APPLICATIONS OF EXPECTATION MAXIMIZATION ALGORITHM FOR COHERENT OPTICAL COMMUNICATION

D. Zibar, O. Winther, R. Borkowski, I. T. Monroy *

Technical University of Denmark
DTU Fotonik/Computing
Kgs. Lyngby
Denmark

L. Carvalho, J. Oliveira

CPqD
Centro de Pesquisa
Desenvol. em Telecomun.
Rod.Campinas, Brazil

ABSTRACT

In this invited paper, we present powerful statistical signal processing methods, used by machine learning community, and link them to current problems in optical communication. In particular, we will look into iterative maximum likelihood parameter estimation based on expectation maximization algorithm and its application in coherent optical communication systems for linear and nonlinear impairment mitigation. Furthermore, the estimated parameters are used to build the probabilistic model of the system for the synthetic impairment generation.

It is shown numerically and experimentally that iterative parameter estimation based on expectation maximization algorithm is a powerful tool in combating system impairments such as non-linear phase noise, inphase and quadrature (I/Q) modulator imperfections and laser linewidth. We show experimentally that for a dispersion managed polarization multiplexed 16-quadrature amplitude modulation (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.

Moreover, we show that joint estimation of carrier frequency, phase, signal means and noise covariance, can be performed iteratively by employing expectation maximization. Using experimental data we show that joint carrier synchronization and detection offers an improvement of 0.5 dB in terms of input power compared to hard decision digital phase-locked loop (PLL) based carrier synchronization and demodulation.

***Index Terms*—** optical communication, machine learning, expectation maximization, nonlinear impairments

1. INTRODUCTION

Optical communication systems are becoming increasingly complex, especially with introduction of advanced modulation formats in combination with digital signal processing based coherent detection [1]. Therefore, advanced tools are necessary in order to perform impairment mitigation, and enable optical communication system to operate close to the channel capacity. Linear signal processing algorithms from wireless communication can be effectively used to compensate for linear fibre channel impairments and have been demonstrated very successfully for higher order QAM signaling. However, one of the major challenges in optical communication is nonlinearity mitigation and characterization. For high capacity 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 nonlinearity is therefore essential.

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

In this paper, we demonstrate that by employing expectation maximization (EM) parameter estimation and subsequent maximum a posteriori (MAP) detection an increase in nonlinear system tolerance can be obtained for coherent optical communication systems. The experimental test-bed consists of polarization multiplexed 16-QAM system operating at 112 Gb/s with transmission distances of up to 800 km. The advantage of the EM algorithm is that it learns the channel from the detected signal. The EM can therefore accommodate for intra (SPM, SPM induced nonlinear phase noise) and -inter (XPM, XPM induced nonlinear phase noise, etc) channel nonlinear impairments that have an imprint on the signal constellation. We show an improvement in the nonlinear tolerance for dispersion managed, unmanaged and mixed line rate scenarios.

*Research leading to these results has received funding from the Danish Council for Independent Research, project CORESON and Villum Foundation Young Investigator program.

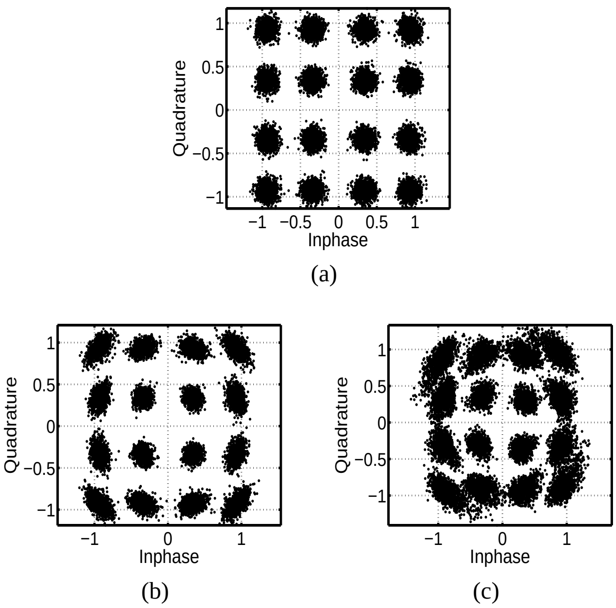


Fig. 2. Impact of different impairments on signal constellation for a 16-QAM signal. (a) Constellation of a signal dominated by additive noise. (b) Constellation of a signal dominated by phase noise. (c) Constellation of a signal dominated by nonlinear phase noise.

Additionally, we show an improvement of performing EM based joint synchronization and data detection compared to a traditional approach employing a digital phase locked loop.

2. EXPERIMENTAL SET UP

The set-up used for the experimental investigations is 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 $\sim$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. For the experimental investigations we first consider dispersion managed link and then we move to dispersion unmanaged ink. The dispersion managed link consists of 80 km of SSMF and 17 km of DCF with inline EDFA amplification. For the dispersion unmanaged link, the DCF is just bypassed. 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 2.1.

2.1. 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. The implemented clock recovery module is a feedback structure and 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. 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. The polarization demultiplexing stage is performed jointly with carrier frequency and phase estimation module. The polarization demultiplexing unit consists of a butterfly structure, and the carrier phase and frequency estimation unit is a decision-directed digital phase-locked loop. 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 demutliplexing and digital phase-locked loop are first trained in then blind mode using constant modulus algorithm and the switched to a decision directed mode as also reported in [1].

The EM algorithm is then applied after the polarization demultiplexing and carrier frequency/phase recovery stage, and used to learn the channel properties from the demodulated data without any prior knowledge. The main idea behind the EM algorithm is that the signal in x/y-polarization can be considered as a Gaussian mixture consisting of a number of components where each of the components can be described by a 2-D Gaussian distribution [4]. For the particular case of 16-QAM, we have 16 components due to 16 constellation points in 2D. The EM then evaluates in the maximum likelihood sense the most likely parameters that generated the Gaussian mixture model in terms of means, variances and mixture components. This information is then used to compensate for system nonlinear impairments by finding optimum (quadratic) decision boundaries for signal detection. In general, the EM algorithm can be used to combat any nonlinear impairments

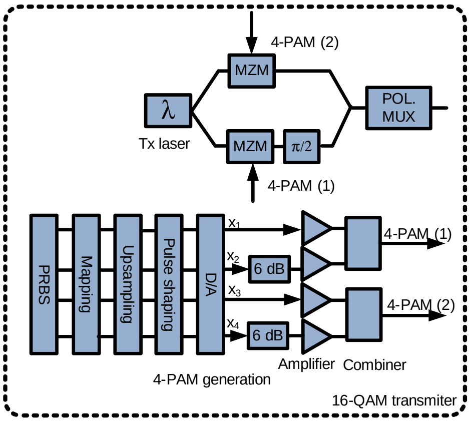


Fig. 1. Experimental setup for generation and detection of polarization multiplexed signals employing 16-QAM. PD: photodiode, PBS: polarization beam splitter, A/D: analog-to-digital converter, LO: local oscillator, EDFA: erbium doped fibre amplifier

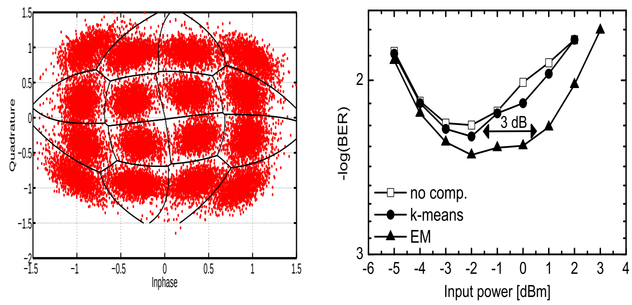
that are imprinted in a constellation. After the EM stage, error counting is performed on ∼100000 received symbols.

3. NONLINEARITY MITIGATION

In this section, it will be illustrated how different impairments impact the signal constellation and which strategy should be employed for the optimum (MAP) symbol detection. An example of the demodulated 16-QAM signal dominated by additive white Gaussian noise is shown in Fig. 2(a). It is observed in Fig. 2(a) that all the clusters look similar. For this particular case, the MAP symbol detection reduces to a simple Euclidean distance metric.

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 severally impaired by non-linear phase noise, see Fig. 2(c). It is observed that in Fig. 2(c) 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.

For the particular case of signal dominated by phase noise and nonlinear phase noise, the optimum MAP symbol detection is obtained by fully evaluating Bayes' theorem:



(a)

(b)

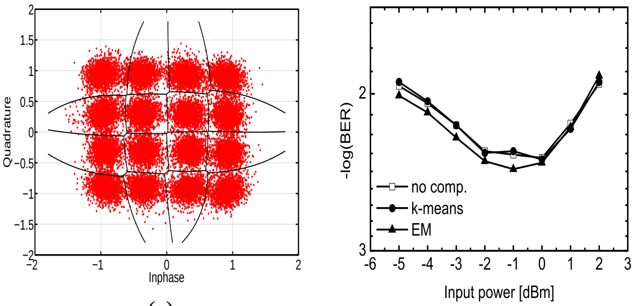
Fig. 3. (a) Constellation diagram of the demodulated signal after 800 km of transmission through dispersion managed link. (b) BER as a function of span input power for dispersion managed link after 800 km of transmission.

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

where we have assumed that the signal can be described as a mixture of Gaussian and $N(\cdot)$ represents multivariate Gaussian distribution. $\mathbf{x}$ is a received complex symbol, Σ_k is a covariance matrix and π_k is a mixing coefficient . k is an index running from 1 to M and M is the number of constellation points. Three parameters are needed to determine the distribution: μ_k, Σ_k and π_k. For finding the parameters of the distribution,μ_k and Σ_k, the EM algorithm is used [4]. Optical fiber channel impairments, intra and -inter, affect equation (1) through the determinant of the covariance matrix Σ_k.

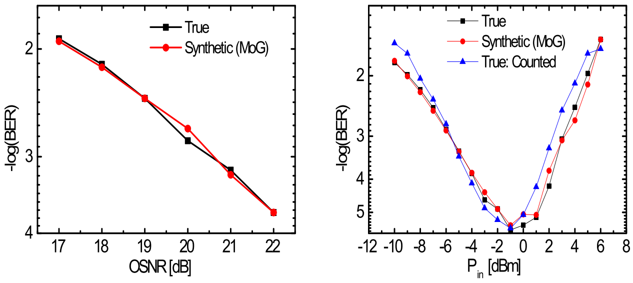
3.1. Dispersion managed link

In this section, it is demonstrated how EM can be effectively used to extract information from a severely distorted constellation and use this information to mitigate the impairments.



(a) (b)

Fig. 4. (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.



(a) (c)

Fig. 5. Comparison of the system performance: full simulation ("true") versus transmitter based impairment emulation "synthetic". (a) BER as a function of OSNR/0.1nm (b) BER as a function of input power for the dispersion unmanaged link.

In Fig. 3(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. 3(a) is distorted and therefore the optimal decision boundaries are nonlinear. In Fig. 3(b), we plot -log(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. 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, i.e. decisions based on Euclidean distances. In general, it can be said that if the clusters are circularly symmetric, as is the case for long-haul dispersion uncompensated link, the benefits of using nonlinear decision boundaries disappear for the current configuration of the EM where the memory is not taken into account.

3.2. Dispersion unmanaged link

For our next investigations, the DCF is removed and we consider 800 km of dispersion unmanaged signal transmission. In Fig. 4(a), the constellation diagram of the recovered signal is plotted together with the optimal decision boundaries for the input power of 0 dBm. It is observed in Fig. 4(a), that the cluster 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. There is however, some slight distortion in the shape of the clusters. In Fig. 4(b), -log(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 then 0 dBm shown in Fig. 4(b), there is a an improvement of ~0.5 dB. The improvement in the BER may be attributed to the imbalance associated with the I/Q modulator as it would results in the distorted constellation.

4. SYNTHETIC IMPAIRMENT GENERATION

In this section, we will explain the concept behind synthetic impairment emulation. The main idea is that we would like to synthesize optical fibre channel impairments at the transmitter. We envision that this would be very useful from a system modeling point of view, because, it would avoid running full Monte Carlo simulations. Additionally, it can be used to predict system performance. From section 3, we have seen that mixture of Gaussians (MoG) can be used to describe the signal which has nonlinear impairments. Therefore, we would like to emulate signals with linear and nonlinear impairments by drawing samples from MoG. This corresponds to performing back-to-back simulation, where the impairments are generated at the transmitter. In Fig. 5(a), -log(BER) is plotted as a function of OSNR. For the curve denoted by "true", we have run full Monte Carlo simulations including the addition of the noise. For a curve denoted "synthetic", we have sampled from MoG at the transmitter. It can be observed that there is a very good agreement between the "true" and "synthetic" curves describing -log(BER). We would like to note that -log(BER) is estimated from the Q-factor.

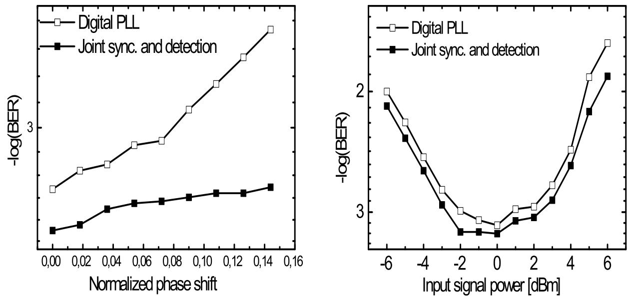
Next, we want to compare if MoG is a good model for the signal when we operate in the nonlinear regime. In Fig. 5(b), -log(BER) is plotted as a function of input power to the transmission link. In Fig. 5(b), we have additionally plotted a curve denoted "true-counted", which corresponds to the counted -log(BER) for the full simulation. Once again, it is observed in Fig. 5(b), that there is a good agreement between the transmitter based impairment emulation compared to when the full simulation is performed. It should be noted that both for "true" and "synthetic" there is a constant deviation from "true-counted" curve in the nonlinear regime. This is mainly because there is a penalty when deriving BER from the Q-factor in the nonlinear region, due to remaining correlation. This is because we are only able to equalize lienal impairment, however, the impact of the nonlinear impairments will still be present in the data.

5. CARRIER RECOVERY

The principle behind the EM algorithm for iterative parameter estimation can be expanded to perform joint carrier synchronization and data detection [5]. This type of joint parameter estimation is especially useful when the signal experiences large distortions. In Fig. 6(a), -log(BER) is plotted as a function of the normalized mean nonlinear phase shift. In order to investigate the impact of the normalized phase shift, input power to the fibre is varied. We compare joint carrier synchronization and signal detection with the PLL based synchronization and demodulation. Fig. 6(a), shows that -log(BER) is affected by the nonlinear phase shift, when digital PLL is employed. This is especially valid when the mean normalized phase shift exceeds 0.08. For the joint synchronization and signal detection -log(BER) is very little impacted by the nonlinear phase shift. Next, we investigate the system performance, in terms of -log(BER), as a function of input signal power to the transmission span for the transmission distance of 278 km. The results are shown in Fig. 6(b). It is observed that an improvement of approximately 0.5 dB is obtained when employing joint carrier synchronization and symbol estimation, compared to when digital PLL based approach is used. This is observed in both linear and nonlinear regime of operation. It should also be noted that for input signal powers above 4 dBm, an improvement of approximately 1 dB is observed. This is because under strong signal degradations, as in the case of high input signal power, the PLL does not give accurate estimates of carrier frequency and phase. One of the advantages of the proposed scheme, which remains to be shown in future work, is that it is very well suited to be integrated with soft decision forward error correction.

6. CONCLUSION

We have shown that the expectation maximization algorithm is a powerful tool for combating and modeling system impairments, (fibre nonlinearities, I/Q imperfections and laser linewidth), which significantly distort the signal constellation. Joint carrier frequency and phase, means and noise variance estimation is demonstrated for coherent optical communication system employing dual polarization 16-QAM. A relevant issue when applying EM based nonlinearity compensation and carrier recovery is the computational complexity of the scheme. In general, the computational complexity of the EM schemes scales linearly with the length of the observation interval, N, constellation order, M, and the number of iterations, P, e.g. $O(PMN)$ [5]. As the final remark, more research is needed in order to determine the impact of interchannel effects of the performance of the EM algorithm.



(a) (b)

Fig. 6. (a) BER as function of normalized nonlinear phase sift. (b) BER as a function of input power.

REFERENCES

[1] Peter J. Winzer, "High-spectral-efficiency optical modulation formats," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3824–3835, 2012.

[2] R.-J. Essiambre, G. Kramer, P.J. Winzer, G.J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *Journal of Lightwave Technology*, vol. 28, no. 1/4, pp. 662, 2010.

[3] Darko Zibar, Ole Winther, Niccolo Franceschi, Robert Borkowski, Antonio Caballero, Valeria Arlunno, Mikkel N. Schmidt, Neil Guerrero Gonzales, Bangning Mao, Yabin Ye, Knud J. Larsen, and Idelfonso Tafur Monroy, "Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged pdm 16-qam transmission," *Opt. Express*, vol. 20, no. 26, pp. B181–B196, Dec 2012.

[4] Christopher M. Bishop, *Pattern recognition and machine learning*, Springer, 2006.

[5] Darko Zibar, Luis Henrique Hecker de Carvalho, Jose Estaran, Edson Silva, Carolina Franciscangelis, Vitor Ribeiro, Robert Borkowski, Ole Winther, Mikkel N. Schmidt, Júlio Oliveira, and Idelfonso Tafur Monroy, "Joint iterative carrier synchronization and signal detection employing expectation maximization," *J. Lightwave Technol.*, vol. 32, no. 8, pp. 1608–1615, Apr 2014.