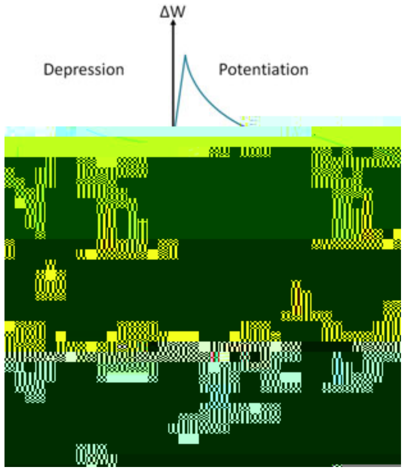

Fig. 1. (a) Illustration of the biological STDP characteristic. (b) Photonic STDP experimental setup.

neuron's inputs (presynaptic spikes) and outputs (postsynaptic spikes). In one of the two STDP outcomes, the presynaptic spike received by a neuron contributes to the firing of a postsynaptic spike, resulting in a strengthening, or "potentiation," of a synaptic connection. Alternatively, "depression" occurs when a neuron fires prior to receiving input from another neuron, resulting in a weakening of the synaptic strength between the two. In this section, a photonic implementation of STDP characteristics, utilizing nonlinear polarization rotation (NPR) and cross gain modulation (XGM) within a single semiconductor optical amplifier (SOA), is designed and experimentally demonstrated [17], improving upon photonic STDP circuits requiring multiple electro-optic devices [16].

The relatively simple experimental setup for optical STDP, consisting primarily of one SOA, two bandpass filters, and a polarization beam splitter (PBS), is shown in Fig. 1(b). A polarizer and polarization controllers (PCs) are used for initial polarization alignment prior to operation. The pre- and post-synaptic spikes are at different wavelengths, which are generated through use of a fiber laser and four-wave mixing [17]. The pulse serving as the presynaptic spike is at 1550.12 nm (λ_{pre}) with a repetition rate of 625 MHz (1600 ps period), solely for experimental pur-

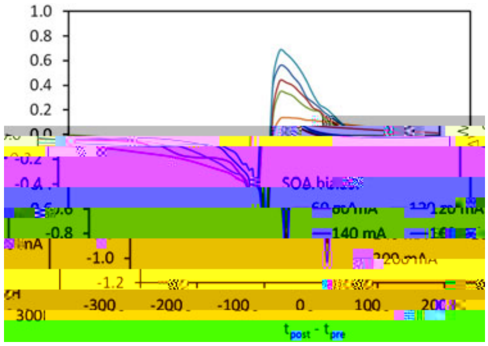

Fig. 3. Normalized photonic-based STDP characteristic.

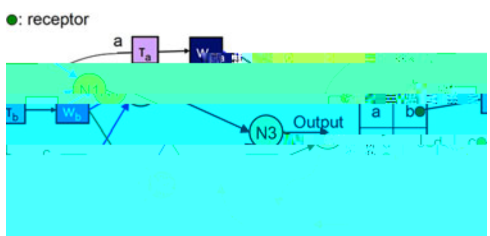

Fig. 4. Schematic illustration of supervised learning.

stronger NPR effect but also serve to restore the device's original birefringence more rapidly. The increasing NPR response and decreasing recovery time compete, and the potentiation window's width remains constant across all driving currents.

III. STDP for Supervised Learning

Learning based on experience is essential for a system to dynamically adapt to an unpredictable or changing environment, or one that is too complex to be characterized a priori. There are several different ways neurons can learn [19], and one such method, supervised learning, programs a system to perform a function defined in terms of its input-output pairs. Supervised learning relies on the use of a set of training pairs, where a sequence of training data is sent to the neuron's input and the corresponding desired output acts as the teacher. The neuron responds to the training input sequence, and the neuron response is compared with the teacher sequence. If the neuron response and teacher sequence are different, the synaptic strength between two neurons changes according to the STDP rule; otherwise, the synaptic strength is unchanged. After the learning phase, the teacher is removed, as the neuron is trained and ready to perform a specific task.

Using a photonic STDP circuit, we have demonstrated supervised learning in a photonic neuron in which a teacher determines the way that the photonic neuron should spike in response to its inputs [16]. Fig. 4 depicts a high-level schematic of supervised learning in photonic neurons. Both photonic neurons, N1 and N2, take several inputs (a-d) through receptors, which are weighted ($w_{a,}\ \ w_d$) and delayed ($\tau_{a,}\ \ \tau_d$) individually, and are temporally integrated at N1 and N2. An electro-optic modulator is used to convert the signal to optical spikes. Here, we are studying the synaptic plasticity between neuron N2 and N3

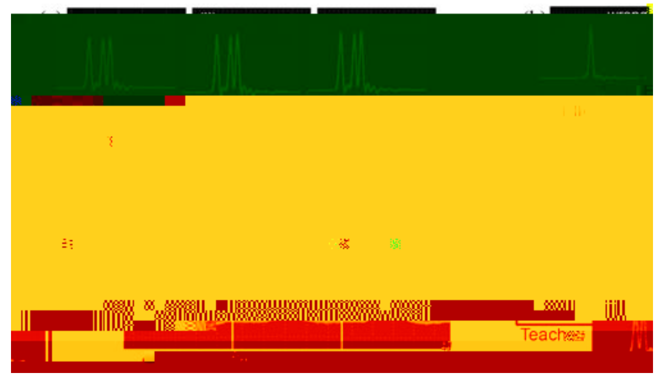

Fig. 5. Experimental results of automatic gain control in pulse-processing device (a photonic neuron) based on optical STDP. (a) Weighted output before integration at the photonic neuron. (b) Photonic neuron outputs. (c) Teacher signal. (d) Instruction signal from the STDP to indicate weight change. (i) Weight is too low. (ii) Weight is too high. (iii) Correct weight after learning.

by incorporating a photonic STDP circuit in the interconnection between N2 and N3. The initial weight w_e is at a random value, depending on the last operation of the photonic neuron, and is launched to the photonic neuron N3, which responds to the inputs depending on the current weight of the synaptic connection.

Since the initial weights are not optimized for detecting the desired signal, N3 is not spiking correctly in response to its inputs. Therefore, learning by means of the STDP circuit at the synaptic connection is desired. During the learning phase, N3's output, serving as a sequence of postsynaptic spikes, is compared with a teacher signal, serving as a sequence of presynaptic spikes, which represents the correct, desirable response to the neuron N3 inputs. If the photonic neuron is not spiking as the teacher is, the STDP circuit will send a signal, as illustrated by the red arrow in Fig. 4, indicating the necessary weight change to be applied to the synaptic connection. To enable a fast weight change response an electro-optic modulator is used as the variable weight device. The weight of each synaptic connection changes until the output matches what the teacher expects. After the learning phase is complete, the teacher is removed, and the photonic neuron still responds to its adjusted inputs in the desirable way as taught by the teacher.

Fig. 5 presents the experimental results of a supervised learning system. Fig. 5(a) depicts the weighted outputs before integration, 5(b) shows the photonic neuron outputs, 5(c) shows the teacher signal, and 5(d) illustrates the corresponding weight change needed as determined by the STDP circuit. Initially, in Fig. 5(a)i the interconnection strength is too weak, and the photonic neuron emits an undesirable output as shown in Fig. 5(b)i. According to the response at the photonic neuron output and the teacher [Fig. 5(c)], the STDP circuit will send a signal

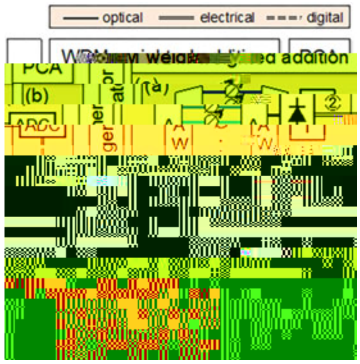

Fig. 6. Experimental setup of WDM weighted addition applied to PCA. (a) WDM weighted addition unit. (b) PCA algorithm, where ADC: analog-digital converter; CPU: central processor; DAC: digital-analog converter. Inset 1: Partially temporally and spatially correlated inputs. Insert 2: weighted sum of correlated channels.

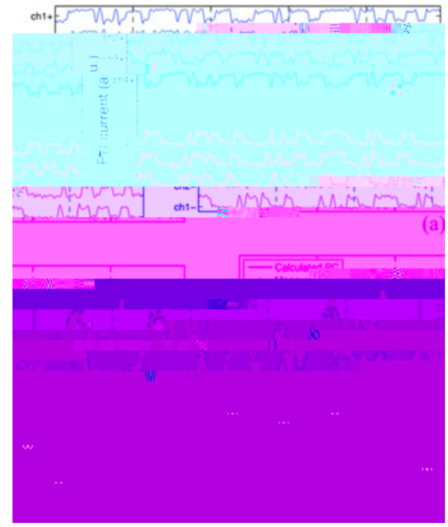

Fig. 7. Experimental results showing a 100 ns time window of a typical epoch. (a) A subset of partially correlated 4 positive input channels and their 4 negative complements on other wavelengths. (b) First principal component output as calculated by a software matrix decomposition-based PCA (red) compared to the measured output after convergence of the iterative algorithm (blue). Both calculated and measured PCA algorithms are applied to the measured inputs.

sending a signal [Fig. 5(d)] to adjust the weight. Fig. 5(a)iii and (b)iii show the photonic neuron output after it has learned from the teacher, indicating that the device has adjusted itself based on the STDP output and is spiking as expected by the teacher. After the learning phase the teacher is removed, and the processor continues to spike as the teacher instructed, as shown in Fig. 5(b)iii.

Photonic neurons are capable of processing information at billions of times a second, but a scalable learning scheme that allows such systems to adapt to a changing environment has yet to be well established. The development of the supervised learning capabilities presented here would enable a large range of new applications for photonics neurons, while taking advantage of the high bandwidths and processing speeds of photonics. Unlike prior approaches [15], our STDP learning system uses a single SOA, reducing the space occupied by a given STDP module. This increased scalability—especially if instantiated in an integrated platform—could enable complex learning operations, including principal component analysis (PCA) and independent component analysis (ICA).

IV. Principal Component Analysis with Optical STDP

The scalability of optical STDP would enable a number of useful algorithms. Here, we focus on a simple instantiation of principal component analysis (PCA) and describe how an STDP module can be implemented in such a system. principal component analysis (PCA) is a technique for unsupervised pattern recognition and dimensionality reduction of multidimensional random variables. In PCA the first principal component (PC) represents the original data in one dimension while maximizing the amount of "information explained."

To achieve PCA using a photonic neuron, a wavelength-division-multiplexing (WDM) weighted addition unit [20], [21] is developed. Weighted addition of an array of inputs is a fundamental component of neuron models. Microelectronic implementations of weighted addition suffer from interconnect limitations, which become more intractable as the dimension of the inputs increases. Digital electronic implementations, which use time-division multiplexing to accumulate summands, trade off bandwidth against fan-in (essentially the number of terms in the addition) [22]. By taking advantage of photonic bandwidths and by multiplexing many hundreds of signals onto a single waveguide, WDM is capable of removing that tradeoff and significantly improving interconnection performance. As explored in other works, WDM has enabled the experimental demonstration of weighted addition [23] with both wide bandwidth and scalable fan-in, and the experimental setup is illustrated in Fig. 6 [23]. In Fig. 6(a), the pair of array waveguide gratings (AWG) is used to separate the inputs and recombine them after weighting. The circles are variable optical attenuators, and their respected level of attenuation changes according to the PCA unit. A photodiode (PD) is used to convert the optical weighted signal back to the electrical domain. Inset (1) illustrates the partially temporally and spatially correlated inputs, and inset (2) shows the electrical output of the PD, which represents the weighted sum of correlated channels.

The proposed setup demonstrates for the first time a generalized Hebbian learning algorithm, implemented electronically, for synaptic modification, demonstrating that it converges iteratively to the first principal component of the inputs, as shown in Fig. 7. In this design, a CPU is responsible for computing correlations between postsynaptic and presynaptic signals, calculating the update rule, and controlling the weight bank, yielding a slow learning rate. The simple pair-wise operations required for the PCA controller described in the paper are feasible for a co-integrated microelectronic processor or other analog hardware, which could raise the learning rate to the hardware's bandwidth.

The Hebbian learning rule implemented in the described PCA system is known in the neuroscience field as activity-dependent synaptic plasticity (ADSP), introduced by Donald Hebb in

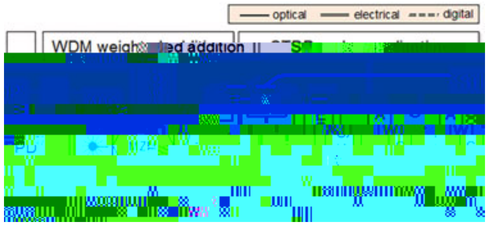

Fig. 8. Optoelectronic setup for scalable learning based on STDP. All-optical STDP circuits operating at the signal bandwidth enable Hebbian-type learning using only low-performance electronic hardware, which operates at the learning bandwidth ($<$ 10 MHz).

[24]. He suggested that synaptic connections should reinforce "when an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it." In our terminology, the axon of cell A represents one input (presynaptic) that is weighted and integrated by cell B, which in turn either releases or does not release an output spike (postsynaptic). This ADSP synaptic function, however, only assesses correlations between presynaptic and postsynaptic spike trains, rewarding positive correlations and penalizing negative correlations. Unfortunately, correlation does not imply causation. The STDP modification function, as described in Section II and implemented in Section III, extends ADSP, taking causation into account.

Neural models require independent tuning of their synaptic weights, which implies that each synaptic afferent requires an STDP module operating at a response time comparable to the picosecond timescales of inter-spike-intervals [25]. Since information is encoded in wavelength-carriers, these modules can be merged with the weighted addition optoelectronic circuit used for PCA and proposed in [23], as illustrated in Fig. 8. The output of the STDP function, which carries information of post/pre-synaptic synchrony, could be used by a lower module responsible for implementing a generalized Hebbian learning rule of the same sort as that used for PCA, shifting much of the processing to the optical domain. The latency of the weight controller (WC) loop defines the learning rate. Since the more complex nonlinear operation is already carried out by the STDP module, the controller's function is solely to translate a pulse to a weight update value. Like the PCA controller, the WC loop is feasible for current analog hardware technology.

Merging the weighted addition and STDP function iswhic(656r[(fat)-367.2(u6bp2A5.6(xtends)-666(xtends)-666(o4.93553 7.593553 754

[13] B. J. Shastri, M. A. Nahmias, A. N. Tait, A. W. Rodriguez, B. Wu, and P. R. Prucnal, “Dynamical laser spike processing,” arXiv:1507.06713, 2015.
[14] A. N. Tait, M. A. Nahmias, Y. Tian, B. J. Shastri, and P. R. Prucnal, “Photonic neuromorphic signal processing and computing,” in *Nanophotonic*