photonics

MDPI

Communication

Efficient Frequency-Domain Block Equalization for Mode-Division Multiplexing Systems

Yifan Shen [1,2], Jianyong Zhang [1,2,*], Shuchao Mi [1,2], Guofang Fan [1,2] and Muguang Wang [1,2]

[1] Key Lab of All Optical Network & Advanced Telecommunication Network, Ministry of Education, Beijing Jiaotong University, Beijing 100044, China; 21120124@bjtu.edu.cn (Y.S.); 20111007@bjtu.edu.cn (S.M.); gffan@bjtu.edu.cn (G.F.); mgwang@bjtu.edu.cn (M.W.)
[2] Institute of Lightwave Technology, Beijing Jiaotong University, Beijing 100044, China
* Correspondence: jyzhang@bjtu.edu.cn

Abstract: In this paper, an adaptive frequency-domain block equalizer (FDBE) implementing the adaptive moment estimation (Adam) algorithm is proposed for mode-division multiplexing (MDM) optical fiber communication systems. By packing all frequency components into frequency-dependent blocks of a specified size B, we define an adaptive equalization matrix to simultaneously compensate for multiple frequency components at each block, which is computed iteratively using the Adam, recursive least squares (RLS) and least mean squares (LMS) algorithms. Simulations show that the proposed FDBE using the Adam algorithm outperforms those using the LMS and RLS algorithms in terms of adaptation speed and symbol error rate (SER) performance. The FDBE using the Adam algorithm with $B = 1$ has the fastest adaption time, requiring about $n_{tr} = 100$ and $n_{tr} = 900$ less training blocks than the RLS algorithm at the SER of 3.8×10^{-3} for the accumulated mode-dependent loss (MDL) of $\xi = 1$ dB and $\xi = 5$ dB, respectively. The Adam algorithm with $B = 16$ and $B = 8$ has 0.4 dB and 0.3 dB SNR better than the RLS algorithm with $B = 4$ for MDL and $\xi = 1$ dB and $\xi = 55$ dB, respectively.

Keywords: frequency-domain block equalization; adaptive moment estimation; mode division multiplexing

check for updates

Received: 15 December 2024
Revised: 8 February 2025
Accepted: 12 February 2025
Published: 17 February 2025

Citation: Shen, Y.; Zhang, J.; Mi, S.; Fan, G.; Wang, M. Efficient Frequency-Domain Block Equalization for Mode-Division Multiplexing Systems. *Photonics* **2025**, *12*, 161. https://doi.org/10.3390/photonics12020161

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The progressive increase in broadband applications, such as cloud computing and high-definition videos, has put forward higher requirements for the channel capacity and transmission performance of optical fiber communications [1]. The typical physical dimensions of light waves, including the amplitude, phase, polarization, and frequency, have been well used and exploited. Recently, mode division multiplexing (MDM) technology, which exploits physical spatial dimensions using few-mode fiber (FMF) or multi-mode fiber (MMF), has been proposed as a promising alternative for future optical communication systems due to its large bandwidth, high speed, and fast link deployment [2–5].

However, with the increase in multiplexed modes D and transmission lengths, mode coupling, mode-dependent loss (MDL), and the large group delay (GD) resulting from modal dispersion (MD) are unavoidable because the orthogonality between transmission modes is destroyed in the actual communication channel [6,7]. Hence, the crosstalk between all multiplexed modes is non-negligible, and full-scale multiple-input multiple-output (MIMO) equalization is essential to compensate for the crosstalk and dispersion at the receiver [8–10]. Moreover, the receiver computational complexity of the MIMO equalization

is also increased significantly with the increase in D. As shown in [11–13], MIMO frequency-domain equalization (FDE) can reduce complexity, enabling the computational complexity per symbol to scale with D and in the GD spread from MD.

To achieve fast adaptation and better performance, many adaptive FDEs have been proposed. These include adaptive FDEs based on the recursive least squares (RLS) algorithm, the least mean squares (LMS) algorithm, and the normalized least mean squares (NLMS) algorithm [14,15]. However, the traditional FDEs only compensate for a single frequency at each time in the receiver [16–18]. To achieve high accuracy and fast convergence, it is important to reduce the interference between different frequency components for FDEs.

In this paper, we propose an adaptive frequency-domain block equalizer (FDBE) using the Adam algorithm to compensate for multiple frequency components to achieve fast adaption and high performance in MDM systems. By packing the entire N_{FFT} frequency components into N_B frequency-dependent blocks within a specified size B, we can define an adaptive equalization matrix $\mathbf{W_B}$ to compensate for the mode-dependent effects at each frequency block, which can be computed iteratively using the Adam, RLS, and LMS algorithms. Simulation results indicate that the FDBE with the Adam adaptation algorithm has the fastest adaptation time and the best SER performance. This suggests that, in comparison to traditional FDEs, the proposed FDBE effectively mitigates the impact of MD and MDL, thereby enhancing the overall performance of the MDM system.

2. System Model

The complex baseband system for an MDM transmission system is shown in Figure 1. Let $x_d[n]$ be the symbol of the d-th mode at the n-th time interval, where $d = 1, ..., D$. The pulse shaping and electrical-to-optical conversion at the transmitter are given by $q(t)$. The channel response of the MDM system is given by $\mathbf{M_{tot}}(\mathrm{t})$, which can be described by a multi-section model, as given in [19]. The dominant noise source is $n_d[t]$, which is contributed by inline amplifiers. At the receiver, the optical-to-electrical conversion, along with the electrical filtering, is performed by $p(t)$. By sampling at N_{FFT} equally spaced frequencies $\Omega = 2\pi r_{os}(k - N_{\mathrm{FFT}}/2)/N_{\mathrm{FFT}}T_s$, where $k = 0, \ldots, N_{\mathrm{FFT}} - 1$, the frequency relationship of the MDM system can be described by the following:

$$\begin{aligned}\mathbf{Y}[k] = \begin{bmatrix} Y_1[k] \\ \vdots \\ Y_D[k] \end{bmatrix} &= \mathrm{P}(\Omega) \cdot \mathbf{M}_{\mathrm{tot}}\,(\Omega) \cdot \mathrm{Q}(\Omega) \cdot \begin{bmatrix} X_1[k] \\ \vdots \\ X_D[k] \end{bmatrix} + \begin{bmatrix} N_1[k] \\ \vdots \\ N_D[k] \end{bmatrix}, \\ &= \mathbf{H}[k] \cdot \mathbf{X}[k] + \mathbf{N}[k]\end{aligned} \tag{1}$$

where $\mathbf{Y}[k] = (Y_1[k], \ldots, Y_D[k])^{\mathrm{T}}$, $\mathbf{X}[k] = (X_1[k], \ldots, X_D[k])^{\mathrm{T}}$, and $\mathbf{N}[k] = (N_1[k], \ldots, N_D[k])^{\mathrm{T}}$ are N_{FFT}-point fast Fourier transform (FFT) blocks of the time-domain signals, and $\mathrm{Q}(\Omega)$, $\mathbf{M_{tot}}(\Omega)$, and $\mathrm{P}(\Omega)$ are the N_{FFT}-point Fourier transforms of $q(t)$, $\mathbf{M_{tot}}(\mathrm{t})$ and $p(t)$, respectively. The overall impulse response of the MDM system is $\mathbf{H}[k] = \mathrm{P}(\Omega) \cdot \mathbf{M_{tot}}(\Omega) \cdot \mathrm{Q}(\Omega)$. The output $\mathbf{Y}[k]$ represents the outputs of the CD equalizers [20].

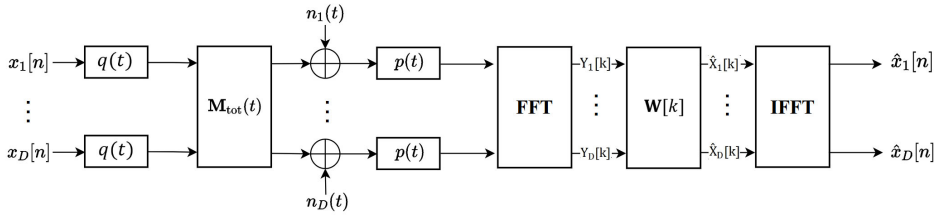

Figure 1. Complex baseband model of MDM transmission with MIMO equalizer.

The channel response of the MDM system composed of S_{sec} sections is given by $\mathbf{M_{tot}}(\Omega)$, which can be described by the multi-section model given in (2).

$$\mathbf{M}_{tot}(\Omega) = \prod_{s=1}^{S_{\text{sec}}} \mathbf{M}^{(s)}(\Omega). \tag{2}$$

The transmission response of each independent segment $\mathbf{M}^{(s)}(\Omega)$ can be expressed as follows:

$$\begin{aligned} \mathbf{M}^{(s)}(\Omega) &= \mathbf{V}^{(s)}\Lambda^{(s)}(\Omega)\mathbf{U}(s)* \\ &= \mathbf{V}^{(s)} \times \text{diag}\left[\exp\left(g_1^{(s)}/2 - j\Omega\tau_1^{(s)}\right), \ldots, \exp\left(g_D^{(s)}/2 - j\Omega\tau_D^{(s)}\right)\right] \times \mathbf{U}(s)*, \end{aligned} \tag{3}$$

where $(\cdot)^*$ represents the matrix Hermitian conjugate transpose; $\mathbf{V}^{(s)}$ and $\mathbf{U}^{(s)}$represent the frequency-independent random unitary matrix, which represents the mode coupling at the input and output of the s-th independent segment, respectively [21]. $\mathbf{\Lambda}^{(s)}(\Omega)$ represents a diagonal matrix, the diagonal elements of which represent various types of damage suffered by each mode, including both the MD and MDL of the s-th independent section [22]. $g_i^{(s)}$ is the uncoupled mode gain of the s-th section for the i-th mode and satisfies $g_1^{(s)} + \ldots + g_D^{(s)} = 0$ with the root mean square (rms) value σ_g. $\tau_i^{(s)}$is the uncoupled mode group delay of the s-th section for the i-th mode and satisfies $\tau_1^{(s)} + \ldots + \tau_D^{(s)} = 0$. Because the dispersion effect is compensated for by the receiver, $\tau_i^{(s)}$ is set to zero in this work. In the case of $S_{\text{sec}} \gg 1$, the accumulated MDL in the strong coupling fiber link is $\xi = \sqrt{S_{\text{sec}}}\sigma_g$ [23].

3. Frequency-Domain Block Equalization

3.1. Frequency-Domain Equalization

A traditional approach of FDE is using an adaptive $D \times D$ matrix $\mathbf{W}[k]$ was used to compensate for the mode-dependent effects at each frequency k [24]. The equalizer output is given by the following:

$$\widetilde{\mathbf{X}}[k] = \begin{bmatrix} \widetilde{X}_1[k] \\ \vdots \\ \widetilde{X}_D[k] \end{bmatrix} = \mathbf{W}[k] \cdot \mathbf{Y}[k], \tag{4}$$

where $\widetilde{\mathbf{X}}[k] = \left(\widetilde{X}_1[k], \ldots, \widetilde{X}_D[k]\right)^{\text{T}}$ is the estimated signal after equalization.

3.2. Frequency Domain Block Equalization

Due to the influence of MD and MDL, the traditional signal interferes with each other between different frequency components [25,26]. Thus, it is insufficient for the equalizer to compensate for a single frequency at each time. Therefore, we considered a frequency-domain block equalization (FDBE) strategy to compensate for multiple frequency components to further improve performance and eliminate interference.

Firstly, we packed the entire $k = 0, \ldots, N_{\text{FFT}} - 1$ frequency components into N_B frequency-dependent blocks within a specified size, $B = \lfloor N_{\text{FFT}}/N_B \rfloor$, where $\lfloor \cdot \rfloor$ represents the operation of the floor function. The packed input is given as follows:

$$\mathbf{Z}[b] = \begin{bmatrix} \mathbf{X}[b \cdot B + 0] \\ \mathbf{X}[b \cdot B + 1] \\ \vdots \\ \mathbf{X}[b \cdot B + (B-1)] \end{bmatrix}, b = 0, \ldots, N_B - 1. \tag{5}$$

The packed output is given as follows:

$$\mathbf{G}[b] = \begin{bmatrix} \mathbf{Y}[b \cdot B + 0] \\ \mathbf{Y}[b \cdot B + 1] \\ \vdots \\ \mathbf{Y}[b \cdot B + (B-1)] \end{bmatrix}, \cdot b = 0, \ldots, N_B - 1. \quad (6)$$

We defined an adaptive equalization matrix $\mathbf{W_B}[b]$ of size $(B \times D) \times (B \times D)$ to implement FDBE. In the system equalization, the output for the b-th block is given in Algorithm 1.

$$\widetilde{\mathbf{Z}}[b] = \mathbf{W_B}[b] \cdot \mathbf{G}[b], \quad (7)$$

where $\widetilde{\mathbf{Z}}[b]$ is the estimated signal after equalization.

After performing FDBE on the N_B blocks, we performed traditional FDE on the remaining $N_{\mathrm{FFT}} - (B \times N_{\mathrm{B}})$ frequency components. The algorithm is given by the following:

Algorithm 1: Frequency-Domain Block Equalization

Input : $\mathbf{X}, \mathbf{Y}, \mathbf{W}$ and $\mathbf{W}_{\mathrm{B}}$;
For all $b = 0, ..., N_B - 1$**do**
  $\mathbf{Z}[b] \leftarrow$ pack $\mathbf{X}[b{\cdot}B + 0], ..., \mathbf{X}[b{\cdot}B + (B-1)]$ according to Equation (5);
  $\mathbf{G}[b] \leftarrow$ pack $\mathbf{Y}[b{\cdot}B + 0], ..., \mathbf{Y}[b{\cdot}B + (B-1)]$ according to Equation(6);
  $\widetilde{\mathbf{Z}}[b] = \mathbf{W}_{\mathrm{B}}[b]{\cdot}\mathbf{G}[b]$;
  $\widetilde{\mathbf{X}}[\mathrm{k}] \leftarrow$ unpack $\widetilde{\mathbf{Z}}[b]$;
End for
For all $k = (B \times N_B), \ldots, N_{\mathrm{FFT}} - 1$ **do**
  $\widetilde{\mathbf{X}}[k] = \mathbf{W}[k]{\cdot}\mathbf{Y}[k]$;
End for
Return $\widetilde{\mathbf{X}}$

4. Adaptive Frequency-Domain Block Equalization

At each frequency block b, the error signal is given by the following:

$$\mathbf{e}[b] = \tilde{\mathbf{Z}}[b] - \mathbf{W_B}[b]\mathbf{G}[b]. \quad (8)$$

The total frequency domain error is given by the following:

$$\mathbf{e}_{\mathrm{tot}} = \sum_{b=0}^{N_{\mathrm{B}}-1} E\{\mathbf{e}[b] * \mathbf{e}[b]\} + \sum_{k=B\times N_{\mathrm{B}}}^{N_{\mathrm{FFT}}-1} E\{\mathbf{e}[k] * \mathbf{e}[k]\}, \quad (9)$$

where $\mathbf{e}[k] = \widetilde{\mathbf{X}}[k] - \mathbf{W}[k]\mathbf{Y}[k]$ [27].

Therefore, minimizing $\mathbf{e_{tot}}$ is equivalent to separately minimizing the error terms at each frequency k or each frequency block b, as the equalization matrix $\mathbf{W_B}[b]$ or $\mathbf{W}[k]$ can be optimized independently.

For MDM systems, the channel state information matrix $\mathbf{H}[k]$ is unknown at the equalizer, and the $\mathbf{W_B}[b]$ or $\mathbf{W}[k]$ must be computed iteratively to update. Furthermore, if $B = 1$, $\mathbf{W_B}[b]$ is reduced to $\mathbf{W}[k]$. Therefore, we only focus on the updating of $\mathbf{W_B}[b]$ for simplicity. The traditional approach is the LMS and RLS algorithms.

4.1. LMS

LMS is a stochastic gradient descent minimization using instantaneous estimates of the error $\mathbf{e}[b]$ [28] and is described by the updated equation below:

$$\mathbf{W_B}[b] \leftarrow \mathbf{W_B}[b] + \left(\tilde{\mathbf{Z}}[b] - \mathbf{W_B}[b]\mathbf{G}[b]\right)\mathbf{G}[b] * \mu, \tag{10}$$

where μ is the step size and $0 < \mu < 2/\lambda_{max}$. The λ_{max} is the largest eigenvalue of the autocorrelation matrix of $\mathbf{G}[b]$.

4.2. RLS

RLS involves an iterative minimization of an exponentially weighted cost function, treating the minimization problem as deterministic [29]. The updated equation for $\mathbf{W_B}[b]$ is given by the following:

$$\mathbf{W_B}[b] \leftarrow \mathbf{W_B}[b] + \left(\tilde{\mathbf{Z}}[b] - \mathbf{W_B}[b]\mathbf{G}[b]\right)\mathbf{G}[b] * \left(\mathbf{R}[b]\kappa^{-1}\right), \tag{11}$$

where $\mathbf{R}[b]$ is given by the following:

$$\mathbf{R}[b] \leftarrow \left(\mathbf{R}[b]\kappa^{-1}\right) - \frac{\left(\mathbf{R}[b]\kappa^{-1}\right) - \left(\mathbf{R}[b]\kappa^{-1}\right)\mathbf{G}[b]\mathbf{G}[b] * \left(\mathbf{R}[b]\kappa^{-1}\right)}{1 + \mathbf{G}[b] * \left(\mathbf{R}[b]\kappa^{-1}\right)\mathbf{G}[b]}. \tag{12}$$

$\mathbf{R}[b]$ is initialized with an identity matrix multiplied by the forgetting factor κ, where $0 \ll \kappa < 1$.

4.3. ADAM

Adam is an algorithm for the first-order gradient-based optimization of stochastic objective functions. Based on adaptive estimates of lower-order moments, it is a popular stochastic optimization method in deep learning, which is straightforward to implement, computationally efficient, and has minimal memory requirements [30].

The i-th update algorithm $F(\cdot)$ of $\mathbf{W_B}[b]$ with the Adam algorithm is given by Algorithm 2. $Re[\cdot]$ and $Im[\cdot]$ are the real and imaginary parts for a complex number, respectively. α is the learning rate for the Adam function. The hyper-parameters $\beta_1, \beta_2 \in [0, 1)$ control the exponential decay rates of moving averages of the gradient $\mathbf{m}$ and the squared gradient $\mathbf{v}$. $\cdot^i$ is the current number of iterations. ε is a constant close to zero to generate a valid initial step size $\alpha \cdot \hat{\mathbf{m}} / \left(\sqrt{\hat{\mathbf{v}}} + \varepsilon\right)$.

Algorithm 2: $F(\cdot)$

Input : $\mathbf{W_B}[b], \widetilde{\mathbf{Z}}[b], \mathbf{G}[b]$;
$\mathbf{g} = -\left(\widetilde{\mathbf{Z}}[b] - \mathbf{W_B}[b]\mathbf{G}[b]\right)\mathbf{G}[b]^*$;
$\mathbf{g_{re}} \leftarrow Re[\mathbf{g}]$, $\mathbf{g_{im}} \leftarrow Im[\mathbf{g}]$;
$\Delta\mathbf{W} = Adam(\mathbf{g_{re}}) + j \cdot Adam(\mathbf{g_{im}})$;
Return $\mathbf{W}_B[b] - \Delta\mathbf{W}$
Function $Adam(\cdot)$
Input : $\mathbf{g}$;
$\mathbf{m} = \beta_1 \cdot \mathbf{m} + (1 - \beta_1) \cdot \mathbf{g}$;
$\mathbf{v} = \beta_2 \cdot \mathbf{v} + (1 - \beta_2) \cdot \mathbf{g}^2$;
$\hat{\mathbf{m}} = \mathbf{m} / (1 - \beta_1^i)$;
$\hat{\mathbf{v}} = \mathbf{v} / (1 - \beta_2^i)$;
Return $\alpha \cdot \hat{\mathbf{m}} / \left(\sqrt{\hat{\mathbf{v}}} + \varepsilon\right)$

4.4. Algorithm Complexity

In this section, the computational complexity of the different algorithms mentioned above is analyzed and calculated by the number of real multiplications and real additions, as shown in Table 1.

Table 1. The computational complexity of different algorithms.

Operation	Number of Real Multiplications	Number of Real Additions
Adaption with LMS (10)	$N_{\mathrm{FFT}}\left(4B^2D^2+4BD\right)$	$N_{\mathrm{FFT}}\left(4B^2D^2+4BD\right)$
Adaption with RLS (11)	$N_{\mathrm{FFT}}\left(20B^2D^2+8BD\right)$	$N_{\mathrm{FFT}}\left(18B^2D^2+4BD\right)$
Adaption with Adam	$N_{\mathrm{FFT}}\left(12B^2D^2+14BD\right)$	$N_{\mathrm{FFT}}\left(6B^2D^2+BD\right)$

The complexity of FDBE using different adaptation algorithms depends mainly on the number of modes D, the size B, and the frequency block length N_{FFT}. For the LMS algorithm, the number of multiplications and additions is $N_{\mathrm{FFT}}\left(4B^2D^2+4BD\right)$ and $N_{\mathrm{FFT}}\left(4B^2D^2+4BD\right)$, respectively, and is proportional to B^2D^2+BD. In summary, the complexity of the RLS algorithm is the highest, and the LMS algorithm is the lowest.

5. Simulation Results

5.1. Transmission System

In this section, a long-haul MDM transmission system is studied, which is described by the system model of Section 2. The system has a total of 2000 sections, each of length L_{sec} = 1km, for a total length of 2000 km. We selected the 12-mode FMFs with low uncoupled GD spread, relying on strong mode coupling [31]. The statistics of the coupled group delay were determined by $\Delta\beta_{1,rms}L_{sec}$, where $\Delta\beta_{1,rms}$ =29 ps/km is the root mean square (rms) uncoupled modal dispersion. The total frequency block length N_{FFT} was 2^{11} to ensure fast adaptation.

At the transmitter, a symbol rate $R_s = 1/T_s$ =32 Gbaud was assumed. The pulse shaping $q(t)$ is a rectangular pulse of duration T_s filtered by a fifth-order Bessel low-pass filter with −3 dB cutoff frequency $0.8/T_s$ [20]. We assumed a quadrature phase-shift keying (QPSK) constellation with average power P_x to obtain the transmitted symbols. At the receiver, the receiver oversampling rate was r_{os} = 2. The SNR per mode is given by SNR = P_x/δ_N^2, where the received noise per mode is calculated by δ_N^2 = SNR/P_x= $N_0R_S\gamma_{os}$. The receiver filter $p(t)$ is a fifth-order Butterworth low-pass filter with −3 dB cutoff frequency r_{os}/T_s [20].

5.2. Adaption Time

We study the adaptation time of FDBE using the LMS, RLS, and Adam algorithms under different SNR and MDL. The parameters of different algorithms were initially optimized using training sequences in order to improve the performance of the algorithms. For the LMS algorithm, we chose the step size $\mu = 1.5\times10^{-5}$ for ξ = 1 dB and $\mu = 7.2\times10^{-6}$ for ξ = 5 dB. For the RLS algorithm, the forgetting factor κ = 0.999 is defined. For the Adam algorithm, we define the step size $\alpha = 0.1$, $\beta_1 = 0.99999$, $\beta_2 = 0.99$ for ξ = 1 dB, and $\beta_2 = 0.999$ for ξ = 5 dB, respectively, $\varepsilon = 10^{-8}$. Meanwhile, the adaptation time in Table 2 represents an idealized simulation scenario, with actual hardware speed potentially varying based on implementation details such as the number of parallel samples and clock speed.

Table 2. Adaption time of FDBE using the LMS, RLS, and Adam algorithms under different SNR and MDL.

Adaptive Algorithm	Number of Block Size B	$\mathbf{n_{tr}}$	
		SNR=10.5 dB, ξ = 1 dB, SER = 3.8 $\times$ 10^{-3}	**SNR=14 dB, ξ = 5 dB, SER = 3.8 $\times$ 10^{-3}**
LMS	1	24,700	-
	4	25,600	-
	8	27,200	-
	16	29,800	-
RLS	1	2600	4800
	4	2650	4950
	8	2700	5200
	16	3050	9500
Adam	1	2500	3900
	4	2600	4100
	8	2650	4900
	16	4850	10,700

Table 2 shows the number of training blocks n_{tr} of the LMS, RLS, and Adam algorithms with $B = 1, 4, 8$, and 16 at the SER of 3.8×10^{-3} for $\xi = 1$ dB and $\xi = 5$ dB, respectively. It can be seen in Table 2 that the Adam algorithm with $B = 1$ has the fastest adaption time, where $n_{tr} = 2500$ and $n_{tr} = 3900$ for $\xi = 1$ dB and $\xi = 5$ dB, respectively. The LMS algorithm has the worst adaptation performance in all cases and cannot meet the convergence of the SER of 3.8×10^{-3} for $\xi = 5$ dB. Furthermore, the number of training blocks increases with B and ξ as shown in Table 2. If $B \leq 8$, the number of training blocks increases a little bit higher for the RLS and Adam algorithms as B increases.

5.3. SER Performance

In this section, the SER performance of FDBE using different algorithms under different SNR and MDL is studied.

Figure 2 shows the SER versus the number of training blocks n_{tr} achieved by the LMS, RLS, and Adam algorithms with $B = 1, 4, 8$, and 16 at SNR= 10.5 dB for $\xi = 1$ dB. As B increases, the SER performance of the Adam algorithm increases. The SER performance of the RLS algorithm increases when $B \leq 4$ and decreases significantly when $4 < B \leq 16$. It is shown in Figure 2 that the Adam algorithm with $B = 16$ has the best SER performance among all the algorithms, where the SER at $n_{tr} = 10^5$ is about 2.1×10^{-5}. The RLS algorithm with $B = 4$ has the best SER performance for $B = 1, 4, 8$, and 16, and the SER at $n_{tr} = 10^5$ is about 3.1×10^{-4}.

Figure 3 presents the SER versus the number of training blocks n_{tr} achieved by the LMS, RLS, and Adam algorithms with $B = 1, 4, 8$, and 16 at SNR= 14 dB for $\xi = 5$ dB. With the increase of B, the SER performance of the Adam algorithm increases when $B \leq 8$ and decreases when $8 < B \leq 16$. Figure 3 shows that the Adam algorithm with $B = 8$ has the best SER performance among all the algorithms, where the SER at $n_{tr} = 10^5$ is about 3.7×10^{-4}. The RLS algorithm with $B = 4$ has the best SER performance for all Bs, and the SER at $n_{tr} = 10^5$ is about 4.7×10^{-4}.

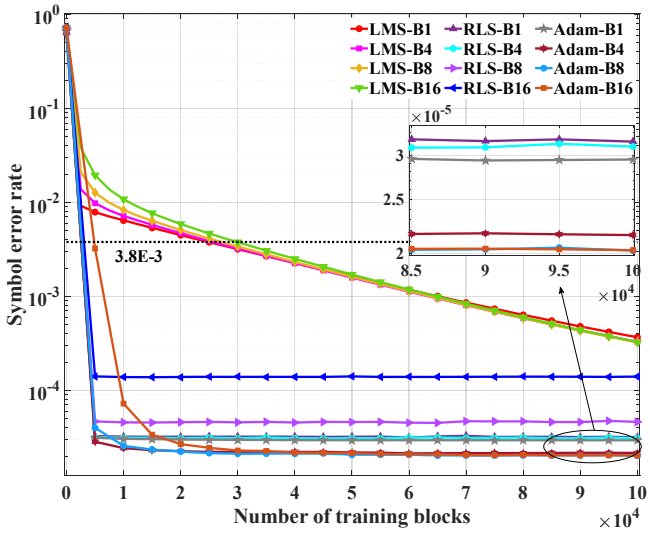

Figure 2. Average symbol–error ratio versus number of training blocks n_{tr} using LMS, RLS, and Adam algorithms with $B = 1, 4, 8$, and 16, assuming $\alpha = 0.1$, $\beta_1 = 0.99999$, $\beta_2 = 0.99$, $\varepsilon = 10^{-8}$, $\kappa = 0.999$, $\mu = 1.5 \times 10^{-5}$, $\xi = 1$ dB, and $D = 12$.

Moreover, the SER performance of the LMS algorithm is almost unchanged when B increases and is much worse than the Adam and RLS algorithms, as shown in Figures 2 and 3.

Figure 4 shows the SER versus SNR in dB achieved by the LMS, RLS, and Adam algorithms at $n_{tr} = 10^5$ for MDL $\xi = 1$ dB and $\xi = 5$ dB. We chose $B = 16$ and $B = 8$ for the Adam algorithm for $\xi = 1$ dB and $\xi = 5$ dB, while $B = 4$ was chosen for the RLS algorithm and $B = 16$ for the LMS algorithm. As expected, the SER performance of the Adam algorithm was always the best among all the algorithms. For $\xi = 1$ dB, the Adam algorithm requires SNR $= 6.8$ dB at the SER of 3.8×10^{-3}, which outperforms the LMS algorithm by about 1.1 dB and outperforms the RLS algorithm by 0.4 dB. For $\xi = 5$ dB, the SNR required for the Adam algorithm at the SER $= 3.8 \times 10^{-3}$ is 11.8 dB, which is about 0.3 dB better than the RLS algorithm. The SER performance of the LMS algorithm is the worst among all the algorithms.

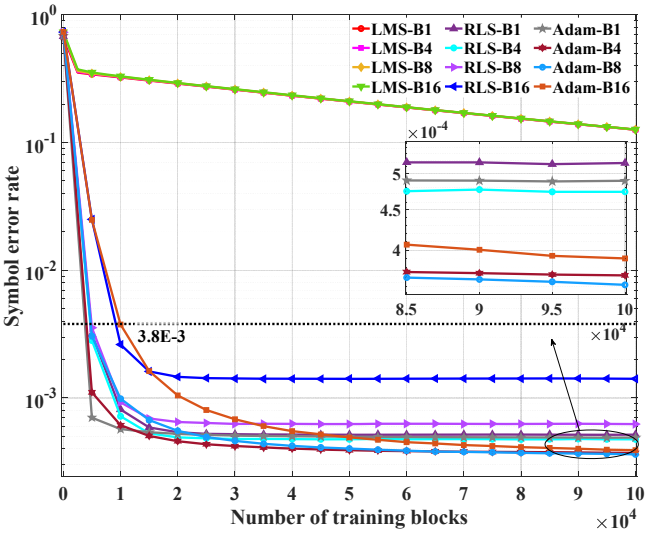

Figure 3. Average symbol–error ratio versus number of training blocks n_{tr} using LMS, RLS, and Adam algorithms with $B = 1$, 4, 8, and 16, assuming $\alpha = 0.1$, $\beta_1 = 0.99999$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$, $\kappa = 0.999$, $\mu = 7.2 \times 10^{-6}$, $\xi = 5$ dB, and $D = 12$.

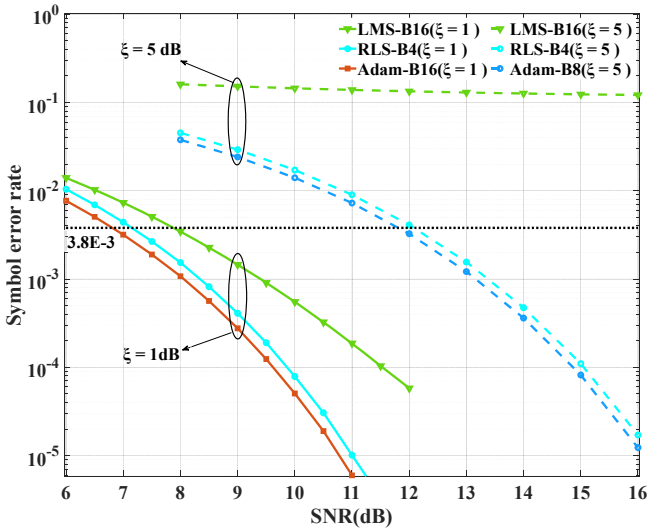

Figure 4. Average symbol–error ratio versus SNR in dB using LMS, RLS, and Adam algorithms, assuming $\alpha = 0.1$, $\beta_1 = 0.99999$, $\beta_2 = 0.99$, $\varepsilon = 10^{-8}$, $\kappa = 0.999$, $\mu = 1.5 \times 10^{-5}$, $\xi = 1$ dB and $D = 12$ (solid line), and $\alpha = 0.1$, $\beta_1 = 0.99999$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$, $\kappa = 0.999$, $\mu = 7.2 \times 10^{-6}$, $\xi = 5$ dB, and $D = 12$ (dotted line).

6. Conclusions

In this work, we propose an adaptive frequency-domain block equalizer (FDBE) based on the Adam algorithm for an MDL-impaired MDM system. The proposed FDBE packs the received signals into frequency blocks and uses the Adam algorithm to update the adaptive equalization matrix to improve the adaptation speed and the SER performance. The simulation results suggest that the Adam algorithm with $B = 16$ and $B = 8$ has the best SER performance for $\xi = 1$ dB and $\xi = 5$ dB, respectively. Furthermore, the SER performance of FDBE increases with the block size. Thus, the proposed adaptive FDBE using the Adam algorithm has the potential application to mitigate the effect of MD and MDL in MDM transmission with fast convergence and high performance.

Author Contributions: Conceptualization, Y.S. and J.Z.; methodology, Y.S. and J.Z.; software, Y.S.; validation, Y.S.; formal analysis, Y.S.; investigation, Y.S.; resources, J.Z.; data curation, S.M.; writing—original draft preparation, Y.S.; writing—review and editing, Y.S., J.Z., S.M., G.F. and M.W.; visualization, Y.S.; supervision, J.Z.; project administration, J.Z.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Fundamental Research Funds for the Central Universities (Grant No.2023JBZY014); the National Key R&D Program of China (Grant No.2021YFF0600905); the Science Foundation of The Chinese Academy of Sciences (Grant No. W24L00380); and the National Natural Science Foundation of China (Grant No.U2006217,62371035).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The 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.

Conflicts of Interest: The authors declare that there are no conflicts of interest.

References

1. Richardson, D.J.; Fini, J.M.; Nelson, L.E. Space-division multiplexing in optical fibres. *Nat. Photonics* **2013**, *7*, 354–362. [CrossRef]
2. Winzer, P.J. Scaling Optical Fiber Networks: Challenges and Solutions. *Opt. Photon News* **2015**, *26*, 28–35. [CrossRef]

3. Winzer, P.J.; Neilson, D.T.; Chraplyvy, A.R. Fiber-optic transmission and networking: The previous 20 and the next 20 years. *Opt. Express* **2018**, *26*, 24190–24239. [CrossRef]
4. Jinno, M. Spatial Channel Network (SCN): Opportunities and Challenges of Introducing Spatial Bypass Toward the Massive SDM Era. *J. Opt. Commun. Netw.* **2019**, *11*, 1–14. [CrossRef]
5. Torres, L.M.; Cañete, F.J.; Díez, L. BER analysis of an optimum MIMO linear receiver in optical SDM systems with mode-dependent loss. *Opt. Express* **2023**, *31*, 31671–31686. [CrossRef] [PubMed]
6. Noboru, Y.; Tsuritani, T. Recent progress in space-division multiplexing optical network technology. In Proceedings of the 2020 International Conference on Optical Network Design and Modeling (ONDM), Barcelona, Spain, 18–21 May 2020; IEEE: New York, NY, USA, 2020.
7. Awaji, Y. Review of Space-Division Multiplexing Technologies in Optical Communications. *IEICE Trans. Commun.* **2019**, *E102.B*, 1–16. [CrossRef]
8. Weng, Y.; Wang, J.; Pan, Z. Recent advances in DSP techniques for mode division multiplexing optical networks with MIMO equalization: A review. *Appl. Sci.* **2019**, *9*, 1178. [CrossRef]
9. Arik, S.O.; Kahn, J.M.; Ho, K.-P. MIMO Signal Processing for Mode-Division Multiplexing: An overview of channel models and signal processing architectures. *IEEE Signal Process. Mag.* **2014**, *31*, 25–34. [CrossRef]
10. Arik, S.O.; Ho, K.-P.; Kahn, J.M. Group delay statistics and management in mode-division multiplexing. In Proceedings of the 2015 49th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 8–11 November 2015; pp. 991–998.
11. Choutagunta, K.; Arik, S.O.; Ho, K.-P.; Kahn, J.M. Characterizing Mode-Dependent Loss and Gain in Multimode Components. *J. Light. Technol.* **2018**, *36*, 3815–3823. [CrossRef]
12. Huang, Y.; Ji, F.; Wei, Z.; Wen, M.; Chen, X.; Tang, Y.; Guo, W. Frequency Domain Analysis and Equalization for Molecular Communication. *IEEE Trans. Signal Process.* **2021**, *69*, 1952–1967. [CrossRef]
13. Shibahara, K.; Kobayashi, T.; Miyamoto, Y. Out-of-band exclusive frequency-domain MIMO equalization with the reduced-overlap-and-save method for scalable mode multiplexed signal transmission. *Opt. Express* **2023**, *31*, 2302–2315. [CrossRef]
14. Ghauri, S.A.; Sohail, M.F. System identification using LMS, NLMS and RLS. In Proceedings of the 2013 IEEE Student Conference on Research and Development, Putrajaya, Malaysia, 16–17 December 2013; IEEE: New York, NY, USA, 2013; pp. 65–69.
15. Shibahara, K.; Hoshi, M.; Hasegawa, T.; Hayashi, T.; Miyamoto, Y. Recursive Least Squares Based Low-Complexity Frequency-Domain MIMO Equalization for MDL-Tolerant Long-Haul Space Division Multiplexing Transmission. *J. Light. Technol.* **2024**, *42*, 4324–4337. [CrossRef]
16. Arik, S.Ö.; Askarov, D.; Kahn, J.M. Effect of Mode Coupling on Signal Processing Complexity in Mode-Division Multiplexing. *J. Light. Technol.* **2012**, *31*, 423–431. [CrossRef]
17. Bai, N.; Li, G. Adaptive Frequency-Domain Equalization for Mode-Division Multiplexed Transmission. *IEEE Photon Technol. Lett.* **2012**, *24*, 1918–1921. [CrossRef]
18. Iqbal, N.; Al-Dhahir, N.; Zerguine, A.; Zidouri, A. Adaptive Frequency-Domain RLS DFE for Uplink MIMO SC-FDMA. *IEEE Trans. Veh. Technol.* **2014**, *64*, 2819–2833.
19. Rottenberg, F.; Mestre, X.; Horlin, F.; Louveaux, J. Efficient Equalization of Time-Varying Channels in MIMO OFDM Systems. *IEEE Trans. Signal Process.* **2019**, *67*, 5583–5595. [CrossRef]
20. Sercan, Ö.A.; Askarov, D.; Kahn, J.M. Adaptive frequency-domain equalization in mode-division multi-plexing systems. *J. Light. Technol.* **2014**, *32*, 1841–1852.
21. Ho, K.-P.; Kahn, J.M. Linear Propagation Effects in Mode-Division Multiplexing Systems. *J. Light. Technol.* **2013**, *32*, 614–628. [CrossRef]
22. Mi, S.; Zhang, J. A low complexity large-scale MIMO detection algorithm for MDL-impaired fiber mode division multiplexing transmission. In Proceedings of the 13th International Conference on Information Optics and Photonics (CIOP 2022), Xi'an, China, 7–10 August 2022; pp. 119–126.
23. Ho, K.-P.; Kahn, J.M.; Kaminow, I.P.; Li, T.; Willner, A. EMode coupling and its impact on spatially multiplexed systems. *Opt. Fiber Telecommun. VI* **2013**, *17*, 1386–1392.
24. Shibahara, K.; Hoshi, M.; Miyamoto, Y. 10-Spatial-Mode 1300-Km Transmission Over 6-LP Graded Index Few-Mode Fiber with 36-Ns Modal Dispersion. *J. Light. Technol.* **2023**, *42*, 1257–1264. [CrossRef]
25. Yi, W.; Sillekens, E.; Lavery, D.; Dzieciol, H.; Zhou, S.; Law, K.; Chen, J.; Bayvel, P.; Killey, R.I. Performance of momentum-based frequency-domain MIMO equalizer in the presence of feedback delay. *Opt. Express* **2020**, *28*, 19133–19143. [CrossRef]
26. Yang, Z.; Yu, W.; Peng, G.; Liu, Y.; Zhang, L. Recent Progress on Novel DSP Techniques for Mode Division Multiplexing Systems: A Review. *Appl. Sci.* **2021**, *11*, 1363. [CrossRef]
27. Mello, D.A.A.; Srinivas, H.; Choutagunta, K.; Kahn, J.M. Impact of Polarization- and Mode-Dependent Gain on the Capacity of Ultra-Long-Haul Systems. *J. Light. Technol.* **2019**, *38*, 303–318. [CrossRef]
28. Benesty, J.; Paleologu, C.; Gänsler, T.; Ciochină, S. *A Perspective on Stereophonic Acoustic Echo Cancellation*; Springer Nature: Dordrecht, The Netherlands, 2011.

29. Paulo, D.S.R.; Diniz, P.S.R. The least-mean-square (LMS) algorithm. In *Adaptive Filtering: Algorithms and Practical Implementation*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 61–102.
30. Diederik, K.; Jimmy, P.; Adam, B. A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
31. Ip, E.; Kahn, J.M. Digital Equalization of Chromatic Dispersion and Polarization Mode Dispersion. *J. Light. Technol.* **2007**, *25*, 2033–2043. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.