Research Article | Vol. 30, No. 20 / 26 Sep 2022 / Optics Express 36343
Optics EXPRESS

Comparison of low-complexity sparse and weight-sharing nonlinear equalizers for C-band 100-Gbit/s DSB PAM-4 transmission over 60-km SSMF

JUNWEI ZHANG,[1,2,7] HEYUN TAN,[1,8] XIAOJIAN HONG,[3] JIE LIU,[1] CHANGJIAN GUO,[4] CHAO FEI,[3] XIONG WU,[2] ALAN PAK TAO LAU,[5,6] SIYUAN YU,[1] AND CHAO LU[2,6]

[1]*State Key Laboratory of Optoelectronic Materials and Technologies, School of Electronics and Information Technology, Sun Yat-Sen University, Guangzhou 510006, China*
[2]*Photonics Research Center, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China*
[3]*State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering, Zhejiang University, Hangzhou 310058, China*
[4]*South China Academy of Advanced Optoelectronics, South China Normal University, Guangzhou 510006, China*
[5]*Photonics Research Center, Department of Electrical Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China*
[6]*The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518057, China*
[7]*zhangjw253@mail.sysu.edu.cn*
[8]*tanhy35@mail.sysu.edu.cn*

Abstract: To cope with the nonlinear distortions and the chromatic dispersion (CD) induced power fading in double-side band (DSB) intensity modulation and direct detection (IM/DD) transmission systems, high-performance Volterra nonlinear equalizers (VNLEs) including Volterra feed-forward equalizer (VFFE) and Volterra decision-feedback equalizer (VDFE) are widely applied. However, the conventional VNLEs have high computational complexity, especially for longer memory lengths. In this paper, based on sparse and weight-sharing strategies for significant kernel reduction, we propose four low-complexity NLEs including a sparse diagonally pruned VDFE (S-DP-VDFE), a sparse diagonally pruned absolute-term DFE (S-DP-ATDFE), a weight-sharing DP-VDFE (WS-DP-VDFE), and a weight-sharing DP-ATDFE (WS-DP-ATDFE), and present a comprehensive comparison among them in terms of computational complexity and bit error ratio (BER) performance in a C-band 100-Gbit/s PAM-4 transmission system over 60-km standard single-mode fiber (SSMF). The experimental results show that the proposed S-DP-VDFE and WS-DP-VDFE not only exhibit comparable performance with the conventional DP-VDFE but also reduce the complexity by 54.5% and 45.9%, respectively. While the proposed S-DP-ATDFE and WS-DP-ATDFE yield lower complexity at the expense of a slight performance degradation. Compared with the proposed S-DP-VDFE, S-DP-ATDFE, and WS-DP-VDFE, the proposed WS-DP-ATDFE with the lowest number of real-valued multiplications of 45 achieves up to 90.9%, 81.6%, and 95.8% complexity reduction, respectively, at the 7% hard-decision forward error correction (HD-FEC) BER limit of 3.8×10^{-3}. The proposed low-complexity WS-DP-ATDFE shows great potential in low-cost and high-performance IM/DD optical transmission systems.

1. Introduction

To support the rapid growth of data traffic in optical data-center interconnections and access networks, while meeting the strict requirements for low-cost, high energy-efficiency, and small

#468635 https://doi.org/10.1364/OE.468635
Journal © 2022 Received 23 Jun 2022; revised 2 Sep 2022; accepted 6 Sep 2022; published 20 Sep 2022

Optics EXPRESS

form factors, the simple intensity modulation and direct detection (IM/DD) systems with double-side band (DSB) modulation are regarded as the most cost-effective option [1,2]. Among different IM/DD modulation schemes, 4-level pulse amplitude modulation (PAM-4) with an optimal balance between performance and complexity is preferred, which has already been standardized in 400G Ethernet and is a promising solution for the next-generation 800G Ethernet [3–5]. Nevertheless, the achievable transmission distance and capacity of DSB based PAM-4 IM/DD systems in C-band are limited by several linear and nonlinear distortions induced by band-limited components, fiber chromatic dispersion (CD), and nonlinearities related to modulation and detection. Due to frequency selective power fading and nonlinear distortions caused by the interaction of CD and square-law detection, the received IM/DD signal is severely corrupted when the transmission distance and data rate are beyond 50 km and 50 Gbit/s, respectively.

To overcome these transmission impairments, numerous digital signal processing (DSP) based equalization techniques have been proposed. One of the most popular approaches to compensate these linear and nonlinear distortions in IM/DD systems is to use 2^{nd}- or 3^{rd}-order Volterra nonlinear equalizers (VNLEs) by modeling and equalizing the complex time-dispersive nonlinear channel response, which include a Volterra feed-forward equalizer (VFFE) [6–10], and a combination of a VFFE and a Volterra decision-feedback equalizer (VDFE) [11–14]. The VDFE with a nonlinear infinite impulse response (IIR) structure can efficiently equalize the CD-induced spectral nulls by pole insertion and a part of nonlinear distortions [13–15], which results in a significant improvement of transmission performance compared with the case of only employing VFFE. Nonetheless, they tend to have high computational complexity, which grows exponentially as the memory length increases, and thus prohibits their real-time implementation in cost-sensitive IM/DD optical transmission systems.

To reduce the computational complexity in VNLEs, one of the popular strategies is to prune some negligible self/ cross beating terms. By pruning the nonlinear terms with large delays, diagonally pruned VFFEs (DP-VFFEs) [6–9], a joint DP-VFFE and DP-VDFE [12], a polynomial nonlinear FFE (PFFE) [7,16], a PFFE-DFE [17], and a joint PFFE and polynomial nonlinear DFE (PDFE) [5,18] have been investigated and demonstrated in PAM-4 IM/DD systems. To prune more negligible linear and nonlinear terms and further reduce the complexity, sparse VFFEs based on threshold pruning [19], L1-regularization [20,21], a combination of L1- and L2- regularizations [22], and an orthogonal matching pursuit (OMP) greedy algorithm [9,23] have been proposed and implemented in IM/DD transmission systems. Apart from the pruning strategy, the weight-sharing strategy based on *k*-means clustering algorithm has been applied in a PFFE [10], a DP-VFFE [24], and a PFFE-DFE [25] for IM/DD transmissions using PAM signals, showing significant reduction of computational complexity, especially for the absolute-term based NLEs by replacing multiplication terms with absolute terms in a PFFE [10], a DP-VFFE [24], and a PFFE-DFE [25]. However, both the sparse and the weight-sharing strategies have not been employed in high-performance VDFE/ DP-VDFE to reduce the equalization complexity for C-band IM/DD transmissions over dispersion-uncompensated links. Moreover, the comprehensive comparison of equalization performance and computational complexity between the sparse and the weight-sharing NLEs in IM/DD systems has not been reported yet. Such comparison is meaningful for low-cost and high-capacity C-band PAM-based IM/DD systems in data-center interconnect applications.

In this paper, to compensate for both the nonlinear distortions and the CD-induced power fading, we propose and experimentally demonstrate four low-complexity nonlinear equalizers including a sparse DP-VDFE (S-DP-VDFE), a sparse diagonally pruned absolute-term DFE (S-DP-ATDFE), a weight-sharing DP-VDFE (WS-DP-VDFE), and a weight-sharing DP-ATDFE (WS-DP-ATDFE) for a DSB PAM-4 based IM/DD transmission system. The kernel reductions of the proposed S-DP-VDFE and S-DP-ATDFE are realized by performing an OMP algorithm to select the important kernels while weight-sharing kernels of the proposed WS-DP-VDFE and

Optics EXPRESS

WS-DP-ATDFE are obtained by a k-means clustering algorithm to reduce the kernel redundancy, which significantly reduce the equalization complexity. Meanwhile, the multiplication operation is replaced by the absolute operation in the proposed S-DP-ATDFE and WS-DP-ATDFE for further complexity reduction. Both computational complexity and transmission performance of the proposed four equalizers are evaluated and compared with the conventional DP-VDFE and DP-ATDFE in a C-band 100-Gbit/s PAM-4 transmission system over 60-km standard single-mode fiber (SSMF). The experimental results show that the proposed S-DP-VDFE and WS-DP-VDFE not only exhibit comparable performance with the conventional DP-VDFE but also reduce the computational complexity by 54.5% and 45.9%, respectively. The proposed S-DP-ATDFE and WS-DP-ATDFE show similar performance as DP-ATDFE, while lower complexity than DP-VDFE, S-DP-VDFE, and WS-DP-VDFE at the expense of a slight performance degradation. Compared with the proposed S-DP-VDFE, S-DP-ATDFE, and WS-DP-VDFE, the proposed WS-DP-ATDFE with the lowest number of real-valued multiplications of 45 achieves up to 90.9%, 81.6%, and 95.8% complexity reduction, respectively, at the 7% hard-decision forward error correction (HD-FEC) BER limit of 3.8×10^{-3}. Therefore, the proposed low-complexity WS-DP-ATDFE shows great potential in low-cost and high-performance IM/DD optical transmission systems.

2. Principle and computational complexity of sparse and weight-sharing nonlinear equalizers

2.1. DP-VDFE and DP-ATDFE

To effectively mitigate linear and nonlinear transmission distortions in IM/DD systems while maintaining relatively low complexity, a 2$^{\text{nd}}$-order DP-VFFE [7–9] can be implemented as a post-equalizer at the receiver side. In DP-VFFE, the nonlinear kernels with large-interval product terms are truncated to reduce the equalization complexity since they have negligible contribution for compensating the signal distortions. However, the spectral nulls caused by CD-induced power fading cannot be compensated by the DP-VFFE even with a long linear/ nonlinear memory length. To effectively compensate for both nonlinear distortions and CD-induced power fading, a combination of a DP-VFFE and a DP-VDFE is a good choice, in which the VDFE with a IIR structure can efficiently equalize the spectral nulls by pole insertion as well as a part of nonlinear distortions [13–15]. For concision, the joint 2$^{\text{nd}}$-order DP-VFFE and 2$^{\text{nd}}$-order DP-VDFE is entitled as DP-VDFE, whose output can be expressed as [8,9,12]:

$$\begin{aligned} y(n) = & \underbrace{\sum_{k=0}^{N_1-1} h_1(k)x(2n-k) + \sum_{q=0}^{Q-1}\sum_{k=0}^{N_2-1-q} h_2(k,q)x(2n-k)x(2n-k-q)}_{\text{DP - VFFE}} \\ & + \underbrace{\sum_{k=1}^{D_1} w_1(k)d(n-k) + \sum_{u=0}^{U-1}\sum_{k=1}^{D_2-u} w_2(k,u)d(n-k)d(n-k-u)}_{\text{DP - VDFE}} \end{aligned} \tag{1}$$

where $x(n)$ with $T/2$ symbol space is the received DD signal corrupted by nonlinear distortions and power fading, $d(n)$ with T symbol space is the previous hard-decision output, $h_K$and N_K are the K^{th}-order ($K = 1, 2$) kernel and memory length of the VFFE, respectively, w_K and D_K are the K^{th}-order kernel and memory length of the VDFE, respectively. Q and U represent the pruning factors, which are utilized to prune the unimportant cross-beating terms of the VFFE and VDFE, respectively. Note that here a fractionally spaced DP-VFFE with $T/2$ symbol space is used to

Optics EXPRESS

provide robustness against clock jitter [17] while T symbol space is applied to DP-VDFE. The DP-VDFE becomes the conventional linear FFE-DFE and PFFE-PDFE when $N_2 = D_2 = 0$ and $Q = U = 1$, respectively.

By replacing the cross-beating terms in DP-VDFE by the terms formed of the absolute value of a sum of two input samples [8,9], the equalized output of the DP-ATDFE can be written as:

$$y(n) = \underbrace{\sum_{k=0}^{N_1-1} h_1(k)x(2n-k) + \sum_{q=0}^{Q-1}\sum_{k=0}^{N_2-1-q} h_2(k,q)|x(2n-k) + x(2n-k-q)|}_{\text{DP - ATFFE}} + \underbrace{\sum_{k=1}^{D_1} w_1(k)d(n-k) + \sum_{u=0}^{U-1}\sum_{k=1}^{D_2-u} w_2(k,u)|d(n-k) + d(n-k-u)|}_{\text{DP - ATDFE}} \tag{2}$$

where $|\bullet|$ denotes the absolute operator. Applying the Taylor series expansion of $|x|$ at $x = 0$, $|x| \approx x^2/2 - x^4/8$ can be obtained. Therefore, the $|x(2n) + x(2n-q)|$ and $|d(n) + d(n-u)|$ not only contain the nonlinear terms $x(2n)x(2n-q)$ and $d(n)d(n-u)$ in DP-VDFE, respectively, but also reduce the computational complexity since the absolute operation can be realized by an addition operation [26] instead of a multiplication operation. The total number of kernels in DP-VDFE or DP-ATDFE is $L = N_1 + Q(2N_2\text{-}Q\text{+}1)/2 + D_1 + U(2D_2\text{-}U\text{+}1)/2$.

For simplicity, we rewrite the DP-VDFE/ DP-ATDFE of Eq. (1)/ Eq. (2) in a matrix form as:

$$y(n) = \mathbf{h}^T\mathbf{x}_n = \mathbf{x}_n^T\mathbf{h} \tag{3}$$

where $(\bullet)^T$ stands for transpose operation, $\mathbf{x}_n$ and $\mathbf{h}$ are $L\times1$ input signal vector and $L\times1$ kernel vector, respectively, which are given by:

$$\mathbf{x}_n = [x_0(n), x_1(n), \cdots, x_{L-1}(n)]^T \tag{4}$$

$$\mathbf{h} = [h(0), h(1), \cdots, h(L-1)]^T \tag{5}$$

where $h(l)$ is the kernel composed of h_1, h_2, w_1, and w_2 in Eq. (1)/ Eq. (2), whose corresponding rewritten term $x_l(n)$ is composed of $x(2n-k)$, $x(2n-k)x(2n-k-q)$, $d(n-k)$, and $d(n-k)d(n-k-u)$ for DP-VDFE, or $x(2n-k)$, $|x(2n-k) + x(2n-k-q)|$, $d(n-k)$, and $|d(n-k) + d(n-k-u)|$ for DP-ATDFE. The detailed derivations of $\mathbf{x}_n$ and $\mathbf{h}$ for DP-VDFE and DP-ATDFE are presented in Appendix.

In the training phase, the kernels in Eq. (3) can be obtained by recursive least-squares (RLS) algorithm [27] using a M-point training sequence $t(n)$ with $0 \le n \le M-1$, which is summarized as follows:

1) Initialize the kernel vector $\mathbf{h}(0) = \mathbf{0}$, the inverse of the correlation matrix $\mathbf{P} = \delta^{-1}\mathbf{I}$, where δ is a small positive constant and $\mathbf{I}$ is an identity matrix with a rank of L, and the iteration counter $n = 1$.
2) Calculate the error $e(n) = t(n) - \mathbf{h}^T(n-1)\mathbf{x}_n$.
3) Compute the gain vector $\mathbf{k}(n) = \mathbf{P}(n-1)\mathbf{x}_n\{\lambda + \mathbf{x}_n^T\mathbf{P}(n-1)\mathbf{x}_n\}^{-1}$ and update the matrix $\mathbf{P}(n) = \lambda^{-1}\{\mathbf{P}(n-1) - \mathbf{k}(n)\mathbf{x}_n^T\mathbf{P}(n-1)\}$.
4) Update the kernel vector $\mathbf{h}(n) = \mathbf{h}(n-1) + \mathbf{k}(n)e(n)$.
5) If $n < M$-1, make the value of the iteration counter $n = n + 1$ and return to Step 2.
6) After obtaining the kernel coefficients, the equalizers are switched to the equalization phase with DP-VDFE or DP-ATDFE.

Optics EXPRESS

2.2. Proposed S-DP-VDFE and S-DP-ATDFE

Since linear and nonlinear responses of the IM/DD transmission systems are sparse [9,19–23], it is feasible to only adopt the most significant kernels while discarding the unimportant ones for equalization. To further reduce the computational complexity of nonlinear equalization, here we propose low-complexity S-DP-VDFE and S-DP-ATDFE, which can be expressed as:

$$y(n) = \mathbf{x}_n^T(\mathbf{\Lambda})\mathbf{h}_s(\mathbf{\Lambda}) \tag{6}$$

where $\mathbf{h}_s$ is $L\times1$ sparse kernel vector, the optimal index set $\mathbf{\Lambda}$ with a size of P is used to determine the sparse kernels and the corresponding input signal terms from the original DP-VDFE/ DP-ATDFE, and P is the sparsity level as well as the total number of sparse kernels. Considering that M-point training sequence is transmitted before data transmission, the input-output relationship of DP-VDFE/ DP-ATDFE can be expressed as:

$$\mathbf{t} = \mathbf{X}\mathbf{h} \tag{7}$$

where $\mathbf{t} = [t(0), t(1), \cdots, t(M-1)]^T$ is the transmitted training vector, $\mathbf{X}$ with a size of $M\times L$ is the measurement matrix defined as:

$$\mathbf{X} = [\mathbf{v}_0, \mathbf{v}_1, \cdots, \mathbf{v}_{L-1}] = [\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_{M-1}]^T \tag{8}$$

where $\mathbf{v}_l$ is the l^{th} column vector of measurement matrix $\mathbf{X}$. The kernels of S-DP-VDFE/ S-DP-ATDFE can be obtained in the training phase using the OMP greedy algorithm [9,23,28,29], which adaptively searches for a good approximation of the transmitted training vector $\mathbf{t}$ by the linear combination of a small number of columns from the measurement matrix $\mathbf{X}$. The detailed procedures of OMP based kernel estimation [9,23,28,29] are summarized as:

1) Initialize the residual $\mathbf{b}_0 = \mathbf{t}$, the index set $\mathbf{\Lambda}_0 = \phi$, the matrix of chosen atoms $\mathbf{S}_0 = \phi$, the iteration counter $p = 1$, and the sparsity level P.
2) Find the index $k_p = \arg\max_l |\mathbf{v}_l^H \mathbf{b}_{p-1}|$, where $(\bullet)^H$ stands for conjugate transpose, and $l = 0, 1, \cdots, L-1$ and $l \notin \mathbf{\Lambda}_{p-1}$.
3) Augment the index set $\mathbf{\Lambda}_p = \mathbf{\Lambda}_{p-1} \cup k_p$ and the matrix of chosen atoms $\mathbf{S}_p = [\mathbf{S}_{p-1}, \mathbf{v}_{k_p}]$.
4) Solve the LS problem by calculating $\mathbf{h'}_p = (\mathbf{S}_p^H \mathbf{S}_p)^{-1} \mathbf{S}_p^H \mathbf{t}$ and update the new residual as $\mathbf{b}_p = \mathbf{t} - \mathbf{S}_p \mathbf{h'}_p$, where $(\bullet)^{-1}$ represents matrix inversion.
5) If $p < P$, make the value of the iteration counter $p = p + 1$ and return to Step 2.
6) The optimal index set $\mathbf{\Lambda} = \mathbf{\Lambda}_P$ and the kernel vector $\mathbf{h}_s(\mathbf{\Lambda}) = \mathbf{h'}_p$ with P nonzero kernels are obtained for sparse equalization with S-DP-VDFE or S-DP-ATDFE.

2.3. Proposed WS-DP-VDFE and WS-DP-ATDFE

In addition to the sparse strategy, the computational complexity of nonlinear equalization can be further reduced by weight-sharing strategy, in which the linear and nonlinear kernels with similar values can be grouped into a new kernel for equalization [10,24,25]. Here the k-means clustering algorithm [25,30] is employed for the estimated kernels of the DP-VDFE/ DP-ATDFE and the resulted cluster centers serve as new kernels for weight-sharing equalization with the proposed WS-DP-VDFE/ WS-DP-ATDFE, which is given by:

$$y(n) = \sum_{i=0}^{C-1} h_c(i) x_c(n,\ i) \tag{9}$$

where $h_c(i)$ and C are the cluster centers (i.e., new kernels after clustering) and the number of clusters obtained by the k-means clustering algorithm, respectively. $x_c(n, i)$ is the sum of the

Optics EXPRESS

corresponding input signal terms in $\mathbf{x}_n$, with their kernels belonging to the i^{th} cluster with a cluster center of $h_c(i)$. In summary, the detailed procedures of k-means clustering algorithm based kernel estimation [25,30] are as follows:

1) Randomly select C kernels from the estimated kernel dataset $H = \{h(0), h(1), \cdots, h(L-1)\}$ as the initial cluster centers dataset $H_c = \{h_c(0), h_c(1), \cdots, h_c(C-1)\}$. Note that the kernels are firstly estimated based on the RLS algorithm similar as in DP-VDFE/ DP-ATDFE.

2) Calculate the Euclidean distance between each kernel $h(l)$ and all C cluster centers $h_c(i)$ by $d_{l,i} = ||h(l) - h_c(i)||^2 = |h(l) - h_c(i)|$, and then assign it to the nearest cluster.

3) Update the cluster centers by calculating the mean of each cluster.

4) Steps 2 and 3 are repeated until the standard metric function, i.e., the within-cluster sum of squares $E = \sum_{i=0}^{C-1} \sum_{h(l)\in h_c(i)} ||h(l) - h_c(i)||^2 = \sum_{i=0}^{C-1} \sum_{h(l)\in h_c(i)} |h(l) - h_c(i)|$, converges.

5) The final cluster-center vector $\mathbf{h}_c = [h_c(0), h_c(1), \cdots, h_c(C-1)]^T$ are obtained for weight-sharing equalization with WS-DP-VDFE or WS-DP-ATDFE.

Figures 1(a), 1(b), and 1(c) show the schematic diagrams of the DP-VDFE/ DP-ATDFE, the proposed S-DP-VDFE/ S-DP-ATDFE, and WS-DP-VDFE/ WS-DP-ATDFE in equalization process, respectively. It can be observed that different from the sparse equalization to realize complexity reduction by discarding the trivial kernels, the weight-sharing equalization explores fewer kernels to realize high-efficiency equalization by sharing the same kernel with different input signal terms. Compared with the conventional DP-VDFE/ DP-ATDFE, the received distorted signals can be equalized by the proposed S-DP-VDFE/ S-DP-ATDFE and WS-DP-VDFE/ WS-DP-ATDFE using sparse strategy and weight-sharing strategy, respectively, with much fewer kernel coefficients. Note that both the sparse and weight-sharing strategies can be applied for full VDFE/ ATDFE, which requires more number of training symbols to achieve high-accuracy kernel estimation in training phase owing to the increased number of kernels.

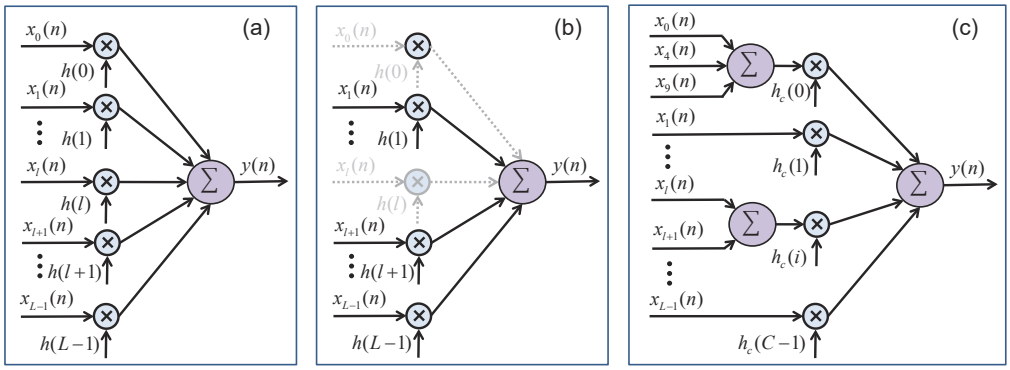

Fig. 1. Schematic diagrams of (a) DP-VDFE/ DP-ATDFE, (b) S-DP-VDFE/ S-DP-ATDFE, and (c) WS-DP-VDFE/ WS-DP-ATDFE.

2.4. Complexity analysis

The computational complexity of the nonlinear equalizer, which is strongly correlated to the power consumption, is one of the most significant factors for its real-time application and can be quantified by the required number of real-valued multiplications per symbol (RNRM). Since the

Optics EXPRESS

fiber channel is quasi-invariant for IM/DD systems, the training process is not necessary once the kernel coefficients of the equalizers are estimated [4–25]. In addition, the sparse and weight-sharing kernel coefficients can be further updated by the RLS algorithm with their corresponding known optimal index set and clusters, respectively, which further reduces the computational complexity in training process. Therefore, we mainly consider the computational complexity in equalization process and the complexity results of the six abovementioned equalizers are presented in Table 1. The DP-VDFE needs $N_1 + Q(2N_2\text{-}Q+1)$ and $D_1 + U(2D_2\text{-}U+1)$ real-valued multiplications in feed-forward and decision-feedback stages, respectively. After performing absolute operation, the RNRM of DP-ATDFE is reduced to $N_1 + Q(2N_2\text{-}Q+1)/2$ and $D_1 + U(2D_2\text{-}U+1)/2$ in feed-forward and decision-feedback stages, respectively. Thanks to significant kernel reduction, the proposed S-DP-ATDFE and WS-DP-ATDFE need only P ($P < L$) and C ($C < L$) real-valued multiplications regardless of the memory length, respectively, which are much lower than those of the conventional DP-VDFE and DP-ATDFE. Moreover, the computational complexity of the proposed S-DP-ATDFE/ WS-DP-ATDFE is also lower than that of the S-DP-VDFE, which requires $W+2(P\text{-}W)$ real-valued multiplications, where W is the number of significant linear kernels in sparse DP-VDFE. With respect to WS-DP-VDFE, $Q(2N_2\text{-}Q+1)/2 + U(2D_2\text{-}U+1)/2$ real-valued multiplications are required to obtain the self/ cross beating terms (i.e., $x(2n)x(2n-q)$ and $d(n)d(n-u)$) before performing weight-sharing equalization. The detailed analyses of equalization performance and complexity among the abovementioned NLEs will be presented in section 4.

Table 1. Complexity comparison of different nonlinear equalizers

Equalizer	RNRM	
	Feed-forward stage	Decision-feedback stage
DP-VDFE	$N_1 + Q(2N_2\text{-}Q+1)$	$D_1 + U(2D_2\text{-}U+1)$
DP-ATDFE	$N_1 + Q(2N_2\text{-}Q+1)/2$	$D_1 + U(2D_2\text{-}U+1)/2$
Proposed S-DP-VDFE	$W+2(P\text{-}W)$	
Proposed S-DP-ATDFE	P	
Proposed WS-DP-VDFE	$C + Q(2N_2\text{-}Q+1)/2 + U(2D_2\text{-}U+1)/2$	
Proposed WS-DP-ATDFE	C	

3. Experimental setup

The performance of the proposed sparse and weight-sharing equalizers is evaluated in a C-band 100-Gbit/s PAM-4 transmission system over 60-km SSMF. Figure 2 illustrates the experimental setup and the DSP block diagram. At the transmitter, a pseudo random bit sequence (PRBS) is firstly generated and mapped to PAM-4 symbols. Then rectangular pulse shaping with 4 samples per symbol is applied and the resulted PAM-4 signal is loaded into an arbitrary waveform generator (AWG, Keysight 8199A) operating at a sample rate of 200 GSa/s. The generated 50-GBaud PAM-4 electrical signal from the AWG is amplified by an electrical amplifier (SHF S807) and used to drive a 35-GHz Mach-Zehnder modulator (MZM, Thorlabs LN05S-FC) for DSB electrical-optical conversion. The optical carrier generated from an external cavity laser (ECL) with a center wavelength of 1550.12 nm is injected to the MZM biased at a quadrature point of around 1.4 V. After 60-km SSMF transmission without any dispersion compensation, a variable optical attenuator (VOA) is employed to adjust the received optical power (ROP) of the received signal, which is then detected by a photodetector (PD, XPDV2120RA). Due to the lack of a trans-impedance amplifier (TIA) after PD, an Erbium doped fiber amplifier (EDFA) followed by an optical band-pass filter (OBPF) is applied to boost the power into the PD up to 7 dBm. Finally, the detected electrical PAM-4 signal is analog-to-digital converted by a digital

Optics EXPRESS

storage oscilloscope (DSO) operating at a sample rate of 160 GSa/s for subsequent off-line DSP procedures including resampling to 2 samples per symbol, synchronization, half-symbol-spaced equalization with the proposed sparse and weight-sharing equalizers, symbol decision, PAM-4 demapping, and BER calculation. In the training process, we use 10000 training symbols to obtain the kernel coefficients of different equalizers, which keep unchanged in the equalization process.

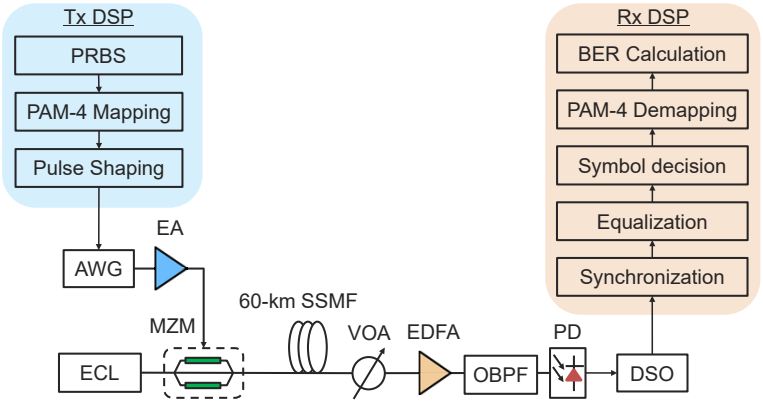

Fig. 2. Experimental setup of 100-Gbit/s PAM-4 IM/DD system. AWG: arbitrary waveform generator; EA: electrical amplifier; MZM: Mach-Zehnder modulator; ECL: external cavity laser; SSMF: standard single-mode fiber; VOA: variable optical attenuator; EDFA: Erbium doped fiber amplifier; OBPF: optical band-pass filter; PD: photodetector; DSO: digital storage oscilloscope.

4. Results and analysis

We first optimize the parameters including memory length, pruning factor, sparsity level, and the number of clusters of different equalizers in a 100-Gbit/s IM/DD PAM-4 system over 60-km SSMF at a ROP of -10 dBm. Figure 3 shows the measured BER versus the memory length N_1 at different D_1 for FFE-DFE (i.e., the linear part of DP-VDFE and DP-ATDFE). It is observed that the BER performance of the FFE-DFE is firstly improved with the increase of N_1 or D_1 and then reaches the optimal value when $N_1 = 74$ and $D_1 = 32$. Besides, the FFE-DFE outperforms the FFE owing to the effective compensation of CD-induced power fading with the DFE.

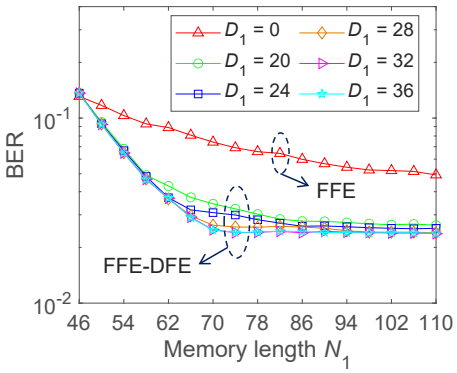

Fig. 3. Measured BER versus linear memory length N_1 at different D_1 for FFE-DFE.

Optics EXPRESS

Next, with $N_1 = 74$ and $D_1 = 32$ fixed, we start to optimize the pruning factors Q and the nonlinear memory lengths N_2 of a DP-VFFE-DFE and a DP-ATFFE-DFE, which are the simplified equalizers of DP-VDFE and DP-ATDFE with $D_2 = 0$, respectively. The measured BER as a function of the pruning factor Q with a sufficiently long nonlinear memory length $N_2 = 74$ is shown in Fig. 4(a). One can see that the BER reduces significantly with a larger pruning factor Q, especially for the DP-VFFE-DFE. Compared with FFE-DFE, the BERs of both the DP-VFFE-DFE and the DP-ATFFE-DFE can be reduced to below 1×10^{-2}. Moreover, the performance of the DP-VFFE-DFE is slightly better than that of DP-ATFFE-DFE, at the cost of higher computational complexity. To balance the performance and complexity, $Q = 13$ is chosen for both the DP-VFFE-DFE and the DP-ATFFE-DFE. Figure 4(b) shows the measured BER versus the nonlinear memory length N_2 with $Q = 13$. The performance is improved with a larger nonlinear memory length N_2 and the BER comes to a standstill at $N_2 = 74$. Similarly, it can be observed from Figs. 4(c) and 4(d) that the optimal BERs are achieved when the pruning factor U and the nonlinear memory length D_2 of the DP-VDFE and the DP-ATDFE are selected as $U = 9$ and $D_2 = 20$, respectively. With nonlinear DFEs performed, the achieved BERs of both the DP-VDFE and the DP-ATDFE are further reduced to below the 7% HD-FEC limit of 3.8×10^{-3}, indicating that the DP-VDFE and the DP-ATDFE are superior to the DP-VFFE-DFE and the DP-ATFFE-DFE in nonlinear compensation.

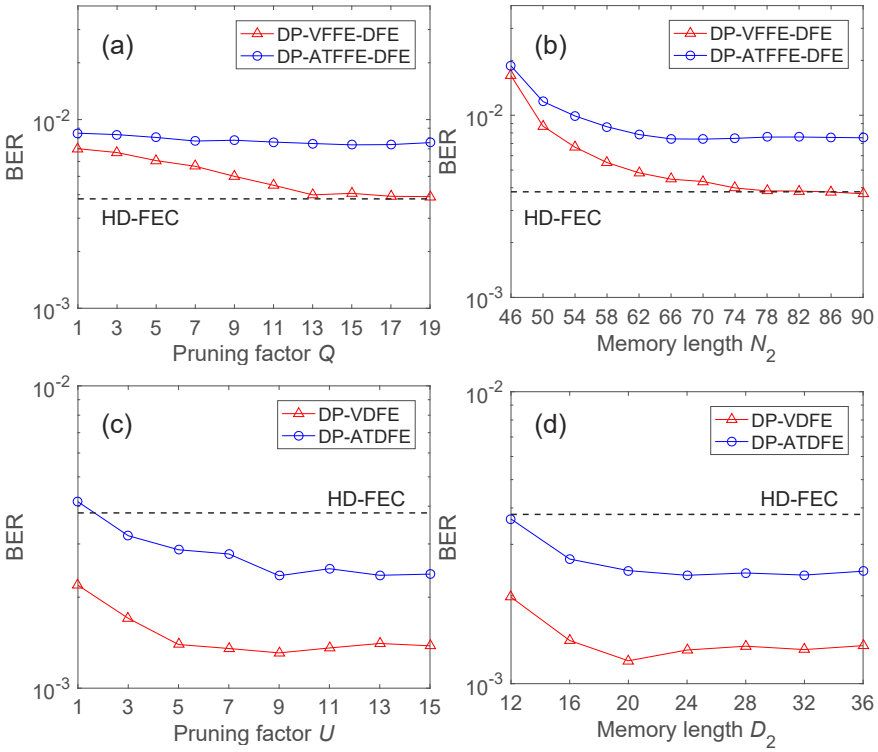

Fig. 4. (a) Measured BER versus pruning factor Q for DP-VFFE-DFE and DP-ATFFE-DFE with $N_1 = 74$, $N_2 = 74$, and $D_1 = 32$; (b) measured BER versus memory length N_2 for DP-VFFE-DFE and DP-ATFFE-DFE with $N_1 = 74$, $Q = 13$, and $D_1 = 32$; (c) measured BER versus pruning factor U for DP-VDFE and DP-ATDFE with $N_1 = 74$, $N_2 = 74$, $Q = 13$, $D_1 = 32$, and $D_2 = 32$; (d) measured BER versus memory length D_2 for DP-VDFE and DP-ATDFE with $N_1 = 74$, $N_2 = 74$, $Q = 13$, $D_1 = 32$, and $U = 9$.

Optics EXPRESS

Based on the parameters $N_1 = 74$, $N_2 = 74$, $Q = 13$, $D_1 = 32$, $D_2 = 20$, and $U = 9$, the measured BER as a function of sparsity level P for S-DP-VDFE and S-DP-ATDFE after 60-km SSMF transmission at a ROP of -10 dBm is shown in Fig. 5(a). It can be seen that the performance of S-DP-VDFE and S-DP-ATDFE are significantly improved with a larger sparsity level P before reaching the optimal BERs of around 1.2×10^{-3} and 2.4×10^{-3} at $P = 535$ and $P = 330$, respectively, which reduce the number of kernels by 52.8% and 70.9% without any performance degradation, compared with the DP-VDFE and DP-ATDFE with 1134 kernels. To reach the 7% HD-FEC threshold of 3.8×10^{-3}, the required number of kernels of S-DP-VDFE and S-DP-ATDFE are 280 and 245 at least, respectively. The corresponding weight distributions of kernel vector $\mathbf{h}/\ \mathbf{h}_s$ for DP-VDFE, DP-ATDFE, S-DP-VDFE with $P = 280$, and S-DP-ATDFE with $P = 245$ are depicted in Figs. 5(b) to 5(e), respectively. It can be observed that after OMP algorithm performed, only the significant kernels are estimated and reserved for sparse equalization.

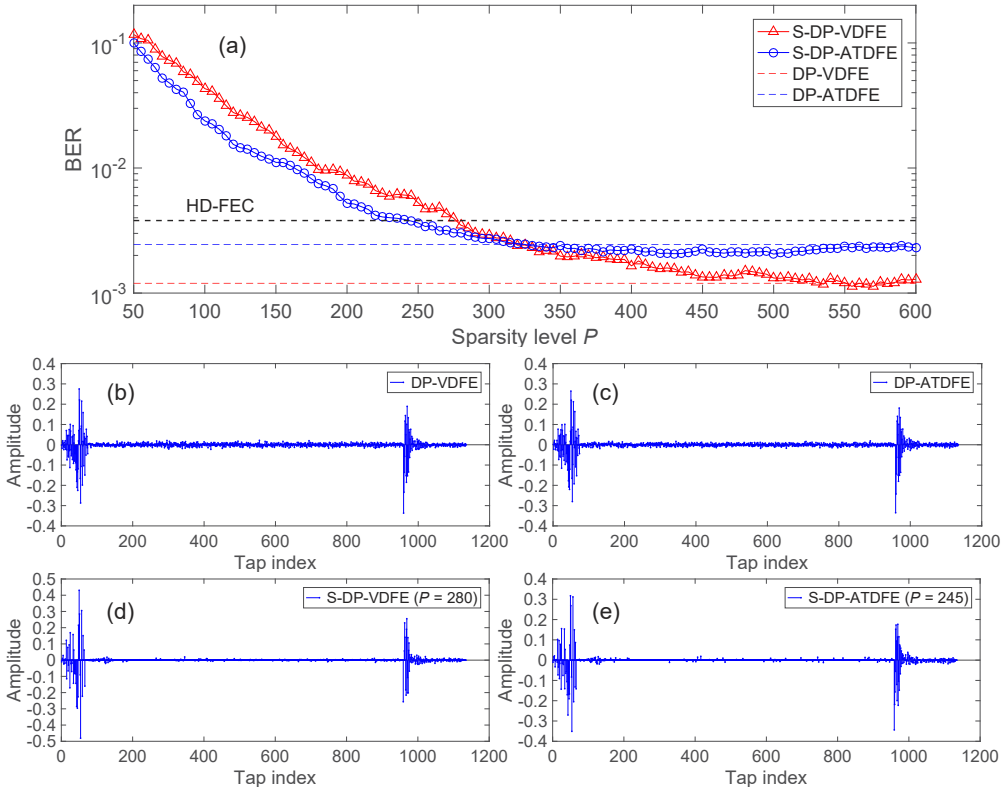

Fig. 5. (a) Measured BER versus sparsity level P for sparse nonlinear equalizers; corresponding weight distributions of kernel vector $\mathbf{h}/\ \mathbf{h}_s$ for (b) DP-VDFE, (c) DP-ATDFE, (d) S-DP-VDFE with $P = 280$, and (e) S-DP-ATDFE with $P = 245$.

The number of cluster centers C and their values (i.e., kernels of weight-sharing equalizers) of the WS-DP-VDFE and the WS-DP-ATDFE are further optimized based on the estimated kernels of the DP-VDFE and the DP-ATDFE, respectively. Figure 6 shows the measured BER versus the number of clusters for the WS-DP-VDFE and the WS-DP-ATDFE using k-means clustering algorithm after 60-km SSMF transmission at a ROP of -10 dBm. To show the superiority of weight-sharing scheme, the weight distributions of all kernels of the DP-VDFE and the DP-ATDFE presented in Figs. 5(b) and 5(c) are rearranged by values as shown in Fig. 6(b),

Optics EXPRESS

and are compared with those of the WS-DP-VDFE with $C = 35$ and the WS-DP-ATDFE with $C = 45$ as shown in Fig. 6(c). The following observations could be made from Figs. 6(a)–6(c): 1) The BER firstly decreases as the number of cluster centers (i.e., kernels) increases for both the WS-DP-VDFE and the WS-DP-ATDFE. 2) To achieve a BER of around 1.2×10^{-3} (2.4×10^{-3}) as same as the DP-VDFE (DP-ATDFE), the number of kernels of WS-DP-VDFE (WS-DP-ATDFE) is significantly reduced from 1134 to 75 (140). 3) To satisfy the 7% HD-FEC BER limit, the number of kernels of the WS-DP-VDFE and the WS-DP-ATDFE can be reduced to 35 and 45 at least, respectively. 4) As shown in Fig. 6(b), this significant reduction of kernels for weight-sharing equalization is attributed to the fact that a large number of kernels in DP-VDFE or DP-ATDFE have similar weight values, which can be clustered into a smaller number of new kernels to reduce the kernel redundancy.

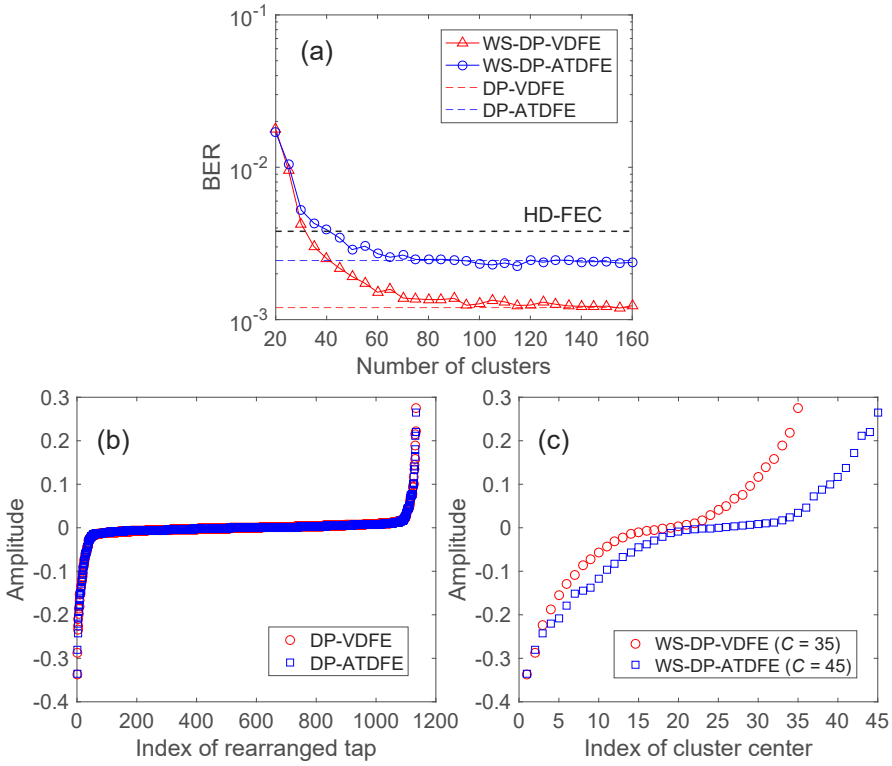

Fig. 6. (a) Measured BER versus the number of clusters for weight-sharing nonlinear equalizers; (b) weight distributions of rearranged kernels for DP-VDFE and DP-ATDFE; (c) weight distributions of cluster centers for WS-DP-VDFE with $C = 35$ and WS-DP-ATDFE with $C = 45$.

The equalization performance and complexity comparisons between the proposed sparse and weight-sharing equalizers are also evaluated after 60-km SSMF at a ROP of -10 dBm. Based on the complexity calculation presented in Table 1, the RNRMs of the proposed sparse and weight-sharing NLEs are obtained by varying the sparsity level and the number of cluster centroids, respectively, with $N_1 = 74$, $N_2 = 74$, $Q = 13$, $D_1 = 32$, $D_2 = 20$, and $U = 9$ fixed. The measured BER as a function of RNRM for different equalizers are shown in Fig. 7. It can be summarized as follows: 1) Thanks to the absolute operation instead of the multiplication operation and the reduction of kernel redundancy, the proposed WS-DP-ATDFE requires the lowest number of real-valued multiplications of 45 among four proposed equalizers, which reduces the RNRM

Optics EXPRESS

by 90.9%, 81.6%, and 95.8% to reach the 7% HD-FEC BER limit, in comparison with 492, 245, and 1063 real-valued multiplications for the proposed S-DP-VDFE, S-DP-ATDFE, and WS-DP-VDFE, respectively. 2) At the BER of around 1.2×10^{-3}, the proposed S-DP-VDFE (WS-DP-VDFE) needs 983 (1169) real-valued multiplications, which lowers the RNRM by 54.5% (45.9%) compared with the DP-VDFE with 2162 real-valued multiplications. 3) At the BER of around 2.4×10^{-3}, the proposed S-DP-ATDFE (WS-DP-ATDFE) needs 330 (75) real-valued multiplications, which saves 70.9% (93.4%) of RNRM compared to the DP-ATDFE with 1134 real-valued multiplications. 4) Among the four proposed equalizers, the RNRM of WS-DP-VDFE is the highest due to the requirement of calculation of the self/ cross beating terms, which consumes 1028 real-valued multiplications.

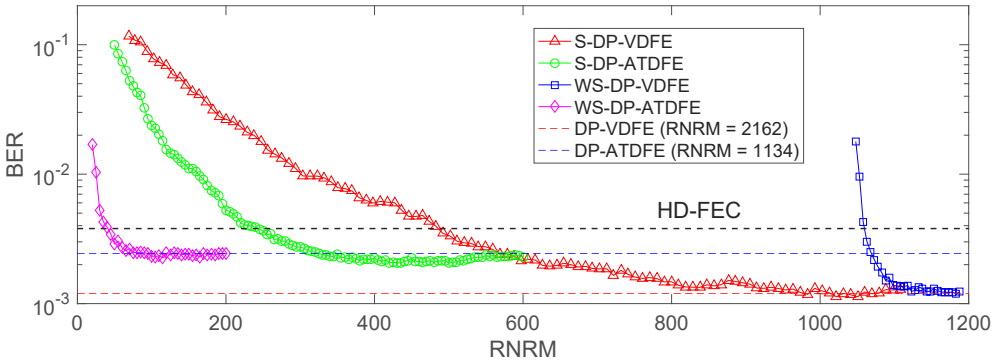

Fig. 7. Measured BER as a function of RNRM for different equalizers after 60-km SSMF transmission at a ROP of -10 dBm.

Finally, we evaluate the system transmission performance based on the proposed S-DP-VDFE with $P = 535$ and WS-DP-ATDFE with $C = 75$. For concision, the results of the WS-DP-VDFE and the S-DP-ATDFE with similar equalization performance but higher computational complexity over the S-DP-VDFE and the WS-DP-ATDFE, respectively, are omitted. Figure 8(a) shows the measured BER versus the ROP for 100-Gbit/s PAM-4 signal transmission over 60-km SSMF using different equalizers, in which the results of the conventional FFE-DFE, DP-VFFE-DFE, DP-ATFFE-DFE, DP-VDFE, and DP-ATDFE are presented for comparison with abovementioned optimized parameters. It can be seen that: 1) The BERs of DP-VFFE-DFE and DP-ATFFE-DFE cannot reach the 7% HD-FEC limit of 3.8×10^{-3} when the ROP is lower than -10 dBm, owing to the fact that the residual nonlinear distortions cannot be compensated by nonlinear FFE. 2) With nonlinear DFE performed, the achieved BERs of DP-VDFE, DP-ATDFE, S-DP-VDFE, and WS-DP-ATDFE are further reduced to below the 7% HD-FEC limit of 3.8×10^{-3}. 3) Compared with the DP-VFFE-DFE, around 3 dB, 2.9 dB, 1.6 dB, and 1.5 dB improvements of the receiver sensitivity are achieved by the DP-VDFE, S-DP-VDFE, DP-ATDFE, and WS-DP-ATDFE, respectively. 4) The achieved BER of the proposed WS-DP-ATDFE (S-DP-VDFE) is close to that of the DP-ATDFE (DP-VDFE), but the former saves 93.4% (54.5%) of RNRM. 5) The proposed S-DP-VDFE outperforms the proposed WS-DP-ATDFE and achieves around 1.4 dB improvement of the receiver sensitivity, at the expense of 92.4% improvement of computational complexity. 6) Taking both equalization performance and computational complexity into account, the proposed WS-DP-ATDFE with RNRM $<$ 100 is a promising choice for real-time implementation of low-complexity applications. To clearly show the performance improvement after using the proposed WS-DP-ATDFE and S-DP-VDFE, the recovered eye diagrams of 100-Gbit/s PAM-4 signals with different equalizers at a ROP of -10 dBm are shown in Figs. 8(b)-(h). Taking the advantages of simultaneously equalizing both the pre-cursor and the post-cursor nonlinear

Optics EXPRESS

distortions, the eye diagrams of DP-VDFE, DP-ATDFE, S-DP-VDFE, and WS-DP-ATDFE are clearer than those of FFE-DFE, DP-VFFE-DFE, and DP-ATFFE-DFE.

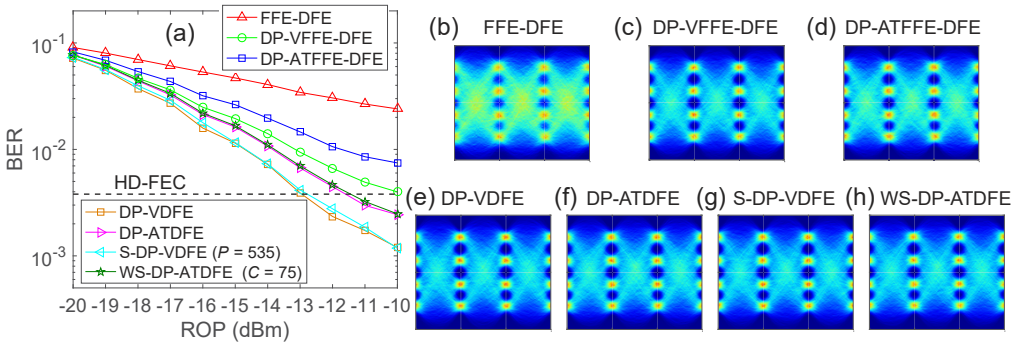

Fig. 8. (a) Measured BER versus ROP using different equalizers after 60-km SSMF transmission; (b)-(h) recovered eye diagrams of PAM-4 signals with different equalizers at a ROP of -10 dBm.

5. Conclusion

In this paper, to reduce the computational complexity of conventional VDFE, we have proposed and experimentally demonstrated four low-complexity NLEs including the S-DP-VDFE, S-DP-ATDFE, WS-DP-VDFE, and WS-DP-ATDFE to eliminate both the CD-induced power fading and the nonlinear distortions in a C-band 100-Gbit/s PAM-4 transmission system over 60-km SSMF. The proposed S-DP-VDFE and WS-DP-VDFE not only exhibit comparable performance with the conventional DP-VDFE but also reduce the complexity by 54.5% and 45.9%, respectively. While the proposed S-DP-ATDFE and WS-DP-ATDFE yield suboptimal performance while further reduce the complexity. Among the proposed four equalizers, the WS-DP-ATDFE with the lowest complexity saves > 80% of RNRM, compared with other three equalizers at the 7% HD-FEC BER limit of 3.8×10^{-3}. Therefore, the proposed sparse and weight-sharing equalizers with low complexity and high equalization performance, especially the WS-DP-ATDFE with ultra-low complexity, are promising to be applied in the low-cost and high-performance IM/DD optical transmission systems.

Appendix

In this section, we present the derivations of the input signal vector $\mathbf{x}_n$ and the kernel vector $\mathbf{h}$. For DP-VDFE as depicted in Eq. (1), we first define the following vectors:

$$\boldsymbol{\alpha}_1 = [x(2n), x(2n-1), \cdots, x(2n-N_1+1)] \tag{10}$$

$$\begin{aligned}\boldsymbol{\alpha}_2 = [&x(2n)x(2n), x(2n-1)x(2n-1), \cdots, x(2n-N_2+1)x(2n-N_2+1),\\ &x(2n)x(2n-1), x(2n-1)x(2n-2), \cdots, x(2n-N_2+2)x(2n-N_2+1), \cdots,\\ &x(2n)x(2n-Q+1), x(2n-1)x(2n-Q), \cdots, x(2n-N_2+Q)x(2n-N_2+1)]\end{aligned} \tag{11}$$

$$\boldsymbol{\beta}_1 = [d(n-1), d(n-2), \cdots, d(n-D_1)] \tag{12}$$

$$\begin{aligned}\boldsymbol{\beta}_2 = [&d(n-1)d(n-1), d(n-2)d(n-2), \cdots, d(n-D_2)d(n-D_2),\\ &d(n-1)d(n-2), d(n-2)d(n-3), \cdots, d(n-D_2+1)d(n-D_2), \cdots,\\ &d(n-1)d(n-U), d(n-2)d(n-U-1), \cdots, d(n-D_2+U-1)d(n-D_2)]\end{aligned} \tag{13}$$

$$\mathbf{h}_1 = [h_1(0), h_1(1), \cdots, h_1(N_1-1)] \tag{14}$$

Optics EXPRESS

$$\mathbf{h}_2 = [h_2(0,0), h_2(1,0), \cdots, h_2(N_2-1,0), h_2(0,1), h_2(1,1), \cdots, h_2(N_2-2,1), \cdots, h_2(0,Q-1), h_2(1,Q-1), \cdots, h_2(N_2-Q,Q-1)] \tag{15}$$

$$\mathbf{w}_1 = [w_1(1), w_1(2), \cdots, w_1(D_1)] \tag{16}$$

$$\mathbf{w}_2 = [w_2(1,0), w_2(2,0), \cdots, w_2(D_2,0), w_2(1,1), w_2(2,1), \cdots, w_2(D_2-1,1), \cdots, w_2(1,U-1), w_2(2,U-1), \cdots, w_2(D_2-U+1,U-1)] \tag{17}$$

Then $L\times1$ $\mathbf{x}_n$ and $L\times1$ $\mathbf{h}$ can be obtained and expressed as:

$$\mathbf{x}_n = [x_0(n), x_1(n), \cdots, x_{L-1}(n)]^T = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2]^T \tag{18}$$

$$\mathbf{h} = [h(0), h(1), \cdots, h(L-1)]^T = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{w}_1, \mathbf{w}_2]^T \tag{19}$$

Similar expression of input signal vector $\mathbf{x}_n$ for DP-ATDFE can be deduced by replacing the terms $x(2n)x(2n-q)$ and $d(n)d(n-u)$ with $|x(2n)+x(2n-q)|$ and $|d(n)+d(n-u)|$ in Eqs. (11) and (13), respectively, which are omitted here for simplicity.

Funding. National Key Research and Development Program of China (2018YFB1801701); National Natural Science Foundation of China (62101602, 62035018, 61875233, 62001415, 62101486); The Hong Kong Government General Research Fund (PolyU 15217620, PolyU 15220120); Project of the Shenzhen Municipal Science and Technology Innovation Commission (SGDX20201103095203030); Hong Kong Polytechnic University (G-SB1P); PolyU postdoc matching fund scheme of the Hong Kong Polytechnic University (1-W150); China Postdoctoral Science Foundation (2022M713559); Natural Science Foundation of Zhejiang Province (LQ21F050013).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. M. Chagnon, "Optical Communications for Short Reach," J. Lightwave Technol. **37**(8), 1779–1797 (2019).
2. K. Zhong, X. Zhou, J. Huo, C. Yu, C. Lu, and A. P. T. Lau, "Digital Signal Processing for Short-Reach Optical Communications: A Review of Current Technologies and Future Trends," J. Lightwave Technol. **36**(2), 377–400 (2018).
3. "IEEE Standard for Ethernet - Amendment 10: Media Access Control Parameters, Physical Layers, and Management Parameters for 200 Gb/s and 400 Gb/s Operation," IEEE Std 802.3bs (2017).
4. T. Wettlin, S. Calabro, T. Rahman, J. Wei, N. Stojanovic, and S. Pachnicke, "DSP for High-Speed Short-Reach IM/DD Systems Using PAM," J. Lightwave Technol. **38**(24), 6771–6778 (2020).
5. J. Zhang, X. Wu, L. Sun, J. Liu, A. P. T. Lau, C. Guo, S. Yu, and C. Lu, "C-band 120-Gb/s PAM-4 transmission over 50-km SSMF with improved weighted decision-feedback equalizer," Opt. Express **29**(25), 41622–41633 (2021).
6. J. Zhang, C. Guo, J. Liu, X. Wu, A. P. T. Lau, C. Lu, and S. Yu, "Low Complexity Frequency-Domain Nonlinear Equalization for 40-Gb/s/wavelength Long-Reach PON," in *OFC* 2018, W1J.8.
7. N.-P. Diamantopoulos, H. Nishi, W. Kobayashi, K. Takeda, T. Kakitsuka, and S. Matsuo, "On the Complexity Reduction of the Second-Order Volterra Nonlinear Equalizer for IM/DD Systems," J. Lightwave Technol. **37**(4), 1214–1224 (2019).
8. Y. Yu, T. Bo, Y. Che, D. Kim, and H. Kim, "Low-complexity nonlinear equalizer based on absolute operation for C-band IM/DD systems," Opt. Express **28**(13), 19617–19628 (2020).
9. J. Zhang, Z. Lin, X. Wu, J. Liu, A. P. Tao Lau, C. Guo, C. Lu, and S. Yu, "Low-complexity sparse absolute-term based nonlinear equalizer for C-band IM/DD systems," Opt. Express **29**(14), 21891–21901 (2021).
10. Q. Zhang, S. Duan, Z. Wang, B. Cao, Y. Wu, J. Chen, and M. Wang, "Low complexity Volterra nonlinear equalizer based on weight sharing for 50 Gb/s PAM4 IM/DD transmission with 10G-class optics," Opt. Commun. **508**, 127762 (2022).
11. Y. Xu, J. Zhang, B. Yu, S. Zhang, T. Zuo, L. Liu, J. Liu, C. Guo, A. P. T. Lau, and C. Lu, "260-Gb/s PAM-6 Transmission Using Joint Optical Pre-equalization and a Low-complexity Volterra Equalizer for Short-Reach Optical Interconnects," in *ACP* 2018, 1–3.
12. H. Xin, K. Zhang, L. Li, H. He, and W. Hu, "50 Gbps PAM-4 Over Up to 80-km Transmission With C-Band DML Enabled by Post-Equalizer," IEEE Photonics Technol. Lett. **32**(11), 643–646 (2020).
13. M. Zhu, J. Zhang, Q. Liu, B. Xu, X. Yi, Z. Yu, P. Zhang, and K. Qiu, "Experimental Demonstration of 80-Gb/s DSB OOK Signal Transmission Over 100-km SSMF with Simplified Volterra Based DFE," in *OECC* 2021, T2B.5.
14. H. Xin, K. Zhang, D. Kong, Q. Zhuge, Y. Fu, S. Jia, W. Hu, and H. Hu, "Nonlinear Tomlinson-Harashima precoding for direct-detected double sideband PAM-4 transmission without dispersion compensation," Opt. Express **27**(14), 19156–19167 (2019).

Optics EXPRESS

15. R. Rath, D. Clausen, S. Ohlendorf, S. Pachnicke, and W. Rosenkranz, "Tomlinson–Harashima Precoding For Dispersion Uncompensated PAM-4 Transmission With Direct-Detection," J. Lightwave Technol. **35**(18), 3909–3917 (2017).
16. J. Zhou, H. Wang, J. Wei, L. Liu, X. Huang, S. Gao, W. Liu, J. Li, C. Yu, and Z. Li, "Adaptive moment estimation for polynomial nonlinear equalizer in PAM8-based optical interconnects," Opt. Express **27**(22), 32210–32216 (2019).
17. X. Tang, S. Liu, Z. Sun, H. Cui, X. Xu, J. Qi, M. Guo, Y. Lu, and Y. Qiao, "C-band 56-Gb/s PAM4 transmission over 80-km SSMF with electrical equalization at receiver," Opt. Express **27**(18), 25708–25717 (2019).
18. X. Wu, J. Zhang, G. Zhou, A. P. T. Lau, and C. Lu, "C-Band 112-Gb/s PAM-4 Transmission over 50-km SSMF Using Absolute-Term Based Nonlinear FFE-DFE," in *ACP* 2021, M4I.3.
19. J. Zhang, Y. Zheng, X. Hong, and C. Guo, "Increase in Capacity of an IM/DD OFDM-PON Using Super-Nyquist Image-Induced Aliasing and Simplified Nonlinear Equalization," J. Lightwave Technol. **35**(19), 4105–4113 (2017).
20. W. Huang, W. Chang, C. Wei, J. Liu, Y. Chen, K. Chi, C. Wang, J. Shi, and J. Chen, "93% Complexity Reduction of Volterra Nonlinear Equalizer by ℓ1-Regularization for 112-Gbps PAM-4 850-nm VCSEL Optical Interconnect," in *OFC* 2018, M2D.7.
21. S.-Y. Lu, C.-C. Wei, C.-Y. Chuang, Y.-K. Chen, and J. Chen, "81.7% Complexity Reduction of Volterra Nonlinear Equalizer by Adopting L1 Regularization Penalty in an OFDM Long-Reach PON," in *ECOC* 2017, 1–3.
22. G. S. Yadav, C. Y. Chuang, K. M. Feng, J. H. Yan, J. Chen, and Y. K. Chen, "Reducing computation complexity by using elastic net regularization based pruned Volterra equalization in a 80 Gbps PAM-4 signal for inter-data center interconnects," Opt. Express **28**(26), 38539–38552 (2020).
23. J. Zhang, C. Guo, J. Liu, X. Wu, A. P. T. Lau, C. Lu, and S. Yu, "Orthogonal Matching Pursuit based Sparse Nonlinear Equalization for 40-Gb/s/wavelength Long-Reach PON," in *CLEO-PR* 2018, W2I.6.
24. J. Zhang, X. Wu, Q. Yan, H. Tan, X. Hong, C. Fei, A. P. T. Lau, and C. Lu, "Nonlinearity-aware PS-PAM-16 transmission for C-band net-300-Gbit/s/λ short-reach optical interconnects with a single DAC," Opt. Lett. **47**(12), 3035–3038 (2022).
25. X. Wu, J. Zhang, A. P. T. Lau, and C. Lu, "Low-complexity absolute-term based nonlinear equalizer with weight sharing for C-band 85-GBaud OOK transmission over a 100-km SSMF," Opt. Lett. **47**(6), 1565–1568 (2022).
26. S. Chen, "An area/time-efficient motion estimation micro core," IEEE Trans. Consumer Electron. **39**(3), 298–303 (1993).
27. A. Zaknich, *Principles of Adaptive Filters and Self-Learning Systems* Springer, Berlin, Germany, 2006.
28. G. Z. Karabulut and A. Yongacoglu, "Sparse Channel Estimation using Orthogonal Matching Pursuit Algorithm," in *Proc. Vehicular Technol. Conference* 2014, 3880–3884.
29. J. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," IEEE Trans. Inf. Theory **53**(12), 4655–4666 (2007).
30. J. Zhang, W. Chen, M. Gao, and G. Shen, "K-means-clustering-based fiber nonlinearity equalization techniques for 64-QAM coherent optical communication system," Opt. Express **25**(22), 27570–27580 (2017).