



CHALMERS

Chalmers Publication Library

Blind Equalization in Optical Communications Using Independent Component Analysis

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

Journal of Lightwave Technology (ISSN: 0733-8724)

Citation for the published paper:

Nafta, A. ; Johannisson, P. ; Shtauf, M. (2013) "Blind Equalization in Optical Communications Using Independent Component Analysis". Journal of Lightwave Technology, vol. 31(12), pp. 2043-2049.

<http://dx.doi.org/10.1109/jlt.2013.2263831>

Downloaded from: <http://publications.lib.chalmers.se/publication/179793>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

Blind Equalization in Optical Communications Using Independent Component Analysis

Alon Nafta, Pontus Johannisson, and Mark Shtaif

Abstract—We propose a multi-tap independent component analysis (ICA) scheme for blind equalization and phase recovery in coherent optical communication systems. The proposed algorithm is described and evaluated in the cases of QPSK and 16-QAM transmission. Comparison with CMA equalization shows similar performance in the case of QPSK and an advantage for the ICA equalizer in the case of 16-QAM. The equalization scheme was evaluated in a multi-span optical communications system impaired by both polarization mode dispersion (PMD) and polarization dependent loss (PDL).

Index Terms—Coherent communication, optical fiber communication, equalization, independent component analysis (ICA).

I. INTRODUCTION

Advanced equalization techniques are key elements in modern coherent polarization multiplexed optical communications systems. Blind equalization, where the equalizer operates autonomously (i.e. without receiving predetermined control sequences from the transmitter) is particularly attractive from the standpoint of system simplicity. To date, with most optical communications systems relying on phase-shift keying (PSK) modulation, the most widely deployed equalizers are based on the constant modulus algorithm (CMA) [1], where the equalizer is designed to minimize variations in the amplitudes of the received samples. Equalizers based on CMA have been shown to be close to optimal in systems using binary phase-shift-keying (BPSK) and quadrature phase-shift-keying (QPSK) modulation [2]. In the case of quadrature amplitude modulation (QAM) with more than one amplitude level, the performance of CMA-based equalization deteriorates noticeably, motivating the search for alternative blind equalization approaches. This is particularly important in view of the rapidly growing rates of communications, which necessitate the transition to more spectrally efficient modulation formats.

In this paper we explore the possibility of using an alternative blind equalization method relying on independent component analysis (ICA) [3]. ICA was developed for situations in which extraction of statistically independent components from a measured mixture is necessary [4]. These algorithms were demonstrated in various fields of signal processing including biomedical informatics, image processing and wireless communications [5]–[7]. In the context of optical communications, ICA has been mostly considered for polarization demultiplexing applications, disregarding channel induced impairments such as polarization mode-dispersion (PMD) and polarization dependent loss (PDL). In [8] it has been shown by Zhang *et al.* that ICA solves the converge-to-the-same-source problem of CMA while maintaining similar polarization tracking capabilities. In [9], Xie *et al.* have shown that ICA can be used for blind demultiplexing of 16-QAM polarization multiplexed

signals, and in a recent paper by Johannisson *et al.* [10] the use of ICA was proposed for arbitrary constellations with multiple amplitude levels and was shown to significantly outperform CMA in terms of convergence rates.

In this paper we extend the work of [8]–[10] by considering the use of ICA not only for polarization demultiplexing, but also for the equalization of channel polarization impairments induced by PMD and PDL. To achieve that we extended the basic single-tap ICA algorithm of [11] to accommodate multi-tap operation [12], and integrated the simulation of the multi-tap ICA-equipped receiver with a detailed model of the link PMD [13] and PDL [14]. The performance of ICA-based equalization was compared with that of CMA in a variety of cases. We show that consistently with [10], the ICA algorithm displays faster convergence. The error-rate performance of ICA was found to be comparable to that of CMA in the case of QPSK transmission, but exceeds it considerably in the case of 16-QAM. This behavior has been observed consistently both in the presence of channel polarization effects as well as in their absence. The results that we present in what follows suggest that a multi-tap ICA algorithm is a viable candidate for replacing CMA, and even CMA combined with decision-directed least-mean-square (DD-LMS) in constellations higher than QPSK.

II. THEORETICAL BACKGROUND

A. Problem Formulation

Throughout our study we assume operation in the linear regime in which the effect of fiber nonlinearities is negligible. As is customary in studies of algorithms for mitigating polarization impairments, we assume that chromatic dispersion is compensated for by a separate module (optical or electronic) and hence it does not interfere with the process of equalization, which combines the procedures of demultiplexing and compensation for polarization impairments. The channel is modeled as consisting of multiple fiber spans separated by inline optical amplifiers. In each amplifier, unpolarized white Gaussian noise is added to the propagating signal. The noise contributions of different amplifiers are statistically independent and identically distributed. The effect of link PDL is accounted for by assigning statistically independent random PDL values to each of the optical amplifiers, jointly producing the prescribed average PDL value. The statistics of PDL in the individual sections is discussed in [14], and the statistics of overall link PDL (in decibels) is nearly Maxwellian, as discussed in [15]. This way of modeling PDL is consistent with the fact that in practical systems the PDL of the link is dominated by the amplifiers' contributions. The effect of

PMD is modeled according to [13] by considering the link as consisting of many statistically independent birefringent sections, jointly producing the prescribed value of the mean differential group delay (DGD).

B. The Instantaneous ICA Algorithm

Following the notation in [14], the optical signal at the receiver can be expressed as

$$\mathbf{x} = \mathbf{T}_0 \mathbf{s} + \mathbf{n} \quad (1)$$

where boldface lowercase symbols denote Jones vectors, and boldface uppercase symbols denote matrices. The two components of the Jones vectors in Eq. (1) represent the two polarization multiplexed signals. The vector \mathbf{x} represents the received signal and \mathbf{s} is the Jones vector representing the transmitted signal. The matrix \mathbf{T}_0 is the channel transfer matrix accounting for signal propagation through the entire link. The term \mathbf{n} represents the contribution of amplifier noise. Since the noise is also affected by the presence of PDL, it is partly polarized, i.e. its coherency function is not proportional to unity [14]. This property is reflected upon the performance of the link, although as in the case of CMA, it is not accounted for in the equalization procedure.

For the sake of simplicity, it is assumed that the receiver and transmitter are synchronized in terms of carrier frequency and sampling clock. This is a reasonable assumption as frequency estimation and clock recovery are typically performed separately from the equalizer, and are of little relevance to the choice of equalization algorithm. The ICA algorithm that we use closely follows the work of Novey and Adali [11], who have proposed an algorithm designed specifically for separation of QAM sources. We start by presenting the single-tap version of this algorithm, whereas the expansion to multi-tap operation (accounting for intersymbol interference effects) will be described in what follows. In the first stage of the algorithm we use the accumulated data samples in order to estimate the covariance matrix of the received Jones vector $\mathbf{\Lambda}_x = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$. We then define $\mathbf{z} = \mathbf{\Lambda}_x^{-1/2} \mathbf{x}$, to which we refer as the whitened data samples, and which can be written explicitly as

$$\mathbf{z} = \mathbf{A}\mathbf{s} + \tilde{\mathbf{n}}, \quad (2)$$

where $\mathbf{A} = \mathbf{\Lambda}_x^{-1/2} \mathbf{T}_0$ and the noise is $\tilde{\mathbf{n}} = \mathbf{\Lambda}_x^{-1/2} \mathbf{n}$. Assuming that the estimation of $\mathbf{\Lambda}_x$ is accurate, the covariance matrix $\mathbb{E}[\mathbf{z}\mathbf{z}^\dagger] \simeq \mathbf{I}$, where \mathbf{I} is the identity matrix, and in the limit of high signal-to-noise ratio (SNR) \mathbf{A} is a nearly unitary matrix [16]. Hence, we extract the original signals \mathbf{s} by iteratively searching for a unitary matrix \mathbf{W} that approximates \mathbf{A} 's inverse [3]. The ability to limit the search to unitary \mathbf{W} reduces the number of degrees of freedom, which explains the motivation for the above described whitening procedure. While a number of approaches for extracting \mathbf{W} have been considered, we follow the one that relies on the principle of maximizing the non-Gaussianity of the estimated signal $\hat{\mathbf{s}} = \mathbf{W}^\dagger \mathbf{z}$. The underlying idea is that the distribution of mixed signals is closer to Gaussian than the distribution of the individual entries, as implied by the central limit theorem.

One way to maximize the non-Gaussianity is by maximizing the negentropy [17] of $\hat{\mathbf{s}}$, which is the summed entropy of the components of $\hat{\mathbf{s}}$ subtracted from the summed entropy of the components of a Gaussian random vector of the same covariance matrix. Since \mathbf{W} is unitary, the covariance matrix of $\hat{\mathbf{s}}$ is identical to that of \mathbf{z} — namely it is constant and approximately equal to the identity matrix. Hence, maximizing the negentropy is equivalent to minimizing the sum of the component entropies of $\hat{\mathbf{s}}$. While this is a viable procedure, it requires estimation of the probability density of the various components in order for their entropy to be calculated. Ways to circumvent this difficulty have been proposed in the ICA literature, and in particular Novey and Adali [11] propose maximization of the following cost-function

$$J_{\text{ICA}}(\mathbf{W}) = \sum_{j=x,y} \mathbb{E}[\log p_a(\hat{s}_j)], \quad (3)$$

where

$$p_a(x) = \frac{1}{M2\pi\sigma^2} \sum_{i=1}^M e^{-\frac{(x^R - m_i^R)^2 + (x^I - m_i^I)^2}{2\sigma^2}}, \quad (4)$$

is an auxiliary function, which would be the probability density function of the samples of an M-QAM signal accompanied by additive Gaussian noise of variance σ^2 . Since $p_a(x)$ is not the true distribution of the signal, $J_{\text{ICA}}(\mathbf{W})$ is not rigorously related to the entropy. Nonetheless, as was claimed in [11], maximization of $J_{\text{ICA}}(\mathbf{W})$ is equivalent to the maximization of the negentropy. The value of σ^2 is selected so as to achieve satisfactory convergence of the optimization procedure. Throughout the simulations presented in this Paper the value of σ^2 was chosen as 0.5 and 0.2 for QPSK and 16-QAM correspondingly, which were found to produce satisfactory results [11]. Systematic optimization of the procedure with respect to the choice of σ^2 has not been performed within our study. The terms m_i are the ideal M-QAM constellation points and the superscripts R and I indicate real and imaginary parts, respectively. Note that the averaging denoted by the symbol \mathbb{E} in Eq. (3) is performed with respect to the received samples and not with respect to the auxiliary distribution p_a . In all the numerical studies presented in what follows the iterative search for the optimal \mathbf{W} was performed on the basis of Newton's method, following [11]. Other numerical approaches for obtaining the \mathbf{W} that maximizes J_{ICA} are of course possible.

C. The Multi-tap ICA Algorithm

We now extend the above described ICA algorithm to account for dispersive effects resulting from PMD. This problem falls into the category of non-instantaneous ICA, and a number of approaches to handling it have been proposed [18], [19]. Here we adopt a simple approach where the vectors and matrices are simply expanded in dimension in order to account for the non-instantaneous effects [3]. In the case of symbol-spaced sampling, and neglecting the possible effects of coding, subsequent samples are statistically independent, as are different polarization components of the launched signal.

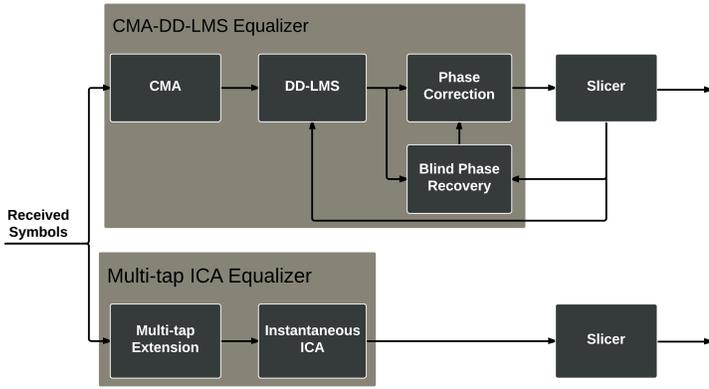


Fig. 1. Block diagram of the simulated system.

In that case, the symbol-separation problem can be reformulated as standard ICA. Specifically in the case of polarization multiplexing, the two dimensional Jones vectors considered earlier are replaced by $2L$ -component vectors, where L is the expected dispersive delay. Following [3], we define a time-extended input vector

$$\tilde{\mathbf{s}}(t) = [s_1(t), \dots, s_1(t-L+1), s_2(t), \dots, s_2(t-L+1)], \quad (5)$$

and similarly define

$$\tilde{\mathbf{x}}(t) = [x_1(t), \dots, x_1(t-L+1), x_2(t), \dots, x_2(t-L+1)], \quad (6)$$

where the first L components are samples of one polarization channel, whereas the subsequent L components are samples of the other polarization channel. Using these definitions, the convolutive mixing model can be written as

$$\tilde{\mathbf{x}}(t) = \tilde{\mathbf{A}}\tilde{\mathbf{s}}(t), \quad (7)$$

where $\tilde{\mathbf{A}}$ is a $2L \times 2L$ matrix containing the FIR filters' coefficients in a suitable order. Using this representation, we can apply the formalism described earlier with no changes to recover $\tilde{\mathbf{A}}$. Hence, in the first stage of the multi-tap algorithm, the samples of the received polarizations are used to generate the extended vectors defined in Eq. 6. In the second stage, the instantaneous ICA algorithm is applied to the extended vectors, recovering $\tilde{\mathbf{A}}$ which is then used to recover $s_1(t)$ and $s_2(t)$. Since the ICA cost function presented in (3) is optimized when the QAM symbols are aligned, the phase shift estimation problem is simplified to the detection of $\pi/2$ multiples. Moreover, since the cost function is not changed, the expression for the fixed-point update as presented in [11] is directly used in the equalizer implementation.

Traditionally, ICA algorithms use symbol-spaced sampling, assuming zero sampling offset. In the case of fractional sampling with more than one sample per symbol, the assumption of independence between adjacent samples is unjustified, but simple adaptations of the above approach can be considered [20]. The explicit evaluation of fractionally spaced ICA algorithms and their possible advantage in terms of sensitivity to timing errors will not be considered within the scope of the present work.

III. NUMERICAL ANALYSIS

A. Comparison with the CMA Algorithm

In what follows we describe the performance of the ICA equalization algorithm in the case of QPSK (4-QAM) and 16-QAM polarization multiplexed systems. We also show the results obtained with CMA-based equalization as a reference [12], [21], [22]. The CMA equalizer in our simulations was followed by a decision directed least-mean square (DD-LMS) equalization stage, as is common in most implementations [22]. The DD-LMS stage was not applied in the ICA equalization results that we present. Unlike the ICA implementation which used symbol-spaced sampling, we used fractional sampling in the case of CMA-DD-LMS with 2 samples per symbol, where the signal was upsampled and filtered using a raised cosine filter. The number of taps was $L = 5$ in the case of ICA and $L = 11$ in the case of CMA-DD-LMS, since its operation rate is twice the symbol rate, so that 4 symbol-delays are accounted for in both cases. This number was chosen to be large enough to mitigate the effects of PMD. The cost function used by the CMA algorithm was

$$J_{\text{CMA}}(\mathbf{W}) = \sum_{j=x,y} E \left[\left(|\hat{s}_j|^2 - R^2 \right)^2 \right], \quad (8)$$

where $R^2 = E[|s_j|^4] / E[|s_j|^2]^2$ with s_j denoting the signal launched into the j -th polarization channel — constant for a given constellation. The optimal matrix \mathbf{W} was found by the gradient descent method, whose update rule was

$$\begin{aligned} \mathbf{W}_{k+1} &= \mathbf{W}_k + \mu \cdot \hat{\nabla} J_{\text{CMA}}(\mathbf{W}_k) \\ &= \mathbf{W}_k + \mu \cdot \mathbf{x} \cdot \mathbf{e}_k^\dagger, \end{aligned} \quad (9)$$

where \mathbf{e}_k denotes the equalization error after the k -th iteration and is given by $\mathbf{e}_k = \mathbf{y}_k(R^2 - |\mathbf{y}_k|^2)$ and $\mathbf{y}_k = \mathbf{W}^\dagger \mathbf{x}$ is the k -th equalizer output. The step size parameter μ determines the convergence rate.

Figure 1 shows the block diagrams of the ICA and CMA equalizers, and the principle setup in which they were used for performance comparison. We note that the parameters of both equalization schemes can be optimized in different manners (emphasizing convergence speed vs. accuracy) and hence the choice of parameters under which the comparison is conducted necessarily involves some degree of arbitrariness. We chose to compare their performance under conditions where the convergence rate is similar, which we find satisfactory in the current scope of this work. Note that since the CMA algorithm is insensitive to the absolute phase, a phase recovery algorithm had to be implemented externally as illustrated in Fig. 1. Specific parameters used in the simulations are given in the section that follows.

B. Simulation Results

To ensure the reliability of our results, we reproduced the CMA results reported in [21] and [22] in certain configurations, obtaining consistent results. Blind phase estimation of QAM-modulated signals (which needs to be implemented externally in the CMA case) is considered a classic problem in digital communications theory, and many algorithms have been

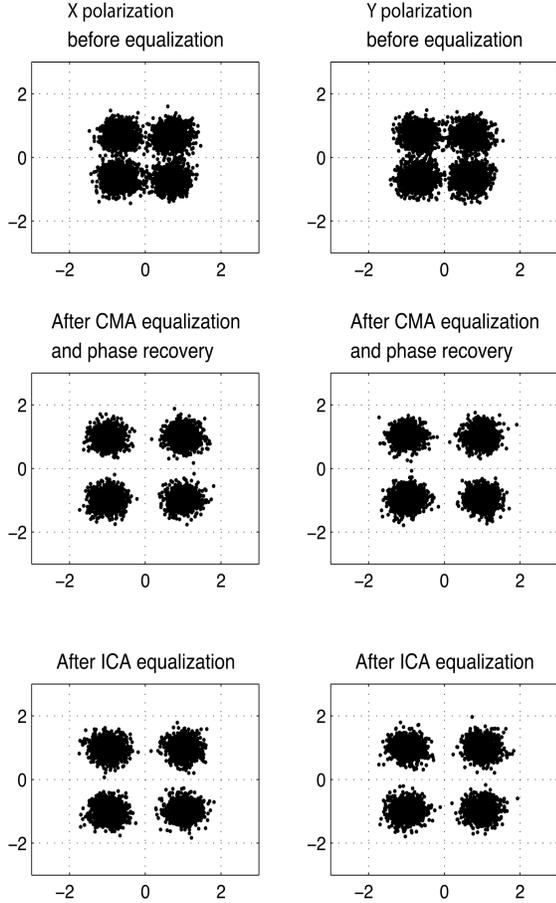


Fig. 2. Illustrated constellation diagrams for QPSK modulated signals, using the parameters described in section III, where each column shows a different polarization. Top panel: The received constellation. Middle panel: After CMA-DD-LMS. Bottom panel: After ICA equalization.

suggested for its solution. Here we used the algorithm of Chen *et al.* [23], which is standard in many practical implementations. The step-size [10], [21] in our implementation of the CMA was set to be $\mu = 0.001$. We also assumed block-type equalization in both the ICA and CMA cases, meaning that the equalizer receives the entire set of samples in each iteration, as opposed to symbol-by-symbol equalization in which the equalizer's coefficients update gradually in every sample. Although typically block-type equalization cannot be used by real-time equalizer implementations, it is plausible when the goal is to compare the performance of two equalization strategies. All simulations were performed with $N_s = 10$ amplified spans and with SNR values of 20 dB in the case of 16-QAM and 13.5 dB in the case of QPSK. [24] These values were chosen as they provide a 2 dB margin with respect to a nominal (back-to-back) bit-error-rate (BER) of 10^{-4} [25], and also since they allow some comparison with the results in [12], [21] and [22] which used similar SNR values. Figures 2 and 3 show the received and equalized constellations in the cases

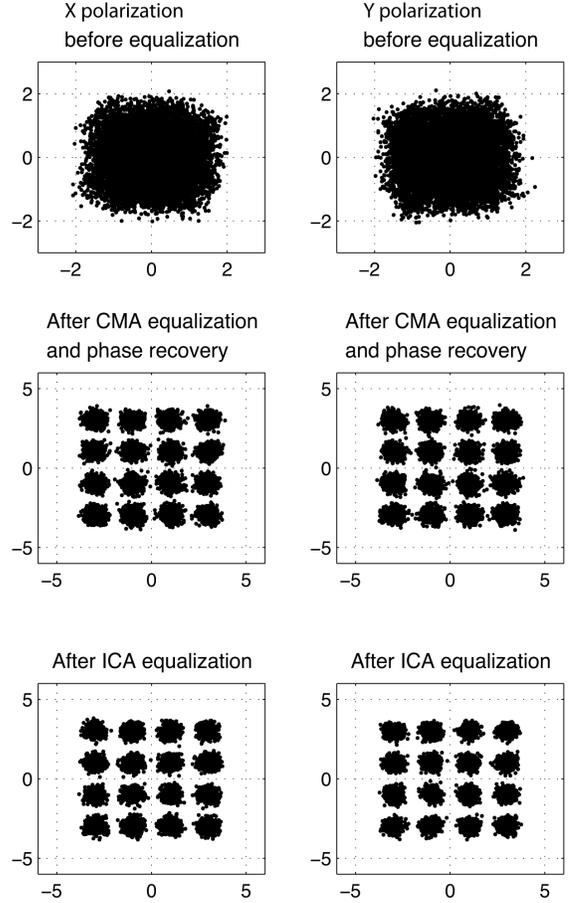


Fig. 3. Illustrated constellation diagrams for 16-QAM modulated signals, using the parameters described in section III, where each column shows a different polarization. Top panel: The received constellation. Middle panel: After CMA-DD-LMS. Bottom panel: After ICA equalization.

of QPSK and 16-QAM, respectively. The top panel in each of the figures shows the received signal prior to equalization. The center panel shows the constellation after CMA-DD-LMS and phase recovery. The third panel shows the constellation obtained after ICA equalization is applied. As can be seen in the figures, the ICA and CMA-DD-LMS equalizers both perform similarly in the case of QPSK, whereas in the 16-QAM case the ICA approach is seen to have somewhat better performance. This property is further quantified in Figs. 4–7. Figure 4 illustrates the convergence rates of the proposed ICA and the CMA-DD-LMS equalizers. The vertical axis shows the measured SNR reduction with respect to the nominal SNR of each scheme, the solid curves were plotted for the case in which the PMD was 0 and hence they represent polarization demultiplexing, similarly to [8]–[10]. The dotted curves were computed in the presence of PMD and PDL, where equalization is meant to mitigate these phenomena (in addition to polarization demultiplexing). In all cases the ICA algorithm is seen to converge in a satisfactory manner, slightly

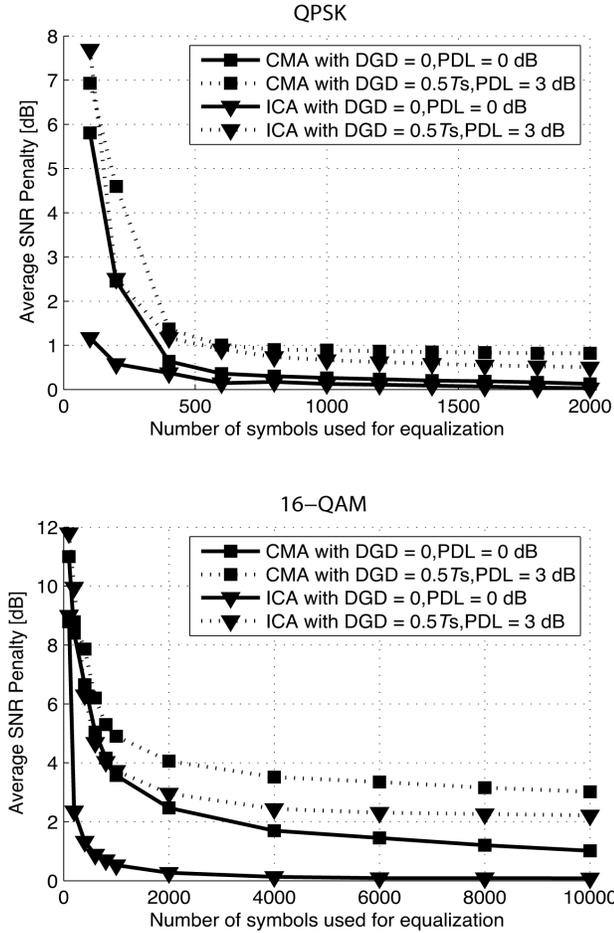


Fig. 4. The averaged (with respect to PMD and PDL) SNR penalty as a function of the number of symbols used for equalization for various combinations of PMD and PDL. The top panel corresponds to QPSK modulation and nominal SNR of 13.5 dB. The bottom panel corresponds to 16-QAM and 20 dB SNR. The SNR values were chosen to provide 2 dB margin in back to back with respect to BER of 10^{-4} .

faster than CMA. We remark that Fig. 4 was plotted only in order to verify acceptable convergence of the ICA algorithm and it is not based on a systematic optimization of parameters with respect to convergence rate for the two equalization methods.

In the top panel of Fig. 5 we show the symbol error-rate (SER) of the 16-QAM system as a function of the SNR for two sets of PDL and DGD values. These curves were plotted after averaging the measured penalties with respect to 150 fiber realizations and their role is to give an idea of the relative performance of the ICA and CMA schemes. In the bottom panel of Fig. 5 we show the SNR penalty as a function of mean DGD for varying PDL values. Here too the displayed SNR penalty is averaged with respect to fiber realizations. The advantage of ICA is evident in all cases. The averaging with respect to fiber realizations in Fig. 5 was performed in order to provide a rough quantitative idea of the comparison between the ICA and CMA schemes. A

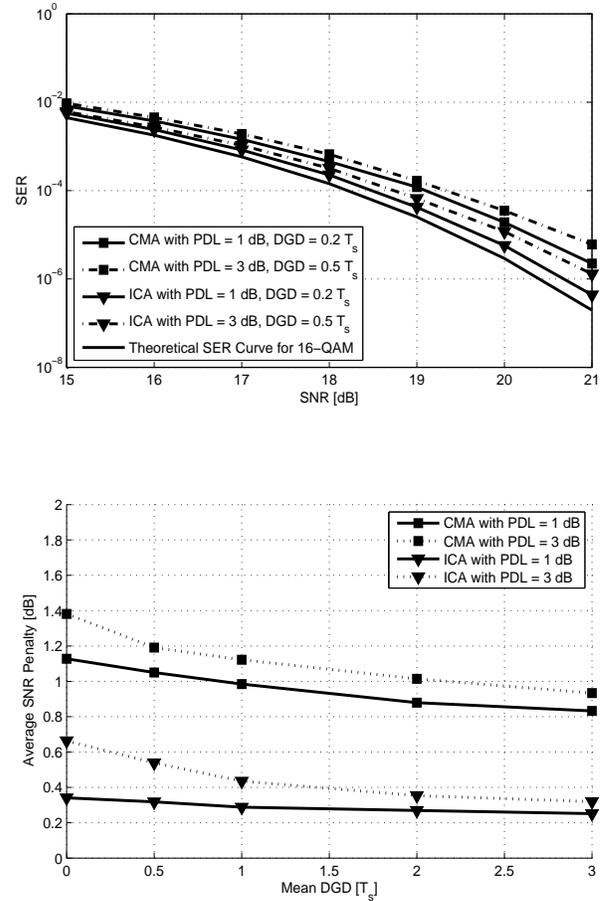


Fig. 5. Top panel: The SER as a function of SNR in the cases of ICA and CMA. The SER values in the plot are averaged with respect to PMD and PDL realizations. Bottom panel: Average SNR penalties as a function of the mean DGD with SNR= 20 dB. All curves were produced for 16-QAM transmission and each point in both panels was averaged with respect to 150 random fiber realizations.

more complete analysis would have to account for the PMD and PDL statistics, while focusing on the penalties that they induce with the probabilities associated with system outage (usually 4×10^{-5}). This procedure would however imply very long computation times and it was not included in the scope of our study. Nonetheless, in order to give some idea on the distribution of the penalties, we plot in Figs. 6 and 7 histograms of the SNR-penalties for a few sets of DGD and PDL values and for a fixed SNR of 13.5 dB, in the case of QPSK, and 20 dB in the case of 16-QAM. The randomness of the SNR penalties in the absence of PMD and PDL results from the randomness of the equalization algorithms in the presence of noise. In the case of QPSK, the variances in the SNR penalties are fairly small for both algorithms (of the order of 0.2 dB), but in the case of 16-QAM modulation the variance of ICA is noticeably smaller than that of CMA. In the large PMD and PDL regime, the SNR penalty randomness is mostly dominated by the randomness of PDL. As the displayed results

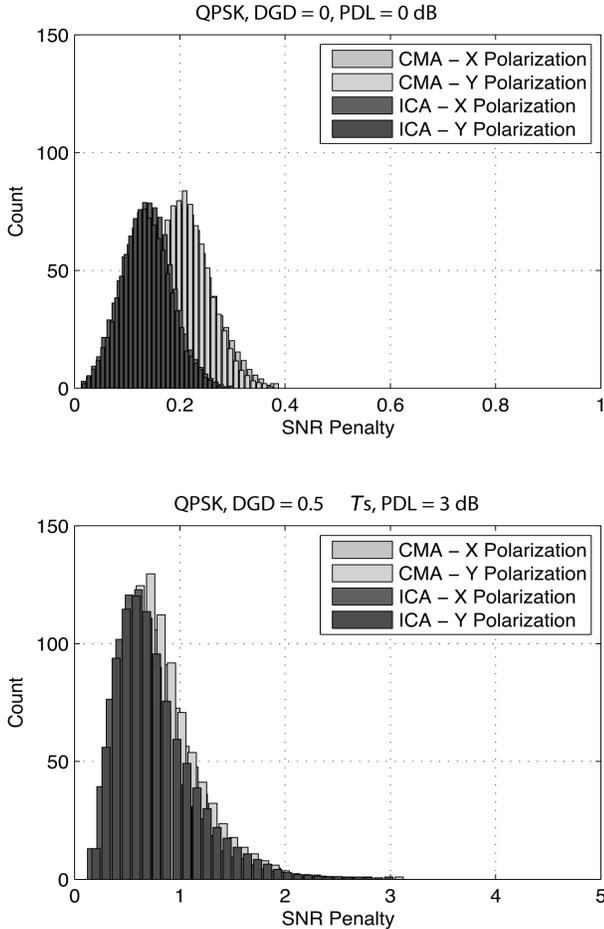


Fig. 6. Histograms of the SNR penalty in the case of QPSK, SNR= 13.5 dB. In the top panel polarization impairments are absent. In the bottom panel the mean DGD is half symbol duration and the average PDL is 3dB.

indicate, the ICA equalizer typically achieves lower SNR penalties, with and without PMD. As expected, the differences in typical SNR penalties are larger for 16-QAM modulated signals than for QPSK modulated signals.

IV. SUMMARY

Although complexity considerations are not a key factor in this study, a rudimentary comparison follows from estimating the number of multiplications that needs to be performed in each iteration. In the case of CMA the number of multiplications is known to be of the order $\mathcal{O}(NL)$ [26], where N is the number of data samples in a block and L is the filter length. The complexity of CMA is independent of the size of the constellation. In the case of ICA the number of multiplications grows linearly with the block length N , with the filter length L and with the number of constellation points M , so that the complexity is $\mathcal{O}(NLM)$. In both cases the dependence on N is linear, which is the most relevant property in the comparison. The extra factor of M appearing in the ICA's complexity should not be of major significance

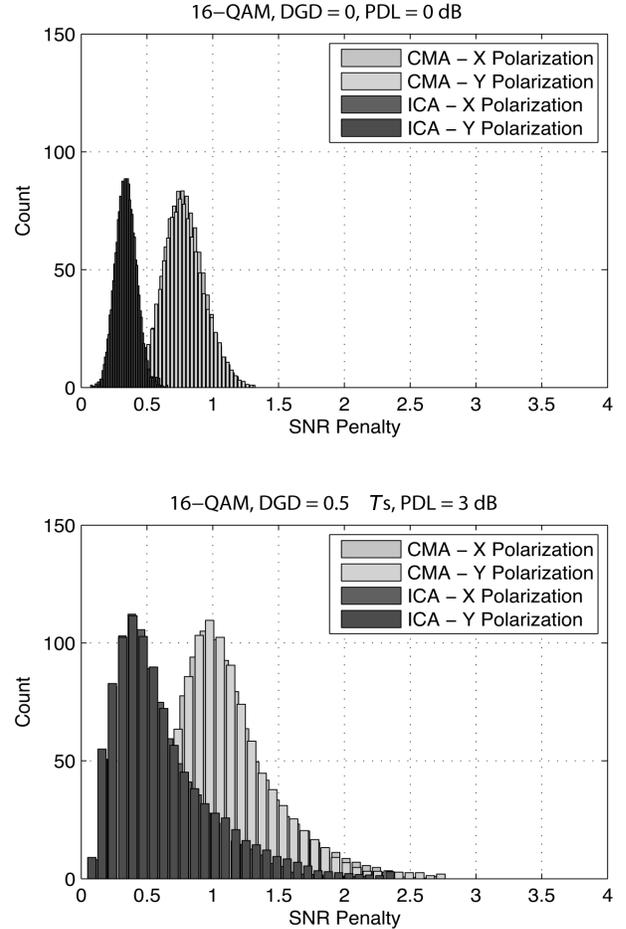


Fig. 7. Histograms of the SNR penalty in the case of 16-QAM, SNR= 20 dB. In the top panel polarization impairments are absent. In the bottom panel the mean DGD is half symbol duration and the average PDL is 3dB.

when comparing between the two schemes. Note that the ICA approach can be optimized for complexity, and as has been shown in [27] ICA with lower complexity than CMA can in principle be deployed. The proposed ICA algorithm recovers the absolute optical phase, thereby in principle eliminating the need for an external phase estimation module. However an explicit characterization of its performance in the presence of dynamic (e.g. laser) phase-noise is deferred for future studies.

To conclude, we proposed and analyzed a blind equalization scheme that relies on the ICA algorithm. It has been shown to perform similarly to CMA-DD-LMS equalization in the case of QPSK transmission and to noticeably outperform it in the case of 16-QAM. Our studies were conducted while taking into account the possible presence of PMD and PDL and hence a non-instantaneous ICA algorithm was used, and shown to successfully mitigate channel polarization effects. The ICA algorithm proposed in this paper can be further optimized in terms of performance, complexity and implementation aspects. Our results suggest that ICA should be a good candidate for replacing CMA and even CMA-DD-LMS in higher-than

QPSK constellations.

REFERENCES

- [1] S. J. Savory, "Digital Coherent Optical Receivers: Algorithms and Subsystems," *IEEE J. Sel. Topics. Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Sept.–Oct. 2010.
- [2] H. H. Zeng, T. Lang, C. R. Johnson Jr., "Relationships between the constant modulus and Wiener receivers," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1523–1538, Jul. 1998.
- [3] A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, Wiley, 2004.
- [4] J.-F. Cardoso, A. Souloumiac, "Blind beamforming for non-Gaussian signals," *Proc. Inst. Electr. Eng. Radar Signal Process.*, vol. 140, no. 6, pp. 362–370, Dec. 1993.
- [5] V. D. Calhoun, T. Adali, "Unmixing fMRI with independent component analysis," *IEEE Eng. Med. Biol. Mag.*, vol. 25, no. 2, pp. 79–90, Mar.–Apr. 2006.
- [6] M. S. Bartlett, J. R. Movellan, T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. Neural Netw.*, vol. 13, no. 6, pp. 1450–1464, Nov. 2002.
- [7] A. Hyvärinen, E. Oja, "Independent component analysis: algorithms and applications", *Neural Networks*, vol. 13, No. 4-5, pp. 411–430, Jun. 2000.
- [8] H. Zhang, Z. Tao, L. Liu, S. Oda, T. Hoshida, J. C. Rasmussen, "Polarization demultiplexing based on independent component analysis in optical coherent receivers," in *Proc. 34th Eur. Conf. Optical Communication (ECOC)*, Brussels, Belgium, Sept. 2008, pp. 1–2.
- [9] X. Xie, F. Yaman, X. Zhou, and G. Li, "Polarization demultiplexing by independent component analysis," *IEEE Photon. Technol. Lett.*, vol. 22, no. 11, pp. 805–807, Jun. 2010.
- [10] P. Johannisson, H. Wymeersch, M. Sjödin, A.S. Tan, E. Agrell, P. A. Andrekson, M. Karlsson, "Convergence Comparison of the CMA and ICA for Blind Polarization Demultiplexing," *J. Opt. Commun. Netw.*, vol. 3, no. 6, pp. 493–501, Jun. 2011.
- [11] M. Novey, T. Adali, "Complex Fixed-Point ICA Algorithm for Separation of QAM Sources using Gaussian Mixture Model," in *Proc. Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, Honolulu, HI, Apr. 2007, pp. II-445–448.
- [12] E. Ip, J. M. Kahn, "Digital Equalization of Chromatic Dispersion and Polarization Mode Dispersion," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, Aug. 2007.
- [13] M. Shtaif, A. Mecozzi, J. A. Nagel, "Mean-square magnitude of all orders of polarization mode dispersion and the relation with the bandwidth of the principal states," *IEEE Photon. Technol. Lett.*, vol. 12, no. 1, pp. 53–55, Jan. 2000.
- [14] M. Shtaif, "Performance degradation in coherent polarization multiplexed systems as a result of polarization dependent loss," *Opt. Exp.*, vol. 16, no.18, Sep. 2008.
- [15] A. Mecozzi, M. Shtaif, "The statistics of polarization dependent loss in optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 14, no. 3, pp. 313–315, Mar. 2002.
- [16] If the estimation of \mathbf{A}_x were exact, $E[\mathbf{z}\mathbf{z}^\dagger]$ was exactly unity and in the absence of noise, \mathbf{A} is rigorously unitary. In our case, these properties are approximate.
- [17] M. Novey, T. Adali, "Complex ICA by Negentropy Maximization," *IEEE Trans. Neural Netw.*, vol. 19, no. 4, pp. 596–609, Apr. 2008.
- [18] L. Parra, C. Spence, "Convolutional blind separation of non-stationary sources," *IEEE Trans. Speech Audio Process.*, vol.8, no.3, pp. 320–327, May 2000.
- [19] J. Thomas, Y. Deville, S. Hosseini, "Time-domain fast fixed-point algorithms for convolutional ICA," *IEEE Signal Process. Lett.*, vol. 13, no. 4, pp. 228–231, Apr. 2006.
- [20] S. R. Curnew, J. How, "Blind Signal Separation in MIMO OFDM Systems Using ICA and Fractional Sampling," in *Proc. Int. Symp. Signals, Systems and Electronics (ISSSE)*, Montreal, Canada, 2007, pp. 67–70.
- [21] I. Fatadin, D. Ives, S.J. Savory, "Blind Equalization and Carrier Phase Recovery in a 16-QAM Optical Coherent System," *J. Lightw. Technol.*, vol. 27, no. 15, pp. 3042–3049, Aug. 2009.
- [22] X. Xu, B. Châtelain, D. V. Plant, "Decision Directed Least Radius Distance Algorithm for Blind Equalization in a Dual-polarization 16-QAM System," in *Proc. Optical Fiber Commun. Conf. (OFC)*, Los Angeles, CA, Mar. 2012, p. OM2H.5.
- [23] L. Chen, H. Kusaka, "Blind Phase Recovery in QAM Communication Systems Using Higher Order Statistics," *IEEE Signal Process. Lett.*, vol. 3, no. 5, pp. 147–149, May 1996.
- [24] Throughout this paper we use the terms SNR as is common in the communications theory literature, i.e. defined as the average signal power divided by the noise variance of the sampled value.
- [25] For 16-QAM the symbol error rate curve can be found for example in J.G. Proakis, *Digital Communications*, 4th ed. p. 281, McGraw Hill, 2000. Conversion to BER assumes Grey coding.
- [26] M. Selmi, P. Ciblat, Y. Jaouën, C. Gosset, "Complexity Analysis of Block Equalization Approach for PolMux QAM Coherent Systems," in *Proc. Signal Process. Photon. Commun. (SPPCom)*, Toronto, Canada, Jun. 2011, p. SPTuC5.
- [27] V. Zarzoso, P. Comon, M. Kallel, "How fast is FastICA?" in *Proc. XIV Eur. Signal Process. Conf. (EUSIPCO)*, Florence, Italy, Sep. 2006.

Alon Nafta received his B.Sc. and B.A. degrees in electrical engineering and physics, respectively, from the Technion in 2004, and his M.Sc. degree in electrical engineering from the Tel Aviv University in 2012. In 2004, he joined an elite technology unit of the Israeli Defense Forces, where he worked on signal processing algorithms, communication technologies and security research. In 2012, he joined a small company in Palo Alto, CA, where he is currently conducting research in areas related to computer security.

Pontus Johannisson received his Ph.D. degree from Chalmers University of Technology, Göteborg, Sweden, in 2006. His thesis was focused on nonlinear intrachannel signal impairments in optical fiber communications systems. In 2006, he joined the research institute IMEGO in Göteborg, Sweden, where he worked with digital signal processing for inertial navigation with MEMS-based accelerometers and gyroscopes. In 2009, he joined the Photonics Laboratory, Chalmers University of Technology, where he currently holds a position as Assistant Professor. His research interests include, e.g., nonlinear effects in optical fibers and digital signal processing in coherent optical receivers.

Mark Shtaif received his M.Sc. and Ph.D. degrees in electrical engineering from the Technion in 1993 and 1997, respectively. In 1997 he joined the Light-wave Networks Research Department at AT&T Labs Research as Senior and then Principal Member of Technical Staff. At AT&T his work centered on modeling and characterization of optical fiber communication systems focusing on propagation effects in optical fibers including fiber nonlinearities, polarization mode dispersion, special modulation formats and interaction of signals and noise. During his employment he consulted to the AT&T business units on the integration of fiber-optic technologies. In 2000 he became a Principal Architect in Celion Networks, an optical networking company, where he worked on the analysis and design of long-haul optical transmission systems. In April 2002 he joined the School of Electrical Engineering at Tel-Aviv University, where he is teaching and conducting research in areas related to fiber-optics and optical communications. Prof. Shtaif is a fellow of the Optical Society of America (OSA).