




Recursive Least Squares Based Low-Complexity Frequency-Domain MIMO Equalization for MDL-Tolerant Long-Haul Space Division Multiplexing Transmission

Kohki Shibahara , *Member, IEEE*, Megumi Hoshi, Takemi Hasegawa , *Member, IEEE*, Tetsuya Hayashi , *Senior Member, IEEE*, and Yutaka Miyamoto, *Member, IEEE*

Abstract—We propose a novel low-complexity, fast-converging frequency-domain multiple-input multiple-output (MIMO) equalization technique suitable for long-haul space division multiplexed (SDM) transmission. Based on the recursive least squares (RLS) method, the proposed scheme, referred to as out-of-band exclusive frequency domain equalization (OBE-FDE), provides a robust learning of MIMO filter coefficients even in the presence of accumulated mode dependent loss (MDL), while achieving reduced complexity by selectively performing computations only within the signal's bandwidth. Numerical evaluation results show that complexity reduction effect is enhanced with increased number of spatial modes in a specific roll-off factor range. Also demonstrated is an improved convergence tolerance against MDL introduced by the RLS adaptation, contrasting with the widely-used least mean squares (LMS) adaptation in scenarios involving long-haul MDL accumulation and/or high-mode-count mode division multiplexing (MDM) transmissions. We also show experimental demonstrations, proving that the proposed low-complexity RLS-adapted MIMO-FDE achieved a ten-fold faster learning of filter coefficients in 4-coupled core multi-core fiber transmission at 5528 km. Additionally, by newly applying the scheme into 10-spatial-mode transmission, a record-long signal transmission distance over 1560 km was achieved while reducing computational complexity requirements by 40%.

Index Terms—Space division multiplexing (SDM), Mode division multiplexing (MDM), Mode dependent loss (MDL), Recursive least squares (RLS).

I. INTRODUCTION

EXPANDING per-link capacity over optical fiber is eagerly anticipated for realizing future optical transport systems.

Manuscript received 15 December 2023; revised 7 March 2024 and 1 April 2024; accepted 2 April 2024. Date of publication 4 April 2024; date of current version 27 June 2024. This work was supported by the National Institute of Information and Communications Technology (NICT), Japan, under the commissioned research under Grant JPJ012368C01001. (*Corresponding author: Kohki Shibahara.*)

Kohki Shibahara, Megumi Hoshi, and Yutaka Miyamoto are with NTT Network Innovation Laboratories, NTT Corporation, Kanagawa 100-8116, Japan (e-mail: kouki.shibahara@ntt.com; megumi.hoshi@ntt.com; yutaka.miyamoto@ntt.com).

Takemi Hasegawa and Tetsuya Hayashi are with Optical Communications Laboratory, Sumitomo Electric Industries, Ltd., Sakae-ku 244-8588, Japan (e-mail: hase@sei.co.jp; t-hayashi@sei.co.jp).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JLT.2024.3385094>.

Digital Object Identifier 10.1109/JLT.2024.3385094

One prime candidate for it is space division multiplexing (SDM) technology allowing spatial channels to be parallelized within single fiber by means of core multiplexing and/or mode multiplexing. Known as a family of mode division multiplexing (MDM) approach, the latter utilizes transmission medium including coupled-core multicore fibers (CC-MCFs) or multi-mode fibers (MMFs), providing high throughput capacity exceeding 1 Pbps [1], [2] and/or long transmission reach over 1000 km [3], [4], [5], [6], while keeping fiber cladding diameter within 125 μm . This is greatly attributed to the help of multiple-input multiple-output equalization technique for undoing intra- and inter mode mixing that occurs during propagation over MDM fiber. A practical issue in applying MIMO equalization for MDM is a slow convergence of filter coefficients affected by the presence of mode dependent loss (MDL) [7]. This is particularly prominent in an use of MIMO equalization with the steepest descent algorithm including the least mean squares (LMS) method, because of a learning rate sensitivity with respect to eigenvalue spread of a correlation matrix of input signals into MIMO equalization [8]. Furthermore, the convergence slowing is enhanced in high-mode-count MDM transmission over long distances due to accumulated MDL. A candidate solution to avoid this effect is an introduction of recursive least squares (RLS) based MIMO equalization filter adaptation [7]. A convergence speed improvement by the frequency-domain RLS has been experimentally demonstrated using three spatial modes over up to 1000 km [9], [10]. One remaining challenge for RLS-based filter adaptation is mitigating the rather enhanced computational complexity.

We have previously proposed a simplified scheme for LMS-adapted MIMO frequency-domain equalization (FDE) that avoids redundant computations outside of signal's bandwidth [11], which was referred to as out-of-band exclusive (OBE) processing. In [12], we further developed OBE's application into RLS-adapted one with the aim to simultaneously attaining fast convergence and low complexity even in the presence of MDL. With extensions of the work presented in [12], the aim of this paper is to provide comprehensive and detailed descriptions of the underlying FDE techniques along with the analysis of the required computational burdens in various MDM transmission scenarios with different signal/system

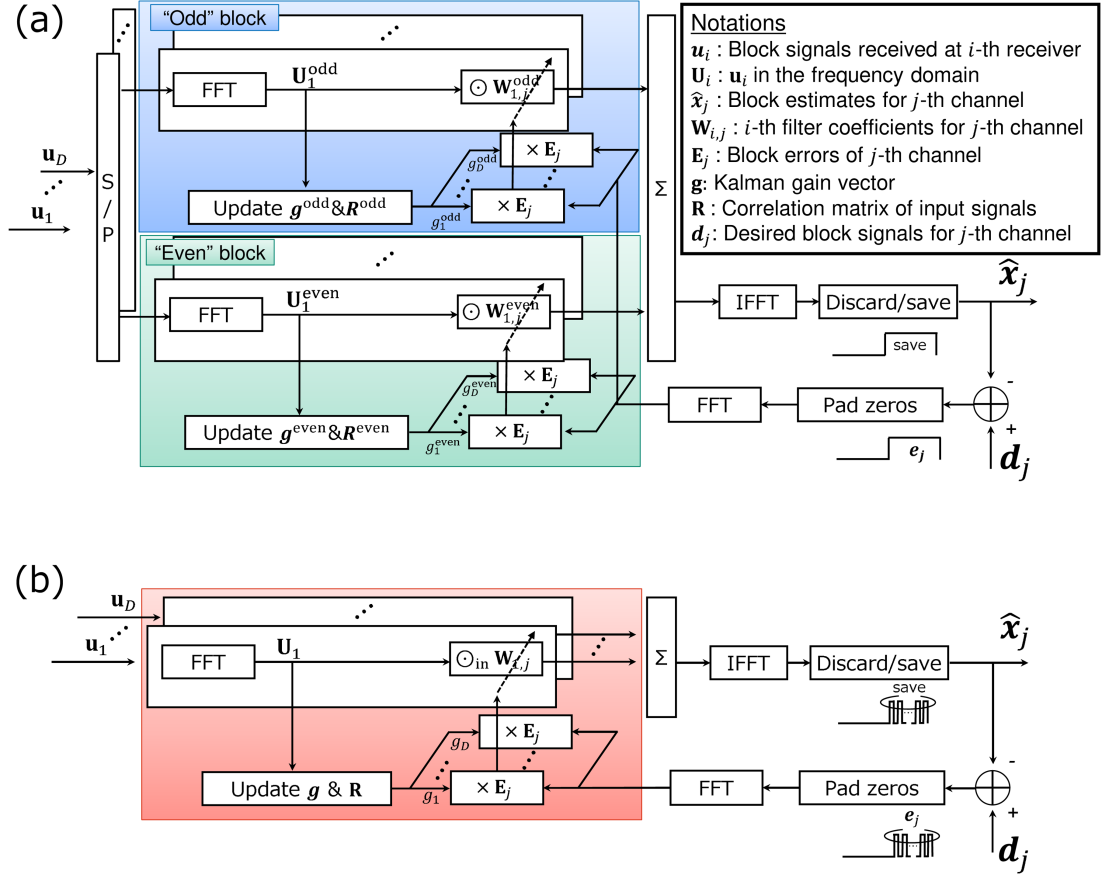


Fig. 1. Configuration comparison of RLS-adapted FDE schemes. (a) Conventional FDE (C-FDE). (b) Out-of-band exclusive FDE (OBE-FDE).

parameters. We revealed that complexity mitigation by the proposed FDE technique is consistently achieved with practical MDM transmission parameters, especially further reduced in higher mode-count scenarios in a specific range of roll-off factor of signals. Advantages of the RLS adaptation with respect to the LMS are also numerically confirmed in high-mode-count MDM transmission even impaired with MDL. Based on experimental results over 4-core CC-MCF, we also show that convergence rate is improved by an order of magnitude at 5500 km. Also demonstrated is a record-long-distance 10-mode WDM/MDM transmission over 1560 km with a decreased complexity by 40%, corresponding to 20% reach extension relative to our previous work [13].

The remaining sections of this paper are structured as follows: Section II provides algorithm descriptions of the conventional and proposed OBE-based FDE adapted by the RLS, and analyses of the required computational complexity for both schemes, followed by performance evaluations through numerical simulation to clarify of an advantage of the proposed scheme in various MDM transmission scenarios. In Section III, we describe an experimental setup for long-haul MDM transmission supporting a spatial channel count up to 10. Section IV presents experimental results over 4-core CC-MCF and 10-mode FMF that verifies an improvement both in convergence speed and MDM signal's transmission performance achieved by the introduction of RLS-adapted MIMO-FDE, specifically showing a

record-long 10-spatial-mode transmission. Finally, in Section V, we conclude the paper with a brief summary of our work.

II. PROPOSED: OUT-OF-BAND EXCLUSIVE (OBE) FDE WITH RLS ADAPTATION

This section briefly reviews the conventional FDE (C-FDE), followed by a description of the proposed low-complexity FDE technique. Throughout this section, we focus on an RLS-based filter adaptation instead of widely used LMS one. We also assume an employment of the overlap-save MIMO-FDE algorithm with a block size of N working at an oversampling rate of 2 and an overlapping ratio of 0.5. Accordingly, each block processing produces $N/4$ output symbols per block processing per spatial channel. In what follows, we consider an MDM system where D spatial channels including polarization ones are transmitted over SDM fibers. An objective of MIMO equalization is to extract a block estimate of a data stream for a j -th channel \hat{x}_j using a convolution of received signal block at i -th receiver u_j with an i -th filter weight vector for a j -th spatial channel $W_{i,j} \in \mathbb{C}^{N \times 1}$ ($i, j \in \{1, 2, \dots, D\}$).

A. Conventional FDE With RLS Adaptation

Fig. 1(a) shows a conventional FDE (C-FDE) [9], [10] for a j -th data stream output. The processing parts requiring complex multiplications/divisions include fast Fourier transform (FFT),

output calculation, and updates for a filter coefficient $\mathbf{W}_{i,j}$, a Kalman gain vector \mathbf{g} , and a correlation matrix \mathbf{R} . At the k -th block time, the i -th input blocks $\mathbf{u}_i(k) \in \mathbb{C}^{N \times 1}$ is configured as

$$\mathbf{u}_i(k) \triangleq \begin{bmatrix} u_i \left(1 + \frac{N}{2}(k-1)\right) \\ u_i \left(2 + \frac{N}{2}(k-1)\right) \\ \vdots \\ u_i \left(N + \frac{N}{2}(k-1)\right) \end{bmatrix}, \quad (1)$$

where superscript T denotes the transpose operation, and u_i is an i -th input signal sampled at the oversampling rate of 2. $\mathbf{u}_i(k)$ is then parallelized into odd and even blocks, denoted as $\mathbf{u}_i^o(k) \in \mathbb{C}^{N/2 \times 1}$, and $\mathbf{u}_i^e(k) \in \mathbb{C}^{N/2 \times 1}$, respectively, as

$$\mathbf{u}_i^o(k) \triangleq \begin{bmatrix} u_i \left(1 + \frac{N}{2}(k-1)\right) \\ u_i \left(3 + \frac{N}{2}(k-1)\right) \\ \vdots \\ u_i \left(N-1 + \frac{N}{2}(k-1)\right) \end{bmatrix}, \quad (2)$$

and

$$\mathbf{u}_i^e(k) \triangleq \begin{bmatrix} u_i \left(2 + \frac{N}{2}(k-1)\right) \\ u_i \left(4 + \frac{N}{2}(k-1)\right) \\ \vdots \\ u_i \left(N + \frac{N}{2}(k-1)\right) \end{bmatrix}, \quad (3)$$

followed by FFT with a size of $N/2$ to obtain frequency-domain input signals $\mathbf{U}_i^o(k) \in \mathbb{C}^{N/2 \times 1}$ and $\mathbf{U}_i^e(k) \in \mathbb{C}^{N/2 \times 1}$:

$$\mathbf{U}_i^{o,e}(k) = \text{FFT}[\mathbf{u}_i^{o,e}(k)], \quad (4)$$

where $\text{FFT}[\cdot]$ denotes the FFT operation. Note that, for the conventional FDE processing, superscript o and e denote ‘‘odd’’ and ‘‘even’’ quantities, respectively. To obtain the output symbols for the j -th spatial channel $\hat{\mathbf{x}}_j \in \mathbb{C}^{N/4 \times 1}$, the sum of the product of these input blocks with relevant filter weight vectors, $\mathbf{W}_{i,j}^{o,e} \in \mathbb{C}^{N/2 \times 1}$, are again converted into the time-domain signals by inverse FFT (IFFT), $\text{IFFT}[\cdot]$, the last half of which is saved:

$\hat{\mathbf{x}}_j =$ the last half of

$$\text{IFFT} \left[\sum_{i=1}^D \mathbf{W}_{i,j}^o \odot \mathbf{U}_i^o(k) + \sum_{i=1}^D \mathbf{W}_{i,j}^e \odot \mathbf{U}_i^e(k) \right], \quad (5)$$

where \odot denotes the element-wise product (or Hadamard product). The time-domain error vector $\hat{\mathbf{x}}_j(k)$ for the j -th spatial channel is obtained as the difference between $\hat{\mathbf{x}}_j$ and the desired symbol vector $\mathbf{d}_j(k) \in \mathbb{C}^{N/4 \times 1}$:

$$\mathbf{e}_j(k) \triangleq \mathbf{d}_j(k) - \hat{\mathbf{x}}_j(k). \quad (6)$$

With $\mathbf{e}_j(k)$, a frequency-domain error vector $\mathbf{E}_j(k) \in \mathbb{C}^{N/2 \times 1}$ is constructed as

$$\mathbf{E}_j(k) \triangleq \text{FFT} \begin{bmatrix} \mathbf{0}_{N/4} \\ \mathbf{e}_j(k) \end{bmatrix}, \quad (7)$$

where $\mathbf{0}_p$ is a vector comprising p zeros. Letting m denoted as an index identifying a frequency bin, input signal vector at each frequency is constructed as $\mathbf{V}^{o,e}[m] = [U_1^{o,e}[m],$

$U_2^{o,e}[m], \dots, U_D^{o,e}[m]]^T \in \mathbb{C}^{D \times 1}$. Subsequent processing at each block time follows with updates for relevant parameters including $\mathbf{g}^{o,e}[m] \triangleq [g_1^{o,e}[m], g_2^{o,e}[m], \dots, g_D^{o,e}[m]]^T \in \mathbb{C}^{D \times 1}$ and $\mathbf{R}^{o,e}[m] \in \mathbb{C}^{D \times D}$:

$$\mathbf{g}^{o,e}[m] \leftarrow \frac{\mathbf{R}^{o,e}[m] \mathbf{V}^{o,e}[m]}{\lambda + (\mathbf{V}^{o,e})^H [m] \mathbf{R}^{o,e}[m] \mathbf{V}^{o,e}[m]}, \quad (8)$$

$$\mathbf{R}^{o,e}[m] \leftarrow \lambda^{-1} \mathbf{R}^{o,e}[m] - \lambda^{-1} \mathbf{g}^{o,e}[m] (\mathbf{V}^{o,e})^H [m] \mathbf{R}^{o,e}[m], \quad (9)$$

where superscripts H denotes a Hermitian conjugate operation, and λ is a forgetting factor satisfying $0 \ll \lambda < 1$. Note that processing of (8) through (10) is performed for each frequency bin m , and also that a correlation matrix $\mathbf{R}^{o,e}$ is initialized with a $D \times D$ identity matrix $\mathbf{I}_{D \times D}$ multiplied by a positive number β . For simplicity in computational notation, we redefine a Kalman gain vector as $\mathbf{h}_i^{o,e} \triangleq [g_i^{o,e}[1], g_i^{o,e}[2], \dots, g_i^{o,e}[N/2]]^T \in \mathbb{C}^{N/2 \times 1}$. Subsequently, a filter weight vector $\mathbf{W}_{i,j}^{o,e}$ is simply updated by

$$\mathbf{W}_{i,j}^{o,e} \leftarrow \mathbf{W}_{i,j}^{o,e} + (\mathbf{h}_i^{o,e})^* \odot \mathbf{E}_j(k), \quad (10)$$

where $*$ denotes a conjugate operation.

The processing flow of the algorithm is summarized in Table I. A drawback is that the computations of the conventional FDE are performed on an one-sample-per-symbol basis due to the ‘‘odd/even’’ FDE configuration, which individually processes odd and even blocks. This approach may lead to neglecting the a priori knowledge of signal bandwidth occupation that is spectrally shaped with a small roll-off factor α approaching 0, information readily available in modern coherent transceivers. Motivated by this consideration, we modify the algorithm to directly handle twofold oversampled received signals (refer to Fig. 1(b)), a method we term as *out-of-band exclusive (OBE) FDE*.

B. Out-of-Band Exclusive (OBE) FDE With RLS Adaptation

This subsection introduces the newly proposed OBE-FDE (Fig. 1(b)). Leveraging a priori knowledge of signal bandwidth occupation, such as the roll-off factor, OBE-FDE selectively performs multiplications exclusively for the in-band frequency components of the signals. Throughout the remainder of this subsection, OBE-FDE is explained, assuming the utilization of a sufficiently small roll-off factor $\alpha \sim 0$. A discussion considering a more general case with an arbitrary α is provided in Section II-C. The processing in OBE-FDE initiates with a direct conversion of the input blocks $\mathbf{u}_i(k)$ into the frequency domain, as follows:

$$\mathbf{U}_i(k) \triangleq \text{FFT}[\mathbf{u}_i(k)]. \quad (11)$$

$\mathbf{W}_{i,j} \in \mathbb{C}^{N \times 1}$ is then exclusively multiplied by $\mathbf{U}_i(k)$, considering only the components corresponding to the frequency bins where a signal is present. This operation effectively reduces the number of complex multiplications from N to approximately $\sim N/2$ for signals with a ‘‘rectangular-shaped’’ frequency profile. For the remaining components (i.e., out-of-band entries of a size $\sim N/2$), the multiplication results are set to zero.

TABLE I
ALGORITHM 1: CONVENTIONAL RLS-ADAPTED FDE AT THE k -TH BLOCK

Line#	Algorithm	Equation	Complexity per step
1:	Input: Input signal blocks $\mathbf{u}_i (i \in \{1, 2, \dots, D\})$		
2:	Output: A block estimate of data streams $\hat{\mathbf{x}}_j (j \in \{1, 2, \dots, D\})$		
3:	Initialize $\mathbf{R}^{o,e} = \beta \cdot \mathbf{I}_{D \times D}$ at each frequency bin		
4:	for $i = 1$ to D do		
5:	$\mathbf{U}_i^{o,e}(k) = \text{FFT} [\mathbf{u}_i^{o,e}(k)]$	(4)	$\frac{ND}{2} \log_2 \frac{N}{2}$
6:	for $j = 1$ to D do		
7:	$\hat{\mathbf{x}}_j =$ the last half of IFFT $[\sum_{i=1}^D \mathbf{W}_{i,j}^o \odot \mathbf{U}_i^o(k) + \sum_{i=1}^D \mathbf{W}_{i,j}^e \odot \mathbf{U}_i^e(k)]$	(5)	$ND^2 + \frac{ND}{4} \log_2 \frac{N}{2}$
8:	$\mathbf{E}_j(k) = \text{FFT} \begin{bmatrix} \mathbf{0}_{N/4} \\ \mathbf{e}_j(k) \end{bmatrix}$	(7)	$\frac{ND}{4} \log_2 \frac{N}{2}$
9:	end for		
10:	end for		
11:	for $m = 1$ to $N/2$ do		
12:	$\mathbf{g}^{o,e}[m] \leftarrow \frac{\mathbf{R}^{o,e}[m] \mathbf{V}^{o,e}[m]}{\lambda + (\mathbf{V}^{o,e})^H [m] \mathbf{R}^{o,e}[m] \mathbf{V}^{o,e}[m]}$	(8)	$ND^2 + 2ND$
13:	$\mathbf{R}^{o,e}[m] \leftarrow \lambda^{-1} \mathbf{R}^{o,e}[m] - \lambda^{-1} \mathbf{g}^{o,e}[m] (\mathbf{V}^{o,e})^H [m] \mathbf{R}^{o,e}[m]$	(9)	$2ND^2$
14:	end for		
15:	for $i = 1$ to D do		
16:	for $j = 1$ to D do		
17:	$\mathbf{W}_{i,j}^{o,e} \leftarrow \mathbf{W}_{i,j}^{o,e} + (\mathbf{h}_i^{o,e})^* \odot \mathbf{E}_j(k)$	(10)	ND^2
18:	end for		
19:	end for		
	Total		$5ND^2 + ND + ND \log_2 N$

The output signals are then obtained by preserving the odd (or, even) components from the latter half of the IFFT output. Mathematically, the processing described above for obtaining $\hat{\mathbf{x}}_j$ is expressed as

$\hat{\mathbf{x}}_j(k) =$ the odd entries at the last half of

$$\text{IFFT} \left[\sum_{i=1}^D \mathbf{W}_{i,j} \odot_{\text{in}} \mathbf{U}_i(k) \right], \quad (12)$$

where \odot_{in} denotes the element-wise product performed only for the in-band components. Corresponding frequency-domain error vector for the OBE processing $\mathbf{E}'_j(k) \in \mathbb{C}^{N \times 1}$ is constructed as

$$\begin{aligned} \mathbf{E}'_j(k) &\triangleq \text{FFT} \begin{bmatrix} \mathbf{0}_{N/2} \\ \text{upsample}(\mathbf{e}_j(k), 2) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0}_{N/2} \\ e_j(k)[1] \\ 0 \\ e_j(k)[2] \\ \vdots \\ e_j(k)[N/4] \\ 0 \end{bmatrix}, \end{aligned} \quad (13)$$

where $\text{upsample}(\mathbf{a}, p)$ is the upsampling function that inserts $p - 1$ zeros between every entry of \mathbf{a} . Filter parameters for the OBE processing, including $\mathbf{g} \triangleq [g_1[m], g_2[m], \dots, g_D[m]]^T \in \mathbb{C}^{D \times 1}$ and $\mathbf{R} \in \mathbb{C}^{D \times D}$, are then updated using the input signal vector at each frequency \mathbf{V} , constructed as $\mathbf{V}[m] = [U_1[m], U_2[m], \dots, U_D[m]]^T \in \mathbb{C}^{D \times 1}$:

$$\mathbf{g}[m] \leftarrow \frac{\mathbf{R}[m] \mathbf{V}[m]}{\lambda + \mathbf{V}^H [m] \mathbf{R}[m] \mathbf{V}[m]}, \quad (14)$$

$$\mathbf{R}[m] \leftarrow \lambda^{-1} \mathbf{R}[m] - \lambda^{-1} \mathbf{g}[m] \mathbf{V}^H [m] \mathbf{R}[m], \quad (15)$$

where correlation matrix \mathbf{R} is initialized as $\mathbf{R} = \beta \cdot \mathbf{I}_{D \times D}$. Similarly to the approach used in the processing for the conventional FDE, a Kalman gain vector is redefined as $\mathbf{h}_i \triangleq [g_i[1], g_i[2], \dots, g_i[N]]^T \in \mathbb{C}^{N \times 1}$. A filter weight vector $\mathbf{W}_{i,j}$ is then updated by

$$\mathbf{W}_{i,j} \leftarrow \mathbf{W}_{i,j} + \mathbf{h}_i^* \odot_{\text{in}} \mathbf{E}_j. \quad (16)$$

The processing flow of the algorithm is summarized in Table II. A notable characteristic of the proposed OBE processing is the selective execution of element-wise computations exclusively for the out-of-band frequency components at Lines #7, #12, #13, and #17 in Table II. This selective approach is anticipated to reduce computational efforts by approximately 50% in scenarios utilizing spectrally shaped signals with $\alpha \sim 0$. A more in-depth analysis of this complexity mitigation is provided in Section II-C.

C. Complexity Analysis

This subsection discusses a computational resource requirement achieved by two FDE schemes presented in Section II-A and II-B. In this work, the computational requirement is assessed based on the number of necessary complex multiplications or divisions to obtain $ND/4$ symbols during one block processing period, ensuring a fair comparison. All FFT/IFFT operations are assumed to be executed using the radix-2 FFT algorithm, requiring $(N/2) \log_2 N$ complex multiplications for an input block of size N . Letting r be an integer satisfying $r = \lfloor \log_2(4BW) \rfloor$ where B , W , and $\lfloor \cdot \rfloor$ are a symbol rate in GBaud, a required window size for MIMO equalization in nanoseconds, and the floor function which outputs a nearest integer smaller than its input, respectively, FFT block size N is designed as $N = 2^k$. Despite the discussions in the previous sections, which exclusively considered the case with $\alpha \sim 0$, we now extend the analysis to a more general scenario with arbitrary α . Denoting N_{OBE} as the number of frequency-bin entries used in the OBE processing,

TABLE II
ALGORITHM 2: PROPOSED RLS-ADAPTED OBE-FDE AT THE k -TH BLOCK

Line#	Algorithm	Equation	Complexity per step
1:	Input: Input signal blocks $\mathbf{u}_i (i \in \{1, 2, \dots, D\})$		
2:	Output: A block estimate of data streams $\hat{\mathbf{x}}_j (j \in \{1, 2, \dots, D\})$		
3:	Initialize $\mathbf{R} = \beta \cdot \mathbf{I}_{D \times D}$ at each frequency bin		
4:	for $i = 1$ to D do		
5:	$\mathbf{U}_i(k) = \text{FFT}[\mathbf{u}_i(k)]$	(11)	$\frac{ND}{2} \log_2 N$
6:	for $j = 1$ to D do		
7:	$\hat{\mathbf{x}}_j =$ the odd entries at the last half of IFFT $[\sum_{i=1}^D \mathbf{W}_{i,j} \odot_{\text{in}} \mathbf{U}_i(k)]$	(12)	$\varepsilon ND^2 + \frac{ND}{2} \log_2 N$
8:	$\mathbf{E}'_j(k) = \text{FFT} \left[\begin{array}{c} \mathbf{0}_{N/2} \\ \text{upsample}(\mathbf{e}_j(k), 2) \end{array} \right]$	(13)	$\frac{ND}{2} \log_2 N$
9:	end for		
10:	end for		
11:	for $m = 1$ to εN do		
12:	$\mathbf{g}[m] \leftarrow \frac{\mathbf{R}[m]\mathbf{V}[m]}{\lambda + \mathbf{V}^H[m]\mathbf{R}[m]\mathbf{V}[m]}$	(14)	$\varepsilon(ND^2 + 2ND)$
13:	$\mathbf{R}[m] \leftarrow \lambda^{-1}\mathbf{R}[m] - \lambda^{-1}\mathbf{g}[m]\mathbf{V}^H[m]\mathbf{R}[m]$	(15)	$2\varepsilon ND^2$
14:	end for		
15:	for $i = 1$ to D do		
16:	for $j = 1$ to D do		
17:	$\mathbf{W}_{i,j} \leftarrow \mathbf{W}_{i,j} + \mathbf{h}_i^* \odot_{\text{in}} \mathbf{E}_j$	(16)	εND^2
18:	end for		
19:	end for		
	Total		$\varepsilon(5ND^2 + 2ND) + \frac{3}{2}ND \log_2 N$

which is designed as

$$N_{\text{OBE}} = \left\lfloor \frac{N(1 + \alpha)}{2} \right\rfloor, \quad (17)$$

the “reduction” ratio of computational efforts introduced by OBE processing for in-band entries of signals is expressed as

$$\varepsilon \triangleq \frac{N_{\text{OBE}}}{N}, \quad (18)$$

which satisfies $0 \leq \varepsilon \leq 1$.

We first take an scenario for the computations with the conventional FDE, where FFT/IFFT with a block size of $N/2$ are required $2D$, D , and D times in (4), (5), and (7), respectively. Additionally, the element-wise product totally requires ND^2 , $ND^2 + 2ND$, $2ND^2$ and ND^2 , for the computations in (5), (8), (9) and (10), respectively. The resultant total computational complexity for conventional FDE, $C_{\text{C-FDE}}$, becomes

$$C_{\text{C-FDE}} = 5ND^2 + ND + ND \log_2 N. \quad (19)$$

Next we consider the total computational complexity for OBE-FDE, $C_{\text{OBE-FDE}}$. The main difference in comparison to conventional FDE is that the proposed OBE-FDE selectively executes multiplications only for the in-band frequency entries of the signals. The rest entries located at the out-of-band frequency bins defaults to zeros. In a most intuitive scenario with a sufficiently small α (~ 0), we can roughly halve a required number of complex multiplications or divisions. Estimation of $C_{\text{OBE-FDE}}$ with an arbitrary α is also possible if we split the OBE processing into steps, comprising FFT/IFFT or multiplications/divisions. For the OBE processing, FFT/IFFT with a block size of N is required D times each in (11), (12), and (13). The element-wise product involves computations in (12), (14), (15), and (16) and requires εND^2 , $\varepsilon(ND^2 + 2ND)$, $2\varepsilon ND^2$, and εND^2 , respectively. Therefore, the computational complexity

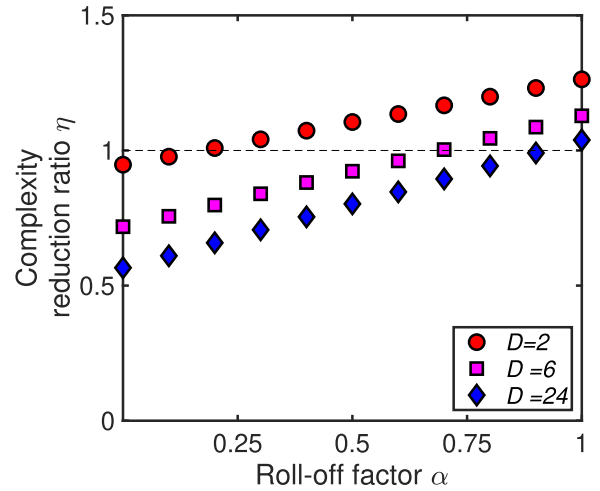


Fig. 2. Complexity reduction ratio η achieved by OBE-based MIMO-FDE with RLS adaptation as a function of roll-off factor α with $N = 256$ and $D \in \{2, 6, 24\}$.

$C_{\text{OBE-FDE}}$ is obtained by summing these up:

$$C_{\text{OBE-FDE}} = \varepsilon(5ND^2 + 2ND) + \frac{3ND}{2} \log_2 N. \quad (20)$$

Comparative computational complexity requirement for both FDE schemes is then evaluated with a complexity reduction ratio $\eta \triangleq C_{\text{OBE-FDE}}/C_{\text{C-FDE}}$. We consider the cases with $N \in \{256, 1024\}$ and $D \in \{2, 6, 8, 12, 20, 24\}$, which should be achievable in a standard-cladding single mode fiber (SMF), a MMF, or a CC-MCF. Fig. 2 depicts achieved η for various α settings when $N = 256$ and $D \in \{6, 20\}$. It is important to note that, for a given α , complexity mitigation with η is evaluated through (17), (18), (19), and (20). The application of OBE into SMF transmission with $D = 2$ does not contribute to decreasing the complexity requirement, especially in a scenarios with larger α . For other MDM transmission scenarios with D of 6 or 20, the reduction in complexity is also less effective with larger α .

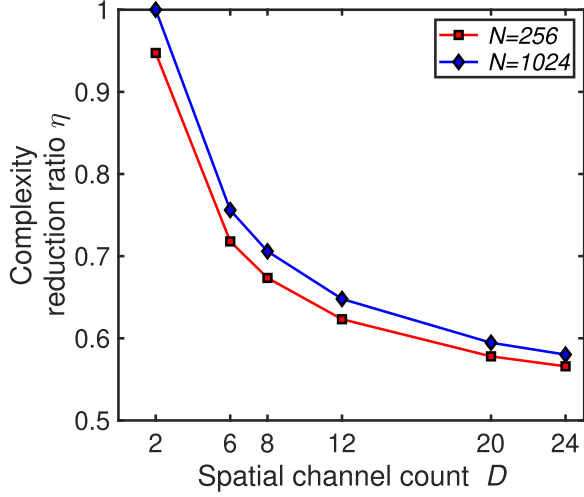


Fig. 3. Complexity reduction ratio η achieved by OBE-based MIMO-FDE with RLS adaptation as a function of spatial channel count D with $N \in \{256, 1024\}$ and $D \in \{2, 6, 8, 12, 20, 24\}$.

This is because the OBE processing contributes to complexity reduction by skipping complex multiplication for frequency bins in the signal's absence. In a specific range of α with $\alpha \leq \sim 0.7$, however, it does relax the complexity requirement for MDM transmission with D of 6 or 20, resulting in η values reaching below 1. In a practical use case of the OBE, small α (~ 0) is expected to be optimum in terms of achieving the most reduced computational complexity and ensuring WDM channel alignment without wasteful bandwidth occupation between neighboring channels.

The dependence of η on the spatial channel count D is also assessed in Fig. 3 when α is fixed at 0.01, while D is varied within the set of $\{2, 6, 8, 12, 20, 24\}$. We find that complexity reduction of η below 1 is achieved for almost all D scenarios, and more depressed in a larger D regime, indicating a suitable application to high-mode-count MDM system rather than SMF systems. Notably, η is less sensitive to a block size N because of the logarithmic dependence on η . In an extreme case where $D \gg \log_2 N$, η would approach to 0.5 (i.e., 50%), while in a practical case, the use of MIMO-FDE with OBE processing reduces computational burdens by $\sim 30\%$ and $\sim 40\%$ in transmissions over 4-core CC-MCF ($D = 8$) and 10-mode FMF ($D = 20$), respectively, both of which will be experimentally evaluated in Section III.

D. Numerical Simulation

In this subsection, we conduct a numerical evaluation of spatially coupled SDM signals processed by a MIMO equalizer adapted with either LMS or RLS adaptation. Our interests encompass two aspects: firstly, a comparison of bit error rate (BER) performance obtained by the OBE-based exclusive processing over a specific frequency range. Secondly, an investigation of whether a robust equalization performance is maintained under scenarios with various transmission parameters including MDL, the spatial channel count D , and signal symbol rate. To emulate an accumulation of modal dispersion and MDL along

a transmission link, we here introduce a random matrix model with strong mode coupling [14], [15]. We consider a strongly coupled MDM transmission system comprises K sections, each having an *uncoupled* matrix $\Lambda^{(k)}(\omega)$ at an angular frequency ω :

$$\Lambda^{(k)}(\omega) \triangleq \left[e^{g_1^{(k)}/2 - j\omega\tau_1^{(k)}}, e^{g_2^{(k)}/2 - j\omega\tau_2^{(k)}}, \dots, e^{g_D^{(k)}/2 - j\omega\tau_D^{(k)}} \right], \quad (21)$$

where $g_i^{(k)} \in \{1, 2, \dots, D\}$ and $\tau_i^{(k)}$ denote an i -th uncoupled mode-averaged loss and i -th uncoupled modal delay at a k -th section, respectively. Each $g_i^{(k)}$ and $\tau_i^{(k)}$ varies with a root mean square (rms) value of σ_g and σ_τ , respectively, which characterizes the statistics of overall MDL and modal dispersion in strongly coupled MDM systems [14], [15]. With an assumption that each section is concatenated by random unitary matrices $\mathbf{P}^{(k)}$ and $\mathbf{Q}^{(k)}$, a transfer matrix for the k -th section, $\mathbf{M}^{(k)}$, is given by

$$\mathbf{M}^{(k)} = \mathbf{P}^{(k)} \Lambda^{(k)} \mathbf{Q}^{(k)H}, \quad (22)$$

and an overall transfer matrix of the system, \mathbf{M} , is constructed as

$$\mathbf{M} = \prod_{k=1}^K \mathbf{M}^{(k)}. \quad (23)$$

An end-to-end rms MDL, σ_{MDL} , grows with a curve of [14], [15]

$$\sigma_{\text{MDL}} = \xi \sqrt{1 + \xi^2 / (12(1 - D^2))}, \quad (24)$$

where ξ is an accumulated rms MDL expressed as $\xi \triangleq \sqrt{K} \sigma_g$. In the subsequent analysis, we simplify the model by neglecting the frequency-dependent effect of $\Lambda^{(k)}$ and the fiber nonlinear effect for the sake of simplicity. The end-to-end signal propagation in a linear transmission regime is described by an MIMO model at each sample time:

$$\mathbf{u} = \mathbf{M}\mathbf{x} + \mathbf{n}, \quad (25)$$

where $\mathbf{n} \in \mathbb{C}^{D \times 1}$ is a noise vector that satisfies $E[\mathbf{n}\mathbf{n}^H] = P_n \cdot \mathbf{I}_{D \times D}$, where $E[\cdot]$ and P_n are an expectation function and a MDL-free noise power for each spatial channel, respectively. Note that for a simpler analysis purpose, \mathbf{n} was added in a lumped manner at the receiver side, and also that no co-propagating WDM channel was assumed. The use of \mathbf{u} as an input signal source into the MIMO equalizer enables us to assess an equalization performance of filter adaptation schemes.

In the conducted simulation, we transmitted signals with QPSK, 16QAM, or 64QAM modulations over an MDM link expressed with Equation (25), and set the total number of sections K and a FFT block size N as 50 and 512, respectively. We also varied rms MDL of each section σ_g , and the spatial channel count D as $\sigma_g \in \{0.1, 0.2, 0.4, 0.8, 1.2\}$ in dB unit, and $D \in \{6, 8, 10, 12, 14\}$. To focus on clarifying MDL impact on convergence behavior of filter coefficient learning, we first analyze them with a zero modal-dispersion model with σ_τ fixed to 0 ps (whose results will be provided in Figs. 4 to 7), followed by an analysis with σ_τ fixed to 70 ps (in Figs. 8 and 9). We also considered signal symbol rates $B \in \{1, 10, 30\}$ GBaud, although the default setting was 10 GBaud unless otherwise

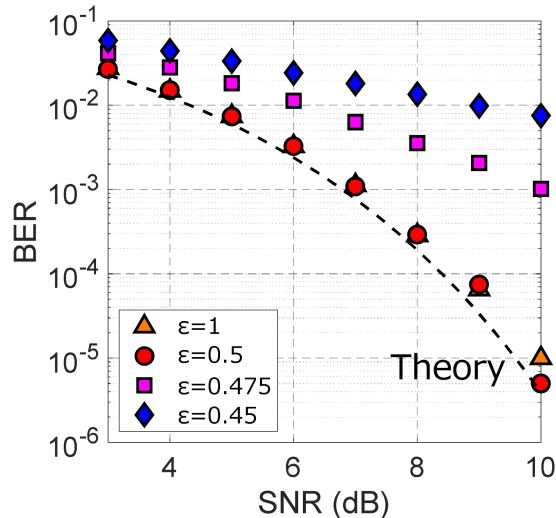


Fig. 4. BER transitions of QPSK signals as a function of SNR obtained by RLS-adapted MIMO-FDE under the scenarios with σ_g of 0.8 dB, σ_τ of 0 ps and D of 6 when ε is changed within the set of $\{0.45, 0.475, 0.5, 1\}$.

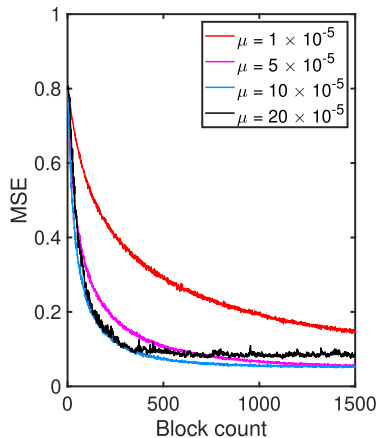


Fig. 5. MSE transition of QPSK signals obtained by LMS-adapted FDE with σ_g of 0.8 dB, σ_τ of 0 ps, and D of 6.

specified. For a fair comparison among different signal formats, signal-to-noise ratio (SNR) was set to 10, 16.7, and 22.7 dB for QPSK, 16QAM, and 64QAM, respectively, to achieve the same BERs. Since M depends on the statistics of σ_g and σ_τ , we obtained filter equalization performance in each scenario by averaging them with 10 transfer matrix realizations.

We commence a numerical evaluation with the equalization performance of OBE-based MIMO-FDE, fixing α , D , and N at 0.01, 6, and 256, respectively, while varying N_{OBE} . Although an optimum (i.e., necessary and sufficient) value of N_{OBE} , and the corresponding ε , can be designed using (17) and (18), we deliberately explore scenarios where ε is changed even below 0.5. BER transitions obtained by RLS-adapted MIMO-FDE in a MDL- and modal-dispersion-free MDM QPSK-modulated signal transmission (i.e., $\sigma_g = 0$ dB and $\sigma_\tau = 0$ ps) with various SNR are shown in Fig. 4 with $\varepsilon \in \{0.45, 0.475, 0.5, 1\}$. An almost perfect overlap of the BER curves is obtained for the cases with ε of 0.5 and 1, indicating that there is a negligible BER penalty even if we skip the computations of MIMO-FDE over the

half bandwidth of the signal. For smaller values of $\varepsilon < 0.5$, however, the BER performance was significantly degraded. Even a modest reduction in ε by 2.5% of the overall signal's bandwidth (specifically, changing ε from 0.5 to 0.475) led to noticeable deterioration in BER performance. Nevertheless, the results presented in Fig. 4 demonstrate the feasibility and effectiveness of the OBE processing, provided that N_{OBE} is properly designed.

In the subsequent numerical evaluation, equalization performance is measured using normalized mean squared error (MSE) that is obtained by averaging the error e_j (defined in (6)) for all spatial modes at each block time divided by the error at the initial block time. The primary parameter when discussing on convergence behaviors of LMS-adapted MIMO filters is the step size parameter, denoted as μ . In Fig. 5, an example of MSE transient behavior of QPSK signals is presented for various μ in the set of $\{1 \times 10^{-5}, 5 \times 10^{-5}, 1 \times 10^{-4}, 2 \times 10^{-4}\}$. The values of other parameters are fixed with σ_g , σ_τ , and D at 0.6 dB, 0 ps, and 6, respectively. The figure illustrates that a smaller value of μ leads to reliable but slower convergence, whereas larger values of μ result in a faster learning. However, as observed in the curve with μ of 2×10^{-4} , it is important to note that an excessively large μ introduces an enhanced excess MSE [8]. Nevertheless, we confirmed that a value of $\mu = 1 \times 10^{-4}$ yields relatively satisfactory MSE performance across various parameter settings. Consequently, this value was employed in the subsequent evaluations for LMS adaptations.

We then evaluate convergence performance against increasing σ_g with a fixed D of 6. Fig. 6 shows MSE transitions obtained by LMS (Fig. 6(a) to (c)) or RLS adapted MIMO-FDE (Figs. 6(d) to (f)). For all modulation format settings, MSE curves with small values of σ_g (≤ 0.2 dB) represent consistently negligible impact on MSE behavior for both adaptations. Conversely, the LMS adaptation demonstrates significantly slower convergence for σ_g values of 0.4 dB or higher, while the behavior of MSE curves with the RLS adaptation was less affected even by these larger σ_g . This property introduced by RLS adaptation arises from the insensitivity to the eigenvalue spread of a correlation matrix of \mathbf{u} . Such insensitivity proves advantageous in multi-span MDM transmission scenarios, where MDL is expected to accumulate. Note that differences in achieved MSE floor after convergence among modulation formats stem from the difference in SNR. Increased σ_g may have little impact on 64QAM signals due to the filter coefficients converging to the zero-forcing solution in high SNR regimes. In the conducted simulation, despite setting a forgetting factor λ to 0.99 for RLS adaptation, a smaller λ (e.g., 0.97) may offer faster learning for MIMO filter coefficients. Unfortunately, although not shown here, we observed faster but unstable learning curves with such a smaller λ , possibly stemming from a numerical instability in the algorithm under finite-precision implementation [8]. This issue might be mitigated by initially employing RLS adaptation for the initial convergence phase and then transitioning to LMS adaptation. Another approach to overcome this problem is to use QR decomposition-based RLS (QRD-RLS) [8], [16], [17], which, however, is beyond the scope of this paper. Here, our focus is on clarifying the impacts on MSE behaviors when σ_g or D is changed.

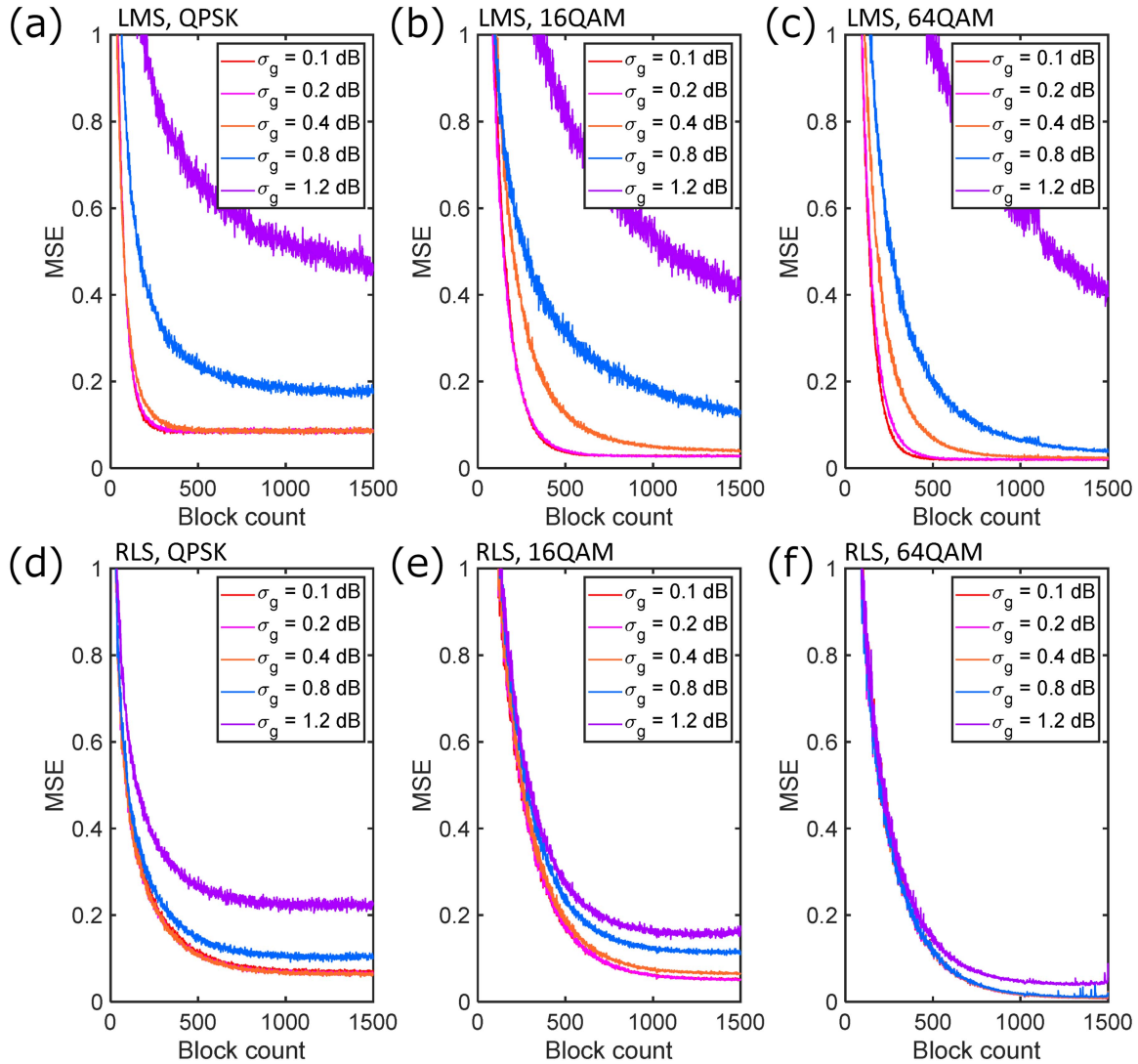


Fig. 6. (a)–(c): MSE transitions as a function of a block count with $D = 6$ and $\sigma_g \in \{0.1, 0.2, 0.4, 0.8, 1.2\}$ in dB obtained by MIMO-FDE with LMS adaptation for (a) QPSK, (b) 16QAM, and (c) 64QAM modulation. (d)–(f): MSE transitions with the same parameter setup in (a) to (c) with RLS adaptation for (d) QPSK, (e) 16QAM, and (f) 64QAM modulation.

Our next focus is to evaluate the convergence performance of two adaptation schemes with a fixed D of 6 in an increased D regime, and the results are compared in Fig. 7. While there is not a drastic difference in MSE behaviors between LMS (Figs. 7(a) to (c)) and RLS adaptations (Figs. 7(d) to (f)), it is evident that RLS adaptation consistently achieves smaller MSE in every D regime, with overlapping MSE curves regardless of the modulation formats used. This indicates that RLS outperforms LMS even in high-mode-count scenarios. On the other hand, the increased excess MSE in the LMS-adapted learning curve for larger D may be attributed to an enhanced MIMO equalizer window size [8], which linearly scales with D .

The convergence results with various baud rate settings B of {1, 10, 30} GBaud and σ_g of {0.1, 0.4, 0.8} in dB under the MDM transmission with a fixed σ_τ of 70 ps and $D = 6$ are depicted in Fig. 8. For all σ_g , the obtained convergence time was independent on B , indicating that, with a fixed parameter set of both MIMO equalization schemes (namely, μ , λ and N), MIMO

filter learning rate is determined only by the number of symbols used for filter training. We also observed that smaller MSE was achieved with higher B (Fig. 8(c) and 8(f)) in a transmission scenario with a large σ_g of 0.8 dB. This may be attributed to the averaging effect of MDL over frequency, with MDM signals occupying a wider bandwidth relative to the coherence bandwidth, which is inversely proportional to the modal dispersion spread [18]. From the practical design perspective of B , it is important to discuss convergence time using absolute time instead of FFT block count. Fig. 9 illustrates MSE transitions evaluated with absolute time, which corresponds to a replot of Fig. 8(a) and 8(d). Faster filter learning was observed with higher symbol-rate signals for both adaptations.

In any case, as seen in Figs. 6 to 9, our simulation results demonstrate that MIMO equalization with RLS adaptation provides robust convergence in scenarios involving long-haul MDL accumulation and/or high-mode-count MDM transmissions.

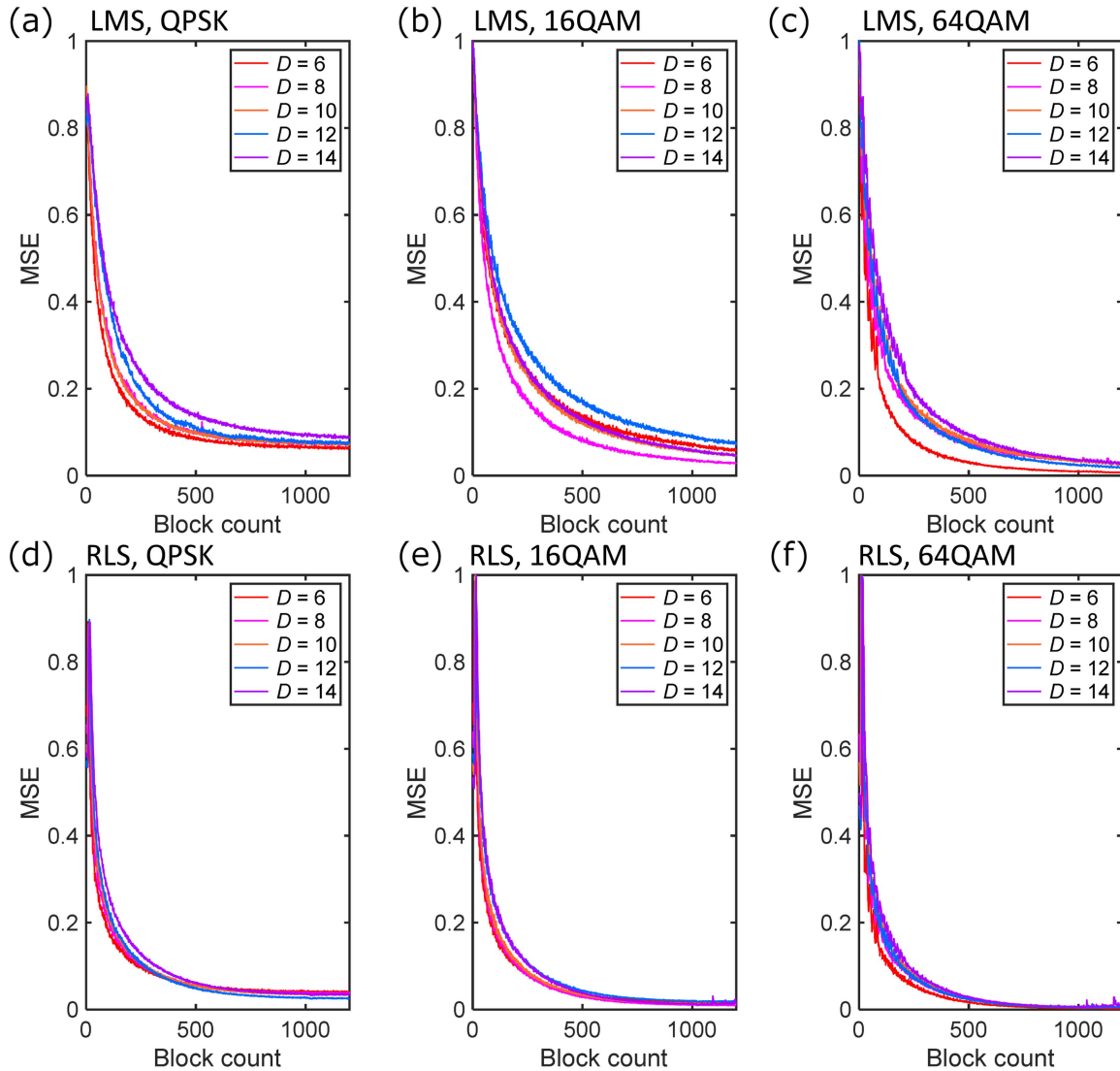


Fig. 7. (a)–(c): MSE transitions as a function of a block count with $\sigma_g = 0.8$ dB and $D \in \{6, 8, 10, 12, 14\}$ obtained by MIMO-FDE with LMS adaptation for (a) QPSK, (b) 16QAM, and (c) 64QAM modulation. (d)–(f): MSE transitions with the same parameter setup in (a) to (c) with RLS adaptation for (d) QPSK, (e) 16QAM, and (f) 64QAM modulation.

III. EXPERIMENTAL SETUP

We conducted MDM transmission experiments to evaluate the proposed OBE-FDE scheme by using almost the same experimental setup of inline amplified mode-multiplexing transmission supporting up to 10 spatial modes [13], as shown in Fig. 10. At the transmitter, a channel under test (CUT) of 2×6 -GBaud dual QPSK signals digitally shaped with α of 0.01 was created from a binary pattern coded by a low density parity check code (LDPC) with a code rate of 4/5 defined in the DVB-S2 standard and a BCH (30832, 30592) code with the hard-decision (HD) forward error correction (FEC) threshold BER of 5×10^{-5} [19] to remove the error floor after LDPC decoding to achieve post-FEC BER of 1×10^{-15} . Each transmission frame was repeated as a 34320-symbol-length QPSK-pattern also containing 1.4%-OH for the training sequence. An advantage of using low symbol-rate signals driven at 6 GBaud is the ability to reduce the required

number of MIMO filter taps to 800, which is necessary to deal with signal pulses broadened with a time window exceeding several tens of nanoseconds after 1000 km transmission induced by modal dispersion [13]. CUT was then decorrelated through polarization division multiplexed (PDM) emulator with a 295-ns delay, combined with wavelength selective switch (WSS) shaped amplified spontaneous emission (ASE) sources to yield 11-WDM 12.5-GHz-spaced signals locating from 1549.556 nm to 1550.558 nm, and again decorrelated for producing MDM signal inputs with delay lines. To achieve fully-decorrelated MDM channels with a channel count of up to 10, delays were set to 567, 1204, 1796, 2388, 2965, 3534, 4149, 4751, 5232 ns for mode 1 through mode 10, respectively. The setup yielded WDM-MDM signals with PDM-QPSK-modulated channels, achieving a throughput of 37.56 Gbps/ λ /mode, a spectral efficiency (SE) of 3.00 bps/Hz/mode, and a normalized generalized mutual information (NGMI) threshold of 0.836 [20]. To perform

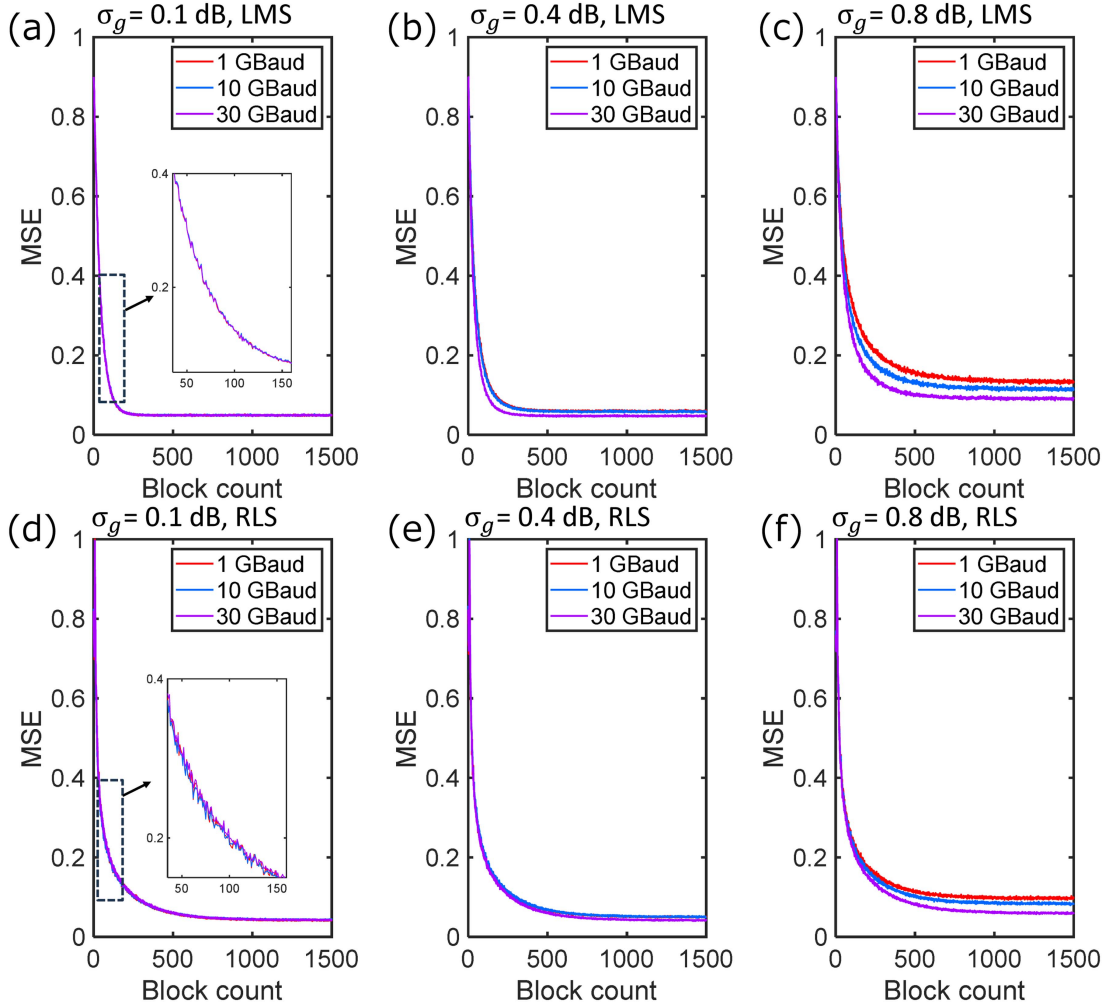


Fig. 8. (a)–(c): MSE transitions of QPSK signals driven at various symbol rate B of {1, 10, 30} GBaud under the MDM transmission link with fixed σ_τ of 70 ps as a function of a block count obtained by MIMO-FDE with LMS adaptation for the transmission cases of $D = 6$ and σ_g of (a) 0.1, (b) 0.4, and (c) 0.8 dB. (d)–(f): MSE transitions with the same parameter setup in (a) to (c) with RLS adaptation for the transmission cases of σ_g of (d) 0.1, (e) 0.4, and (f) 0.8 dB. Insets in (a) and (d) are a zoomed view of MSEs during the block count of around 50 to 150.

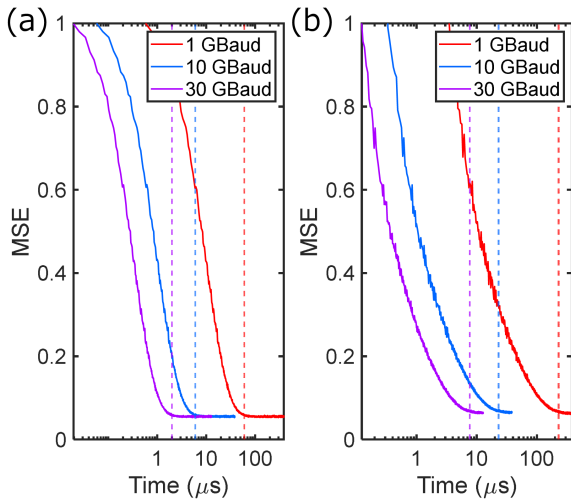


Fig. 9. MSE transitions of QPSK signals driven at various symbol rate B of {1, 10, 30} GBaud for the transmission cases of $D = 6$ and $\sigma_g = 0.1$ dB as a function of an absolute time with (a) LMS and (b) RLS adaptations. Vertical broken lines in each panel are the time when the convergence is achieved.

MDM long-haul transmission, ten-fold recirculating loop systems were also constructed, each comprising multi-core/discrete Erbium-doped fiber amplifiers (EDFAs), optical bandpass filters (OBPFs), and acousto-optic modulator (AOM) loop switches. Inline optical amplifiers for fully compensating span loss were employed with discrete EDFAs and integrated core-pumped 7-core multicore EDFAs [21].

We tested two kinds of MDM fibers listed in Fig. 10(b) and 10(c): one was a pure-silica 4-core CC-MCF [22] with a total length of 53.9 km, an effective area of $112 \mu\text{m}^2$ at 1550 nm, an attenuation of 0.157 dB/km, and a spatial mode dispersion of $8 \text{ ps}/(\text{km})^{1/2}$. Another was a trench-assisted graded-index (GI) 6-LP FMFs supporting 10 spatial modes [13] with a total length of 52 km, an effective area of $80 \mu\text{m}^2$ at 1550 nm for LP_{01} , an attenuation of 0.25 dB/km, and a spatial mode dispersion of 157 ps/km. Devices responsible for signal conversion into the spatial domain were fan-in/fan-out (FIFO) devices for the 4-core CC-MCF transmission, and multi-plane light conversion-based mode-selective multiplexer/demultiplexer (MUX/DEMUX) for the 6-LP FMF transmission. When FMFs were tested, we

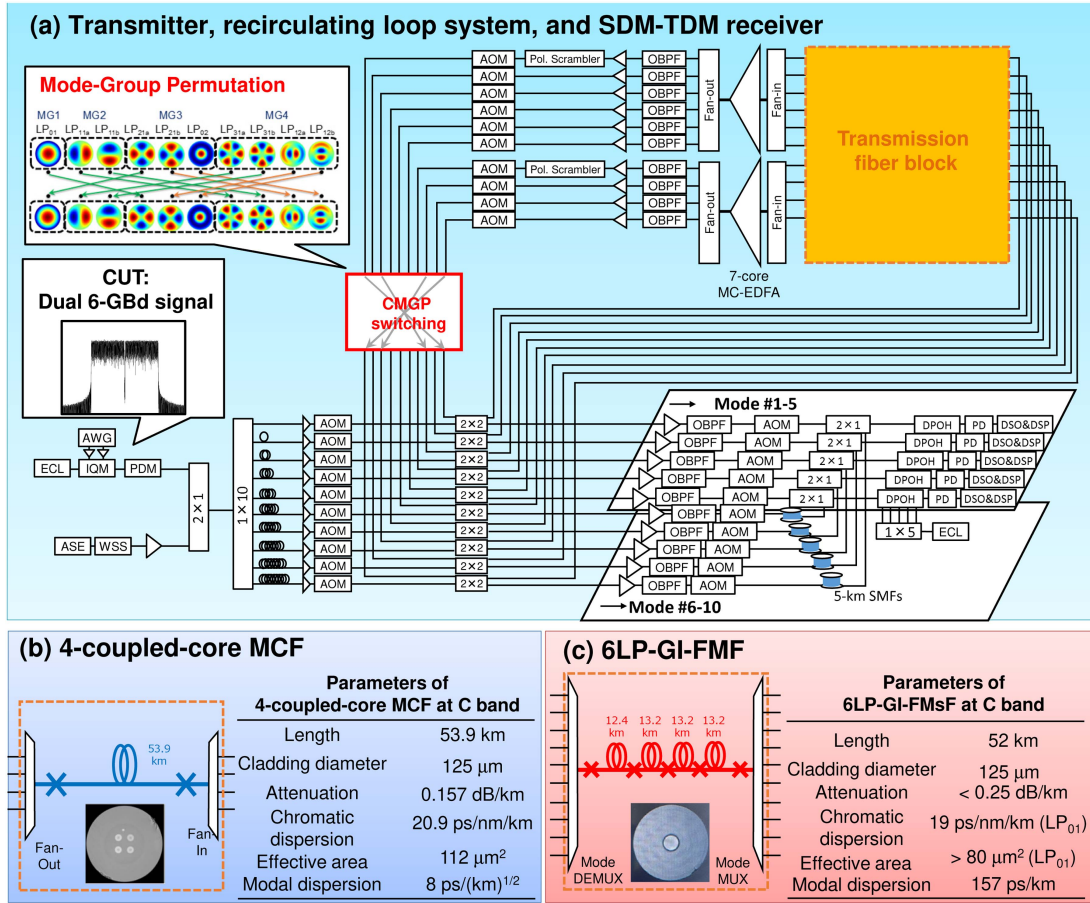


Fig. 10. (a) Experimental setup for transmitter, ten-fold recirculating loop system, and SDM-TDM receiver to support a long-haul MDM signal reception with a spatial channel count D up to 20. (b) Transmission fiber block configuration of 4-coupled core MCF along with fiber parameters. (c) Transmission fiber block configuration of 6LP-GI-FMF along with fiber parameters.

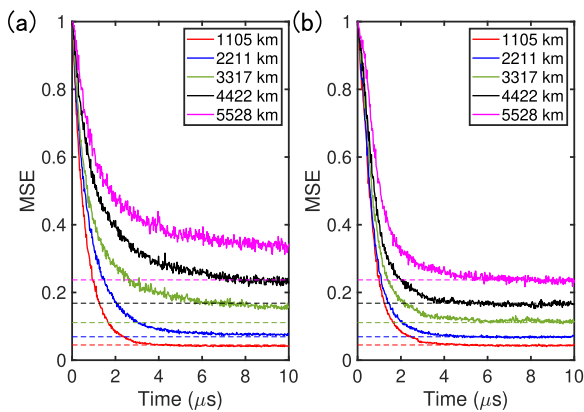


Fig. 11. Learning curves of MIMO equalization coefficients in 4-core CC-MCF transmission for (a) LMS adaptation and for (b) RLS adaptation. Dotted are MSE baselines achieved by the RLS after $\tau_{\text{RLS}} = 6 \mu\text{s}$.

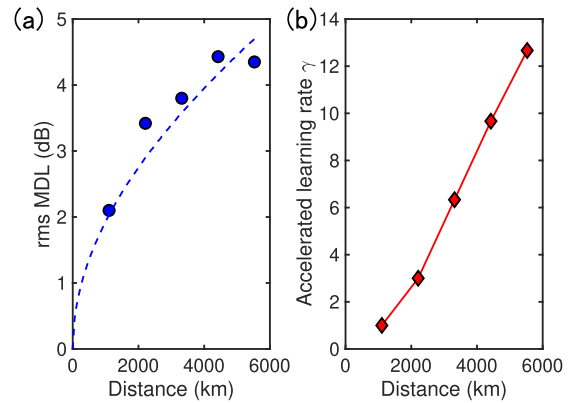


Fig. 12. (a) rms MDL evolution as a function of distance. (b) Accelerated MIMO equalization learning rate γ introduced by switching from LMS adaptation to RLS adaptation.

introduced a cyclic mode-group permutation technique (CMGP) [20] at the ends of recirculating loop systems to minimize the signal pulse spread induced by modal dispersion. Note that, in an evaluation for the FMF transmission, each spatial channel with integer (i.e., spatial channels #1 through #10) corresponds to those launched at the first span as LP₀₁, LP_{11a}, LP_{11b}, LP_{21a},

LP_{21b}, LP₀₂, LP_{31a}, LP_{31b}, LP_{12a}, and LP_{12b}, respectively. Loop-synchronous polarization scramblers were incorporated in both the first and seventh recirculating loops, considering the cyclic excursion of the signal over all the recirculating loops. This ensures a continuous randomization of the polarization state as the signal traverses each span.

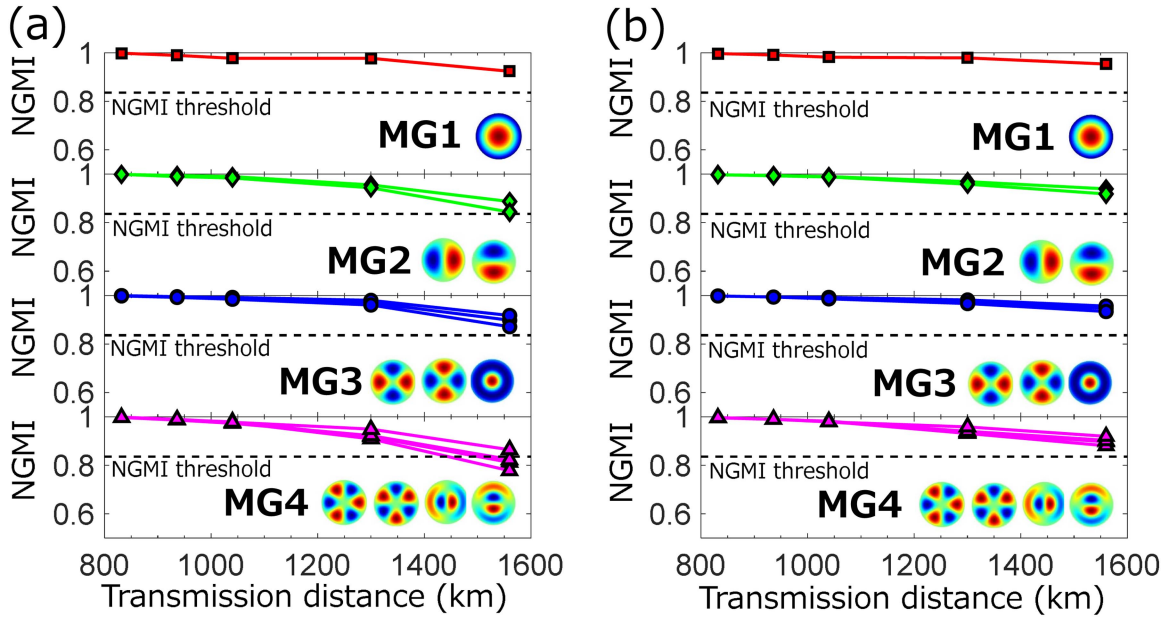


Fig. 13. NGMI for each MG in 10-mode FMF transmission at wavelength channel of λ_7 obtained by (a) LMS-adapted. (b) RLS-adapted MIMO-FDE. NGMI threshold of 0.836 assumes a combined use of LDPC and BCH codes [13], [20].

After transmissions, MDM signals were detected using a coherent SDM time division multiplexing (SDM-TDM) receiver setup, which shared optics/electronics for coherent reception, as described in [23]. In the configured SDM-TDM receivers, relative delays were introduced using 5-km single mode fibers (SFMs) before the detection process. These delays were then combined by a 2×1 coupler to ensure proper alignment in the time domain. The signals were subsequently received by reduced-count coherent receivers with spatial channel pairs #1, #6, #2, #7, #3, #8, #4, #9, and #5, #10 (Fig. 10(a)). Thereby each SDM-TDM receiver was assigned for receptions of 2 spatial channels, saving receiver setup counts from 10 to 5. Note that the laser linewidth both of Tx and local oscillator lasers was 25 kHz. Multiple unsynchronized phase errors, resulting from the direct input configuration of the LOs into the coherent receivers and superimposed over the detected data signals, were effectively mitigated using the MIMO carrier phase recovery (CPR) technique [24]. The detected signals were then stored for an off-line processing, performing front-end error correction, chromatic dispersion compensation, and MIMO-CPR-embedded MIMO-FDE. Signals were then processed by OBE-enabled reduced-complexity MIMO-FDE either with LMS or RLS adaptations. In both evaluations with the CC-MCF transmission and the FMF transmission, the FFT block size, the step-size parameter for LMS adaptation, and the forgetting factor for RLS adaptation were consistently set to 800, 10^{-4} , and 0.99, respectively.

IV. EXPERIMENTAL RESULTS

We start an experimental evaluation with transmission results over the 4-core CC-MCF. Fig. 11 compares learning curves for MIMO equalization filter coefficients obtained by LMS (Fig. 11(a)) or RLS adaptations (Fig. 11(b)) in terms of

normalized MSE. While the rate of convergence was almost identical for both adaptation schemes at 1105 km, it exhibited a slower pace for LMS adaptation at longer distances. This is attributed to the direct impact of accumulated MDL on the eigenvalue spread of input signals into the MIMO equalization, as illustrated in Fig. 12(a) with an estimation of rms MDL based on the singular value decomposition of a transfer matrix. Also shown in Fig. 12 is a theoretical curve given in (24)(a) [14], [15], giving σ_g of 0.45 dB. On the contrary, the learning curves by RLS adaptation was insensitive to increased transmission reach, fully converged by $\tau_{\text{RLS}} = 6 \mu\text{s}$ at all distances (Fig. 11(b)). We also estimated a convergence time for the LMS-adapted cases, τ_{LMS} , by extrapolating the curves until reaching to baselines of MSE achieved by the RLS (dotted lines in Fig. 11(a) and 11(b)). Fig. 12(b) represents an accelerated learning rate, γ , defined by the ratio of τ_{RLS} to τ_{LMS} evaluated at each distance, showing that RLS adaptation provided a faster convergence at longer distance. Particularly, ten-fold faster learning of the filter coefficients was obtained at 5528 km. We also applied both FDE schemes to a long-haul 10-mode FMF transmission where we employed the CMGP technique. With this technique, we have verified that the accumulated modal dispersion effects can be suppressed up to 82% [13]. Regarding mode-dependent loss (MDL), based on the results of MDL estimation obtained from the MIMO equalization filter coefficients, we could not observe a mitigation effect of MDL itself in our experiment. Nevertheless, we emphasize that the impact of MDL on MDM signal performance can be improved by CMGP. This is evident as the BER difference between spatial modes is clearly reduced when mode permutation technique is applied [5]. NGMI results of each mode group are represented in Figs. 13 with LMS adaptation (Fig. 13(a)) and RLS adaptation (Fig. 13(b)). Little NGMI difference between both FDE schemes was observed

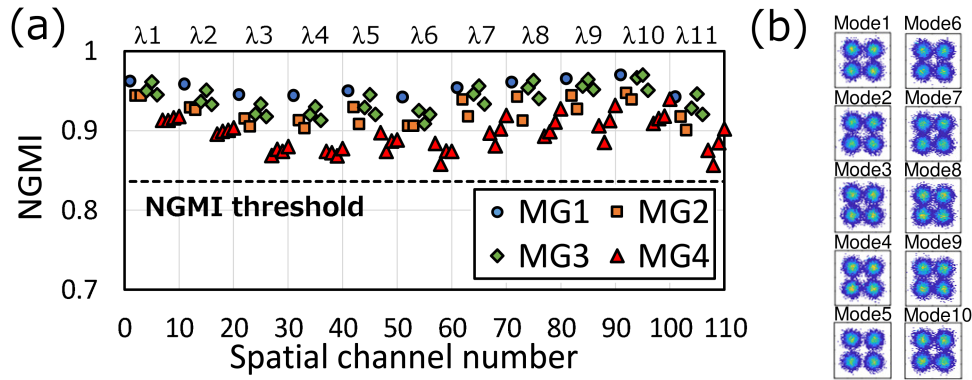


Fig. 14. (a) NGMI for all wavelength/spatial channels at 1560 km obtained by MIMO-FDE with RLS adaptation. (b) Representative constellations of wavelength channel of λ_7 at 1560 km.

at distances shorter than 1000 km, while decoupling for the mixed MDM signals was successfully performed at 1560 km only by the RLS adaptation even in the presence of a peak-to-peak MDL exceeding 10 dB. From the NGMI results for all wavelength/spatial channels after 1560-km WDM/MDM transmission shown in Fig. 14, we confirmed the achievement of a record-long-distance 10-mode Few-Mode Fiber (FMF) transmission using RLS-adapted OBE-FDE, extending the reach by 20% compared to our previous work covering 1300 km [13]. This accomplishment was made while simultaneously reducing the computational burden by 40%, as discussed in Section II-C.

In the conducted experiment, due to limited experimental resources, we unfortunately employed around 1 nm bandwidth for evaluating MDM transmission. However, in [13], we confirmed that the CMGP technique is applicable over a wide bandwidth in the C band, covering around 20 nm. As reported in [25], full C-band long-haul transmission with a capacity of 273.6 Tbps was achieved with the help of the CMGP technique. We believe that high-capacity long-distance 10-mode transmission could also be achieved by combining our novel key enablers of CMGP and OBE-based RLS-adapted FDE techniques.

V. CONCLUSION

We proposed a computationally efficient implementation technique for recursive least squares (RLS) adapted MIMO frequency-domain equalization (FDE). Referred to as out-of-band exclusive (OBE) FDE, the scheme achieves complexity reduction by omitting redundant computations on the outside bandwidth of signals. Simulation results highlight the enhanced complexity reduction with an increased number of spatial modes, particularly within a specific roll-off factor range, rendering it highly suitable for high-mode-count mode division multiplexed (MDM) systems. Our study further reveals that, regardless of signal modulation formats, MIMO equalization with OBE-based RLS adaptation introduced the improved convergence tolerance against mode-dependent loss (MDL), especially in scenarios involving long-haul MDL accumulation and/or high-mode-count MDM transmissions. The advantages of RLS-adapted OBE-based MIMO-FDE were demonstrated in MDM transmissions, with achievements of 10-fold convergence

speed acceleration over 4-core coupled-core multi-core fiber (CC-MCF) and the world's longest MDL-tolerant 10-mode few-mode fiber (FMF) transmission over 1560 km with a reduced computational complexity by 40%.

REFERENCES

- [1] G. Rademacher et al., "Randomly coupled 19-core multi-core fiber with standard cladding diameter," in *Proc. Opt. Fiber Commun. Conf.*, 2023, Paper Th4A.4.
- [2] G. Rademacher et al., "153 peta-bit/s C-band transmission in a 55-mode fiber," in *Proc. Eur. Conf. Opt. Commun.*, 2022, Paper Th3C.3.
- [3] E. Ip et al., "146λ × 6 × 19-Gbaud wavelength-and mode-division multiplexed transmission over 10 × 50-km spans of few-mode fiber with a gain-equalized few-mode EDFA," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 790–797, Nov. 2014.
- [4] R. Ryf et al., "Long-distance transmission over coupled-core multicore fiber," in *Proc. 42nd Eur. Conf. Opt. Commun.*, 2016, Paper Th.3.C.3.
- [5] K. Shibahara et al., "DMD-unmanaged long-haul SDM transmission over 2500-km 12-core 3-mode MC-FMF and 6300-km 3-mode FMF employing intermodal interference canceling technique," *J. Lightw. Technol.*, vol. 37, no. 1, pp. 138–147, Jan. 2019.
- [6] D. Soma et al., "50.47-tbit/s standard cladding ultra-low-loss coupled 4-core fiber transmission over 9,150 km," in *Proc. Opt. Fiber Commun. Conf.*, 2021, Paper W7D.3.
- [7] S. O. Arik, D. Askarov, and J. M. Kahn, "Adaptive frequency-domain equalization in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 32, no. 10, pp. 1841–1852, May 2014.
- [8] S. S. Haykin, *Adaptive Filter Theory*, NJ, USA, 2008.
- [9] Z. Yang et al., "Experimental demonstration of adaptive VFF-RLS-FDE for long-distance mode-division multiplexed transmission," *Opt. Exp.*, vol. 26, pp. 18362–18367, Jul. 2018.
- [10] Z. Pan, Y. Weng, X. He, and J. Wang, "Adaptive frequency-domain equalization and MIMO signal processing in mode division multiplexing systems using few-mode fibers," in *Proc. Adv. Photon. (IPR, NOMA, Sensors, Netw., SPPCom, SOF)*, 2016, Paper SpW2G.1.
- [11] K. Shibahara, T. Kobayashi, and Y. Miyamoto, "Out-of-band exclusive frequency-domain MIMO equalization with the reduced-overlap-and-save method for scalable mode multiplexed signal transmission," *Opt. Exp.*, vol. 31, pp. 2302–2315, Jan. 2023.
- [12] K. Shibahara, M. Hoshi, T. Hasegawa, T. Hayashi, and Y. Miyamoto, "Recursive least squares based low-complexity frequency-domain MIMO equalisation for MDL-tolerant long-haul space division multiplexed transmission," in *Proc. Eur. Conf. Opt. Commun.*, 2023, Paper Th.B.5.3.
- [13] K. Shibahara, M. Hoshi, and Y. Miyamoto, "10-spatial-mode 1300-km transmission over 6-LP graded index few-mode fiber with 36-ns modal dispersion," in *Proc. Opt. Fiber Commun. Conf.*, 2023, Paper M2B.2.
- [14] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing," *Opt. Exp.*, vol. 19, pp. 16612–16635, Aug. 2011.

- [15] K. Choutagunta, S. O. Arik, K.-P. Ho, and J. M. Kahn, "Characterizing mode-dependent loss and gain in multimode components," *J. Lightw. Technol.*, vol. 36, no. 18, pp. 3815–3823, Sep. 2018.
- [16] W. M. Gentleman and H. Kung, "Matrix Triangularization by Systolic Arrays," in *Proc. SPIE, Real-Time Signal Process. IV*, vol. 298, pp. 19–26, 1982.
- [17] J. McWhirter, "Recursive Least-Squares Minimization Using a Systolic Array," in *Proc. SPIE, Real-Time Signal Process. VI*, vol. 431, pp. 105–113, 1983.
- [18] K.-P. Ho and J. M. Kahn, "Frequency diversity in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 29, no. 24, pp. 3719–3726, Dec. 2011.
- [19] D. S. Millar et al., "Detection of a 1 TB/s superchannel with a single coherent receiver," in *Proc. 41st Eur. Conf. Opt. Commun.*, 2015, pp. 1–3.
- [20] K. Shibahara, T. Mizuno, H. Ono, K. Nakajima, and Y. Miyamoto, "Long-haul dmd-unmanaged 6-mode-multiplexed transmission employing cyclic mode-group permutation," in *Proc. 43rd Opt. Fiber Commun. Conf. Exhib.*, 2020, Paper Th3H.3.
- [21] K. Takenaga et al., "Multicore EDF optimized for remotely pumped amplification system over multicore fiber," in *Proc. 18th OptoElectronics Commun. Conf.*, 2013, Paper Tu1S-2.
- [22] T. Hayashi, Y. Tamura, T. Hasegawa, and T. Taru, "Record-low spatial mode dispersion and ultra-low loss coupled multi-core fiber for ultra-long-haul transmission," *J. Lightw. Technol.*, vol. 35, no. 3, pp. 450–457, Feb. 2017.
- [23] R. G. H. van Uden, C. M. Okonkwo, H. Chen, H. de Waardt, and A. M. J. Koonen, "Time domain multiplexed spatial division multiplexing receiver," *Opt. Exp.*, vol. 22, pp. 12668–12677, May 2014.
- [24] K. Shibahara, T. Mizuno, and Y. Miyamoto, "Mimo carrier phase recovery for carrier-asynchronous SDM-MIMO reception based on the extended kalman filter," *Opt. Exp.*, vol. 29, pp. 17111–17124, May 2021.
- [25] M. Van Den et al., "273.6 TB/s transmission over 1001 km of 15-mode fiber using 16-QAM C-band signals," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, 2023, Paper Th4B.5.

Kohki Shibahara (Member, IEEE) received the B.S. degree in physics, the M.S. degree in geophysics, and the Ph.D. degree in informatics from Kyoto University, Kyoto, Japan, in 2008, 2010, and 2017, respectively. In 2010, he joined NTT Network Innovation Laboratories.

His research interests include spatial division multiplexing transmission systems and advanced multiple-input and multiple-output signal processing. He is also a Member of the IEICE and the IEEE/Photonics Society. He was the recipient of the Tingye Li Innovation Prize from the OSA in 2016, Young Researcher's Award from the IEICE in 2017, and Maejima Hisoka Encouragement Award from the Communications Association in 2024.

Megumi Hoshi received the B.S. and M.S. degrees in mechanical engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 2018 and 2021, respectively. In 2021, she joined NTT Network Innovation Laboratories. Her research interests include spatial division multiplexing systems and their light propagation properties. She is also a Member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

Takemi Hasegawa (Member, IEEE) received the B.E. and M.E. degrees in electronic engineering from The University of Tokyo, Tokyo, Japan, in 1997 and 1999, respectively. In 1999, he joined Sumitomo Electric Industries (SEI) Ltd., Yokohama, Japan. He has been engaged in research and development on optical fibers and photonic sensors. He is also a Group Manager engaged in the research and development of new optical fibers. Mr. Hasegawa was the recipient of the 2002 Hasunuma Prize from the Society of Instrument and Control Engineers.

Tetsuya Hayashi (Senior Member, IEEE) received the B.E. and M.E. degrees in electronic engineering from The University of Tokyo, Tokyo, Japan, in 2004 and 2006, respectively, and the Ph.D. degree in engineering from Hokkaido University, Sapporo, Japan, in 2013. Since 2006, he has been with the Optical Communications Laboratory, Sumitomo Electric Industries Ltd., Yokohama, Japan, where he has been engaged in research and development on optical fibers and fiber optic technologies, and is currently the Group Leader for optical fiber research and development for space-division multiplexing. He has authored or coauthored more than 100 publications in international journals and conference proceedings, and three book chapters. Dr. Hayashi is also a Senior Member of Optica (formerly OSA) and the Institute of Electronics, Information and Communication Engineers (IEICE). He was the recipient of various awards and prizes, including the Tingye Li Innovation Prize from OSA at the Optical Fiber Communication Conference (OFC) 2017. He was on and chaired many conference subcommittees, including OFC and the OptoElectronics and Communications Conference (OECC). He is also an Associate Editor for the JOURNAL OF LIGHTWAVE TECHNOLOGY.

Yutaka Miyamoto (Member, IEEE) received the B.E. and M.E. degrees in electrical engineering from Waseda University, Tokyo, Japan, in 1986 and 1988, respectively, and the Dr. Eng. degree in electrical engineering from Tokyo University, Tokyo, in 2016. In 1988, he joined the NTT Transmission Systems Laboratories, Yokosuka, Japan, where he engaged in research and development on high-speed optical communications systems including the 10-Gbit/s first terrestrial optical transmission system (FA-10 G) using EDFA inline repeaters. He then joined NTT Electronics Technology Corporation between 1995 and 1997, where he engaged in the planning and product development of high-speed optical module at the data rate of 10 Gbps and beyond. Since 1997, he has been with NTT Network Innovation Labs, where he has contributed in the research and development of optical transport technologies based on 40/100/400-Gbit/s channel and beyond. He is now NTT Fellow and director of Innovative Photonic Network Research Center of NTT Network Innovation Laboratories, where he has been investigating and promoting the future scalable Optical Transport Network with the Pbit/s-class capacity based on innovative transport technologies, such as digital signal processing, space division multiplexing and cutting-edge integrated devices for photonic pre-processing. He is also a Fellow of IEICE.