

Sparse Adaptive Frequency Domain Equalizers for Mode-Group Division Multiplexing

Kai Shi, *Member, IEEE*, and Benn C. Thomsen, *Member, IEEE*

Abstract—In this paper, we develop low complexity sparse frequency domain equalizers (FDEs) that exploit the sparsity that we observe in the graded index multimode fiber (GI-MMF) multiple-input multiple-output (MIMO) channel. The sparse channel impulse response is caused by the strong crosstalk at the mode MUX/DEMUX and weak coupling in the fiber between different mode groups (MGs). Two sparse FDE designs are proposed in order to compensate the crosstalk with relatively low computational complexity. The first method is based on a priori knowledge of the channel impulse response, which is used to generate a mask of taps with significant magnitudes. The second method is based on the improved proportionate normalized least-mean-square (IPNLMS) algorithm, where the active and non-active taps are adjusted at different rates of convergence. The computational complexity and the system performance of the proposed algorithms are analyzed. It is shown that the sparse FDEs offer low complexity relative to the sparse equalizers that use delay buffers, while maintaining improved performance over non-sparse equalizers in the presence of noise.

Index Terms—Sparse adaptive equalizer, mode-group multiplexing, active taps mask.

I. INTRODUCTION

IN order to meet the annual growth in traffic rate, space-division multiplexing (SDM) has been widely studied recently as the only known option to scale the network capacity in a single fiber [1]. Multiplexing in both multicore [2] and multimode fiber [3], [4] has been demonstrated, where the coupling between the channels is cancelled by using MIMO digital signal processing (DSP) [5]. Degenerate mode-group division multiplexing (MGDM) using conventional GI-MMF has been shown to offer relatively low coupling in the fiber due to the large differential group delay (DGD) between individual mode groups. MIMO DSP can be eliminated if the channels are selectively launched into different mode groups by using spatial light modulators (SLMs) [6], which can significantly reduce the DSP complexity [7]. However, the crosstalk at the mode MUX/DEMUX of MGDM systems still exists due to the impairments of the SLM based optics system. Previously, we have shown that the channel impulse response of degenerated MGDM systems using GI-MMF is sparse as a result of the strong crosstalk at the mode MUX/DEMUX and weak coupling in the fiber [8]. Consequently, the total length of the channel impulse response increases as the length or the DGD of the fiber increases.

K. Shi and B. C. Thomsen are with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, U.K. (e-mail: k.shi@ucl.ac.uk; b.thomsen@ucl.ac.uk).

Copyright (c) 2013 IEEE. Personal use of this material is permitted.

Manuscript received April 19, 2005; revised December 27, 2012. This work was supported by the U.K. Engineering and Physical Sciences Research Council under Grant EP/J008842/1.

Thus, for a conventional uniformly spaced tap-delay-line equalizer, at least the same number of taps as the total number of channel coefficients in the sampled impulse response are required to achieve reasonable performance. Hence, the computational complexity increases significantly for longer transmission distances resulting from the large value of the DGD of the GI-MMF. However, due to the sparsity of the channel, only a small portion of the tap coefficients are non-zero or with significant magnitudes, while the rest are zero or small. Various algorithms have been developed for sparse channels aiming for low-complexity and cost, as well as noise reduction [9]. A priori information based optimization using the known channel impulse response can be used to determine the optimal or sub-optimal selection of the active tap positions [10], or an adaptive approach can be taken using the IPNLMS algorithm. In this algorithm, different convergence rates are applied on the active and inactive taps during the update of the equalizer coefficients [11]. The sparse equalizer design was first introduced for fiber SDM systems by using a nonuniformly spaced tap-delay-line equalizer, where the taps with small magnitudes were replaced by delay buffers [12]. Based on this method, we proposed the delayed FDE which reduces the computational complexity compared with the delayed time domain equalizer (TDE) when the tap length exceeds a certain number [8]. In both cases, a priori knowledge of the channel is required, where the delays and number of taps are manually selected by observing the channel impulse response. Therefore, it is important to design a sparse equalizer where the active taps can be selected either automatically based on the training sequence or adaptively. Furthermore, the complexity of the delayed FDE is still found to be approximately twice higher than the conventional FDE. Hence, it is desirable to develop sparse FDEs that fully exploit the complexity reduction offered by the frequency domain implementation.

In this paper, we propose two sparse FDEs with low complexity for the degenerated MGDM systems. The first method is the conventional FDE with an active taps mask, which is obtained from the least squares (LS) tap coefficients combined with a simple active tap selection method. The second method is a sparse adaptive equalizer using the IPNLMS algorithm. The performance of these two FDEs are compared, along with the delayed TDE and FDE in a 2×2 MGDM system using a 28 Gbaud quadrature phase shift keying (QPSK) signal. This paper is organized as follows: In section II, the channel impulse response of the MGDM system is presented and the complexity of different equalizers for the system is examined. This is followed by the algorithms of the two proposed sparse FDEs in section III. The system performance of different types

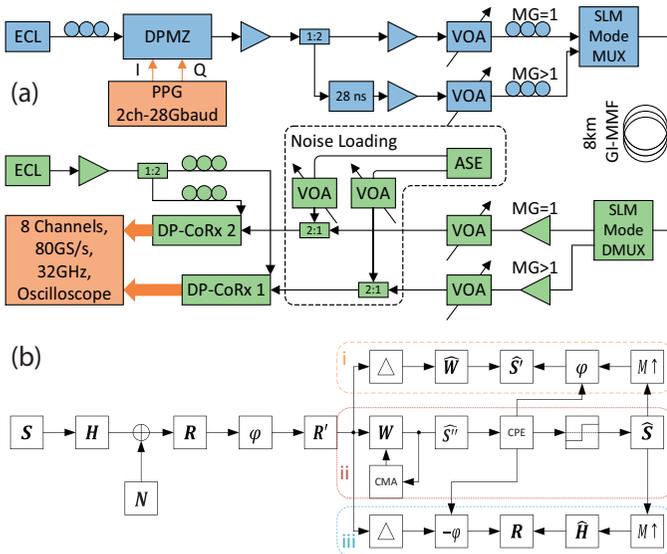


Fig. 1. (a) Experimental setup. ECL: external cavity laser, PPG: pulse pattern generator, DPMZ: dual parallel Mach-Zehnder modulator, VOA: variable optical attenuator, ASE: Amplified spontaneous emission noise, DP-CoRx: dual polarization coherent receiver. (b) Simplified block diagram of the MGDM system with the offline DSP blocks. (i) LS equalizer. (ii) Data recovery. (iii) LS channel estimation. \mathbf{S} : transmitted signal; \mathbf{H} : MIMO channel; \mathbf{R} : received signal w/o carrier phase; \mathbf{R}' : received signal with carrier phase; \mathbf{W} : tap-delayed-line equalizer using LMS algorithm updated according to the CMA. Δ : delay buffers; $\hat{\mathbf{W}}$: tap-delayed-line equalizer using LS algorithm; $\hat{\mathbf{S}}'$: estimated signal with noise and carrier phase; φ : carrier phase; $\hat{\mathbf{S}}''$: estimated signal \mathbf{S} with carrier phase; $\hat{\mathbf{H}}$: estimated MIMO channel; $\hat{\mathbf{S}}$: estimated signal \mathbf{S} . $M \uparrow$: upsample by 2.

of equalizers is investigated in section IV. Finally, a conclusion is given.

II. MIMO BASED DEGENERATE MGDM SYSTEM

Standard 28 Gbaud QPSK transmitters and receivers are used for the 2×2 degenerate MGDM system as shown in Fig. 1 (a). The in-phase (I) and quadrature (Q) components of the QPSK signal are generated using the two data outputs of a pulse pattern generator (PPG) to drive a dual parallel Mach-Zehnder (DPMZ) modulator. Each output is a 28 Gb/s data signal, which is derived from two interleaved 14 Gb/s pseudo random binary sequences (PRBSs) of length $2^{15} - 1$. A delay of 9980 symbols is inserted between the two output signals to decorrelate the IQ components. The optical output of the DPMZ is split into two channels and decorrelated with a relative delay of 28 ns (738 symbols). The optical output of the DPMZ is split into two channels and decorrelated with a relative delay of 28 ns (738 symbols). The SLM based MUX is only capable of launching the channels into two MGs, $LP_{0,1}$ (MG1) and an arbitrary higher order mode, here $LP_{0,2}$ (MG3), each in a single polarization [13]. The launched signals are transmitted through an 8 km GI-MMF. At the mode DEMUX, the optimal weighting of each degenerate mode within the mode group is determined based on the measured optical amplitude and phase of each mode within the mode group and the resulting composite phase mask is displayed on the SLM [6]. The optically demultiplexed signals of different mode groups are combined with amplified spontaneous emission (ASE) noise to vary the received optical signal to noise ratio (OSNR), which is measured by an optical

spectrum analyzer (OSA). The noise loaded signal is then detected by two integrated coherent receivers and an 8-channel real time oscilloscope respectively. An OSNR compensation of 3 dB is incorporated into the measurements to account for the fact that only a single polarization is transmitted, however all optical noise is equally distributed to both polarizations as the DP-CoRx (dual polarization coherent receivers) are used in this experiment.

The 2×2 MIMO system shown in Fig. 1 (a) can be represented in matrix notation as $\mathbf{R} = \mathbf{H}\mathbf{S} + \mathbf{N}$ (see Fig. 1 (b)). Additional noise (φ) is introduced by the phase noise of the transmit and local oscillator lasers and the frequency offset between the transmitter and the receiver. Offline DSP is employed for the data recovery (see Fig. 1 (b-ii)) and the channel estimation (see Fig. 1 (b-iii)). The recovered data sequence $\hat{\mathbf{S}}$ is then correlated with the transmitted PRBS pattern to reconstructed the transmitted signal, which is used to emulate the training sequence for the channel estimation. The LS method is subsequently employed to calculate the channel impulse response of the system as $\hat{\mathbf{H}} = \mathbf{R}\hat{\mathbf{S}}^H (\hat{\mathbf{S}}\hat{\mathbf{S}}^H)^{-1}$ [5]. A 2^{15} twofold oversampled sequence is used to estimate the 1024-sample-long channel impulse response and the result is plotted in Fig. 2. The large central impulse in h_{11} and h_{22} represents the amount of energy that remains in the same mode group as it is selectively launched, which corresponds to the ideal channel response. The remaining impulses, mainly in the anti-diagonal matrices, are caused by the strong crosstalk at the mode MUX/DEMUX. For example, the impulse located at the time of 0 ns in h_{21} represents the crosstalk from $LP_{0,2}$ to MG1 at the mode DEMUX as it propagates at the same speed as the ideal channel impulse in h_{11} . As the magnitudes

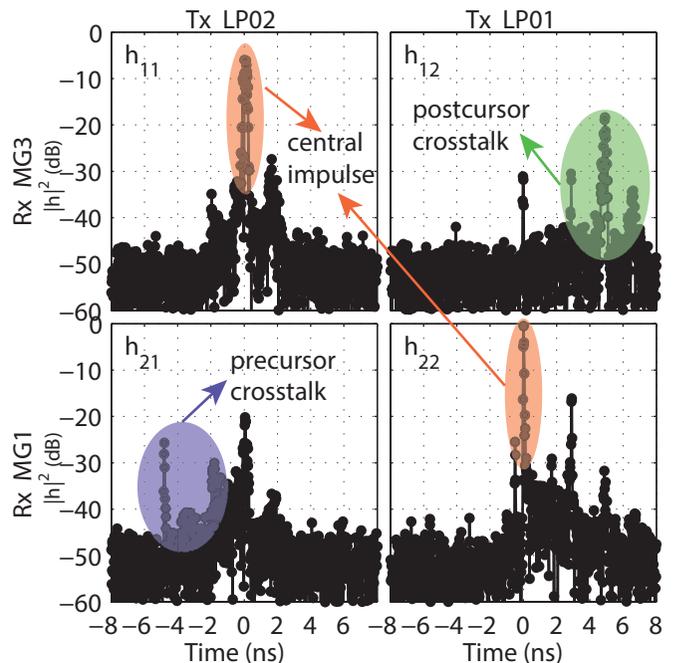


Fig. 2. The estimated 2×2 channel impulse response ($h_{i,j}$) after 8 km GI-MMF.

of the taps between the impulses that arise from the coupling at the MUX/DEMUX are approximately 35 dB lower than the central impulse due to the weak coupling between the MGs in the fiber, it can be considered as a sparse channel. The DGD of the fiber can be calculated from the separation between the impulses, which is approximately 0.36 ns/km for MG2 and 0.61 ns/km for MG3.

In order to undo the crosstalk and coupling in the MIMO system, the length of the uniformly spaced adaptive equalizer \mathbf{W} is required to be at least the same length as the channel impulse response. As a result, the computational complexity, which is defined as the total number of complex multiplications per symbol per mode, increases as a function of the transmission distance. The complexity of various types of equalizers are calculated and plotted in Fig. 3 using the estimated DGD above. The sparse equalizers offer a complexity reduction compared with the conventional TDE, as shown by the delayed TDE and the delayed FDE in Fig. 3, where the inactive taps with small magnitudes are replaced by the delay buffers. In general, the FDEs have a significant advantage of lower complexity over the TDEs when the transmission distance is greater than 0.5 km. This is due to the high efficiency of the fast Fourier transform (FFT). The FFT size is also optimized as a power of two in the plot, which provides the lowest complexity. Ideally, if the DGDs between the degenerate modes in each mode group are zero, the complexity of the delayed equalizers should remain constant for a given number of multiplexed mode groups in the system. The reason is that the coupling in the GI-MMF is weak, therefore expanding the transmission distance mainly introduces extra low magnitude coefficients in the sparse channel response. However, due to impairments

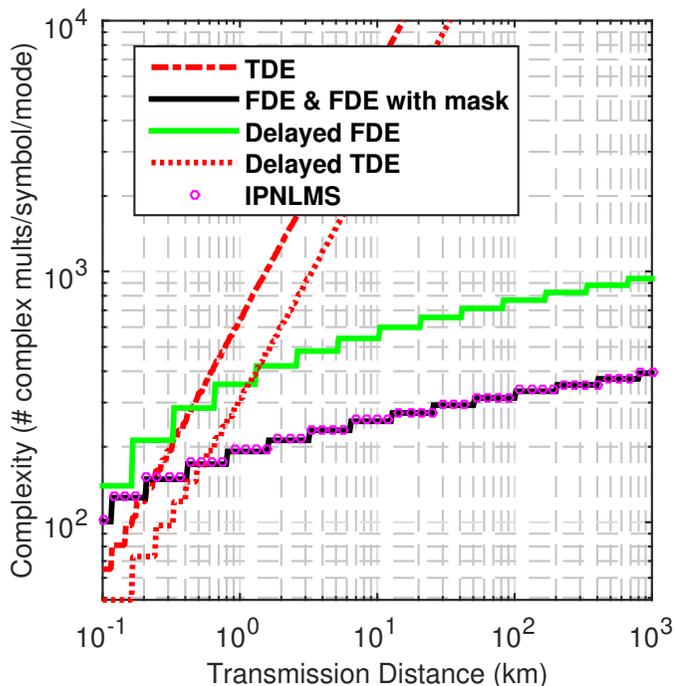


Fig. 3. Computational complexity per symbol per mode for various types of equalizers.

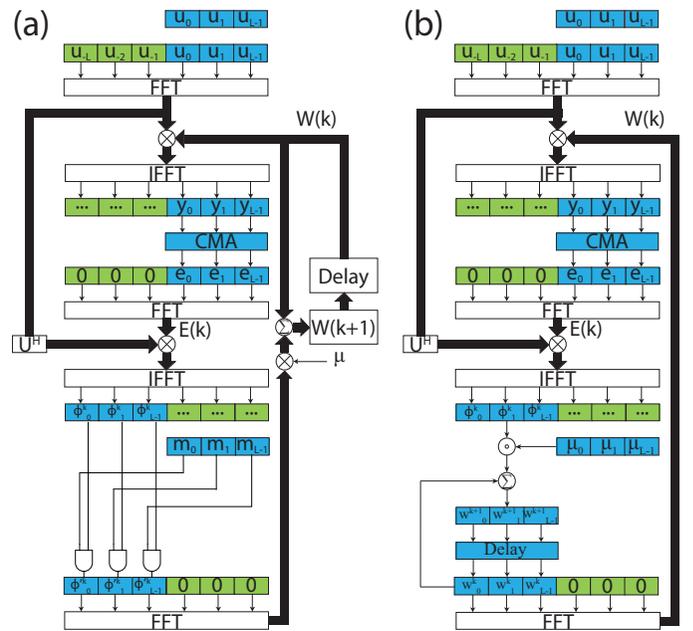


Fig. 4. (a) The FDE with an active taps mask. (b) The IPNLMS FDE. \mathbf{u} : inputs; \mathbf{y} : outputs; \mathbf{e} : error; Φ : cross-correlation between inputs and error signals; \mathbf{w} : equalizer coefficients (time domain); \mathbf{W} : equalizer coefficients (frequency domain); \dots : discard signals.

during the fiber fabrication process, the DGD between the degenerate modes in each mode group is not negligible as shown by the pulse width of the impulses in Fig. 2. Therefore, the required number of taps still increases as a function of the transmission distance for the delayed equalizers due to the broadening of the impulses in the channel response. In order to approximate the necessary active taps in each of the impulses at different distances, a value of 0.22 ns/km is used for the calculation in Fig. 3. The details of the complexity of different equalizers are shown in appendix A. It should be noted that the lowest complexity among all the equalizers is achieved by the conventional uniformly spaced FDE, which is approximately twice lower than the delayed FDE. However, as shown in [8], the performance of the conventional FDE is degraded by the noise accumulated from the inactive taps.

III. SPARSE FDE

In the following section, we introduce two sparse FDEs, which eliminate the noise accumulation while maintaining low complexity when compared to the conventional FDE. In this paper, the overlap-save method with 50% overlap is used in the FDEs, as shown in Fig. 4 [14]. For an equalizer with a length of L , the linear convolution of the tap inputs and the tap weights of the equalizer is performed in the frequency domain as:

$$\mathbf{y}^T(k) = \text{last } L \text{ elements of IFFT} [\mathbf{U}(k)\mathbf{W}(k)], \quad (1)$$

where \mathbf{W} and \mathbf{U} denotes the FFT of the zero-padded tap-coefficient vector \mathbf{w} and the inputs \mathbf{u} respectively. Both of the proposed algorithms are applied on the cross-correlation signals (Φ) between the inputs (\mathbf{u}) and the errors (\mathbf{e}), which

can also be calculated in the frequency domain as:

$$\boldsymbol{\phi}(k) = \text{first } L \text{ elements of IFFT } [\mathbf{U}^H(k)\mathbf{E}(k)], \quad (2)$$

where the cost-function of the Least-Mean-Square (LMS) algorithm is defined as $|\mathbf{e}|^2 = (1 - |\mathbf{y}^2|)^2$ according to the constant-modulus algorithm (CMA) [14].

A. FDE with an active taps mask

The accumulated noise can be eliminated by masking the taps, where the outputs of the taps with small magnitudes are set to zero [15]. In order to implement the mask in the FDE with the same complexity as the conventional FDE, we propose a method with only extra logical AND operations (\wedge). The two inputs of the AND gate are the cross-correlation signal $\boldsymbol{\phi}$ and the active taps mask \mathbf{m} as shown in Fig. 4 (a). In this case, only the taps with significant magnitude are updated every iteration of the LMS equalization by:

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \mu_{\text{FFT}} \begin{bmatrix} \boldsymbol{\phi}(k) \wedge \mathbf{m} \\ \mathbf{0} \end{bmatrix}, \quad (3)$$

The mask \mathbf{m} in eq. (3) is a sparse matrix containing 1s in the position of the active taps and 0s elsewhere (see Fig. 5 (b)). Since the AND operation is used instead of additional multiplications, the complexity of the FDE with an active taps mask follows the same trend as the conventional FDE in Fig. 3.

Various algorithms have been proposed to obtain the optimal position of the active taps based on a priori information of the channel [10]. Instead of complex computation to find the optimal solution, here we propose a simple method as shown in Fig. 5. First the tap coefficients are estimated by using a

training sequence and the LS algorithm as shown in Fig. 1 (b-i):

$$\hat{\mathbf{W}} = \hat{\mathbf{S}}'\mathbf{R}'^H (\mathbf{R}'\mathbf{R}'^H)^{-1}. \quad (4)$$

The magnitudes of the taps of one of the 2×2 equalizer are plotted in Fig. 5 (a). A total number of 2^{15} training symbols with twofold oversampling is used in the LS algorithm to get a close estimation to the optimum coefficients according to the *Wiener-Hopf* equation [14]. The tap position is then sorted by the magnitudes of the tap coefficients, and the first N_m taps with the highest magnitudes are set to 1 while the rest are set to 0 to form the active taps mask as displayed in Fig. 5 (b). The mask is subsequently applied in eq. (3) for the tap update of the FDE. The tap coefficients after the convergence of the adaptive FDE with a total number of 288 active taps is displayed in Fig. 5(c). The performance of this scheme is shown in section IV. There are three groups of active taps at around index 250, 400 and 500, which corresponds to the two selectively launched MGs, i.e. MG1 and MG3 and the intermediate mode group (MG2). The inactive taps between these three impulses are mostly removed while some residual taps remain, which are used to compensate the weak coupling in the fiber.

B. IPNLMS method

When the channel impulse response is not available at the receiver of the system, it is important to develop a sparse adaptive equalizer, where the magnitudes of the inactive taps are further reduced adaptively. Instead of using a constant step-size parameter μ for all the taps, the values of μ vary based on the inputs and the tap coefficients in the IPNLMS as:

$$\mu_l(k) = \frac{\mu_0 L \cdot g_{ip,l}(k)}{\sum_{j=0}^{L-1} |u(k-j)|^2 g_{ip,j}(k) + \delta_{\text{IPLMS}}}, \quad (5)$$

where, $0 < \mu_0 < 1/L$ is the overall step-size factor, L is the equalizer length, and δ_{IPLMS} is the regulation factor proportional to the power of the inputs \mathbf{u} . The parameter $g_{ip,l}(k)$ is the gain distributor which adjusts the step sizes of the individual taps of the equalizer according to:

$$g_{ip,l}(k) = \frac{1 - \beta}{2L} + (1 + \beta) \frac{|w_l(k)|}{2\|\mathbf{w}(k)\|_1 + \epsilon}. \quad (6)$$

Parameter $-1 \leq \beta < 1$ controls the amount of proportionality in the IPNLMS algorithm [11]. We can see from eq. (6), for $\beta = -1$, $g_{ip,j}(k) = 1/L$, which means all the taps are updated with the same step size. When the value of β increases, the second term in eq. (6) gets larger, therefore the taps with higher magnitudes ($|w_l(k)|$) are updated faster than those with lower magnitudes. Hence, the inactive taps remain suppressed close to their initial values, which are set to zeros at the start of the equalization. This is desirable for sparse channels with long impulse response but small number of taps with significant magnitudes as the noise introduced by the inactive taps in the equalizer is reduced.

Similar to the method in section III-A, the variable step size for each tap is introduced through the cross-correlation

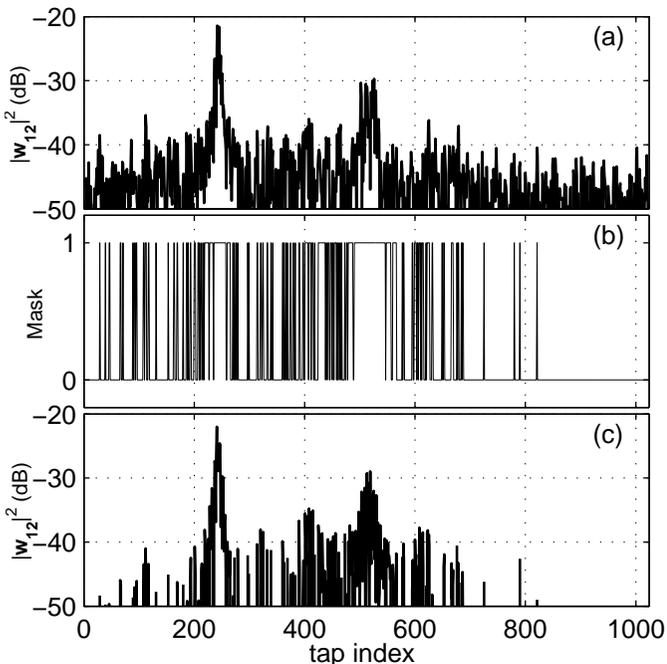


Fig. 5. The magnitudes the tap coefficients of the equalizer from input 2 to output 1 (w_{12}). (a) LS equalizer. (b) Active taps mask with $N_m = 288$. (c) FDE using a mask as shown in (b).

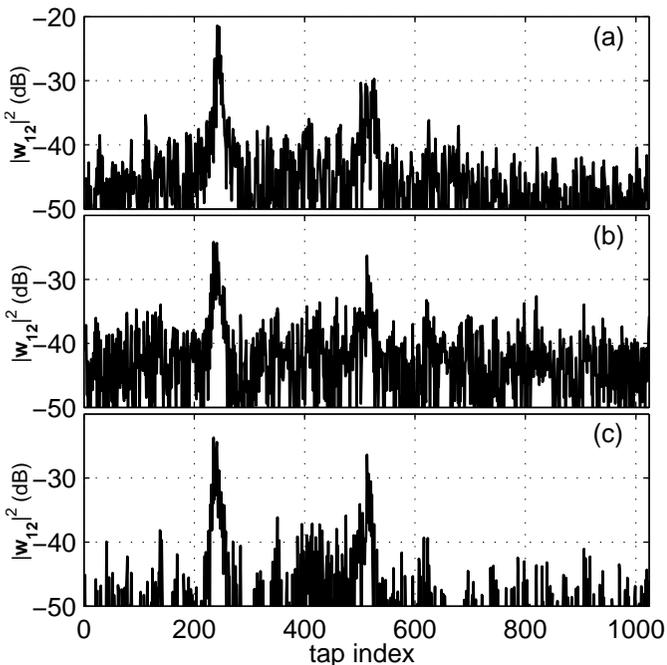


Fig. 6. The magnitudes the tap coefficients of the equalizer from input 2 to output 1 (w_{12}). (a) LS equalizer. (b) IPNLMS equalizer with $\beta = -1$. (c) IPNLMS equalizer with $\beta = -0.3$.

function as shown in Fig. 4 (b). However, the updates of the equalizer are changed from eq. (3) to:

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \text{FFT} \begin{bmatrix} \Phi(k) \circ \boldsymbol{\mu} \\ \mathbf{0} \end{bmatrix}, \quad (7)$$

where $\boldsymbol{\mu} = [\mu_0, \mu_1, \dots, \mu_{L-1}]$ and the symbol \circ represents the entry-wise product. As shown in Fig. 4 (b), the updates of the tap coefficients are performed in the time domain in order to save an extra FFT. Using the notations in appendix A, $2M^2L$ extra real multiplications are required compared with the conventional FDE. If we consider one complex multiplication can be carried out using only three real multiplications, the added complexity on top of the FDE in eq. (9) is $\frac{2ML}{3(N-L+1)}$. However, for the 2×2 MGDM system with $M = 2$, this number is negligible compared with the FDE complexity. Hence, the complexity of the IPNLMS is approximately the same as the conventional FDE as shown in Fig. 3. The extra complexity can be further reduced by switching the equalizer from the IPNLMS to the FDE with a mask after the convergence. While the locations of the active taps can be decided by the IPNLMS coefficients instead of using the LS algorithm with a priori information of the channel.

The resulting magnitudes of the w_{12} taps are shown in Fig. 6. The magnitudes of the coefficients derived from the LS algorithm are plotted in Fig. 6 (a) as a reference. It is seen that when $\beta = -1$, the magnitudes of the inactive taps are approximately 5 dB higher than the case of the LS algorithm. When the value of β increases, the level of the noise floor is reduced accordingly. We have found that the best performance is achieved by the adaptive sparse equalizer for the given MGDM system when $\beta = -0.3$ and the corresponding

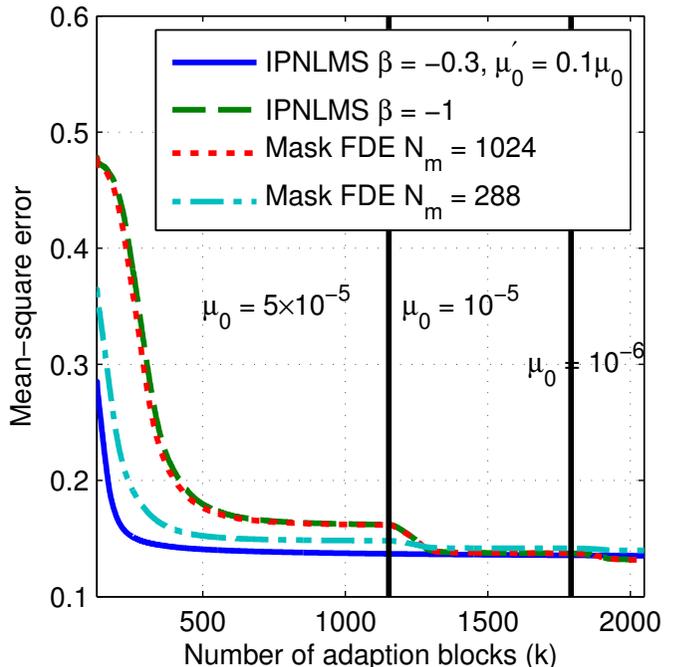


Fig. 7. Learning curves of the sparse FDEs, where the FFT size $N_{FFT} = 1024$.

magnitudes of taps are plotted in Fig. 6 (c). Comparing the tap coefficients in Fig. 5 (c), a similar result is observed, where the magnitudes of the inactive taps are significantly suppressed.

The learning curves of the two proposed methods are plotted in Fig. 7. The mean square errors (MSE) are calculated over each block of data, which is equal to the FFT size, as a 50 % overlap save method is used. For the sparse IPNLMS, the fastest convergence rate is achieved for $\beta = -0.3$. However, the value of μ is required to be much smaller ($\mu'_0 = 0.1\mu_0$) than the other equalizers in the figure in order to assure the convergence of the algorithm. The convergence rate decreases as β increases, and for $\beta = -1$, the MSE converges at the same rate as FDE with a mask containing constant 1s ($N_m = N_{FFT}$), which operates equally as the conventional FDE. It is clearly seen that the FDE with an active taps mask also converges faster when the number of active taps (N_m) is reduced. It should be noted that three discrete stages of the μ parameter are also adapted in the equalizer to improve the steady state performance as indicated in Fig. 7.

IV. SYSTEM PERFORMANCE

The performance of the system using the conventional TDE and FDE is compared by plotting the Q^2 factor as a function of the total number of taps at a received OSNR of 16.8 dB (see Fig. 8). There are two discrete steps in the curve of the TDE when the equalizer length increases beyond 322 and 546 taps for MG1, and 224 and 546 taps for MG3, which corresponds to the equalizer lengths that are required to include the pre/post-cursors in the channel impulse response as shown in Fig. 2. The required equalizer length is twice longer than the delay due to the DGD in order to allow for both pre/post-cursor responses to be compensated. The degradation

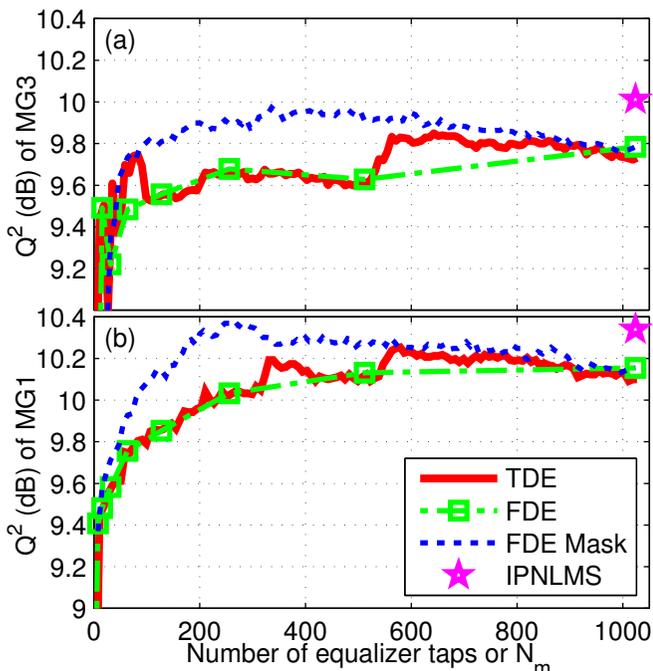


Fig. 8. Q^2 factor for (a) MG3 and (b) MG1 as a function of the total number of taps (or active taps for the FDE with mask). OSNR = 16.8 dB

of the performance shown as the decrease of the Q^2 factor between the two steps is due to the introduction of noise in the equalizer arising from the inactive taps between the central and the pre/post-cursor impulses. The performance of the FDE, shown in Fig. 8, follows the same trend as the TDE, which indicates that it also suffers from noise added from the inactive taps. In addition, the best performance of the conventional TDE can not be achieved using the FDE, as the number of taps is constrained to a power of 2 to achieve high FFT operational efficiency.

As discussed in section III, the noise arising from the inactive taps can be removed by using sparse equalizers. In order to compensate the channel impulse response, the total number of taps for both the FDE with an active taps mask and the IPNLMS algorithm are chosen as 1024. Here, the Q^2 factor is plotted against the total number of active taps (N_m) for the masked FDE. The highest Q^2 factor is achieved with approximately 288 active taps distributed at different locations of the equalizer as shown in Fig. 5 (b). The best performance of the system is achieved when the number of active taps is sufficient for the pulse width of each impulses in the channel response. As the number of active taps increases beyond the optimum number, the performance is degraded gradually as the additional active taps only introduce noise. The performance of the IPNLMS method is also shown in Fig. 8, where the proportional parameter $\beta = -0.3$. In this equalizer the inactive taps are adaptively suppressed, resulting in the same performance as the FDE with an optimum number of the active taps. Overall, both schemes give an improvement of approximately 0.2 dB compared with the conventional FDE of the same length.

As the sparsity is affected by the noise level in the channel

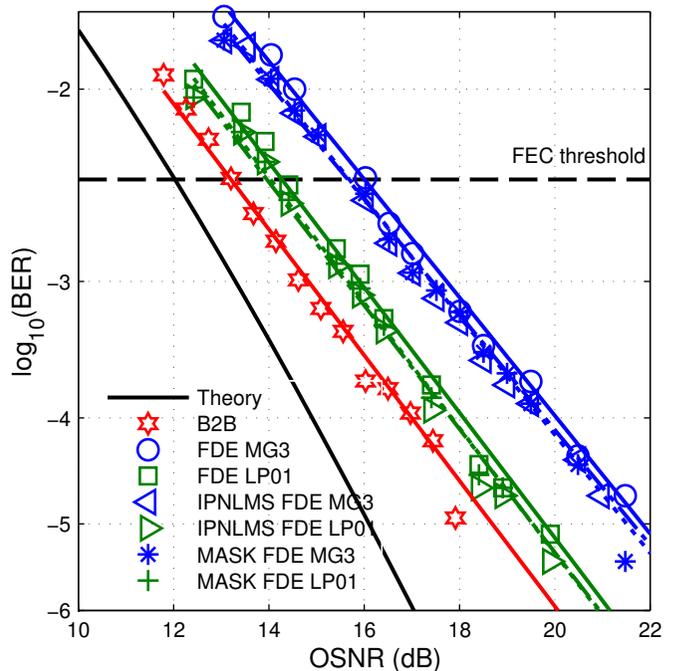


Fig. 9. BER vs. OSNR for different equalizers.

response, it is important to investigate the performance of the sparse FDEs at different values of the OSNR. The bit error rate (BER) as a function of the OSNR is plotted in Fig. 9. It can be seen that both the IPNLMS and the mask method work at different OSNR regions. The BER curve of the conventional FDE is also shown in Fig. 9. However, the improvement gained by using the sparse FDEs is relatively small (~ 0.2 dB) and can only be distinguished by the linearly fitted lines (solid line for the conventional FDE, dotted line for the sparse FDEs). Also, we can see that the IPNLMS provides a slightly better performance over the FDE with an active taps mask, especially in the high OSNR region. This is attributed to the fact that the optimum number of active taps may vary at different OSNRs and is required to be adjusted accordingly. While in Fig. 9, a constant number of 288 is used for all the values of the OSNR.

V. CONCLUSION

The channel impulse response of the MGDM system using conventional GI-MMF is shown to exhibit a sparse response. Two sparse FDEs are introduced for MIMO DSP, one that uses a fixed mask determined using a training sequence to estimate the channel response and another that uses a blind adaptive tap weight update. The computational complexity of the FDE with an active taps mask is the same as the conventional FDE as only logical AND gates are used. However, a priori knowledge of the channel is required to generate the mask. Although, extra real multiplications are required for the IPNLMS algorithm, which can adaptively equalize the sparse channel without a training sequence, the increase in complexity is negligible, especially for a small number of channels. In addition, the accumulated noise from the taps with small magnitudes is eliminated by using the sparse FDEs. A fast convergence rate

is achieved by the IPNLMS FDE. An improvement of approximately 0.2 dB in Q^2 factor is demonstrated by the sparse FDEs compared with the non-sparse equalizers at different values of the OSNR. Similar channel impulse responses have also been discovered in other SDM systems arising from the mode mixing at splice points, which means that the sparse FDE can also be employed to reduce the equalizer induced noise and maintain low computational complexity in these systems as well. As GI-MMF has been widely deployed in data centers due to its low cost, the MGDM system using sparse equalizer can be very useful to meet the growing demand of capacity for these short reach high data capacity applications.

APPENDIX A

CALCULATION OF THE COMPUTATIONAL COMPLEXITY

The complexity in this paper is calculated as the total number of complex multiplications per symbol per mode. For the conventional TDE with M MIMO channels and an equalizer with L T/2 fractional-spaced taps, to obtain one output symbol of each mode, M^2L multiplications are needed. M multiplications are then required for computing the errors and M^2L multiplications are needed for the update of the tap coefficients. Therefore the complexity for the conventional TDE C_{TDE} is expressed as [14]:

$$C_{TDE} = \frac{M^2 \cdot L + M + M^2 \cdot L}{M} = 2ML + 1 \quad (8)$$

For the conventional FDE with a block size of $N-L+1$ and an FFT size of N , $2M^2N$ multiplications and $2M$ FFTs and M inverse FFTs (IFFTs) are needed for calculating $N-L+1$ output symbols for each channel. $M(N-L+1)$ multiplications and M FFT are required for the error calculation. $2M^2N$ multiplications and $\alpha \cdot 2M$ FFTs/IFFTs are needed for updating the tap coefficients, where $\alpha = 2$ for the constrained FDE and $\alpha = 0$ for the unconstrained FDE. Thus, the complexity of the conventional FDE C_{FDE} can be derived as [5]:

$$C_{FDE} = \frac{4MN + (N-L+1) + (4+2\alpha M)C_{FFT}}{N-L+1} \quad (9)$$

The complexity of the delayed TDE is the same as the conventional TDE for a given number of sub-equalizer tap length. Meanwhile the masked FDE is the same as the conventional FDE since AND operation is used instead of additional multiplications. The added complexity of the IPNLMS algorithm is $\frac{2ML}{3(N-L+1)}$ for the entry-wise product.

For the delayed FDE with B sub-equalizers and $B-1$ delays, $2M^2NB$ multiplications and $2MB$ FFTs and M IFFTs are needed for calculating $N-L+1$ output symbols for each channel. The same number of multiplications and FFTs/IFFTs are needed for the error calculation as the conventional FDE and $2M^2NB$ multiplications and $\alpha \cdot 2MB$ FFTs/IFFTs are needed for updating the tap coefficients. Thus, the resulting complexity for the delayed FDE C_{DFDE} is:

$$C_{DFDE} = \frac{4MNB + (N-L+1) + ((2B+2) + 2\alpha MB)C_{FFT}}{\div (N-L+1)} \quad (10)$$

ACKNOWLEDGMENT

The authors would like to thank Teledyne LeCroy for the loan of the multi-channel oscilloscope and Corning Inc. for supplying the fiber.

REFERENCES

- [1] P. Winzer, "Spatial Multiplexing: The Next Frontier in Network Capacity Scaling," in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*, 2013, pp. 372–374.
- [2] M. Suzuki, "Recent progress in space-division multiplexing R&D in Japan," in *2013 IEEE Photonics Society Summer Topical Meeting Series*. IEEE, Jul. 2013, pp. 93–94.
- [3] N. Cvijetic, E. Ip, N. Prasad, M.-J. Li, and T. Wang, "Experimental Time and Frequency Domain MIMO Channel Matrix Characterization versus Distance for 628Gbaud QPSK Transmission over 4025km Few Mode Fiber," in *Optical Fiber Communication Conference*, 2014, p. Th1J.3.
- [4] V. A. J. M. Sleiffer, H. Chen, Y. Jung, P. Leoni, M. Kuschnerov, A. Simperler, H. Fabian, H. Schuh, F. Kub, D. J. Richardson, S. U. Alam, L. Grüner-Nielsen, Y. Sun, A. M. J. Koonen, and H. de Waardt, "Field demonstration of mode-division multiplexing upgrade scenarios on commercial networks." *Optics express*, vol. 21, no. 25, pp. 31 036–46, Dec. 2013.
- [5] S. Randel, A. Sierra, S. Mumtaz, A. Tulino, R. Ryf, P. Winzer, C. Schmidt, and R. Essiambre, "Adaptive MIMO signal processing for mode-division multiplexing," in *Optical Fiber Communication Conference*, 2012, p. OW3D.5.
- [6] J. Carpenter, B. C. Thomsen, and T. D. Wilkinson, "Degenerate Mode-Group Division Multiplexing," *Journal of Lightwave Technology*, vol. 30, no. 99, pp. 3946–3952, Dec. 2012.
- [7] S. Randel, P. Winzer, M. Montoliu, and R. Ryf, "Complexity Analysis of Adaptive Frequency-Domain Equalization for MIMO-SDM Transmission," in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*, 2013, pp. 801–803.
- [8] K. Shi, G. Gordon, and B. C. Thomsen, "Degenerate Mode-Group Division Multiplexing using Delayed Adaptive Frequency-Domain Equalization," in *Optical Fiber Communication Conference*, 2014, p. WIH.2.
- [9] G. Kutz and D. Raphaeli, "Determination of Tap Positions for Sparse Equalizers," *IEEE Transactions on Communications*, vol. 55, no. 9, pp. 1712–1724, Sep. 2007.
- [10] D. Wei, C. K. Sestok, and A. V. Oppenheim, "Sparse Filter Design Under a Quadratic Constraint: Low-Complexity Algorithms," *IEEE Transactions on Signal Processing*, vol. 61, no. 4, pp. 857–870, Feb. 2013.
- [11] Y. Huang, J. Benesty, and J. Chen, *Acoustic MIMO signal processing*. Springer, 2006.
- [12] T. Sakamoto, T. Mori, T. Yamamoto, N. Hanzawa, S. Tomita, F. Yamamoto, K. Saitoh, and M. Koshihara, "Mode-Division Multiplexing Transmission System With DMD-Independent Low Complexity MIMO Processing," *Journal of Lightwave Technology*, vol. 31, no. 13, pp. 2192–2199, Jul. 2013.
- [13] K. Shi, G. Gordon, M. Paskov, J. Carpenter, T. D. Wilkinson, and B. C. Thomsen, "Degenerate mode-group division multiplexing using MIMO digital signal processing," in *2013 IEEE Photonics Society Summer Topical Meeting Series*. IEEE, Jul. 2013, pp. 141–142.
- [14] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, New Jersey, USA: Prentice-Hall, Inc, 2002.
- [15] J. PIETRASKI, Philip, M. BELURI, A. DEMIR, J.-L. PAN, S. STERNBERG, Gregory, R. YANG, and B. LI, "ADAPTIVE EQUALIZER WITH A DUAL-MODE ACTIVE TAPS MASK GENERATOR AND A PILOT REFERENCE SIGNAL AMPLITUDE CONTROL UNIT," 2006.