PHOTONICS Research

Equalization of a 10 Gbps IMDD signal by a small silicon photonics time delayed neural network

Emiliano Staffoli,* Mattia Mancinelli, Paolo Bettotti, and Lorenzo Pavesi

Nanoscience Laboratory, Department of Physics, University of Trento, 38123 Trento, Italy
**Corresponding author: emiliano.staffoli@unitn.it*

Received 12 December 2022; revised 10 March 2023; accepted 16 March 2023; posted 20 March 2023 (Doc. ID 483356); published 1 May 2023

A small 4-channel time-delayed complex perceptron is used as a silicon photonic neural network (PNN) device to compensate for chromatic dispersion in optical fiber links. The PNN device is experimentally tested with non-return-to-zero optical signals at 10 Gbps after propagation through up to 125 km optical fiber link. During the learning phase, a separation-loss function is optimized in order to maximally separate the transmitted levels of 0s from the 1s, which implies an optimization of the bit-error-rate. Testing of the PNN device shows that the excess losses introduced by the PNN device are compensated by the gain in the transmitted signal equalization for a link longer than 100 km. The measured data are reproduced by a model that accounts for the optical link and the PNN device. This allows simulating the network performances for higher data rates, where the device shows improvement with respect to the benchmark both in terms of performance and ease of use. © 2023 Chinese Laser Press

https://doi.org/10.1364/PRJ.483356

1. INTRODUCTION

Optical fibers are the backbone of the Internet since they allow data transmission at large bandwidths and long distances. To increase the capacity of the optical links, large input optical power signals are needed to compensate for fiber losses [1]. In these conditions of high-power transmission, both linear and nonlinear effects alter the shape of the transmitted optical pulses [2], which implies the necessity of distortion compensation in the optical network. Today, signal recovery (equalization) is mostly accomplished by digital devices that introduce latency, delay, and power consumption [3]. A clear example is observed in the trend to replace simple intensity-modulation direct-detection (IMDD) transceivers with more high performing but costly and power-hungry coherent transceivers [4], where digital signal processing (DSP) devices allow running algorithm to restore the data [5]. Different numerical approaches to correct for both linear and nonlinear optical fiber impairments exist with an emerging trend to use artificial intelligence-based algorithms [6].

To reduce the cost and power consumption of optical links, it is desirable to introduce equalization techniques also for simple IMDD systems. Even linear impairments, such as chromatic dispersion (CD), polarization mode dispersion (PMD), symbol timing offset, and optical filtering, severely distort the transmission [1]. Among these impairments, one of the most severe is CD, which causes a broadening of the optical pulse and the associated intersymbol interference [7]. To compensate or correct for CD, several types of equalization techniques have been introduced, among which dispersion compensated optical fiber and Bragg gratings are the most diffused ones [1]. These are based on the use of dispersion-compensated units, which recover the initial undispersed signal by counter-acting the CD effect. Another approach relies on the use of a dispersion compensating photonic-integrated programmable lattice filter formed by cascaded Mach–Zehnder interferometers [8]. An alternative to these approaches is the use of integrated photonic neural networks [9]. Their advantages derive from operating the corrections directly in the optical domain, drastically reducing the power demand and the latency, as well as in the flexibility of the equalization, which can be learned directly on the deployed link and, as a result, can be easily adapted to optical link variations. Few hardware implementations of this concept exist [10–16].

Here, we propose and validate the use of a small silicon photonics 4-channels delayed complex perceptron [17] to equalize a 10 Gbps IMDD 100-km-long optical link. In the proposed photonic neural network (PNN), the input signal is split into 4 channels where the combined actions of delay lines and tunable phase shifters create the desired interference pattern at the output that counteracts the intersymbol interference. This working principle has been applied to compensate for distortions induced by linear effects during propagation in a single-mode fiber. Equalization is performed on-chip and no external data processing is thus needed, except for the training phase. PNN

training is based on a particle swarm optimizer (PSO) [18]. Moreover, being the PNN of the feed-forward type [17], the latency induced in signal processing is maximally reduced.

2. PROCEDURES

The small PNN device, whose design is shown in Fig. 1, is based on a delayed complex perceptron [17]. The input signal [$u(t)$, the input complex field] is split into four waveguides by a cascade of 1×2 multimode interferometers (MMIs). On each kth waveguide, but only on the first ($k = 1$), a spiral forms a delay line that adds a delay of $\Delta_k = (k-1)\Delta_t$ to the input signal where $\Delta_t = 50$ ps has been determined from the signal bitrate that, in this case, is 10 Gbps in the non-return-to-zero (NRZ) modulation. After the delay stage, the kth waveguide hosts a delayed copy of the input $u(t)$, i.e., $u_k(t) = u[t - \Delta_t(k-1)]$, with $k = 1, \ldots, 4$. Now, the signal undergoes a phase modulation performed by phase shifters realized with current-controlled heaters. In this way, the signals in each waveguide are weighted with $w_k = a_k \exp(i\phi_k)$, where a_k stands for the spiral losses and ϕ_k for the added phase. After the weighting section, the four signals are recombined by means of a 1×4 combiner, realized using a cascade of 2×1 MMIs, which performs the operation $\sum_{k=1}^{4} u_k(t)w_k$. The output signal is then detected by a fast photodetector, which closes the processing by performing a nonlinear transformation, i.e., the detected signal intensity is $y(t) = |\sum_{k=1}^{4} u_k(t)w_k|^2$.

The delayed complex perceptron acts as a 4-tap filter. The complexity of the layout in terms of the number of taps (N_T) and the delay unit (Δ_t) is determined in relation to the input bitrate (B) and the target propagation distance (L). This relation can be empirically described as

$$N_T = \mathrm{int}\left(\frac{1/B + |L\beta_2\Delta\omega|}{\Delta_t}\right). \quad (1)$$

Here, the numerator represents an estimate of the new pulse width, obtained as the sum of the initial bit time slot ($1/B$) and the pulse broadening (ΔT) induced by CD on a Gaussian pulse propagating in a fiber [1]. β_2 represents the group velocity dispersion parameter and $\Delta\omega$ the pulse bandwidth. Substituting in Eq. (1) the parameters for the propagation of a 10 Gbps NRZ pseudo-random binary sequence (PRBS) ($\Delta\omega \approx 2\pi \times 10$ GHz) through an $L = 100$ km long standard SM G.652D fiber ($\beta_2 = -0.021$ ps^2/m), one obtains a pulse broadening of $\Delta T = 130$ ps, that for $\Delta_t = 50$ ps corresponds to $N_T \approx 4$. The choice of Δ_t is the result of a trade-off between a sufficient sampling of the information of a single-bit time slot at recombination (at least 2 samples per bit) and the aim of having a restricted number of channels to contain the excess losses.

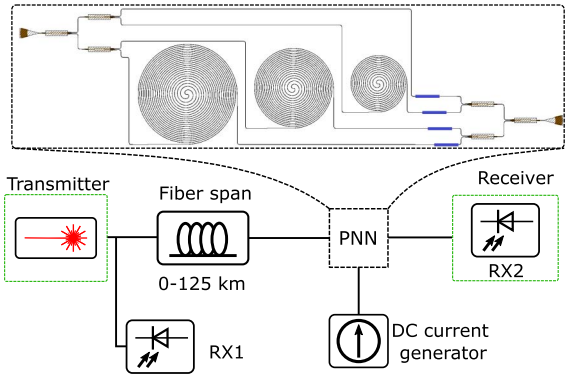

Fig. 1. Experimental setup. The full link consists of a transmission stage, the optical fiber, the photonic neural network (PNN) device, and the receiver stage. Two fast photodetectors (RX1 and RX2) allow for measuring the input and the transmitted signals. The inset shows the actual design of the PNN device, where one can observe the cascaded 1×4 and 4×1 splitter and combiner, the three spirals, and the four phase shifters (small blue rectangles) connected to the external DC current controller. Details are given in Appendix A or in Ref. [17].

The PNN device is fabricated on a silicon-on-insulator (SOI) platform within a multi-project wafer run at IMEC-Belgium. The waveguides are 220-nm thick and have a width of 450 nm, allowing for single-mode operation on both polarizations at 1550 nm. The input and output gratings have a footprint of 50 μm × 30 μm and fix the polarization to the transverse electric (TE). The 1×2 and 2×1 MMIs used for splitting and recombination of the optical signal have a footprint of 20 μm × 100 μm. The gratings and the MMIs are implemented using IMEC's proprietary process design kits (PDKs). Phase shifters are based on current-driven heaters realized as 60-μm-long and 0.6-μm-wide with a resistance of 60 Ω placed on top of an 800-nm-thick silica cladding. Δ_k are realized with spirals of a length kth multiple of 3.56 mm (corresponding to a delay of $\Delta_t = 50$ ps). The optical losses of the spirals due to surface roughness present on the waveguide and to the bends in the curved optical paths have been measured to be 6 dB/cm. These result in an attenuation of 2.1 dB, 4.3 dB, and 6.4 dB for $k = 2, \ldots, 4$, respectively. The PNN device's insertion losses have been estimated to be 8.2 dB at 1550 nm. The chip is placed on a proportional-integral-derivative-controlled Peltier cell that keeps its temperature at 21°C.

The experimental setup is represented in Fig. 1. In the transmission stage, a tunable laser source (TLS) operating at 1550 nm is modulated as an NRZ 10 Gbps PRBS of order 10 and a period of 2^{10} bits. A 50:50 fiber optic splitter sends half of the signal to a fast photodiode (RX1). The other half is coupled to an optical fiber span, where distortions induced by CD are accumulated. The length of the span goes from a minimum of 0 km to a maximum of 125 km, with a granularity of 25 km. The distorted signal enters the PNN device for optical processing. DC current controllers set the currents in the heaters. The output signal from the PNN device is coupled to a fast photodiode (RX2) at the receiver stage. Both RX1 and RX2, which monitor $y_{\mathrm{in}}(t)$ and $y_{\mathrm{out}}(t)$, respectively, are connected to a 40 GSa/s oscilloscope with a 16 GHz bandwidth. The signal-to-noise ratio (SNR) at receiver RX2 can be varied by using a variable optical attenuator (VOA2) inserted after the PNN device (see Appendix A).

For each measurement, the DC controller sends preset currents to the PNN device and a triggering signal to the oscilloscope. The acquisition is delayed by 1 ms from the arrival of the triggering signal to let the optical signal stabilize at the output of the PNN device, according to the thermal relaxation time of the heaters (few tens of μs). The observation window of the

oscilloscope is 1 μs wide, allowing for the observation of at least 9 periods of PRBS at each acquisition. Four samples per bit are available because of the 40 GSa/s sampling rate. In what follows, the samples in each bit are labeled from 1 to 4, the 4th being the most recent. Acquired sequences are then under-sampled, obtaining a sub-sequence constituted by the nth sample in each bit of the full trace and being the chosen sample that is the most representative of the actual bit value (typically, the closest to the center of the bit). An operation performed over the under-sampled sequence at the nth sample in each bit will be shortly referred to as performed over the nth sample.

Input (from RX1) and output (from RX2) signals are aligned by exploiting their cross-correlation, obtaining y_{in} and y_{out}. The undersampling at the nth sample of the two sequences provides $\bar{y}_{\text{in}}$ and $\bar{y}_{\text{out}}$, respectively. The output signal $\bar{y}_{\text{out}}$ is compared with $\bar{y}_{\text{in}}$ to label the 1 level ($\bar{y}_{\text{out},H}$) or 0 level ($\bar{y}_{\text{out},L}$).

The PNN training procedure is performed off-chip using fully automatized software. An analog loss function $\mathcal{L}$ is created to obtain the largest possible separation between the distributions of signal levels expected as 1s or 0s in the output signal. This quickly minimizes the associated bit-error-rate (BER) since it is directly linked to the overlap between these two distributions. Indeed, in the presence of random Gaussian noise characterized by standard deviations σ_0 and σ_1 affecting 0s and 1s in the bit sequence, the BER can be computed as [1]

$$\text{BER} = \frac{1}{4}\left[\text{erfc}\left(\frac{I_1 - I_D}{\sigma_1\sqrt{2}}\right) + \text{erfc}\left(\frac{I_D - I_0}{\sigma_0\sqrt{2}}\right)\right]. \tag{2}$$

Here, $I_{(0,1)} = \langle \bar{y}_{\text{out},(L,H)} \rangle$ are the average levels for 1s and 0s, $\sigma_{(0,1)}$ are their standard deviations, I_D is the decision threshold, and erfc is the complementary error function. A BER reduction can thus be obtained by maximizing $I_1 - I_0$. $\mathcal{L}$ measures the spacing between the tails of the distributions related to $\bar{y}_{\text{out},H}$ and $\bar{y}_{\text{out},L}$. The training's goal is the maximization of this spacing. Therefore, we call it the separation loss function. Having $\bar{y}^i_{\text{out},(L,H)}$, the measured signal values in a sequence, the separation loss function is expressed as

$$\mathcal{L} = E[0] - E[1] = \frac{1}{N_L}\left(\sum_{i=1}^{N_L} \bar{y}^i_{\text{out},L}\right) - \frac{1}{N_H}\left(\sum_{i=1}^{N_H} \bar{y}^i_{\text{out},H}\right), \tag{3}$$

where $E[0]$ and $E[1]$ are estimates of the tail position in the two distributions. In $E[0]$, i runs over the samples such that $\bar{y}^i_{\text{out},L} > I_0 + 1.28\sigma_0$, i.e., $\bar{y}^i_{\text{out},L}$ is part of the group of the rightmost N_L points corresponding to the 10% of the population of the $\bar{y}_{\text{out},L}$ distribution. Similarly, $\bar{y}^i_{\text{out},H} < I_1 - 1.28\sigma_1$ is part of the group of the leftmost N_H points corresponding to the 10% of the population of the $\bar{y}_{\text{out},H}$ distribution. The PSO is adopted for training [18], which is performed in a condition of no attenuation in front of RX2, i.e., an average optical power at RX2 of about 0 dBm and an SNR of 11.2 dB.

In light of the differentiability of $\mathcal{L}$ with respect to the currents controlling the induced phase shifts in the device, other choices for the training algorithm are possible, including a back-propagation (BP) technique. During the experimental phase, we performed some tests using an adapted version of the Adam algorithm [19], which is a gradient-based alternative in which the descent proceeds with a memory of the previous iterations. Such weighted adaptation of the gradient is well known for making the trajectory towards the local minimum in the presence of noise more robust and is often preferred over the standard BP algorithm. The algorithm proved to be more time efficient but is possibly limited by premature endings of the research at a local minimum. Therefore, here we chose to rely on the PSO, which guaranteed the robustness and the repeatability of the final outcomes.

After the training phase, the testing phase is performed via a scan over the power at the receiver (PRX) made by varying the attenuation of VOA2 in front of RX2, which corresponds to a scan over the SNR at the RX2. For each PRX value, 100 acquisitions for a total of 6×10^6 bits with the trained currents set are performed, evaluating the BER for each measure. The BER is defined here as the cumulative error between the digitized input and output signals. The digitized signals are obtained by applying a threshold to $\bar{y}_{\text{in}}$ and $\bar{y}_{\text{out}}$. At each evaluation, the optimal sample for the generation of $\bar{y}_{\text{out}}$ and the optimal threshold, which minimize the BER, are selected. The threshold is chosen among 10 possible equally spaced levels between the minimum and maximum of the signal. Training and testing procedures are performed for multiple lengths of the fiber span and then compared with the corresponding reference curves obtained without the PNN device.

The full optical link (from the transmission to the receiver stages) is simulated to model the effect of the PNN device (see Appendix B). Also in the simulation, the PNN's training is performed by optimizing the separation loss function with the PSO. Noise is added, as described in the Appendix C. The sampling of the oscilloscope is modeled as well. The BER is computed as in the experimental case.

3. RESULTS

The equalization effect of the PNN device for a span of 125 km is summarized in Fig. 2. The eye diagrams show the three aperture conditions reached after the modulation at the transmitter [Fig. 2(a)], after the fiber propagation [Fig. 2(b)], and after the equalization performed by the PNN device [Fig. 2(c)], respectively. As a consequence of the intersymbol interference, the CD generates a closure of the eye diagram. Particularly evident in Fig. 2(b) is a high-density region between the normalized amplitude values of 0.3 and 0.4 crossed by the red dashed line, which represents a rise of the zero-level induced by the interference of a low bit with neighboring bits in the high state. The action of the PNN device partially restores the aperture [Fig. 2(c)], eliminating the intermediate level seen in Fig. 2(b). The same scenario is presented in Figs. 2(d)–2(f), where the histograms report the distributions of the optical power levels expected as 0s or 1s associated with the 2nd sample in the bit in the input [Fig. 2(d)], non-corrected [Fig. 2(e)], and corrected [Fig. 2(f)] output signals, respectively. An example of their time evolution is reported in Fig. 2(g) with normalized amplitudes. Data are collected with an SNR = 11.2 dB at RX2. In this regime, the evaluation of the BER is not limited by the SNR but by the fiber dispersion that generates an intersymbol interference. The distorted output in Fig. 2(g) clearly shows the presence of pulse broadening and the consequent generation of

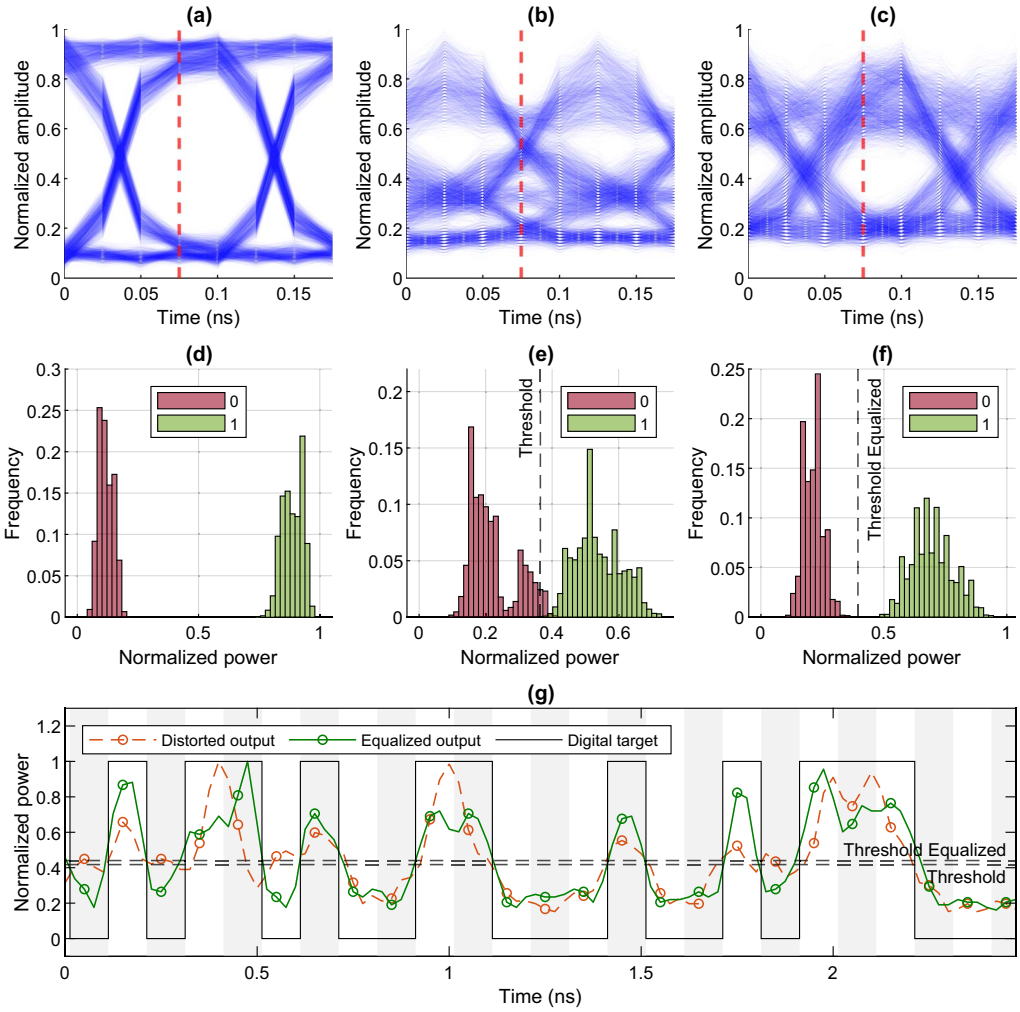

Fig. 2. Testing of the photonic neural network device on a 125 km fiber link. (a)–(c) Eye diagrams of (a) the input signal, (b) the transmitted signal without the PNN device in the link, and (c) the transmitted signal with the trained PNN device in the link. The dashed red lines refer to the sampling time that is used to generate the histograms. (d)–(f) Histograms of the power levels associated with the 2nd sample in the input bit sequence for (d) the input signal, (e) the transmitted signal without the PNN device in the link, and (f) the transmitted signal with the trained PNN device in the link. The red and green columns refer to the input 0s and 1s, respectively. The dashed vertical lines set the decision threshold. The sub-optimal choice of the threshold in (e) derives from discretization of the possible threshold values. (g) The actual temporal sequences for the transmitted signal as recorded by the RX2 (line with the PNN device in the link, dashed line without). The light black line refers to the digital input. Circles highlight the 2nd sample in each bit that is used in the BER calculations. Horizontal dashed lines show the different thresholds for the equalized and non-equalized data used to minimize the BER. Threshold values are rescaled according to the normalization used to plot the curves.

an intersymbol. Bits expected as 0s and preceded or followed by a 1 are raised close to the 1s, thus increasing the probability for errors. As a consequence, the distributions of power levels for 0s and 1s widen, and the gap between the distributions reduces, as shown in Fig. 2(e). This leads to an increased BER. As clear from Fig. 2(f), the corrective action of the PNN device partially restores the two distorted distributions of Fig. 2(e), thus decreasing the BER.

The training returns a set of 3 optimal currents, associated with the channels in the PNN device. One can then model the relative recombination phase shift ϕ_k (with $k = 2, \ldots, 4$) used for the weight w_k in the kth channel with respect to the first channel (chosen as reference) as

$$\phi_k = \phi_k^0 + i_k^2 \gamma_k, \tag{4}$$

where ϕ_k^0 is the relative phase measured at zero currents, i_k is the optimal current in the kth channel, and γ_k is the conversion factor between the dissipated thermal power in a resistor and the induced phase shift in the underneath waveguide. Measurements conducted on test resistor structures yield $\gamma_k \approx 0.01$ rad/mA^2 [20], while finding ϕ_k^0 is cumbersome due to the uncertainties in the optical path lengths and widths caused by the finite fabrication resolution. Therefore, for the sake of clarity, we show in Fig. 3(a) the currents used for the trained PNN at different fiber lengths and in Fig. 3(b) the corresponding phase shifts obtained from the simulation. In Fig. 3(b) it appears that longer optical links require an increase of the phase shift to about 2π in each channel, meaning that the delayed copies constructively contribute to the output [17]. Thus, the PNN device weights more associated to the

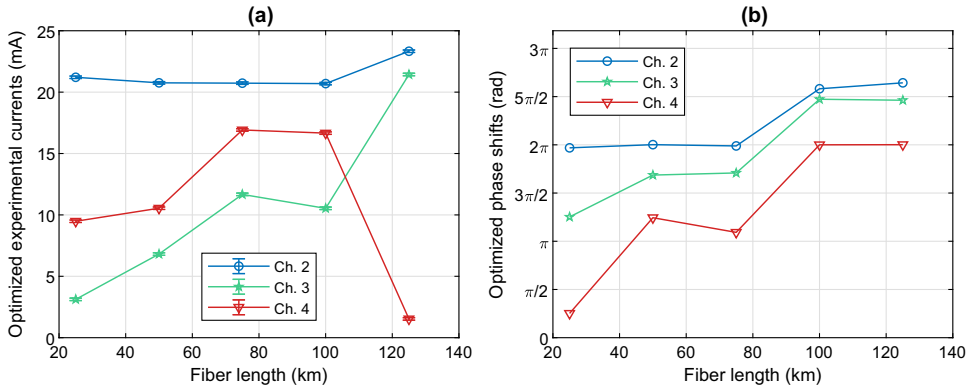

Fig. 3. Training outcomes for signal equalization as a function of the fiber link length. (a) The optimized experimental currents and (b) the simulated optimized relative phase shifts in channels 2 (blue dots), 3 (green stars), and 4 (red triangles) after the training. Error bars in (a) (barely visible) derive from the instrument output precision, not from statistics. The phase shifts in (b) are measured in each channel with respect to the 1st channel (no spiral). PSO is chosen as the training algorithm in both cases.

longer delay lines since a larger pulse broadening is to be compensated.

Figures 4(a)–4(c) report the simulated and experimental BER versus the PRX profiles obtained for different fiber spans. The back-to-back (BTB) configuration, i.e., with no PNN device and no fiber (black curve) measures the TX/RX performance. For the low PRX values, i.e., a low SNR, the BER is dominated by noise, which is present in the output signal at the receiver regardless of the length of the fiber. Thus, all the profiles overlap in this region. On the contrary, for a higher PRX,

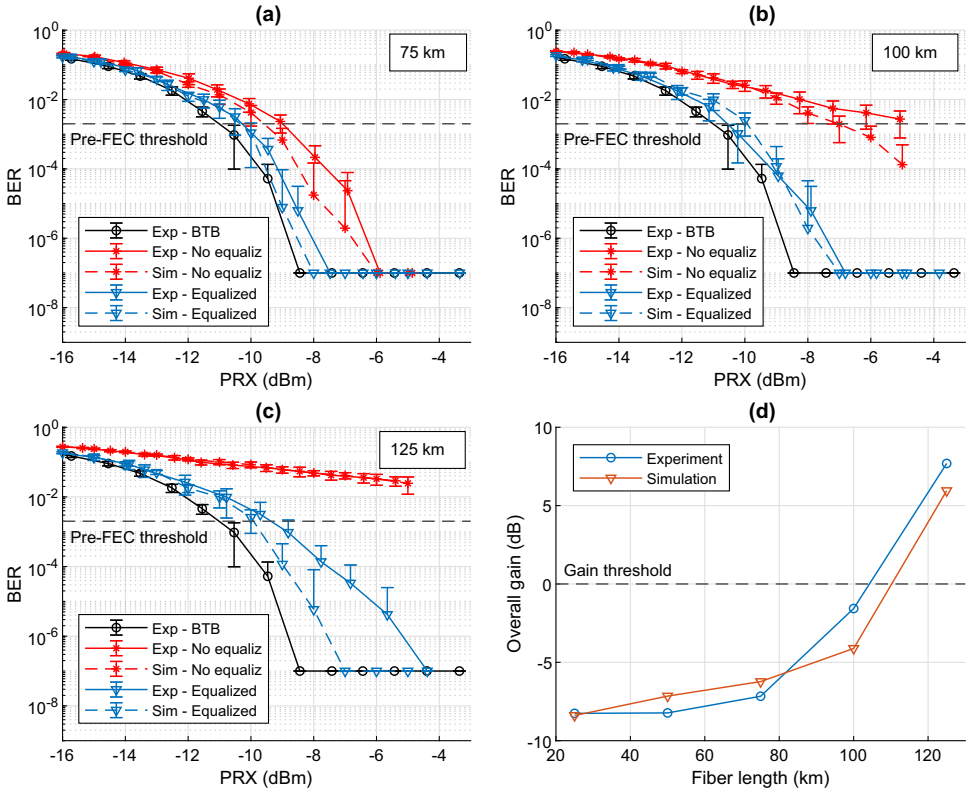

Fig. 4. (a)–(c) Experimental (full lines) and simulated (dashed lines) BER versus the PRX curves: the black discs refer to the back-to-back (BTB) configuration (the transmission stage is directly interfaced to the receiver), the red stars to the transmission by a fiber link, and the blue triangles to the transmission by a fiber link with the PNN device. Dashed horizontal black lines refer to the pre-FEC threshold value. The error bars are calculated as the standard deviation over multiple acquisitions (see Appendix A). The error-free points are replaced by 1×10^{-7} due to the finite dimension of the data set. The used fiber link is (a) 75 km long, (b) 100 km long, or (c) 125 km long. (d) The overall gain provided by the PNN device as a function of the fiber link length (blue discs experiment, orange triangles simulation). The gain is given as the improvement of the PRX at a BER $= 2 \times 10^{-3}$ when the PNN device is used with respect to the results without the PNN device. The dashed line marks the threshold above which the gain, guaranteed by the equalization, is greater than the PNN device excess loss of about 8.5 dB.

the SNR increases, too, and the most dominant contribution in the BER is provided by distortions in the signal induced by cumulated chromatic dispersion. These distortions become more important for longer fibers, causing a worsening of the BER even at a high PRX. In fact, the dispersion length for this system is $L_D = T_0^2/|\beta_2| = 2\pi c_0 T_0^2/(\lambda^2|D|) = 450$ km, using $D = 17.2$ ps/(nm km) and $T_0 = 100$ ps. The effects of the corrections operated by the PNN device are evident for long fiber lengths (≥100 km) when the amount of distortion to be compensated is significant. The PNN device almost recovers the BER versus the PRX curves to the reference optimal case (BTB).

The gain brought in by the action of the PNN device can be described starting from the PRX values corresponding to the same BER in the experimental curves. The reference BER value is considered to be 2×10^{-3}, being a typical BER threshold value for pre-forward error correction (pre-FEC threshold) [21]. The corresponding PRX values are interpolated for each BER versus the PRX profile obtained both with and without the PNN device, producing PRX(w) and PRX(w/), respectively. Figure 4(d) reports the corresponding experimental and simulated overall gain obtained as PRX(w) – PRX(w/) subtracted with the excess loss (EL) introduced by the trained PNN device. Note that the EL depends on the actual weights configuration since the output signal results from the interference of the weighted and delayed copies of the input [17]. We use the best-case scenario, and we neglect in the EL calculations the 8.2 dB contribution of the grating losses. Note also that the values of PRX(w/) for 100 km and 125 km fiber spans are extrapolated from the corresponding BER versus the PRX since no data at the pre-FEC threshold are available. The horizontal line at the null value highlights the point where the gain generated by the PNN device compensates for its excess loss. This happens for fiber lengths above 100 km for the used bit rate. The PNN device has to be considered as underperforming, being that the 6 dB/cm spiral propagation losses are unusually higher than the expected nominal value of 2 dB/cm for IMEC processing [22]. Improvements in the fabrication could further increase the performance of this already working PNN device.

4. CONCLUSIONS

The model of the PNN device allows accessing working conditions that are not explorable with the present integrated version of the PNN device. Indeed, its versatility is limited by the fixed delay lines that are set for a 10 Gbps data rate. On the contrary, the simulations allow adopting a higher modulation frequency by adapting the delay lines to different bit rates. We can thus compare the performance of our PNN device with the results obtained in Refs. [10,12], which can be considered a benchmark for the current state-of-the-art short reach (up to 25 km) access link applications. The first approach [10] is based on the reservoir computing paradigm where a photonic integrated circuit composed of delay lines and beam splitters arranged in a swirl topology forms the reservoir. The second approach [12] is based on the spectral decomposition technique where the spectral content of the optical carrier is divided into slices and analyzed following an all-optical/hybrid approach.

The comparison starts by tuning the parameters of our simulation in order to reproduce the BER versus the fiber length profile of Ref. [10] at 40 Gbps. In particular, the SNR has been fixed to 12 dB, and each BER value is obtained as an average over 1.024×10^6 transmitted bits. The parameters are kept unaltered for the other runs, too, including the training and subsequent testing phase for the PNN device. The PNN has been modeled with delay lines introducing a shift of half (12.5 ps) or three-quarters (18.75 ps) of a bit. To compare our PNN device with the ones in Refs. [10,12], we used the same representative performances as in these works. First, the BER as a function of the link length for the PNN device at 40 Gbps is reported in Fig. 5(a). A clear BER improvement is observed where the equalization provided by the trained PNN device ensures an extension of the link reach up to almost 20 km when a delay of 18.75 ps is used. A comparison with the results in Ref. [10] shows that the present PNN device provides better

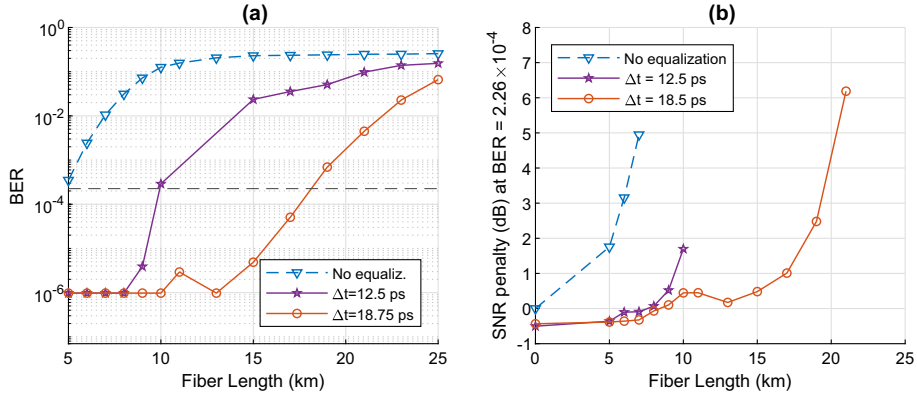

Fig. 5. (a) BER versus the link length with a SNR = 12 dB at the receiver for the link without the PNN device (dashed blue line and triangles), with the PNN device and a delay granularity of 12.5 ps (full purple line and stars), and with the PNN device and a delay granularity of 18.75 ps (full orange line and circles). (b) The SNR penalty at a BER of 2.26×10^{-4} as a function of the link length without the PNN device (dashed blue line and triangles), with the PNN device and a delay granularity of 12.5 ps (full purple line and stars), and with the PNN device and a delay granularity of 18.75 ps (full orange line and circles). The penalty is calculated from the back-to-back performance. Curves are interrupted at the last fiber length value for which it was possible to interpolate the chosen BER threshold in the corresponding BER versus the PRX profile.

BER performances up to 20 km fiber length. Despite its simplicity, the PNN device outperforms the swirl-based reservoir without the need for electrical data post-processing [10].

In Ref. [12], the SNR penalty is used as a figure of merit. This is defined as an increase in the SNR needed to achieve the same BER as that of the BTB configuration and calculated at the pre-FEC threshold of 2.26×10^{-4}. The SNR penalty of our trained PNN device for an NRZ 40 Gbps data rate as a function of the fiber link length is shown in Fig. 5(b). When the PNN device is used with a delay line granularity of 18.75 ps, the SNR penalty stays below 1 dB up to a link length of 18 km. Compared to the performances of the devices discussed in Ref. [12], the present PNN device is doing better than the 1- and 2-stage fully optical devices but worse than the 4-stage fully optical devices, which have, however, a significantly larger complexity (it requires 30 Mach–Zehnder interferometers) than our PNN design.

The performances of our device validate its use for signal equalization, suggesting further studies for the optimization of the layout for in-line applications. We foresee next-generation devices equipped with an augmented number of channels and amplitude modulators in each tap to allow for much larger adaptability to the different transmission scenarios (bitrate, modulation format, etc.). A transceiver-packaged version of these optimized devices would provide significant advantages even at high modulation frequencies (up to 100 Gbps) at metro propagation distances (up to 100 km). In addition to a latency reduction, these in-line transceivers relieve the computational efforts of complex DSPs both in coherent and IMDD systems. The most important for short-reach applications is a significant reduction in power consumption, which for the present PNN accounts for 70 mW, to be compared with the typical >1 W for DSP (Table 19.1 in Ref. [23]). Thus, simplified DSPs (e.g., less power hungry) will be required to achieve the same BER over longer distances without reducing the carrier frequency.

In summary, we demonstrated a simple concept of a feed-forward neural network device that is able to correct linear signal distortion both on a metro network (10 Gbps, 100 km) and on a high-speed short-reach access link (40 Gbps, 20 km). For different applications that have different data rates, proper tuning of the nodes' delays is needed.

APPENDIX A: EXPERIMENTAL SETUP

Figure 6 presents the experimental setup. The tunable laser source (TLS) is constituted by an InGaAs-based semiconductor laser that can be thermally tuned to around 1550 nm. The source is modulated by a nested Mach–Zehnder interferometer (NMZI) with 30 GHz of electro-optical bandwidth and driven by an arbitrary waveform generator (AWG) with 30 GHz of electrical bandwidth and a sampling rate of up to 64 GSa/s. After the modulation stage, a fiber optic coupler with 50% coupling ratio addresses half of the optical signal to a fast photodiode (RX1, 20 GHz bandwidth), which detects the input signal. A polarization controller allows the tuning of the local compression and torque applied to the fiber itself, inducing a polarization change.

An erbium-doped fiber amplifier (EDFA1) amplifies the optical signal to 20 dBm and then a variable optical attenuator (VOA1) controls the effective power level launched into the fiber link made of an SM G.652D fiber with a nominal loss coefficient of 0.2 dB/km. The fiber link length is varied during the experiments from a minimum of 0 km to a maximum of 125 km with a granularity of 25 km. A semiconductor optical amplifier (SOA) with a small signal gain of 13.4 dB is inserted at the end of the fiber link to partially recover the fiber link attenuation. The amplified optical signal is then sent to a switch that allows for addressing the optical signal to an optical spectrum analyzer (OSA) or to the input grating of the PNN device, respectively. Here, the optical signal is processed by the PNN and, via the output grating coupler, is coupled to the output fiber. Currents sent to the PNN device are provided by a terminal-controlled DC current generator.

The output fiber is connected to a fiber optic coupler with a 99.9:0.1 coupling ratio to address 0.1% of the optical signal to a power monitor (PM1). The other 99.9% is sent to a second EDFA (EDFA2) with a small-signal gain of 30 dB followed by a second VOA (VOA2). The combined action of these last two elements regulates the amount of optical power detected to a level below the damage threshold of the fast photodiode RX2. Then, a tunable optical filter with 30 GHz bandwidth and 5 dB insertion loss cleans up the signal from the out-of-band amplified spontaneous emission noise added by the amplification stages. Another fiber optic coupler with 99.9:0.1 coupling ratio

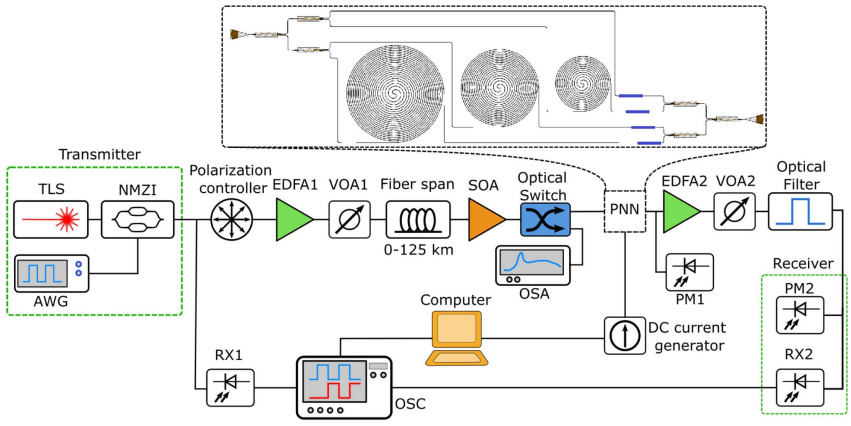

Fig. 6. Experimental setup. The different symbols are self-explanatory and are discussed in the text. The inset shows the design of the PNN device.

splits the signal towards a second power monitor (PM2) and to another fast photodiode (RX2, 20 GHz bandwidth), which measures the output signal. Both the RX1 and RX2 fast photodiodes are connected to a 40 GSa/s oscilloscope (OSC) with a 16 GHz bandwidth. The bandwidth limit of the transmission line is thus fixed by the oscilloscope, having the narrowest bandwidth in the line.

For each measure, the DC current generator sends pre-set currents to the PNN device and a triggering signal to the oscilloscope. The acquisition is delayed by 1 ms from the arrival of the triggering signal to stabilize the PNN device response, according to the thermal relaxation time of the heaters. The observation window of the oscilloscope is 1 μs wide, allowing the observation of at least 9 periods of PRBS at each acquisition. Four samples per bit are available because of the 40 GSa/s sampling rate. Each point in the BER versus the PRX profiles (as those in Fig. 4) is obtained as an average over $N = 100$ measurements, which obtains a minimum non-null measurable value of $1/(N \times$ number of bits in the sequence). Error bars are obtained as the standard deviation of the measured BER values for that point.

APPENDIX B: MODELING OF THE EXPERIMENT

The main elements of the optical link have been modeled, simulating the modulation apparatus, the structure of the PNN device, the propagation of the optical PRBS in the fiber link, and the noise contributions at the receiver. In the following, the reported numerical values for the parameters refer to the simulation performed at 10 Gbps (40 Gbps).

The model includes a Mach–Zehnder modulator driven by an electrical signal to imprint a PRBS of order of 10 and a period of 2^{10} bits on the optical carrier with an analog bandwidth of 30 GHz, an extinction ratio of 13.9 dB, and a null chirp. In order to preserve the information over a sufficiently large frequency range, the sampling frequency of the electrical signal is fixed to 320 GSa/s. Thus, the same sampling frequency is also maintained in the resulting modulated optical signal propagating across the simulated setup until it reaches the final detection process and the simulated oscilloscope, which reduces the sampling frequency to 40 (160) GSa/s.

The evolution of the signal is simulated by solving the linear Schrödinger equation in the Fourier domain. This can be derived following the approach proposed in Ref. [1], which reduces to

$$\tilde{A}(z,\omega) = \exp\left(iz\frac{\beta_2}{2}\omega^2 - \frac{\alpha}{2}z\right)\tilde{A}(0,\omega), \qquad \textbf{(B1)}$$

where $\tilde{A}(z,\omega)$ is the Fourier transform of the temporal optical field envelope, z is the propagation distance, β_2 is the group-velocity dispersion (GVD) parameter, and α stands for the fiber losses. The result of this operation is the propagated temporal optical field envelope $A(z,t)$, obtained by applying the inverse Fourier transform to $\tilde{A}(z,\omega)$.

This complex optical field signal is then provided as input to a model of the PNN device. The model simulates the action of four delay lines, each of them associated with a fixed attenuation value measured for the actual spiral length in the PNN device. Before the output combiner, a tunable phase shift is applied to each channel to simulate the action of the heater-actuated phase shifters. Note that Eq. (B1) was not used to simulate the signal propagation inside the PNN device, since the dispersion length $L_D \approx 2$ km [1] associated with the spirals is much longer than the length of the spirals themselves (∼1 cm). The effects deriving from chromatic dispersion can thus be neglected inside the PNN device. Therefore, the relative delay between the 4 channels has been emulated by inserting a shift of the proper number of samples between the 4 sequences.

After the combiner, the complex optical field signal is converted into the detected optical power (output signal) through the modulus square operation and then is treated to account for the noise measured experimentally at the receiver (noise modeling is discussed in the next section). A band-pass filter with a bandwidth of 16 GHz (28 GHz) obtained with a 5th-grade Bessel polynomial is then applied to the detected output signal, simulating the electronic bandwidth of the oscilloscope. An 8-bit vertical sampling with a 100 mV full-scale is then applied to the output signal, together with a 40 GSa/s (160 GSa/s) horizontal sampling. For each simulated acquisition, the position of the first sample in the first bit is randomly chosen in the first quarter of the duration of the bit itself. Indeed, in the experimental setup, the triggering signal for the oscilloscope comes from the DC generator that controls the phase shifters too. The oscilloscope is then asynchronous with the AWG. Therefore, the position of the first sample in the sequence is different in each acquisition. Depending on where the first sample falls in each bit, the contrast level in the acquired curve changes, possibly leading to a different BER result.

The sampled output signal is then compared with the input signal according to the same modalities described in Section 2 for the real experiment. The only difference is in regard to the training phase, during which the loss function is always evaluated at the 3rd sample, this being closer to the center of the bit and more distant from the transients. Different fiber length scenarios are simulated using the PSO training algorithm.

After each run, the BER versus PRX curves are calculated. Each BER value appearing in the profiles is obtained as an average over $N = 1000$ measurements, corresponding to a minimum non-null measurable value of $1/(N \times$ number of bits in the sequence).

APPENDIX C: NOISE MODELING

In the experimental setup, the optical amplifiers (EDFAs and SOA) act as noise sources, but the presence of the 30 GHz band-pass optical filter reduces their impact on deteriorating the SNR at the receiver. In the studied configuration, their contribution is negligible with respect to that introduced by the fast photodiode (RX2, receiver). The fluctuations in its response to the input optical power can be modeled as follows [1]:

$$\sigma^2 = \langle(\Delta I)^2\rangle = \sigma_s^2 + \sigma_T^2$$
$$= 2q(I_p + I_d)\Delta f + (4k_B T/R_L)F_n\Delta f.$$

The first term, with q the electron charge, accounts for the contribution coming from shot noise. I_p is the average current, I_d is the dark current, and Δf represents the effective noise bandwidth of the detector. The second term, with k_B the

Boltzmann constant, describes fluctuations induced by thermal noise. T is the temperature, R_L is the load resistor of the detector, and F_n is the noise figure of its internal amplifier. For the current experimental setup, the previous equation becomes $\sigma^2 = \langle(\Delta V)^2\rangle = mV_{\text{meas}} + q$, where V_{meas} is the measured voltage at the oscilloscope, m accounts for the proportional term due to shot noise, and q includes the noise contributions deriving from the thermal and shot noise associated with the dark current. A characterization of the setup provided us with $m = 0.0189$ mV and $q = 0.2263$ mV^2.

Funding. H2020 European Research Council (788793, 963463).

Acknowledgment. We acknowledge a fruitful discussion with Stefano Biasi.

Disclosures. M.M., P.B., and L.P. have filed a patent on the technology here described.

Data Availability. The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1. G. P. Agrawal, *Fiber-Optic Communication Systems*, 4th ed. (Wiley, 2012).
2. E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," J. Lightwave Technol. **26**, 3416–3425 (2008).
3. J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, "Digital signal processing for fiber nonlinearities," Opt. Express **25**, 1916–1936 (2017).
4. K. Kikuchi, "Fundamentals of coherent optical fiber communications," J. Lightwave Technol. **34**, 157–179 (2016).
5. J. Zhao, Y. Liu, and T. Xu, "Advanced DSP for coherent optical fiber communication," Appl. Sci. **9**, 4192 (2019).
6. J. Mata, I. de Miguel, R. J. Durán, N. Merayo, S. K. Singh, A. Jukan, and M. Chamania, "Artificial intelligence (AI) methods in optical networks: a comprehensive survey," Opt. Switching Netw. **28**, 43–57 (2018).
7. A. F. Elrefaie, R. E. Wagner, D. A. Atlas, and D. G. Daut, "Chromatic dispersion limitations in coherent lightwave transmission systems," J. Lightwave Technol. **6**, 704–709 (1988).
8. G. M. Brodnik, C. Pinho, F. Chang, and D. J. Blumenthal, "Extended reach 40 km transmission of C-band real-time 53.125 Gbps PAM-4 enabled with a photonic integrated tunable lattice filter dispersion compensator," in *Optical Fiber Communications Conference and Exposition (OFC)* (IEEE, 2018), pp. 1–3.
9. P. Stark, F. Horst, R. Dangel, J. Weiss, and B. J. Offrein, "Opportunities for integrated photonic neural networks," Nanophotonics **9**, 4221–4232 (2020).
10. A. Katumba, X. Yin, J. Dambre, and P. Bienstman, "A neuromorphic silicon photonics nonlinear equalizer for optical communications with intensity modulation and direct detection," J. Lightwave Technol. **37**, 2232–2239 (2019).
11. S. M. Ranzini, R. Dischler, F. Da Ros, H. Bülow, and D. Zibar, "Experimental investigation of optoelectronic receiver with reservoir computing in short reach optical fiber communications," J. Lightwave Technol. **39**, 2460–2467 (2021).
12. S. M. Ranzini, F. Da Ros, H. Bülow, and D. Zibar, "Tunable optoelectronic chromatic dispersion compensation based on machine learning for short-reach transmission," Appl. Sci. **9**, 4332 (2019).
13. A. Argyris, J. Bueno, and I. Fischer, "Photonic machine learning implementation for signal recovery in optical communications," Sci. Rep. **8**, 8487 (2018).
14. S. Li, S. Dev, S. Kühl, K. Jamshidi, and S. Pachnicke, "Micro-ring resonator based photonic reservoir computing for PAM equalization," IEEE Photon. Technol. Lett. **33**, 978–981 (2021).
15. S. Boshgazi, A. Jabbari, K. Mehrany, and M. Memarian, "Virtual reservoir computer using an optical resonator," Opt. Mater. Express **12**, 1140–1153 (2022).
16. S. Li, S. Ohlendorf, and S. Pachnicke, "100 km 56 GBd PAM-4 transmission using photonic reservoir computing," in *45th European Conference on Optical Communication (ECOC)* (IET, 2019), pp. 1–4.
17. M. Mancinelli, D. Bazzanella, P. Bettotti, and L. Pavesi, "A photonic complex perceptron for ultrafast data processing," Sci. Rep. **12**, 4216 (2022).
18. J. F. Schutte, J. A. Reinbolt, B. J. Fregly, R. T. Haftka, and A. D. George, "Parallel global optimization with the particle swarm algorithm," Int. J. Numer. Methods Eng. **61**, 2296–2315 (2004).
19. D. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2017).
20. S. Biasi, R. Franchi, D. Bazzanella, and L. Pavesi, "On the effect of the thermal cross-talk in a photonic feed-forward neural network based on silicon microresonators," Front. Phys. **10**, 1093191 (2022).
21. T. Mizuochi, "Recent progress in forward error correction and its interplay with transmission impairments," IEEE J. Sel. Top. Quantum Electron. **12**, 544–554 (2006).
22. P. P. Absil, P. De Heyn, H. Chen, P. Verheyen, G. Lepage, M. Pantouvaki, J. De Coster, A. Khanna, Y. Drissi, D. Van Thourhout, and J. Van Campenhout, "Imec iSiPP25G silicon photonics: a robust CMOS-based photonics technology platform," Proc. SPIE **9367**, 93670V (2015).
23. D. C. Kilper, "Energy efficiency in optical networks," in *Springer Handbook of Optical Networks* (Springer, 2020), pp. 631–664.

193.205.210.74 {ts '2023-05-02 06:16:47'}