A modified constant modulus algorithm for polarization-switched QPSK

Pontus Johannisson[1], Martin Sjödin[1], Magnus Karlsson[1], Henk Wymeersch[2], Erik Agrell[2], and Peter A. Andrekson[1]

[1]*Photonics Laboratory, Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-412 96 Göteborg, Sweden*
[2]*Communication Systems Group, Department of Signals and Systems, Chalmers University of Technology, SE-412 96 Göteborg, Sweden*

pontus.johannisson@chalmers.se

Abstract: By using a generalized cost function, a modified constant modulus algorithm (CMA) that allows polarization demultiplexing and equalization of polarization-switched QPSK is found. An implementation that allows easy switching between the conventional and the modified CMA is described. Using numerical simulations, the suggested algorithm is shown to have similar performance for polarization-switched QPSK as CMA has for polarization-multiplexed QPSK.

OCIS codes: (060.2330) Fiber optics communications; (060.1660) Coherent communications; (060.4080) Modulation.

References and links

1. H. Bülow, "Polarization QAM modulation (POL-QAM) for coherent detection schemes," in *Optical Fiber Communication Conference (OFC)* (2009), p. OWG2.
2. M. Karlsson and E. Agrell, "Which is the most power-efficient modulation format in optical links?" Opt. Express **17**, 10814–10819 (2009).
3. E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," J. Lightw. Technol. **27**, 5115–5126 (2009).
4. P. Poggiolini, G. Bosco, A. Carena, V. Curri, and F. Forghieri, "Performance evaluation of coherent WDM PS-QPSK (HEXA) accounting for non-linear fiber propagation effects," Opt. Express **18**, 11360–11371 (2010).
5. P. Serena, A. Vannucci, and A. Bononi, "The performance of polarization switched-QPSK (PS-QPSK) in dispersion managed WDM transmissions," in *European Conference on Optical Communication (ECOC)* (2010), p. Th.10.E.2.
6. M. Sjödin, P. Johannisson, H. Wymeersch, P. Andrekson, and M. Karlsson, "Experimental comparison of polarization-switched QPSK and polarization-multiplexed QPSK at 30 Gbit/s," Opt. Express **accepted with revisions**.
7. D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," IEEE Trans. Commun. **28**, 1867–1875 (1980).
8. A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," IEEE Trans. Signal Processing **55**, 2740–2746 (2007).
9. S. J. Savory, "Digital filters for coherent optical receivers," Opt. Express **16**, 804–817 (2008).
10. K. Kikuchi, "Polarization-demultiplexing algorithm in the digital coherent receiver," in *IEEE/LEOS Summer Topical Meetings (LEOSST)* (2008), p. MC2.2.
11. I. Roudas, A. Vgenis, C. S. Petrou, D. Toumpakaris, J. Hurley, M. Sauer, J. Downie, Y. Mauro, and S. Raghavan, "Optimal polarization demultiplexing for coherent optical communications systems," J. Lightw. Technol. **28**, 1121–1134 (2010).
12. L. Liu, Z. Tao, W. Yan, S. Oda, T. Hoshida, and J. C. Rasmussen, "Initial tap setup of constant modulus algorithm for polarization de-multiplexing in optical coherent receivers," in *Optical Fiber Communication Conference (OFC)* (2009), p. OMT2.
13. U. Madhow, *Fundamentals of Digital Communication* (Cambridge Univ Press, 2008).

14. T. Duthel, C. R. S. Fludger, J. Geyer, and C. Schulien, "Impact of polarisation dependent loss on coherent POLMUX-NRZ-DQPSK," in *Optical Fiber Communication Conference (OFC)* (2008), p. OThU5.
15. M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "DSP for coherent single-carrier receivers," J. Lightw. Technol. **27**, 3614–3622 (2009).

1. Introduction

Polarization-multiplexed (PM) transmission combined with coherent reception and impairment mitigation by digital signal processing has recently attracted considerable interest. In particular, PM quadrature phase-shift keying (PM-QPSK) is currently considered to be very promising since this format combines good spectral efficiency with manageable implementation complexity. However, although PM-QPSK has many favorable features, it has been shown that by viewing the real and imaginary parts of the two polarization components of the electric field as a four-dimensional (4D) signal space, modulation formats with even better performance can be found. The first such format to be suggested was a 24-point constellation and a sensitivity improvement over PM-QPSK was demonstrated [1]. The topic of 4D modulation has then been investigated by sphere-packing simulations, which has resulted in a number of suggestions for new modulation formats [2, 3]. Among these, an 8-point constellation is particularly interesting due to its optimal power efficiency. This format has been named polarization-switched QPSK (PS-QPSK) and can be described in the following way: Starting from a single-polarization QPSK signal we can, for every transmitted symbol, choose the launch polarization. This allows one extra bit to be transmitted without any power increase. Comparing with PM-QPSK, a (4D) PS-QPSK symbol carries 3 instead of 4 bits per symbol and has a sensitivity improvement of 1.76 dB at high SNR and 0.97 dB at a bit-error rate (BER) of 10^{-3}.

Numerical studies have confirmed the sensitivity improvement of PS-QPSK. In a comparison with PM-QPSK in a system using wavelength-division multiplexing (WDM), it was suggested that PS-QPSK can serve as a fall-back alternative to PM-QPSK when an optical link is degraded by additional losses [4]. The link could then continue to operate at 75% of the data rate with a 3 dB sensitivity improvement without any hardware adjustments. Furthermore, it was found that PS-QPSK has significantly improved nonlinear tolerance, in particular when using electronic dispersion compensation, and it was suggested that PS-QPSK may be competitive for 100 Gigabit Ethernet links. A second numerical study also reported improved nonlinear tolerance for PS-QPSK and the largest difference was seen for inter-channel nonlinear effects in a WDM system [5].

We have performed transmission experiments comparing the performance of PS-QPSK and PM-QPSK and found that the sensitivity improvement of PS-QPSK can be realized with readily available components [6]. During this work, we found that the conventional constant-modulus algorithm (CMA) cannot be used for PS-QPSK. We have investigated this problem and we here suggest a modified CMA that allows polarization demultiplexing and equalization of PS-QPSK. We will refer to the conventional CMA as *PM-CMA* and to the new algorithm as *PS-CMA*.

This paper is organized as follows: In Section 2, we explain why PM-CMA cannot be used for PS-QPSK and in Section 3 we describe PS-CMA. In Section 4 we show how we have measured the performance of PS-CMA and PM-CMA and numerical results are reported in Section 5. A discussion about the impact of polarization-mode dispersion (PMD) and polarization-dependent loss (PDL) is found in Section 6 and this is followed by the conclusions.

Notation: Vectors are denoted in bold letters (e.g., $\mathbf{a}$), and matrices in capital bold letters (e.g., $\mathbf{A}$). Transposition is written as $\mathbf{a}^{\mathrm{T}}$, conjugation as $\mathbf{a}^{*}$, and conjugate transpose is denoted by $\mathbf{a}^{\mathrm{H}}$. The identity matrix is written as $\mathbf{I}$ and the expectation operator is denoted by $\mathbb{E}[\cdot]$.

2. The problem with PM-CMA when using PS-QPSK

We write 4D symbols as Jones vectors by combining the real and imaginary parts of the electric field's x- and y-polarized components according to $(E_{x,r}+jE_{x,i},E_{y,r}+jE_{y,i})^{\mathrm{T}}$ [2, 3]. The above description of PS-QPSK then corresponds to a symbol alphabet with unit energy 4D pulses $\mathscr{M}_{\mathrm{PS}}=\{(\pm1\pm j,0)^{\mathrm{T}}/\sqrt{2};(0,\pm1\pm j)^{\mathrm{T}}/\sqrt{2}\}$ for any choice of signs. In total, there are 8 constellation points, to compare with the 16 points in the PM-QPSK alphabet $\mathscr{M}_{\mathrm{PM}}=\{(\pm1\pm j,\pm1\pm j)^{\mathrm{T}}/2\}$ for any choice of signs. It may seem obvious that PM-CMA cannot be used for PS-QPSK since the two polarizations do not have a constant modulus, but PS-QPSK can alternatively be described as a subset of PM-QPSK [2]. Viewed in this way, PS-QPSK does have a constant modulus and the PM-CMA failure therefore needs a more detailed explanation.

One alternative symbol alphabet for PS-QPSK is $\tilde{\mathscr{M}}_{\mathrm{PS}}=\{(\pm1\pm j,\pm1\pm j)^{\mathrm{T}}/2\}$ for any sign combination with even parity, i.e., $E_{x,r}E_{x,i}E_{y,r}E_{y,i}>0$. The mapping from $\tilde{\mathscr{M}}_{\mathrm{PS}}$ to $\mathscr{M}_{\mathrm{PS}}$ can be done with a unitary Jones matrix and can therefore be performed both by the fiber and by the demultiplexing algorithm. To show this, we introduce the two unitary matrices

$$\mathbf{T}_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\-1&1\end{pmatrix}\quad\text{and}\quad\mathbf{T}_2=\begin{pmatrix}1&0\\0&e^{j\varphi}\end{pmatrix}. \tag{1}$$

The $\mathbf{T}_1$ matrix is a $\pi/4$ rotation of linearly polarized light and $\mathbf{T}_2$ is a phase retardation of the y polarization. Applying the $\mathbf{T}_1$ rotation to $\tilde{\mathscr{M}}_{\mathrm{PS}}$ and $\mathscr{M}_{\mathrm{PM}}$, respectively, the constellations in Figs. 1a and 1b are obtained. (Only the x polarization is plotted since the y polarization is identical.) The constellation in Fig. 1a corresponds to $\mathscr{M}_{\mathrm{PS}}$ and in Fig. 1b we see four additional points since all points no longer overlap after a $\mathbf{T}_1$ rotation.

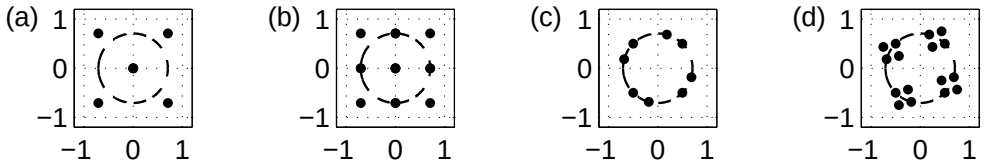

Fig. 1. (a) PS-QPSK after a $\mathbf{T}_1$ rotation. (b) PM-QPSK after a $\mathbf{T}_1$ rotation. (c) PS-QPSK after a $\mathbf{T}_1\mathbf{T}_2\mathbf{T}_1$ rotation. (d) PM-QPSK after a $\mathbf{T}_1\mathbf{T}_2\mathbf{T}_1$ rotation.

To explain the failure of PM-CMA for PS-QPSK, we construct the unitary Jones matrix $\mathbf{T}_1\mathbf{T}_2\mathbf{T}_1$ and for this example we choose $\varphi=\pi/6$. The resulting constellations for PS-QPSK and PM-QPSK are plotted in Figs. 1c and 1d, respectively. For PM-QPSK, the constellation no longer has a constant modulus and this will be corrected by PM-CMA. However, for PS-QPSK the constellation conserves its constant modulus property since all points lie on the circle. Thus, the problem with PM-CMA for PS-QPSK is *not* that there is no demultiplexing matrix that gives a constant modulus constellation. Instead, the problem is that obtaining a constant modulus constellation after equalization is not sufficient to have proper demultiplexing for PS-QPSK.

3. Description of the PS-CMA

3.1. Obtaining the cost function for PS-CMA

The cost function for PM-CMA is

$$J_{\text{PM-CMA}}=\mathbb{E}\left[\left(|E_x|^2-P_0\right)^2+\left(|E_y|^2-P_0\right)^2\right], \tag{2}$$

where, for QPSK modulation, the constant P_0 is the average power level for each polarization [7]. We rewrite this expression in terms of the Stokes parameters

$$s_0 = |E_x|^2 + |E_y|^2, \quad s_1 = |E_x|^2 - |E_y|^2, \quad s_2 = 2\,\mathrm{Re}(E_x E_y^*), \quad s_3 = 2\,\mathrm{Im}(E_x E_y^*), \tag{3}$$

with the result

$$J_{\text{PM-CMA}} = \mathbb{E}\left[\frac{(s_0 - 2P_0)^2}{2} + \frac{s_1^2}{2}\right]. \tag{4}$$

Thus, PM-CMA simultaneously makes the total power equal to $2P_0$ and s_1^2 as small as possible. The latter agrees with the fact that $s_1 = 0$ for all symbols in $\mathscr{M}_{\text{PM}}$. However, for PS-QPSK the alphabet $\mathscr{M}_{\text{PS}}$ is mapped to only two points on the Poincaré sphere since $(s_0, s_1, s_2, s_3) = (1, \pm 1, 0, 0)$. Thus, minimizing s_1^2 is no longer meaningful. Instead, we should minimize $s_2^2 + s_3^2 = s_0^2 - s_1^2$, where $s_0^2 = s_1^2 + s_2^2 + s_3^2$ was used. The two criteria that i) the total power is equal to a given constant P and ii) that $s_2^2 + s_3^2$ is minimal are combined into the cost function

$$J = \mathbb{E}\left[\frac{(s_0 - P)^2}{2} + Q\frac{s_2^2 + s_3^2}{2}\right] = \mathbb{E}\left[\frac{(|E_x|^2 + |E_y|^2 - P)^2}{2} + 2Q|E_x|^2|E_y|^2\right], \tag{5}$$

where the trade-off parameter Q must be positive for PS-QPSK. This more general cost function can be used to find update expressions for both PS-CMA and PM-CMA and switching between these two algorithms is done simply by changing the numerical parameters. This is convenient if PS-QPSK serves as a fall-back for PM-QPSK [4]. We set $Q = 1/2$ for PS-CMA and should take $Q = -1/2$ for PM-CMA. The parameter P is equal to the total average power for PS-QPSK and must be set to the average power of each polarization P_0 to reproduce the PM-CMA. We may also need to adjust the step size.

3.2. Equalizer update expressions

We denote the signal after equalization by $\mathbf{y} = (y_1, y_2)^{\mathrm{T}}$ and the samples of the received signal in time domain by the column vectors $\mathbf{x}_1$ and $\mathbf{x}_2$. These are related by the multitap equalizer column vectors according to

$$y_1 = \mathbf{h}_{11}^{\mathrm{T}}\mathbf{x}_1 + \mathbf{h}_{12}^{\mathrm{T}}\mathbf{x}_2, \quad \text{and} \quad y_2 = \mathbf{h}_{21}^{\mathrm{T}}\mathbf{x}_1 + \mathbf{h}_{22}^{\mathrm{T}}\mathbf{x}_2. \tag{6}$$

The direction in which the cost function changes most rapidly is found by differentiating with respect to the complex conjugated tap coefficients, see Theorem 3 of [8]. Performing the differentiation for the cost function in (5), we find the update rules

$$\mathbf{h}_{11}^{(k+1)} = \mathbf{h}_{11}^{(k)} - \mu\left[|y_1|^2 + (1+2Q)|y_2|^2 - P\right] y_1 \mathbf{x}_1^*, \tag{7}$$

$$\mathbf{h}_{12}^{(k+1)} = \mathbf{h}_{12}^{(k)} - \mu\left[|y_1|^2 + (1+2Q)|y_2|^2 - P\right] y_1 \mathbf{x}_2^*, \tag{8}$$

$$\mathbf{h}_{21}^{(k+1)} = \mathbf{h}_{21}^{(k)} - \mu\left[(1+2Q)|y_1|^2 + |y_2|^2 - P\right] y_2 \mathbf{x}_1^*, \tag{9}$$

$$\mathbf{h}_{22}^{(k+1)} = \mathbf{h}_{22}^{(k)} - \mu\left[(1+2Q)|y_1|^2 + |y_2|^2 - P\right] y_2 \mathbf{x}_2^*, \tag{10}$$

where μ is the step size and k is the iteration number. It is easily seen that the well-known update rules for PM-CMA [9] are recovered with $Q = -1/2$ and $P = P_0$.

3.3. Singularities in the PS-CMA and the PM-CMA

PM-CMA sometimes produces the same channel two times as output which is commonly referred to as the *singularity problem*. This behavior is not penalized by the cost function and

various ways to work around the problem have been suggested. For example, the problem is avoided in the single-tap case by enforcing the condition that the demultiplexing matrix should be unitary [10, 11]. Another way is to reinitialize the taps after some iterations [12].

To discuss the singularity problem we introduce the fiber Jones matrix $\mathbf{A}$, a demultiplexing matrix $\mathbf{B}$, and two singular constant matrices according to

$$\mathbf{M}_1 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{M}_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}. \tag{11}$$

The total transfer matrix $\mathbf{BA} = \mathbf{I}$ when $\mathbf{B} = \mathbf{A}^{-1}$ and this corresponds to zero cost in the absence of noise. However, the PM-CMA cost function does not define $\mathbf{B}$ uniquely since also $\mathbf{B} = \mathbf{M}_1\mathbf{A}^{-1}$ leads to zero cost for any symbol from $\mathscr{M}_{\mathrm{PM}}$. This is the singularity problem for PM-CMA and there is a similar problem for PS-CMA, which we see by studying $\mathbf{B} = \mathbf{M}_2\mathbf{A}^{-1}$. Choosing any symbol from $\mathscr{M}_{\mathrm{PS}}$, the cost is zero and the second channel will carry zero power. The third bit, which can be viewed as the selection of the launch polarization, has therefore been lost. It is fair to say that PS-CMA is neither better nor worse than PM-CMA in this respect but it is easy to detect the failure for PS-CMA by checking the output channel power.

4. Numerical evaluation of PS-CMA and PM-CMA

We here perform a numerical investigation of the convergence and tracking capabilities of PS-CMA and PM-CMA for the single-tap demultiplexing case. A symbol sequence is drawn from either $\mathscr{M}_{\mathrm{PS}}$ or $\mathscr{M}_{\mathrm{PM}}$ and complex white Gaussian noise (AWGN) is added. For the convergence study, the fiber Jones matrix, $\mathbf{A}$, is drawn uniformly from the set of 2×2 unitary matrices and is then held constant. The demultiplexing matrix $\mathbf{B}$ is initially set to $\mathbf{I}$. Running many simulations using different $\mathbf{A}$ matrices we compute the probability of being above a given SNR penalty threshold at every iteration. We have set the convergence threshold value to be 1 dB SNR penalty and used the found probabilities as a measure of the convergence rate. The step size has been selected as to maximize the final probability of being below 1 dB penalty in all cases. The polarization tracking study is made by choosing the fiber matrix time evolution to be

$$\mathbf{A} = \begin{pmatrix} \cos\phi(t) & \sin\phi(t) \\ -\sin\phi(t) & \cos\phi(t) \end{pmatrix}, \tag{12}$$

where $\phi(t)$ is a linearly evolving phase. Running the algorithms on long sequences of symbols, we find the averaged value of the SNR penalty. Plotting this penalty as a function of the angular frequency of ϕ provides a quantitative measurement of the algorithm tracking capability.

The most relevant performance norm for a communication system is the BER. Although this would have been possible to use, it is quite computationally demanding. We have therefore instead used the SNR penalty, which is closely related to the BER. As is often done for multiple-input multiple-output (MIMO) systems, we use an SNR definition that includes the interference from the other PM channel. This method is also known as *signal to interference plus noise ratio* or *signal to interference ratio*, see for example [13]. The SNR is then calculated in the following way: Defining $\mathbf{C} = \mathbf{BA}$, we have $\mathbf{y} = \mathbf{Bx} = \mathbf{B}(\mathbf{Aa} + \mathbf{n}) = \mathbf{Ca} + \mathbf{Bn}$, where $\mathbf{a}$ is drawn from either $\mathscr{M}_{\mathrm{PS}}$ or $\mathscr{M}_{\mathrm{PM}}$ and $\mathbf{n}$ is the complex AWGN. We then have

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}. \tag{13}$$

Viewing the interference from the a_2-term in y_1 as part of the noise, we obtain

$$\mathrm{SNR}_1 = \frac{\mathbb{E}\left[|C_{11}a_1|^2\right]}{\mathbb{E}\left[|C_{11}n_1 + B_{12}a_2 + B_{12}n_2|^2\right]} = \frac{|C_{11}|^2 P_s}{|C_{12}|^2 P_s + 2\sigma^2(|B_{11}|^2 + |B_{12}|^2)} \tag{14}$$

for the first channel, where $P_s = \mathbb{E}[|a_1|^2] = \mathbb{E}[|a_2|^2]$. For perfect demultiplexing in the case without PDL, i.e., $C_{12} = 0$ and $|B_{11}|^2 + |B_{12}|^2 = 1$ since $\mathbf{B}$ is then unitary, we recover the nominal $\mathrm{SNR}^{\mathrm{nom}} = P_s/(2\sigma^2)$. The expression for SNR_2 is obtained in an analogous manner. Assuming that the channels are processed separately after the polarization demultiplexing, it is reasonable to define the SNR penalty using the channel with the largest penalty. Thus, assuming that all SNR values are given in dB, we define the SNR penalty after iteration k as

$$\mathrm{SNR}_k^{\mathrm{pen}} = \mathrm{SNR}^{\mathrm{nom}} - \min(\mathrm{SNR}_{1,k}, \mathrm{SNR}_{2,k}). \tag{15}$$

5. Numerical results for the single-tap equalizer

Figure 2a shows the convergence result for PS-CMA and PM-CMA when performing symbol-by-symbol updates. The thin lines correspond to an implementation suggested by Kikuchi [10], which can be used in the single-tap demultiplexing case. This method avoids the singularity problem by orthogonalizing the rows of $\mathbf{B}_k$ in every iteration. The blue and the red curves show the result for PS-QPSK and PM-QPSK, respectively, when noise has been added to make the BER $= 10^{-3}$. The black curve shows the case for PS-QPSK with equal SNR as for PM-QPSK.

In general, the performance for PS-CMA is similar as for PM-CMA. At equal BER, the PS-CMA converges marginally slower than PM-CMA, which is a consequence from the extra noise added for PS-QPSK to obtain BER $= 10^{-3}$. Comparing the results at equal amount of added noise, PS-CMA performs somewhat better than PM-CMA. It is seen that without orthogonalization, both algorithms have a probability floor due to the singularity. At the SNR values used, the probability of failure is 2–4%. The step size used for each case is indicated close to each curve.

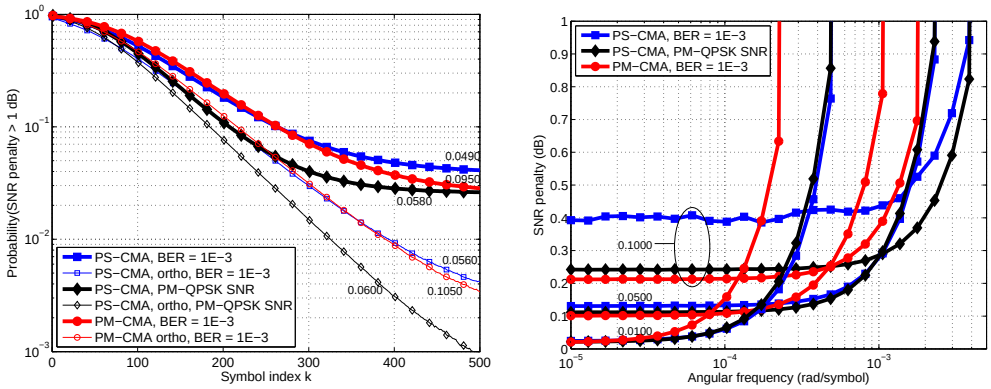

Fig. 2. (a) The probability to be above the SNR penalty limit of 1 dB as a function of the number of processed symbols. (b) The SNR penalty (in dB) as a function of the angular drift frequency of the fiber matrix. The step sizes are indicated close to each curve.

The result of the tracking comparison is seen in Fig. 2b. As indicated, simulations have been run with three different step sizes. In general, a large step size corresponds to better tracking capability but worse steady-state performance. We find that the tracking capability of PS-CMA well matches that of PM-CMA and both algorithms can track fast polarization changes. For example, 10^{-3} rad/symbol corresponds to 10 Mrad/s at 10 Gbaud, which is a very high polarization rotation rate. For the longest step size at BER $= 10^{-3}$, PS-CMA shows an increased steady-state SNR penalty. However, as seen in Fig. 2b, both better steady-state and tracking performance than PM-CMA can be obtained by reducing the step size.

As seen in Figs. 2a and b, the step sizes are longer than what is typically used in experimental investigations. Part of the reason for this is that, as seen from (7)–(10), the parameter μ is not dimensionless but depends on the signal power. Since PS-QPSK is a 4D modulation format, we have used unit energy 4D symbols. Thus, the energy per polarization is $1/2$, which leads to an increase of μ by a factor of four. For the convergence phase, it is beneficial to use a long step size in order to rapidly converge close to the demultiplexing matrix but for the following tracking phase, the step size can be reduced to improve the steady state performance. As seen from the tracking study, both algorithms are capable of tracking polarization changes that are much faster than what is typically seen in practice. Therefore, the step size should be decreased in order to reduce the steady-state penalty to a negligible value.

6. The impact from PMD and PDL

In addition to separating the two PM channels, the demultiplexing algorithm is also used to compensate for inter-symbol interference due to, e.g., residual chromatic dispersion (CD) and PMD. Furthermore, the algorithm should be able to operate in the presence of PDL. The effect of CD and differential group delay (DGD) can be described as an all-pass filter, i.e., in the frequency domain only the phase but not the amplitude is affected. Provided that the transfer function can be identified, such an effect can be compensated without penalty. We have investigated the ability of the PS-CMA to compensate for CD and DGD and found that close to ideal compensation is indeed possible. The conditions for this are that the number of taps is sufficiently large to account for the memory effect of the channel and that sufficiently much data is used to allow proper convergence. On the other hand, PDL introduces loss and leads to an SNR penalty that cannot be compensated by any equalizer.

To investigate the algorithm performance in the presence of PDL, the system model from Section 4 was changed by introducing PDL into the $\mathbf{A}$ matrix. This modified matrix, $\tilde{\mathbf{A}}$, was set up as the product of a random unitary matrix and a PDL matrix that had one transparent axis and one orthogonal lossy axis. The angle for the lossy axis was randomly selected from a uniform distribution. The obtained average SNR penalty has been plotted as a function of the amount of PDL in Fig. 3. In the worst possible alignment of the PDL [14], only one of the channels is attenuated and the following noise loading will lead to an SNR degradation equal to the amount of PDL. In the best case, the PDL is equally divided between the two PM channels. The worst and best cases are indicated by the dashed black lines. The averaged penalty is almost identical for all three simulated cases. As an example, a PDL of 3 dB causes an average SNR penalty of 2.4 dB. In these simulations, the convergence has been guaranteed by using an initial matrix that is sufficiently close to the system transfer matrix.

The penalty found above is caused by a combination of the signal power loss and the impact from PDL on the algorithm performance. It is therefore interesting to investigate how much of the penalty that is caused by the latter. To do this, we have compared the CMA algorithms with the minimum mean square error (MMSE) equalizer, which is known to maximize the SNR as defined by (15) [13]. In the numerical simulations we assume the channel response to be known in order to set up the channel-dependent MMSE demultiplexing matrix according to

$$\mathbf{B}_{\mathrm{MMSE}} = \left(\tilde{\mathbf{A}}^{\mathrm{H}} \tilde{\mathbf{A}} + \frac{2\sigma^2}{P_s} \mathbf{I} \right)^{-1} \tilde{\mathbf{A}}^{\mathrm{H}}. \tag{16}$$

Comparing the results from PS-CMA/PM-CMA and the MMSE equalizer we found that the difference in obtained SNR is very small. Indeed, both CMA algorithms seem to converge to the MMSE result when the step size is decreased. However, it is known [15] that the probability for failure to converge increases in presence of PDL. For example, introducing 3 dB PDL into the simulation shown in Fig. 2a increases the probability to fail to get within 1 dB of the

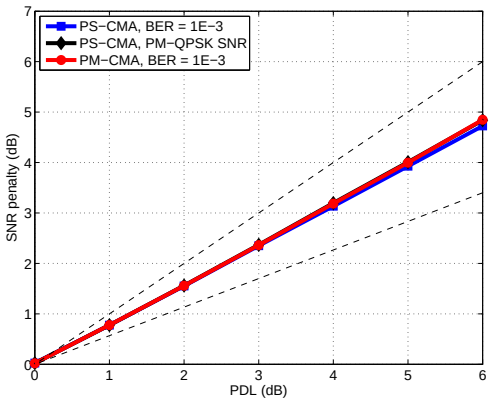

Fig. 3. The averaged SNR penalty as a function of the amount of PDL. The PDL alignments leading to the maximum and minimum penalties, respectively, have been used to plot the dashed black lines.

MMSE SNR penalty from the previous 2–4% to roughly 20% for both PS-CMA and PM-CMA. This increases the importance of monitoring to detect the potential convergence problems but provided this is done, both CMA algorithms are capable of obtaining SNR penalty values that are close to the MMSE result.

7. Conclusion

We have presented a generalized CMA cost function that allows polarization demultiplexing and equalization of both PM-QPSK and PS-QPSK. The resulting algorithm can be implemented in such a way that a switch between PS-CMA and PM-CMA is easy. Unfortunately, PS-CMA also shows a singularity problem. Numerical simulations have shown similar convergence and tracking performance for PS-CMA compared to PM-CMA. Thus, PS-CMA is in several ways a natural replacement for PM-CMA when using PS-QPSK.

Acknowledgment

This work was supported in part by the Swedish research council, the Swedish Governmental Agency for Innovation Systems (VINNOVA), and the EU EURO-FOS project. This work was performed within the Fiber Optic Communications Research Center (FORCE) at Chalmers.