PHOTONICS

On-chip photonic synapse

Zenguang Cheng,1 Carlos Rios,1 Wolfram H. P. Pernice,2 C. David Wright,3 Harish Bhaskaran1*

The search for new “neuromorphic computing” architectures that mimic the brain’s approach to simultaneous processing and storage of information is intense. Because, in real brains, neuronal synapses outnumber neurons by many orders of magnitude, the realization of hardware devices mimicking the functionality of a synapse is a first and essential step in such a search. We report the development of such a hardware synapse, implemented entirely in the optical domain via a photonic integrated-circuit approach. Using purely optical means brings the benefits of ultrafast operation speed, virtually unlimited bandwidth, and no electrical interconnect power losses. Our synapse uses phase-change materials combined with integrated silicon nitride waveguides. Crucially, we can randomly set the synaptic weight simply by varying the number of optical pulses sent down the waveguide, delivering an incredibly simple yet powerful approach that heralds systems with a continuously variable synaptic plasticity resembling the true analog nature of biological synapses.

INTRODUCTION

In stark contrast to conventional computing systems based on the von Neumann architecture where the central computing unit is separated from the main memory, the human brain contains a large number of neurons with synapses, each of them acting as both the computing and the memory unit (1). This unique structure makes the brain energy-efficient (2) in dealing with emotions, learning, and thinking—actions that are nearly impossible, or at least far less efficient and effective, for traditional computers (3). As shown in Fig. 1A, a neuron (pre-neuron) generates action potentials (spikes, fire time $t_{\text{pre}}$) that propagate along the axon and are transmitted through a junction to the next neuron (post-neuron) that generates the postsynaptic action potentials (fire time $t_{\text{post}}$). The junction is called a synapse (inset in Fig. 1A), with the synaptic weight ($w$) determining the communication strength between the two neurons. The synaptic plasticity (that is, the change in synaptic weight) $\Delta w$ is determined by neural activities, for example, $\Delta w = f(t_{\text{post}} - t_{\text{pre}})$ based on the Hebbian learning rule (4), and is believed to be the primary mechanism for memory and learning in the human and animal brain (5). Inspired by the brain, neuromorphic computing that attempts to imitate the neural system at the physical level has gained significant attention, recently driven by the needs of “big data” (6), artificial intelligence (7), and a supporting computing concept for the internet of things (8). As a first and key step to the construction of neuromorphic architecture, it is essential to develop suitable plastic synapse-like devices (1, 9, 10)—not least because synapses are by far the most numerous component of real brains, outnumbering neurons by several orders of magnitude (2).

Several electronic devices have recently been investigated to achieve synaptic function, such as those based on electrically induced resistive changes in phase-change chalcogenides (11–13), metal-insulator-metal structures (14, 15), and ferroelectric materials (16), as well as nanomaterial-based field-effect transistors (17, 18). A photonic synapse based on microfibers (19) and an optoelectronic synapse using carbon nanotubes (20) have also been demonstrated with potential benefits of large bandwidth (21, 22) and no electrical interconnect power loss (22) inherently from the optical domain, although they tend to be either difficult to integrate and speed-limited, or still rely on electrical excitation signals. Here, we demonstrate a fully integrated all-photonic synapse based on phase-change materials (PCMs) that resembles the neural synapse at the physical level and can achieve synaptic plasticity compatible with the well-known Hebbian learning or spike timing-dependent plasticity (STDP) rule.

RESULTS

Concept of on-chip photonic synapse

The concept of a photonic synapse is shown schematically in Fig. 1B. A waveguide with discrete PCM structures on top acts as the photonic synapse with the input and output of the waveguide connected with a pre-neuron and a post-neuron. An optical circulator is used for connecting the output of the synapse and the post-neuron (from port 2 to port 3) and for applying optical pulses to alter the synaptic weight (from port 1 to port 2). Low-energy optical transmission can be measured from the pre-neuron to the post-neuron, with the transmission level dependent on the synaptic weight.

It has previously been demonstrated that optical pulses can switch PCMs integrated on waveguides to provide non-volatile photonic memories storing up to eight levels in a single cell (23). To move between levels in that memory implementation required single pulses with varying powers for amorphization and a complicated multipulse, multipower format for recrystallization. However, for the realization of a practicable synapse mimic, not only precise control of synaptic weighting but also a device whose weight is controlled by fixed pulse characteristics is crucial. Here, we achieve this by designing a tapered waveguide structure and then incorporating multiple, small, discrete PCM islands (Fig. 1C and D). This enabled us to have much improved control of the electric field (of the optical pulses propagating along the waveguide) that interacts with the PCM. This resulted in a very effective method for synaptic weight control that is based entirely on the number of fixed-duration, fixed-power excitation pulses applied, as we demonstrate later.

Synaptic-mimic design with FEM analysis

We use the well-studied chalcogenide Ge2Sb2Te5 (GST) (24–27) for the PCM cells, with indium tin oxide (ITO) as the capping layer (see Materials and Methods). Figure 1C shows the optical image of a typical device. Two diffraction grating couplers are used to connect the device with fiber arrays for signal transmission (Fig. 1C and fig. S1). The central part of the waveguide (with the discrete GST islands on top) is tapered to a 0.8-μm width from a nominal width of 1.3 μm elsewhere (Fig. 1D).

The effectiveness of the combination of the tapered structure with the discrete PCM islands for controlling the interaction field along the waveguide is exemplified in Fig. 2 (and fig. S3). Here, a transverse...
electric (TE) optical field at 1580 nm is “launched” into the left side of several alternative waveguide designs (Fig. 2, A and B, and fig. S3, A and B), and the strength of the electric field (E-field) at the waveguide surface is simulated for the PCM in the crystalline (Fig. 2) and amorphous (phase (s3). With the GST in the crystalline phase, a simple rectangular waveguide with a single large GST cell on top (Fig. 2A, as used in the memory application of Rios et al. (23)), the E-field decays rapidly because of the very strong absorption and exhibits many resonant peaks and troughs (standard design in Fig. 2C). However, for the tapered waveguide with discrete GST islands (synapse-mimic design in Fig. 2C), the interaction of the E-field (and so the absorption) is much more “controlled,” with a much more gradual decay and little evidence of resonance effects. To quantitatively compare the electric field distributions in the various structures, we calculated the average, SD, and range of the E-field inside the GST regions, as well as the ratio between the fields at the right ($E_{\text{GST}}^{\text{in}}$) and left sides ($E_{\text{GST}}^{\text{out}}$) of the GST cells (Fig. 2D). The SD and the range of the electric field in the GST regions of the synapse-mimic structure are the smallest, and this structure also has the highest $E_{\text{GST}}^{\text{out}}/E_{\text{GST}}^{\text{in}}$ ratio, thus demonstrating the smoothest distribution of the E-field inside and that most energy transmitted past the GST cells in this structure (detailed comparisons between four structures are elucidated in section S2 of Supplementary Text).

**Synaptic weighting and plasticity**

The enhanced E-field control engendered by the use of the structure comprising a tapered waveguide with discrete PCM islands leads directly to the implementation of a simple yet most effective integrated photonic synapse, as we now show. Before optical measurements, the devices were annealed on a hot plate (~250°C) for 10 min to completely crystallize the GST. The optical transmission ($T_0$) of the device with the GST in the fully crystalline state is defined as the baseline of the readout and assigned to a synaptic weight “0.” Any subsequent change of the readout ($\Delta T = T - T_0$) during the measurement is normalized as the relative change in percentage ($\Delta T/T_0$) to the baseline (Fig. 3A). Changes in synaptic weighting, as shown in Fig. 3, were then achieved by sending fixed-duration, fixed-energy optical pulses down the waveguide. Using one single pulse of 50 ns at 243 pJ (section S1 of Supplementary Text), the transmission readout changed by ~7% to weight “3” (arbitrarily defined, but predetermined weight numbers). The weight could then be decreased from weight “3” to “1” with 100 identical pulses (repetition rate, 1 MHz; total weighting time, 100 μs) using the exact same pulse parameters (that is, 50 ns, 243 pJ). When we increased the pulse number to 1000 (total weighting time, 1 ms), the device went from weight “1” to “0.” This represents a crucial advance, because it allows one to arbitrarily adjust the synaptic weight using a set of known pulses without having previous knowledge of the current actual weight. This finding is worth discussing further. For example, weight “1” in Fig. 3A can be switched from weight “3,” “2,” or “0” with exactly the same parameters of pulses (100 pulses of 243 pJ and 50-ns length); similarly, weight “0,” “2,” and “3” can be set with 1000, 50, and 1 pulses, respectively, from any (undetermined) previous weight. Moreover, we examined the long-term durability of the switching between different weights, as shown for example in Fig. 3B. Even after 38 cycles of switching, carried out over a period of many minutes, the individual weights are clearly distinguishable with the deviation of each weight below 0.77% in readout transmission. These results illustrate a direct and unambiguous achievement of deterministic synaptic weighting using a known number of optical pulses (section S3 of Supplementary Text).

To uncover the relationship between the synaptic weight and the pulse number in more detail, we increased the power of the pulse to obtain five stable weights with high signal-to-noise ratio (Fig. 3C and fig. S5). The device started at weight “0” (baseline), which was then followed by pulses of 200, 100, 50, and 1 (404.5 pJ, 50 ns) to reach weights “1,” “2,” “3,” and “4,” respectively; the pulse sequence was then reversed (that is, 50, 100, 200, and 1000 pulses applied corresponding to a total weighting update time of 50 μs, 100 μs, 200 μs, and 1 ms, respectively) to access weights “3,” “2,” “1,” and “0,” and the whole process was repeated 10 times. We extracted the mean value of the transmission change for each weight (monitored for ~10 s) from the 1st and 10th weighting cycle and plotted them versus the corresponding pulse...
numbers, shown in Fig. 3D. First, as previously stated, we see that the change in synaptic weight is exponentially and monotonically dependent on the number of pulses applied. This relation is very stable and is sustained for long periods of time (it took us 7 min to complete the experiments in Fig. 3C and fig. S5, for example). Moreover, we note that, by tuning of the pulse parameters, we can readily achieve an increased number of synaptic weights. For example, with a decreased pulse width of 20 ns and a pulse energy of 216 pJ, we readily achieved 11 synaptic weights, with a similar exponential relationship between the change of synaptic weights, with a similar exponential relationship between the change of synaptic weight and the number of pulses (fig. S6). By further pulse control, and/or by increasing the signal-to-noise ratio [for example, by increasing the readout power (23)] and/or by further device optimization, it will be possible to access considerably more synaptic weighting levels or even achieve continuous weighting to mimic the true analog nature of synaptic weight change in biological systems (28).

All-optical STDP plasticity

Finally, we point out that the exponential dependence of synaptic weight on the number of applied pulses in our photonic synapse leads to a very simple and compact implementation of the STDP rule using photonic integrated-circuit techniques. The STDP rule to describe synaptic weight change in a biological system has the form (4) of $\Delta w = A e^{-\Delta T/\tau}$. Here, $\Delta w$ and $\Delta T$ are the synaptic weight change and the time delay between pre- and postsynaptic signals, whereas $A$ and $\tau$ are constants. This STDP behavior could be achieved quite simply with an all-optical structure based on our photonic synapse (Fig. 4). As shown in Fig. 4A, the presynaptic signal with power $P_{\text{pre}}$ is split into two beams with 50% coupled into a photonic synapse similar to that of Fig. 1, and the other 50% ($P_{\text{in1}}$) is connected to an interferometer via a phase modulator. The postsynaptic signal ($P_{\text{post}}$) is also divided into two parts, with 50% transmitted and the remainder ($P_{\text{in2}}$) fed back to the interferometer. By adjusting the phase modulator, the net output power ($P_{\text{out}}$) of the interferometer obtains the tunability between zero and ($P_{\text{in1}} + P_{\text{in2}}$), and this output is used to update the weight of the synapse. In this particular design, the powers of pre- and postsynaptic signals are chosen to lie between $P_{\text{th}}$, and $2 \times P_{\text{th}}$ where $P_{\text{th}}$ is the threshold power for switching PCMs (23) such that $P_{\text{th}}/2 < P_{\text{in1}}, P_{\text{in2}} < P_{\text{th}}$ as shown in Fig. 4B. The pulse widths and repetition rates of pre- and postsynaptic signals are intentionally set differently. When there is no time delay ($\Delta T = 0$) between the pre- and postsynaptic signals, the net output power from the interferometer applied to the synapse (the red trace in Fig. 4B) has a single pulse larger than $P_{\text{th}}$. With increasing the time delay ($\Delta T = \Delta t_1, \Delta t_2, \Delta t_3$, and $\Delta t_4$), the number of output pulses with net power above $P_{\text{th}}$ gradually increases to 2, 3, 4, and 5 in this example, as shown in Fig. 4 (C to F). By an appropriate design of the pre- and postsynaptic signals, the number of output pulses with power larger than $P_{\text{th}}$ could be linearly dependent on $\Delta t$, leading to the required exponential dependence of the synaptic weight change on the time delay (between pre- and postsynaptic signals), thus mimicking the STDP behavior in a simple and effective manner.
In conclusion, with a specially designed structure of discrete phase-change islands on a tapered waveguide, we have obtained a brain-inspired, on-chip biomimetic photonic synapse with analog and cumulative programmability, essential requirements for neuromorphic computing. Optical field simulations demonstrate that the distribution of the electric field in the photonic synapse is much more homogeneous than in conventional waveguide designs. This feature allows for the deterministic adjustment of synaptic weights with a predetermined number of identical, fixed-energy, fixed-duration pulses. Moreover, by intentionally arranging the pre- and post-neuron signals, we have elucidated an all-optical method to modulate the synaptic weight based on the time delay between the pre- and post-neuron signals that can mimic the STDP rule in biological systems. This feature allows for the deterministic adjustment of synaptic weights with a predetermined number of identical, fixed-energy, fixed-duration pulses. Furthermore, by intentionally arranging the pre- and post-neuron signals, we have elucidated an all-optical method to modulate the synaptic weight based on the time delay between the pre- and post-neuron signals that can mimic the STDP rule in biological systems.

MATERIALS AND METHODS

Device fabrication and characterization

The photonic synapses were fabricated on a Si$_3$N$_4$/SiO$_2$ platform, as reported previously. A JEOL JBX-5500ZX 50-kV electron beam lithography system was used to define the photonic devices on the wafer spin-coated with MAM-N 2403 negative-tone resist, followed by reactive ion etching in CHF$_3$/O$_2$/Ar to etch down 300 nm of Si$_3$N$_4$. A second step of electron beam lithography using poly(methyl methacrylate)–positive resist was used to define the pattern of discrete PCM islands, and 10-nm GST/10-nm ITO was deposited.
subsequently sputter-deposited. The structure of the photonic synapse was characterized by a scanning electron microscope (Hitachi S-4300) with a low accelerating voltage (1 to 3 kV). The images were obtained using the secondary electron detector at a working distance of ~13 mm.

**Finite element method simulations**

The finite element method simulations were carried using the COMSOL Multiphysics software with the RF module. A TE mode optical field with a nominal power of 1 W was injected into the waveguide. The electromagnetic field distribution in the frequency domain was simulated inside a three-dimensional model of the waveguide with GST films or islands. The results shown in the text are the amplitudes of the electric fields recorded in the central cross section in the $x$-$y$ plane of the structure.

**Optical measurement**

Real-time transmission measurements of the photonic synapse during optical pulse switching were performed using a probe-pump technique, as described previously (23) and in section S4 of Supplementary Materials and Methods. The measurement setup is illustrated in fig. S2. Briefly, a low-power probe laser and a high-power pump laser working at different wavelengths were routed through the photonic synapse from opposite directions. Two optical circulators were used to guide one laser into the device while directing the other laser out to detectors. To suppress interference between the two signals, we used two optical band-pass filters (OTF-320, Santec) in the probe and pump lines. A continuous-wave (CW) diode laser (TSL-550, Santec) as a probe laser was used to interrogate the transmission of the photonic synapse. The pump pulse was generated from a CW diode laser (N7711A, Keysight) combined with an electro-optic modulator (EOM) (2623NA, Lucent) that was controlled by an electrical pulse generator (AFG 3102C, Tektronix), and subsequently amplified by a low-noise erbium-doped fiber amplifier (AEDFA-CL-23, Amonics).

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/3/9/e1700160/DC1

**Supplementary Materials and Methods**

Supplementary Text

fig. S1. Design of photonic synapse.
fig. S2. Optical measurement scheme.
fig. S3. Optical field distribution in photonic synapses.
fig. S4. Optical field distributions in S2 and synapse-mimic designs.
fig. S5. Full trace of five-level weighting.
table S1. Dimensions of photonic synapses.
table S2. Fitting of synaptic weight on pulse number.

**REFERENCES AND NOTES**


Acknowledgments: We acknowledge discussions with G. Rodriguez Hernandez and A. Ne. Funding: This research was supported via the Engineering and Physical Sciences Research Council Manufacturing Fellowship EP/K018694/1, the Wearable and Flexible Technologies (WAFIT) collaboration (EP/M015173/1), and the Chalcogenide Advanced Manufacturing Partnership (EP/M015130/1). Author contributions: All authors contributed substantially. Z.C. and H.B. conceived, planned, and supervised the project. Z.C. and C.R. fabricated the samples and performed the optical measurements and simulations. W.H.P.P., C.D.W., and H.B. led the project. All authors discussed the results and wrote the manuscript.

Competing interests: H.B. notes that he serves on the board of directors of Bodle Technologies Ltd. C.R., W.H.P.P., and H.B. are authors on a pending international patent application related to this work filed by the Oxford University Innovation Ltd. (application no. PCT/GB2016/052871; publication no. WO2017046590 A1; filing date, 15 September 2016; publication date, 23 March 2017). All other authors declare that they have no competing interests.

Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 16 January 2017
Accepted 5 September 2017
Published 27 September 2017
10.1126/sciadv.1700160

On-chip photonic synapse
Zengguang Cheng, Carlos Ríos, Wolfram H. P. Pernice, C. David Wright and Harish Bhaskaran

Sci Adv 3 (9), e1700160.
DOI: 10.1126/sciadv.1700160