



Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged PDM 16-QAM transmission

Zibar, Darko; Winther, Ole; Franceschi, Niccolo; Borkowski, Robert; Caballero Jambrina, Antonio; Arlunno, Valeria; Schmidt, Mikkel Nørgaard; Guerrero Gonzalez, Neil; Mao, Bangning; Ye, Yabin

Total number of authors:
12

Published in:
Optics Express

Link to article, DOI:
[10.1364/OE.20.00B181](https://doi.org/10.1364/OE.20.00B181)

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):

Zibar, D., Winther, O., Franceschi, N., Borkowski, R., Caballero Jambrina, A., Arlunno, V., Schmidt, M. N., Guerrero Gonzalez, N., Mao, B., Ye, Y., Larsen, K. J., & Tafur Monroy, I. (2012). Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged PDM 16-QAM transmission. *Optics Express*, 20(26), B181-B196. <https://doi.org/10.1364/OE.20.00B181>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged PDM 16-QAM transmission

Darko Zibar,^{*1} Ole Winther,² Niccolo Franceschi,^{1,2} Robert Borkowski,¹ Antonio Caballero,¹ Valeria Arlunno,^{1,2} Mikkel N. Schmidt,² Neil Guerrero Gonzales,³ Bangning Mao,³ Yabin Ye,³ Knud J. Larsen,¹ and Idelfonso Tafur Monroy¹

¹ DTU Fotonik, Technical University of Denmark, Build. 343, DK-2800, Denmark

² DTU Informatics, Technical University of Denmark, Build. 305, DK-2800, Denmark

³ European Research Center, Huawei Technologies Duesseldorf GmbH, Riesstrasse 25, Munich, Germany

* dazi@fotonik.dtu.dk

Abstract: In this paper, we show numerically and experimentally that expectation maximization (EM) algorithm is a powerful tool in combating system impairments such as fibre nonlinearities, inphase and quadrature (I/Q) modulator imperfections and laser linewidth. The EM algorithm is an iterative algorithm that can be used to compensate for the impairments which have an imprint on a signal constellation, i.e. rotation and distortion of the constellation points. The EM is especially effective for combating non-linear phase noise (NLPN). It is because NLPN severely distorts the signal constellation and this can be tracked by the EM. The gain in the nonlinear system tolerance for the system under consideration is shown to be dependent on the transmission scenario. We show experimentally that for a dispersion managed polarization multiplexed 16-QAM system at 14 Gbaud a gain in the nonlinear system tolerance of up to 3 dB can be obtained. For, a dispersion unmanaged system this gain reduces to 0.5 dB.

© 2012 Optical Society of America

OCIS codes: (060.0060) Fiber optics and optical communications; (060.1660) Coherent communications.

References and links

1. S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J Sel. Top. Quantum Electron* **16**, 1164–1179 (2010).
2. R.-J. Essiambre, G. Kramer, P. Winzer, G. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightwave Technol.* **28**, 662–701 (2010).
3. A. Lau and J. Kahn, "Signal design and detection in presence of nonlinear phase noise," *J. Lightwave Technol.* **25**, 3008–3016 (2007).
4. E. Ip and J. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightwave Technol.* **26**, 3416–3425 (2008).
5. Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *J. Lightwave Technol.* **29**, 2570–2576 (2011).
6. N. Stojanovic, Y. Huang, F. N. Hauske, Y. Fang, M. Chen, C. Xie, and Q. Xiong, "Mlse-based nonlinearity mitigation for wdm 112 gbit/s pdm-qpsk transmissions with digital coherent receiver," in *Proc. of OFC*, paper OTu3C.5, Los Angeles, California, USA, (2011).

7. D. Rafique, J. Zhao, and A. D. Ellis, "Compensation of nonlinear fibre impairments in coherent systems employing spectrally efficient modulation formats," *IEICE Trans. on Commun.* **E94-B**, 1815–1822 (2011).
8. D. Zibar, O. Winther, N. Franceschi, R. Borkowski, A. Caballero, A. Valeria, N. M. Schmidt, G. G. Neil, B. Mao, Y. Ye, J. K. Larsen, and T. I. Monroy, "Nonlinear impairment compensation using expectation maximization for pdm 16-qam systems," in *Proc. of ECOC*, paper Th1D2, Amsterdam, The Netherlands, (2012).
9. P. Winzer, A. Gnauck, C. Doerr, M. Magarini, and L. Buhl, "Spectrally efficient long-haul optical networking using 112-gb/s polarization-multiplexed 16-qam," *J. Lightwave Technol.* **28**, 547–556 (2010).
10. D. Zibar, J. C. R. F. de Olivera, V. B. Ribeiro, A. Paradisi, J. C. Diniz, K. J. Larsen, and I. T. Monroy, "Experimental investigation and digital compensation of dgd for 112 gb/s pdm-qpsk clock recovery," *Opt. Express* **19**, 429–437 (2011).
11. H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers / Synchronization, Channel Estimation, and Signal Processing* (Wiley, 1998).
12. J. Kurzweil, *An Introduction to Digital Communications* (John Wiley, 2000).
13. C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
14. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. Roy. Stat Soc. Series B.* **39**, 1–38 (1977).
15. N. G. Gonzalez, D. Zibar, A. Caballero, and I. T. Monroy, "Experimental 2.5-gb/s qpsk wdm phase-modulated radio-over-fiber link with digital demodulation by a k-means algorithm," *IEEE Photon. Technol. Lett.* **22**, 335–337 (2010).
16. A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links," *J. Lightwave Technol.* **30**, 1524–1539 (2012).
17. F. Vacondio, O. Rival, C. Simonneau, E. Grellier, L. Lorcy, J.-C. Antona, S. Bigo, and A. Bononi, "On nonlinear distortions of highly dispersive optical coherent systems," *Opt. Express* **20**, 1022–1032 (2012).
18. A. Bononi, N. Rossi, and P. Serena, "Transmission limitations due to fiber nonlinearity," in *Proc. of OFC*, paper OWO7, Los Angeles, California, USA, (2011).
19. E. Ip, N. Bai, and T. Wang, "Complexity versus performance tradeoff for fiber nonlinearity compensation using frequency-shaped, multi-subband backpropagation," in *Proc. of OFC*, paper OThF4, Los Angeles, California, USA, (2011).
20. S. Makovejs, D. S. Millar, D. Lavery, C. Behrens, R. I. Killey, S. J. Savory, and P. Bayvel, "Characterization of long-haul 112gb/s pdm-qam-16 transmission with and without digital nonlinearity compensation," *Opt. Express* **18**, 12939–12947 (2010).

1. Introduction

The application of digital signal processing (DSP) based coherent detection has allowed optical communication systems to operate closer to the nonlinear Shannon capacity limit by employing spectrally efficient modulation formats. Therefore, there is currently a lot of ongoing research on DSP based algorithms for signal detection and optical fibre channel impairment compensation. Linear signal processing algorithms can be effectively used to compensate for linear fibre channel impairments and have been demonstrated very successfully for higher order quadrature amplitude modulation (QAM) signaling [1]. However, for long-haul systems employing higher order QAM, nonlinear optical fibre impairments can severely limit the transmission distance as well as the achievable total capacity [2]. Mitigation of optical fibre nonlinearities is therefore very crucial as it will allow launching more power into the fibre and thereby enhancing the transmission distance. Additionally, mitigation of fibre nonlinearities will help us reduce the nonlinear crosstalk from the neighboring channel in a multi-channel transmission system.

It has been shown that nonlinear fibre impairments can be compensated by various techniques: digital backpropagation (DBP), maximum-likelihood sequence estimation, nonlinear polarization crosstalk cancellation, nonlinear pre- and post-compensation, RF-pilot, etc, [3–7] and references therein. Some of the mentioned methods suffer from complexity and, additionally, the achievable gain in the nonlinear tolerance is dependent on particular transmission scenarios. Therefore, efficient and widely applicable DSP algorithms for nonlinearity compensation are still open for research.

We have already demonstrated that for the dispersion managed links, the expectation maximization algorithm can be used to enhance system tolerance towards nonlinearities [8]. However, dispersion managed link will impact the signal propagation in a different way compared to

the dispersion unmanaged link, and it is therefore essential to investigate the benefits of EM for dispersion unmanaged link as well. In this paper, we consider both numerically and experimentally dispersion unmanaged link. Transmission distances of 240 km, 400 km and 800 km are investigated experimentally for the dispersion unmanaged link. For the consistency of the paper, results obtain for the dispersion managed link are included as well. We consider dispersion managed (link consisting of multiples stages of standard single mode fibre (SSMF) in combination with dispersion compensating fibre (DCF)) and dispersion unmanaged (link consisting of multiples stages of SSMF) polarization multiplexed 16-QAM single channel transmission. For the transmission link erbium doped fibre amplifier (EDFA) amplification is employed. First, it is investigated numerically, for the back-to-back case, if the expectation maximization (EM) algorithm can be effective in combating inphase and quadrature (I/Q) modulator nonlinearity, imbalance and laser linewidth. We investigate, also by numerical simulations, an improvement in nonlinear system tolerance that can be gained for PDM 16-QAM transmission. To begin with, we consider the case in which the chromatic dispersion is neglected and the system is impaired by self phase modulation induced nonlinear phase noise only. Finally, we move to a dispersion managed and unmanaged transmission system. As a proof of concept, an experimental set up employing dispersion managed and unmanaged PDM 16-QAM at 14 Gbaud is constructed and an improvement in the nonlinear system tolerance is investigated. The paper is ended with a conclusion and the future prospects of EM algorithm.

2. Numerical and experimental system set-up

In this section, a numerical and experimental set-up is presented. We first start by describing the numerical set-up used for simulations, and then we move to the experimental set-up. The section is concluded by a subsection describing the DSP algorithms used for signal equalization and demodulation.

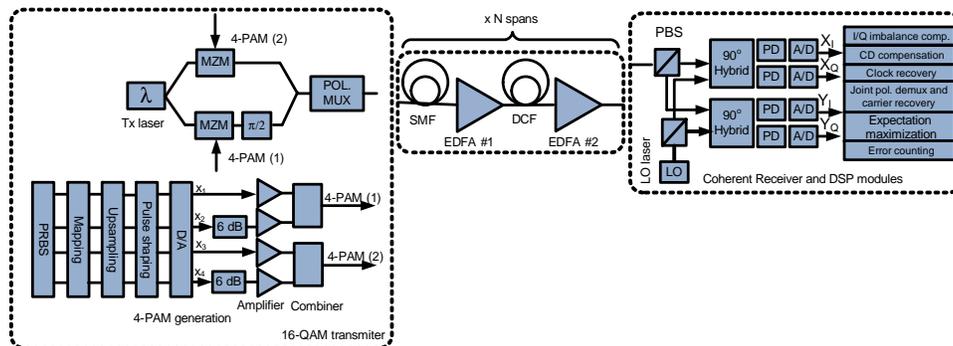


Fig. 1. Schematic diagram of the set-up used for simulations and experiment. PD: photodiode, PBS: polarization beam splitter, A/D: analog-to-digital converter, LO: local oscillator

2.1. Numerical set-up

The set-up used for the numerical investigations is shown in Fig. 1. All simulations are done using MATLAB (R2010a). For all numerical simulations, the baud rate is kept at 28 Gbaud resulting in the total bit rate of 224 Gb/s for the system under consideration. The transmitter and local oscillator (LO) laser phase noise is modeled as a random walk Wiener process. The output of the laser is then passed through an optical I/Q modulator. The I/Q modulator is driven by two four level pulse amplitude modulated (4-PAM) electrical signals, 4-PAM(1) and 4-PAM(2), in order to generate the optical 16-QAM signal. The module for generating 4-PAM

signals is shown in Fig. 1. It consists of pseudo-random binary sequence (PRBS) generator (generates four independent sequences of length $2^{15} - 1$), signal mapping, upsampling, pulse shaping filter (raised cosine), digital-to-analog converter (DAC), attenuators and electrical amplifiers. The actual impulse response of the driving amplifiers is not taken into consideration. It is assumed that the electrical amplifiers have sufficient bandwidth such that they don't induce any signal distortion. The method of 4-PAM signal generation is very similar to the one reported in [9]. The output of the I/Q modulator is then passed through a polarization multiplexing stage with a delay of 10 symbols, and the output is then amplified (EDFA). For the back-to-back numerical investigations, the generated PDM 16-QAM signal is coherently detected in a 90 degrees optical hybrid, photodetected and sampled at twice the baud rate by the analog-to-digital converter. We assume that the sampling frequency and the phase is not synchronized to the incoming signal and that the clock recovery is thereby performed by the DSP. The response of the analog-to-digital converter is modeled as a fourth-order Butterworth filter with a 3 dB bandwidth corresponding to 75% of the signal symbol rate. The sampled signal is then sent to the DSP modules which are described in subsection 2.3.

In this paper, we will also perform numerical investigations involving fibre transmission, and will therefore consider dispersion managed and unmanaged link. The dispersion managed link consists of a different number of stages where each stage consists of 80 km of SSMF and 17 km of a DCF. EDFA amplification is employed after the SSMF and the DCF spans, respectively, as shown in Fig. 1. For the SSMF we have the following fibre parameters: $\alpha_{smf}=0.2$ dB/km, $D_{smf}=17$ ps/nm/km and nonlinear coefficient is $\gamma_{smf}=1.3$ W⁻¹km⁻¹. For the DCF, we have the following parameters: $\alpha_{dcf}=0.5$ dB/km, $D_{dcf}=-80$ ps/nm/km and nonlinear coefficient is $\gamma_{dcf}=5.3$ W⁻¹km⁻¹.

2.2. Experimental set-up

The set-up used for the experimental investigations is very similar to the one shown in Fig. 1. For the experiment, the baud rate is kept at 14 Gbaud resulting in the total bit rate of 112 Gb/s. The transmitter and LO laser are both tunable external cavity lasers with a linewidth of ~ 100 kHz. The wavelength of the transmitter and LO laser is set to 1550 nm. Pulse-pattern generator outputs four copies, (x_1, x_2, x_3, x_4) , of a true PRBS of length $2^{15} - 1$. The PRBS sequences are first decorrelated by 270 bits, amplified and combined into a 4-PAM electrical signal. The two PRBS sequences x_1 and x_3 are independent, while x_2 and x_4 are inverted versions of x_1 and x_3 , respectively. The peak-to-peak amplitude of the 4-PAM signal used to drive an optical I/Q modulator is approximately 3 V. The delay in the polarization multiplexing stage is 10 symbols. Also, for the experiment we consider a dispersion managed and unmanaged link. For the experimental investigations we first consider dispersion managed link and then we move to dispersion unmanaged link. The dispersion managed link consists of 80 km of SSMF and 17 km of DCF with inline EDFA amplification. For the dispersion managed link, the DCF is just bypassed. The fibre parameters for SSMF and DCF used for the experiment are the similar to the one we have used in the numerical set-up. At the receiver, the 14 Gbaud PDM 16-QAM signal is then sampled at 50 Gs/s using a sampling scope with a nominal resolution of 8-bits and analog bandwidth of 17 GHz. The sampled signal is then send to DSP modules for the offline processing described in section 2.3.

2.3. Digital signal processing algorithms

The DSP modules consists of an I/Q imbalance compensation, interpolation (clock recovery) module, joint polarization demultiplexing and carrier recovery stage. We apply expectation maximization algorithm for nonlinearity compensation after joint polarization demultiplexing and carrier recovery. Within the expectation maximization algorithm, symbol demodulation is

embedded. The I/Q imbalance compensation algorithm employs Gram-Schmidt orthogonalization and is similar to the one reported in [1]. In order to make sure that we have a control signal for the clock recovery module, irrespective of the polarization mixing angle and the differential group delay, (DGD), a method reported in [10] is implemented. The implemented clock recovery module is a feedback structure similar to the one reported in [11]. It consists of an interpolator, timing error detector, loop filter and a number controlled oscillator (NCO). The timing error detector, which is the most crucial component, is a modified Gardner algorithm [10]. For the loop filter, an averaging filter is used. After, the clock recovery module a decimator is used in order to downsample the signal to one sample per symbol. The algorithm used for signal decimation is based on the maximum search method as proposed in [11]. The polarization demultiplexing stage is performed jointly with carrier frequency and phase estimation module. The polarization demultiplexing unit consists of a butterfly structure as the one reported in [1], and the carrier phase and frequency estimation unit is a decision-directed digital phase-locked loop [11]. The decisions from the digital phase-locked loop are then used as the error signal for the polarization demultiplexing. We found that the significant gain in the performance of the phase-locked loop can be obtained by properly designing the digital loop filter. We found that for the considered case proportional integrator filter was the best choice. We emphasize that the polarization demultiplexing and digital phase-locked loop are first trained in the blind mode using constant modulus algorithm and the switched to a decision directed mode as also reported in [1]. This type of joint equalization and phase/frequency estimation is described in more details in [12]. The EM algorithm is then applied after the polarization demultiplexing and carrier frequency/phase recovery stage. The task of the EM algorithm is then to learn the channel properties from the demodulated data without any prior knowledge. The information extracted from the channel is then used to compensate for the channel impairments and perform subsequent signal demodulation. After the EM stage, error counting is performed on ~ 100000 received symbols.

3. Theory

3.1. Statistical signal representation - mixture of Gaussians

In this section, we will first describe how a received signal can be modeled as a so called "Mixture of Gaussians, (MoG)" and then we will move into basic principles of the EM algorithm. For a more detailed treatment of MoG see [13], Chapter 9.2. Throughout, the entire section we will assume that the signal that is input to the EM algorithm contains one sample per symbol, and is obtained after polarization demultiplexing, frequency and phase recovery stage. We will refer to this signal as the demodulated signal.

The demodulated signal in x/y-polarization can be considered as a mixture of Gaussian densities (MoG) consisting of a number of components (clusters), where each of the components (clusters) can be described by a 2-D Gaussian distribution. For instance, in the case of 16-QAM, we have 16 clusters. For a 16-QAM signal constellation, in the absence of any impairment, we will have 16 distinct constellation points. However, in the presence of additive white Gaussian noise, around each of the 16 constellation points there will occur spread of symbols. The cluster is then defined as a grouping of the points/symbols around a mean value. Irrespective of the modulation format applied, (PSK or QAM), the demodulated signal can mathematically be expressed as a superposition of M Gaussian densities, where M is the number of clusters and corresponds to the number of constellation points. The probability density function of the demodulated signal is then expressed as:

$$p(\mathbf{x}) = \sum_{k=1}^M \pi_k N(\mathbf{x} | \mu_k, \Sigma_k), \quad (1)$$

where k refers to each cluster in the constellation, π_k is a mixing coefficient (for the considered case of a signal where symbols have uniform distribution $\pi_k = 1/M$). $\mathbf{x} = [x_1, x_2]$ is a 2-D vector, corresponding to a detected symbol in the constellation (Inphase/Quadrature) plane and $N(\mathbf{x}|\mu_k, \Sigma_k)$ is a 2-D Gaussian density with mean μ_k and a 2×2 covariance matrix Σ_k [13]:

$$N(\mathbf{x}|\mu_k, \Sigma_k) = \frac{1}{2\pi|\Sigma_k|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mu_k)^T \Sigma_k^{-1}(\mathbf{x}-\mu_k)}, \quad (2)$$

where $|\Sigma_k|$ is the determinant of the covariance matrix and it expresses the area covered by the specific cluster k . The covariance matrix is defined as:

$$\Sigma_k = \begin{bmatrix} \text{var}(x_1) & \text{cov}(x_1, x_2) \\ \text{cov}(x_1, x_2) & \text{var}(x_2) \end{bmatrix} \equiv \begin{bmatrix} \sigma_{1,1}^2 & \sigma_{1,2}^2 \\ \sigma_{2,1}^2 & \sigma_{2,2}^2 \end{bmatrix}. \quad (3)$$

For the most general case, each cluster is described by its specific covariance matrix Σ_k . However, when the signal is mostly dominated by the additive Gaussian noise, the covariance matrices will be equal and diagonal, i.e. there is no correlation among symbols within the cluster. Additionally, the clusters will be circularly symmetric (equal variances) and the covariance matrix is then expressed as:

$$\Sigma_k = \Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}. \quad (4)$$

An example of the demodulated 16-QAM signal dominated by additive Gaussian noise is shown in Fig. 2(a). It is observed in Fig. 2(a) that all the clusters look similar. An example of the demodulated signal strongly impaired by laser phase noise is shown in Fig. 2(b). It is observed in Fig. 2(b) that the clusters are not similar. Indeed, the clusters belonging to the outer ring are elliptical. Here, we will distinguish between two cases: (1) the covariance matrix is still diagonal and $\sigma_{1,1}^2 \neq \sigma_{2,2}^2$; the clusters are stretched in either vertical or horizontal direction, (2) the covariance matrix is non-diagonal and in this case the shape and orientation of the cluster is arbitrary, all depending if there is positive or negative correlation. Finally, let's look at third case when the demodulated signal is severely impaired by non-linear phase noise, see Fig. 2(c). It is observed that in Fig. 2(c) not only outer cluster are affected but all the clusters experience distortion. It should also be noticed that the entire constellation is tilted (phase offset introduced), and the outer points have been compressed. This compression means that the mean values μ_k have been altered compared to the reference constellation. By reference constellation, it is meant the constellation which is free of any impairment.

In general, different optical channel impairments will have a different imprint on the received signal constellation. This information can then be used to determine the impairment and make optimal signal detection as explained next.

The optimal signal detection in maximum likelihood sense is obtained by maximizing a posteriori probability of the received symbol \mathbf{x} belonging to one of the clusters k , where $k = 1, \dots, M$:

$$\hat{k} = \underset{k}{\operatorname{argmax}} p(k|\mathbf{x}) \quad (5)$$

or in another words find a cluster k for which $p(k|\mathbf{x})$ is maximized. The a posteriori probability $p(k|\mathbf{x})$ is obtained from Bayes' theorem [13]:

$$p(k|\mathbf{x}) = \frac{\pi_k N(\mathbf{x}|\mu_k, \Sigma_k)}{\sum_{l=1}^M \pi_l N(\mathbf{x}|\mu_l, \Sigma_l)}. \quad (6)$$

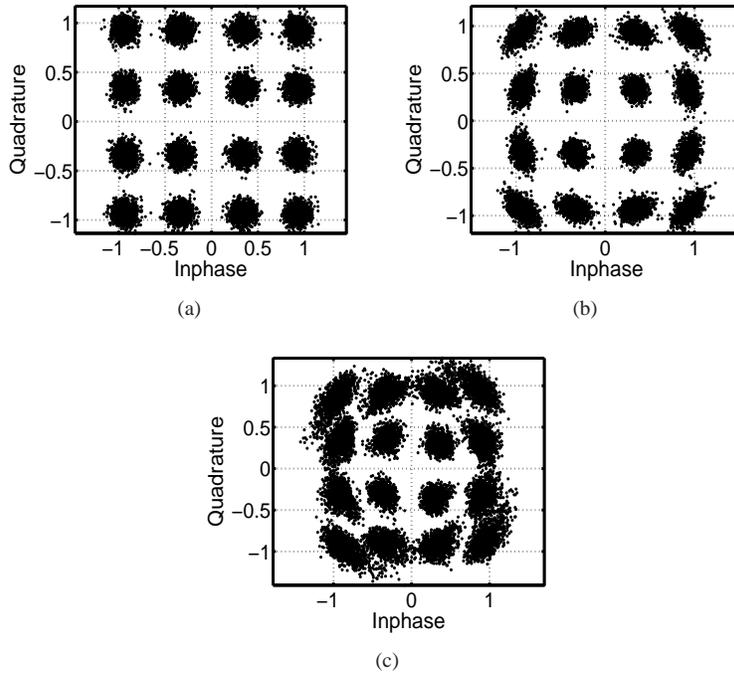


Fig. 2. Impact of different impairments on signal constellation for a 16-QAM signal. (a) Constellation of a signal dominated by additive noise. (b) Constellation of a signal dominated by phase noise. (c) Constellation of a signal dominated by non-linear phase noise.

Inserting the expression for the Gaussian distribution, the optimal decision, Eq. (5), reduces to a quadratic decision rule:

$$\hat{k} = \operatorname{argmax}_k \left\{ -\frac{1}{2} \mathbf{x}^T \Sigma_k^{-1} \mathbf{x} + \mathbf{w}_k^T \mathbf{x} + w_{k0} \right\} \quad (7)$$

with $\mathbf{w}_k = \Sigma_k^{-1} \boldsymbol{\mu}_k$ and $w_{k0} = \log \pi_k - \log |\Sigma_k|/2 - \boldsymbol{\mu}_k^T \Sigma_k^{-1} \boldsymbol{\mu}_k/2$. In the case, when the covariance matrices are equal, the quadratic term in optimal decision rule in Eq. (7) are the same for all k and the decision rule becomes linear:

$$\hat{k} = \operatorname{argmax}_k \left\{ \mathbf{w}_k^T \mathbf{x} + w_{k0} \right\} . \quad (8)$$

However, in case when the signal constellation is distorted by nonlinear phase noise, laser phase noise, etc, Eq. (6) needs to be used in order to make optimum signal detection. In order to evaluate Eq. (6), M Gaussian densities, $N(\mathbf{x}|\boldsymbol{\mu}_k, \Sigma_k)$ and thereby parameters $\boldsymbol{\pi} \equiv \{\pi_1, \dots, \pi_k\}$, $\boldsymbol{\mu} \equiv \{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_k\}$ and $\boldsymbol{\Sigma} \equiv \{\Sigma_1, \dots, \Sigma_k\}$ describing Gaussian densities need to be determined. Next, we will show how to use a powerful method of EM in order to determine the parameters that generate the Gaussian mixture model. The EM will determine in a maximum likelihood sense the most likely parameters $\boldsymbol{\Xi} = [\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}]$ that generated Gaussian densities.

3.2. Expectation maximization algorithm

In general, the EM is a numerical method of producing a solution to a maximum likelihood estimation for problems which can be simplified by introducing latent variables [13, 14]. In the

mixture model context the latent variables are the assignments. The set of parameters Ξ to be estimated, in the maximum likelihood sense, from the demodulated signal \mathbf{X} is governed by the following expression:

$$\hat{\Xi} = \underset{\Xi}{\operatorname{argmax}} p(\mathbf{X}|\Xi), \quad (9)$$

where $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$, N is the length of the observation interval and the likelihood function of Ξ , $p(\mathbf{X}|\Xi)$, for independent identically distributed data is expressed as:

$$p(\mathbf{X}|\Xi) = \prod_{n=1}^N p(\mathbf{x}_n|\Xi) = \prod_{n=1}^N \sum_{k=1}^M \pi_k N(\mathbf{x}_n|\mu_k, \Sigma_k). \quad (10)$$

No closed-form analytical solution for Eq. (9) is available. Therefore, the iterative EM framework can be used to find a solution. The EM is a two step iterative procedure which is guaranteed to converge to the (local) maximum likelihood solution given in Eq. (9) [14]. The two step procedure, so called expectation (E) step and maximization (M) step for the particular case considered in this section is as follows [13]:

$$\mathbf{E}\text{-step} : \gamma_{nk} \equiv p(k|\mathbf{x}_n) = \frac{\pi_k N(\mathbf{x}_n|\mu_k, \Sigma_k)}{\sum_{l=1}^M \pi_l N(\mathbf{x}_n|\mu_l, \Sigma_l)} \quad \text{for } n = 1, \dots, N \quad \text{and } k = 1, \dots, M \quad (11)$$

$$\mathbf{M}\text{-step} : N_k = \sum_{n=1}^N \gamma_{nk} \quad (12)$$

$$\pi_k = \frac{N_k}{N} \quad (13)$$

$$\mu_k = \frac{1}{N_k} \sum_{n=1}^N \gamma_{nk} \mathbf{x}_n \quad (14)$$

$$\Sigma_k = \frac{1}{N_k} \sum_{n=1}^N \gamma_{nk} (\mathbf{x}_n - \mu_k)(\mathbf{x}_n - \mu_k)^T \quad \text{for } k = 1, \dots, M, \quad (15)$$

where γ_{nk} is called the responsibility and is nothing but a posteriori probability, Eq. (6), needed for optimal decisions.

The flow-chart describing the algorithm is shown in Fig. 3. To begin with, we initialize the EM with initial parameters for the means, covariance matrices and mixing coefficients and then the algorithms start to iterate in order to find most likely parameters. In the E-step, the current values of the parameters, Ξ^i at the iteration i , are used to evaluate the Eq. (11). The E-step expressed by Eq. (11) computes the probability of the received symbol belonging to one of the clusters, i.e. posterior probability. In the M-step we use those probabilities to re-estimate the parameters Ξ . In other words, in the M-step we are trying to find the parameters that maximize the probability that the data has been generated by a particular cluster. When making a parameter update resulting from the E step and followed by the M step, the likelihood function, $P(\mathbf{X}|\Xi^i)$, on the parameters will increase and will flatten out when the algorithm has converged. The convergence properties of the EM strongly dependent on the initialization. For the considered cases throughout the paper, we found that the EM algorithm will converge after 3 iterations. Once the EM algorithms has converged ($i > N_{iter}$), we use the results to perform the optimum signal detection governed by Eq. (5).

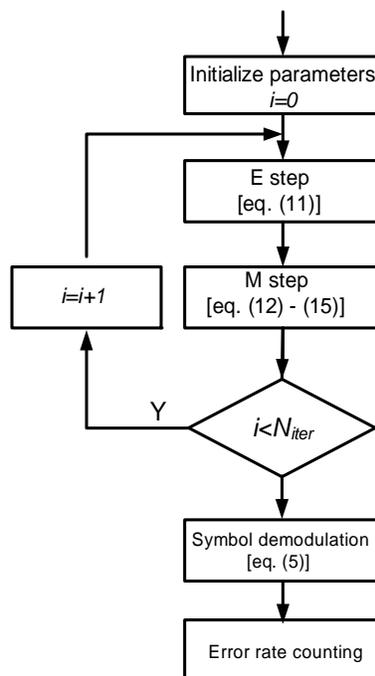


Fig. 3. Flow-chart illustrating the steps performed by the EM algorithm and subsequent signal demodulation. N_{iter} denotes the number of specified iteration.

4. Simulation results

4.1. Back-to-back investigation

First, we consider back-to-back case. It is investigated how the EM algorithm can be used to combat the combined effects of I/Q modulator nonlinearities and imperfection, and combined laser linewidth. We deliberately drive the I/Q modulator with large peak-to-peak amplitude, V_{pp} , of the electrical 4-PAM signal such that the constellation diagram of the 16-QAM signal is distorted. The resulting modulation depth of the modulator is then $m = V_{pp}/V_{\pi} = 2.12$. Furthermore, it is assumed that the phase shift between the inphase and quadrature branch of the I/Q modulator deviates from $\pi/2$ by 5%. In Fig. 4, $-\log[\text{BER}]$ is plotted as function of the combined laser linewidth for the optical signal to noise ratio (OSNR) of 25 dB. Figure 4 shows that compared to the case when no compensation is used (linear decision boundaries), the EM is very efficient in combating the combined impairments originating from I/Q modulator nonlinearities, non-ideal phase shift between the I and Q branches and combined laser linewidth.

4.2. Dispersion managed link

In order to investigate the effects of nonlinear phase noise only, the dispersion is numerically set to zero, i.e. transmission link consisting of 12 spans of SSMF with dispersion set to zero. The dominant nonlinear impairment is therefore the nonlinear phase noise originating from the EDFA amplification along the link. In order to determine the effectiveness of the EM algorithm, we also plot in the same figure $-\log(\text{BER})$ when k-means algorithm is used [13, 15]. The k-means algorithm, is another widely used clustering algorithm in nonlinear signal processing. It computes the means of the clusters and the information is used to perform minimum distance signal detection expressed by Eq. (8). In Fig. 5(a), $-\log(\text{BER})$ is plotted as a function of the

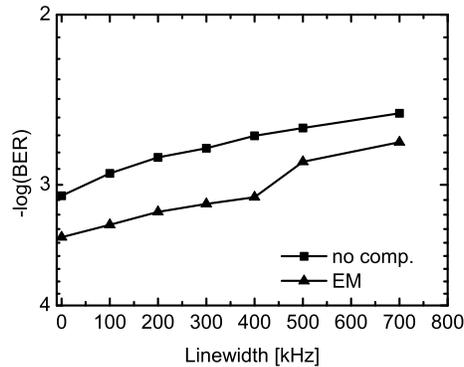


Fig. 4. BER as a function of combined laser linewidth for the back-to-back-case. The modulator modulation depth, $m = V_{pp}/V_{\pi} = 2.12$, I/Q imbalance: 5% and OSNR is 25 dB.

input signal power to the transmission span. In general, there is a relatively large improvement in the nonlinear system tolerance when using the EM algorithm. Figure 5(a) shows that by only using the k-means algorithm the system tolerance can be increased by approximately 1 dB for $-\log(\text{BER})$ of 3, while by using the EM, the system tolerance can be increased by 2.2 dB compared to the case when neither the EM nor k-means is used.

Next, we consider dispersion managed link consisting of SSMF and DCF as described in section 2.1. The total number of spans is 12. The results of $-\log(\text{BER})$ as a function of span input power are shown in Fig. 5(b). It is observed from Fig. 5(b) that by employing the EM system tolerance towards nonlinearities can be increased. The improvement in the nonlinear system tolerance is approximately 1.2 dB for the $-\log(\text{BER})$ of 3.

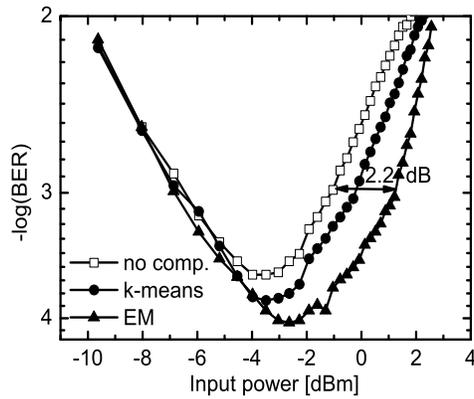
4.3. Dispersion unmanaged link

Next, we move to the case of dispersion unmanaged link. In Fig. 6(a), $-\log(\text{BER})$ is plotted as a function of the span input signal power. The total number of spans is 15. It is observed that there is no benefit of using the EM algorithm. This is because for the dispersion unmanaged links, the impact of the nonlinearity can be modeled as additive Gaussian white noise [16–18]. Therefore, all 16 covariance matrices are circularly symmetric and the optimum decision boundaries become linear as explained in Section 3. As the laser linewidth will also have an imprint on the signal constellation, next, we want to investigate if dispersion unmanaged links can benefit from the EM when impaired by laser linewidth and operate in the nonlinear regime. In Fig. 6(b), the $-\log(\text{BER})$ is plotted as a function of the combined laser linewidth. The number of spans for the unmanaged link is set 15 and the input signal power to the span is 1.4 dBm. It is observed in Fig. 6(b) that some improvement in the laser linewidth tolerance can be obtained when applying EM.

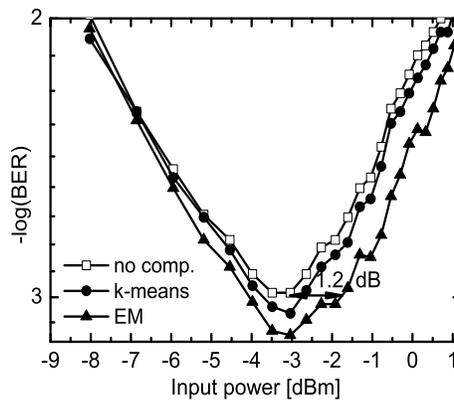
5. Experimental results

5.1. Dispersion managed link

First, we will demonstrate how EM can be effectively used to extract information from a severely distorted constellation and use this information to mitigate the impairments. Figure 7, shows a constellation diagram of a signal impaired by nonlinear phase noise after one span transmission through the dispersion managed link. We plot the recovered constellation diagram



(a)



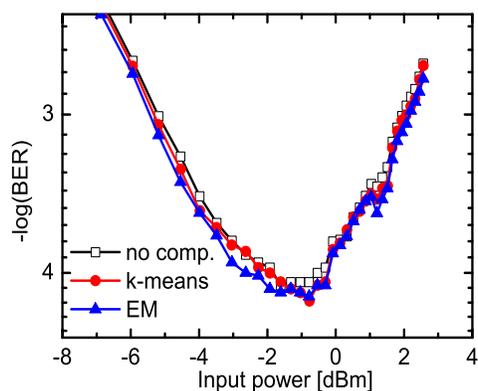
(b)

Fig. 5. The total number of spans is 12 and the combined laser linewidth is 200 kHz. (a) BER as a function of span input power for NLPN dominated transmission link, (dispersion numerically set to zero). (b) BER as a function of span input power for dispersion managed link. Transmission link consists of SSMF and DCF.

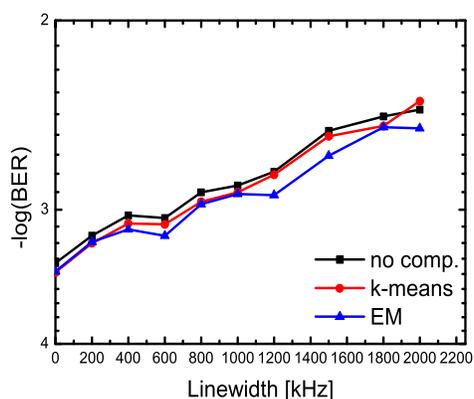
of the x polarization after the carrier recovery stage. Together with the constellation, we have also plotted optimal (nonlinear) decision boundaries obtained by applying Eq. (6) in conjunction with the EM algorithm.

It is observed from Fig. 7, that due to the nonlinear impairments, especially the outer constellation points are distorted. The inner constellation points are less affected due to lower power, however, all clusters experience a significant phase shift. In the case, no compensation is applied $-\log(\text{BER})$ is 1.30. For the case when the k-means is used the respective $-\log(\text{BER})$ is 2.04 while when the EM is applied $-\log(\text{BER})$ gets down to 3. This example demonstrates the capabilities of EM of compensating distorted and phase shifted constellations.

In Fig. 8(a), we plot the demodulated signal constellation for the input power of $P_{in} = 0$ dBm and the corresponding optimal decision boundaries after 800 km of transmission. It is observed that the demodulated signal constellation shown in Fig. 8(a) is distorted and therefore



(a)



(b)

Fig. 6. The total number of spans is 12. (a) BER as a function of span input power for dispersion unmanaged link. The combined laser linewidth is 200 kHz. (b) BER as a function of combined laser linewidth for dispersion unmanaged link.

the optimal decision boundaries are nonlinear. In Fig. 8(b), we plot $-\log(\text{BER})$ as a function of span input power after 800 km of transmission through dispersion managed link. It is observed that there is an improvement in the nonlinear system tolerance by employing the EM algorithm which is in accordance with simulation results. We observe up to 3 dB of improvement in nonlinear tolerance compared to the case when no compensation is used. The reason why we get more improvement for the experimental data may be attributed to the fact that the EM is also effective in compensating residual distortion induced on the signal. It is observed from the figure that only very little improvement can be obtained by using the k-means algorithm, and this is also in good agreement with the simulation results.

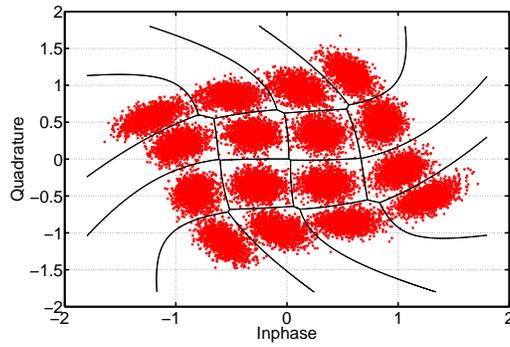
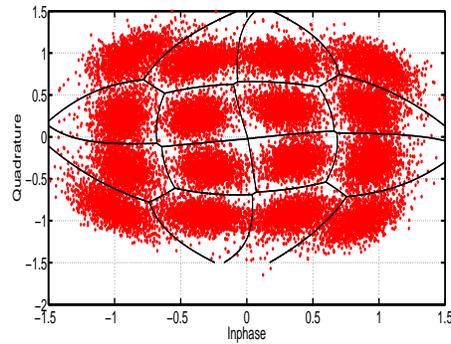
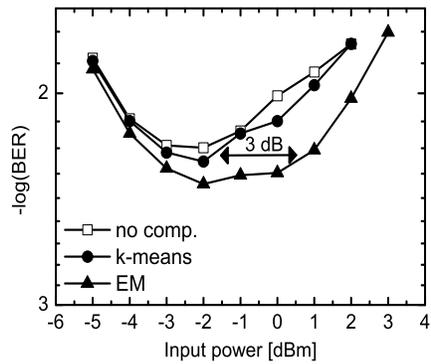


Fig. 7. Recovered constellation diagram impaired by nonlinear phase noise. Only a single transmission span is considered



(a)



(b)

Fig. 8. (a) Constellation diagram of the demodulated signal after 800 km of transmission through dispersion managed link. (b) BER as a function of span input power for dispersion managed link after 800 km of transmission.

5.2. Dispersion unmanaged link

For our next investigations, the DCF is removed and we consider 800 km of dispersion unmanaged signal transmission. In Fig. 9(a), the constellation diagram of the recovered signal is

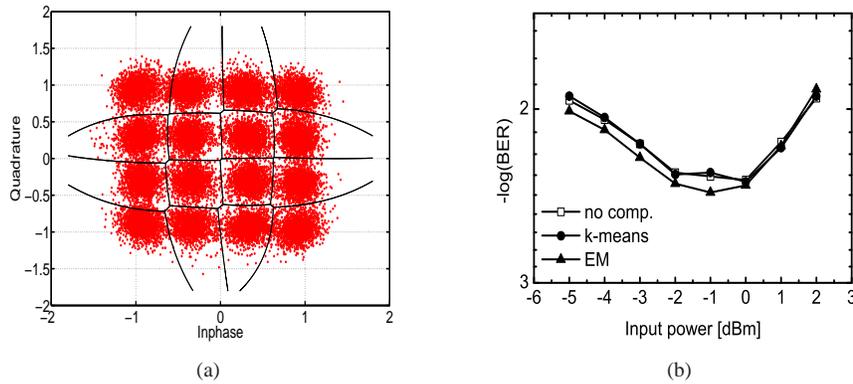


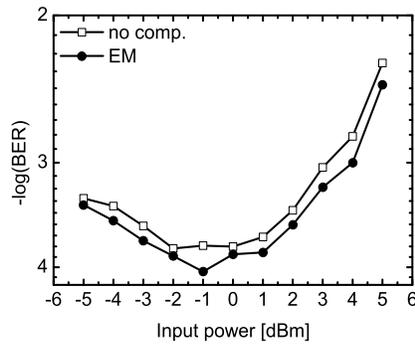
Fig. 9. (a) Constellation diagram of the demodulated signal after 800 km of transmission through dispersion unmanaged link. (b) BER as a function of span input power for dispersion unmanaged link after 800 km of transmission.

plotted together with the optimal decision boundaries for the input power of 0 dBm. It is observed in Fig. 9(a), that the clusters are not distorted in the same way as for the dispersion managed link. Indeed, the clusters seem to be more circularly symmetric and the optimal decision boundaries are very close to linear. This confirms the earlier observations reported in [16–18], that for the dispersion unmanaged links the effect of fibre nonlinearity corresponds to adding white Gaussian noise. There is however, some slight distortion in the shape of the clusters. In Fig. 9(b), $-\log(\text{BER})$ is plotted as a function of span input power. It is observed that for input power exceeding 0 dBm, there is no improvement in the nonlinear system tolerance when k-means or EM is used. However, for the input power less than 0 dBm shown in Fig. 9(b), there is an improvement of ~ 0.5 dB.

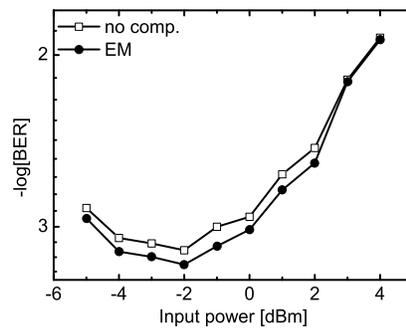
In order to investigate how nonlinearity compensation by the EM scales with transmission distance, $-\log(\text{BER})$ is plotted as a function of span input signal power for transmission distances of 240 km and 400 km, respectively, see Fig. 10. It is observed in Fig. 10(a) that an improvement in the nonlinear system tolerance of approximately 0.5 dB is observed for the entire range of the considered input optical power. For the transmission distance of 400 km, Fig. 10(b), a similar improvement of 0.5 is observed, however, the gain disappears for input power exceeding 2 dBm. One of the explanations could be that for input power exceeding 2 dBm the distortion is large and the EM cannot properly estimate the parameters.

6. Conclusion

We have shown that the expectation maximization algorithm is a powerful tool for combating system impairments, (fibre nonlinearities, I/Q imperfections and laser linewidth), which significantly distort the signal constellation. By the distortion, it is meant the deviation from circular symmetric shape. Additionally, we have demonstrated, experimentally that by using expectation maximization, up to 3 dB of improvement in nonlinear system tolerance can be obtained for 800 km long dispersion managed transmission link. For the considered dispersion unmanaged link of 240 km, 400 km and 800 km, an improvement of approximately 0.5 dB, is obtained. For a multi-channel WDM transmission system, the EM algorithm can potentially be beneficial as the inter-channel nonlinear effects will have an imprint on the constellation of the signal under constellation. The improvement in the nonlinear system tolerance offered by the EM, will depend on the modulation format of the neighboring channels, the spacing, as well as



(a)



(b)

Fig. 10. (a) BER as a function of span input power for dispersion unmanaged link after 240 km of transmission. (b) BER as a function of span input power for dispersion unmanaged link after 400 km of transmission.

the number of neighboring channels. However, this is the topic that needs further investigations.

In order to see what benefits EM brings we need to relate the performance of the EM to a digital backpropagation, which has become a benchmark for nonlinearity compensation techniques. For the dispersion managed links, it has been numerically shown that up to 4 dB of improvement can be obtained [19]. The gain in the nonlinear system tolerance will though depend on the implementation of the digital backpropagation. For instance this gain reduces to 2 dB if 1 step/fiber digital backpropagation is used [19]. For the dispersion unmanaged links, an improvement of 3 dB for PDM 16-QAM signal has been shown experimentally [20]. As a rule of thumb, we may conclude that for dispersion unmanaged system digital backpropagation offers better performance, while for dispersion managed systems, depending on the implementation of digital backpropagation, the performance of the EM and digital backpropagation may be comparable. However, one technique does not have to exclude the other, as one may consider the combination of both techniques in combating optical fibre channel nonlinearities.

Acknowledgments

This work has been partly supported by the Danish Council for Independent Research, project CORESON and the CHRON (Cognitive Heterogeneous Reconfigurable Optical Net-

work) project, with funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement n° 258644, <http://www.ict-chron.eu>.