Improved polarization dependent loss tolerance for polarization multiplexed coherent optical systems by polarization pairwise coding

Chen Zhu,1 Binhuang Song,1 Bill Corcoran,1,2 Leimeng Zhuang1 and Arthur James Lowery1,2,*

1Electro-Photonics Laboratory, Electrical and Computer Systems Engineering, Monash University, Clayton, VIC 3800, Australia
2Centre for Ultrahigh-bandwidth Devices for Optical Systems (CUDOS), Australia
*arthur.lowery@monash.edu

Abstract: Polarization dependent loss (PDL) causes imbalanced optical signal to noise ratio (OSNR) of the two polarizations, thus remains one of the major bottlenecks for next-generation polarization-division-multiplexed (PDM) coherent optical transmission systems. In this paper, we investigate Pairwise Coding for adaptive PDL mitigation in PDM coherent optical systems. By pre-coding across two polarizations, the PDL-induced performance degradation can be largely mitigated without any coding overhead. We present details of the coding and de-coding design, and also derive the analytical symbol/bit error rate of the Polarization Pairwise Coding scheme, which can be used to predict the performance gain as well as for optimal rotation angle calculation. Simulation results verify that Pairwise Coding achieves substantial system performance gains over a wide range of PDL values. Compared with other digital coding techniques, Polarization Pairwise Coding shows improved performance than Walsh-Hadamard transform since it maximizes the coordinate diversity; and also Pairwise Coding is computationally much simpler to decode compared with the Golden and Silver Codes, therefore is practical for current 100-Gb/s and future 400-Gb/s and 1-Tb/s digital coherent transceivers.

OCIS codes: (060.2330) Fiber optics communications; (060.1660) Coherent communications.

References and links


1. Introduction

The combination of advanced modulation formats, polarization-division-multiplexing (PDM) and digital coherent receivers enables next-generation high-capacity optical transmission [1]. By obtaining the full optical field information with coherent detection, powerful digital signal processing (DSP) techniques enable effective compensation of most of the system impairments such as I/Q imbalance, chromatic dispersion (CD), polarization mode dispersion (PMD), laser phase noise and local oscillator frequency offset [2]; even fiber nonlinearity can be effectively mitigated to a certain level [3]. However, polarization dependent loss (PDL), which refers to two orthogonal polarizations being attenuated differently, remains an unsolved problem due to its non-unitary nature. Although PDL is negligible in fibers, it is significant in discrete devices such as amplifiers, wavelength division multiplexers, circulators, and isolators [4]. After long-haul transmission, an accumulated PDL of several dB can be easily observed, which may become one of the major bottlenecks for the high-speed long-haul PDM coherent optical systems.

In a PDM coherent optical system, PDL causes non-orthogonality of the PDM signals and imbalance in the received optical signal to noise ratios (OSNR) of the two signal polarizations [5, 6]. Signal non-orthogonality can be equalized with an adaptive multiple-input-multiple-output (MIMO) polarization mode dispersion (PMD) equalizer; however, OSNR imbalance
will eventually limit the overall system performance. In practical systems, the data bits for both polarizations are encoded together so that there is only one FEC decoder at the receiver, where then the total system performance is the average of both “good” and “bad” polarizations; however, the overall system performance will still be dominated by the lossy polarization.

The degraded OSNR of the lossy polarization is similar to frequency selective fading in a wireless system; thus, PDL could be viewed as “polarization selective fading”. Therefore, coding concepts from wireless systems might be able to be applied for PDL mitigation. Recently, polarization-time coding, in the form of the Golden and Silver Codes, has been introduced for PDM coherent optical systems to achieve superior PDL tolerance [7–9]. Golden and Silver Codes encode across four symbols (two in each polarization) with a non-invertible coding matrix, and they require computationally expensive maximum likelihood sequence estimators at the receiver. A semi-Silver Code has been demonstrated in [10], using a 4 × 4 adaptive filter for simpler joint channel equalization and decoding; however, the performance is compromised and the computational effort is still double that of a conventional 2 × 2 adaptive filter. The Walsh-Hadamard transform [11], on the other hand, has a simpler coding/decoding transform matrix and maintains the normal constellation decision region at the receiver. It is also able to equalize the OSNR difference between two polarizations; however, the performance improvement is limited because it does not fully utilize the coordinate diversity. The optical implementation of low-complexity space-time pre-coding [12] and disjoint receiver detection schemes [13] either provides limited system performance gain, or increases the system’s hardware complexity.

Pairwise Coding originates from the scheme of maximizing the signal space diversity by rotating the conventional signal constellations [14], i.e., rotating the constellation maximizes the performance when the powers of the I and Q noise components are not identical. Pairwise Coding was first applied to single-input-single-output systems to improve the performance over fading channels [14], where the I and Q components are interleaved to allow different channel gains, mitigating the imbalanced SNRs of the I and Q components. It was then extended to MIMO wireless systems, where pairing of sub-channels with different signal-to-interference-and-noise-ratios (SINR), using the same rotation angle and exchanging the real/imaginary parts between different sub-channels, improved the overall BER performance [15]. This Pairwise Coding also improved the receiver sensitivity for direct-detection optical orthogonal-frequency-division multiplexing (OFDM) [16]. At OFC 2015 we reported Pairwise Coding in a PDM coherent optical system for PDL mitigation [17]; our experimental results showed that the pairing of two polarizations significantly improves the system performance over a wide range of worst-case PDL values. Importantly, Pairwise Coding does not require an overhead that would reduce the payload data rate, and needs only a few extra computations per symbol, because at the receiver end, after I/Q de-interleaving, only symbol-by-symbol decision processing is required.

In this paper, we present a theoretical analysis for the Pairwise Coding and decoding design, and derive the analytical symbol error rate (SER) and bit error rate (BER) for pairwise coded signals, which leads to accurate prediction of the performance gain and makes the determination of the optimal rotation angle easier. By comparing with alternative techniques including the Walsh-Hadamard transform and Golden and Silver Codes, using numerical simulations, we show that Pairwise Coding is a good candidate to be integrated into current 100 Gb/s, future 400 Gb/s and 1 Tb/s digital coherent transceivers.

2. Pairwise coding theory and design

2.1 Pairwise pre-coding

Figure 1 shows the structure of polarization pairwise pre-coding. Firstly, two data streams are mapped to quadrature amplitude modulation (QAM) symbols \( X_n = a_n + b_n j \) and \( Y_n = c_n + d_n j \), and then a constant phase shift, \( \theta \), is applied to all symbols, which leads to...
After this angular rotation, I/Q component interleaving is used to generate the transmitted signals for each polarization:

\[
\begin{align*}
TX_n &= \Re\{X_{\theta,a}\} + \Re\{Y_{\theta,b}\} = a_n \cos \theta - b_n \sin \theta + (c_n \cos \theta - d_n \sin \theta) j \\
TY_n &= \Im\{X_{\theta,a}\} + \Im\{Y_{\theta,b}\} = a_n \sin \theta + b_n \cos \theta + (c_n \sin \theta + d_n \cos \theta) j
\end{align*}
\]

(1)

where \(\Re(*)\) and \(\Im(*)\) represent the real and imaginary parts of the signals, respectively. The three constellation diagrams in Fig. 1 indicate the evolution of the transmitted data symbols with QPSK modulation and \(\theta = 45^\circ\).

The angular rotation and I/Q interleaving can also be described in matrix form as:

\[
\begin{bmatrix}
\Re(TX_n) \\
\Re(TY_n)
\end{bmatrix} = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
a_n & c_n \\
b_n & d_n
\end{bmatrix}
\]

(2)

It is clear that the transform matrix \(\begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}\) is an orthogonal matrix that can be easily inverted, therefore such a coding scheme will not introduce extra sensitivity penalty compared to conventional QAM modulation if there is no PDL. When PDL is present, intuitively, after receiver I/Q de-interleaving, the signal to noise ratio (SNR) difference between two polarizations is converted into the SNR difference between the I and Q components of each polarization, and then the angular rotation maximizes the system performance. The optimal angle estimator is based on a minimum-error-rate search and its trade-off for practical implementation considerations will be presented in the later section.

2.2 Channel model and receiver decoding process

After single-mode fiber transmission, the frequency representation of the received signals can be described by the following equation:

\[
\begin{bmatrix}
RX(f) \\
RY(f)
\end{bmatrix} = H_{\text{linear}}(f) \cdot H_{\text{PDL}}(f) \cdot \begin{bmatrix}
TX(f) \\
TY(f)
\end{bmatrix} + \begin{bmatrix}
NX(f) \\
NY(f)
\end{bmatrix}
\]

(3)

where \(H_{\text{linear}}\) represents the combination of system linear transfer functions including: chromatic dispersion (CD), high-order polarization mode dispersion (PMD), and all electrical and optical filtering effects. \(NX\) and \(NY\) are the complex noise samples in the X and Y polarizations, which are a combination of optical amplified spontaneous emission (ASE), fiber nonlinearity noise and the background noise, including shot noise, electrical thermal noise and quantization noise. Furthermore, by applying the additive Gaussian model to the
nonlinear noise [18], we define the complex noise as Gaussian distributed with zero mean and variance $2\sigma^2$. $\mathbf{H}_{\text{PDL}}$ is governed by the following equation [5]:

$$\mathbf{H}_{\text{PDL}} = \prod_{i=1}^{N} \mathbf{R}_{\alpha,i} \left( \begin{bmatrix} \sqrt{1-\gamma} & 0 \\ 0 & \sqrt{1+\gamma} \end{bmatrix} \right) \mathbf{R}_{\beta,i}$$

(4)

where there are, in total, $N$ PDL elements contributing to the total PDL. $\mathbf{R}_{\alpha,i}$ and $\mathbf{R}_{\beta,i}$ are two random rotation matrices with uniform distribution in the range $[0, 2\pi]$, $\gamma$ defines the actual PDL for each PDL element, where $PDL_i = \frac{1+\gamma_i}{1-\gamma_i}$. Typically PDL causes one polarization to have a higher OSNR (therefore SNR) than the other, unless the two polarizations share the same loss. The overall system performance is then dominated by the polarization with the lowest SNR. We can combine the PDL rotation matrices with linear $\mathbf{H}$ to rewrite Eq. (1) as:

$$\begin{bmatrix} \mathbf{RX}(f) \\ \mathbf{RY}(f) \end{bmatrix} = \mathbf{H}_{\text{linear}}(f) \begin{bmatrix} \sqrt{1-\eta} & 0 \\ 0 & \sqrt{1+\eta} \end{bmatrix} \begin{bmatrix} \mathbf{TX}(f) \\ \mathbf{TY}(f) \end{bmatrix} + \begin{bmatrix} \mathbf{NX}(f) \\ \mathbf{NY}(f) \end{bmatrix}$$

(5)

where $\eta$ is the power factor and $-1 < \eta < 1$. The actual PDL value for each PDL element can be considered as a constant; however, due to the random rotation matrices, the principal state of the lossy polarization will evolve between the X and Y polarizations, and so $\eta$ will vary over time. Fortunately, this variation is much slower than the symbol rate, e.g. kHz [19] vs. 28 Gbaud, so it is appropriate to perform analysis over tens of thousands of symbols based on the same $\eta$ value. Therefore, in this equation, we assume that the signals have a constant polarization state within a short time frame, therefore the same PDL characteristics. This assumption allows the PDL to be separated from the PMD matrix. A more general expression that encapsulates all of the polarization effects can be found in [20]. We will show later that our scheme can adapt to random changes of the signal polarization state.

After channel equalization, and frequency and phase correction, the normalized recovered time-domain signals can be described as [2]:

$$\begin{bmatrix} \mathbf{EX}_n \\ \mathbf{EY}_n \end{bmatrix} = \begin{bmatrix} \mathbf{TX}_n \\ \mathbf{TY}_n \end{bmatrix} + \begin{bmatrix} \mathbf{NX}_n / \sqrt{1-\eta} \\ \mathbf{NY}_n / \sqrt{1+\eta} \end{bmatrix}$$

(6)

where: $\mathbf{EX}_n$ and $\mathbf{EY}_n$ are the equalized signals, $\mathbf{NX}_n / \sqrt{1-\eta}$ and $\mathbf{NY}_n / \sqrt{1+\eta}$ are the filtered noise samples with both zero mean and variance of $2\sigma^2 / (1-\eta)$ and $2\sigma^2 / (1+\eta)$, respectively. The noise terms are primed because of the receiver optical and electrical filters. In this case, the total SNR is given by $1/\sigma^2$ (assuming QPSK modulation with average symbol power of 2, and the SNR difference between two polarizations is:

$$\Delta\text{SNR}_\text{diff} = 10\log_{10} \left( \frac{1+\eta}{1-\eta} \right)$$. Practically, $\Delta\text{SNR}$ has a random distribution [21], and with a total PDL value $L$, the highest and lowest possible SNR differences are $L$ and 0, corresponding to the worst and best case PDLs, respectively. In the following, we focus on the system performance for different $\Delta\text{SNR}$ values. The performance improvement with $\Delta\text{SNR} > 0$ can be thought of as an increased tolerance of PDL.

The pairwise decoding process is shown in Fig. 2: the SNR estimation is first performed for each polarization for $\eta$ estimation using the statistical moments method [22], and then the equalized signals are rescaled accordingly, which essentially balances the noise variances of
both polarizations. After I/Q de-interleaving, the resulting real and imaginary parts on each polarization now have different SNR levels:

\[
\begin{align*}
\mathbf{E}_x & = \sqrt{1-\eta} (a_c \cos \theta - b_c \sin \theta) + \Re \{N_x\} + \sqrt{1+\eta} (a_a \sin \theta + b_a \cos \theta) + \Re \{N_y\} \\
\mathbf{E}_y & = \sqrt{1-\eta} (c_c \cos \theta - d_c \sin \theta) + \Im \{N_x\} + \sqrt{1+\eta} (c_a \sin \theta + d_a \cos \theta) + \Im \{N_y\}
\end{align*}
\]

Maximum likelihood detection (MLD) is then applied for symbol decision:

\[
\begin{align*}
\mathbf{D}_x & = \arg\min_{C_i} \left\{ | \mathbf{E}_{\theta_n} - \zeta_k |^2 \right\}, \quad \mathbf{D}_y = \arg\min_{C_i} \left\{ | \mathbf{E}_{\theta_n} - \zeta_k |^2 \right\},
\end{align*}
\]

where \( \zeta_k \), \( k = \{1,2,3,4\} \) are the rotated and rescaled constellation points, and \( C_i \) are the conventional data symbols, i.e. \( C_k = \{1+1j, 1-1j, -1+1j, -1-1j\} \) for \( k = \{1,2,3,4\} \) with QPSK modulation.

The essential idea of Polarization Pairwise Coding is that by interleaving and de-interleaving the I/Q components at the transmitter and receiver, the PDL-induced SNR difference between two polarizations is translated into a SNR imbalance between the real and imaginary components of both polarizations, which then enables constellation rotation and rescaling to create an optimal decision region to achieve improved system performance.

3. Error rate and optimal rotation angle analysis

In the following sections we answer the questions: (a) what is the expected benefit that applying Pairwise Coding in the presence of PDL, and (b) which rotation angle leads to the largest benefit?

3.1 Analytical SER and BER

As described in Eq. (7), in each polarization, one of the I and Q components will have a worse SNR than the other, but as both polarizations have similar overall characteristics, we need only analyze one polarization. Also for the sake of simplicity, we use QPSK as the modulation format. As shown in Fig. 3, we need to compare the four Euclidean distances \( (D_k) \) between a given signal and \( \zeta_k \), and make the decision based on the one that gives the minimum value.
The SER for pairwise coded QPSK signals can be calculated as:

$$\text{SER}_{\text{pairwise QPSK}} = \sum_{i=1}^{4} 0.25 \times \text{Prob}(D_{x} \neq C_{i} | X_{n} = C_{i})$$

$$= 0.5 \times \text{Prob}(D_{x} \neq C_{i} | X_{n} = C_{i}) + 0.5 \times \text{Prob}(D_{x} \neq C_{i} | X_{n} = C_{i})$$

$$+ 0.5 \times (\text{Prob}(D_{x} > D_{y} | X_{n} = C_{i}) + \text{Prob}(D_{x} > D_{z} | X_{n} = C_{i}) + \text{Prob}(D_{x} > D_{z} | X_{n} = C_{i}))$$

where the case of the signal having two or three distances that are simultaneously smaller than the desired decision point is ignored since it is negligible when the SER is less than 0.01, and $$D_{k} = |\text{EX}_{\eta, n} - \zeta_{n}|^2$$, the calculation of $$\text{Prob}(D_{x} > D_{y} | X_{n} = C_{i})$$ can be shown as:

$$D_{k} = \left[\Re(NX_{x}) + \text{Re}(NY_{x})\right]^2 + \left[\Re(NX_{x}) - \text{Re}(NY_{x})\right]^2$$

$$D_{k} = -2\sqrt{1 - \eta} \sin \theta + \Re(NX_{x}) + \left(2\sqrt{1 + \eta} \cos \theta + \text{Re}(NY_{x})\right)$$

$$= \left(4(1 - \eta) \sin \theta - 4\sqrt{1 - \eta} \sin \theta \Re(NX_{x}) + \Re(NX_{x})\right)$$

$$+ \left(4(1 + \eta) \cos \theta + 4\sqrt{1 + \eta} \cos \theta \Re(NY_{x}) + \Re(NY_{x})\right)$$

$$\text{Prob}(D_{x} > D_{y} | X_{n} = C_{i})$$

$$= \text{Prob}(4\sqrt{1 + \eta} \cos \theta \Re(NY_{x}) - 4\sqrt{1 - \eta} \sin \theta \Re(NX_{x}) < -4(1 - \eta) \sin \theta - 4(1 + \eta) \cos \theta)$$

$$= 0.5 \left[1 + \text{erf} \left(\frac{-1 + \eta \left(\sin \theta - \cos \theta\right)}{2 + 2\eta \left(\cos \theta - \sin \theta\right)}\right)\right]$$

With similar operations, the SER of pairwise coded QPSK signals can be written as:

$$\text{SER}_{\text{pairwise QPSK}}$$

$$= 0.5 \left[1 + \text{erf} \left(\frac{-1 + \eta \left(\sin \theta - \cos \theta\right)}{2 + 2\eta \left(\cos \theta - \sin \theta\right)}\right)\right] + 0.5 \left[1 + \text{erf} \left(\frac{-1 + \eta \left(\cos \theta - \sin \theta\right)}{2 + 2\eta \left(\sin \theta - \cos \theta\right)}\right)\right]$$

$$+ 0.25 \left[1 + \text{erf} \left(\frac{-\sqrt{1 + 2\eta} \sin \theta \cos \theta}{\sigma}\right)\right] + 0.25 \left[1 + \text{erf} \left(\frac{-\sqrt{1 - 2\eta} \sin \theta \cos \theta}{\sigma}\right)\right]$$

Equation (11) can be simplified when $$\theta = 45^\circ$$. 

---

#246126 Received 16 Jul 2015; revised 23 Sep 2015; accepted 5 Oct 2015; published 9 Oct 2015

© 2015 OSA 19 Oct 2015 | Vol. 23, No. 21 | DOI:10.1364/OE.23.027434 | OPTICS EXPRESS 27440
As a comparison, without Pairwise Coding, the average SER can be approximated to:

\[
SER_{\text{QPSK}} = 0.5 \left( 1 - \text{erf} \left( \frac{1 - \eta}{2\sigma^2} \right) \right) + 0.5 \left( 1 - \text{erf} \left( \frac{1 + \eta}{2\sigma^2} \right) \right).
\] (13)

The optimal bit-to-symbol mapping scheme for pairwise coded signals is still an open question, because it depends on the actual SNR, \( \Delta \text{SNR} \) and rotation angle. Assuming Gray coding, the BER can be described by simply modifying Eq. (11) to:

\[
\begin{align*}
\text{BER}_{\text{QPSK}} &= 0.25 \left( 1 + \text{erf} \left( -\frac{1 + \eta (\sin^i \theta - \cos^i \theta)}{\sqrt{2 + 2\eta (\cos^i \theta - \sin^i \theta)\sigma}} \right) \right) + 0.25 \left( 1 + \text{erf} \left( -\frac{1 + \eta (\cos^i \theta - \sin^i \theta)}{\sqrt{2 + 2\eta (\sin^i \theta - \cos^i \theta)\sigma}} \right) \right) \\
&+ 0.25 \left( 1 + \text{erf} \left( -\frac{1 + 2\eta \sin \theta \cos \theta}{\sigma} \right) \right) + 0.25 \left( 1 + \text{erf} \left( -\frac{1 - 2\eta \sin \theta \cos \theta}{\sigma} \right) \right).
\end{align*}
\] (14)

It is important to note that the rotation angle that minimizes the SER will not necessarily minimize the BER performance when Gray coding is used. On the other hand, the BER for conventional PDM-QPSK can be approximated by \( \text{BER}_{\text{QPSK}} = \frac{SER_{\text{QPSK}}}{2} \).

3.2 Optimal rotation angle

The optimal rotation angle for QPSK modulation has been derived based on maximizing the mutual information between two information vectors [15]:

\[
\theta_{\text{opt}} = \begin{cases} 
\pi / 4 & \Delta \text{SNR} \leq 3 \\
\tan^{-1} \left( (\Delta \text{SNR} - 1) - \sqrt{(\Delta \text{SNR} - 1)^2 - \Delta \text{SNR}} \right) & \Delta \text{SNR} > 3
\end{cases}
\] (15)

where knowledge of \( \Delta \text{SNR} \) is required for this approach. Note that, because the signal’s SOP evolves, the receiver’s \( \Delta \text{SNR} \) estimation is only accurate within a group of symbols that have similar polarization states. Because the SOP rotation rate is much lower than the baud rate, it is advantageous to estimate the SNR, so find the optimal \( \theta \), over more symbols with larger symbol rate. An alternative solution is to test across a number of \( \theta \) candidates using Eqs. (11) and (14) to find the one that provides the minimum SER or BER; however, this approach requires information about \( \Delta \text{SNR} \) and also the total SNR. The advantage of the SER or BER searching method is that it can be easily extended to high-order QAM modulation formats, whereas only the solution for QPSK is provided in [15].

Figure 4 plots the derived optimal angles based on Eq. (15), SER and BER searching, using a fine search resolution of 0.045°. In Fig. 4(a), at \( \text{SNR} = 12 \text{ dB} \), the difference between the theoretical and SER searching curves is more than 5° for \( \Delta \text{SNR} > 5 \text{ dB} \). This difference is reduced by about 50% when Gray-coded BER searching is used, because the minimum SER does not correspond to the minimum BER. Figure 4(b) shows that the SER and BER searching methods estimate different optimal rotation angles over a range of SNRs; for example, there is a 13°-difference at 5-dB SNR. At high SNRs the two methods converge.
3.3 Simulation verifications

We verified the above derivations by conducting a single-channel PDM-QPSK system simulation using VPItransmissionMaker. The symbol rate is 12.5 GHz with 0.01 roll-off root raised cosine filtering. We simulated different $\Delta$SNR values by using a single PDL element with the lossy polarization aligned to the X-polarization of the PDM signals, and then followed this by an OSNR-setting module to define the OSNR (specified for a 0.1-nm noise bandwidth). After coherent reception, training-aided channel equalization [23] and pilot-aided maximum likelihood phase estimation [24] were used to recover the signals. Gray coding was applied for bit-to-symbol mapping.

Figure 5 plots BER and SER performances at 10-dB SNR, against rotation angle at a resolution of 4.5° for 3- and 9-dB $\Delta$SNR. For QPSK, the performance is periodic with rotation angle, with a period of 90°; thus, we need only sweep in a range of 0° to 90°. It is clear that the optimal rotation angle is the same for the minimum BER and the minimum SER at 3-dB $\Delta$SNR, because in this case the BER is almost half of the SER, similar to the normal QPSK case. When the $\Delta$SNR is increased to 9 dB, the optimal angles that lead to best BER and SER performance are slightly different. This illustrates that if we choose the rotation angle from the SER search, Gray coding is not the optimal bit-to-symbol mapping scheme. Also, different SNR values lead to different optimal rotation angles; therefore, to pursue the best system performance, we need to jointly consider the bit-to-symbol mapping scheme, system SNR (OSNR), and $\Delta$SNR when determining the optimum rotation angle.

Figure 6 shows the single-channel OSNR versus $\Delta$SNR, where the rotation angle is calculated based on either Eq. (15) or the BER and SER searching methods. Only results for BERs and SERs worse than $10^{-6}$ are shown, due to the limited number of symbols that can be simulated. For the BER results in Figs. 6(a)-6(d), for $\Delta$SNRs of 0 and 3 dB, the optimal rotation angle estimated by both methods is the same; for $\Delta$SNRs of 6 and 9 dB, the two optimal rotation angle calculation methods achieve a similar performance. With increased

Received 16 Jul 2015; revised 23 Sep 2015; accepted 5 Oct 2015; published 9 Oct 2015

© 2015 OSA

19 Oct 2015 | Vol. 23, No. 21 | DOI:10.1364/OE.23.027434 | OPTICS EXPRESS 27442
ΔSNRs between the two polarizations, Pairwise Coding provides a larger performance enhancement compared with the unpaired signals – the pairwise coded signals require about 1/2.5/4 dB lower OSNR to achieve BERs of $10^{-3}$ with 3/6/9 dB ΔSNR. There is no performance penalty when ΔSNR is zero, i.e. Pairwise Coding achieves same performance as unpaired signals if there is zero PDL, or if both polarizations are attenuated equally with best-case PDL (45° between the signal polarizations and the PDL lossy axis).

The BER results with 6-dB and 9-dB ΔSNR are shown in Figs. 6(a)-(d), respectively. Pairwise Coding requires about 3/4.5 dB lower OSNR to achieve an BER of $10^{-3}$ with 6/9 dB ΔSNRs; this again illustrates that ≈0.5-dB penalty is incurred due to the sub-optimal Gray bit-to-symbol mapping for Pairwise Coding. With all BER and SER results, the simulations agree well with the theoretical curves derived from Eqs. (14) and (11). By comparing Pairwise Coding with different optimal angle calculation methods, generally speaking, the optimal angle from Eq. (15) produces a similar performance as the SER and BER searching methods, but the magnified insets in Figs. 6(c) and 6(e) show that the SER/BER searching methods can achieve a minor improvement. This illustrates that, based on the optimal rotation angle calculated from Eq. (15), the system performance approximates to the theoretical optimal performance.

Overall, the simulation results verify that, based on the theoretical models, the performance gain given by Pairwise Coding can be predicted precisely.

Fig. 6. Simulation results for single-channel OSNR sweep (signal bandwidth 12.5 GHz) with various ΔSNR between two polarizations and different optimal rotation angle search schemes: (a)-(d) BER results for 0/3/6/9 dB SNR difference; (e), (f): SER results for 6/9 dB SNR difference.
To demonstrate the effectiveness of the proposed technique within a more practical scenario, we first simulated a single PDL element with 4-dB PDL, fix the system SNR to 10-dB, and then swept the angle between the input signals’ SOP and the PDL element from 0 to 360°. As shown in Fig. 7(a), without Pairwise Coding, the good and bad polarizations switch between the X and Y axes every 45°. The overall system performance (black curve) also swings with the same period. With Pairwise Coding, the worst performance appears when the SNR is the same for both polarizations (the crossing points of blue and red curves): the best performance can be attained with largest ΔSNR. Therefore Pairwise Coding improves the lower bound of the achievable system performance to be the same as the uncoded system with ‘best case PDL’.

We then split the single 4-dB PDL element into four cascaded 1-dB PDL elements, with the same lossy axes and randomly varying signal SOPs between the four PDL elements. We collected 1000 simulation results. The uncoded and coded system performances are shown in Figs. 7(b) and 7(c), respectively. It is clear that Polarization Pairwise Coding improves the worst, best and average system performances significantly. It is also worth mentioning that, although we have only considered the system performance without forward error correction (FEC) codes, it has been reported that a combination of constellation rotation and FEC codes can achieve further system performance gains over fading wireless channels [25]; therefore, there may also be some potential benefits by combining advanced FEC techniques with Polarization Pairwise Coding.

Fig. 7. (a): Simulated performance with a single 4-dB PDL element, sweeping the angle between signals’ SOP and PDL lossy axis from 0 to 360°; (b), (c): unpaired and paired signal performance with 1000 simulation runs of four cascaded PDL elements, each with 1-dB PDL and the signals’ SOPs randomly rotated between the PDL elements.

4. Discussion on practical implementation issues

4.1 Choosing a fixed rotation angle

We have presented the approaches to identify the optimal rotation angle \( \theta \) in the previous section. However, the methods based on Eq. (15) and SER/Ber searching, need values of ΔSNR and sometimes the system SNR to be communicated from the receiver to the transmitter, which complicates the system architecture. Also, because the ΔSNR and system SNR are time varying, the rotation angle needs to updated adaptively to cope with the evolution of these two parameters. This is especially true for the ΔSNR since the state of polarization of PDM signals can change completely within one millisecond, requiring frequent feedback. Furthermore, different rotation angles generate different constellations after transmitter-side I/Q interleaving, which may set higher requirements for number of bits of the transmitter side DAC. The transmitted constellations that are based on optimal rotation angle calculation using Eq. (15) for 3, 6 and 9-dB ΔSNR are shown in Figs. 8(a), 8(b) and 8(c), respectively, indicating that at least a 2-bit DAC is required, rather than a 1-bit DAC for PDM-QPSK. Therefore, a fixed rotation angle is preferred, as it gives a reasonable gain over a wide range of ΔSNR and SNRs for simpler transmitted constellations.
As 45° is the optimal rotation angle for ΔSNR ≤ 4.7 dB, and has 9-QAM like constellation requiring only a 2-bit DAC, θ = 45° is attractive for Pairwise Coding. Moreover, the receiver digital channel equalization and phase estimation processes for 9-QAM format have been thoroughly investigated [26], thus the adaptive filter and phase estimator structure of conventional PDM-QPSK systems can be used, with small changes to the error-update algorithms. As the main feature of the θ = 45° solution is a simplified pre-coding/decoding process, it is worth comparing it to another well-known coding scheme, the Walsh-Hadamard transform (WHT), which also has simple encoding matrix with single-carrier PDM signals as:

$$\begin{bmatrix}
    TX_x \\
    TY_x
\end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix}
    1 & 1 \\
    1 & -1
\end{bmatrix} \begin{bmatrix}
    X_x \\
    Y_x
\end{bmatrix}.$$  \hspace{1cm} (16)

Since the inverse of the transform matrix is itself, and the normal QPSK decision region is maintained, the decoding process of WHT is even simpler than Pairwise Coding scheme. Indeed the WHT is very similar to Pairwise Coding with the θ = 45° case, except that Pairwise Coding exchanges the I of X/Y-polarization and the Q of Y/X-polarization, while WHT mixes all of the I and Q components of two polarizations. If we consider the ΔSNR as the worst-case PDL with a single PDL element, applying the WHT is equivalent to rotating the SOP of the incoming signals by θ = 45° before passing through the PDL element, which changes the worst-case PDL to the best-case PDL. In conclusion, the major difference between these two schemes is that Pairwise Coding not only balances the SNR between two polarizations, it also manipulates the constellation to maximize the coordinate diversity, while WHT only aims to equalize SNR.

We simulated the performance with three different coding schemes: (1) using θ = 45° for Pairwise Coding; (2) adaptively using optimal angle calculated by Eq. (15) for Pairwise Coding, and (3) using WHT coding scheme. Figure 9 shows the results. The normal PDM-QPSK performance is also shown as a reference. All three schemes show improved performance compare to the uncoded signals with worst-case PDL. For Pairwise Coding, clearly θ = 45° provides similar performance gain as the optimal angle case with less than 7-dB ΔSNR, and then the performance penalty for the θ = 45° case is about 2.8 dB with 12-dB ΔSNR. At the same time, the θ = 45° solution shows a consistent advantage over the Walsh-Hadamard transform method. The reason can be revealed by comparing the recovered constellations for three schemes (12 dB SNR and 10 dB ΔSNR), as shown in Figs. 9(b)-9(d). With θ = 45°, the symbol errors mainly come from the overlap between the two quadrants that sit on the real axis, and also in this case one symbol error translates to two bit errors due to Gray coding; while with optimal θ the four quadrants interact with each other quite in a similar fashion, which leads to better performance. In contrast, WHT only aims to equalize the SNR between two polarizations without altering the decision region, resulting in the least improvement, and it shows good agreement to the uncoded system performance with best case PDL. Note that at high ΔSNR, the performance with θ = 45° Pairwise Coding can be improved by optimizing the bit-to-symbol mapping scheme to halve the BER, e.g. coding the two interacting quadrants (1-1j and −1 + 1j) with only one bit difference (e.g. 10 01).
Fig. 9. (a) Performance comparison between 45° and optimal angle; recovered signal constellation 12 dB SNR and 10 dB ΔSNR when using: (b) θ = 45° and (c) optimal angle for Pairwise Coding, and (d) WHT coding scheme.

An extreme example to understand these three schemes is if there is no noise on Y-polarization (infinite ΔSNR). Here, Pairwise Coding with an optimal rotation angle will create constellations that all sit on imaginary axis, achieving infinite capacity. With the θ = 45° solution, part of the coordinate diversity is missed, which results in moderate performance. Using WHT, the system’s capacity is limited to being equivalent to having twice the SNR of the X-polarization, which is the lowest among the three schemes. We can therefore conclude that Pairwise Coding with θ = 45° provides large gains for a wide range of ΔSNR values, while greatly simplifying the system design.

4.2 Comparison with other polarization time codes

Besides WHT, there are also some other digital coding methods have been proposed to improve the PDL tolerance; for example, the Alamouti code, as a well-known code with best performance in 2 × 1 MIMO system, has been applied for PDL penalty mitigation [27]:

$$\begin{bmatrix} TX_{n-1} & TX_n \\ TY_{n-1} & TY_n \end{bmatrix} = \begin{bmatrix} X_{n-1} & -\text{conj}(Y_{n-1}) \\ Y_{n-1} & \text{conj}(X_{n-1}) \end{bmatrix}. \quad (17)$$

Unfortunately, 50% redundancy is required, which degrades the spectral efficiency. Golden and Silver Codes have also been proposed to improve the PDM system performance in the presence of PDL [8], the transfer matrix for Golden Code can be shown as [28]:

$$\begin{bmatrix} TX_{n-1} & TX_n \\ TY_{n-1} & TY_n \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} (1+j-\alpha)(X_{n-1}+\alpha X_n) & (1+j-\alpha)(Y_{n-1}+\alpha Y_n) \\ (1-j-\bar{\alpha})(Y_{n-1}+\bar{\alpha} Y_n) & (1+j-\bar{\alpha})(X_{n-1}+\bar{\alpha} X_n) \end{bmatrix}. \quad (18)$$

where $\alpha = \frac{1+\sqrt{5}}{2}$ and $\bar{\alpha} = \frac{1-\sqrt{5}}{2}$. For a Silver Code [29]:

$$\begin{bmatrix} TX_{n-1} & TX_n \\ TY_{n-1} & TY_n \end{bmatrix} = \begin{bmatrix} X_{n-1}+\bar{X}_n & -\text{conj}(Y_{n-1})-\text{conj}(\bar{Y}_n) \\ Y_{n-1}-\bar{Y}_n & \text{conj}(X_{n-1})-\text{conj}(\bar{X}_n) \end{bmatrix}$$

$$\begin{bmatrix} \bar{X}_n \\ \bar{Y}_n \end{bmatrix} = \frac{1}{\sqrt{7}} \begin{bmatrix} 1+j & -1+2j \\ 1+2j & 1-j \end{bmatrix} \begin{bmatrix} X_n \\ Y_n \end{bmatrix}. \quad (19)$$

The Golden Code achieves the best performance in 2 × 2 fading wireless system; therefore, this scheme may also lead to the best achievable result in PDM coherent optical system in the presence of PDL. It has been reported that the Silver Code achieves even better performance for PDL mitigation with reduced decoding complexity [8]; however, compared with Pairwise Coding, a large computational effort is required for the equalization and detection in these two
schemes, which may be impractical at high data rates. Since the Golden/Silver Codes encode across two time slots of two polarizations, the complexity for symbol detection is large. Unlike Pairwise Coding, which only requires symbol-by-symbol detection, Golden and Silver Codes need $2 \times 2$ matrix maximum likelihood sequence estimation over all possible signal combinations. For instance, with $M$-QAM modulation format, Pairwise Coding needs to compare between $M$ single-entry Euclidean distances for each symbol, while for Golden and Silver Codes, $M^4$ and $2M^3 \times 2$ matrix likelihood calculations are required to decode every four symbols, respectively. Therefore, compared with Golden/Silver Codes, Pairwise Coding is significantly less complex during equalization and symbol detection, so could be added into existing digital coherent optical systems.

7. Conclusions

In this paper, we have presented the design of a zero-overhead digital coding method, Pairwise Coding, to improve the PDL tolerance of PDM coherent optical systems. The pre-coding and decoding structure have been discussed in detail. We derived the analytical SER and BER for pairwise coded signals, which can accurately predict the performance gain and also can be used to find the optimal rotation angle for the transmitter side pre-coding. Simulation results illustrate the benefit of Pairwise Coding over a wide range of PDL values, which agree well with the analytical SER and BER models. Although the investigation of this paper is limited to QPSK modulation, the principle can be easily extended to higher-order QAM formats by constructing the analytical SER and BER analysis accordingly, to select the correct rotation angle for system performance enhancement. By comparing with other coding methods, we prove that using a fixed rotation angle, $\theta = 45^\circ$, greatly simplifies the system coding and decoding process while still providing a large performance gain. Thus, Pairwise Coding with a fixed rotation angle can be easily integrated into current and next-generation commercial digital coherent transceivers to give a beneficial performance gain.

Acknowledgments

We thank VPIphotonics (www.vpiphotonics.com) for the use of their simulator, VPItransmissionMakerWDM V9.1. This work is supported under the Australian Research Council’s Laureate Fellowship (FL130100041) scheme and under CUDOS – ARC Centre of Excellence for Ultrahigh bandwidth Devices for Optical Systems (CE110001018).