Low-complexity cluster-assisting look-up-table-based Volterra decision-feedback equalizer for IM/DD systems

Junwei Zhang,[1,*] ◉ **Heyun Tan,**[1] ◉ **Alan Pak Tao Lau,**[2] **Zhaohui Li,**[1,3,4] **and Chao Lu**[1,2]

[1]*School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510725, China*
[2]*Photonics Research Institute, Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China*
[3]*Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519000, China*
[4]*lzhh88@mail.sysu.edu.cn*
*zhangjw253@mail.sysu.edu.cn

To address the challenges posed by chromatic dispersion (CD)-induced power fading and nonlinear signal distortions in double-sideband (DSB) intensity modulation and direct detection (IM/DD) transmission systems, a combination of a Volterra feed-forward equalizer (VFFE) and a Volterra decision-feedback equalizer (VDFE) is widely employed. However, the conventional VFFE–VDFE exhibits significant computational complexity, particularly for longer memory lengths. In this Letter, a low-complexity cluster-assisting look-up table-based VDFE (CLUT-VDFE) is proposed to effectively reduce the computational complexity associated with compensating for CD and nonlinear distortions. By utilizing CLUTs, all multiplication operations required for the implementation of the VDFE are completely eliminated. To validate the effectiveness of the proposed CLUT-VDFE, experiments on a C-band 100-Gb/s PAM-4 transmission system over a 60-km standard single-mode fiber (SSMF) are conducted. The experimental results show that the CLUT-VDFE not only achieves comparable equalization performance to the conventional VDFE but also effectively eliminates multiplication operations and significantly saves 48.8% real-valued additions. © 2025 Optica Publishing Group.

Propelled by the proliferation of bandwidth-consuming services such as cloud computing, data storage, and Internet of things (IoT), the rapid growth of network traffic necessitates higher throughput demands for optical data-center interconnects and access networks, covering distances of up to 100 km. Compared to coherent detection schemes, intensity modulation and direct detection (IM/DD) systems are the preferred choice for most optical data-center interconnections and access network applications due to their advantages of low cost, low power consumption, and small footprint [1,2]. For optical interconnects spanning distances under 40 km, IM/DD systems operating in the O band are preferred due to their minimal chromatic dispersion (CD) [3]. When transmission distances exceed 40 km, C-band transmissions demonstrate superior performance in many aspects such as fiber loss, the maturity of wavelength division multiplexing (WDM) optics, and optical amplifiers.

However, high-speed IM/DD signals suffer from distortions caused by CD and nonlinearities associated with modulation and detection [4,5]. The interaction between CD and square-law detection leads to power fading, which degrades system performance. One of the most widely adopted strategies for compensating for CD-induced power fading and nonlinear distortions in IM/DD systems is to employ digital signal processing (DSP)-based equalization techniques, as they offer flexibility and cost-effectiveness. A combination of a feed-forward equalizer (FFE) and a decision feedback equalizer (DFE) is an effective and low-complexity solution for post equalization, in which the DFE can efficiently mitigate the CD-induced power fading [6]. To enhance equalization performance by compensating for both CD-induced power fading and nonlinear distortions, joint Volterra FFE (VFFE) and Volterra DFE (VDFE) schemes have been extensively investigated in IM/DD systems [7–9], providing a superior balance between performance and computational complexity when compared to deep learning nonlinear equalizers such as neural network (NN) and long short-term memory (LSTM) nonlinear equalizers. However, the equalization performance of receiver-side linear and nonlinear DFEs is limited by the error propagation. To address this issue, the DFE and VDFE can be relocated to the transmitter side named Tomlinson–Harashima pre-coding (THP) [6] and nonlinear THP (NTHP) [10], respectively. Alternatively, the performance of DFE and VDFE can be enhanced by implementing improved weighted DFE (IWDFE) [11] and improved weighted VDFE (IWVDFE) [12], which replace hard-decision outputs with soft-decision ones based on a compressed sigmoid nonlinear function.

Another challenge to implement the VDFE is the high computational complexity, which grows exponentially as the memory length increases. Greedy algorithm-based sparse pruning and *k*-means clustering-based weight sharing schemes are

two popular strategies to reduce the equalization complexity of VDFE [7]. In addition to the two strategies mentioned above, a recently proposed cluster-assisting look-up-table-based DFE (CLUT-DFE) offers a multiplication-free decision-feedback equalization scheme with performance comparable to traditional linear DFE [13]. Besides, the CLUT technique has also been introduced on the transmitter side to reduce the computational complexity of a finite impulse response (FIR) filter-based electrical dispersion pre-compensation scheme for IM/DD systems [14]. Nevertheless, there is currently no investigation into implementing CLUT in nonlinear DFE/precoding, which is significant for low-cost and high-capacity double-sideband (DSB) IM/DD systems.

In this Letter, we propose a low-complexity CLUT-VDFE to effectively compensate for both CD and nonlinear distortions in IM/DD systems. The proposed CLUT-VDFE incorporates two key strategies to reduce complexity: (1) utilizing the k-means clustering algorithm to cluster the linear and second-order nonlinear kernels of VDFE separately, thereby reducing the kernel redundancy; and (2) establishing a LUT for each cluster center (i.e., new kernel), resulting in a reduced total table size. By implementing the CLUT followed by a summation operation, the total table size can be significantly reduced and all multiplication operations required for the VDFE are completely eliminated. Based on post-equalization with a VFFE and the proposed CLUT-VDFE, C-band 100-Gb/s four-level pulse amplitude modulation (PAM-4) transmission over a 60-km standard single-mode fiber (SSMF) is realized with the achieved bit error ratio (BER) below 7% hard-decision forward error correction (HD-FEC) limit of 3.8×10^{-3}. The proposed CLUT-VDFE achieves comparable equalization performance to conventional VDFE, while eliminating all multiplication operations and significantly reducing the number of real-valued additions by 48.8%.

To effectively mitigate the transmission impairments in IM/DD systems while maintaining relatively low complexity, a combination of a second-order VFFE and a second-order VDFE, both with diagonal pruning, can be implemented at the receiver side [7,8]. In this configuration, the VFFE can equalize the precursor linear and nonlinear impairments while the VDFE can compensate for the CD-induced power fading as well as a portion of nonlinear distortions. The nth sample of the output from the VFFE–VDFE can be expressed as follows:

$$y(n) = y_{\text{VFFE}}(n) + y_{\text{VDFE}}(n) =$$

$$\underbrace{\sum_{k=-L_1}^{L_1} h_1(k)x(2n-k) + \sum_{q=0}^{Q-1}\sum_{k=-L_2}^{L_2-q} h_2(k,q)x(2n-k)x(2n-k-q)}_{\text{VFFE}}$$

$$+ \underbrace{\sum_{k=1}^{D_1} \omega_1(k)d(n-k) + \sum_{u=0}^{U-1}\sum_{k=1}^{D_2-u} \omega_2(k,u)d(n-k)d(n-k-u)}_{\text{VDFE}},$$

$$\tag{1}$$

where $x(2n)$ with $T/2$ symbol space denotes the received DD signal, $d(n)$ with T symbol space represents the previous hard-decision output. h_K ($K = 1, 2$) and $N_K = 2L_K + 1$ are the Kth-order kernel and memory length of the VFFE, respectively. ω_K and D_K are the Kth-order kernel and memory length of the VDFE, respectively. Q and U represent pruning factors, which are used

to prune unimportant kernels and their cross-beating terms having large relative delays. Consequently, the total number of kernels for VFFE and VDFE are given by $N_1 + Q(2N_2-Q+1)/2$ and $D_1 + U(2D_2-U+1)/2$, respectively, while their required numbers of real-valued multiplications per symbol (RNRMs) are $N_1 + Q(2N_2-Q+1)$ and $D_1 + U(2D_2-U+1)$, respectively. In Eq. (1), these kernels can be estimated using a recursive least squares (RLS) algorithm [7] based on a prior-to-signal training sequence during the training phase.

The focus of this work is to eliminate all multiplication operations in VDFE by implementing CLUT, thereby maintaining the VFFE unchanged during equalization. The computational complexity of the proposed CLUT-DFE can be significantly reduced through the implementation of CLUT, in which the Kth-order kernels with similar values are grouped together to form a new Kth-order kernel for equalization. A k-means clustering algorithm [7] is employed to cluster the estimated D_1 linear kernels and $U(2D_2-U+1)/2$ nonlinear kernels into C_1 and C_2 cluster centers (i.e., new kernels), respectively. Therefore, the output of the proposed CLUT-VDFE with two large CLUTs can be expressed as follows:

$$y_{\text{CLUT-VDFE}}(n) = \underbrace{\sum_{i=0}^{C_1-1} g_1(i)d_1(n,i)}_{\text{L-CLUT}} + \underbrace{\sum_{i=0}^{C_2-1} g_2(i)d_2(n,i)}_{\text{NL-CLUT}}, \quad \tag{2}$$

where $g_K(i)$ ($K = 1, 2$) and C_K denote the cluster centers and the number of clusters obtained by k-means clustering algorithm for Kth-order kernels, respectively. $d_K(n,i)$ represents the sum of the corresponding Kth-order decision-feedback terms in VDFE, whose kernels belong to the ith cluster with a cluster center of $g_K(i)$. The first and second terms in Eq. (2) represent the outputs of a linear CLUT (L-CLUT) and a nonlinear CLUT (NL-CLUT), respectively. Compared with L-CLUT, a larger table size is often required for NL-CLUT due to the significantly more nonlinear kernels in VDFE. To reduce the overall table size, a sub-LUT is established for each cluster center in the proposed CLUT-VDFE through exhaustive traversal, which differs from the large CLUT presented in [13]. The mth output elements of the ith sub-LUTs in L-CLUT and NL-CLUT are respectively denoted as follows:

$$y_{\text{L-CLUT}}^i(m) = g_1(i) \sum_{j=1}^{N_L(i)} c_1(m,j) \text{ and} \tag{3}$$

$$y_{\text{NL-CLUT}}^i(m) = g_2(i) \sum_{j=1}^{N_{NL}(i)} c_2(m,j)c_2(m,j+N_{NL}(i)), \tag{4}$$

where $N_L(i)$ and $N_{NL}(i)$ represent the number of elements in the ith clusters of linear and nonlinear kernels, respectively. $c_1(m, j)$ and $c_2(m, j)$ denote the jth element of the mth pattern in ith sub-LUTs of the L-CLUT and NL-CLUT, respectively. The ith sub-LUTs of the L-CLUT and NL-CLUT contain $M^{N_L(i)}$ and $M^{2N_{NL}(i)}$ patterns, respectively, resulting in a total table size of $\sum_{i=0}^{C_1-1} M^{N_L(i)} + \sum_{i=0}^{C_2-1} M^{2N_{NL}(i)}$ for the PAM-M signal. For example, for a cluster center $g_1(i)$ with $N_L(i) = 3$, there are 4^3 patterns in the ith sub-LUT of the L-CLUT for the PAM-4 signal, as shown in Fig. 1(a), where $c_1(2,3)$ is -1. On the other hand, for a cluster center $g_2(i)$ with $N_{NL}(i) = 3$, the ith sub-LUT of the NL-CLUT for PAM-4 signal contains 4^6 patterns, ranging from $\{-3, -3, -3, -3, -3, -3\}$ to $\{3, 3, 3, 3, 3, 3\}$. To further reduce the total table
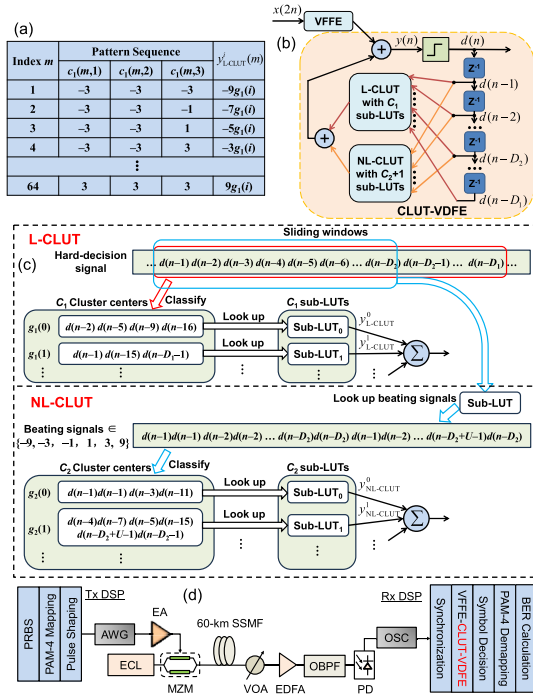
Fig. 1. (a) Example of the ith sub-LUT of L-CLUT with $N_L(i) = 3$. (b) and (c) Schematic diagram of the CLUT-VDFE for the PAM-4 signal. (d) Experimental setup and DSP block diagram.

size of PAM-4 based CLUT-VDFE, we also establish a sub-LUT with 4^2 patterns to look up the beating signal $c_2(m,j)c_2(m,j + N_{NL}(i))$, whose output is $c_3(m,j) \in \{-9, -3, -1, 1, 3, 9\}$. Thus Eq. (4) is rewritten as follows:

$$y^i_{\text{NL - CLUT}}(m) = g_2(i) \sum_{j=1}^{N_{NL}(i)} c_3(m,j). \tag{5}$$

In this situation, the total number of sub-LUTs changes from $C_1 + C_2$ to $C_1 + C_2 + 1$, resulting in a reduced overall table size of $\sum_{i=0}^{C_1-1} 4^{N_L(i)} + \sum_{i=0}^{C_2-1} 6^{N_{NL}(i)} + 16$ for the PAM-4 signal.

The schematic diagram of the proposed CLUT-VDFE for the PAM-4 signal is depicted in Figs. 1(b) and 1(c). The previous hard-decision output symbols with memory lengths of D_1 and D_2 are served as inputs to L-CLUT and NL-CLUT, respectively. In L-CLUT and NL-CLUT, the hard-decision symbols $d(n-k)$ and their beating signals $d(n-k)d(n-k-u)$ undergo classification based on the clustering results of their linear and nonlinear kernels, respectively. Subsequently, distinct signal groups are employed to search for the corresponding sub-LUTs. Finally, the output of the proposed CLUT-VDFE is obtained by summing the retrieved results from C_1 sub-LUTs of L-CLUT and from C_2 sub-LUTs of NL-CLUT, based on Eqs. (3) and (5).

The computational complexity of the proposed CLUT-VDFE is compared to that of VDFE in terms of the RNRM and required a number of real-valued additions per symbol (RNRA). While conventional VDFE necessitates $D_1 + U(2D_2 - U + 1)/2 - 1$ real-valued additions and $D_1 + U(2D_2 - U + 1)$ real-valued multiplications, respectively, the RNRM is completely eliminated in the proposed CLUT-VDFE, and its RNRA is significantly reduced to $C_1 + C_2 - 1$, depending on the total number of cluster centers $C_1 + C_2$.

The performance of the proposed CLUT-VDFE is evaluated in a C-band 100-Gb/s PAM-4 transmission system over a 60-km SSMF. Figure 1(d) illustrates the experimental setup and the DSP utilized in the transmission experiment. At the transmitter, a pseudo random bit sequence (PRBS) is initially generated and mapped to the PAM-4 symbols. Subsequently, the PAM-4 symbols are up-sampled to achieve four samples per symbol for simple rectangular pulse shaping. The resulting PAM-4 signal is then loaded into an arbitrary waveform generator (AWG, Keysight 8199A) operating at a sample rate of 200 GSa/s. The 50-GBaud PAM-4 electrical signal generated 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). The MZM is biased at a quadrature point and injected with an optical carrier generated from an external cavity laser (ECL) with a center wavelength of 1550.12 nm. After being transmitted over a 60-km SSMF without any dispersion compensation, the received optical power (ROP) of the received signal is swept by inserting a variable optical attenuator (VOA), which is then detected by a photodetector (PD, XPDV2120RA). The power into the PD is boosted up to 7 dBm by utilizing an erbium-doped fiber amplifier (EDFA) in conjunction with an optical band-pass filter (OBPF). Finally, the detected electrical PAM-4 signal is captured and stored by a real-time oscilloscope (OSC) operating at 160 GSa/s for subsequent offline DSP procedures, including resampling to two samples per symbol, synchronization, equalization using the proposed VFFE-CLUT-VDFE, symbol decision, PAM-4 demapping, and BER calculation. In the experiment, 10,000 training symbols are initially transmitted to obtain the kernel coefficients based on the RLS algorithm. Subsequently, these kernels are clustered using the k-means clustering algorithm for CLUT-VDFE equalization.

We first optimize the parameters, including memory lengths, pruning factors, and the number of cluster centers of VFFE-VDFE/CLUT-VDFE in a 100-Gb/s PAM-4 transmission system over a 60-km SSMF at a ROP of -10 dBm. It is found that the BER can be minimized to an optimal value for FFE–DFE when the memory lengths of the FFE and DFE reach 109 and 28, respectively. Based on the optimized linear memory lengths $N_1 = 109$ and $D_1 = 28$, the measured BER as a function of the nonlinear memory length N_2 for a VFFE–DFE with $Q = N_2$ is shown in Fig. 2(a). The performance is improved by increasing the nonlinear memory length N_2 and the BER reaches saturation at $N_2 = 45$. Figure 2(b) illustrates the measured BER versus the pruning factor Q of the VFFE–DFE with $N_2 = 45$. It can be observed that a larger pruning factor Q leads to a significant reduction in BER until it reaches a plateau at $Q = 13$. Similarly, it can be observed from Figs. 2(c) and 2(d) that the optimal BERs are achieved when selecting a nonlinear memory length D_2 of 20 and a pruning factor U of 13 for the VFFE–VDFE. Furthermore, compared to the VFFE–DFE featuring a linear DFE structure, the VFFE–VDFE employing a nonlinear DFE structure achieves a significant reduction in BER.

With $N_1 = 109$, $N_2 = 45$, $Q = 13$, $D_1 = 28$, $D_2 = 20$, and $U = 13$ fixed, we keep the VFFE unchanged during equalization and separately optimize the numbers of cluster centers for L-CLUT and NL-CLUT in the proposed CLUT-VDFE. Figure 3(a) illustrates the measured BER and total table size versus the number of cluster centers for L-CLUT, in which only linear DFE is replaced with L-CLUT while keeping the nonlinear part of VDFE unchanged. Similarly, Fig. 3(b) presents the measured BER and total table size against the number of cluster centers for
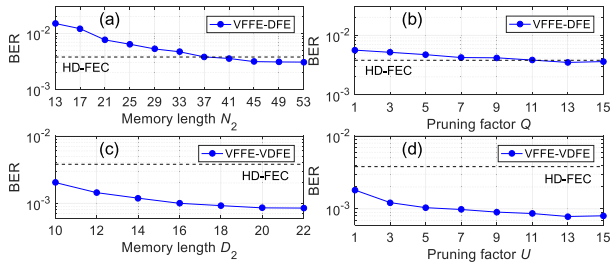
Fig. 2. (a) BER versus memory length N_2 for a VFFE-DFE with $Q = N_2$. (b) BER versus pruning factor Q for a VFFE-DFE with $N_2 = 45$. (c) BER versus memory length D_2 for a VFFE–VDFE with $N_2 = 45$, $Q = 13$, and $U = D_2$. (d) BER versus pruning factor U for a VFFE–VDFE with $N_2 = 45$, $Q = 13$, and $D_2 = 20$.
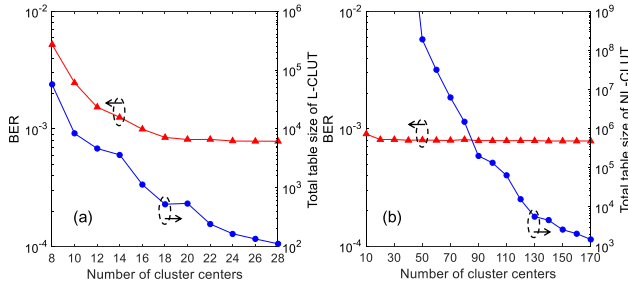

Fig. 3. BER and total table size versus the number of cluster centers for (a) L-CLUT and (b) NL-CLUT.

Fig. 4. (a) BER versus ROP using different equalizers after 60-km SSMF transmission. (b)–(g) Recovered eye diagrams of PAM-4 signals with different equalizers at a ROP of -10 dBm.

NL-CLUT. The BERs of both L-CLUT and NL-CLUT exhibit an initial reduction as the number of cluster centers increases before eventually reaching a saturated performance level. Simultaneously, both cases show a significant decrease in total table sizes with an increasing number of cluster centers. When the numbers of cluster centers reach 18 and 20 for L-CLUT and NL-CLUT, respectively, only a marginal deterioration in BER performance is observed. It should be noted that a larger number of cluster centers leads to a decrease in total table size while simultaneously increasing RNRA. Therefore, the numbers of cluster centers C_1 and C_2 are set to 18 and 90, respectively, for L-CLUT and NL-CLUT, ensuring that the total table size remains below 10^6. As a result, the RNRM (RNRA) of CLUT-VDFE is 0 (107), which is reduced by 100% (48.8%) compared with 392 (209) for the conventional VDFE. Considering the total computational complexity of equalization, the RNRM of VFFE-CLUT-VDFE is reduced by 25.9% compared to that of VFFE-VDFE, which can be further reduced by implementing sparse pruning or weight sharing strategy [7] for VFFE.

Finally, we assess the performance of the proposed CLUT-VDFE under different ROPs. Figure 4 shows the BER versus the ROP using different equalizers, along with their corresponding recovered eye diagrams for 100-Gb/s PAM-4 signals after 60-km SSMF transmission. In the absence of nonlinear equalization or DFE, both FFE–DFE and VFFE fail to meet the 7% HD-FEC BER limit of 3.8×10^{-3}. Regarding DFE-based nonlinear equalization, the proposed VFFE-CLUT-VDFE achieves comparable performance to VFFE–VDFE and improves receiver sensitivity by more than 3 dB over VFFE–DFE. Compared with conventional VDFE, the computational complexity of the proposed CLUT-VDFE is significantly reduced, resulting in a 48.8% reduction in RNRA while eliminating RNRM. Note that
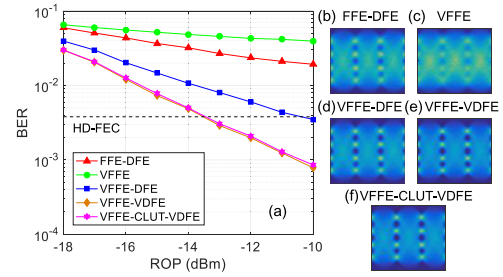
relocating the CLUT-VDFE at the transmitter side as CLUT-NTHP or combining it with a soft-decision operation at the receiver as CLUT-IWVDFE can further enhance the equalization performance of the proposed multiplication-free nonlinear equalization approach.

In conclusion, we have proposed and demonstrated a low-complexity CLUT-VDFE to compensate for CD and nonlinear distortions in IM/DD systems. By utilizing CLUTs, all multiplication operations required for the implementation of the VDFE are completely eliminated. The experimental results show that the CLUT-VDFE not only achieves comparable equalization performance to the conventional VDFE but also saves the RNRM and RNRA by 100% and 48.8%, respectively, in a 100-Gb/s PAM-4 60-km transmission system. Therefore, the proposed WS-TDFE offers a promising solution for compensating CD and nonlinear distortions in cost-effective IM/DD systems.

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, J. Lightwave Technol. **37**, 1779 (2019).
2. X. Pang, O. Ozolins, R. Lin, et al., J. Lightwave Technol. **38**, 492 (2020).
3. C. St-Arnault, R. Bernalet, E. Berikaa, et al., J. Lightwave Technol. (2024).
4. X. Wu, J. Zhang, A. P. T. Lau, et al., Opt. Lett. **47**, 5144 (2022).
5. X. Tang, Y. Qiao, Y.-W. Chen, et al., J. Lightwave Technol. **38**, 4683 (2020).
6. R. Rath, D. Clausen, S. Ohlendorf, et al., J. Lightwave Technol. **35**, 3909 (2017).
7. J. Zhang, H. Tan, X. Hong, et al., Opt. Express **30**, 36343 (2022).
8. H. Xin, K. Zhang, L. Li, et al., IEEE Photonics Technol. Lett. **32**, 643 (2020).
9. L. Huang, Y. Xu, W. Jiang, et al., J. Lightwave Technol. **40**, 4528 (2022).
10. H. Xin, K. Zhang, D. Kong, et al., Opt. Express **27**, 19156 (2019).
11. J. Zhang, X. Wu, L. Sun, et al., Opt. Express **29**, 41622 (2021).
12. J. Zhang, H. Tan, X. Hong, et al., IEEE Photonics Technol. Lett. **35**, 163 (2023).
13. F. Xie, X. Huang, S. Liu, et al., J. Lightwave Technol. **42**, 3118 (2024).
14. W. Ni, D. Zou, Y. Chen, et al., Opt. Lett. **49**, 6417 (2024).