European Conference on Optical Communication (ECOC) 2022 © Optica Publishing Group 2022

A Sparse-Readout Reservoir-Computing Based Equalizer for 100 Gb/s/λ PON

Xiaoan Huang, Dongxu Zhang, Xiaofeng Hu, and Kaibin Zhang
Nokia Bell Labs, Shanghai, 201206, China, xiaoan.huang@nokia-sbell.com

Abstract *A low-power sparse-readout reservoir-computing based equalizer is proposed and evaluated by experiments on a 100-Gbps/λ PON testbed. Results demonstrate that it is feasible to greatly reduce the readout layer's complexity while achieving a 29-dB power budget. Integrated photonics implementation issues are also discussed. ©2022 The Author(s)*

Introduction

To meet the demand of emerging applications, such as fixed-mobile convergence for the fifth generation of mobile networks and beyond it, a 100-Gbit/s/λ access network becomes the next priority for the passive optical network (PON) roadmap [1]. For PON systems with such high link speed, equalization would become one of the core techniques. The conventional neural networks (NN), such as fully-connected NN and recurrent NN (RNN), have been regarded as promising equalization tools to compensate for the transmission impairments including both linear and nonlinear distortions in the high-speed optical communication system [2]. But NN equalizers usually require an energy and time-consuming training stage to work appropriately as they require iterative training operations (i.e. backpropagation) through the whole network.

Compared with the conventional NN models, recently the Reservoir-Computing (RC), which is more suitable to be implemented as optical neuromorphic hardware instead of digital application-specific integrated circuit (ASIC), was proposed as a solution to relieve the computing burden of the training stage. The reservoir layer can be left untrained, but only the weights of the readout layer need to be optimized by a simple regression method [3]-[12]. More importantly, a photonics-based hardware implementation of RC allows to fully exploit light's advantages (high speeds and low power consumption) for computational purposes. For example, using silicon photonics technology, the efficiency and effectiveness of a passive swirl reservoir chip have been experimentally verified [5]-[8].

In terms of equalization performance, the RC-based nonlinear equalizer has been investigated in a few literatures targeting short-reach optical communication applications. For example, in [9]-[11], an RC equalizer composed of a random-structured reservoir matrix and a full-readout layer with hundreds (~500) of readout connections was proposed and experimentally exploited on a 32-GBd OOK intensity-modulation and direct-detection (IMDD) short-reach optical system. To some extent, the RC neural network can be seen as a special type of RNN, but it has the potential of achieving comparable signal equalization performance with lower power consumption and faster training speed.

However, it should be noticed that the conventional RC equalizer with a full-readout layer tends to require a large number of readout connections to guarantee its performance [9]-[11]. When implemented in hardware, the large full-readout layer translates to many active components (e.g., optical modulators (OM)) which are both power-consuming and costly.

In this paper, we firstly propose a sparse-readout deterministic reservoir (SDR) based equalizer and evaluate its performance for PON applications. Using experimental data from a real 100-Gbps/λ PAM4 IMDD PON testbed, we verify that it can achieve similar performance as the conventional full-readout RC equalizers but saves 22~59% readout layer connections. Finally, some considerations regarding integrated photonics implementations are also discussed.

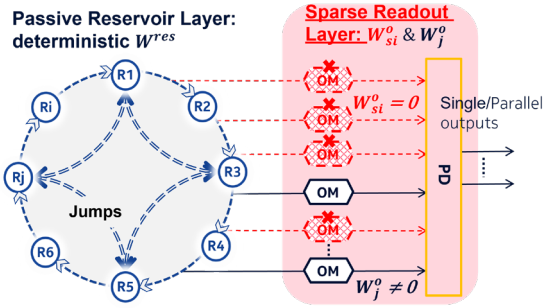

Fig. 1 A sparse-readout RC with a deterministic reservoir structure and a sparse readout layer implemented with OMs.

Architecture and Principles

The nonlinear dynamic system that describes the evolution of the reservoir node's state (x_t) can be expressed as:

$$x_t = \alpha \cdot f_{AF}(W^{in} \cdot u_t + W^{res} \cdot x_{t-1}) + (1 - \alpha) \cdot x_{t-1} \quad (1)$$

, where α is the leaking rate (set to 0.9), u_t is the neurons' input, and f_{AF} denotes the nonlinear activation function (*tanh*). W^{in} is the MxN input

weight matrix where M is the input size and N is the number of reservoir nodes. W^{res} is the "reservoir matrix" whose elements represent the weights of the connections among the nodes. The weights of W^{res} can be constructed in a randomized manner as in literatures [9]-[11], which lacks implementation practicality. Here, for the simple hardware implementation purpose, a "deterministic reservoir with jumps" (DRJ) structure is employed by first connecting all nodes into a circle shape and then adding "jump connections" along the circle with a certain skip ratio [12]. The W^{in} and W^{res} are assigned with fixed values, while the training process only applies to the equalizer's readout layer. For an RC, this is typically accomplished by a Ridge linear regression operation aiming at minimizing the squared error of the loss function:

$$min\{||Y_{target} - XW^o||_2^2 + \lambda||W^o||_2^2\} \quad (2)$$

, where the X, Y_{target}, and λ represent the reservoir state vector, the target vector, and the ridge regularization factor, respectively. Based on the Ridge linear regression, the readout weights W^o are all non-zero [9]-[11].

In this paper, we intend to reduce the complexity of the readout layer by sparsifying W^o, which means W^o should contain as many *0*s as possible after training. This can be done by cost function engineering, i.e., replacing Eq. (2) with the Eq. (3), which is known as an elastic-net linear regression method [13].

$$min\{||Y_{target} - XW^o||_2^2 + \lambda \cdot \rho \cdot ||W^o||_1 + \lambda \cdot (1-\rho) \cdot ||W^o||_2^2\} \quad (3)$$

In Eq. (3), ρ and λ are the elastic-net regression penalty parameter and regularization factor, respectively. After sufficient iterations, the readout weights W^o would be composed of zero $(W_{si}^o = 0)$ and non-zero ones $(W_j^o \neq 0)$. Thus, the readout layer connections corresponding to W_{si}^o (denoted as red-dashed lines in Fig. 1) can be removed safely without influencing the computing result. In a real hardware chip, this means that fewer OMs are required.

Experiment and Results

To evaluate the equalizers' performance, we conduct an experiment of 50-GBaud PAM4-based 100-Gbps/λ IMDD PON using 25G-class optics in the O-band. The testbed is elaborated in detail in our previous work [2]. Here, the proposed SDR-based equalizer training is performed by PyTorch in offline processing. In the experiment, we used datasets generated by Mersenne-Twister based random sequences, which can get rid of the pattern effects of PRBS [14]. The length of the whole data set is 100,000 symbols, including 40,000 and 60,000 symbols for training and test, respectively. The equalizers under investigation can be instantly trained in only a matter of seconds on a laptop CPU due to the simplicity of elastic-net linear regression.

As for the reservoir layer, one jump connection is added for every 8 nodes along the outer circle. The jump-size can be a parameter that impacts the equalizer performance, which is not elaborated here due to the limited paper

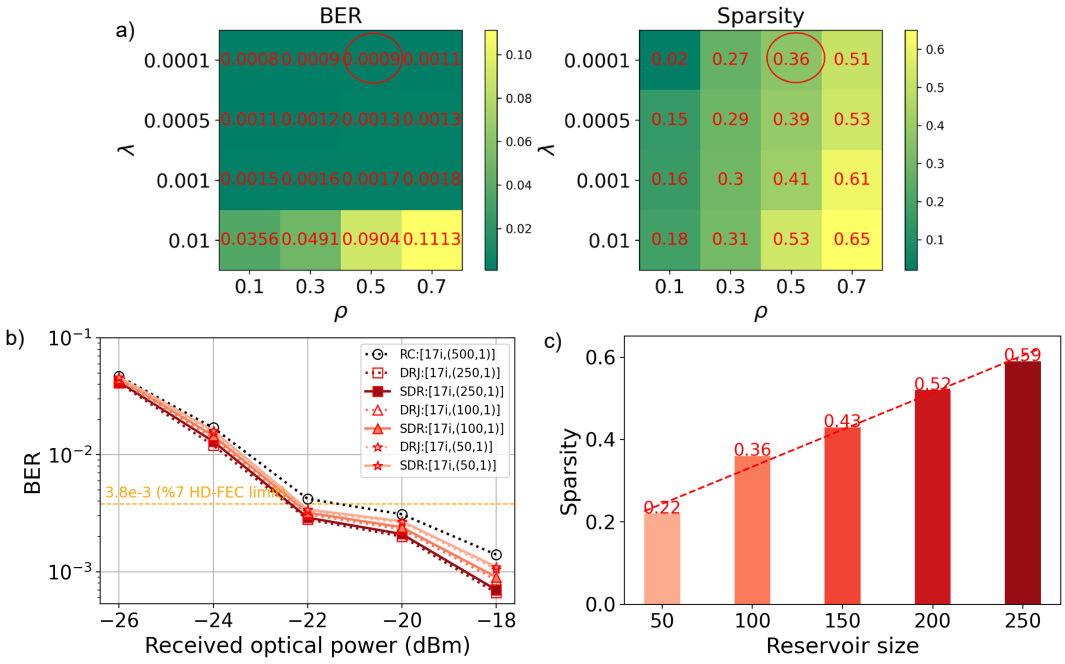

Fig. 2 a) BER & Sparsity against ρ and λ by our proposed SDR of [17i,(100,1)] @-18dBm ROP; b) BER against the ROP (proposed: solid vs. conventional: dotted); c) Sparsity of proposed SDR against the reservoir-size @-18dBm ROP.

length and is left for future publication. Here, we focus more on the key hyperparameters of ρ and λ in the sparse-readout layer. The basic rule of the hyperparameters optimization is to seek a large sparsity under a BER constraint that satisfies the PON requirements. Assuming that the BER constraint should be "smaller than 1e-3 @-18-dBm received optical power (ROP)", using a Grid-Search method, the optimal (ρ, λ) of an SDR with 17-inputs-100-nodes-1-output ([17i, (100,1)]) would be (0.5,0.0001), as shown in Fig. 2 a). In this case, a sparsity of 36% is obtained while the BER of 0.0009 satisfies the system constraint (marked with red circles in Fig. 2 a)).

As shown in Fig. 2 b), all the proposed sparse-readout SDR equalizers with 250, 100, and 50 nodes achieve a similar BER performance as the conventional full-readout ones with the same reservoir size. The proposal can all achieve a good BER under the 7% hard-decision forward error correction (HD-FEC) limit of 3.8e-3 at -21-dBm ROP. In our experiment, the launch power was 8-dBm, thus a 29-dB power budget has been achieved after 20-km fiber. Even with only 50 nodes ([17i, (50,1)]), the SDR can achieve a slightly better BER performance than a conventional 500-node random-structured RC ([17i, (500,1)]) with a full-readout layer, which demonstrates the effectiveness of the DRJ structure. Fig. 2 c) shows that, compared with the full-readout RC with the same reservoir structure and size, the sparsity of the proposed SDR grows almost linearly from 22% to 59% as the reservoir-size increases from 50 to 250. This is true as there is much room for more exploitation of the sparsity when the reservoir-size in our evaluation increases. Also note that when the reservoir-size is almost squeezed to its limit of around 50, the proposed SDR can even obtain a good sparsity (22%) while maintaining the BER performance, thus can further reduce the whole complexity and related power consumption.

Integrated Photonics Discussion

To implement the proposed SDR equalizer, many photonics integration technology options can be considered [15]. From the high-level design perspective, depending on how the readout layer is implemented, Fig. 3 depicts two possible solutions: all-optical and hybrid optoelectronic. An all-optical readout layer (Fig. 3 a)), for example, could be a modulator array followed by a single photodetector (PD), which realizes weighted sum operation ($X \cdot W^o$) in the analog signal domain. For the optoelectronic design (Fig. 3 b)), the reservoir's outputs are converted to electrical signals and digitized individually, and then the weighted sum is done in the digital domain. Since the W^o is sparse, signal paths corresponding to 0s are switched off.

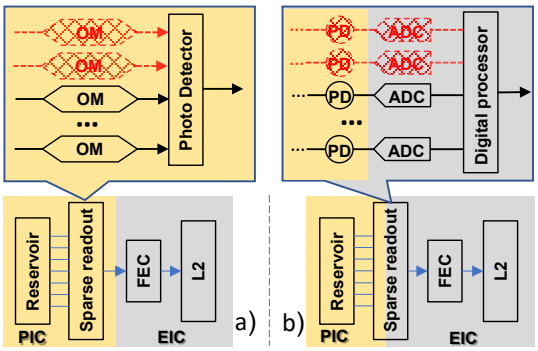

Fig. 3 Implementation options. (PIC: photonics integrated circuit; EIC: electrical integrated circuit; L2: layer 2 functions)

Due to production variations and the difficulty of observing all states in the reservoir, some additional tricks are needed to calibrate the all-optical readout weights [6]. If the hybrid optoelectronic solution is used, the weights can be tuned digitally with the additional flexibility of calibrating imperfections in the optics. But the energy consumption of the PD and analog-to-digital converter (ADC) array may be problematic. Regarding energy efficiency, the all-optical option in Fig. 3 a) is preferred as it eliminates energy-consuming devices including high-speed PDs, ADCs, and the digital processer. Thanks to the W^o sparsity, the energy consumption can be further saved.

Another aspect of concern is the technology platform [15]. Silicon photonics is promising at integrating both PIC and EIC in a single die but suffers higher optical loss. The III-V platform is more versatile in supporting active components but needs advanced packaging with EIC. The low-loss silicon nitride is also an appealing option for optical RC, especially if hybrid integration with other materials becomes matured, such as non-volatile optical memories [16] for weighting, etc.

Conclusions

We introduced a new low-power sparse-readout deterministic reservoir based equalizer for PON applications. Results verify that, applying the proposed equalizer to a 100-Gbps/λ signal after 20-km transmission, a 29-dB power budget can be achieved. Compared with the conventional RC, a large sparsity (22~59%) can be achieved without sacrificing the BER performance. The related integrated photonics implementation issues are also discussed, and we believe that it would be very promising for future PON as well as other low-power and high-speed applications.

Acknowledgements

We thank Dr. Jochen Maes in Nokia Bell Labs, Antwerp, for inspiring discussion.

References

[1] R. Borkowski et al., "World's First Field Trial of 100 Gbit/s Flexible PON (FLCS-PON)," 2020 European Conference on Optical Communications (ECOC 2020), 2020, pp. 1-4. DOI: 10.1109/ECOC48923.2020.9333413.

[2] X. Huang, et al., "Low-Complexity Recurrent Neural Network Based Equalizer with Embedded Parallelization for 100-Gbit/s/λ PON," in Journal of Lightwave Technology, vol. 40, no. 5, pp. 1353-1359, 1 March1, 2022, DOI: 10.1109/JLT.2021.3128579.

[3] Brunner, D., Soriano, M., Mirasso, C. et al. "Parallel photonic information processing at gigabyte per second data rates using transient states". Nature Communications 4, 1364 (2013). https://doi.org/10.1038/ncomms2368.

[4] A. Argyris, J. Bueno, and I. Fischer, "Photonic machine learning implementation for signal recovery in optical communications," Scientific Reports, vol. 8, no. 1, 2018, Art. no. 8487. https://doi.org/10.1038/s41598-018-26927-y.

[5] Vandoorne, K., Mechet, P., Van Vaerenbergh, T. et al. "Experimental demonstration of reservoir computing on a silicon photonics chip". Nature Communications 5, 3541 (2014). https://doi.org/10.1038/ncomms4541.

[6] A. Katumba, et al., "Neuromorphic Computing Based on Silicon Photonics and Reservoir Computing," in IEEE Journal of Selected Topics in Quantum Electronics, vol. 24, no. 6, pp. 1-10, Nov.-Dec. 2018, Art no. 8300310, DOI: 10.1109/JSTQE.2018.2821843.

[7] A. Katumba, X. Yin; J. Dambre, and P. Bienstman, "A neuromorphic silicon photonics nonlinear equalizer for optical communications with intensity modulation and direct detection," Journal of Lightwave Technology, vol. 37, no. 10, pp. 2232–2239, May 2019. DOI: 10.1109/JLT.2019.2900568.

[8] S. Sackesyn, C. Ma, J. Dambre and P. Bienstman, "Experimental Demonstration of Nonlinear fibre Distortion Compensation with Integrated Photonic Reservoir Computing," 2021 European Conference on Optical Communication (ECOC), 2021, pp. 1-4. DOI: 10.1109/ECOC52684.2021.9606162.

[9] F. Da Ros, S. Ranzini, H. Búlow, and D. Zibar, "Reservoir-computing based equalization with optical pre-processing for short-reach optical transmission," IEEE Journal on Selected Topics in Quantum Electronics, vol. 26, no. 5, Sep./Oct. 2020, Art no. 7701912. DOI: 10.1109/JSTQE.2020.2975607.

[10] S. Ranzini, R. Dischler, F. Da Ros, H. Bülow, and D. Zibar, "Experimental demonstration of optoelectronic equalization for short-reach transmission with reservoir computing," in European Conference on Optical Communications, Brussels, Belgium, Sep. 2020. DOI: 10.1109/ECOC48923.2020.9333372.

[11] S. Ranzini, et al. "Experimental investigation of optoelectronic receiver with reservoir computing in short reach optical fiber communications." Journal of Lightwave Technology 39.8 (2021): 2460-2467. DOI: 10.1109/JLT.2021.3049473.

[12] A. Rodan and P. Tino, "Simple Deterministically Constructed Cycle Reservoirs with Regular Jumps," in Neural Computation, vol. 24, no. 7, pp. 1822-1852, July 2012, DOI: 10.1162/NECO_a_00297.

[13] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 67, no. 2, pp. 301–320, 2005. DOI: 10.1111/j.1467-9868.2005.00527.x.

[14] T. A. Eriksson, H. Bülow, and A. Leven, "Applying neural networks in optical communication systems: possible pitfalls," IEEE Photonics Technology Letters 29, 2091–2094 (2017). DOI: 10.1109/LPT.2017.2755663.

[15] P. Stark, F. Horst, R. Dangel, J. Weiss, and B.J. Offrein, "Opportunities for integrated photonic neural networks," Nanophotonics, vol. 9, no. 13, 2020, pp. 4221-4232. https://doi.org/10.1515/nanoph-2020-0297.

[16] C. Rios, M. Stegmaier, P. Hosseini, D. Wang, T. Scherer, C. Wright, H. Bhaskaran, and W. Pernice, "Integrated all-photonic non-volatile multi-level memory," Nature Photonics, vol. 9, no. 11, pp. 725–732, 2015. https://doi.org/10.1038/nphoton.2015.182.