Artificial Synapse with Mnemonic Functionality using GSST-based Photonic Integrated Memory

Mario Miscuglio, Jiawei Meng, Armin Mehrabian, Volker J. Sorger
Department of Electrical and Computer Engineering
George Washington University
Washington DC 20052, USA
mmiscuglio@gwu.edu, sorger@gwu.edu

Omer Yesiliurt, Ludmila J. Prokopeva, Alexander V. Kildishev
Birck Nanotechnology Center
School of ECE, Purdue University
West Lafayette, IN 47907, USA,
kildishev@purdue.edu

Yifei Zhang, Juejun Hu
Department of Materials Science & Engineering
Massachusetts Institute of Technology
Cambridge, MA, USA
hujuejun@mit.edu

***Abstract*—Here we present a multi-level discrete-state non-volatile photonic memory based on an ultra-compact (<4μm) hybrid phase change material GSST-silicon Mach Zehnder modulator, with low insertion losses (3dB), to serve as node in a photonic neural network. Emulating an opportunely trained 100 × 100 fully connected multilayered perceptron neural network with this weighting functionality embedded as photonic memory, shows up to 92% inference accuracy and robustness towards noise when performing predictions of unseen data.**

Keywords—GSST, phase change memories, photonic memories.

I. Introduction

Computing AI-systems and machine-learning (ML) tasks, while transferring and storing data exclusively in the optical domain, is highly desirable because of the inherently large bandwidth, low residual crosstalk, and short-delay of optical information transfer [1]. However, the functionality of memory for storing the trained weights is not straightforwardly achieved in optics [2], [3], or at least in its non-volatile implementation, and therefore requires additional circuitry and components (i.e., DAC, memory) and related consumption of static power, sinking the overall benefits (energy efficiency and speed) of photonics. The non-volatile retention of information in integrated photonics can be provided by the light-matter interaction in phase change memory (PCM) [4]–[8]. Here, we leverage on a recently engineered class of optical PCMs, based on $Ge_2Sb_2Se_4Te_1$ (GSST) alloy [9], whose amorphous state is not characterized by high absorption coefficient, like the commonly used Ge–Sb–Te, and upon phase change, its refractive index is still subjected to unitary modulation. The optimized alloy combines broadband transparency (1–18.5 μm), large optical contrast ($\Delta n = 2.0$), and significantly improved glass forming ability. Thence, we design a low loss non-volatile multilevel photonic memory, using a balanced GSST-based Mach Zehnder Modulator and develop a numerical framework for optimizing the heaters configuration and evaluating the temporal switching response of such GSST-based photonic memory. Ultimately, the novel photonic memories are used as artificial synapses, of an all-optical (AO) photonic Neural Network (NN), which stores and process data in memory. The off-chip trained NN can effortlessly stores and perform dot-product functionality in the optical domain and classify handwritten digits with an accuracy of 92%, considering 3-bit (low precision) weights and network noise.

II. Results

The combination of phase change material (PCM) with silicon photonics platforms can lead to fast non-volatile, reconfigurable memories for optical communications applications. The most used PCM for these devices is GST ($Ge_2Sb_2Te_5$), which is characterized by large refractive index contrast between the amorphous and crystalline state, but also significant propagation losses which can hinder the implementation of large networks which does not need cumbersome O-to-E-to-O conversions [7]. This, however, requires a wisely engineered material process, e.g., interfacial PCM ($GeTe/Sb_2Te_3$) [7] and optimized alloys [9], [10]. Contrary to regularly used GST (**Fig. 1 (a)**), $Ge_2Sb_2Se_4Te_1$ [GSST] exhibits a 3 orders of magnitude lower absorption coefficient while preserving a large Δn of 2.1 to 1.7 across the near- to mid-IR bands. Additionally, the GSST film is characterized by a high index ratio, $\Delta n/\kappa = 5$, due to its low index contrast and relatively small κ, even without any metal-insulator transition. Interferometric modulators based on hybrid GSST-Silicon waveguide operating at 1550 nm would allow to fabricate photonic memory devices with remarkably large extinction ratio and contained insertion losses. For this reason, we decide to model the synapses of our photonic neural network as GSST-based Mach Zehnder Interferometer (MZI). Refractive index of GSST used in our model are taken from [9], however a full multivariate time-domain model can be built using the generalized dispersive material model [11].

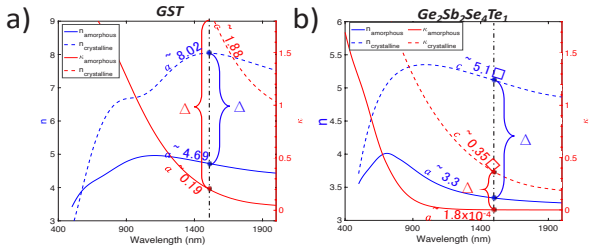

Fig. 1. Experimentally obtained (ellipsometry) optical properties of phase change material (PCM) films: GST (**a**), and Ge2Sb2Se4Te1 (**b**). Real (n, left y-axis) and imaginary (κ, right y-axis) parts of the refractive indices of the amorphous (solid line) and crystalline alloys. (dashed line). For our study, we consider GSS4T1.

In this work, the synaptic weights are set by selectively 'writing' portions of the GSST deposited on the waveguides, by heat induced laser irradiation, or local electrothermal heating, which promotes crystallization and alter the effective refractive index of the hybrid waveguide. In this way, the synapse, which stores the quantized weight, can modulate - weight- the intensity of an input signal accordingly. In details, the active part of the modulator consists of 30 nm of GSST deposited atop of a planarized waveguide and local heaters which induce phase transition through Joule heating. Firstly, we design a balanced passive Mach Zehnder modulator (MZM) configuration (**Fig. 2 (a)**), in which the GSST material is deposited in the amorphous condition (aGSST) on both of

This work is supported by the Presidential Early Career Award for Scientist and Engineers (PECASE) nominated by AFOSR.

Submitted On: September 00, 2000
Accepted On: September 00, 2000

https://doi.org/10.47037/2020.ACES.J.351192

the arms of the modulator. The MZI is purposely unbalanced by thermally writing a portion of the GSST film deposited in the "programmable arm" of the MZI. For a TM mode (**Fig. 2 (b)**), for instance, the length of the active part of the modulator is just 3.8 μm short for achieving a π phase shift, when the entire film on the 'recordable' branch has changed to its crystalline phase. To our knowledge, the device is one order of magnitude smaller than one of the most compact MZM ever reported [12], with positive effects on the electrical capacitance improving both response time and power consumption. The lateral section of the written part of the material corresponds to a "quantized weight", and assuming a stabile writing resolution of about 500 nm [8], same achieved by optical writing, which would avoid crosstalk among multiple states, the total amount of available discrete resolution is given by 8 distinct states (3-bit). This condition on the resolution can be relaxed by extending the device length by multiple of L_π -minimum modulator length- (**Figs. 2 (d-e)**) or by physically separating portions of the film. Additionally, this is a reversible process, which allows to update the weights after many execution times. Interestingly, this solution is not hindered by insertion losses, which are negligible due to the rather low absorption coefficient of GSST at 1550 nm, and the total losses (~3dB) are mainly caused by the balancing mechanism (in this first analysis straightforwardly obtained achieved by placing a gold contact on the balancing arm). The GSST needs to be locally written according to the weights obtained during the training phase. For performing this function, we consider using electro-thermal local switching using heaters. 3D time-dependent multi-physics simulations, including model for heat transfer in solids coupled with the electric currents model, have been carried out. Preliminary thermal characterization conducted by our group shows that the conductivity of GSST is 0.17±0.02 W/m/K for amorphous phase and 0.43±0.04 W/m/K for crystalline phase, while the heat capacity in amorphous and crystalline phase for GSST film are 1.45±0.05 MJ/m^3/K and 1.85±0.05 MJ/m^3/K, respectively. The heating element (**Fig. 2 (c)**) considered is tungsten, which, when properly biased, dissipate energy in the form of Joule heat in the surrounding media. Two tungsten resistive heaters placed directly in contact with the GSST film, 100 nm away from the waveguide, provides heat to the film locally (lowering the switching threshold), and storing heat for successive pulses [13]. The temperature at which the GSST reaches the amorphous state is considered around 900 K, whereas for inducing re-crystallization, the GSST needs to be heated above the crystallization temperature (~523 K) but below the melting point, for a critical amount of time, therefore multiple pulses are needed [9]. This configuration is affected by limited insertion losses due to the non-plasmonic properties of tungsten (additional ~0.1dB/μm for a propagating TM mode) and reversible amorphization by applying a 15 V pulse of 1 μs and crystallization using 20 pulses of 1 μs for a low threshold voltage (5 V). A resistive heater optimized for efficient and low-loss switching, can be made in doped silicon, ITO or in silicide, currently used in p-n modulator, positioned next to the waveguide [15]. The NN architecture that exploits the proposed non-volatile weighed addition can be emulated on an open source ML framework, i.e., Tensorflow [14]. As preliminary study, we estimated the functionality of the proposed synapse as main unit of the NN, by emulating its behavior in a 100 neurons 3-layer fully connected NN emulated in the Google Tensorflow and, as an initial example, for the MNIST data set. Nonlinear activation functions (here considered as electro-optic[15]) are placed between two consecutive layers on each input connection. The network is trained both without and with noise of the weights and NLAF. Our hypothesis, confirmed in a recent publication [16] and from preliminary studies on the network is that, when we allow for a certain amount of noise during the training, the model during the inference stage becomes more robust; the effect of adding a noise equivalent to 0.01% of the maximum signal swing at the output of neurons significantly improves inference up to 92%.

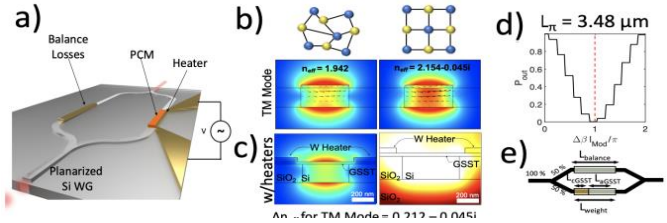

Fig. 2. Photonic memory (**a**). Schematic of an electro-optic modulator based on a balanced Mach Zehnder interferometer (MZI), such as used to program weights (dot-products) of a photonic NN (**b**). Fundamental transverse magnetic mode profiles (normalized electric field) of the GSST-Silicon hybrid waveguide at 1550 nm for amorphous and crystalline GSST show a strong index (real-part) difference ~ 0.2, while incurring a relatively low loss $\Delta\kappa = 0.4$. Black arrows represent the direction and intensity of the magnetic field (Hx,Hy) (**c**). Mode and heat profile produced by Joule heating of a hybrid waveguide with tungsten heating element, considering the aGSST (**d**). Normalized output power considering compensated losses for the MZI. The weighting is quantized since the resolution of the phase changing process is 500nm for altering the active arm of the MZI (**e**).

III. Conclusions and Outlook

In summary, we have investigated through numerical simulations a low losses programmable Mach Zehnder modulator, based on a hybrid GSST-silicon waveguide, with a $L\pi < 4\mu m$. The device showed a coherent quantized response, which is function of the portion of the phase change material that has been written by means of electrothermal switching (at sub MHz speed). We tested the quantized transfer function of the reporgrammable photonic synapse on a standardized neural network training set and we showed that the neural network reaches very high level of accuracy (>91%) in the inference phase and it is sufficiently robust to network noise.

IV. Acknowledgment

V.S. is supported by the Presidential Early Career Award for Scientist and Engineers (PECASE) nominated by the Department of Defense (AFOSR).

References

[1] D. A. B. Miller, "Optical interconnects to electronic chips," Appl. Opt., AO, vol. 49, no. 25, pp. F59-F70, Sep. 2010. doi: 10.1364/AO.49.000F59.

[2] G. Heinze, C. Hubrich, and T. Halfmann, "Stopped light and image storage by electromagnetically induced transparency up to the regime of one minute," Phys. Rev. Lett., vol. 111, no. 3, p. 033601, July 2013. doi: 10.1103/PhysRevLett.111.033601.

[3] L. Ma, O. Slattery, and X. Tang, "Optical quantum memory based on electromagnetically induced transparency," J. Opt., vol. 19, no. 4, Apr. 2017. doi: 10.1088/2040-8986/19/4/043001.

[4] M. Stegmaier, C. Ríos, H. Bhaskaran, C. D. Wright, and W. H. P.

Pernice, "Nonvolatile all-optical 1 × 2 switch for chipscale photonic networks," Advanced Optical Materials, vol. 5, no. 1, p. 1600346, Jan. 2017. doi: 10.1002/adom.201600346.

[5] P. Xu, J. Zheng, J. K. Doylend, and A. Majumdar, "Low-loss and broadband nonvolatile phase-change directional coupler switches," ACS Photonics, vol. 6, no. 2, pp. 553-557, Feb. 2019. doi: 10.1021/acsphotonics.8b01628.

[6] L. Waldecker, et al., "Time-domain separation of optical properties from structural transitions in resonantly bonded materials," Nature Materials, vol. 14, no. 10, pp. 991-995, Oct. 2015. doi: 10.1038/nmat4359.

[7] R. E. Simpson, et al., "Interfacial phase-change memory," Nature Nanotechnology, vol. 6, no. 8, pp. 501-505, Aug. 2011. doi: 10.1038/nnano.2011.96.

[8] M. Rudé, et al., "Optical switching at 1.55 μ m in silicon racetrack resonators using phase change materials," Appl. Phys. Lett., vol. 103, no. 14, p. 141119, Sep. 2013. doi: 10.1063/1.4824714.

[9] Y. Zhang, et al., "Broadband transparent optical phase change materials for high-performance nonvolatile photonics," Nat. Commun., vol. 10, no. 1, pp. 1-9, Sep. 2019. doi: 10.1038/s41467-019-12196-4.

[10] M. Delaney, I. Zeimpekis, D. Lawson, D. W. Hewak, and O. L. Muskens, "A new family of ultralow loss reversible phase-change materials for photonic integrated circuits: Sb2S3 and Sb2Se3," Advanced Functional Materials, vol. 30, no. 36, p. 2002447, 2020. doi: 10.1002/adfm.202002447.

[11] L. J. Prokopeva, et al., "Time domain modeling of bi-anisotropic media and phase change materials with generalized dispersion (Conference Presentation)," in Metamaterials, Metadevices, and Metasystems 2019, Sep. 2019, vol. 11080, p. 1108006. doi: 10.1117/12.2529097.

[12] R. Amin, et al., "0.52 V mm ITO-based mach-zehnder modulator in silicon photonics," APL Photonics, vol. 3, no. 12, p. 126104, Dec. 2018. doi: 10.1063/1.5052635.

[13] M. Miscuglio and V. J. Sorger, "Photonic tensor cores for machine learning," Applied Physics Reviews, vol. 7, no. 3, p. 031404, July 2020. doi: 10.1063/5.0001942.

[14] N. C. Harris, et al., "Efficient, compact and low loss thermo-optic phase shifter in silicon," Opt. Express, vol. 22, no. 9, p. 10487, May 2014. doi: 10.1364/OE.22.010487.

[15] R. Amin, et al., "ITO-based electro-absorption modulator for photonic neural activation function," APL Materials, vol. 7, no. 8, p. 081112, Aug. 2019. doi: 10.1063/1.5109039.

[16] A. Mehrabian, M. Miscuglio, Y. Alkabani, V. J. Sorger, and T. El-Ghazawi, "A Winograd-based integrated photonics accelerator for convolutional neural networks," IEEE Journal of Selected Topics in Quantum Electronics, vol. 26, no. 1, pp. 1-12, Jan. 2020. doi: 10.1109/JSTQE.2019.2957443.