Science Bulletin 68 (2023) 2500–2502

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

Research Highlight

In-memory computing based on photonic-electronic hybrid phase-change cells

Bai-Qian Wang [a], Xian-Bin Li [a,*], Hong-Bo Sun [b]

[a] State Key Laboratory of Integrated Optoelectronics, College of Electronic Science and Engineering, Jilin University, Changchun 130012, China
[b] State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instrument, Tsinghua University, Beijing 100084, China

In-memory computing based on emerging non-volatile memory arrays holds a great promise to cope with the drastically increased demand for data processing. Phase-change material (PCM) is a leading candidate for in-memory computing, which utilizes the amorphous-crystalline phase transition and the associated changes in electrical resistance or optical transmission for data encoding [1]. The flagship PCM is $Ge_2Sb_2Te_5$ (GST). In comparison with PCM electronic devices, the data throughput can be largely boosted by leveraging parallel computing in the photonic waveguide devices based on the wavelength division multiplexing scheme. Recently, a GST-based all-optical photonic crossbar array was successfully designed and fabricated to realize the matrix–vector multiplication [2] that takes a heavy workload in artificial intelligence computation. However, optical programming and computing were performed in two separated steps due to sharing the same photonic integrated circuits (PICs). To avoid extra time waste in optical programming, electrically reprogrammable GST cells [3] as a hybrid optoelectronic interface could enable decoupled electrical programming via electronic integrated circuits (EICs) and optical computing in the PICs, and which can be performed simultaneously. Most recently publishing in *Nature Communications*, Zhou et al. [4] achieved a milestone by developing a dot-product engine based on a high-performance photonic-electronic GST cell array. A hybrid computing system was developed using non-volatile waveguide memory elements, which may speed up photonic computing landscape transformation to non-von Neumann architecture [5].

Waveguide microheater with partial ion implantation region in silicon acts as a current conductive path for local Joule heating of the deposited GST thin film [3]. By applying appropriate pulse profiles, structural transitions between crystalline and amorphous phases can be initialized based on annealing and melt-quenching processes [6]. Fig. 1a shows a schematic of an improved design in electrothermal switching of the GST cell, below which the width of doping region is tapered down. Thus, a temperature gradient can be generated with its highest temperature overlapping with the waveguide core and GST cell. Fig. 1b shows optical micrographs of the fabricated devices on a silicon-on-insulator (SOI) chip. Contact pads of the GST devices are wire bonded to an external printed circuit board. Input light power (X_i) is partially absorbed and attenuated by a GST cell with an encoded weight (W_i) depending on the amorphous-to-crystalline ratio by electrical programming with an encoding precision higher than 4 bits. This light transmission process mathematically represents an in-memory multiplication operation ($W_i \cdot X_i$). Output from multiple devices can be incoherently combined together to realize a multiply-accumulate (MAC) operation ($\Sigma W_i \cdot X_i$), which is essential for neuromorphic computing in signal processing and image recognition. For the developed GST cell, it achieves one of the lowest energy consumptions per unit modulation depth (1.7 nJ/dB). Operating time required for the Write and Erase processes are 232 and 556 ns, respectively.

Fig. 1c shows one application using in-memory computing in brightness scaling of an image. By leveraging large switching contrasts of (64%, 128%), the measured standard deviation (σ) is as low as 0.019 by comparing processed and calculated images. However, change of hue and noises are observed for the processed image using low switching contrasts of (4%, 8%). Standard deviation versus switching contrast was systematically investigated. As shown in Fig. 1d, by enlarging the switching contrast, contrast-to-noise ratio (CNR) can be enhanced with significantly reduced standard deviation in computing errors. For high-precision processing, a large switching contrast is preferred, which is an advantage of using electrothermal switching of GST compared with those based on the optical switching with relatively small contrasts [7,8]. Convolutional operation was also performed in edge extraction of MNIST database images. The standard deviation is less than 0.12. These processed images were fed into convolutional neural networks with inferencing accuracies higher than 86%. In addition to a high computing accuracy, large switching contrasts also promise reliable driving of the artificial neurons crossing their thresholds, which are crucial for the photonic deep neural networks [9].

This work exhibits merits of optoelectronic hybrid integration for the in-memory photonic computing. Compute density, efficiency, and energy per MAC operation are predicted to be 7.3 TOPS/mm^2, 10 TOPS/W, 0.2 pJ/MAC, respectively, for a medium sized 16 × 16 waveguide crossbar array. The present in-memory photonic–electronic hybrid platforms clearly demonstrate attractive advantages like scalable and flexible programming, parallel

* Corresponding author.
E-mail address: lixianbin@jlu.edu.cn (X.-B. Li).

https://doi.org/10.1016/j.scib.2023.08.058

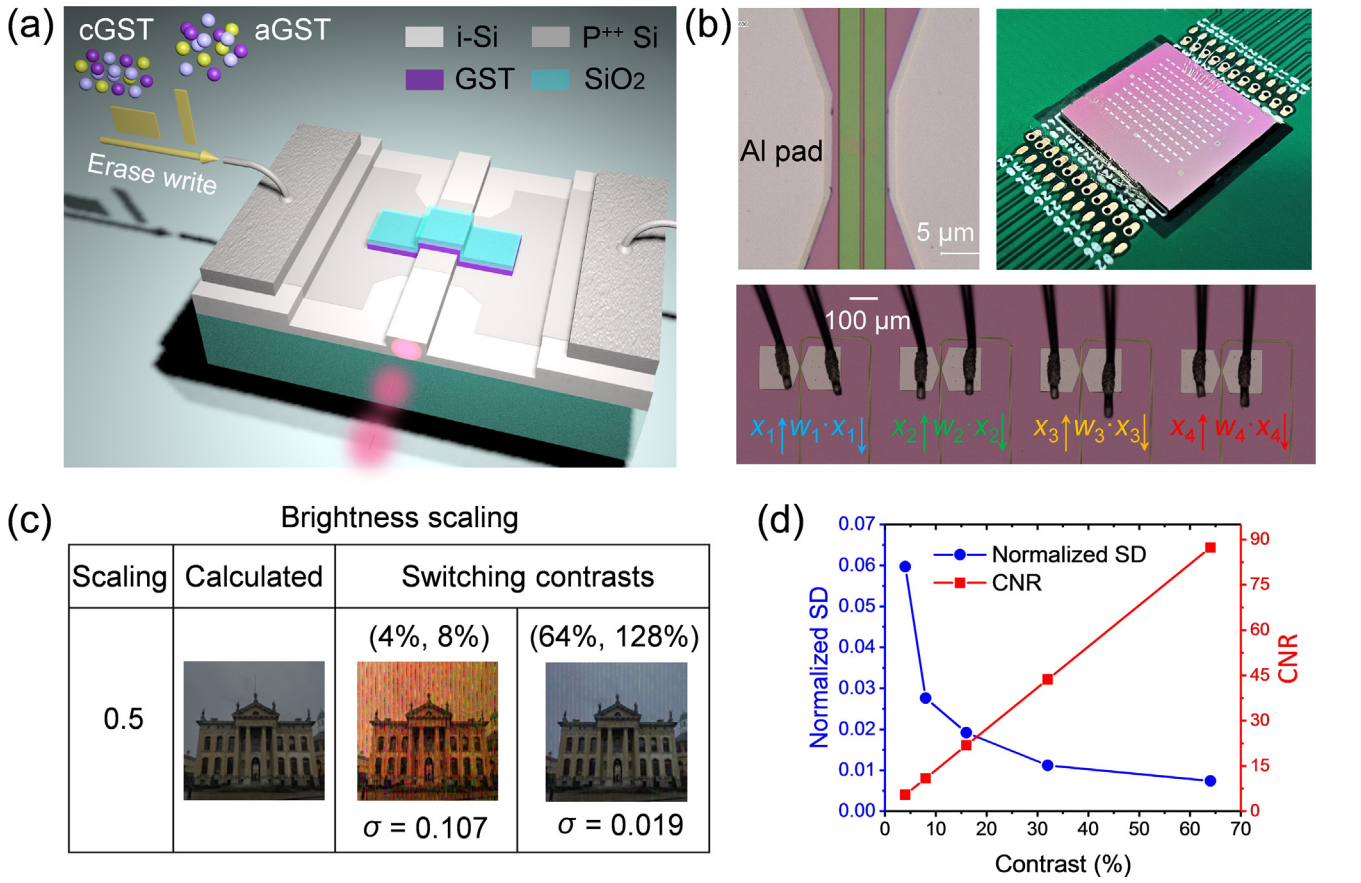

Fig. 1. (Color online) Electrically reconfigurable in-memory photonic dot-product engine. (a) Schematic of a GST memory element for variable optical attenuation. (b) Optical micrographs of a fabricated silicon photonic chip consisting of GST devices. (c) Image brightness scaling application. (d) Computational error in standard deviation (SD) and contrast-to-noise ratio (CNR) versus switching contrast of the GST device. Reproduced with permission from Ref. [4], Copyright © 2023 Springer Nature.

optical signal processing, and CMOS-compatible wafer-scale fabrication. Undoubtedly, this study by Zhou et al. [4] has made an important step to accelerate practical applications of PCM based optical computing technology for the big-data era.

In future, reducing insertion losses in devices, system scaling up using crossbar array networks, design or searching of high-performance PCMs [10] could be further studied. Besides, device innovation is expected by focusing on low-energy switching with an operating voltage compatible with the CMOS architecture by employing graphene heaters [11]. Improvement in storage capability is anticipated from the current record 5-bit operation reported in electrically controlled PCM devices [12]. Survey on PCMs with ultrafast nucleation mechanism may be conducted to speed up device programming. Cycling endurance and programming consistency are another two issues, which may be improved by exploiting phase-change heterostructures [13]. Considering wafer-scale manufacturing, design and fabrication of the computational PCM memory array should be compatible with the standardized silicon photonic CMOS line with a robust back-end-of-line process for integrating PCM functional layers on the PICs. Photonic-electronic codesigned systems consisting of high-speed modulators, PCM crossbar arrays, and photodetectors potentially pave the way for the next-generation photonic deep neural networks operating at GHz clock rate. In addition to inferencing presented in Ref. [4], *in-situ* training of weight banks is also expected for supervised and unsupervised learning in AI tasks. Many more learning algorithms, e.g., associative learning, reinforcement learning, can be implemented based on the photonic-electronic in-memory computing architecture.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the National Key R&D Program of China (2022ZD0117600) and the National Natural Science Foundation of China (12274172 and 12274180).

References

[1] Wuttig M, Bhaskaran H, Taubner T. Phase-change materials for non-volatile photonic applications. Nat Photonics 2017;11:465–76.
[2] Feldmann J, Youngblood N, Karpov M, et al. Parallel convolutional processing using an integrated photonic tensor core. Nature 2021;589:52–8.

[3] Zhang H, Zhou L, Lu L, et al. Miniature multilevel optical memristive switch using phase change material. ACS Photonics 2019;6:2205–12.
[4] Zhou W, Dong B, Farmakidis N, et al. In-memory photonic dot-product engine with electrically programmable weight banks. Nat Commun 2023;14:2887.
[5] Shastri BJ, Tait AN, Ferreira de Lima T, et al. Photonics for artificial intelligence and neuromorphic computing. Nat Photonics 2021;15:102–14.
[6] Zhou W, Farmakidis N, Feldmann J, et al. Phase-change materials for energy-efficient photonic memory and computing. MRS Bull 2022;47:502–10.
[7] Li X, Youngblood N, Cheng Z, et al. Experimental investigation of silicon and silicon nitride platforms for phase-change photonic in-memory computing. Optica 2020;7:218–25.
[8] Zhou W, Farmakidis N, Li X, et al. Artificial biphasic synapses based on nonvolatile phase-change photonic memory cells. Phys Status Solidi RRL 2022;16:2100487.
[9] Ashtiani F, Geers AJ, Aflatouni F. An on-chip photonic deep neural network for image classification. Nature 2022;606:501–6.
[10] Liu YT, Li XB, Zheng H, et al. High-throughput screening for phase-change memory materials. Adv Funct Mater 2021;31:2009803.
[11] Fang Z, Chen R, Zheng J, et al. Ultra-low-energy programmable non-volatile silicon photonics based on phase-change materials with graphene heaters. Nat Nanotechnol 2022;17:842–8.
[12] Chen R, Fang Z, Perez C, et al. Non-volatile electrically programmable integrated photonics with a 5-bit operation. Nat Commun 2023;14:3465.
[13] Ding K, Wang J, Zhou Y, et al. Phase-change heterostructure enables ultralow noise and drift for memory operation. Science 2019;366:210–5.

Bai-Qian Wang received his B.S. (2021) degree from Jilin University. He is now a doctoral candidate at the State Key Laboratory of Integrated Optoelectronics, College of Electronic Science and Engineering, Jilin University. His research mainly focuses on the high-throughput screening and large-scale molecular dynamics studies of phase-change memory materials, based on machine-learning methods and first-principles calculations.

Xian-Bin Li received his B.S. (2005), M.S. (2007), and Ph.D. (2010) degrees in Microelectronics from Jilin University. He is currently a professor and the group leader of the Laboratory of Computational Semiconductor Physics in the State Key Laboratory of Integrated Optoelectronics, College of Electronic Science and Engineering, Jilin University. He also serves as the head of the Department of Microelectronics Science and Engineering. His research mainly focuses on the key problems in microelectronics and optoelectronics with first-principles method, including phase-change memory physics, semiconductor defect physics, and light-matter interaction.