Reversible training of waveguide-based AND/OR gates for optically driven artificial neural networks using photochromic molecules

OPEN ACCESS
IOP Publishing
Journal of Physics D: Applied Physics
J. Phys. D: Appl. Phys. **55** (2022) 044002 (10pp)
https://doi.org/10.1088/1361-6463/ac2d62

Reversible training of waveguide-based AND/OR gates for optically driven artificial neural networks using photochromic molecules

Seon-Young Rhim[1], Giovanni Ligorio[1], Felix Hermerschmidt[1], Michael Pätzel[2], Martin Herder[2], Stefan Hecht[2,3,4] and Emil J W List-Kratochvil[1,5,*]

[1] Institut für Physik, Institut für Chemie & IRIS Adlershof, Humboldt-Universität zu Berlin, Zum Großen Windkanal 2, 12489 Berlin, Germany
[2] Institut für Chemie & IRIS Adlershof, Humboldt-Universität zu Berlin, Brook-Taylor-Straße 2, 12489 Berlin, Germany
[3] DWI—Leibniz Institut für interaktive Materialien e.V., RWTH Aachen University, Forckenbeckstraße 50, 52074 Aachen, Germany
[4] Institut für Technische und Makromolekulare Chemie, RWTH Aachen University, Worringerweg 2, 52074 Aachen, Germany
[5] Helmholtz-Zentrum für Materialien und Energie GmbH, Hahn-Meitner-Platz 1, 14109 Berlin, Germany

E-mail: emil.list-kratochvil@hu-berlin.de

Received 12 July 2021, revised 13 September 2021
Accepted for publication 6 October 2021
Published 21 October 2021

Abstract

Artificial neural networks (ANNs) are inspired by the biological nervous system. The high performance of such ANNs is achieved through the dynamic change of the synaptic weights by applying self-optimizing learning algorithms. Despite the simple operations for each single element in an ANN, a network with a huge number of simulated elements consumes lots of computing capacity using von Neumann computer architectures. To overcome this issue, neuromorphic devices facilitate the design of hardware ANNs that emulate the synaptic functions. Here we demonstrate the viability of such an approach using photonic waveguides in combination with a photochromic diarylethene (DAE) molecule. By positioning and irradiating DAE molecules on single waveguides, we modulate the intensity and thereby emulate the plasticity of the synaptic weights. To run the photonic device as an ANN we firstly characterize the modulation range and encode a learning procedure accordingly. As the proof of concept, we operate a y-shaped waveguide performing basic AND/OR logic gate functions, with the capability to switch between these two gate functions by using specific training sets.

Keywords: artificial neural networks, backpropagation, gradient descent, neuromorphic devices, photochromic molecules, synaptic plasticity, waveguides

(Some figures may appear in colour only in the online journal)

* Author to whom any correspondence should be addressed.

1. Introduction

Biological nervous systems are highly complex structures primarily formed by neurons, which interconnect and communicate via synapses. While neurons mainly participate in transducing, processing, and transmitting signals, synapses are responsible for wiring of the neurons and therefore determine the network formed through their individual synaptic wiring strength. The ability to enhance (potentiate) and inhibit (depress) the signal transduction capability between neurons (synaptic weight) is called synaptic plasticity. In such biological neural networks, signals, which are mostly triggered by external stimuli, propagate in parallel through the system and in the entanglement of the synapses a communication pattern of the interworking neurons is created. Inspired by these biological processes, artificial neural networks (ANNs) have been implemented to solve complex tasks, such as pattern recognition or complex data processing, by emulating the functions of neurons and synapses in software or complementary metal–oxide–semiconductor (CMOS) based hardware network simulations [1–3].

However, the simulation of complex tasks requires fast computing capacities such that commonly used CMOS-based hardware architectures face their physical limits [4, 5]. In order to overcome these limitations, recent progress in neuromorphic engineering offers novel neuro-inspired electronic devices that directly emulate the synaptic plasticity functions and can therefore serve as building blocks for ANNs [6–9]. A well-established example are two-terminal resistive switching elements that mimic the synaptic transduction process by a change in resistance due to a change in applied voltage bias. Yet, ANNs composed of such electrical devices need a conversion of the current to the voltage after passing every neuron in the signaling process [10]. In contrast, photonic devices and consequently networks can run the same operations without conversion [11–14]. By employing the same physical quantity such as light intensity for the input and the output, an uninterrupted cascade of the signal through the network is possible. Therefore, switchable materials such as phase-change materials and photochromic molecules became the focus of the research to enable intensity modulation of guided modes [15–18].

Common ANNs are schematically built up with artificial neurons, typically organized in layers [19], in which each neuron is connected with all other neurons in the prior and next layer as shown in figure 1(a). As first proposed by McCulloch and Pitts [20], figure 1(b) displays the model of a single artificial neuron, which consists of the synaptic weights w_i, a transition function and an activation function φ. Each input branch represents a single synapse with its synaptic weight w_i, while the outgoing branch is the output. Thus, the incoming values x_i are multiplied by the associated synaptic weights w_i. The transition function sums up the multiplied values and transfers it to the activation function φ with a threshold a, which gives 1 as the output when the transferred value exceeds the threshold a, otherwise it gives 0 [21]. Mathematically, the output y of a single artificial neuron can therefore be described as follows

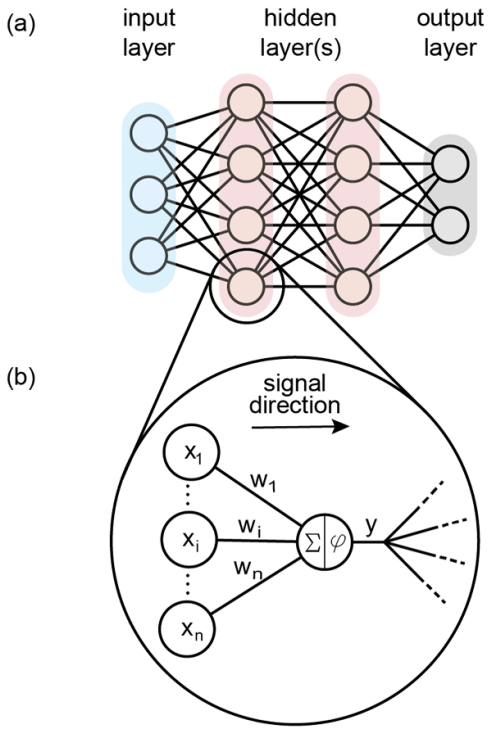

Figure 1. (a) The schematic illustration of a multi-layered artificial neural network (input layer (blue), hidden layer (red), output layer (grey)). The nodes here presented as circles are the artificial neurons and the connections between the neurons are the synapses. (b) Basic structure of a single artificial neuron with n inputs.

$$y = \varphi\left(\sum_i w_i x_i\right) = \begin{cases} 1, \text{ if } \sum w_i x_i \geqslant a \\ 0, \text{ else} \end{cases} . \qquad (1)$$

In this model of the artificial neuron, the synaptic weight is the only free parameter affecting the output y. Hence, controlling these synaptic weights with a learning algorithm such as backpropagation enables the network to adapt to new situations and to solve new tasks [22–24]. Actually, the learning process is the crucial part of the ANNs and is usually conducted by training sets, which consists of certain inputs x_i and a target output t. By frequently feeding the network with training sets, the synaptic weight values successively change depending on the error between the actual and the target output. The method of these gradual changes is called gradient descent and adjusts the synaptic weights w_i by minimizing the error at the output [25]. Trained networks allow variations of the reminded (trained) inputs on the same outcome. This is the case, for example in software used for handwriting recognition, where letters and numbers are interpreted correctly and substituted with the appropriate corresponding ASCII character.

In general, to design analogous hardware ANNs, the physical quantity of the devices must be able to process the output as per equation (1). Such hardware networks can be realized in electronics and in photonics. In the case of electronics, the synaptic functions can be emulated by the current flowing through the device, while in the case of

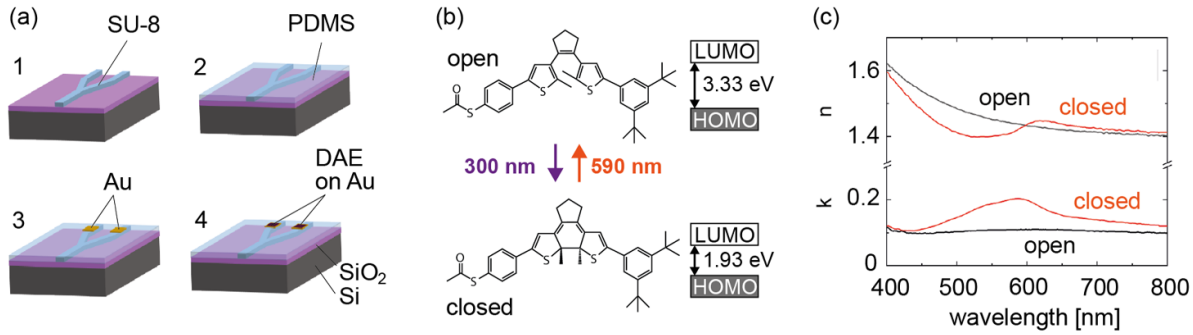

Figure 2. (a) Fabrication steps of the y-branch device. (b) The chemical structures of the photochromic diarylethene (DAE) molecule in its open (top) and closed (bottom) configuration with the associated energy gaps. (c) Refractive index n and the extinction k spectrum of the DAE-film in open (black) and closed (red) configuration measured by ellipsometry.

photonics, the synaptic functions can be emulated by light transmitted through a waveguide. Hardware ANNs capable of learning have received the attention of much research focus because of their promising ability to run these operations very efficiently [26, 27]. However, limitations are quickly reached when using resistive switching devices or transistor ANNs implemented in a crossbar structure to perform ANN functions. Though these devices are able to perform the ANN operations according to equation (1), such implementations are not suitable to directly implement analogous learning processes.

Especially when using a resistively switching two-terminal device, one cannot write and read the device state simultaneously, which makes a (short-term) storage of the output error necessary to recall it later during the learning process. In contrast to such electronic implementations, optical devices allow a change to their optical properties while being irradiated externally. Thus, using different wavelengths allows to conduct write and read processes simultaneously [18, 28, 29]. This enables a direct analogous adjusting of the synaptic weights w_i depending on the current output. Photochromic and phase-change materials can change their dielectric function upon irradiation. With switching responses down to nanoseconds for both material classes, they are promising candidates for signal modulation to realize competitive photonic networks [30–33]. Moreover, light propagation in a waveguide is (almost) not affected by waveguide crossings and therefore photonic structures provide a high design freedom to implement ANN optical systems [34, 35]. Several approaches with phase-change materials were reported in recent years to realize photonic ANNs by using different operating wavelengths [36] and phase coding [37]. Finally, ANNs using photonic implementations also benefit in terms of energy consumption by less heat dissipation and high bandwidth due to lower signal pulse dispersion.

In this contribution we utilize a y-branch waveguide as shown in figure 2(a) to demonstrate the capability of a photonic network to carry out the fundamental ANN processes. In our previous work we investigated the use of a photochromic diarylethene (DAE) thin film to emulate long-term synaptic plasticity using surface plasmon polariton switching in a Kretschmann configuration [38]. In that work, surface plasmon polaritons were excited on a thin gold film with a DAE layer on top and were modulated by irradiating the DAE layer with the appropriate wavelength. The modulation causes an intensity change of the reflected light from the prism. A cycle endurance at least up to 50 cycles and a state retention at least up to one hour was sufficient for the lab scale investigations presented herein. Note that all measurements were performed in ambient atmosphere at room temperature, and it is reasonable to expect that the cycle endurance and stability can be drastically improved upon encapsulation of the entire optical chip. Here we expand the usage of the DAE molecules for optical ANNs to waveguide systems, by depositing a DAE film on a reverse symmetry waveguide (with a higher refractive index of the upper cladding as compared to the bottom cladding). With this structure the penetration depth of an evanescent field, which is excited by the total internal reflection of the guided light, increases up to several micrometers into the upper cladding [39, 40]. The larger penetration depth enables the evanescent field to interact with the molecules, which are not deposited in the direct vicinity of the waveguide. This effect is also called frustrated total internal reflection and causes absorption of the guided light according to the energy gap of the DAE molecule [41]. We exploit this absorption to modulate the light intensity in each single arm of the y-branch waveguide by irradiating the DAE-film to emulate the synaptic function. The resulting recombination of the light intensities determines the logical output value as per equation (1). In the proof of concept for the full viability of the y-branch waveguide as a photochromic-photonic ANN, we train the y-branch using supervised learning for AND/OR logic gates, by applying the gradient descent delta rule (the special backpropagation case for one layered ANNs).

2. Results and discussion

2.1. Device fabrication and measurement setup

The fabrication steps of the y-branch device are shown in figure 2(a). Firstly, the photoresist SU-8 TF6005 (70 cSt

viscosity, 1.083 mg ml^{-1}, Kayaku Advanced Materials, Inc.) was spin-coated on a clean 3 μm thick silicon dioxide substrate (Siegert Wafer GmbH) at 3500 rpm with 500 rpm s^{-1} acceleration for 50 s to form the waveguide material. After spin coating, the substrate was soft baked at 110 °C for 5 min. For the structure patterning, a direct writing laser lithograph (μPG 101, Heidelberg Instruments) with 375 nm wavelength was used. Afterwards the substrate was post-exposure baked at 110 °C for 2 min then put in the developer solvent (mr-600 dev, micro resist technology GmbH) for 3 min. Subsequently the device was rinsed in isopropanol, then gently blown dry with nitrogen and hard baked at 200 °C for 20 min. The thickness (5 μm) of the waveguide was measured with a profilometer (Dektak, Bruker Corp.) and the width (30 μm) with optical microscopy. For the cladding, polydimethylsiloxane (PDMS, Dowsil 184) was prepared with the curing agent (Sylgard 184) in a 10:1 ratio according to the recipe provided by Dow Chemical. After treating the structured device in oxygen plasma for 5 min, the PDMS was spin-coated with 4500 rpm for 40 s to obtain a ca. 6 μm thick layer. For the deposition of the photochromic DAE-film, the molecules were dissolved in ethanol with a concentration of 1 mg ml^{-1}. A 2×2 mm gold area with 40 nm thickness was evaporated on PDMS over each waveguide arms of the y-branch structure. Finally, 10 μl of the DAE solution was deposited via drop casting on the two gold areas. The synthesis of the DAE molecules has previously been described in [29]. Note that due to the better wetting of the DAE solution on gold as compared to PDMS, the DAE solution remains on the gold areas, which hinders the diffusion of the DAE molecules into the PDMS.

As depicted in figure 2(b), when irradiated with light of the appropriate wavelength, the DAE molecules undergo a photochemical isomerization in which the molecule changes between a fully conjugated (closed) and a disconnected, cross-conjugated (open) structure. This valence-bond tautomerization substantially alters the energy gap between the highest occupied molecular orbital and the lowest unoccupied molecular orbital of the DAE and thereby affects the absorption and hence the refractive index of the DAE-film. Figure 2(c) shows the refractive index n and the extinction coefficient k of a DAE film in which the molecules are in their open (black) and closed (orange) configurations. Both display a significant reversible modulation in the region where the absorption band of the molecule appears from 500 nm to 600 nm upon optical cycling.

For the characterization of the y-branch optical device, light from a tungsten halogen lamp (DH-2000, Ocean Optics, Inc.) was guided to a 100× objective (Olympus) through a fiber. The objective was mounted on an *xyz*-stage for motion control. By focusing the light on the entrance of both waveguide arms, which we define as port 1 and port 2, the light couples to waveguide modes and propagates through the structure to the output, as shown in figure 3(a). In this configuration, the DAE film on the waveguide allows to modulate the propagating light in the waveguide through the interaction with the evanescent field. At the end of the waveguide the light gets collected by a 20× objective which is connected with a charge-coupled

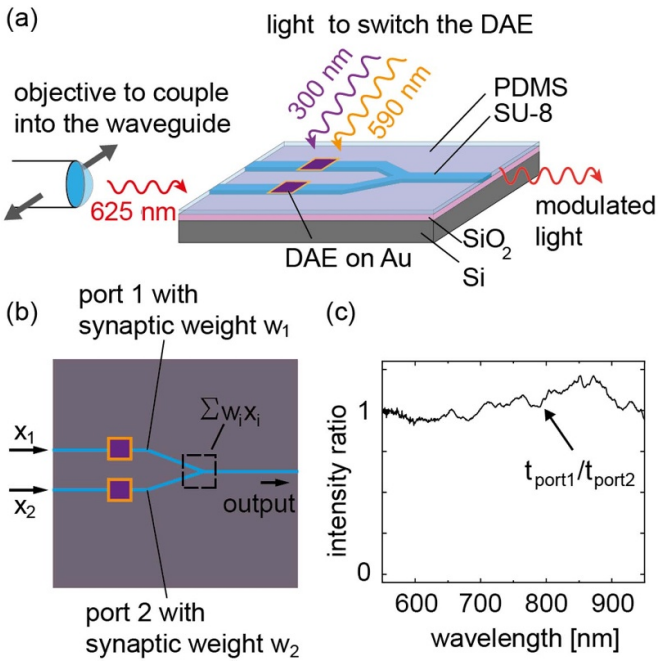

Figure 3. (a) Schematic of the device measurement setup. The objective for light coupling at the inputs can be moved sideways to address the single port. The DAE molecules of both ports were irradiated individually with 300 nm and 590 nm to perform depression and potentiation respectively. (b) The schematic illustration of the working principle of the fabricated y-branch waveguide with two ports. (c) Spectrum of the transmission ratio between port 1 and port 2.

device detector through a fiber. The modulation of the transmitted light for each port is now accomplished by irradiating the DAE molecules using 300 nm and 590 nm LED lamps (Thorlabs Inc.) with an irradiance in the 1 μW mm^{-2} regime. The irradiation time was controlled by a digital multimeter (Keithley 2461). Figure 3(b) displays the working principle of the fabricated y-branch including the position of the DAE films, the ports and the output. The input values (x_1 and x_2) as the light intensities are modulated by the DAE-films. Since the modulation changes the transmission ratio of the propagating light, the outgoing intensity of each port can be described as the multiplication of the initial intensity with the synaptic weight. The output of the device is the sum of the outgoing intensities in the single ports at the branch. For the deployment of the y-branch as a logic gate, the two ports need an equal influence on the total output intensity. For this purpose, we recorded the transmission of each port individually by coupling light first into port 1 then into port 2 and compared their intensities. Figure 3(c) shows the intensity ratio of the two ports near 1 in the visible wavelength range. These transmission spectra in the pristine state of the device were used as references for the following measurements to estimate the DAE absorption and thus the synaptic weights.

2.2. *Characterization of the y-branch waveguide*

In the first step we determine and quantify the modulation range for the waveguide transmission upon switching the

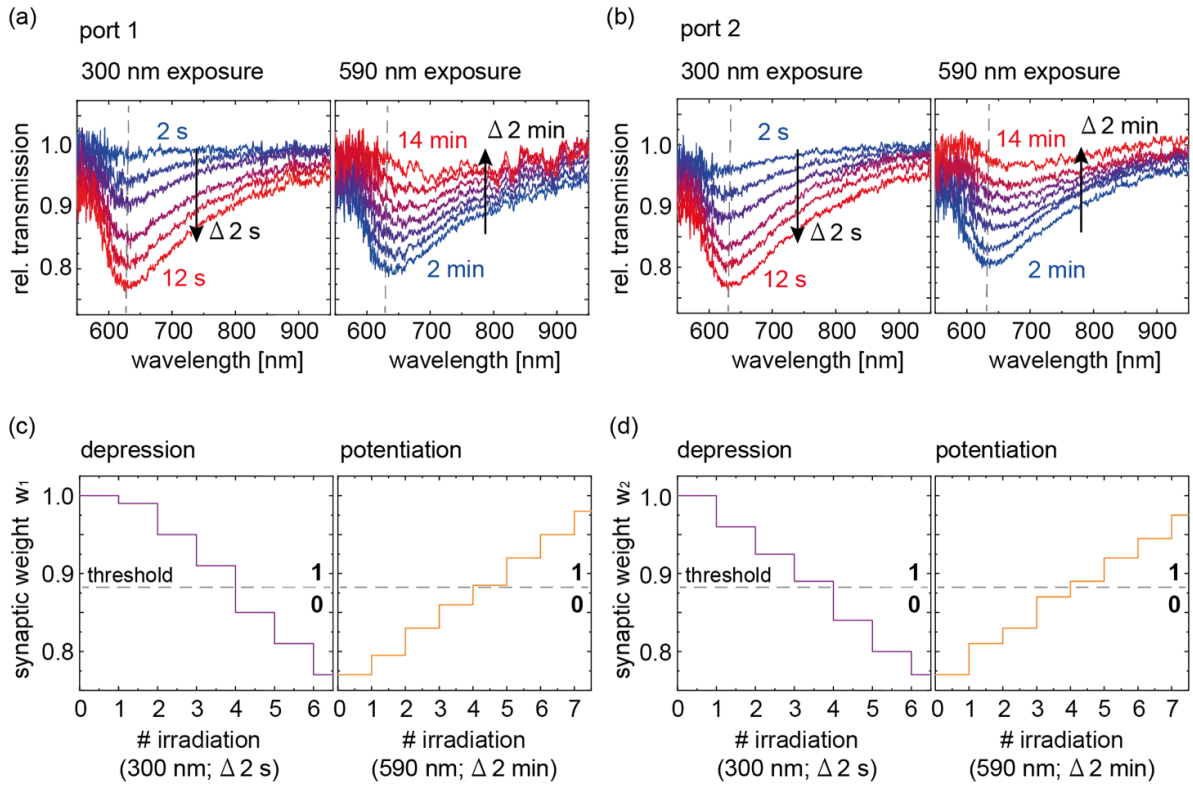

Figure 4. (a), (b) Relative transmission spectra of both ports as a function of the irradiation time on DAE films for 300 nm and 590 nm. (c), (d) Plots of the synaptic weight values for both ports as a function of the number of irradiation steps at the operation wavelength 625 nm.

DAE-films for each port. We therefore irradiated the DAE films on the waveguides with 300 nm and 590 nm light to switch the molecules to the closed and to the open configuration, respectively. In our previous work [38] we found out that the DAE film switches with least fatigue when the UV irradiation time stays below 20 s. Based on this we set the maximum total UV irradiation time to 12 s. All spectra in this section are normalized by using the transmission spectra of the pristine device state as references.

The left graph in figure 4(a) shows the transmission spectra of port 1 for stepwise irradiation with 300 nm light for 2 s. In the spectra we observe a broad absorption peak at 625 nm (black dashed line), which continuously increases. This absorption is caused by the refractive index change from the open to closed configuration of the DAE-film, as shown in the extinction coefficient k of figure 2(c). For a total irradiation time of 12 s, we achieve a transmission reduction at 625 nm by 23%. To switch back to the starting point, the DAE-film was irradiated with visible light (590 nm) in 2 min steps for a total of 14 min as depicted in the right graph of figure 4(a). By changing the molecules from the closed to the open configuration, the transmission in the waveguide can be reversibly switched back to the starting point. The same experimental steps were conducted to characterize the reversible switching of the transmitted light collected at port 2. In figure 4(b) we show that upon closing the DAE molecules one can reduce the transmitted light by 23% and switch back to the starting point by stepwise irradiation with visible light, which shows that both waveguide branches are fully symmetric with respect to their switching behavior. Both ports show the maximum modulation of the transmitted light at 625 nm. For all further discussions and considerations, we now assign the value of the transmitted light at 625 nm to the synaptic weight w_i of each individual port.

With this assignment, figures 4(c) and (d) show the evolution of the synaptic weight w_1 (w_2) versus the individual DAE irradiation steps with 300 nm (left) and with 590 nm light (right). Here the left graphs in figures 4(c) and (d) show a decrease or depression of the synaptic weights w_1 and w_2, while on the right-hand side we show the according potentiation characteristics. According to the achieved synaptic weight modulation range, we define the mean value 0.88 as the threshold a (grey dashed lines). Consequently, if the output value y exceeds this threshold value a the output is interpreted as the logical 1, otherwise 0 as stated in equation (1). Applying the gradient descent delta rule [22], the target value t, learning rate ε and irradiation time b need to be matched for a successful training to adapt such 2×1 (2 inputs × 1 output) ANN functions. With the delta rule the synaptic weight change is defined as follows

$$\Delta w_i = \varepsilon \cdot \delta \cdot x_i$$
$$\text{with } \delta = (t - \sum w_i x_i). \tag{2}$$

Without loss of generality, we set the learning rate ε at 0.5 and the target value t at 0.79 and 0.96 for the logical 0 and

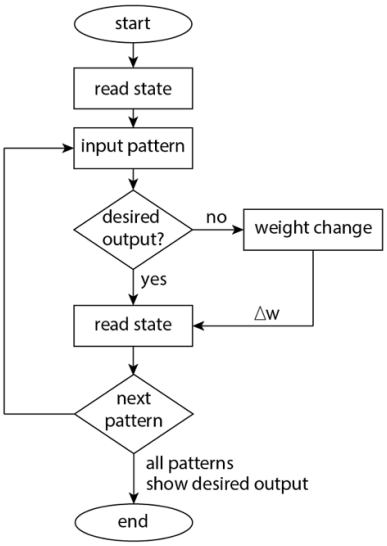

Figure 5. The flow diagram of the training algorithm depicts the procedure described in the main text.

1 respectively. According to equation (2) the synaptic weight change Δw_i is determined by δ, which is here defined as the difference between the sum of the synapses (device ports) $\sum w_i x_i$ and the target value t. To assign the calculated weight changes Δw_i in (2) to irradiation time b, we define

$$b = \Delta w_i \cdot d \tag{3}$$

with d as a constant whose value depends on the negative or positive sign of the weight change Δw_i from equation (2). The constant d is chosen to be $d = -20$ for negative synaptic weight change ($\Delta w_i \leqslant 0$; depression), which triggers an DAE irradiation on port i with 300 nm light according to the time b. A positive weight change ($\Delta w_i \geqslant 0$; potentiation) triggers a DAE irradiation with 590 nm light of which the irradiation time b is calculated with $d = 1200$. Additionally we give the condition only to change the synaptic weights if the output y differs from the desired value. The schema of the utilized training algorithm is shown in figure 5. The training was based on a set of input and output patterns to teach the device a desired function. According to the algorithm in figure 5 we feed the device with training patterns one after another. The synaptic weight is changed every time the output y differs from the desired value according to the threshold value $a = 0.88$, learning rate $\varepsilon = 0.5$, the target values t and equations (2) and (3). The whole learning process ends when all the patterns in the training set show the right output y. In the next section we demonstrate the capability of the optical y-branch to carry out AND/OR logic gate functions by using specified training sets in combination with the delta rule and the algorithm in figure 5.

2.3. Device training for AND/OR logic gate

As a proof of concept, we demonstrate the ability of the y-branch to perform simple AND/OR logic gate functions.

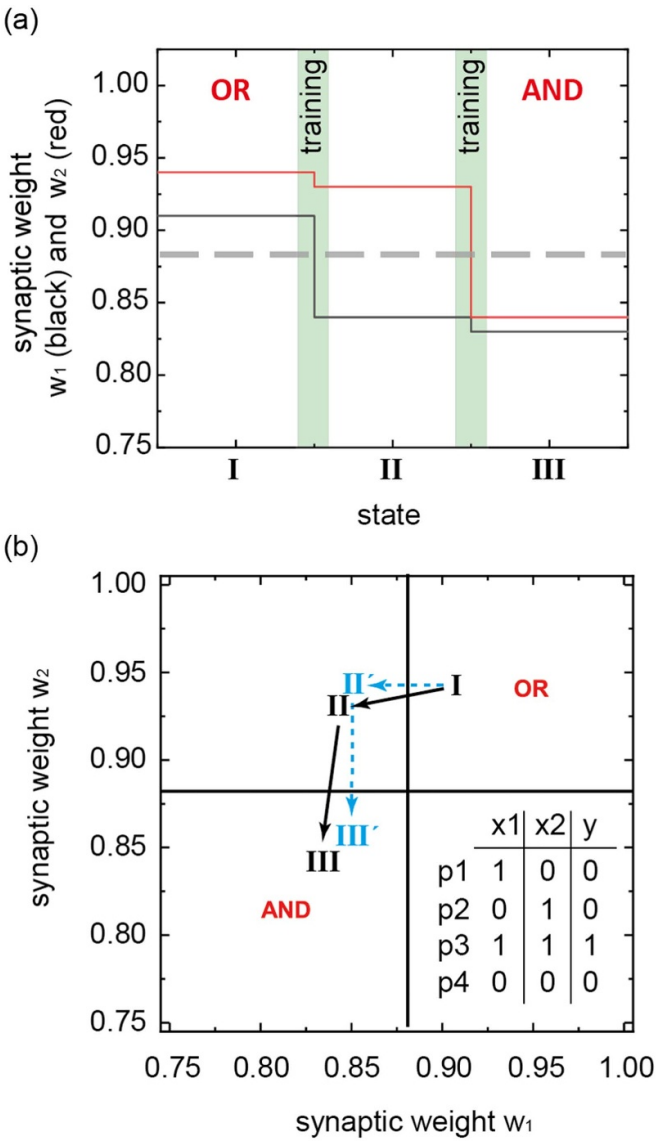

Figure 6. (a) Plots of synaptic weight of port 1 (black) and port 2 (red) during the AND gate training after every training (I to III). The grey dashed line indicates the threshold value at 0.88, while the green marked areas indicates the training. (b) The state map displays the training steps for ideal (blue) and measured (black) device states with the corresponding training paths.

Depending on the conducted training, the y-branch adapts to become either an AND or an OR gate, by applying the gradient descent delta rule with customized training sets. Starting from a random device state, we firstly train the y-branch to run AND logic gate operations by using the equations (2) and (3) and the training algorithm in figure 5. The training set for the AND logic gate is displayed in figure 6 with x_1, x_2 as the inputs and with y as the associated desired output. The training patterns are fed according to the numbered order and are repeated until all patterns show the desired output y.

Figure 6(a) shows the evolution of the synaptic weights w_1 (black line) and w_2 (red line) during learning. Each training process, which is triggered by a training pattern, causes a

Table 1. Training results of the AND-gate training.

Pattern no.	State	Desired output y?	Error δ	Irradiation	Weight changes	
					Δw_1	Δw_2
p1	**I**→**II**	No	−0.12	1.2 s (uv)	−0.07	−0.01
p2	**II**→**III**	No	−0.14	1.4 s (uv)	−0.01	−0.09
p3	**III**	Yes	—	—	—	—
p4	**III**	Yes	—	—	—	—
p1	**III**	Yes	—	—	—	—
p2	**III**	Yes	—	—	—	—

synaptic weight change according to the equation (2). By starting the learning process from an initial state **I**, the synaptic weights show values which exceed the threshold $a = 0.88$, with $w_1 = 0.91$ and $w_2 = 0.94$ as shown in figure 6(a). On the state map in figure 6(b), the initial state **I** is therefore marked in the OR gate region. Table 1 shows the error δ, the irradiated time b and the synaptic weights (w_1 and w_2) for each used training pattern during learning of the AND gate. By feeding the device with the first training pattern (p1) according to equation (1) we obtain the Boolean value 1 ($w_1x_1 + w_2x_2 = 0.91{\cdot}1 + 0.94{\cdot}0 = 0.91 \geqslant 0.88 \Rightarrow y = 1$) as the output. Since it is not the desired output according to p1 ($y = 0$), a weight correction is triggered. Therefore, we set the target value t with 0.79 to calculate a negative error $\delta = (0.79 - 0.91) = -0.12$. A weight change of $\Delta w'_1 = 0.5{\cdot}{-}0.12{\cdot}1 = -0.06$ and $\Delta w'_2 = 0.5{\cdot}{-}0.12{\cdot}0 = 0$ are calculated. According to equation (3), these values trigger an irradiation time b of 1.2 s ($b = \Delta w_i{\cdot}d = -0.06{\cdot}-20 = 1.2$) at 300 nm wavelength on the waveguide port 1. The resulting state **II** ($w_1 = 0.84$ and $w_2 = 0.93$) shows for the synaptic weight w_1 a value below the threshold a. In both ports, the actual changes to the synaptic weight ($\Delta w_1 = -0.07$; $\Delta w_2 = -0.01$) caused by the 300 nm light irradiation, deviate from the calculated changes ($\Delta w'_1 = -0.06$; $\Delta w'_2 = 0$). On the state map in figure 6(b), the device state therefore moves out of the OR gate region (black arrow) to the marked state **II**. Continuing the process, the second training pattern p2 triggers a 1.4 s 300 nm light irradiation on the waveguide port 2, which causes synaptic weight changes of $\Delta w_1 = -0.01$ and $\Delta w_2 = -0.09$. With the synaptic values $w_1 = 0.83$ and $w_2 = 0.84$, the final state **III** is located in the AND gate region, as shown in figure 6(b). All training patterns now show the right outputs for the AND gate function, which leads to an end of the whole learning process. By analyzing the path on the state map, the actual path (black) and the calculated path (blue) coincide sufficiently, confirming the good prediction of the device states through equations (2) and (3).

To obtain the OR gate, the same learning algorithms were applied. By changing the training set, which is displayed in figure 7, we conducted the synaptic weight changes by using equations (2) and (3) again. Table 2 shows the error δ, the irradiated time b and the synaptic weights (w_1 and w_2) to each used training pattern during the OR gate learning. Figure 7(a) shows the synaptic weights w_1 (black line) and w_2 (red line) after every training step. In contrast to the AND gate learning, we notice positive changes of the weights ($\Delta w_i \geqslant 0$), which

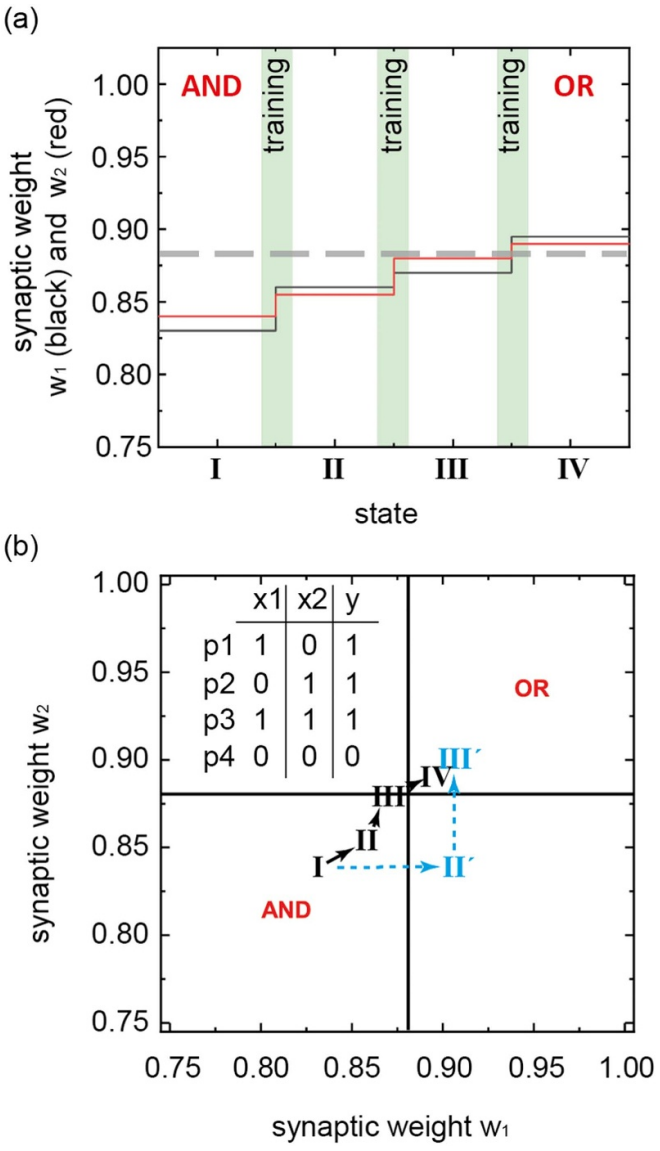

Figure 7. (a) Plots of synaptic weight of the port 1 (black) and the port 2 (red) during the OR gate training after every training (I to IV). The grey dashed line indicates the threshold value at 0.88, while the green marked areas indicate training. (b) The state map displays the training steps for ideal (blue) and measured (black) device states with the corresponding training paths.

trigger visible irradiation of the DAEs. Also, the required training steps (3 steps) are increased to achieve the desired OR gate function. The state map in figure 7(b) shows a large

Table 2. Training results of the OR-gate training.

Pattern no.	State	Desired output y?	Error δ	Irradiation	Weight changes	
					Δw_1	Δw_2
p1	**I→II**	No	0.13	78 s (vis)	0.03	0.015
p2	**II→III**	No	0.11	63 s (vis)	0.01	0.025
p3	**III**	Yes	—	—	—	—
p4	**III**	Yes	—	—	—	—
p1	**III→IV**	No	0.09	54 s (vis)	0.025	0.01
p2	**IV**	Yes		—	—	—
p3	**IV**	Yes		—	—	—
p4	**IV**	Yes		—	—	—
p1	**IV**	Yes		—	—	—

difference between the actual path (black) and the calculated path (blue). The positive changes conducted by DAE irradiation with 590 nm light are more gradual than the 300 nm light irradiation. In order to obtain a better match between the actual Δw_i and the calculated synaptic weight changes $\Delta w'_i$, the constant d in equation (3) can be adjusted to force longer 590 nm light irradiation time b. An explanation for this large deviation is the switching fatigue of the DAE molecules reported in several publications [42–44]. Therefore, the performance of the presented photochromic-photonic device crucially depends on the utilized photochromic materials with regard to retention and switching performance. By employing materials with larger alteration of their optical properties, a higher modulation range and thus a better signal-to-noise ratio can be achieved by the proposed approach. Since the experiments were performed in ambient atmosphere at room temperature, an encapsulation is expected to be able to improve the cycle endurance and reduce switching fatigue. Nevertheless, we observe a successful OR-gate training, indicating the robustness of the gradient descent delta rule on physical neuromorphic devices. Especially the structure of the fabricated optical device, in which the DAE film is located over the waveguide, allows to read (guided transmitted light) and to write (DAE irradiation) simultaneously. By also using different wavelengths for the DAE irradiation (300 nm; 590 nm) and for the waveguide operation (625 nm), an additional separation between writing and reading signals is provided.

To carry out more complex logic functions such as XOR, the optical device (which is the smallest ANN possible in a 2×1 configuration) can be extended with more synaptic input branches and more such elements in series. Thus, a whole layer, as displayed in figure 1(a), can be designed with waveguides in a crossbar structure and thereby expand equation (1) to a matrix calculation. Even a cascaded linking of these layers is possible in this optical approach. It is preferred to rearrange the signal after each layer to facilitate the integration of an activation function and therefore exploit the full potential of an optical ANN. To address single elements of a possible large-scale photonic chip, different approaches based on stereolithography can be exploited. Particularly the employment of so-called digital light processing chips exhibits great spatial and temporal control, which is essential for the performance of such large-scale photonic ANNs. Furthermore, positioning a liquid crystal matrix panel on top of photonic ANNs can be used to control the network with appropriate backlights [45]. However, to achieve competitive integration density of this optical device towards large system integration, the photonic waveguide components need to be exchanged with active plasmonic counterparts to bypass the diffraction limit by confining the propagating light in metallic nanostructures. Combining photochromic molecules with surface plasmons would enable a high modulation range, due to their high sensitivity to the dielectric environment. The feasibility of such approaches in the future can be deduced from previous works studying photochromic materials for plasmonic switching [38, 46, 47]. As for the optical waveguides here, the propagating surface plasmons show a modulation effect in the intensity by switching the photochromic materials, which enables emulation of synaptic functions.

3. Conclusion

In this work we have used a waveguide-based optical y-branch to demonstrate the ANN capability by adapting logic gate functions through learning. By utilizing the switching character of photochromic DAE molecules, we have trained the waveguide device to perform AND/OR gate functions. The utilized training sets were based on the so-called frustrated total internal reflection, which causes an absorption of the guided light according to the energy gap of the DAE molecule. By modulating the light intensity in the waveguide with DAE film, we could therefore prove the viability of an optical neuromorphic device to be compliant to the common learning algorithms such as the gradient descent delta rule. As a result of this concept, we envision a possible utilization of an optical ANN with photochromic molecules for analogue computing.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

We would like to acknowledge Paul Zybarth and Bodo Kranz for their continuous support in the lab. Thanks to Jutta Schwarz for her help with the purification of the DAE derivative. The authors gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft through project 182087777—SFB 951. This work was carried out in the framework of the Joint Lab GEN_FAB and was supported by the HySPRINT Innovation Lab at Helmholtz-Zentrum Berlin.

ORCID iDs

Seon-Young Rhim https://orcid.org/0000-0002-5030-2678
Giovanni Ligorio https://orcid.org/0000-0001-9277-6903
Felix Hermerschmidt https://orcid.org/0000-0001-8292-4124
Martin Herder https://orcid.org/0000-0002-7250-588X
Stefan Hecht https://orcid.org/0000-0002-6124-0222
Emil J W List-Kratochvil https://orcid.org/0000-0001-9206-800X

References

[1] Graves A and Schmidhuber J 2009 Offline handwriting recognition with multidimensional recurrent neural networks *Adv. Neural Inf. Process. Syst. 21—Proc. 2008 Conf.*
[2] Kim T H 2010 Pattern recognition using artificial neural network: a review *Commun. Comput. Inf. Sci.* vol **76** and (Berlin: Springer) 138–48
[3] Lecun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436–44
[4] Theis T N and Philip Wong H S 2017 The end of Moore's Law: a new beginning for information technology *Comput. Sci. Eng.* **19** 41–50
[5] Khan H N, Hounshell D A and Fuchs E R H 2018 Science and research policy at the end of Moore's law *Nat. Electron.* **1** 14–21
[6] Strukov D B, Snider G S, Stewart D R and Williams R S 2008 The missing memristor found *Nature* **453** 80–83
[7] Li Y, Wang Z, Midya R, Xia Q and Joshua Yang J 2018 Review of memristor devices in neuromorphic computing: materials sciences and device challenges *J. Phys. D: Appl. Phys.* **51** 503002
[8] Müller L *et al* 2017 Electric-field-controlled dopant distribution in organic semiconductors *Adv. Mater. Weinheim* **29** 1701466
[9] Van De Burgt Y, Melianas A, Keene S T, Malliaras G and Salleo A 2018 Organic electronics for neuromorphic computing *Nat. Electron.* **1** 386–97
[10] Bayat F M, Prezioso M, Chakrabarti B, Nili H, Kataeva I and Strukov D 2018 Implementation of multilayer perceptron network with highly uniform passive memristive crossbar circuits *Nat. Commun.* **9** 2331
[11] Gu M, Zhang Q and Lamon S 2016 Nanomaterials for optical data storage *Nat. Rev. Mater.* **1** 1–14
[12] Wei H, Wang Z, Tian X, Käll M and Xu H 2011 Cascaded logic gates in nanophotonic plasmon networks *Nat. Commun.* **2** 387
[13] Pan D, Wei H and Xu H 2013 Optical interferometric logic gates based on metal slot waveguide network realizing whole fundamental logic operations *Opt. Express* **21** 9556
[14] Zasedatelev A V, Baranikov A V, Urbonas D, Scafirimuto F, Scherf U, Stöferle T, Mahrt R F and Lagoudakis P G 2019 A room-temperature organic polariton transistor *Nat. Photon.* **13** 378–83
[15] Wuttig M, Bhaskaran H and Taubner T 2017 Phase-change materials for non-volatile photonic applications *Nat. Photon.* **11** 465–76
[16] Stegmaier M, Ríos C, Bhaskaran H and Pernice W H P 2016 Thermo-optical effect in phase-change nanophotonics *ACS Photonics* **3** 828–35
[17] Ebisawa F, Hoshino M and Sukegawa K 1994 Self-holding photochromic polymer Mach-Zehnder optical switch *Appl. Phys. Lett.* **65** 2919–21
[18] Jiang L, Wu J, Li Q, Deng G, Zhang X, Li Z, Chen K and Chiang K S 2019 A photochromic dye doped polymeric Mach-Zehnder interferometer for UV light detection *J. Mater. Chem.* C **7** 6257–65
[19] Gardner M W and Dorling S R 1998 Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences *Atmos. Environ.* **32** 00447–0
[20] McCulloch W S and Pitts W 1943 A logical calculus of the ideas immanent in nervous activity *Bull. Math. Biophys.* **5** 115–33
[21] Almási A D, Woźniak S, Cristea V, Leblebici Y and Engbersen T 2016 Review of advances in neural networks: neural design technology stack *Neurocomputing* **174** 31–41
[22] Widrow B and Hoff M E 1960 Adaptive switching circuits *Wescon Conf. Rec.*
[23] Dreyfus S E 1973 The computational solution of optimal control problems with time lag *IEEE Trans. Automat. Control* **18** 383–5
[24] Rumelhart D E, Hinton G E and Williams R J 1986 Learning representations by back-propagating errors *Nature* **323** 533–6
[25] Curry H B 1944 The method of steepest descent for non-linear minimization problems *Q. Appl. Math.* **2** 258–61
[26] Prezioso M, Merrikh-Bayat F, Hoskins B D, Adam G C, Likharev K K and Strukov D B 2015 Training and operation of an integrated neuromorphic network based on metal-oxide memristors *Nature* **521** 61–64
[27] Ambrogio S *et al* 2018 Equivalent-accuracy accelerated neural-network training using analogue memory *Nature* **558** 60–67
[28] Kang J W, Kim J J and Kim E 2002 All-optical Mach-Zehnder modulator using a photochromic dye-doped polymer *Appl. Phys. Lett.* **80** 1710–2
[29] Venkatakrishnarao D, Mohiddon M A, Chandrasekhar N and Chandrasekar R 2015 Photonic microrods composed of photoswitchable molecules: erasable heterostructure waveguides for tunable optical modulation *Adv. Opt. Mater.* **3** 1035–40
[30] Sasaki K and Nagamura T 1997 Ultrafast all-optical switch using complex refractive index changes of thin films containing photochromic dye *Appl. Phys. Lett.* **71** 434–6
[31] Ishibashi Y, Nakai S, Masuda K, Kitagawa D, Kobatake S and Asahi T 2020 Nanosecond laser photothermal effect-triggered amplification of photochromic reactions in diarylethene nanoparticles *Chem. Commun.* **56** 7088–91
[32] Rudé M, Pello J, Simpson R E, Osmond J, Roelkens G, Van Der Tol J J G M and Pruneri V 2013 Optical switching at 1.55 μm in silicon racetrack resonators using phase change materials *Appl. Phys. Lett.* **103** 141119
[33] Gong Z, Yang F, Wang L, Chen R, Wu J, Grigoropoulos C P and Yao J 2021 Phase change materials in photonic devices *J. Appl. Phys.* **129** 030902

[34] Ding W, Tang D, Liu Y, Chen L and Sun X 2010 Compact and low crosstalk waveguide crossing using impedance matched metamaterial *Appl. Phys. Lett.* **96** 111114
[35] Wu S, Mu X, Cheng L, Mao S and Fu H Y 2020 State-of-the-art and perspectives on silicon waveguide crossings: a review *Micromachines* **11** 326
[36] Feldmann J *et al* 2021 Parallel convolutional processing using an integrated photonic tensor core *Nature* **589** 52–58
[37] Feldmann J, Youngblood N, Wright C D, Bhaskaran H and Pernice W H P 2019 All-optical spiking neurosynaptic networks with self-learning capabilities *Nature* **569** 208–14
[38] Rhim S-Y, Ligorio G, Hermerschmidt F, Hildebrandt J, Pätzel M, Hecht S and List-Kratochvil E J W 2020 Using active surface plasmons in a multi-bit optical storage device to emulate long-term synaptic plasticity *Phys. Status Solidi* **217** 2000354
[39] Park J W 2010 Waveguide photodiode (WGPD) with a thin absorption layer *Adv. Opt. Photonic Devices* (London: IntechOpen) 161–72
[40] Horváth R, Niels S and Henrik B L 2005 Reverse Symmetry Waveguide for Optical Biosensing *Front. Chem. Sens.* (Berlin: Springer) 279–301
[41] Zhu S, Yu A W, Hawley D and Roy R 1986 Frustrated total internal reflection: a demonstration and review *Am. J. Phys.* **54** 601–7
[42] Mendive-Tapia D, Perrier A, Bearpark M J, Robb M A, Lasorne B and Jacquemin D 2014 New insights into the by-product fatigue mechanism of the photo-induced ring-opening in diarylethenes *Phys. Chem. Chem. Phys.* **16** 18463–71
[43] Herder M, Schmidt B M, Grubert L, Pätzel M, Schwarz J and Hecht S 2015 Improving the fatigue resistance of diarylethene switches *J. Am. Chem. Soc.* **137** 2738–47
[44] Pariani G, Quintavalla M, Colella L, Oggioni L, Castagna R, Ortica F, Bertarelli C and Bianco A 2017 New insight into the fatigue resistance of photochromic 1,2-diarylethenes *J. Phys. Chem.* C **121** 23592–8
[45] Huang J, Qin Q and Wang J 2020 A review of stereolithography: processes and systems *Processes* **8** 1138
[46] Pala R A, Shimizu K T, Melosh N A and Brongersma M L 2008 A nonvolatile plasmonic switch employing photochromic molecules *Nano Lett.* **8** 1506–10
[47] Rudé M, Simpson R E, Quidant R, Pruneri V and Renger J 2015 Active control of surface plasmon waveguides with a phase change material *ACS Photonics* **2** 669–74