Reduction in Complexity of Volterra Filter by Employing ℓ_0-Regularization in 112-Gbps PAM-4 VCSEL Optical Interconnect

Yi-Yu Lin[1], Chun-Jui Chen[1], Hong-Minh Nguyen[2], Chun-Yen Chuang[2], Chia-Chien Wei[1,*], Jyehong Chen[2] and Jin-Wei Shi[3]

[1] Department of Photonics, National Sun Yat-sen University, Kaohsiung 804, Taiwan
[2] Department of Photonics, National Chiao Tung University, Hsinchu 300, Taiwan
[3] Department of Electrical Engineering, National Central University, Taoyuan 320, Taiwan
**ccwei@mail.nsysu.edu.tw*

Abstract: We employ ℓ_0-regularization to reduce Volterra filter complexity by up to 90% in 112-Gbps PAM-4 VCSEL transmission. Compared to ℓ_1-regularization, ℓ_0-regularization achieves lower complexity and more precise weights without retraining after sparse identification.

OCIS codes: (060.4230) Multiplexing; (060.4510) Optical communications.

1. Introduction

Given the continued demand for higher capacity in cloud computation and storage, 400-Gbps optical interconnects are the current main development goals for short-distance optical communication systems. 4-level pulse amplitude modulation (PAM-4) has become an important approach due to its ability to double the data rate of non-return to zero (NRZ) signals and been adopted for 500-m, 2-km and 10-km single-mode fiber (SMF) physical medium dependents (PMDs) in IEEE 802.3bs 400GbE [1]. Specifically, 56-Gbaud PAM-4 signaling is the primary objective of the SMF PMDs. However, for short-reach (SR) applications [2], i.e., 100-m multi-mode fiber (MMF) PMDs based on vertical-cavity surface-emitting lasers (VCSELs), 56-Gbaud PAM-4 signaling is still challenging mainly because of the limited bandwidth and significant modulation nonlinearity of VCSELs. In addition, the transmission over MMF will additionally cause detrimental nonlinear distortion and a reduction in bandwidth. While VCSELs would still be an inevitable solution due to low power consumption, large coupling tolerances and low assembly cost, digital signal processing (DSP) provides a viable solution to enable high-speed PAM-4 VCSEL transmission [3, 4]. A Volterra filter is a promising approach to DSP in optical systems due to its extraordinary ability to restore distorted signals by generating a nonlinear inverse system [5–7]. However, increasing the memory length of a Volterra filter leads to an exponential increase in the computational complexity due to numerous higher-order cross-terms in the Volterra series, potentially limiting practical applicability. To overcome this problem, several traditional schemes of sparse identification have been applied in a high-speed PAM-4 VCSEL systems, such as the least absolute shrinkage and selection operator (LASSO; i.e., the ℓ_1-regularization) algorithm [5] and a coefficient pruning algorithm [6]. Alternatively, sparsity identification can be realized by the ℓ_0-regularization [7].

This work demonstrates the reduction in the complexity of Volterra filter in a 112-Gbps PAM-4 VCSEL transmission system using the ℓ_0-regularization. Since sparse identification could disturb the weights of filter, retraining a sparse Volterra filter may be necessary. We show that the gain of retraining after the ℓ_0-regularization is limited, dissimilar to the case using the ℓ_1-regularization [5]. Specifically, the ℓ_0-regularization achieves more reduction in the complexity and obtains more precise weights, compared to the ℓ_1-regularization; thus, retraining could be omitted in the sparse Volterra filter with the ℓ_0-regularization. Using the ℓ_0-regularization without retraining, the experiment results show that up to 90% complexity can be removed when satisfying the KP4 forward error correction (FEC) limit; i.e., the bit error rate (BER) of 2×10^{-4}.

2. Volterra Filter with Regularization

The non-degenerate Volterra filter can be written as [8],

$$y(n) = \sum_{p=1}^{P} \sum_{l_1=0}^{L-1} \sum_{l_2 \geq l_1}^{L-1} \cdots \sum_{l_p \geq l_{p-1}}^{L-1} w(l_1, \ldots, l_p) \prod_{i=1}^{p} x(n - l_i) = \mathbf{x}_n^{\mathsf{T}} \mathbf{w} \tag{1}$$

where $x(n)$ is the nth input sample; $y(n)$ is the nth output sample; $w(l_1, \ldots, l_p)$ denotes the weight of a pth-order tap; P is the order of filter; L is the memory length, and $\mathbf{x}_n$ and $\mathbf{w}$ represent the input vector and weight vector, respectively, composed of all inputs $\prod_{i=1}^{p} x(n - l_i)$ and weights $w(l_1, \ldots, l_p)$. The total number of multiplications

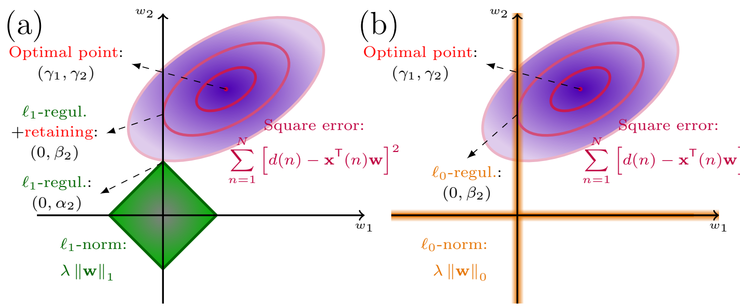

Fig. 1: The concepts of (a) ℓ_1- and (b) ℓ_0-regularization for sparsity.

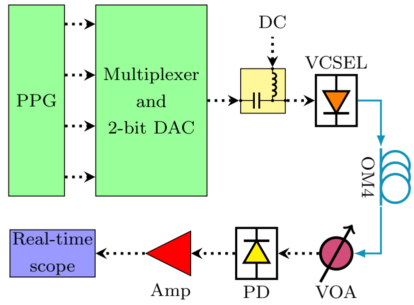

Fig. 2: Experiment setup.

in (1) gives a measure of the computational complexity, $C(P,L) = \sum_{p=1}^{P}(L-1+p)!/(p-1)!/(L-1)!$ [8], where $(\cdot)!$ denotes the factorial operation. Note that the order P is set to 3 in this work to limit the computational complexity. The required weights are estimated through minimizing the cost function; i.e., the square error of training symbols: $J_{\text{se}}(\mathbf{w}) = \sum_{n=1}^{N}[d(n) - \mathbf{x}_n^{\mathsf{T}}\mathbf{w}]^2$, where N is the number of samples in training symbols, and $d(n)$ denotes the desired sample. When the ℓ_0- or ℓ_1-regularization is applied, the cost function becomes $J_0(\mathbf{w}) = J_{\text{se}}(\mathbf{w}) + \lambda\|\mathbf{w}\|_0$ or $J_1(\mathbf{w}) = J_{\text{se}}(\mathbf{w}) + \lambda\|\mathbf{w}\|_1$, respectively, where $\|\mathbf{w}\|_0$ denotes the ℓ_0-norm of $\mathbf{w}$; i.e., the number of non-zero components in $\mathbf{w}$, and $\|\mathbf{w}\|_1$ is the ℓ_1-norm of $\mathbf{w}$; i.e., the sum of the absolute value of all components in $\mathbf{w}$. The contribution of the penalty to the cost functions can be controlled by the value of the penalty parameter λ, and increasing λ will encourage the sparsity of filter. Fig. 1 schematically draws how the regularization promotes the sparsity, and only two taps w_1 and w_2 are included for the sake of simplicity. When using the ℓ_1-regularization, the minimum cost would be at the coordinates $(0, \alpha_2)$, thereby forcing the weight of w_1 to be zero and enabling the sparsity, as shown in Fig. 1(a). However, considering the lowest square error achieved by w_2 alone, the weight should be β_2; thus, retraining the sparse filter would be helpful to reach a lower cost and better performance. On the other hand, $\rho\|\mathbf{w}\|_0$ in $J_0(\mathbf{w})$ tends to reduce the number of nonzero taps. In this case, it would be easier to identify the more critical taps of a filter and attain better weights without retraining, as shown in Fig. 1(b).

3. Experiment and Discussion

Fig. 2 plots the experiment setup of 112-Gbps PAM-4 VCSEL transmission. A pulse pattern generator (PPG, Anritsu MP1800A) generated four-channel 28-Gbps NRZ signals of a $2^7 - 1$ pseudo-random bit sequence. Every two NRZ channels were firstly multiplexed to a 56-Gbps NRZ signal, and two 56-Gbps NRZ signals were then combined as a 56-Gbaud PAM-4 signal by a 2-bit digital-to-analog converter (DAC). The 112-Gbps PAM-4 driving signal was applied to directly modulate a homemade VCSEL [9]. The center wavelength and the 3-dB bandwidth of the VCSEL were $\sim$850 nm and 22 GHz, respectively. Using the optical launch power of 3.2 dBm, the optical PAM-4 signal transmitted over 100-m or 200-m OM4 fiber, and a variable optical attenuator (VOA) was applied to control the received optical power (ROP) before being detected by a 22-GHz photo-detector (PD, Newport 1484-A-50). After amplifying the received electrical signal by a 20-GHz amplifier (Newport 1422), the signal was recorded by a real-time oscilloscope (Tektronix MSO73304DX) with the 3-dB bandwidth of 33 GHz and the sampling rate of 100 GS/s. The BER was measured offline by bit-by-bit comparison. Note that the memory

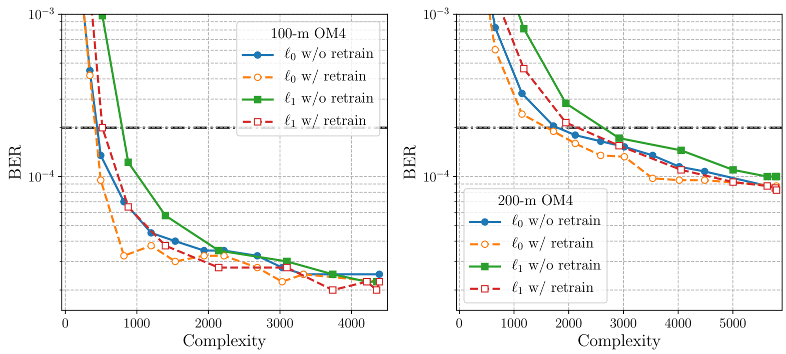

Fig. 3: The BER at the ROP of -4 dBm versus the complexity after (a) 100-m and (b) 200-m OM4.

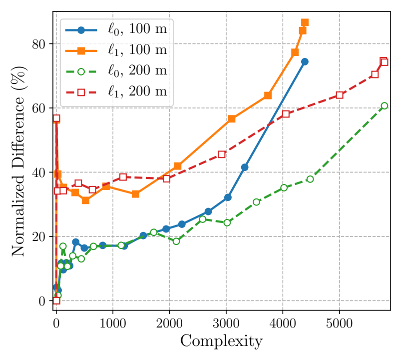

Fig. 4: Normalized difference at the ROP of -4 dBm after retraining.

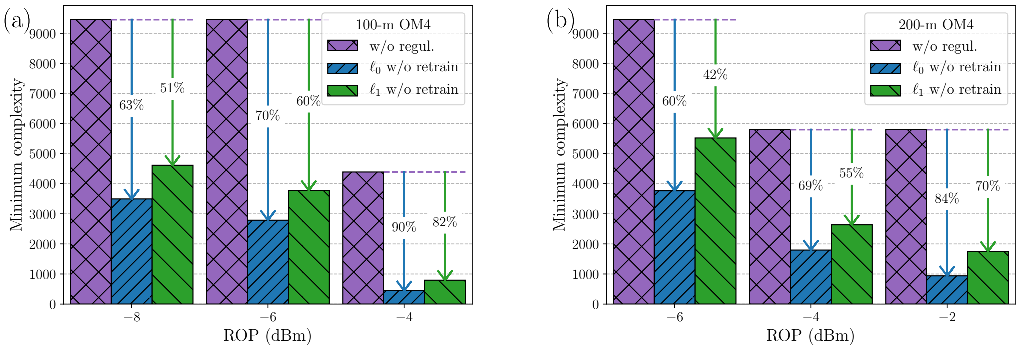

Fig. 5: The minimum complexities at different ROPs after (a) 100-m and (b) 200-m OM4.

length of a full Volterra filter in this work was the minimum value of L in (1) to meet the KP4 FEC limit.

After 100-m and 200-m OM4, the BERs at the ROP of -4 dBm are plotted as functions of the complexity in Fig. 3(a) and (b), respectively. In Fig. 3(a), the memory length is 19 corresponding to the full complexity of $C(3,19) = 4389$. Using the ℓ_1-regularization, the required complexity to meet the FEC limit is lowered by 82% and 88% to 794 and 518 before and after retraining, respectively. On the other hand, the complexity using the ℓ_0-regularization can be reduced by 90% and 90.5% to 445 and 418 before and after retraining, respectively. It is obvious that the ℓ_0-regularization can achieve more reduction in the complexity and exhibit much less difference in the complexity after retraining, compared to the ℓ_1-regularization. Similar trends can be found in Fig. 3(b), where the full complexity is $C(3,21) = 5796$, and the reduction in the complexity is 55% (63%) or 69% (72%) as using the ℓ_1- or ℓ_0-regularization (after retraining), respectively. To examine the change in weights after retraining, the normalized difference in weights is defined by $R = \|\mathbf{w}' - \mathbf{w}\|^2 / \|\mathbf{w}'\|^2$, where $\mathbf{w}'$ stands for the weight vector after retraining. Using the results in Fig. 3, the normalized difference in weights is shown in Fig. 4, and the ℓ_0-regularization indeed shows a smaller difference in weights, implying the less necessity of retraining and the better accuracy of weights. Based on the same criterion, Fig. 5(a) and (b) plot the minimum complexities using different ROPs after 100-m and 200-m OM4, respectively. When the ROP and signal-to-noise ratio are lower, a Volterra filter of more complexity would be required to perform nonlinear equalization, in order to meet the same FEC limit; in this case, the regularization scheme also shows the lower ratios of reduction in the complexity. Because avoiding retraining can save extra computational complexity, retraining is not applied in Fig. 5. Compared to the ℓ_1-regularization, the ℓ_0-regularization achieves more reduction in the complexity, and the average differences in the ratio of reduction are 10% and 15% in Fig. 5(a) and (b), respectively, revealing the superiority of the l_0-regularization.

4. Conclusions

In a 112-Gbps PAM-4 VCSEL transmission system, we demonstrate the ℓ_0-regularization as well as the ℓ_1-regularization for the reduction in the complexity of a Volterra filter. The experiment results show that the ℓ_0-regularization can attain better sparse identification, such that the minimum complexity to meet the FEC limit is lower, compared to the ℓ_1-regularization. Besides, when using the ℓ_0-regularization, retraining the weights of a Volterra filter brings less improvement in performance and exhibits less difference in weights, indicating that the ℓ_0-regularization can obtain weights of better accuracy. In the cases without retraining, the ℓ_0-regularization realizes up to 90% reduction in complexity and shows more reduction of 8–18%, compared to the ℓ_1-regularization.

References

[1] "The 2018 Ethernet Roadmap," https://ethernetalliance.org/the-2018-ethernet-roadmap/.
[2] P. Dawe, "400GBASE-SR4.2 optical penalties," http://www.ieee802.org/3/cm/public/adhoc/dawe_3cm_adhoc_01_101118.pdf
[3] T. Zuo et al., "112-Gb/s duobinary 4-PAM transmission over 200-m multi-mode fibre," in *ECOC* 2015.
[4] F. Karinou et al., "112 Gb/s PAM-4 Optical Signal Transmission over 100-m OM4 Multimode Fiber for High-Capacity Data-Center Interconnects," in *ECOC* 2016.
[5] W.-J. Huang et al., "93% Complexity Reduction of Volterra Nonlinear Equalizer by ℓ_1-Regularization for 112-Gbps PAM-4 850-nm VCSEL Optical Interconnect," in *OFC* 2018, paper M2D.7.
[6] C. Chuang et al., "Sparse Volterra Nonlinear Equalizer by Employing Pruning Algorithm for High-Speed PAM-4 850-nm VCSEL Optical Interconnect," in *OFC* 2019, paper M1F.2.
[7] K. Shi et al., "Adaptive sparse Volterra system identification with ℓ_0-norm penalty," Signal Process. **91**, 2432–2436 (2011).
[8] J. Tsimbinos et al., "Computational complexity of Volterra based nonlinear compensators," Electron. Lett. **32**, 852–854 (1996).
[9] J. Yen et al., "850 nm Vertical-Cavity Surface-Emitting Laser Arrays With Enhanced High-Speed Transmission Performance Over a Standard Multimode Fiber," J. Lightw. Tech. **35**, 3242–3249 (2017).