HAL
open science

Polarization-time coding for PDL mitigation in long-haul PolMux OFDM systems

Elie Awwad, Yves Jaouën, Ghaya Rekaya-Ben Othman

To cite this version:

Elie Awwad, Yves Jaouën, Ghaya Rekaya-Ben Othman. Polarization-time coding for PDL mitigation in long-haul PolMux OFDM systems. Optics Express, 2013, 21 (19), pp.22773-22790. hal-00868656

HAL Id: hal-00868656

https://hal.science/hal-00868656v1

Submitted on 31 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Polarization-time coding for PDL mitigation in long-haul PolMux OFDM systems

Elie Awwad,* Yves Jaouën and Ghaya Rekaya-Ben Othman

Institut Mines-Télécom, TELECOM ParisTech, CNRS UMR5141,
46 rue Barrault, 75634 Paris, France
**awwad@telecom-paristech.fr*

Abstract: In this paper, we present a numerical, theoretical and experimental study on the mitigation of Polarization Dependent Loss (PDL) with Polarization-Time (PT) codes in long-haul coherent optical fiber transmissions using Orthogonal Frequency Division Multiplexing (OFDM). First, we review the scheme of a polarization-multiplexed (PolMux) optical transmission and the 2×2 MIMO model of the optical channel with PDL. Second, we introduce the Space-Time (ST) codes originally designed for wireless Rayleigh fading channels, and evaluate their performance, as PT codes, in mitigating PDL through numerical simulations. The obtained behaviors and coding gains are different from those observed on the wireless channel. In particular, the Silver code performs better than the Golden code and the coding gains offered by PT codes and forward-error-correction (FEC) codes aggregate. We investigate the numerical results through a theoretical analysis based on the computation of an upper bound of the error probability of the optical channel with PDL. The derived upper bound yields a design criterion for optimal PDL-mitigating codes. Furthermore, a transmission experiment of PDL-mitigation in a 1000km optical fiber link with inline PDL validates the numerical and theoretical findings. The results are shown in terms of Q-factor distributions. The mean Q-factor is improved with PT coding and the variance is also narrowed.

© 2013 Optical Society of America

OCIS codes: (060.0060) Fiber optics and optical communications; (060.1660) Coherent communications; (060.4080) Modulation; (060.4230) Multiplexing.

References and links

1. P. J. Winzer, "High-Spectral-Efficiency Optical Modulation Formats," J. Lightwave Technol. **30**(24), 3824–3835 (2012).
2. S. Savory, "Digital filters for coherent optical receivers," Opt. Express **16**, 804–817 (2008).
3. S. L. Jansen, I. Morita, T. C. W. Schenk, and H. Tanaka, "121.9-Gb/s PDM-OFDM Transmission With 2-b/s/Hz Spectral Efficiency Over 1000 km of SSMF," J. Lightwave Technol. **27**(3), 177–188 (2009).
4. J. P. Gordon, H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," in Proc. Natl. Acad. Sci. U.S.A. **97**(9), 4541–4550 (2000).
5. T. Duthel, C. R. S. Fludger, J. Geyer, and C. Schulien, "Impact of polarization dependent loss on coherent POLMUX-NRZ-DQPSK," in proc. of OFC/NFOEC'08, 1–3.
6. A. Juarez, C. Bunge, S. Warm, and K. Petermann, "Perspectives of principal mode transmission in mode-division-multiplex operation," Opt. Express **20**(13), 13810–13824 (2012).
7. W. Shieh, X. Yi, Y. Ma, and Y. Tang, "Theoretical and experimental study on PMD-supported transmission using polarization diversity in coherent optical OFDM systems," Opt. Express **15**(16), 9936–9947 (2007).

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013

8. S. Mumtaz, G. Rekaya, Y. Jaouën, "Space-Time codes for optical fiber communication with polarization multiplexing," in proc. of ICC'10, 1–5.
9. S. Mumtaz, J. Li, S. Koenig, Y. Jaouën, R. Schmogrow, G. Rekaya-Ben Othman, and J. Leuthold, "Experimental demonstration of PDL mitigation using Polarization-Time coding in PDM-OFDM systems," in proc. of SPPCom'11, paper SPWB6.
10. S. Mumtaz, G. Rekaya-Ben Othman, Y. Jaouën, J. Li, S. Koenig, R. Schmogrow, and J. Leuthold, "Alamouti code against PDL in polarization multiplexed systems," in proc. of SPPCom'11, paper SPTuA2.
11. E. Meron, A. Andrusier, M. Feder, and M. Shtaif, "Use of space-time coding in coherent polarization-multiplexed systems suffering from polarization-dependent loss," Opt. Letters **35**(21), 3547–3549 (2010).
12. V. Tarokh, A. Naguib, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths," IEEE Transactions on Communications **47**(2), 199–207 (1999).
13. J.-C. Belfiore, G. Rekaya, E. Viterbo, "The golden code: a 2x2 full-rate space-time code with nonvanishing determinants," IEEE Transactions on Information Theory **51**(4), 1432–1436 (2005).
14. H. Bulow, W. Baumert, H. Schmuck, F. Mohr, T. Schulz, F. Kuppers, and W. Weiershausen, "Measurement of the maximum speed of PMD fluctuation in installed field fiber," OFC/IOOC'99 **2**, 83–85.
15. S. Ben Rayana, H. Besbes, G. Rekaya-Ben Othman, and Y. Jaouën, "Joint equalization and polarization-time coding detection to mitigate PMD and PDL impairments," in proc. of SPPCom'12, paper SpW2B.3.
16. A. Andrusier, E. Meron, M. Feder, and M. Shtaif, "An optical implementation of a space-time-trellis code for enhancing the tolerance of systems to polarization-dependent loss," Opt. Letters **38**(2), 118–120 (2013).
17. S. R. Desbruslais, P. R. Morkel, "Simulation of polarisation mode dispersion and its effects in long-haul optically amplified lightwave systems," IEE Colloquium on International Transmission Systems, 6/1–6/6 (1994).
18. O. Vassilieva, I. Kim, Y. Akasaka, M. Bouda, and M. Sekiya, "Interplay between PDL and nonlinear effects in coherent polarization multiplexed systems," Opt. Express **19**(26), B357–B362 (2011).
19. A. Mecozzi, M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," IEEE Photonics Technol. Lett. **14**(3), 313–315 (2002).
20. N. Gisin, "Statistics of polarization dependent loss," Optics Communications **114**, Elsevier (1995).
21. L. Nelson, C. Antonelli, A. Mecozzi, M. Birk, P. Magill, A. Schex, and L. Rapp, "Statistics of polarization dependent loss in an installed long-haul WDM system," Opt. Express **19**(7), 6790–6796 (2011).
22. A. Lima, I. Lima Jr., C. Menyuk, and T. Adali, "Comparison of penalties resulting from first-order and all-order polarization mode dispersion distortions in optical fiber transmission systems," Opt. Letters **28**(5), 310–312 (2003).
23. W. Shieh, "PMD-Supported Coherent Optical OFDM Systems," IEEE Photonics Technol. Lett. **19**(3), 134–136 (2007).
24. X. Liu, F. Buchali, "Intra-symbol frequency-domain averaging based channel estimation for coherent optical OFDM," Opt. Express **16**(26), 21944–21957 (2008).
25. M. Shtaif, "Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss," Opt. Express **16**(18), 13918–13932 (2008).
26. J. Proakis, M. Salehi, *Digital Communications*, Fifth Edition. Mc Graw - Hill International Edition (2008).
27. P. Delesques, E. Awwad, S. Mumtaz, G. Froc, P. Ciblat, Y. Jaouën, G. Rekaya, and C. Ware, "Mitigation of PDL in coherent optical communications: How close to the fundamental limit?," in proc. of ECOC'12, paper P4.13.
28. C. Xie, "Polarization-dependent loss induced penalties in PDM-QPSK coherent optical communication systems," in proc. of OFC/NFOEC'10, 1–3.
29. E. Awwad, Y. Jaouën, G. Rekaya-Ben Othman, and E. Pincemin, "Polarization-Time Coded OFDM for PDL Mitigation in Long-Haul Optical Transmission Systems," in proc. of ECOC'13, paper P3.4 (to be published).
30. E. Awwad, Y. Jaouën, and G. Rekaya-Ben Othman,"Improving PDL Tolerance of Long-Haul PDM-OFDM Systems Using Polarization-Time Coding," in proc. of SPPCom'12, paper SpTu2A.5.

1. Introduction

Taking advantage of the numerous degrees of freedom in an optical fiber is a key solution to the exponentially increasing demand of capacity [1] over long-haul and metropolitan optical fiber networks. Over the last years, coherent detection [2] replaced direct detection schemes offering an enhanced receiver sensitivity and introducing multilevel modulation formats. Besides, wavelength division multiplexing (WDM) enabled the transmission of many independent wavelengths over the same fiber providing scalability and an important capacity growth to the optical networks. Furthermore, coherent polarization division multiplexed (PDM) systems [1,3] were considered to double the capacity of the optical link by sending two independent signals on two orthogonal polarization states, thus using the amplitude, phase and polarization state of

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

the transmitted signal. Moreover, in order to cope with the increasing number of internet users and the outgrowth of new bandwidth-demanding services such as cloud computing and real-time multimedia services, emerging research projects are studying new physical dimensions in the fiber such as space-division multiplexing (SDM) that consists in using parallel transmission paths per wavelength [1], ideally multiplying the capacity of the link by the number of paths.

However, the increase in the spectral efficiency of a multiplexed optical transmission comes at the detriment of a reduction in performance of the transmission scheme. This is mainly due to coupling effects and differential group delays such as polarization mode dispersion and rotations of the polarization states in PDM systems [4], as well as differential losses between the multiplexed channels such as polarization dependent loss (PDL) in PDM systems [5] and mode dependent loss (MDL) in SDM systems [6]. These effects can be modeled using a multi-input multi-output (MIMO) formalism [7] which paves the way to the adaptation of mature digital signal processing (DSP) tools, initially designed for wireless MIMO transmissions and the search for new MIMO techniques specific for the optical fiber channel.

In this paper, we will study the simplest multiplexed scheme which consists in a 2×2 polarization-multiplexed (PolMux) MIMO system. The implementation of these systems was made possible with the use of DSP algorithms and the development of high-speed electronics to recover the transmitted data at the receiver side in the electrical domain. The impairments undergone by the high bitrate transmitted signal can be categorized into two classes: linear and nonlinear. The major linear impairments are chromatic dispersion (CD), polarization-mode dispersion (PMD), and polarization-dependent loss (PDL) that can all be modeled using 2×2 transfer matrices (Jones matrices) straightforwardly leading to a 2×2 MIMO system representation. As for nonlinear effects, they will not be considered in the numerical and theoretical studies where we assess the penalties induced by the linear impairments. However, in the experimental study, we will observe the limitations caused by non-linear effects and determine the optimal operating point of the transmission (in terms of optimal launched signal power).

Dispersive effects preserve the total energy of the transmitted signal and the equalization of CD and PMD is very effective in a single-carrier context [2] and in multi-carrier systems such as orthogonal frequency division multiplexing (OFDM) based systems [7]. On the other hand, PDL remains the main limiting linear impairment. In current systems (both coherent and non-coherent), PDL is not mitigated and system margins are considered to absorb the induced penalties and ensure a target performance. In long-haul optical links, PDL is introduced by imperfect optical elements that unequally attenuate the polarization states of the incident signal and induce energy loss and optical signal-to-noise ratio (OSNR) fluctuations. Recently [8–11], space-time (ST) codes were suggested to mitigate PDL. In wireless MIMO channels, ST coding improves the reliability of the transmission by sending multiple copies of a data symbol to the receiver on different antennas and at different time slots expecting that the symbol will survive the impairments of the channel in a good enough state to allow reliable decoding [12]. The performance of ST codes, or more appropriately Polarization-Time (PT) codes in this context, such as the Golden code [13], the Silver code and the Alamouti code were numerically evaluated on an optical channel with PDL [8]. A preliminary experimental demonstration validated the numerical results [9]. It was found that the Silver code performed better than the Golden code, unlike the case of a wireless Rayleigh fading channel. Moreover, the performance of the Alamouti code was found to be independent of the amount of PDL in the link [10]. In this work, we will recall these recent numerical results, theoretically explain them and demonstrate the possibility to implement PT codes in a PolMux OFDM system and their potential in mitigating PDL through a transmission experiment.

Accordingly, we start by defining the transmission scheme of a PT-coded coherent optical OFDM system in Section 2 as well as defining the channel model of an optical link with PDL.

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

In Section 3, we present the simulation results of the investigated ST codes showing their performance and the coding gains in a linear regime. We will also examine the achieved coding gains when a forward-error-correcting (FEC) code is added to the transmission scheme. Afterwards, in Section 4, we derive an upper bound of the error probability of an optical link with PDL to interpret the simulated performance and check whether the used coding schemes are optimal or better codes can be found. Finally, in Section 5, we experimentally validate the numerical and theoretical results in the previous sections and also demonstrate the ability of PT codes to mitigate in-line PDL in a 1000km-long optical link.

2. PT-coded coherent OFDM system model

In this section, we describe the general transmission scheme in which PT codes will be implemented. We also define the channel model of an optical transmission with PDL that will be considered in the simulations and the theoretical study.

2.1. PolMux coherent OFDM transmission scheme

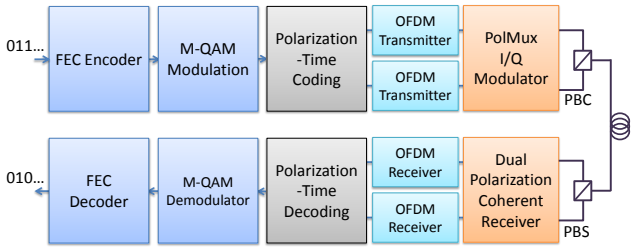


Fig. 1. General scheme of a PolMux OFDM transmission with Polarization-Time coding.

2.1.1. Transmitter side

In a conventional PolMux OFDM transmission [7], two independently modulated OFDM signals where each OFDM subcarrier carries an M-QAM symbol, are sent on two orthogonal polarization states; while in a PT-coded OFDM system, the modulated polarization states are correlated and carry a combination of M-QAM symbols. Indeed, PT coding consists in sending a combination of different modulated symbols on two polarization states (Pol_1,Pol_2) during several time slots. Hence, if we consider two consecutive time slots (T_1,T_2), the PT codeword matrix $\mathbf{X}$ will be:

$$\mathbf{X} = \begin{bmatrix} X_{Pol_1,T_1} & X_{Pol_1,T_2} \\ X_{Pol_2,T_1} & X_{Pol_2,T_2} \end{bmatrix} \tag{1}$$

This operation of combining modulated symbols is represented by the block "Polarization-Time coding" in Fig. 1. The codeword matrices of the investigated coding schemes will be examined in the next section. The first row of $\mathbf{X}$ modulates the first polarization state and the second modulates the second orthogonal polarization state. The columns of $\mathbf{X}$ will be carried by the same subcarrier of two consecutive OFDM symbols. After the assignment of symbols to the different subcarriers, conventional OFDM processing is realized including an inverse Fast-Fourier Transform (iFFT) and the insertion of a well-dimensioned cyclic prefix (CP) to absorb all CD and PMD [7]. The main advantage of choosing OFDM to deal with dispersion instead of single-carrier based solutions is the reduced complexity of the equalization step, at the receiver side. Indeed, the use of a suitable CP allows to replace the multi-tap equalizer usually used to extract the useful symbol from the interfering symbols, with a single-tap frequency

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

domain equalizer needed to correct any phase or amplitude distortion induced by the channel and separate the two polarizations. Next, the signal is up-converted to the optical domain and a polarization beam combiner combines the OFDM signals on two orthogonal polarization states.

2.1.2. Receiver side

At the receiver side, a polarization beam splitter splits the incident optical signal on two orthogonal states and a dual polarization coherent receiver down-converts the signal to the electrical domain. Next, OFDM processing is carried including CP removal and an FFT. Each OFDM subcarrier sees a non-dispersive channel and the received symbol is given by [7]:

$$\mathbf{Y}_{k,i} = \mathbf{H}_k(\omega_k)\mathbf{X}_{k,i} + \mathbf{N}_{k,i} \tag{2}$$

where $\mathbf{X}_{k,i}$ is the 2×2 matrix of transmitted symbols on the kth subcarrier ($k = 1 \ldots n$) of the ith OFDM symbol, $\mathbf{Y}_{k,i}$ is the 2×2 matrix of the received symbols. $\mathbf{H}_k(\omega_k)$ is the 2×2 channel matrix including laser phase noise, CD, PMD and PDL. $\mathbf{N}_{k,i}$ is a 2×2 matrix representing the additive noise. A training sequence known at the receiver is used to estimate the channel matrix for each subcarrier [3]. Then, the common phase error induced by the laser phase noise is corrected and finally the transmitted data symbols can be detected and demodulated. It is important to note that ST codes were designed to bring coding and diversity gains to the MIMO scheme when they are optimally decoded according to the maximum likelihood (ML) criterion instead of using sub-optimal linear decoders. Considering the channel model in Eq. (2) and assuming all codeword matrices $\mathbf{X}$ equiprobable, an ML decoder estimates the codeword $\mathbf{X}$ with $\mathbf{X}'$ according to the following criterion:

$$\mathbf{X}' = \underset{\mathbf{X} \in \mathscr{C}}{argmin} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 \tag{3}$$

where $\mathscr{C}$ is the set of all possible codewords. The indices k and i were dropped for clarity. The PT decoding can only be performed under the assumption of a constant $\mathbf{H}$ during the codeword duration (2 time slots in this case) which is the case of the optical channel [14]. ML decoding is performed by an exhaustive research with reasonable complexity when the considered modulation symbols come from a 4-QAM constellation. We point out that the decoding complexity is quite reduced with the use of OFDM at the cost of an increased overhead. PT coding could have been implemented in a single-carrier context with time domain equalization [15, 16]. However, multi-taps channel matrices would be required increasing the decoding complexity.

2.2. *Long-haul optical link*

In this section, we focus on the properties of the long-haul optical channel. We consider a multi-span optical link with N_S spans where each span contains a fiber modeled as a concatenation of random PMD elements and inline components with PDL as seen in Fig. 2. Each span is followed by an Erbium doped fiber amplifier (EDFA) operating in constant power mode to raise the signal power to the initially injected power at the transmitter. These amplifiers also add noise to the signal caused by amplified spontaneous emission (ASE). The accumulated ASE noise at the end of the link is the dominant noise source in the whole transmission scheme. In the following, the channel model of the link will be described and simplified to carry the numerical and theoretical studies of PDL mitigation with PT codes. We will, in particular, look at the statistics of PDL, its frequency dependence and the noise properties at the receiver side. We do not consider CD and phase noise because these two effects are polarization independent.

Each fiber span is modeled with a concatenation of M random PMD elements where each element has a transfer matrix $\mathbf{H}_{PMD}(\omega_k)$ defined as the product of a delay matrix and a rotation

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

matrix as in [17]. These unitary effects cause no loss of energy to the transmitted signal unlike PDL that induces OSNR fluctuations. The transfer matrix of a PDL element is given by:

$$\begin{aligned}\mathbf{H}_{PDL} &= \mathbf{R}_\alpha \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\varepsilon} \end{bmatrix} \mathbf{R}_\alpha^{-1} \\ &= \frac{1}{\sqrt{1+\gamma}} \mathbf{R}_\alpha \begin{bmatrix} \sqrt{1+\gamma} & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \mathbf{R}_\alpha^{-1}\end{aligned} \tag{4}$$

The diagonal matrix gives the imbalanced attenuation values of the least and most attenuated polarization states and $\mathbf{R}_\alpha$ is a random rotation matrix. α is uniformly drawn in $[0:2\pi]$. ε and γ are defined through $\Gamma_{dB} = 10log_{10}\frac{1+\gamma}{1-\gamma} = -10log_{10}(\varepsilon)$ where $\Gamma_{dB} \geq 0$ is the PDL coefficient in dB (or simply referred to as PDL) and consists of the ratio between the highest and the lowest losses. The normalization factor in the model using the variable γ is commonly dropped in literature [5, 18] because the diagonal matrix and the rotation matrices are enough to take into account the OSNR inequality between the polarization states as well as their crosstalk.

The individual PDL value of each inline component is kept as low as possible , i.e.$0 \leq \gamma \ll 1$. However, many of these components are interspersed in long-haul optical links leading to a significant accumulated PDL. An overview of previous works on PDL, especially by Mecozzi et al. [19] and Gisin [20], shows statistical descriptions of PDL. In [19], it is found that the probability density function of the global Γ_{dB} of an optical link is a Maxwellian distribution when we consider many low-PDL components. Yet, field measurements of PDL in [21] showed that the probability distribution of Γ_{dB} was truncated because of the presence of a limited number of elements in the link having an appreciable PDL. In all cases, the global Γ_{dB} is a random variable depending on the number of PDL elements in the link and their individual PDL.

For each PDL element, γ is drawn from identical independently distributed (iid) normal distributions with mean and variance determined by the desired mean global Γ_{dB} and the number of PDL elements in the link as in [19]. The choice of the normal distribution is arbitrary because regardless of the distributions of the individual PDL elements, the statistics of the global PDL [19] of the whole link approach a Maxwellian distribution when enough PDL elements are inserted. These statistics were simulated with $N_S > 10$ spans and $M = 20$ and the obtained distributions perfectly fitted the expected theoretical Maxwellian distributions.

Fig. 2. Structure of the considered long-haul optical link.

Second, we look at the frequency dependence of the channel due to PMD which implies a frequency dependence of the global PDL. Current optical fibers have a typical PMD coefficient PMD_c as low as $0.05ps/\sqrt{km}$. If we consider a $L = 2000km$ multi-span link, the root mean square accumulated DGD σ_{DGD} is equal to $PMD_c\sqrt{L} = 2.2ps$ which induces no fluctuations to the PDL experienced by the subcarriers of an OFDM signal occupying a bandwidth less than the coherence bandwidth B_{coh} of the optical channel ($B_{coh} \propto \sigma_{DGD}^{-1}$ [22]). Hence, in the rest of our work, we will consider a narrow-band frequency independent channel model. Note that the frequency dependence of Γ_{dB} due to PMD in the wide-band case was reported in [23, 24]. In [24], authors found that an increased DGD (up to an accumulated DGD of $100ps$) in the optical link reduced the BER variations caused by PDL and not the degraded mean BER.

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

Third, we look at the noise in a long-haul optical link. The noise which is added on each amplification stage is modeled as an additive white Gaussian noise (AWGN). However, the PDL elements in the link act as partial polarizers. The signal and the in-line injected noise propagating through the link are both polarized by PDL. Given that the optical amplifiers are operated in a constant output power mode, the received signal $\mathscr{Y}_k$ at the output of the link can be written in function of the input signal $\mathbf{X}_k$ and the in-line injected noise:

$$\begin{aligned}\mathscr{Y}_k =& \frac{\mathscr{H}_{1\to N_S}\sqrt{2}}{(\mathrm{Tr}(\mathscr{H}_{1\to N_S}\mathscr{H}^{\dagger}_{1\to N_S}))^{1/2}}\mathbf{X}_k \\ &+ \Sigma_{j=1}^{N_S-1}\frac{\mathscr{H}_{j+1\to N_S}\sqrt{2}}{(\mathrm{Tr}(\mathscr{H}_{j+1\to N_S}\mathscr{H}^{\dagger}_{j+1\to N_S}))^{1/2}}\mathbf{N}_j + \mathbf{N}_{Ns} \\ &= \mathscr{H}_k\mathbf{X}_k + \mathscr{N}_k\end{aligned} \tag{5}$$

with $\mathscr{H}_{j\to N_S} = \mathscr{H}_{Ns}\mathscr{H}_{Ns-1}...\mathscr{H}_{j+1}\mathscr{H}_j$ and $\mathscr{H}_j$ being the transfer matrix of the jth span. Each component of the noise $\mathbf{N}_j$ added after the jth span is white and Gaussian distributed with a zero mean and a variance N_0 per real dimension. Because of PDL, $\mathscr{N}_k$ will be polarized and its coherency matrix $\mathbf{Q} = \mathbb{E}\left[\mathscr{N}_k\mathscr{N}_k^{\dagger}\right]$ will not be proportional to the identity matrix. However, we can still define an equivalent system given by [25]:

$$\mathbf{Y}_k = \mathbf{Q}^{-1/2}\mathscr{Y}_k = \mathbf{H}_k\mathbf{X}_k + \mathbf{N}_k \tag{6}$$

where $\mathbf{N}_k$ is a zero-mean white Gaussian noise of variance $\sigma^2 = N_S N_0$ per real dimension. The effective PDL experienced by the signal and given by the ratio of the eigenvalues of $\mathbf{H}_k\mathbf{H}_k^{\dagger}$ will be smaller than the PDL given by the ratio of the eigenvalues of $\mathscr{H}_k\mathscr{H}_k^{\dagger}$ but will also follow a Maxwellian distribution.

Using a singular value decomposition, $\mathbf{H}_k$ can be written as a product of two unitary matrices $\mathbf{U}$ and $\mathbf{V}$ and a diagonal matrix with $\sqrt{1+\gamma_{eq}}$ and $\sqrt{1-\gamma_{eq}}$ on its diagonal representing the different loss coefficients between the least and the most attenuated polarization states at the end of the link:

$$\mathbf{Y}_k = \mathbf{U}\begin{bmatrix}\sqrt{\lambda_{max}} & 0 \\ 0 & \sqrt{\lambda_{min}}\end{bmatrix}\mathbf{V}\mathbf{X}_k + \mathbf{N_k} = \sqrt{a}\mathbf{U}\begin{bmatrix}\sqrt{1+\gamma_{eq}} & 0 \\ 0 & \sqrt{1-\gamma_{eq}}\end{bmatrix}\mathbf{V}\mathbf{X}_k + \mathbf{N}_k \tag{7}$$

with $\gamma_{eq} = (\lambda_{max} - \lambda_{min})/(\lambda_{max} + \lambda_{min})$ and $a = (\lambda_{max} + \lambda_{min})/2$ where λ_{max} and λ_{min} are the eigenvalues of $\mathbf{H}_k\mathbf{H}_k^{\dagger}$. a is an inevitable loss coefficient induced by PDL. Being particularly interested in the differential attenuation and crosstalk effects due to PDL and polarization rotations, we will consider unitary rotation matrices. Moreover, knowing that the dynamics of the optical channel present rather slow variations [14], we will only regard constant values of $\Gamma_{dB} = 10log_{10}\frac{1+\gamma_{eq}}{1-\gamma_{eq}}$ for the numerical and theoretical investigations. Hence, in the rest of the paper, we will consider the following channel matrix $\mathbf{H}$ to model an optical link with PDL:

$$\mathbf{H} = \mathbf{R}_\alpha\begin{bmatrix}\sqrt{1+\gamma_{eq}} & 0 \\ 0 & \sqrt{1-\gamma_{eq}}\end{bmatrix}\mathbf{R}_\alpha^{-1} \tag{8}$$

3. Numerical investigation of PDL mitigation

Many ST codes were designed for wireless MIMO channels. We will consider the most famous ones that have proven to be the best-performing codes on 2×2 and 2×1 Rayleigh fading channels: the Golden code, the Silver code and the Alamouti code. In the following, the codeword matrix $\mathbf{X}$ and the rate of each coding scheme is given. The rate of a code is defined as the number of transmitted symbols per time slot (ts). In an uncoded case, we simply fill the matrix with 4 different M-QAM symbols and the rate is equal to 2 M-QAM symbols/ts.

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

3.1. PT codeword matrices

3.1.1. The Golden code

The Golden code [13] has the best performance on 2×2 MIMO Rayleigh fading channels. The codeword matrix of the Golden code is:

$$\mathbf{X}_{\mathscr{G}} = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha\left(S_1 + \theta S_2\right) & \alpha\left(S_3 + \theta S_4\right) \\ i\overline{\alpha}\left(S_3 + \overline{\theta} S_4\right) & \overline{\alpha}\left(S_1 + \overline{\theta} S_2\right) \end{bmatrix} \tag{9}$$

where $\theta = \frac{1+\sqrt{5}}{2}$, $\overline{\theta} = \frac{1-\sqrt{5}}{2}$, $\alpha = 1 + i - i\theta$, $\overline{\alpha} = 1 + i - i\overline{\theta}$ and S_1, S_2, S_3, S_4 are M-QAM symbols. The Golden code achieves a full rate of 2 symbols/ts because 4 symbols are transmitted during 2 symbol times. Hence, this code introduces no redundancy. Moreover, the determinant of the codeword matrix (corresponding to a coding gain on a Rayleigh fading channel [12]) is proportional to $\frac{1}{5}$ which is the highest obtained value for a 2×2 ST code.

3.1.2. The Silver code

The Silver code has a slightly weaker performance than the Golden code but it has the advantage of having a reduced decoding complexity due to its particular structure. The codeword matrix of the Silver code is:

$$\mathbf{X}_{\mathscr{S}} = \begin{bmatrix} S_1 + Z_3 & -S_2^* - Z_4^* \\ S_2 - Z_4 & S_1^* - Z_3^* \end{bmatrix}$$

$$\begin{bmatrix} Z_3 \\ Z_4 \end{bmatrix} = \frac{1}{\sqrt{7}} \begin{bmatrix} 1+i & -1+2i \\ 1+2i & 1-i \end{bmatrix} \begin{bmatrix} S_3 \\ S_4 \end{bmatrix} \tag{10}$$

where S_1, S_2, S_3, S_4 are M-QAM symbols. The determinant of this code is proportional to $\frac{1}{7}$. The Silver code also achieves a full rate of 2 symbols/ts because 4 symbols are transmitted during 2 time slots making it also a redundancy-free code.

3.1.3. The Alamouti code

The Alamouti code has the optimal performance on a 2×1 MIMO Rayleigh fading channel and can be also used for 2×2 MIMO channels. The codeword matrix of the Alamouti code is defined by:

$$\mathbf{X}_{\mathscr{A}} = \begin{bmatrix} S_1 & -S_2^* \\ S_2 & S_1^* \end{bmatrix} \tag{11}$$

where S_1 and S_2 are M-QAM symbols. Note that the codeword matrix has an orthogonal structure which makes the decoding straightforward. However, the Alamouti code introduces a redundancy as it has a rate of only 1 symbol/ts (half-rate code).

3.2. Performance analysis of PT codes

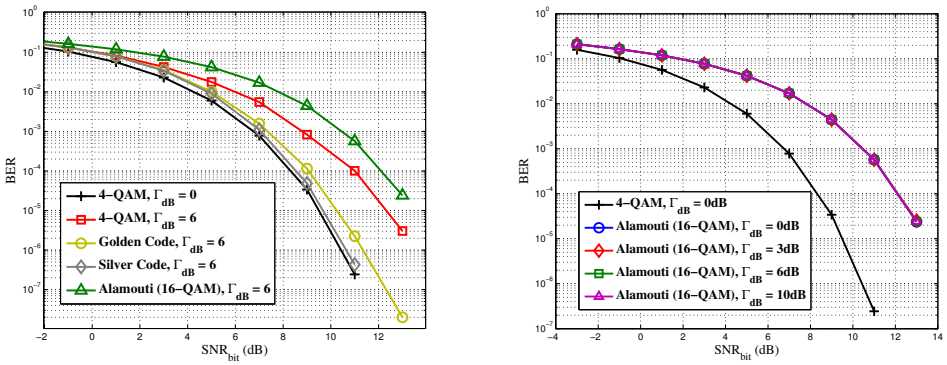
The performance of a PT-coded OFDM transmission on an optical channel with PDL will be investigated using the following frequency non-selective channel model with $\mathbf{H}$ defined in Eq. (8):

$$\mathbf{Y}_k = \mathbf{H}\mathbf{X}_k + \mathbf{N}_k \tag{12}$$

The noise $\mathbf{N}_k$ is modeled as additive white with iid circularly-symmetric complex Gaussian components $\mathscr{N}(0, 2\sigma^2)$. At the transmitter, 4-QAM symbols will be used for the uncoded case as well as to fill the codeword matrices of the Silver and the Golden code (full-rate codes). On the other hand, 16-QAM symbols will be used for the Alamouti code (half-rate code) in order to compare the schemes at the same spectral efficiency of 4 information bits/ts. At the receiver,

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

ML decoding is performed by an exhaustive search with a reasonable complexity. In the case of an uncoded scheme, there are 16 possible codewords in $\mathscr{C}$. In case of the Silver or the Golden code, $\mathscr{C}$ has $4^4 = 256$ codewords. The Alamouti code benefits from its orthogonal structure and achieves the ML criterion with a single decoding operation. Monte Carlo simulations are carried in order to evaluate the performance of the different coding schemes in terms of BERs.



(a) Performance comparison of 4-QAM, Silver, Golden and Alamouti codes when $\Gamma_{dB} = 6$.

(b) BER as a function of SNR_{bit} for the Alamouti code at different Γ_{dB}.

Fig. 3. Performance of PT codes obtained through Monte Carlo simulations.

Figure 3(a) shows the BER evolution versus the SNR_{bit} for a PDL of 6dB. Without PT coding, the SNR degradation induced by PDL is about 2.3dB at BER $= 10^{-3}$. The SNR penalty - defined as the SNR gap to the PDL-free case at a given BER - is only 0.6dB with the Golden code and 0.3dB with the Silver code corresponding to a coding gain of 2dB. Apart from the important coding gains, these codes do not introduce any spectral efficiency penalties compared to the uncoded case as they are by construction redundancy-free. On the other hand, looking at the performance of the Alamouti code, we notice an SNR penalty of 3.6dB at BER $= 10^{-3}$. This is due to the use of 16-QAM symbols. In Fig. 3(b), we observe the performance of the Alamouti code for 3 different levels of PDL. We notice that the code performs the same independently of the amount of PDL. The SNR penalties at BER $= 10^{-3}$ for different PDL values had been evaluated for the first time in [8] and the results were confirmed in [11]. For PDL values less than 4dB, the Silver code can mitigate almost all PDL effects and always performs better than the Golden code. In [10], the performance of the Alamouti code was also found to be independent of PDL. Moreover, experimental demonstrations of PDL mitigation with PT coding were also realized in a back-to-back (no transmission) scenario [9]. However, all these observations were still analytically unexplained.

3.3. Concatenation of FEC and PT coding

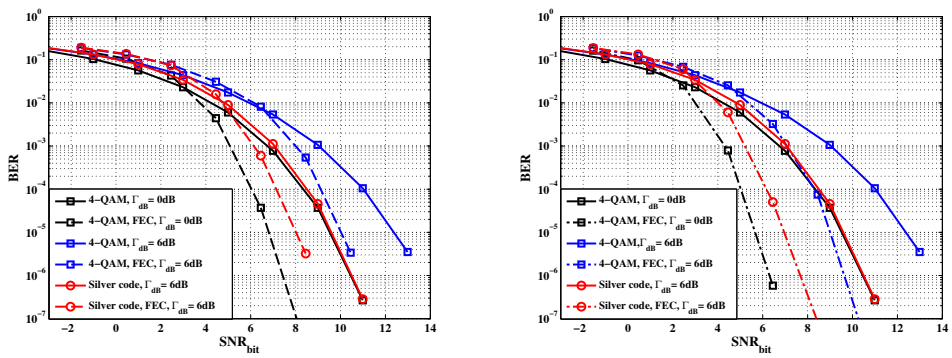
While PT coding technique uses the modulated symbols to form a codeword matrix and mitigate some channel effects, other coding techniques, known as forward error-correcting (FEC) or channel codes, operate on the information bits and add some redundancy in order to enhance the performance over the noisy channel. The FEC block at the transmitter side precedes the M-QAM modulation block and a corresponding FEC decoding unit at the receiver side follows the demodulation block. Linear block codes are a major class of channel codes where a binary information sequence of length k is mapped to a binary sequence of length n called a codeword.

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

The code rate of a block code is denoted by R_c:

$$R_c = \frac{k}{n} < 1 \tag{13}$$

The optimal decoding strategy of FEC codes is soft decision decoding (SDD) [26, Chap.7] that receives at its input a real value indicating the reliability of each coded bit (e.g. log-likelihood ratios) and computes the original information bits. A sub-optimal, yet computationally simple, decoding strategy is hard decision decoding (HDD) [26, Chap.7] that consists in quantizing the samples at the output of the PT demodulator before sending them to the FEC decoder.



(a) HDD (b) SDD

Fig. 4. Bit Error Rate as a function of the SNR_{bit} for the uncoded scheme and the Silver code, with or without FEC: (a) Hard decision decoding (HDD), (b) Soft decision decoding (SDD). The simulated FEC is a BCH(63,45) code.

We are interested, in this section, in evaluating the total gain provided by FEC coding and PT coding through Monte-Carlo simulations. The performance on an optical channel with $\Gamma_{dB} = 6dB$ of an uncoded 4-QAM scheme (no FEC, no PT) and a Silver-coded scheme is compared with a Bose-Chaudhuri-Hocquenghem (BCH) coded binary sequence where $n = 63$, $k = 45$ and an error correction capability $t = 3$, followed by 4-QAM modulation and PT coding (Silver code). The obtained BER curves are plotted in Fig. 4. At a $BER = 10^{-4}$, the gain provided by the Silver code alone is $2.4dB$. When HDD is considered, the gain provided by the BCH code alone is $1.8dB$. The concatenation of both codes offer a total coding gain of $3.9dB$, approximately equal to the sum of the separate coding gains $2.4 + 1.8 = 4.2dB$. The summation of the FEC and PT coding gains is also observed when using SDD. In this case, the gain offered by the BCH code alone at a $BER = 10^{-4}$ is $2.8dB$. The concatenation with the Silver code provides a total gain of $4.9dB$ and the sum of the separate coding gains is $2.4 + 2.8 = 5.2dB$.

4. Theoretical analysis of PDL mitigation

Three main results are retained when observing the previous numerical investigations:

- the Silver code performs better than the Golden code, unlike the case of the wireless Rayleigh fading channel.
- the performance of the Alamouti code is independent of the amount of PDL in the link.
- the total coding gain obtained when using both a FEC code and a PT code is equal to the sum of the separate coding gains.

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

These findings show different behaviors of MIMO codes on the optical channel. Further exploration of the mitigation of PDL with MIMO codes is required to understand the obtained results and check whether the investigated codes are optimal or better codes can be found. Many previous works have investigated the performance of PolMux schemes in presence of PDL in terms of the outage probability [11, 25, 27]. However, an outage analysis only gives the performance limits of a transmission scheme. The only properties of the transmitted signal that appear in an outage analysis are the average symbol energy and the transmission rate. On the other hand, an error probability analysis based on the calculus of an upper bound of the probability that an ML decoder erroneously decodes the transmitted symbol takes into account the codeword matrix of the transmitted symbols. It is an appropriate tool to extract the criteria that a codeword matrix should satisfy in order to minimize or mitigate the channel impairments. A well-known and extensively used error probability calculus in the case of a MIMO Rayleigh fading channel can be found in [12]. This calculus revealed the design criteria that were used to construct the Golden and the Silver codes for the wireless channel. In this section, we will derive an upper bound of the error probability in the case of an optical channel with PDL.

4.1. Upper bound of the error probability

The error probability is defined as:

$$\begin{aligned} P_{error} &= Pr\{\mathbf{X}' \neq \mathbf{X}\} \\ &= \sum_{\mathbf{X}\in\mathscr{C}} Pr\{\mathbf{X}\} Pr\{\mathbf{X}' \neq \mathbf{X}|\mathbf{X}\} \end{aligned} \tag{14}$$

For equiprobable codewords, using the union bound, the error probability is upper-bounded by:

$$P_{error} \leq \frac{1}{\mathrm{card}(\mathscr{C})} \sum_{\substack{\mathbf{X},\mathbf{X}'\in\mathscr{C}, \\ \mathbf{X}'\neq\mathbf{X}}} Pr(\mathbf{X} \to \mathbf{X}') \tag{15}$$

where $\mathrm{card}(\mathscr{C})$ is the cardinality of $\mathscr{C}$ and $Pr(\mathbf{X} \to \mathbf{X}')$ is the pairwise error probability (PEP) supposing that $\mathbf{X}$ and $\mathbf{X}'$ are the only possible codewords in the codeword space. To compute the PEP, we define the conditional PEP that we average over all the possible channel realizations. For an ML decoder, the conditional PEP is defined as:

$$Pr(\mathbf{X} \to \mathbf{X}'|\mathbf{H}) = Pr\{\|\mathbf{Y} - \mathbf{H}\mathbf{X}'\|^2 \leq \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 | \mathbf{X}, \mathbf{H}\} \tag{16}$$

Using the Gaussian properties of the noise N and applying Chernoff's bound, the PEP can be upper-bounded by [26, Chap.4]:

$$Pr(\mathbf{X} \to \mathbf{X}') \leq \mathbb{E}_H[\exp(-\frac{\|\mathbf{H}(\mathbf{X} - \mathbf{X}')\|^2}{8\sigma^2})] \tag{17}$$

where $\mathbb{E}_H[]$ is the averaging operation over all possible channel realizations.

Averaging over $\mathbf{H}$ given in Eq. (8) where we consider constant values of Γ_{dB} and a random rotation angle that varies uniformly in $[0 : 2\pi]$, we get:

$$Pr(\mathbf{X} \to \mathbf{X}') \leq \exp(-\frac{\|\mathbf{X}_\Delta\|^2}{8\sigma^2}) I_0(\frac{\gamma_{eq}}{8\sigma^2}\sqrt{\mathrm{a}^2 + \mathrm{b}^2}) \tag{18}$$

where $I_0(z)$ is the 0th order modified Bessel function of the first kind. $\mathbf{X}_\Delta = \mathbf{X} - \mathbf{X}' = \begin{pmatrix} \vec{x}_1 \\ \vec{x}_2 \end{pmatrix}$, $\vec{x}_{1,2}$ being line vectors and:

$$\mathrm{a} = \|\vec{x}_2\|^2 - \|\vec{x}_1\|^2 \tag{19}$$

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

$$\mathrm{b} = 2Re(\langle \vec{x_1}, \vec{x_2} \rangle) \tag{20}$$

We can approximate the error probability expression in Eq. (18) for high SNR values by using a first order approximation of $I_0(z)$ when $z \to \infty$ and get:

$$Pr(\mathbf{X} \to \mathbf{X}') \approx \exp(-\frac{\|\mathbf{X}_\Delta\|^2 - \gamma_{eq}\sqrt{\mathrm{a}^2 + \mathrm{b}^2}}{8\sigma^2}) \tag{21}$$

4.2. Design criterion

$I_0(z)$ being monotonously increasing for $z \geq 0$, its minimum is at $z = 0$. This corresponds to null 'a' and 'b' and the obtained error probability will be independent of PDL. Consequently:

Proposition 1 *A Polarization-Time code completely mitigates PDL if and only if all codeword differences satisfy:*

1. $a = \|\vec{x_2}\|^2 - \|\vec{x_1}\|^2 = 0$ *and*
2. $b = 2Re(\langle \vec{x_1}, \vec{x_2} \rangle) = 0$.

When this design criterion is met, we recover the performance over two parallel AWGN channels which is the best achievable performance:

$$p_{error,AWGN} \approx \exp(-\frac{\|\mathbf{X}_\Delta\|^2}{8\sigma^2}) \tag{22}$$

The resulting criterion is completely different from the rank and the minimum determinant criteria derived for a Rayleigh fading channel [12] and defining respectively a diversity gain and a coding gain on the BER curves. If we compare the approximation of the error probability expression at high SNR in Eq. (21) to the one obtained in the case of a 2×2 MIMO Rayleigh fading channel, we notice different behaviors: the error probability of the Rayleigh fading channel decays as SNR^{-2r} [12] where r is the minimum rank of the matrix $\mathbf{X}_\Delta$. The diversity of the system is defined as the power of SNR^{-1} ($2r$ in this case) and can be graphically discerned as the slope of the BER curves at high SNR values. While the error probability of the PDL channel decays exponentially as a function of the SNR as in the case of an additive white Gaussian channel in Eq. (22). Hence, Space-Time codes, used as Polarization-Time codes, bring no diversity gain to the optical channel with PDL. A coding gain that will be evaluated in the following, is only brought, reducing the penalty induced by PDL.

4.3. Performance analysis of PT codes

The performance of both coded and uncoded schemes will be examined using the derived upper bound of the pairwise error probability expression in Eq. (21). To compare the different schemes, we will compute the squared distance $d^2 = \|\mathbf{X}_\Delta\|^2 - \gamma_{eq}\sqrt{\mathrm{a}^2 + \mathrm{b}^2}$ of all combinations of codeword differences and compare the minimum values of d^2, denoted d^2_{min}, of the codes. The best code is the one that maximizes this minimum value. In Table 1, we report the minimum values of $\|\mathbf{X}_\Delta\|^2 - \gamma_{eq}\sqrt{\mathrm{a}^2 + \mathrm{b}^2}$ analytically computed for each investigated PT code at four different PDL values. To fill the table, we set a spectral efficiency of 4 bits per time slot for all coding schemes and consider an average symbol energy $E_s = 1$ for all constellations.

First, we notice that the Alamouti code has the same minimal distance for all PDL values which explains why it performs the same independently of PDL in Fig. 3(b). This is due to the orthogonality of its codeword matrix (Eq. (11)) that induces $\mathrm{a} = \mathrm{b} = 0$ for all possible codeword differences. Hence, this code satisfies the criterion of *Proposition 1*. However, its performance is affected by the use of 16-QAM symbols giving a squared minimal distance of 0.8.

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

Table 1. d^2_{min} for different coding schemes at different PDL values

	$\Gamma = 0dB$	$\Gamma = 3dB$	$\Gamma = 6dB$	$\Gamma = 10dB$
Silver Code	2	2	2	1.23
Golden Code	2	1.7	1.46	1.07
No Coding	2	1.34	0.8	0.38
Alamouti Code	0.8	0.8	0.8	0.8

Second, we note that the Silver code is not optimal in mitigating PDL since it does not satisfy the derived design criterion. Unlike the Alamouti code, the Silver code has only some codeword differences having a = b = 0. In Fig. 5, we plot the performance of the Silver code for different PDL values. We see that the code mitigates almost all PDL when the PDL coefficient is equal to 3dB and 6dB whereas for a PDL of 10dB, the code is not able to completely palliate PDL. The computed d^2_{min} in Table 1 explain the behavior of the Silver code. d^2_{min} is given by the same codeword difference with a = 0 and b = 0 at a PDL of 3dB and 6dB, and is equal to 2. Whereas at a PDL of 10dB, it falls to 1.23 given by another codeword difference where a $\neq$ 0 and b $\neq$ 0.

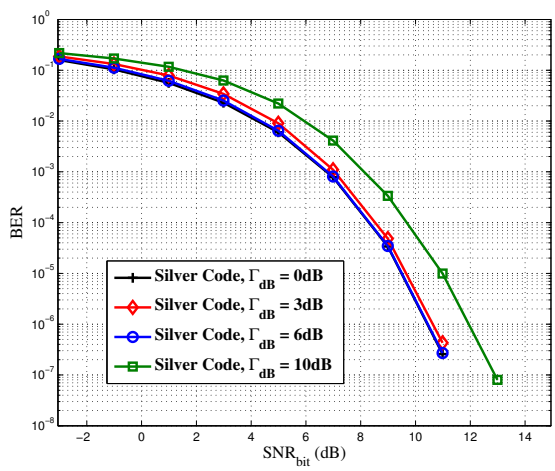


Fig. 5. Bit Error Rate as a function of the SNR_{bit} for the Silver code, obtained through Monte Carlo simulations.

Third, in Fig. 3(a), we saw that the Silver code outperforms the Golden code for $\Gamma_{dB} = 6dB$, and both reduce the penalty that PDL causes to the uncoded scheme. Again, this result can be explained by looking at Table 1. Indeed, d^2_{min} is the greatest for the Silver code followed by the Golden code and then the uncoded scheme.

In conclusion, we were able to explain, in terms of error probability bounds, the performance of the Alamouti, the Silver and the Golden codes on an optical channel with PDL. These codes were designed to satisfy the rank and the minimum determinant criteria for a wireless channel that are no more relevant for the optical channel.

4.4. Concatenation of FEC and PT coding

The numerical investigation of the performance of a PolMux scheme using both a FEC code and a PT code showed that the total coding gain is equal to the sum of the separate coding

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

gains brought by each. This result can be also theoretically explained and is mainly due to the exponential decrease of the error probability as a function of SNR.

The bit error probability of a linear block code over an AWGN channel, and after hard decision decoding, is determined by the minimum distance of the considered FEC code $d_{min,FEC}$ and the crossover probability p of the equivalent binary symmetric channel (BSC) [26, Chap.7]:

$$P_{e,HDD} \leq \sum_{m=t+1}^{n} \binom{n}{m} p^m (1-p)^{n-m} \tag{23}$$

where $t = \lfloor \frac{d_{min,FEC}-1}{2} \rfloor$ is the error correction capability of the linear block code and $d_{min,FEC}$ is the minimum Hamming distance between distinct codewords of this code. At high SNR, p tends towards zero and Eq. (23) is dominated by the first term where $m = t+1$:

$$P_{e,HDD} \leq \binom{n}{t+1} p^{t+1} = \binom{n}{d_{FEC,HDD}} p^{d_{FEC,HDD}} \tag{24}$$

where $d_{FEC,HDD} = \lceil \frac{d_{min,FEC}}{2} \rceil$.

Had we considered soft decision decoding, the bit error probability would have been [26, Chap.7]:

$$P_{e,SDD} \leq (2^k - 1) p^{d_{FEC,SDD}} \tag{25}$$

where $d_{FEC,SDD} = d_{min,FEC}$.

p can be upper-bounded using Eq. (21) where the error probability is a codeword error probability. Given that a full-rate 2×2 PT codeword has $4\log_2 M$ bits, an error in decoding one PT codeword implies that one bit is erroneous in the best case or all $4\log_2 M$ bits are erroneous in the worst case, hence:

$$\frac{1}{4\log_2 M} Pr(\mathbf{X} \rightarrow \mathbf{X}') \leq p \leq Pr(\mathbf{X} \rightarrow \mathbf{X}') \tag{26}$$

At high SNR, the pairwise error probability of the closest neighbors predominates the other terms in the union bound (Eq. (15)) and p is upper-bounded by:

$$p \leq A_{d_{PT}} \exp(-\frac{d_{PT}^2}{8\sigma^2}) \tag{27}$$

where $d_{PT}^2 = \underset{\mathbf{X},\mathbf{X}' \in \mathscr{C}}{argmin}(\|\mathbf{X}_\Delta\|^2 - \gamma_{eq}\sqrt{\mathrm{a}^2 + \mathrm{b}^2})$ and $A_{d_{PT}}$ being the average number of codewords located at the distance d_{PT} of a given codeword. $A_{d_{PT}}$ is usually called the kissing number.

Replacing p by its upper bound and substituting the average symbol energy E_S for the average energy per information bit E_b using $E_S = 1 = r_{PT} R_c E_b \log_2 M$, we get:

$$P_e \leq K (A_{d_{PT}})^{d_{FEC}} \exp(-\frac{d_{FEC} d_{PT}^2 \mathrm{SNR}_{bit} r_{PT} R_c \log_2 M}{4}) \tag{28}$$

where $\mathrm{SNR}_{bit} = \frac{E_b}{2\sigma^2}$. K and d_{FEC} are the constant and the power of the bit error probability of p in Eq. (23) or (25) depending on the chosen FEC decoding strategy. r_{PT} is equal to 1 for a full-rate PT code and 0.5 for a half-rate code.

In order to evaluate the asymptotic gains provided by the concatenation of a FEC code and a PT code, we compare the following two schemes: a first scheme NC without FEC ($d_{FEC} = 1$ and $R_c = 1$) and without PT coding (independent M-QAM symbols) and a second scheme $FEC+$

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

PT using a linear block code and a full-rate PT code. At the same achieved error probability, the coding gain G of the coded scheme is given by:

$$G = \frac{\mathrm{SNR}_{bit,FEC+PT}}{\mathrm{SNR}_{bit,NC}} = \frac{R_c d_{FEC} d_{PT}^2}{d_{NC}^2} \tag{29}$$

In decibels, we obtain G_{dB}:

$$\begin{aligned} G_{dB} &= 10log_{10}(R_c d_{FEC}) + 10log_{10}(\frac{d_{PT}^2}{d_{NC}^2}) \\ &= G_{dB,FEC} + G_{dB,PT} \end{aligned} \tag{30}$$

The first term denote the coding gain of the FEC code and the third term denotes the coding gain provided by PT coding. Equation (30) shows that the total asymptotic gain obtained when concatenating a FEC code and a PT code is the sum of the gains provided by each code separately, validating the third result of our numerical investigation.

5. Experimental validation of PDL mitigation

The numerical and theoretical validations of PDL mitigation using PT coding are limited to a linear channel with a single-element or lumped PDL and AWGN noise at the receiver side. Initial experimental results limited to the linear regime also showed the efficiency of PT codes [9]. These simplifications allowed us to analyze the potential of MIMO codes for PolMux optical transmission using a handy and well defined channel model. The final remaining step is to test the ability of PT codes to mitigate inline PDL through a transmission experiment with distributed PDL taking into account the interactions of PDL with distributed ASE noise [28] and non-linear effects [18].

5.1. Experimental setup

The transmission scheme, shown in Fig. 6, consists of a PolMux OFDM transmitter, a 1000 km optical link and a dual-polarization coherent receiver. The OFDM signals are made of 256 subcarriers including 194 data subcarriers and 10 pilot subcarriers for common phase error estimation. The data subcarriers are alternatively modulated with 4-QAM symbols, Silver-, Golden- and 16-QAM Alamouti- coded symbols. An 18-sample CP is appended to each OFDM symbol (7% overhead). A time-multiplexed training sequence [3] is periodically inserted for channel estimation and time and frequency synchronizations. The total effective bitrate is 25 Gbit/s while the raw bit rate is 36 Gbit/s taking into account the different transmission overheads. The OFDM signals are generated by an arbitrary waveform generator (AWG) with two outputs. Hence, to emulate PolMux, we consider OFDM signals satisfying the Hermitian symmetry property. Next, up-conversion to the optical domain is done using two single-drive Mach-Zehnder modulators (MZM). Then the two optical signals are combined by a Polarization Beam Combiner and multiplexed in the optical link among seven 50 GHz-spaced wavelengths. Further details on the experimental setup can be found in [29].

The transmission line consists in a recirculating loop which contains 2 spans of 100 km of SMF, a PDL element of 2dB, a polarization scrambler to randomize the polarization state between successive loops and an optical band-pass filter. After additional ASE noise loading at the receiver, the desired wavelength is filtered and the signal is detected with a dual-polarization coherent receiver. The received signal is then acquired with a real-time Tektronix oscilloscope and offline-processing [3, 7] ending with ML decoding is carried to measure the BER and the corresponding Q-factor in dB, commonly used in optical communications:

$$BER = 0.5erfc(Q/\sqrt{2}) \qquad \text{and} \qquad Q_{dB} = 20\log_{10}(Q) \tag{31}$$

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

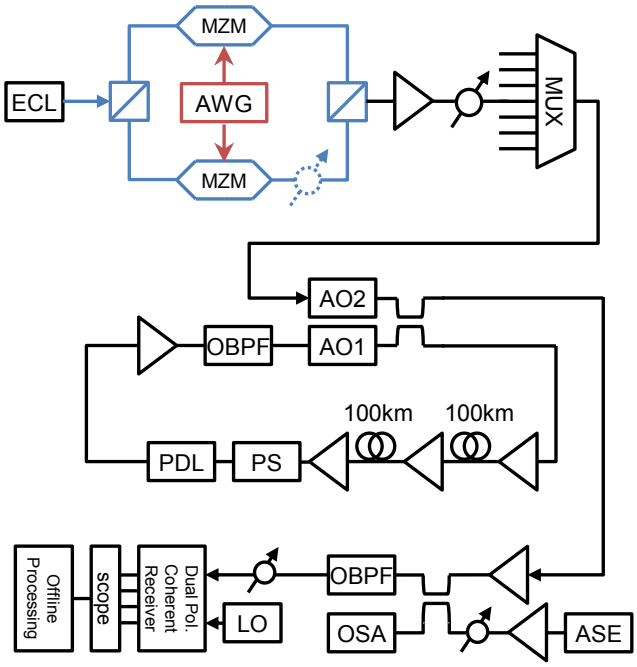


Fig. 6. Experimental setup. (ECL: External Cavity Laser, AWG: Arbitrary Waveform Generator, MUX: Multiplexer, AO: Acousto-Optical Modulator, PS: Polarization Scrambler, OBPF: Optical Band-Pass Filter, ASE: Accumulated Spontaneous Emission source, LO: Local Oscillator, OSA: Optical Spectrum Analyzer, OSC: Tektronix 50GS/s Oscilloscope).

5.2. *Experimental Results*

5.2.1. PT coding in non-linear propagation regime

We will transmit the PT-coded OFDM in the recirculating loop of Fig. 6 and replace the in-line PDL element by an attenuator having the same insertion loss and emulate PDL at the transmitter side by inserting an attenuator in one branch of the PolMux transmitter (this corresponds to the worst case of OSNR degradation of one polarization tributary [18]). First, BER measurements are carried out in a back-to-back setup [29] validating the previous simulated performance of all coding schemes. Then, BERs are measured after a 1000km-transmission for input powers ranging from -15 to 5dBm. The obtained bathtub-shaped curves are plotted in Fig. 7 for the 4-QAM and Silver-coded schemes at a PDL of 0, 3 and 6dB. First, we notice that all curves have their optimum operating point at an input power of -3dBm. However, the uncoded scheme suffers severely from PDL ($Q = 10.5dB$ at -3 dBm for $PDL = 6dB$) while the Silver code completely compensates the penalty at $PDL = 3dB$ and nearly all penalty at $PDL = 6dB$ ($Q = 12.3dB$ at -3dBm). Second, the curves can be separated into three regimes: a linear regime, up to -6dBm, where non-linear effects are negligible and the coding gains match the numerical results obtained with a linear channel model; a moderate non-linear regime, between -6dBm and 0dBm, where a minimum BER is reached; and a severe non-linear regime where the performance of all schemes is deteriorated. An important result is that the PT-coded OFDM does not induce any extra penalties in presence of non-linear effects. This experimental validation confirms a previous numerical investigation of the behavior of PT codes in non-linear propagation regime [30].

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

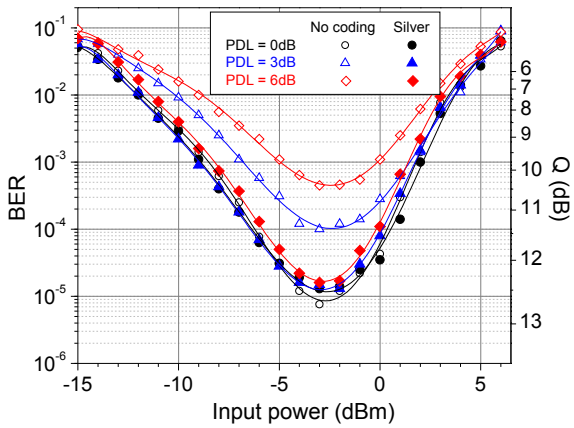


Fig. 7. BER evolution versus launched input power after $5 \times 200km$ for the Silver-coded and 4-QAM schemes, at three different PDL values at the transmitter: 0, 3 and 6dB.

5.2.2. Mitigation of distributed in-line PDL

We remove the emulated PDL at the transmitter side and insert the PDL element of 2dB into the loop. 2000 Q-factor measurements are then carried out for the different coding schemes at the optimum operating point. The measured OSNR at the receiver is 12dB after additional ASE noise loading in order to evaluate the complete Q-factor distributions. The OSNR is proportional to SNR_{bit} that we used along this paper:

$$\mathrm{OSNR} = \frac{R_b}{2B_{ref}}\mathrm{SNR}_{bit} \tag{32}$$

where R_b is the total transmitted bitrate and B_{ref} is the reference spectral bandwidth of $0.1nm$.

The obtained Q-factor distributions are shown in Fig. 8. We also show, in the inset, the measured probability distribution of PDL and the fitted theoretical distribution. PDL is measured using the estimated 2×2 channel matrices and a Maxwellian probability distribution function of PDL is obtained with a 4.2dB mean (the expected theoretical mean being 4dB [19]).

The Q-factor distribution for the 4-QAM-coded subcarriers is asymmetric with a large tail towards the worst Q-factors and has a mean of 11.2dB and a standard deviation (std) of 0.55dB whereas the distribution is symmetric and the mean Q-factor of the Silver- and Golden- coded subcarriers is 11.8dB (the same observed value for the PDL-free case when OSNR = 12dB). The Q-factor distributions are also narrower when PT codes are used. The Silver code gives a distribution slightly narrower than the one obtained with the Golden code: std of 0.35dB and 0.37dB respectively. For the Alamouti code, we observe a mean Q-factor of 8.3dB due to the use of 16-QAM symbols to guarantee the same spectral efficiency of the other schemes. However, its Q factor distribution is the narrowest of all with a std of 0.24dB. This is due to the powerful orthogonal structure of the Alamouti codeword matrix that makes its performance independent of the amount of accumulated PDL in the link, as found in the theoretical analysis in section 4. The observed small variance of its Q-factor distribution can be ascribed to the increased sensitivity of 16-QAM modulation to phase noise and non-linear effects.

#195015 - $15.00 USD Received 31 Jul 2013; revised 29 Aug 2013; accepted 30 Aug 2013; published 20 Sep 2013
(C) 2013 OSA

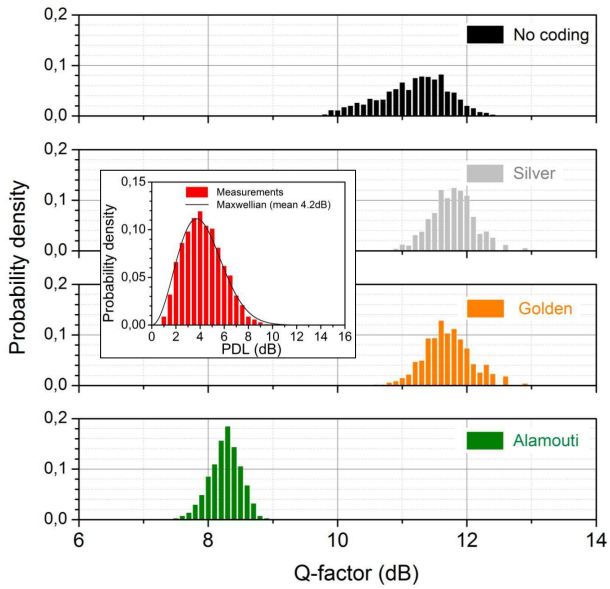


Fig. 8. Q-factor distribution after $5 \times 200km$ at $P_{in} = -3dBm$ ($\mathrm{OSNR}_{0.1nm} = 12dB$). Inset: Experimental and theoretical probability distributions of PDL.

6. Conclusion

We have presented numerical and theoretical studies of PDL mitigation using redundancy-free coding schemes. The Silver code can efficiently mitigate the PDL in current long-haul optical links where PDL values rarely exceed 6dB [21]. However, the theoretical study of PDL mitigation showed that this code is not optimal. The established upper bound of the error probability for an optical channel with PDL yields the design criterion required to construct codes that guarantee a PDL-independent error probability, restoring the performance over an additive white Gaussian channel. These design rules open the way for the construction of optimal codes dedicated to PDL mitigation in the optical channel. We have also shown that the total coding gain of a concatenation of a FEC code and a PT code is the sum of the separate coding gains. A transmission experiment with distributed in-line PDL was accomplished as well, validating the results obtained with a simplified channel model. PT coding does not induce extra penalties when non linear effects are considered and PT codes can both improve the mean and the variance of the Q-factor distributions.

Acknowledgments

The experimental work has been partially supported by the Celtic-Plus SASER-SIEGFRIED project.