Real-Time Deployment of Simplified Volterra Nonlinear Equalizer in High-Speed PON

Wenqing Jiang, Luyao Huang, Yongxin Xu, Zhixue He, Weisheng Hu, *Member, IEEE*, and Lilin Yi

***Abstract*— In cost-sensitive passive optical networks (PON), economical channel equalization is a promising technique. The electronic equalization is realized by a digital signal processing (DSP) chip with an integrated analog-to-digital converter (ADC) commonly implemented at the receiver side. Field programmable gate array (FPGA), one of the major hardware platforms nowadays, can provide prototype verification for chip design due to the programmability and short iteration cycle. In this letter, using the FPGA and a cascaded 30-GSa/s ADC, a real-time 1310nm band-limited intensity modulation and direct detection PON system is constructed. With a powerful Volterra nonlinear equalizer (VNLE) running on FPGA, a 31-dB power budget is achieved at a 30-Gb/s line rate. Meanwhile, by pruning and clustering the weights in VNLE, the power consumption is declined by 71.6%.**

***Index Terms*— Digital signal processing, analog-to-digital converter, field programmable gate array, Volterra nonlinear equalizer, pruning, clustering.**

I. Introduction

IN ORDER to meet the increasing bandwidth requirement, passive optical networks (PON) are gradually extended to higher signal rates [1], [2]. Compared with wavelength division multiplexing (WDM)-PON based on coherent detection technique, time division multiplexing (TDM)-PON based on intensity modulation and direct detection (IMDD) has become one of the main solutions for PON with the advantages of low cost, simple structure, and easy management [3]. For high-speed PON, the limited device bandwidth, chromatic dispersion, and nonlinearities degrade the signal quality. Therefore, the deployment of electrical channel equalization in the application-specific integrated circuit (ASIC) with cost advantages at the receiver is necessary. The analog-to-digital converter (ADC) and digital signal processing (DSP) chip architecture [4] is appropriate for equalization, which provides higher link margins with cost efficiency. In the initial stage of DSP chip design, the pre-verification and analysis of the chip need the help of a hardware platform. Field programmable gate array (FPGA), attributed to its programmability and short iteration cycle, becomes one of the most popular hardware platforms [5]. Hence, it is logical to deploy electronic equalizers in FPGA for performance and cost evaluation.

In related works, the signal is sampled by the digital sampling oscilloscope (DSO) and then loaded into FPGA for equalization [6], [7]. Compared with DSO receiving, the integrated ADC and FPGA receiving is closer to the practical applications, where the signal sampled by ADC is delivered directly to FPGA. Against to high-rate commercial DSO, ADC suffers more severe quantization noise due to the low resolution of 6-bit and the input range is fixed which leads to insufficient filling. So, the equalization performance evaluation of high-speed PON based on ADC and FPGA is more challenging.

An effective equalization algorithm should be chosen for system impairment compensation. By detailed comparison of performance, complexity, and optimization difficulty, Volterra nonlinear equalizer (VNLE) and Volterra-DFE nonlinear equalizer (VDFE) have good performance offline [8]. However, the high-parallel hardware deployment of VDFE is implausible because a small number of feedback loops [9] result in significant resource expenditure. Considering performance and energy consumption, VNLE with a simple feedforward structure is more hardware-friendly for real-time deployment in FPGA. Moreover, pruning [10], [11] and clustering [12] in neural networks are available to reduce the complexity of VNLE for cheaper receiver costs.

In this letter, high-parallel VNLE is running on FPGA. Meanwhile, to achieve power efficiency, pruning and clustering are applied to simplify the structure of VNLE. The effectiveness of real-time ADC-FPGA receiver is verified in an O-band IMDD system transmitting 30-Gb/s 4 pulse amplitude modulation (PAM4) signal. The resource and power consumption of VNLE are evaluated. The experimental results show that VNLE deployed in FPGA is capable to achieve a power budget of 31 dB with a sensitivity of -19 dBm in the 30-Gb/s IMDD-PON, which verifies the efficiency of the real-time receiver. Besides, pruning and clustering cut power usage by 71.6% for energy conservation. To the best of our knowledge, this is the first work to achieve a simplified VNLE based on the real-time ADC-FPGA receiver in the IMDD PON system.

Manuscript received 7 May 2023; accepted 22 July 2023. Date of publication 27 July 2023; date of current version 9 August 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFB1803803, in part by the National Natural Science Foundation of China under Grant 62025503, and in part by the Major Key Project of Peng Cheng Laboratory (PCL). *(Corresponding author: Lilin Yi.)*

Wenqing Jiang, Luyao Huang, Yongxin Xu, Weisheng Hu, and Lilin Yi are with the State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: sjtujwq@sjtu.edu.cn; luyaohuang@sjtu.edu.cn; xuyongxin@sjtu.edu.cn; wshu@sjtu.edu.cn; lilinyi@sjtu.edu.cn).

Zhixue He is with the Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: hezhx01@pcl.ac.cn).

Color versions of one or more figures in this letter are available at https://doi.org/10.1109/LPT.2023.3299300.

Digital Object Identifier 10.1109/LPT.2023.3299300

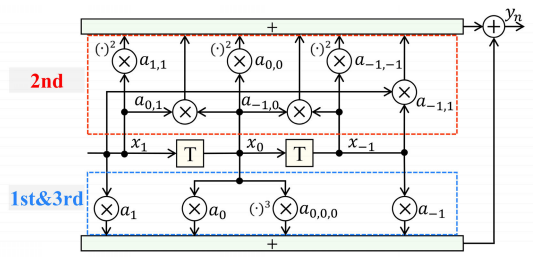

Fig. 1. Block diagram of VNLE deployment when $[L_1, L_2, L_3] = [3, 3, 1]$.

II. Principle

VNLE is a nonlinear feedforward equalizer with a simple structure and excellent performance. It has both linear and nonlinear terms [13] and can compensate for the inter-symbol interference and nonlinearities effectively in O-band system with almost no chromatic dispersion. Therefore, VNLE is widely applied for channel equalization due to its powerful nonlinear fitting ability.

The expression of the third-order VNLE adopted in this letter is shown in (1).

$$\begin{aligned} y_i = & \sum_{j_1=-m_1}^{m_1} a_{j_1} x_{i-j_1} \\ & + \sum_{j_1=-m_2}^{m_2} \sum_{j_2=-m_2}^{j_1} a_{j_1,j_2} x_{i-j_1} x_{i-j_2} \\ & + \sum_{j_1=-m_3}^{m_3} \sum_{j_2=-m_3}^{j_1} \sum_{j_3=-m_3}^{j_2} a_{j_1,j_2,j_3} x_{i-j_1} x_{i-j_2} x_{i-j_3} \end{aligned} \quad (1)$$

The x_i is the received symbol. The $L_1 = 2m_1 + 1$, $L_2 = 2m_2 + 1$, and $L_3 = 2m_3 + 1$ are the memory lengths of the first, second, and third order of VNLE, respectively. The a_{j_1}, a_{j_1,j_2}, and a_{j_1,j_2,j_3} are the corresponding tap coefficients. The y_i is the symbol after equalization. To evaluate the algorithmic complexity of VNLE, the number of multiplications required for each symbol is applied as shown in (2).

$$L_1 + L_2 (L_2 + 1) + \frac{L_3 (L_3 + 1) (L_3 + 2)}{2} \quad (2)$$

The block diagram of VNLE deployment is shown in Fig. 1. The T means the time delay. The $\times$ is the multiplier. The $(\cdot)^n$ denotes the n-th power of the signal. The input signal is delayed to generate the first, second, and third-order terms and then multiplied with the corresponding tap coefficients. Finally, the output of VNLE is obtained by an accumulator.

During VNLE deployment, three optimization methods—data multiplexing, pruning, and clustering—are utilized.

A. Data Multiplexing

To meet the requirement of a 30-Gb/s line rate, VNLE is realized with the parallelism P of 64 when the clock frequency of FPGA is 234.375MHz. Because of the existence of parallel channels, data multiplexing is used to reduce the consumption of multipliers. Figure 2 shows the multiplexing of second-order terms when $P = 64$. The terms are convolved

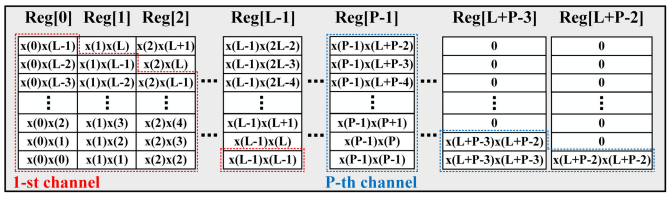

Fig. 2. Data multiplexing of second-order terms.

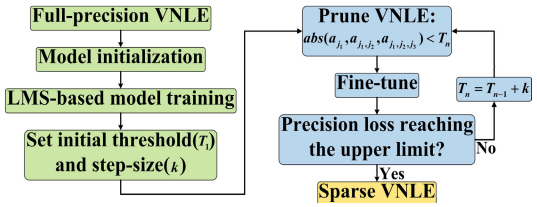

Fig. 3. The principle of pruning.

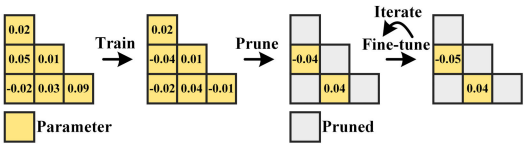

Fig. 4. Pruned second-order convolutional kernel when $L_2 = 3$.

by the corresponding coefficients in the convolutional kernel to generate the required inputs for each channel accumulator.

B. Pruning

There are plenty of redundant weights in VNLE that have a tiny impact on the final result. To further reduce resource expenditure, pruning is adopted to diminish the complexity. The principle of pruning is shown in Fig. 3. The process is divided into five steps: (a) The weights of VNLE are initialized. (b) The model is trained based on the least mean square (LMS) algorithm. (c) Pruning is performed and can be described as (3).

$$a_{j_1}, a_{j_1,j_2}, a_{j_1,j_2,j_3} = 0, if \left| a_{j_1}, a_{j_1,j_2}, a_{j_1,j_2,j_3} \right| < T_n \quad (3)$$

The T_n is the threshold. Then, fine-tuning is executed to compensate for the precision loss. (d) If the precision loss does not reach the set upper limit, the threshold is increased for iterative pruning. (e) Finally, the sparse VNLE is obtained.

The pruned second-order convolutional kernel when $L_2 = 3$ is shown in Fig. 4. After pruning, undersized weights are set to zero. They are not involved in any subsequent calculations thus saving resources.

C. Clustering

The weights in FPGA are denoted as fixed-point numbers. The bit-width of weights is 12-bit in this letter when considering both complexity and precision. It includes the sign, integer, and decimal bits. To lower the computational complexity, clustering is employed. The clustering centers are described as (4).

$$2^{-n} + 2^{-m} or 2^{-n} \ (1 \le n, m \le 10, n \ne m) \quad (4)$$

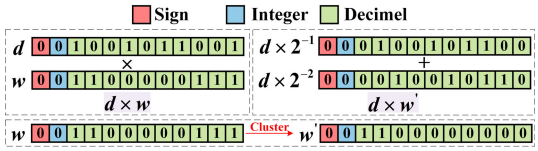

Fig. 5. Illustration of weight clustering.

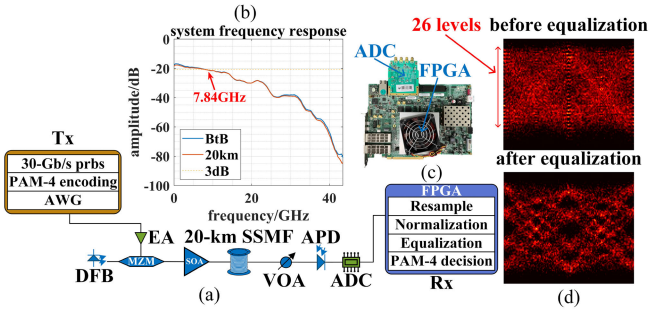

Fig. 6. (a) Experimental setup for 30-Gbps O-band IMDD PON real-time system. (b) Frequency response of the system. (c) Real-time ADC-FPGA receiver. (d) Eye diagram of received signal before and after equalization.

Weights within the range (-0.875,0.875) are replaced by the closest clustering centers. As a result, multiplications are converted to simpler shift operations and additions.

The illustration of weight clustering is shown in Fig. 5. It depicts the multiplication of d and w, which are all 12-bit fixed-point numbers. The value of w is 0.7568359375. After clustering, w is converted to w'. The value of w' is 0.75, which is the sum of 2^{-1} and 2^{-2}. When performing the multiplication of d and w', the result can be obtained from shifting d to right by 1-bit and 2-bit respectively, and adding them together. In this way, the computational cost is further shrunk.

III. Experiment and Result

A. Experimental Setup

The real-time FPGA deployment of VNLE is realized in a 30-Gb/s O-band IMDD PON system. The experimental setup is shown in Fig. 6(a). At the transmitter, the PAM4 downstream signal is sent by the Keysight 8195A arbitrary waveform generator (AWG). Then the signal is amplified by an electrical amplifier with a gain of 26 dB and loaded into a 10-GHz Mach–Zehnder modulator (MZM) for modulation. A 1310 nm distributed feedback (DFB) laser is used as the light source. The modulated optical signal is amplified by a semiconductor optical amplifier (SOA) to reduce the insertion loss and then fed into a 20-km standard single-mode fiber (SSMF) for transmission. At the receiver, a variable optical attenuator (VOA) is used as a splitter to adjust the received optical power. Then the optical signal is detected by a 30-GHz avalanche photodiode (APD). The system frequency response is shown in Fig. 6(b), and the 3-dB bandwidth is 7.84 GHz. In addition, the back-to-back (BtB) spectrum and the one after 20-km link are similar because there is almost no chromatic

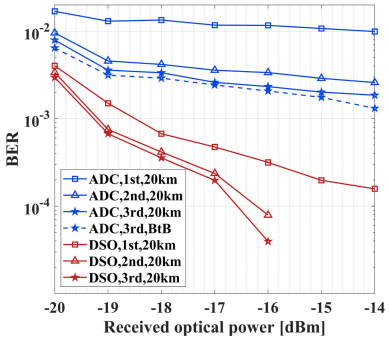

Fig. 7. Comparison of DSO and ADC performance.

dispersion in our system. Then the signal is converted from analog to digital by a 30-GSa/s ADC with 15-GHz bandwidth and Tektronix DPO70000SX DSO which is operated as a reference for performance comparison. The digital signal sampled by ADC with a resolution of 6-bit is delivered into FPGA through the FPGA mezzanine card (FMC) and received by the gigabit transceivers. The FPGA chip is xcvu9p-flga2104-2L-e. The 15-GHz clock of ADC is provided by the Keysight E8257D analog signal generator, which also serves as the clock resource of AWG. The 234.375-MHz clock of FPGA is obtained by dividing the ADC clock. The real-time receiver is shown in Fig. 6(c). After pre-processing, the signal is compensated by VNLE operating at 1 sample per symbol in FPGA.

The eye diagram before and after equalization is shown in Fig. 6(d). Before equalization, the eye diagram is completely closed due to the severe degradation of signal quality. After the impairment is compensated, the eye is opened. In Fig. 7, the bit-error-rate (BER) decreases gradually as the order of VNLE increases. The BtB performance is only slightly improved because the system works in O-band. Besides, it is obvious that the performance of the ADC receiver operating at 30GSa/s with a resolution of 6-bit is worse than the DSO receiver operating at 100GSa/s with a resolution of 8-bit. In particular, the ADC receiver has a poor error floor as the blue line in Fig. 7. It is because the quantization noise, band limit and low resolution of ADC cause a more serious impact on the signal quality compared to the system with DSO as the receiver. Moreover, a distinct horizontal line represents a quantization level and there are only 26 quantization levels in the eye diagram before equalization in Fig. 6(d). In other words, the performance of ADC with 64 quantization levels (6-bit) is not fully utilized due to the fixed input range. The output voltage of APD cannot fill the range of ADC. The actual resolution is under 5-bit and more susceptible to the quantization noise. However, these problems are capable to be solved by the use of higher bandwidth chips or specific circuit design in the subsequent integration of transceiver modules. Therefore, when ADC is used, it is not possible to recover the signal by the first-order VNLE due to the poor performance of ADC and the nonlinearities of SOA and MZM. The third-order VNLE is required to effectively compensate for the serious impairment for best performance.

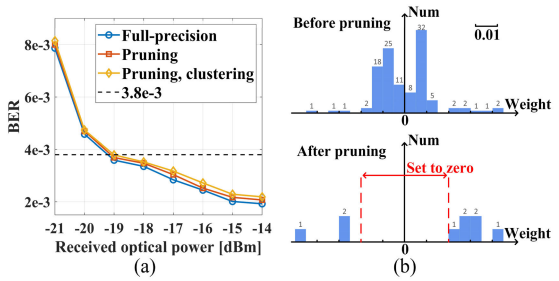

Fig. 8. (a) Sensitivity curve for 30-Gbps O-band PAM4 signal when $[L_1, L_2, L_3] = [121, 15, 5]$. (b) Distribution of first-order kernels within the range (-0.05,0.05) before and after pruning.

TABLE I
RESOURCE AND POWER CONSUMPTION

VNLE with L = [121,15,5]	Look-up tables	Registers	Carry chains	DSPs	Power(W)
Full-precision	618184(100%)	639355(100%)	93568(100%)	1189(100%)	37.539(100%)
Pruning	193372(31.3%)	202872(31.7%)	27684(29.6%)	663(55.8%)	14.113(37.6%)
Pruning&Clustering	127220(20.6%)	177336(27.7%)	19392(20.7%)	663(55.8%)	**10.657(28.4%)**

B. Results

The sensitivity of the real-time receiver is measured in an O-band IMDD PON system with a 30-Gb/s line rate, and the results are shown in Fig. 8(a). Firstly, to obtain the hyperparameters which can achieve a high-power budget while maintaining low complexity and the weights for real-time deployment, the offline performance of VNLE with various complexities is evaluated in MATLAB by using the signal sampled by ADC. The power budget of 31.2 dB is obtained at the launch power of 12 dBm when $[L_1, L_2, L_3] = [121, 15, 5]$. Continuing to increase the complexity cannot get significant power budget improvement but results in large additional resource depletion. Then, the same VNLE is deployed in FPGA as the blue line in Fig. 8(a). It can also achieve a power budget of 31.2 dB at the 3.8e-3 BER threshold, thus verifying the efficiency of the real-time VNLE. When pruning is in effect, the result is shown as the red line in Fig. 8(a). The power budget achieved by the sparse VNLE appears a 0.1-dB degradation. However, the weight quantity of the sparse VNLE is only 25% of the full-precision one. The distribution of first-order kernels within the range (-0.05,0.05) before and after pruning is shown in Fig. 8(b). After pruning and clustering, the sensitivity curve is displayed as the yellow line in Fig. 8(a). It still gets a power budget of 31 dB because the performance deteriorates slightly.

When analyzing the resource and power consumption, the parts only used by VNLE are considered to assess the impact of pruning and clustering. In TABLE I, there are four main types of resources consumed by VNLE in FPGA, including look-up tables, registers, carry chains, and DSPs. According to the results in TABLE I, the resource and power consumption decrease significantly after pruning and clustering in order to overcome the problems in energy supply, cooling, and package for the receiver chip. Finally, the total power drops by 71.6% and will decline further if VNLE is made into an ASIC.

IV. CONCLUSION

It is the first work that a simplified VNLE is deployed for equalization of the real-time ADC-FPGA receiver in IMDD PON. The performance is verified in an O-band system with a 30-Gb/s line rate transmitting through a 20-km fiber. The signal is received by a band-limited and low-resolution ADC operating at 30GSa/s and recovered by VNLE in FPGA. In addition, pruning and clustering are used to further reduce the cost. Total power consumption drops by 71.6% with only a 0.2-dB deterioration in the power budget. Finally, the power budget of 31dB is obtained. In conclusion, we demonstrate feasibility of the proposed real-time receiver based on ADC-FPGA with a high-power budget. This is an important prototype validation for subsequent efficient implementation on ASIC, which can be integrated into the transceiver module. In the future, the capacity and performance are able to be further improved with higher bandwidth chips.

REFERENCES

[1] D. Zhang, D. Liu, X. Wu, and D. Nesset, "Progress of ITU-T higher speed passive optical network (50G-PON) standardization," *J. Opt. Commun. Netw.*, vol. 12, no. 10, pp. 99–108, Oct. 2020, doi: 10.1364/JOCN.391830.

[2] B. Li et al., "DSP enabled next generation 50G TDM-PON," *J. Opt. Commun. Netw.*, vol. 12, no. 9, pp. 1–8, Sep. 2020, doi: 10.1364/JOCN.391904.

[3] Y. Zhu et al., "Comparative study of cost-effective coherent and direct detection schemes for 100 Gb/s/λ PON," *J. Opt. Commun. Netw.*, vol. 12, no. 9, pp. 36–47, Sep. 2020, doi: 10.1364/JOCN.390911.

[4] M.-A. LaCroix et al., "A 60 Gb/s PAM-4 ADC-DSP transceiver in 7 nm CMOS with SNR-based adaptive power scaling achieving 6.9 pJ/b at 32 dB loss," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2019, pp. 114–116.

[5] R. Thakur and K. Khare, "High speed FPGA implementation of FIR filter for DSP applications," *Int. J. Model. Optim.*, vol. 3, pp. 92–94, Feb. 2013, doi: 10.7763/IJMO.2013.V3.242.

[6] N. Kaneda et al., "FPGA implementation of deep neural network based equalizers for high-speed PON," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, 2020, pp. T4D.2.

[7] J. Chen et al., "High-performance low-complexity Volterra decision feedback equalizer based on FPGA for C-band PAM-4 transmission," *Opt. Fiber Technol.*, vol. 64, Jul. 2021, Art. no. 102555, doi: 10.1016/j.yofte.2021.102555.

[8] L. Huang, Y. Xu, W. Jiang, L. Xue, W. Hu, and L. Yi, "Performance and complexity analysis of conventional and deep learning equalizers for the high-speed IMDD PON," *J. Lightw. Technol.*, vol. 40, no. 14, pp. 4528–4538, Jul. 15, 2022, doi: 10.1109/JLT.2022.3165529.

[9] K. K. Parhi, "Pipelining of parallel multiplexer loops and decision feedback equalizers," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Montreal, QC, Canada, 2004, p. 21.

[10] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding," in *Proc. Int. Conf. Learn. Represent.*, San Juan, Puerto rico, 2016, pp. 1135–1143.

[11] L. Ge, W. Zhang, C. Liang, and Z. He, "Threshold-based pruned retraining Volterra equalization for 100 Gbps/lane and 100-m optical interconnects based on VCSEL and MMF," *J. Lightw. Technol.*, vol. 37, no. 13, pp. 3222–3228, Jul. 1, 2019, doi: 10.1109/JLT.2019.2912911.

[12] A. E. Ezugwu et al., "A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects," *Eng. Appl. Artif. Intell.*, vol. 110, Apr. 2022, Art. no. 104743, doi: 10.1016/j.engappai.2022.104743.

[13] R. D. Nowak and B. D. Van Veen, "Volterra filter equalization: A fixed point approach," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 377–388, Feb. 1997, doi: 10.1109/78.554302.