Research Article | Vol. 28, No. 13 / 22 June 2020 / Optics Express 19133
Optics EXPRESS

Performance of momentum-based frequency-domain MIMO equalizer in the presence of feedback delay

WENTING YI,[1,*] ERIC SILLEKENS,[1] DOMANIÇ LAVERY,[1] HUBERT DZIECIOL,[1] SUJIE ZHOU,[2] KEVIN LAW,[3] JACK CHEN,[2] POLINA BAYVEL,[1] AND ROBERT I. KILLEY[1]

[1]*Optical Networks Group, Department of Electronic & Electrical Engineering, UCL, London WC1E 7JE, UK*
[2]*Huawei Chengdu Research Institute U1, Chengdu, Sichuan Province, China*
[3]*Huawei Base G6, Dongguan, Guangdong Province, China*
[*]*w.yi.17@ucl.ac.uk*

Abstract: A frequency-domain multiple-input multiple-output (FD-MIMO) equalizer employing a momentum-based gradient descent update algorithm is proposed for polarization multiplexing coherent receivers. Its performance in operation with dynamically varying optical channels is investigated and the impact of filter update delays, arising from the latency of the fast Fourier transforms (FFTs) and other digital signal processing (DSP) operations in the feedback loop, is assessed. We show that the proposed momentum-based gradient descent algorithm used to control the equalizer response has significantly greater tolerance to feedback delay than the conventional gradient descent algorithm. We considered a 92 Gbaud dual-polarization 64 QAM receiver, with DSP operating at two samples per symbol, and with the equalizer operating on blocks of 512 and 1024 samples (i.e., 512/1024-point FFT). We found that at an optical signal-to-noise power ratio (OSNR) of 35 dB, the momentum-based gradient descent algorithm can successfully track state-of-polarization (SOP) rotation at frequencies of up to 50 kHz and with filter update delays of up to 14 blocks (39 ns). In comparison, using the conventional gradient descent algorithm in an otherwise identical receiver, the equalizer performance starts to deteriorate at SOP rotation frequencies above 20 kHz.

1. Introduction

The use of digital signal processing (DSP) in coherent receivers is a powerful technique to mitigate static and time-varying impairments from optical channels, such as chromatic dispersion (CD), polarization mode dispersion (PMD) and state-of-polarization (SOP) rotation, and a variety of linear filters and adaptive equalizers have been proposed [1–5]. For applications in short-reach links, it is desirable to minimize the complexity and power consumption. It has been shown that frequency-domain (FD) equalization, utilizing the fast Fourier transform (FFT), is more efficient than time-domain (TD) implementation [6–10]. However, in a practical real-time implementation of an adaptive equalizer with block-based processing, filter update delays will be induced due to the FFTs and other DSP operations in the feedback loop, which can degrade the equalizer's tracking ability, particularly during rapid SOP fluctuations [11–14].

In this paper, we propose a momentum-based frequency-domain multiple-input multiple-output (FD-MIMO) equalizer for polarization-multiplexing coherent receivers, and carried out simulations of dual-polarization 92 Gbaud 64 QAM systems over 40 km and 80 km standard single-mode fiber (SSMF) links, with state-of-polarization rotation frequencies of up to 60 kHz.

#394304 https://doi.org/10.1364/OE.394304
Journal © 2020 Received 6 Apr 2020; revised 9 May 2020; accepted 12 May 2020; published 12 Jun 2020

Optics EXPRESS

The FD-MIMO equalizer combines chromatic dispersion compensation (CDC), matched filtering (MF) and adaptive equalization (AQ), so that frequency to time domain conversions can be avoided to minimize the complexity. The adaptive equalizer is implemented in a blind manner, with the decision-directed least mean square (DD-LMS) algorithm being used for updating the filter coefficients [15]. The mean square error is calculated after carrier phase estimation (CPE) and then fed back to the adaptive equalizer. We show that the momentum-based gradient descent approach is significantly more robust to feedback latency than the conventional gradient update.

2. Operating principle of the FD-MIMO equalizer

Figure 1 shows the block processing of the proposed FD-MIMO equalizer with the overlap-save method being adopted due to its lower complexity compared to overlap-add approach [8,16]. The input time-domain signal is first partitioned into blocks of M samples. In the overlap-save method, the overlap is achieved by taking L samples from the last block, giving a total block length of $L + M = N_{FFT}$. After performing FD-MIMO equalization, the output blocks are combined together in the time domain with overlap samples being removed.

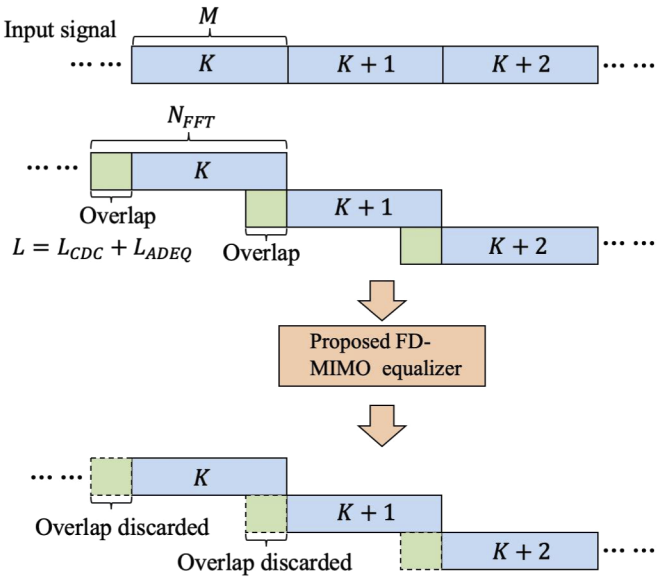

Fig. 1. Block processing of FD-MIMO equalizer with overlap-save method.

The schematic of proposed FD-MIMO equalizer is shown in Fig. 2. It is composed of two static equalizers and an adaptive equalizer. The static equalizers perform chromatic dispersion compensation and matched filtering. The adaptive equalizer mitigates polarization mode dispersion and tracks the SOP. The input time-domain signal, sampled at 2 samples per symbol, is transformed into the frequency domain using a pair of FFTs. As the chromatic dispersion is a static effect, it can be mitigated by multiplying the signal with the inverse transfer function of the dispersive link given by:

$$\mathbf{G}_{CDC} = \exp\left(j\frac{D\lambda^2 z}{4\pi c}\omega^2\right) \tag{1}$$

where D denotes the dispersion parameter, λ is the reference wavelength, z represents the transmission distance, c is the speed of light and ω denotes the angular frequency. It has been shown that in the frequency-domain block processing, the number of overlap samples is critical for

Optics EXPRESS

effective chromatic dispersion compensation [16]. The minimum overlap is calculated assuming the Gaussian pulse broadening [16]:

$$\begin{aligned} L_{CDC} &= 2 \times \lceil \frac{T_p}{2T_s} \rceil + 2 \\ &= 2 \times \lceil \frac{1}{\pi c T_s^2} \sqrt{\pi^2 c^2 T_s^4 + 4\lambda^4 D^2 z^2} \rceil + 2 \\ T_p &= \frac{2}{\pi c T_s} \sqrt{\pi^2 c^2 T_s^4 + 4\lambda^4 D^2 z^2} \end{aligned} \tag{2}$$

where T_p refers to the width of the broadened Gaussian pulse and T_s is the sample duration. For the system we simulated, 92 Gbaud signal transmitted over 40 km and the 80 km links with $D = 16.8$ ps $\cdot$ nm^{-1} $\cdot$ km^{-1}, the minimum overlaps L_{CDC} were 234 and 466 samples respectively. Directly multiplying the frequency-domain signal with $\mathbf{G}_{CDC}$, the first $\frac{L_{CDC}}{2}$ and last $\frac{L_{CDC}}{2}$ samples should be discarded after converting back to the time domain, which accounts for inter-block interference (IBI) [17,18]. For the convenience of implementation, a time delay (i.e., $exp(-j\frac{L_{CDC}}{2} \times \omega)$) is added to the normal CDC transfer function (i.e., $\mathbf{G}_{CDC}$), so that the first L_{CDC} samples are now responsible for IBI and should be removed. This time-delayed CDC is described as:

$$\begin{aligned} \mathbf{H}_{CDC} &= \mathbf{G}_{CDC} \times \exp\left(-j\frac{L_{CDC}}{2}\omega\right) \\ &= \exp\left(j\frac{D\lambda^2 z}{4\pi c}\omega^2 - j\frac{L_{CDC}}{2}\omega\right) \end{aligned} \tag{3}$$

Following CDC, a root-raised cosine (RRC) filter is used to perform matched filtering, which is a complex multiplication between the frequency-domain transfer function of the RRC filter and the signal [19].

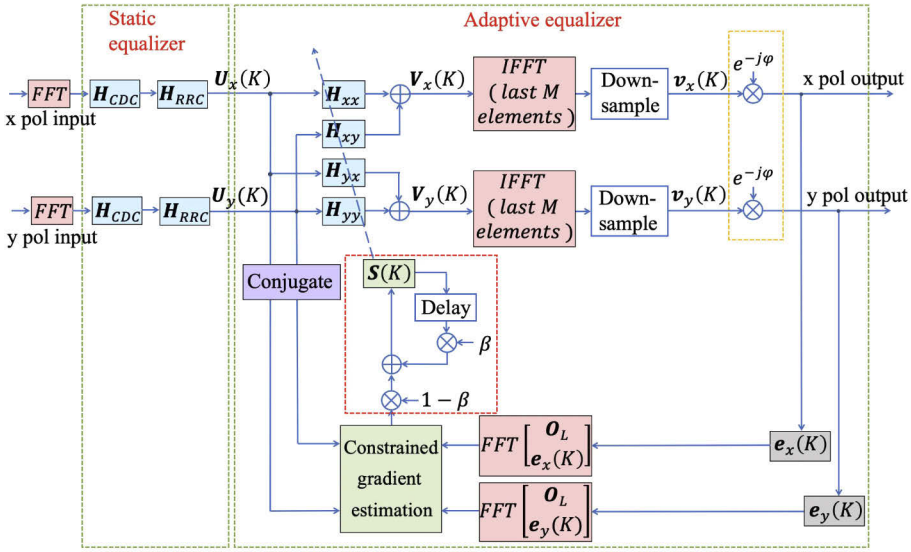

Fig. 2. The schematic of FD-MIMO equalizer. The red-dash box displays the implementation of momentum-based gradient descent and the yellow-dash box shows pilot-based CPE. The $\mathbf{O}_L$ denotes a L by 1 zero vector.

The implementation of frequency-domain adaptive equalizers is well documented [8,9]. The configuration of the adaptive equalizer in this work is similar to the frequency-domain equalizer

Optics EXPRESS

implementation proposed in [9]. Nevertheless, it is worth noting that in our setup, the adaptive equalizer is updated using DD-LMS algorithm while constant modulus algorithm (CMA) method was employed in [9]. In addition, we do not divide the input signal into odd and even samples while in [9], the input signal was divided into odd and even samples to reduce the computational complexity of the adaptive equalizer.

In the proposed FD-MIMO equalizer, the four sets of filter tap weights are initialized in the time domain, each of length L_{ADEQ}. In the simulation, $L_{ADEQ} = 16$. Together with the overlap samples used for chromatic dispersion compensation in the 40 km and 80 km links, the total overlap lengths L were 250 and 482. The adaptive equalizer is operated in the constrained sense with the DD-LMS algorithm being employed to update the filter coefficients [15]. In order to make correct decisions on symbols using DD-LMS algorithm, phase noise needs to be removed first. In the simulation, QPSK pilot symbols inserted at the rate of $\frac{1}{32}$ are used for carrier phase estimation, giving an overhead of 3.1% [20]. The first five QPSK pilots with known phase and polarization state at the transmitter are used to obtain an initial estimation of the polarization rotation, and used to set the center tap of the filter coefficients, with the rest of taps being set to zero. Following this, FFTs are applied to convert the time-domain filter coefficients into the corresponding frequency response. After using pilot symbols for initialization of the adaptive filter coefficients, the adaptive equalizer is updated in a blind manner. Note that it is possible to use pilot symbols in updating the equalizer, which can potentially reduce the equalizer update instability and accelerate the convergence, and this will be investigated in the future.

The equalization process for the two polarizations is described by:

$$\begin{aligned} \mathbf{V}_x(K) &= \mathbf{H}_{xx}(K) \otimes \mathbf{U}_x(K) + \mathbf{H}_{xy}(K) \otimes \mathbf{U}_y(K) \\ \mathbf{V}_y(K) &= \mathbf{H}_{yx}(K) \otimes \mathbf{U}_x(K) + \mathbf{H}_{yy}(K) \otimes \mathbf{U}_y(K) \end{aligned} \tag{4}$$

where $\mathbf{H}_{xx}$, $\mathbf{H}_{xy}$, $\mathbf{H}_{yx}$ and $\mathbf{H}_{yy}$ are the frequency-domain filter coefficients, $\mathbf{U}_x$ and $\mathbf{U}_y$ are the frequency-domain input signals, $\mathbf{V}_x$ and $\mathbf{V}_y$ are the frequency-domain output signals, K is the block number and $\otimes$ denotes element-by-element multiplication.

After converting the equalized signals back to the time domain and down-sampling to 1 sample per symbol, the phase noise is first estimated from pilot symbols and then linearly interpolated between the Wiener filter coefficients and pilot symbols. The Wiener filter is implemented as an FIR filter [21]. The inverse of the phase noise profile (i.e., $e^{-j\varphi}$) is then applied to the time-domain output signal. After removing the phase noise, the time-domain error vectors are calculated, considering only the payload symbols, given by:

$$\mathbf{e}_{x,y}(K) = (\mathbf{d}_{x,y}(K) - e^{-j\varphi}\mathbf{v}_{x,y}(K))e^{j\varphi} \tag{5}$$

where $\mathbf{v}_{x,y}(K)$ denotes the time-domain adaptive equalizer output after down-sampling for x or y polarization as shown in Fig. 2, $\mathbf{d}_{x,y}$ denotes the vector of the most likely symbols after hard decisions. The term $e^{j\varphi}$ in Eq. (5) is used to feed back the estimated phase noise after making hard decisions as the input signals (i.e., $\mathbf{U}_x(K)$ and $\mathbf{U}_y(K)$) of adaptive equalizer are still affected by the phase noise. The errors are then up-sampled and fed back to adjust the MIMO filter coefficients.

The gradient is calculated in the constrained sense which ensures that the frequency-domain filter coefficients and time-domain filter coefficients are equivalent [8,9]. The time-domain gradients are given by:

$$\begin{aligned} \Delta_{xx}(K) &= \text{first } L_{ADEQ} \text{ elements of IFFT}(\mathbf{E}_x(K) \otimes \mathbf{U}_x^*) \\ \Delta_{xy}(K) &= \text{first } L_{ADEQ} \text{ elements of IFFT}(\mathbf{E}_x(K) \otimes \mathbf{U}_y^*) \\ \Delta_{yx}(K) &= \text{first } L_{ADEQ} \text{ elements of IFFT}(\mathbf{E}_y(K) \otimes \mathbf{U}_x^*) \\ \Delta_{yy}(K) &= \text{first } L_{ADEQ} \text{ elements of IFFT}(\mathbf{E}_y(K) \otimes \mathbf{U}_y^*) \end{aligned} \tag{6}$$

Optics EXPRESS

where the asterisk denotes the complex conjugate. $\mathbf{E}_x(K)$ and $\mathbf{E}_y(K)$ represent the frequency-domain error vectors given by:

$$\mathbf{E}_{x,y}(K) = \mathrm{FFT}\begin{bmatrix} \mathbf{O}_L \\ \mathbf{e}_{x,y}(K) \end{bmatrix} \tag{7}$$

where $\mathbf{O}_L$ denotes an L by 1 zero vector and $\mathbf{e}_{x,y}(K)$ is defined in Eq. (5).

Using conventional gradient descent for updating the filter coefficients, $\mathbf{H}_{xx}(K+1)$ can be calculated in the frequency domain as:

$$\mathbf{H}_{xx}(K+1) = \mathbf{H}_{xx}(K) + \mu\mathbf{G}_{xx}(K) \tag{8}$$

where μ denotes the step size. $\mathbf{G}_{xx}(K)$ is the frequency-domain representation of the time-domain gradient given by:

$$\mathbf{G}_{xx}(K) = \mathrm{FFT}\begin{bmatrix} \Delta_{xx}(K) \\ \mathbf{O}_c \end{bmatrix} \tag{9}$$

where $\mathbf{O}_c$ is a zero vector with length of $L_{CDC} + M$ and $\Delta_{xx}(K)$ is defined in Eq. (6).

Equations (6), (7) and (9) are known as constrained gradient estimation. Removing the gradient constraint can lower the calculation complexity, but will result in unstable equalization [8,9]. Therefore, the constrained gradient calculation has been employed in this work. The same operation can be applied to obtain $\mathbf{H}_{xy}(K+1)$, $\mathbf{H}_{yx}(K+1)$, $\mathbf{H}_{yy}(K+1)$.

The implementation of the momentum-based gradient descent algorithm is shown in the red-dash box in Fig. 2. In contrast to the conventional gradient descent approach which only takes the current gradient into account, the momentum-based gradient estimation considers all the previous gradient updates, so that it can smooth out the gradient update oscillations [22]. Updating $\mathbf{H}_{xx}(K+1)$ using momentum-based gradient is given by:

$$\mathbf{H}_{xx}(K+1) = \mathbf{H}_{xx}(K) + \mu\mathbf{S}_{xx}(K) \tag{10}$$

where $\mathbf{S}_{xx}(K)$ denotes the frequency-domain momentum-based gradient. It is calculated as [22]:

$$\mathbf{S}_{xx}(K) = (1-\beta)\mathbf{G}_{xx}(K) + \beta\mathbf{S}_{xx}(K-1) \tag{11}$$

where $\mathbf{G}_{xx}(K)$ is the frequency-domain conventional gradient defined in Eq. (9), β is the momentum parameter from $[0, 1]$, $\mathbf{S}_{xx}(K-1)$ denotes the momentum-based gradient calculation from the previous block [22]. Updating filter coefficients $\mathbf{H}_{xy}(K+1)$, $\mathbf{H}_{yx}(K+1)$, and $\mathbf{H}_{yy}(K+1)$ using momentum-based gradient descent can be implemented in a similar approach.

3. Simulation setup

The simulation was carried out using Matlab. We considered a 92 Gbaud dual-polarization 64 QAM receiver, with DSP operating at two samples per symbol, and with the equalizer operating on blocks of 512 samples or 1024 samples. The FFTs and other DSP operations in the feedback loop were assumed to require multiple clock cycles [23], giving rise to a delay in the feedback loop in the equalizer, and a consequent impact on the equalizer's convergence rate. A long random input signal sequence of 9.8304×10^5 symbols was generated, allowing the adaptive filter to converge to and then accurately track the signal polarization and compensate the link's PMD. The system performance was evaluated on the last 55, 536 symbols of the signal once the adaptive equalizer had converged. At the transmitter, the dual-polarization 92 Gbaud 64 QAM signal was generated with the insertion of QPSK pilot symbols at the rate of $\frac{1}{32}$. The signal was

Optics EXPRESS

up-sampled to 2 samples/symbol and spectrally shaped by an RRC filter with 0.01 roll-off factor, following which an ideal dual-polarization IQ modulator was applied. The transmitter laser was centered at 1550 nm and had a linewidth of 100 kHz. The launch power was set to -8 dBm.

The optical channel considered in this work was a single span of standard single-mode fiber (SSMF), with a length of 40 km or 80 km. A nonlinear wavelength-division multiplexing (WDM) system (7 channels) assuming -8 dBm launch power per channel and 100 GHz channel spacing was simulated. The impact of fiber nonlinearity was found to be negligible, and so, since the focus of the paper is on the equalizer's tracking ability of SOP rotations, only the linear effects of chromatic dispersion, PMD and SOP rotations were included in the rest of the simulations described in this paper. The chromatic dispersion was assumed to be 16.8 $\text{ps} \cdot \text{nm}^{-1} \cdot \text{km}^{-1}$. The effect of PMD was included by modeling the 40 km and 80 km fiber span as 40 and 80 cascaded 1 km sections respectively, each with random birefringence. For a given section i, it can be described using the Jones matrices R_i and D_i given by [24]:

$$\mathbf{R}_i = \begin{bmatrix} \cos\theta_i e^{-j\phi_i} & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i e^{j\phi_i} \end{bmatrix} \quad \mathbf{D}_i = \begin{bmatrix} e^{-j\omega\frac{\tau_i}{2}} & 0 \\ 0 & e^{j\omega\frac{\tau_i}{2}} \end{bmatrix} \tag{12}$$

where θ_i and ϕ_i refer to the random rotation and random phase shift between two polarizations following the uniform distribution between $[-\pi, \pi]$ and $[-\pi/2, \pi/2]$ respectively, τ_i denotes the local differential group delay (DGD) [24]. Given a PMD parameter of 0.5 $ps/\sqrt{km}$, τ_i was set to 0.5427 ps [1]. Continuous SOP rotation J_{SOP} was added at the input of the fiber span to test the dynamic tracking ability of the FD-MIMO equalizer given by [1]:

$$\mathbf{J}_{SOP} = \begin{bmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{bmatrix} \tag{13}$$

Hence, the dynamic channel was modelled as:

$$\mathbf{J} = \mathbf{R}_N\mathbf{D}_N \ldots \mathbf{R}_i\mathbf{D}_i \ldots \mathbf{R}_1\mathbf{D}_1\mathbf{J}_{SOP} \tag{14}$$

where N denotes the total number of cascaded sections. Assuming optical amplification at the receiver, amplified spontaneous emission (ASE) noise was added to the signal. The received optical signal-to-noise power ratio (OSNR) is defined in a 0.1 nm resolution bandwidth.

At the receiver, the signal was combined with a local oscillator (LO) centered at 1550 nm in an ideal polarization-diverse 90 degree optical hybrid coherent receiver. The power of the LO was 13 dBm and its linewidth was 100 kHz. Photodetection and balanced detection were assumed to be ideal to cancel out common modes. The signal then passed through the proposed FD-MIMO equalizer to mitigate impairments from the optical channel (i.e., chromatic dispersion, PMD and SOP rotation) and the laser induced phase noise. Finally, the system performance was evaluated using generalized mutual information (GMI) once the adaptive equalizer had converged.

In the ideal case without delay in the filter update, the signal in block K is equalized as shown in Eq. (4). However, in the presence of latency which is unavoidable in practical real-time implementations, the adaptive equalization for two polarizations will be:

$$\begin{aligned} \mathbf{V}_x(K) &= \mathbf{H}_{xx}(K-P) \otimes \mathbf{U}_x(K) + \mathbf{H}_{xy}(K-P) \otimes \mathbf{U}_y(K) \\ \mathbf{V}_y(K) &= \mathbf{H}_{yx}(K-P) \otimes \mathbf{U}_x(K) + \mathbf{H}_{yy}(K-P) \otimes \mathbf{U}_y(K) \end{aligned} \tag{15}$$

where P denotes the delay in terms of the number of blocks.

4. Results and discussion

The SOP rotation frequency was varied from 0 to 60 kHz . The ASE noise was loaded, varying the received OSNR value. The performance was quantified using generalized mutual information

Optics EXPRESS

(GMI)/m where m denotes the number of bits per symbol. For the 64 QAM signal, m equals to 6. The (GMI)/m was obtained over ten runs of the simulation with random rotation and random phase shift (i.e., random θ and ϕ) of the received signal at each SOP rotation frequency and OSNR value. The momentum parameter β was set to 0.9 [22].

In the simulation of the 40 km link, the FFT size, corresponding to the block length, was 512 points. Figure 3 shows the performance of the FD-MIMO equalizer at an OSNR of 35 dB. Both conventional gradient descent and momentum-based gradient descent were implemented in the frequency domain as described in Section 2. In the absence of feedback latency (i.e., zero delay), the FD-MIMO equalizers using both the conventional gradient descent and the momentum-based gradient estimation successfully mitigate the phase noise, the static and dynamic distortions from the fiber, giving a GMI/m of approximately 1 (12 bits/4D-symbol). When including a delay in the feedback used for the filter update, the tracking ability of the equalizer employing the conventional gradient calculation is impaired. Its performance starts to break down at a delay of 4 blocks and at a SOP rotation frequency of 60 kHz. In contrast, applying momentum-based gradient estimation, the tracking ability of the equalizer is significantly improved and is robust to delays up to 14 blocks (39 ns) with SOP rotation frequencies up to 50 kHz. At a SOP rotation frequency of 60 kHz, the equalizer's tracking ability becomes impaired at feedback delays above 13 blocks.

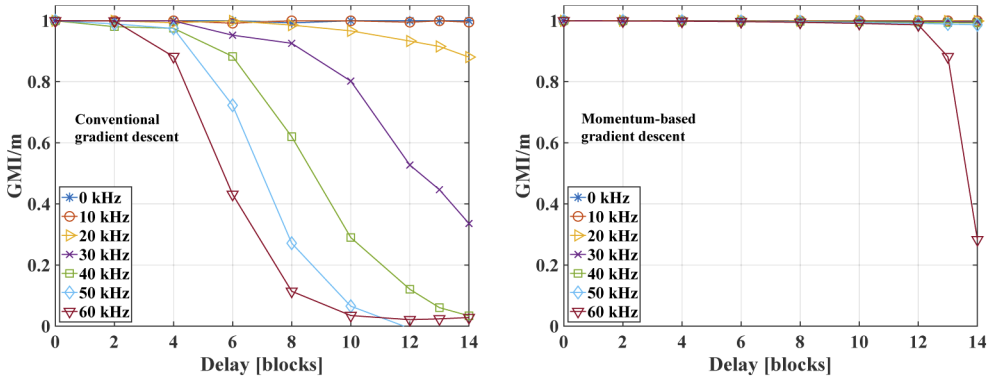

Fig. 3. The performance of FD-MIMO equalizer at an OSNR of 35 dB. Left: conventional gradient descent. Right: momentum-based gradient descent.

At 50 kHz SOP rotation frequency and zero block delay, the received signal scatter plots for the x polarization using conventional gradient descent and momentum-based gradient descent are compared in Fig. 4. The y polarization performs similarly. Without a delay in the feedback loop, both methods achieve good channel equalization at 50 kHz SOP rotation frequency resulting a clean scatter plot. The noisy clouds around the ideal constellation points are mainly due to ASE noise. The scatter plots when employing the conventional gradient descent method with 8 blocks delay in the feedback loop, and the momentum-based gradient descent with 14 blocks delay are presented in Fig. 5. In the case of conventional gradient descent, the equalizer starts to fail to track 50 kHz SOP rotation at 8 blocks delay, resulting in a more distorted constellation (GMI/m $\approx$ 0.28 as shown in Fig. 3). In contrast, employing momentum-based gradient descent, even at 14 blocks delay as shown in Fig. 5 (right), the equalizer is still capable of recovering the signal and the scatter plot is relatively clear (GMI/m $\approx$ 0.98 as shown in Fig. 3).

The results with the 40 km link at an OSNR of 30 dB are shown in Fig. 6, and the maximum achievable GMI/m decreases slightly to 0.98 due to the additional ASE noise. The ASE noise also degrades the performance of the FD-MIMO equalizer. Similarly to the trend in Fig. 3, the conventional gradient estimation fails to track high SOP rotation ($\geq$ 30 kHz) when there are

Optics EXPRESS

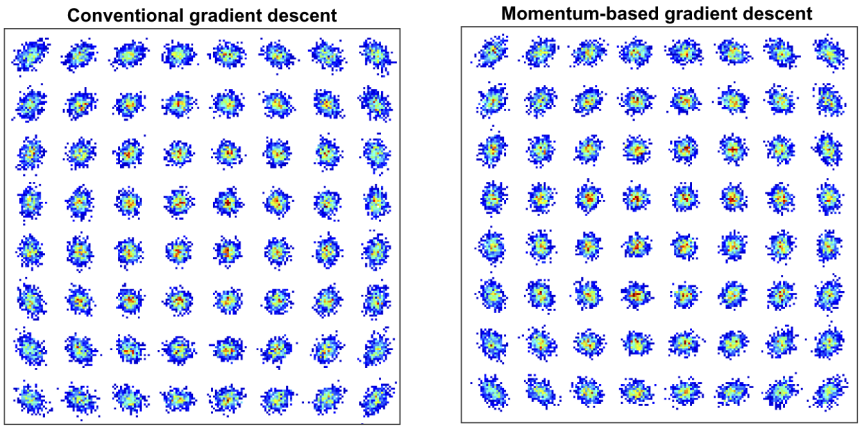

Fig. 4. Without feedback delay, the received signal constellations employing conventional gradient descent (left) and momentum-based gradient descent (right) at 35 dB OSNR and 50 kHz SOP rotation speed.

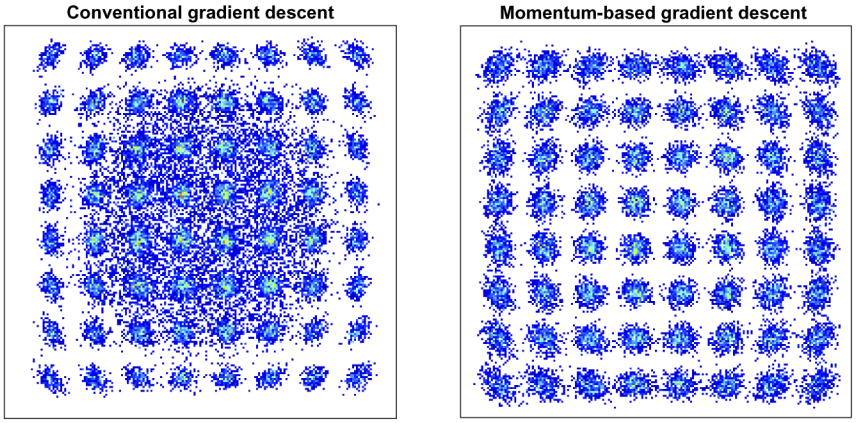

Fig. 5. The received signal constellations of using conventional gradient descent at 8 blocks delay (left), and momentum-based gradient descent at 14 blocks delays (right) at 35 dB OSNR and 50 kHz SOP rotation speed.

delays of 2 blocks or more. In contrast, the performance of the momentum-based gradient descent algorithm only starts to deteriorate at a SOP rotation frequency of 60 kHz and at feedback delays of 8 blocks and above.

In Fig. 7, simulation results for the 80 km link with received OSNR of 27 dB are plotted, showing the performance for longer distance and lower OSNR. According to Eq. (2), the required overlap samples for chromatic dispersion compensation of 80 km SSMF are 466. The overlap samples used for the adaptive equalizer were kept at 16, and therefore, the total overlap length was 482. The FFT size was increased to 1024 points.

At the OSNR of 27 dB, the initial GMI/m is around 0.91. The trends in the results are similar to those observed in the 40 km link as discussed above. However, due to the lower OSNR, the amount of ASE noise causes more perturbations to the equalizer's tracking ability. The equalizer fails to track 10 kHz SOP rotation with feedback delay of 8 blocks or more when employing conventional gradient descent. In contrast, momentum-based gradient descent can still successfully track SOP rotations of up to 20 kHz with 13 blocks delay.

Optics EXPRESS

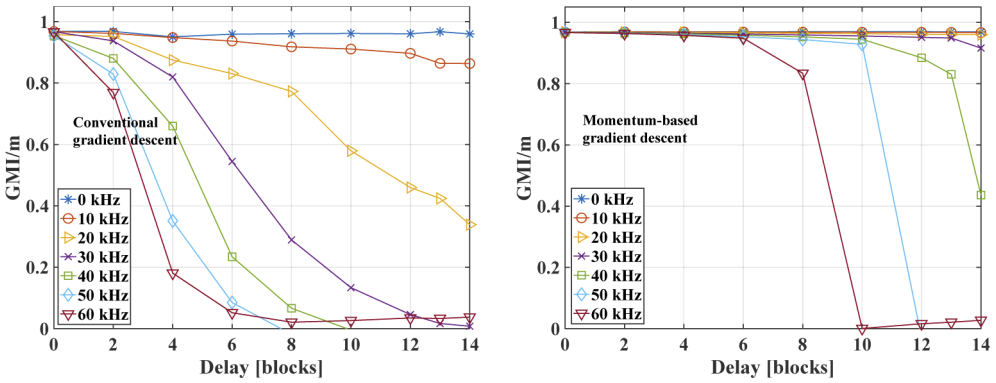

Fig. 6. The performance of FD-MIMO equalizer at an OSNR of 30 dB. Left: conventional gradient descent. Right: momentum-based gradient descent.

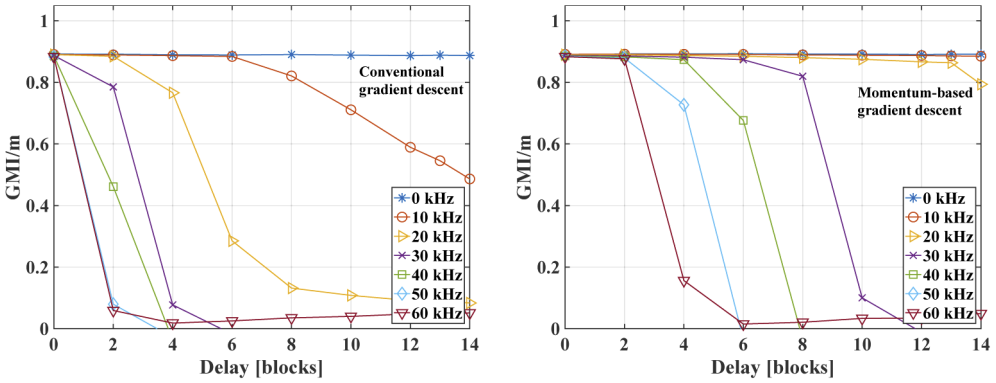

Fig. 7. The performance of FD-MIMO equalizer at an OSNR of 27 dB. Left: conventional gradient descent. Right: momentum-based gradient descent.

The evolution of SOP rotation can be analysed using Stokes parameters, and visualized on the Poincaré sphere [25]. Assuming a linearly polarized input (i.e., $s = [1, 0, 0]^T$) and 50 kHz SOP rotation frequency, the trajectory of the first 2^{16} time instances obtained using Eq. (13) are shown in Fig. 8 (left). To more exhaustively test the equalizer, we also simulated SOP fluctuations with time as random walks of the polarization angle and phase. In these simulations, J_{SOP} in Eq. (13) was replaced by:

$$\mathbf{J}_{SOP} = \begin{bmatrix} \cos\theta(t)e^{-j\phi(t)} & \sin\theta(t) \\ -\sin\theta(t) & \cos\theta(t)e^{j\phi(t)} \end{bmatrix} \tag{16}$$

where the sampled values $\theta(t)$ and $\phi(t)$ were generated by random walks with the change in θ and ϕ between samples being normally distributed random values. An example of the resulting trajectory of the Stokes vector is plotted in Fig. 8 (right). It can been seen that the state of polarization traces a random path around the surface of the Poincaré sphere. In this case, the momentum-based gradient descent algorithm was still found to be more robust to feedback delays and fast SOP fluctuations than the conventional gradient descent approach. Simulation results show that in the 40 km link, with a received OSNR of 35 dB and with the standard deviation of the SOP rotation frequency set to 1.1 MHz, the momentum-based gradient descent method can successfully track SOP rotation and tolerate delays of up to 14 blocks (GMI/$m \approx 1$). However, the conventional gradient descent starts to break down at a delay of 10 blocks (GMI/$m \approx 0.64$).

Optics EXPRESS

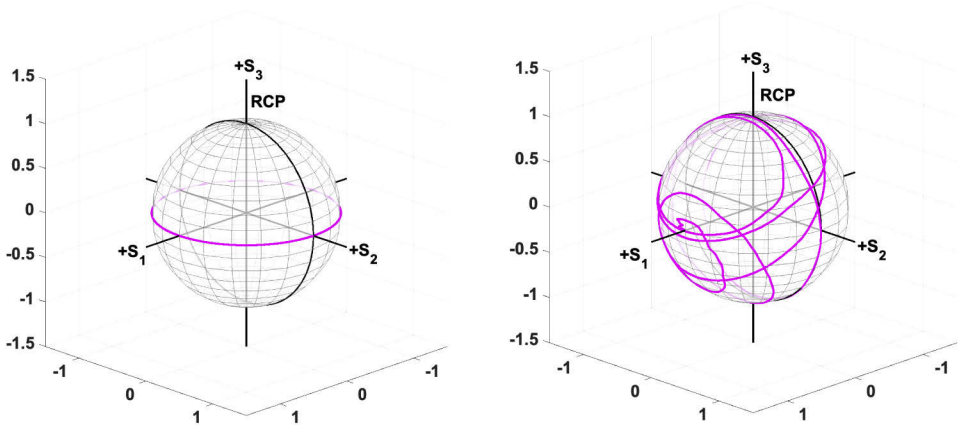

Fig. 8. The trajectory of the Stokes parameters, with the continuous time varying polarization angle (left) and with polarization angle (θ) and phase (ϕ) following random walks (right). RCP: right circularly polarized.

5. Conclusion

We have presented a frequency-domain multiple-input multiple-output equalizer for dual-polarization coherent receivers updated using a momentum-based gradient descent algorithm. Its performance was tested in a dynamically varying optical channel. The impact of feedback delays due to FFTs and other DSP operations in the feedback loop was also assessed. The proposed frequency-domain equalizer achieves SOP tracking with a high tolerance to feedback latency. In simulations of the dual-polarization 92 Gbaud 64 QAM signal transmitted over 40 km SSMF, with OSNR = 35 dB, and using 512-sample block processing, it is capable of tracking SOP rotation at frequencies of up to 50 kHz and compensating PMD even with delays of up to 14 blocks. In transmissions over 80 km SSMF, with OSNR = 27 dB, and using 1024-sample block processing, it is capable of tracking SOP rotation at frequencies of up to 20 kHz and compensating PMD even with delays of up to 13 blocks.

Funding

Engineering and Physical Sciences Research Council (EP/R035342/1).

Disclosures

The authors declare no conflicts of interest.

References

1. S. J. Savory, "Digital filters for coherent optical receivers," Opt. Express **16**(2), 804–817 (2008).
2. E. Ip, A. P. T. Lau, D. J. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," Opt. Express **16**(2), 753–791 (2008).
3. M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats," J. Lightwave Technol. **35**(5), 1125–1141 (2017).
4. K. Zhong, X. Zhou, J. Huo, C. Yu, C. Lu, and A. P. T. Lau, "Digital signal processing for short-reach optical communications: A review of current technologies and future trends," J. Lightwave Technol. **36**(2), 377–400 (2018).
5. P. Winzer, A. Gnauck, C. Doerr, M. Magarini, and L. Buhl, "Spectrally efficient long-haul optical networking using 112-Gb/s polarization-multiplexed 16-QAM," J. Lightwave Technol. **28**(4), 547–556 (2010).
6. B. Spinnler, F. N. Hauske, and M. Kuschnerov, "Adaptive equalizer complexity in coherent optical receivers," in *European Conference on Optical Communication*, (2008) p. We.2.E.4.
7. J. Leibrich and W. Rosenkranz, "Frequency domain equalization with minimum complexity in coherent optical transmission systems," in *Optical Fiber Communication Conference*, (2010), p. OWV1.

Optics EXPRESS

8. J. J. Shynk, "Frequency-domain and multirate adaptive filtering," IEEE Signal Process. Mag. **9**(1), 14–37 (1992).
9. M. S. Faruk and K. Kikuchi, "Adaptive frequency-domain equalization in digital coherent optical receivers," Opt. Express **19**(13), 12789–12798 (2011).
10. C. Fougstedt, P. Johannisson, L. Svensson, and P. Larsson-Edefors, "Dynamic equalizer power dissipation optimization," in *Optical Fiber Communication Conference*, (2016), p. W4A.2.
11. D. Lavery, B. C. Thomsen, P. Bayvel, and S. J. Savory, "Reduced complexity equalization for coherent long-reach passive optical networks [invited]," J. Opt. Commun. Netw. **7**(1), A16–A27 (2015).
12. Q. Guo, B. Xu, and K. Qiu, "Studies on effects of feedback delay on the convergence performance of adaptive time-domain equalizers for fiber dispersive channels," Opt. Eng. **55**(4), 046110 (2016).
13. P. M. Krummrich, D. Ronnenberg, W. Schairer, D. Wienold, F. Jenau, and M. Herrmann, "Demanding response time requirements on coherent receivers due to fast polarization rotations caused by lightning events," Opt. Express **24**(11), 12442–12457 (2016).
14. T. Zeng, Z. he, L. Meng, J. Li, X. Li, and S. Yu, "The real time implementation of a simplified 2-section equalizer with supernal SOP tracking capability," in *Optical Fiber Communication Conference*, (2020), p. M2J.7.
15. I. Fatadin, D. Ives, and S. J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," J. Lightwave Technol. **27**(15), 3042–3049 (2009).
16. T. Xu, G. Jacobsen, S. Popov, M. Forzati, J. Mårtensson, M. Mussolin, J. Li, K. Wang, Y. Zhang, and A. T. Friberg, "Frequency-domain chromatic dispersion equalization using overlap-add methods in coherent optical system," J. Opt. Commun. **32**(2), 131–135 (2011).
17. R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, and Y. Miyamoto, "Coherent optical single carrier transmission using overlap frequency domain equalization for long-haul optical systems," J. Lightwave Technol. **27**(16), 3721–3728 (2009).
18. K. Ishihara, R. Kudo, T. Kobayashi, A. Sano, Y. Takatori, T. Nakagawa, and Y. Miyamoto, "Frequency-domain equalization for coherent optical transmission systems," in *Optical Fiber Communication Conference*, (2011), p. OWW4.
19. K. Kikuchi, "Fundamentals of coherent optical fiber communications," J. Lightwave Technol. **34**(1), 157–179 (2016).
20. Y. Wakayama, E. Sillekens, L. Galdino, D. Lavery, R. I. Killey, and P. Bayvel, "Increasing achievable information rates with pilot-based DSP in standard intradyne detection," in *European Conference on Optical Communication*, (2019), p. W.1.B.5.
21. E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," J. Lightwave Technol. **25**(9), 2675–2692 (2007).
22. S. Ruder, "An overview of gradient descent optimization algorithms. arxiv 2016," arXiv preprint arXiv:1609.04747 (2016).
23. J. S. Bruno, V. Almenar, and J. Valls, "FPGA implementation of a 10 GS/s variable-length FFT for OFDM-based optical communication systems," Microprocess Microsy. **64**, 195–204 (2019).
24. G. P. Agrawal, "Polarization effects," *Nonlinear fiber optics* (Elsevier, 2013).
25. E. Collett, "The wave theory of light," *Field guide to polarization* (Spie, Bellingham, WA, 2005).