation factors of each memory cell arranged as a 4×9 matrix and normalized on a per column basis. Normalization is performed by considering in every column the highest power level as the 0dBm reference value and then representing the value of every memory cell within the same column as the additional attenuation level enforced by this cell. The resulting power difference heatmap indicates efficient differentiation between the values of the SiPROM cells, with power differences ranging from 3.6 dB up to 20.9 dB between the two memory states when using a CW input signal at 1575nm, as has been extracted based on the experimentally measured spectral responses, presented at the section 2 of Supplementary material: *Optical spectrum characterization and memory-state contrast of the photonic SiPROM*. The variations between the memory cells attenuation factors can be ascribed to the fabrication imperfections of MZIs, as well as to non-uniformity in the UV exposure process. The minimum power difference can increase from 3.6 to `5dB by using a 1560nm wavelength for the input signal, at the expense of slightly higher insertion losses. By setting a decision threshold near the worst-case power difference (3.6 dB at 1575nm) and declaring as logical "1" and "0" values the power levels above and below the threshold, respectively, the analog SiPROM implementation turns into a robust digital memory, as illustrated by the resulting binary-valued matrix in Fig. 2(d).

Compute-In-memory Circuit: Layout and experimental validation

Photonic in-memory computing can be realized by the same SiPROM PIC prototype by allowing the incoming optical signal to propagate simultaneously through all four Xbar rows. In this case, assuming that a different value X_i is carried by every input signal that reaches a different row, the proposed layout gets transformed into a CIM circuit that performs a MVM operation. Fig. 3(a) depicts the experimental setup used to evaluate the performance of the 4×9 Xbar-based CIM MVM. A laser source is employed to produce a CW signal, which is launched into the Xbar integrated circuit. The injected signal is subsequently split using a 1:4 splitter, with each resulting component directed to the input-vector stage, which consists of four identical thermo-optic (TO) Mach–Zehnder modulators (MZMs), denoted by the red rectangle in Fig. 3(a) and in the microscope image of Fig. 2(a). The electrical driving state of each modulator encodes a corresponding binary value of '0' or '1', collectively forming a 4-bit input vector. The modulated signal from each Xbar row (X_1-X_4) is uniformly broadcast to all columns of the array using two cascaded 1:3 multimode interference (MMI) splitters and is then delivered to each UV-based MZI memory cell, where the multiplication operation is realized. Subsequently, each individual product is coherently combined on a per-column basis using thermo-optic PSs, yielding

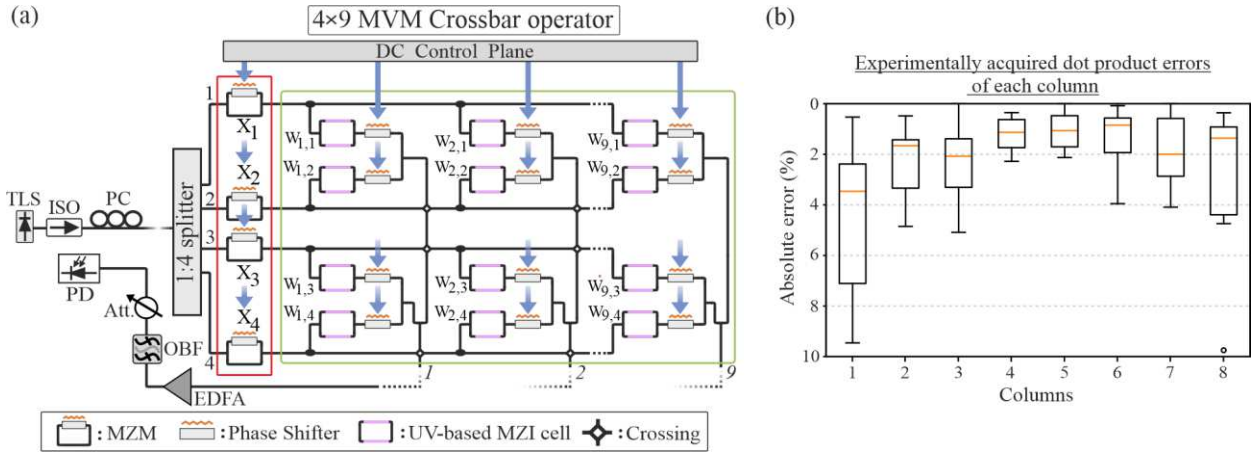


Fig. 3. a. Experimental setup of the 4 × 9 MVM operator. The yellow highlighted area depicts the input vector stage, that includes 4 on chip TO MZMs, used to generate the input vectors. The green highlighted area shows the crossbar-based configured weight matrix, including all the respective UV-exposable weighting elements. **b.** The mean error at each Xbar column output, calculated as the absolute difference between the experimentally obtained dot products and the corresponding values acquired from the proposed simulation model.

the final dot-product value at each Xbar column output, which is recorded by a PD. A simulation model was developed to predict the output products for each column based on the originally targeted memory cell values, so as to compare between the theoretically expected and the experimentally obtained MVM products. Further details on the model are provided in the Methods section: *Simulation model of the in-memory MVM computing engine*. Figure 3(b) depicts the normalized absolute errors (E) expressed in percentage (%) between the experimental and the simulated values of the power levels obtained at every column output, calculated using equation (1):

$$E = \frac{|P_{experimental} - P_{simulated}|}{|P_{simulated}|} \cdot 100\% \qquad (1)$$

, where the $P_{experimental}$ corresponds to the experimentally measured power at the output column for a given input vector and $P_{simulated}$ represents the respective simulated value. Each box plot in Fig. 3(b) illustrates the mean absolute error distribution for 11 indicative input vector combinations at each Xbar output column. The data show that absolute errors are consistently below 10% for all output cases, with mean values and standard deviations ranging between 2.3% and 2.41, respectively, revealing good agreement between the experimentally obtained and theoretically expected matrix values. Further details on the experimental matrix values are provided in the section 3 of the Supplementary material: *Photonic SiPROM Matrix Vector Multiplication experimental matrix values and error analysis*.

Experimental validation of a 2-bit ternary-CAM operating at 50 Gb/s

Given that the proposed architecture has been experimentally validated to perform MVMs with high accuracy, this computational capability can be directly exploited to unlock applications that rely on fast and parallel dot-product calculations. A representative example is the analog implementation of TCAMs, where the combination of unary encoding and dot-product calculation enables look-up operations [20]. However, the current layout employs thermo-optic (TO) MZMs at the input stage, which limits the achievable speed and consequently the performance of the TCAM prototype. To overcome this constraint and demonstrate high-speed search functionality, we employ off-chip high-speed MZMs to encode the search words via a wavelength-encoding scheme. During this demonstration, the TO MZMs are biased to remain optically transparent, ensuring that they do not interact with the externally modulated signals. The 4×9 crossbar layout can support up to 2-bit search words [20] for content matching across the TCAM array. The TCAM stores 9 distinct words, one per memory column, covering all possible combinations of the alphabet {0, 1, X}, where '0' and '1' denote binary values and 'X' represents the ternary 'don't-care' state.

Figure 4(a) presents the principle of operation of the optical TCAM cell. This fundamental building block enables

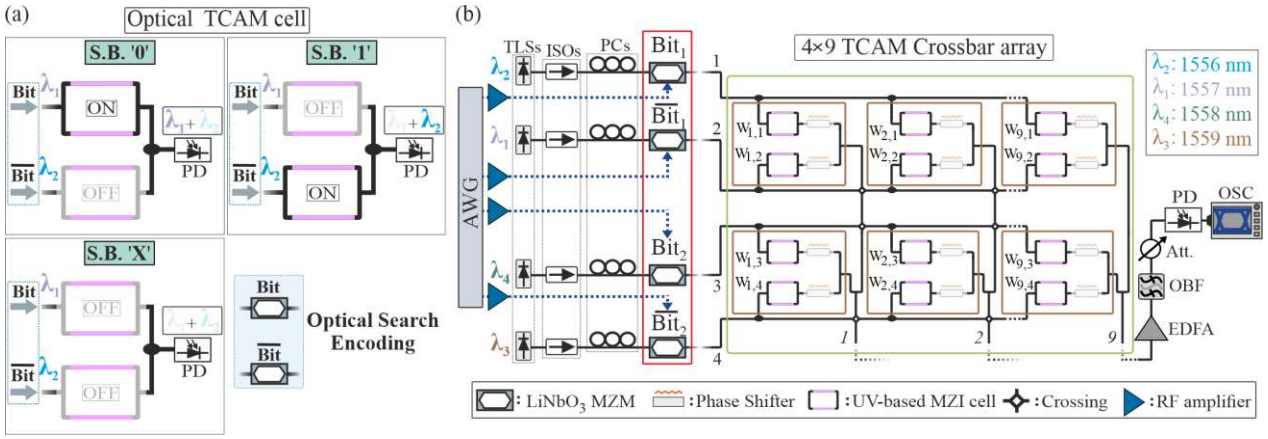


Fig. 4. a. Optical TCAM cell principle of operation for all three logical values of "0", "1", and "X". **b.** Experimental setup for the ternary content addressable memory (TCAM) comparison operation, at 50 Gb/s. The red highlighted shows the SVEU, while the green highlighted area showcases the 4×9 TCAM crossbar array, along with the respective statically configurable UV-based MZI memory cells (brown highlighted area).

non-volatile assignment of the three possible ternary stored values through incoherent power addition of different wavelengths. Two distinct wavelengths (λ_1 and λ_2) carrying the search encoded values (Bit and $\overline{\text{Bit}}$) are injected into two UV-based MZI cells. Depending on the phase difference induced during the UV exposure process, each MZI can be programmed to an "ON" or "OFF" state, as described in the Methods section: *UV-process: Setup and basic cell characterization*. The two MZI outputs are then incoherently combined at a photodiode module, yielding a dot product that indicates the matching state of the memory cell. When the TCAM cell is programmed to the logical "0" state, the corresponding interferometers are configured in the [ON, OFF] state. For an incoming search input signal of a logical "0" (Bit=0, $\overline{\text{Bit}}$=1), the TCAM cell produces a zero-output power, signifying a match between the search input and the store value. Conversely, for a logical "1" search input (Bit=1, $\overline{\text{Bit}}$=0), the programmed cell yields a non-zero output power, indicating a "mismatch". The same operating principle applies to all three stored states: the logical "1" state is implemented as [OFF, ON], while the ternary "don't care" state ("X") is defined as [OFF, OFF], which always produces a "match" regardless of the incoming search input.

Figure 4(b) depicts the experimental setup that was employed for the evaluation of the TCAM matchline operation at the rate of 50 Gb/s. Four laser sources produce four discrete CW optical carriers ($\lambda_1 - \lambda_4$), which are subsequently routed to an equal number of high-speed off-the-shelf MZMs to generate the 2-bit search words. The modulators were driven by a 4-channel arbitrary waveform generator (AWG) with non-return-to-zero (NRZ) time shifted data sequences, to generate all 2-bit search combinations (i.e. "00", "01", "10" and "11"). Each of the modulated signals are inserted into one of the four rows of the 4×9 Xbar and subsequently broadcast to all 9 columns. The output of the comparison between the incoming search values and all the stored binary/ternary values of the photonic array is monitored at each output column of the Xbar array with the wavelength responses of the signals being presented in the Supplementary section 4: *TCAM optical spectrum characterization.* All signals emerging from the Xbar outputs were subsequently amplified using an erbium-doped fiber amplifier (EDFA) and passed through a 4-nm optical bandpass filter (OBF) to suppress the amplified spontaneous emission (ASE) noise. The filtered signal is then launched into a PD where the individual product components are added incoherently, to realize the computation of the dot product between the stored and search words. The signals are finally captured by a sampling oscilloscope, for processing and signal analysis. Additional information about the experimental setup can be found in the Methods section: *Experimental setup of 2-bit T-CAM*.

Figure 5 (a) and (c) present 10-bit long time sequences of both the 2-bit encoded search words as well as the produced comparison output traces of each search sequence with the respective stored words (S.W.). The blue highlighted areas indicate the instances in which the search word corresponds to "10" (Bit_1=1, Bit_2=0). In each trace, the cross-shaped

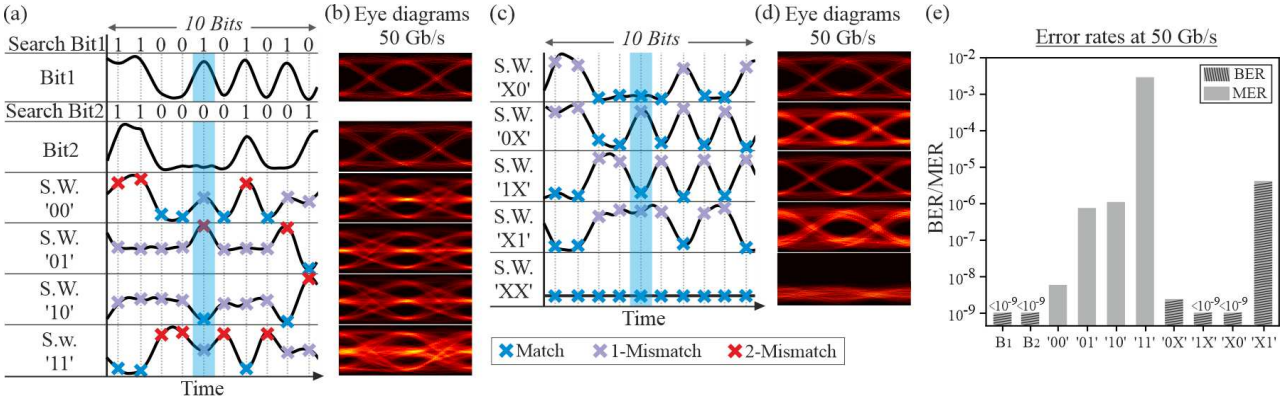


Fig. 5 Experimental results for the validation of the 2-bit comparison operation, at the data rate of 50 Gb/s. **a** Synchronized time traces (20 ps/div) of 10 optically encoded search Bit_1 and Bit_2 values, along with the binary S.W. cases of "00", "01", "10" and "11". **b** Corresponding eye diagrams at 50 Gb/s. **c** Time traces (20 ps/div) of the ternary S.W. cases of "X0", "0X", "X1" and "XX". **d** Eye diagrams for the ternary cases at 50 Gb/s. **e** Error rate values for the 50 Gb/s operation for the search values (B_1 and B_2) and all 2-bit S.Ws.

scatter points denote the different logic states of the respective "match" and "mismatch". The S.W.s "10", "X0", "1X" and "XX" produce a zero-output power value, indicating a "match" with the incoming search information, where the ternary "X" value yields a matching condition for all possible states of the incoming search words. The S.W.s of "00", "11", "0X" and "X1" result in a 1-level output power, while the case of "01" produces a 2-level output power. A 1-level "mismatch" corresponds to a one-bit difference between the search and stored values, while a 2-level "mismatch" indicates a two-bit difference between the compared contents. The eye diagrams at 50 Gb/s of each case are illustrated in Fig. 5 (b) and (d), demonstrating the high quality of the optical signals across all 2-bit comparison operations. Figure 5 (e) presents the projected bit error rate (BER) and match error rate (MER) metrics for all the search and stored word combinations, computed by employing the method described in [20]. Specifically, the MER measurements refer to the projected error rate when a single threshold is applied to the 3-level signal to differentiate between the match and no-match state, regardless of the degree of mismatch, whereas the BER metric is used to evaluate the performance of the 2-level signal cases. As can be observed, the BER values for the 2-level signals range from the order of 10^{-9} to 10^{-6}, while the 3-level signals reveal MERs varying from 10^{-9} to 10^{-3}. All the cases indicate efficient comparisons, with the degradation being mainly attributed to the opto-electronic noise induced by the high-speed components.

Methods

Photonic platform and waveguide specifications

The photonic integrated prototype was fabricated on a wafer-scale silicon photonic platform with 3.2 μm of thermally grown oxide. The photonic waveguides used in the current layout are strip waveguides with a width of 1 μm and a thickness of 400 nm and are designed for a wavelength of 1550 nm and Transverse Electric (TE) polarization. The waveguide core consists of nitrogen-rich SiNx with a refractive index of 1.9, deposited via plasma-enhanced chemical vapor deposition (PECVD) process, while the waveguide is clad with 2 μm of PECVD SiO2.

UV-process: Setup and basic cell characterization

The UV-exposure post-process that is used allows for a change in the attenuation state of the MZI, through phase change. A 244 nm frequency-doubled Argon-Ion laser (Coherent Innova) has been utilized to alter the refractive index of the n-rich waveguide core. The laser is expanded and then split using a beamsplitter, with the two beams being recombined to form a 14 μm spot on the surface of the PIC. The photonic structure is located on an Aerotech ABL9000 4-axis translation stage controlled by a software-based machine controller (Aerotech 3200) with programs written in Aerotech Basic, allowing precise control of the laser spot relative to the PIC. In order to tune the output power of each symmetric MZI cell, a fluence of 1.5 kJ/cm^2 was chosen based on the results in Figure 1 (a). Before exposure of the actual MZI memory cells, the process was first tested on set of MZI memory cell test structures. In one of the arms of the test structure MZI, a raster pattern is exposed, with each line of the raster pattern targeting a 14 μm section of the arm according to the 14 μm spot size. Each exposure reduces the refractive index of the n-rich core and induces a phase difference between the arms of the interferometer. The number of lines in each raster pattern varied from 0 to 24 corresponding to a total exposed length of 0 to 336 μm. Each MZI was measured before and after the exposure process, monitoring the output power difference for every step. The measurements indicate that a maximum extinction ratio (ER) of 15 dB is observed for 20 exposures, with a total exposed length of 280 μm and with an exposure duration of 7 seconds. Based on this, the crossbar memory nodes receive 20 exposures in order to render the interferometer in the "OFF" state defining the "0" bit value, where the signal receives the maximum attenuation and zero exposures for the "ON" state where the incoming signal will receive no attenuation when propagating through the interferometer, creating a "1" bit value. A schematic illustration of the experimental setup employed for UV-exposure and the relationship between achieved ER and number

of UV-exposure are provided in Supplementary Section: *Post fabrication UV-exposure process of MZI cells*.

Experimental setup of the 36-bit SiPROM

A tunable laser source (IDPhotonics tunable laser source) is used to generate a continuous wave signal at λ=1575 nm. The signal is injected into every respective crossbar row (Row_1-Row_4) of the SiPROM array with the use of an angled fiber array (FA) unit, via TE grating couplers, in order to perform the read memory functionality. The respective memory content of each cell is received via a grating coupler positioned at the switching port of each inner MZI and is measured using a PD. The high-power values indicate a logic "1" case, while the low-power values denote a logic "0" case. The normalization of each memory cell is performed offline.

Experimental setup of the Compute-In-memory Circuit

A tunable laser source (IDPhotonics tunable laser source) is employed in order to generate a continuous wave at λ=1575 nm. The signal gets inserted into the photonic integrated circuit, through an angled FA unit and a TE grating coupler. The signal is then split into 4 respective components, via a 1:4 MMI-based splitter, and each one is forwarded to a TO MZM of the input vector encoding state. The modulators are statically driven by a multi-channel direct current (DC) voltage control plane (Rigol DP831 LXI), altering the switching state of the interferometer. In this way the optical signals are modulated to imprint the logical states of "0"s and "1"s, enabling the encoding of all possible 4-bit input vector combinations. The modulated signals are then propagated through each column's UV-based cell of the array, performing a multiplication operation between the input and matrix values of each row/column. The TO PSs of each column are also driven by DC voltage source, ensuring constructive interference at the output of each column of the Xbar. At the output of each column, the produced dot products are captured and monitored by a photodiode.

Simulation model of the Compute-In-memory Circuit

The simulation model that has been developed is based on the equation (2) that provides the field amplitude of an output column of the Xbar architecture, with detailed explanation of the equation modeling found in [18].

$$E_{out} = \frac{E_{in}}{4} \cdot \sum_{i=1}^{4} (x \cdot |w| \cdot e^{j\varphi}) \quad (2)$$

The x represents the field amplitude that is received based on the input modulator combinations, while the $|w|$ refers to the weight values of the cells and φ to the corresponding phase that the light acquires at each propagating path. The total field amplitude (E_{out}) is then defined by the summation of all 4 input and weight multiplication combinations for a given output column of the crossbar array. Based on the experimental measurements that have been performed, the input stage MZMs showcase an ER of 14 dB. According to this ER, the input x values can be normalized to a range, where the x = 1 refers to the logic "1" case and x = 0.2 to the "0" case. The $|w|$ values used in the model are the same as those obtained from the respective experimental measurements, more details about the experimental values are provided in the Supplementary section 3 : *Photonic SiPROM Matrix Vector Multiplication experimental matrix values and error analysis.* Following this model, 11 out of 16 different input vector cases have been experimentally evaluated for each output column. To extract the phase differences (φ) acquired by the coherent signals across an input/output crossbar path, a heuristic least mean squares algorithm (LMS) was employed to estimate the corresponding phase values. In this way, the expected output values can be calculated using equation (2) through this experimentally driven model and directly compared with the experimental results to estimate the error differences, as shown in Fig. 3b.

Experimental setup of 2-bit WDM T-CAM

Four different laser sources (IDPhotonics tunable laser source) are utilised in order to generate four CWs at λ_1 = 1556, λ_2 = 1557, λ_3 = 1558 and λ_4 = 1559 nm. The corresponding wavelength cases have been selected in order to comply with the optimum response of the GCs of the integrated structure offering the minimum insertion losses (ILs), as well as to ensure the best amplification conditions of the EDFA unit. The 4 lambdas are fed to 40 GHz commercially available

$LiNbO_3$ MZMs. An arbitrary waveform generator (AWG Keysight M8194A) is employed to generate 4 distinct NRZ pseudo random data sequences (PRBS^9) at 50 Gb/s, corresponding to Bit_1, Bit_2 and their complementary values $\overline{Bit}_1$, $\overline{Bit}_2$ of the search signals. The time sequences of $Bit_1/\overline{Bit}_1$ and $Bit_2/\overline{Bit}_2$ have a relative time shift in order to ensure all the possible 2-bit search input combinations. The electrical signals are then amplified by a series of RF amplifiers (SHF L806A) to reach 5.5 Vpp each, enabling properly driving conditions for every MZM. The modulated signals are then propagated into the photonic Xbar structure, via TE grating couplers, using an angled FA unit. The search incoming signals are forwarded to all 9 crossbar columns for the 2-bit comparison operation between the search and stored contents. The integrated TO phase shifters (PSs) on the chip are not used in this implementation, as it relies on WDM and incoherent power addition. At each output column of the Xbar, the multi-wavelength signals are first amplified by an EDFA to compensate for the photonic device losses and then filtered by a 4 nm OBPF that encompasses the four wavelength channels, effectively isolating the additional noise introduced by the EDFA. The different signals perform an incoherent power addition in a 70 GHz photodiode module (Finisar XPDV3120), producing a dot product operation that defines the matching condition of the comparison operation. Finally, the output signals are captured by a 75 GHz sampling oscilloscope (OSC Keysight DCA-X N1000A) and are then processed by an additional digital signal processing (DSP) stack. Specifically, a 4th order Butterworth software-based filter is implemented through the sampling oscilloscope module with an adjusted 3 dB bandwidth at 50 GHz, mitigating the additional noise bandwidth response of the channel. Concurrently, a linear equalizer is applied to minimize intersymbol interference arising from the bandwidth limitations of the transmitting channel. Pattern correlation and the calculation of MER/SER values for the experimental signals are performed offline.

Data Availability

The data cases that support the experimental results and findings of the study, as well as all the validation methods and procedures for the additional analysis performed are provided in Supplementary information files. In case of need for additional information and material of the study the authors can provide them upon request.

Code Availability

All software and source code files, as well as documentation and instructions for the analysis and validation of the experimental and theoretical findings of the study are provided in Supplementary information files and documents. Any additional specifications regarding the methods that have been followed can be provided upon request by the authors.

References

[1] Feldmann, J., Youngblood, N., Karpov, M. *et al.* Parallel convolutional processing using an integrated photonic tensor core. *Nature* **589**, 52–58 (2021). https://doi.org/10.1038/s41586-020-03070-1

[2] Ríos, C. *et al.* In-memory computing on a photonic platform. *Sci. Adv.* **5**, eaau5759 (2019). https://doi.org/10.1126/sciadv.aau5759

[3] Lian, C., Vagionas, C., Alexoudi, T., Pleros, N., Youngblood, N. & Ríos, C. Photonic (computational) memories: tunable nanophotonics for data storage and computing. *Nanophotonics* **11**, 3823–3854 (2022). https://doi.org/10.1515/nanoph-2022-0089

[4] Zhou, W., Dong, B., Farmakidis, N. *et al.* In-memory photonic dot-product engine with electrically programmable weight banks. *Nat Commun* **14**, 2887 (2023). https://doi.org/10.1038/s41467-023-38473-x

[5] Feldmann, J., Stegmaier, M., Gruhler, N. *et al.* Calculating with light using a chip-scale all-optical abacus. *Nat Commun* **8**, 1256 (2017). https://doi.org/10.1038/s41467-017-01506-3

[6] Jung, Y. J., Park, N., Jhon, Y. M. & Lee, S. Design of all-optical read-only memory. *Appl. Opt.* **48**, G21–G27 (2009). https://doi.org/10.1364/AO.48.000G21

[7] Mukherjee, K. All-optical read-only memory with frequency-encoded addressing technique. *Optik* **122**, 1437–1440 (2011). https://doi.org/10.1016/j.ijleo.2010.09.024

[8] Sludds, A. *et al.* Delocalized photonic deep learning on the internet's edge. *Science* **378**, 270–276 (2022). https://doi.org/10.1126/science.abq8271

[9] Brunner, D. *et al.* Roadmap on Neuromorphic Photonics. arXiv:2501.07917v2 https://doi.org/10.48550/arXiv.2501.07917

[10] Alexoudi, T., Kanellos, G.T. & Pleros, N. Optical RAM and integrated optical memories: a survey. *Light Sci Appl* **9**, 91 (2020). https://doi.org/10.1038/s41377-020-0325-9

[11] Pappas, C., Moschos, T., Alexoudi, T., Vagionas, C. & Pleros, N. Caching with light: A 16-bit capacity optical cache memory prototype. *IEEE J. Sel. Top. Quantum Electron.* **29**, 1–11 (2023). https://doi.org/10.1109/JSTQE.2023.3247032

[12] Wuttig, M., Bhaskaran, H. & Taubner, T. Phase-change materials for non-volatile photonic applications. *Nature Photon* **11**, 465–476 (2017). https://doi.org/10.1038/nphoton.2017.126

[13] Tossoun, B., Liang, D., Cheung, S. *et al.* High-speed and energy-efficient non-volatile silicon photonic memory based on heterogeneously integrated memresonator. *Nat Commun* **15**, 551 (2024). https://doi.org/10.1038/s41467-024-44773-7

[14] Pintus, P., Dumont, M., Shah, V. *et al.* Integrated non-reciprocal magneto-optics with ultra-high endurance for photonic in-memory computing. *Nat. Photon.* **19**, 54–62 (2025). https://doi.org/10.1038/s41566-024-01549-1

[15] Geler-Kremer, J., Eltes, F., Stark, P. *et al.* A ferroelectric multilevel non-volatile photonic phase shifter. *Nat. Photon.* **16**, 491–497 (2022). https://doi.org/10.1038/s41566-022-01003-0

[16] Xia, J., Wang, T., Wang, Z., Gong, J., Dong, Y., Yang, R. & Miao, X. Seven-bit nonvolatile electrically programmable photonics based on phase-change materials for image recognition. *ACS Photonics* **11**, 723–730 (2024). https://doi.org/10.1021/acsphotonics.3c01598

[17] De Paoli, G., Jantzen, S. L., Dominguez Bucio, T., Skandalos, I., Holmes, C., Smith, P. G. R., Milosevic, M. M. & Gardes, F. Y. Laser trimming of the operating wavelength of silicon nitride racetrack resonators. *Photon. Res.* **8**, 677–683 (2020). https://doi.org/10.1364/PRJ.382529

[18] Giamougiannis, G. *et al.* A coherent photonic crossbar for scalable universal linear optics. *J. Lightwave Technol.* **41**, 2425–2442 (2023). https://doi.org/10.1109/JLT.2023.3234689

[19] Moralis-Pegios, M., Giamougiannis, G., Tsakyridis, A. et al. Perfect linear optics using silicon photonics. Nat Commun 15, 5468 (2024). https://doi.org/10.1038/s41467-024-49768-y

[20] Moschos, T., Pappas, C., Kovaios, S. *et al.* Nonlinear optical vector processing using linear silicon photonic circuits for 50 Gb/s memory and string similarity functions. *Nat Commun* **16**, 11416 (2025). https://doi.org/10.1038/s41467-025-66286-7

Acknowledgements

This work has been supported by the European Commission (EC) through the projects OCTAPUS (101070009) and AMBROSIA (101093166), as well as EPSRC under grants EP/M013294/1 and EP/W524621/1.

Author contributions

S. Simos, T. Moschos, M. Moralis-Pegios and S. Kovaios, conducted the characterization of the photonic Xbar device and all experimental system-scale implementations presented in this manuscript. S. Simos, T. Moschos, M.

Moralis-Pegios, A. Tsakyridis, and N. Pleros performed all post-processing of the data as well as the additional simulation analysis and modelling. I. Johnson, T. Rangarajan, T. D. Bucio, J. Gates, and F. Gardes manufactured the silicon photonic device and characterized the performance of a single MZM-based memory cell for different UV exposure fluence values. F. Gardes and N. Pleros conceived the original idea and all authors contributed to the discussion of the final circuit layouts and the different applications, the organization of the results and manuscript, and the writing process.

Competing interest

The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SUPPL1.doc