Pseudo-Newton based Equalization Algorithms for QAM Coherent Optical Systems

Mehrez Selmi, Philippe Ciblat, Yves Jaouën, Christophe Gosset

Institut Telecom/Telecom ParisTech, CNRS UMR LTCI
46 rue Barrault 75634 Paris CEDEX 13 FRANCE
mehrez.selmi@telecom-paristech.fr

Abstract: In 112Gbit/s optical coherent transmission context, we develop decision-directed and blind fractionally-spaced adaptive equalizer using variable step size suitable for 16-QAM. The considered algorithms offer better steady state performance and convergence speed than standard adaptive equalizers.

OCIS codes: (060.0060) Coherent communications; (060.4510) Optical communications

1. Introduction

Coherent detection combined with multilevel modulation such as M-ary quadrature amplitude modulation formats are one of the most promising techniques to increase the spectral efficiency and reach higher bit rates. Therefore, it is a strong candidate for the implementation of the next generation optical networks (such as, 100Gbit Ethernet). However, those modulation formats are more sensitive to signal distortions due to intersymbol interference (ISI) and the crosstalk between the polarizations [2]. In [6], fractionally spaced equalizers (FSE) based on constant modulus (CMA introduced in [1]) and radius directed (RDE initially developped by [3, 4]) algorithms are proposed to jointly compensate for the residual chromatic dispersion (CD), PMD and the polarization dependent effects (PDE). Nevertheless, due to a constant step size implementation of the gradient algorithm to adapt the equalizer components, these algorithms suffer from slow convergence. In this paper, we propose a variable step-size version of the gradient algorithms for CMA, RDE, and DD (Decision Directed) algorithms. The step-size is updated via the Pseudo-Newton method. We investigate their performance through Monte Carlo simulations. We show that our approach enables us to get faster convergence and better steady-state performance. For instance, we can compensate for a CD of 1500ps/nm with a 6-taps equalizer.

2. Gradient-descent and Pseudo-Newton based equalizers

Let $y_a(t)$ be the continuous-time signal associated with one polarization at the receiver. In order to improve the performance, sampling the received signal at twice the symbol rate is of interest. Therefore we focus on $y(n) = y_a(nT_s/2)$ where T_s is the symbol period. As the transmitted symbol sequence $x(n)$ does not have the same rate as the oversampled received signal, a linear filter can not represent the link between these both sequences. To overcome this problem, we stack two consecutive received samples into a bivariate process as follows: $\mathbf{y}(n) = [y_a(nT_s), y_a(nT_s + T_s/2)]^{\mathrm{T}}$ where the superscript $(.)^{\mathrm{T}}$ stands for the transposition. Now, if the received signal is only disturbed by linear operations, the received signal is obtained by the convolution of symbol sequence $x(n)$ with a SIMO (Single Input/Multiple Ouput) filter. Therefore, to compensate for a SIMO filter, we need to introduce a MISO (Multiple Input/Single Output) equalizer. Let $\mathbf{w}(n)$ be the n^{th} component of the MISO equalizer. $\mathbf{w}(n)$ is actually a 1×2 vector. As the MISO filter is assumed to be an FIR of length L, we have

$$z(n) = \sum_{k=0}^{L} \mathbf{w}(k)^* \mathbf{y}(n-k) = \mathbf{W}^* \mathbf{Y}_L(n) \tag{1}$$

where $(.)^*$ stands for complex conjugaison, $\mathbf{W} = [\mathbf{w}(0), \mathbf{w}(1), \cdots, \mathbf{w}(L-1)]$ and $\mathbf{Y}_L(n) = [\mathbf{y}(n)^{\mathrm{T}}, \mathbf{y}(n-1)^{\mathrm{T}}, \cdots, \mathbf{y}(n-L+1)^{\mathrm{T}}]^{\mathrm{T}}$. Notice that the rate of the sequence $z(n)$ is $1/T_s$ as that of $x(n)$. We now would like to exhibit the filter $\mathbf{W}$ enabling us to have $z(n)$ close to $x(n)$. In blind context, one can based the equalizer derivations on the CMA criterion [1] which looks for the minimization of the cost function $J_{\mathrm{CMA}}(\mathbf{W}) = \mathbb{E}[J_{\mathrm{CMA},n}(\mathbf{W})]$ with $J_{\mathrm{CMA},n}(\mathbf{W}) = (|z(n)|^2 - R)^2$ and $R = \mathbb{E}[|x(n)|^4]/\mathbb{E}[|x(n)|^2]$ or on the RDE criterion corresponding to an adaptation of CMA to QAM constellations [3, 4]. The DD approach leads to the minimization of the following cost function $J_{\mathrm{DD}}(\mathbf{W}) = \mathbb{E}[J_{\mathrm{DD},n}(\mathbf{W})]$ with $J_{\mathrm{DD},n}(\mathbf{W}) = |z(n) - \hat{x}(n)|^2$ and $\hat{x}(n)$ the current decision of symbol $x(n)$.We recall that the minimization of a cost function $\mathbf{W} \mapsto J(\mathbf{W}) = \mathbb{E}[J_n(\mathbf{W})]$ can be implemented using the gradient descent algorithm.

$$\mathbf{W}_{n+1} = \mathbf{W}_n - \mu \nabla J_n(\mathbf{W})|_{\mathbf{W}_n} \tag{2}$$

where μ is the constant step-size parameter, ∇J_n is the gradient at time n, and $\mathbf{W}_n$ is the equalizer at time n. The choice of the step-size is a crucial task for gradient algorithm and arises from a trade-off between convergence speed and steady-state performance. To overcome this problem, we propose to implement variable step-size approach by replacing μ with μ_n in Eq. (2). To derive μ_n, we consider the Pseudo-Newton algorithm [5] that exploits the Hessian matrix as follows

$$\mu_n = \mu H_n^{-1}(\mathbf{W}), \quad \text{with} \quad H_n(\mathbf{W}) = \frac{\partial^2 J_n(\mathbf{W})}{\partial \mathbf{W}^{\mathrm{H}} \partial \mathbf{W}} \tag{3}$$

where the $(.)^{\mathrm{H}}$ denotes the complex-transpose operator. In order to reduce the computational load, the inverse Hessian matrix can be updated as follows

$$H_n^{-1}(\mathbf{W}) = \lambda^{-1} H_{n-1}^{-1}(\mathbf{W}) - \frac{\lambda^{-2} H_{n-1}^{-1}(\mathbf{W}) \mathbf{y}(n) \mathbf{y}^{\mathrm{H}}(n) H_{n-1}^{-1}(\mathbf{W})}{[(1-\lambda) f(n)]^{-1} + \lambda^{-1} \mathbf{y}^{\mathrm{H}}(n) H_{n-1}^{-1}(\mathbf{W}) \mathbf{y}(n)} \tag{4}$$

where λ is a forgetting factor ($0 \leq \lambda \leq 1$ and assuming $\lambda + \mu = 1$ [5]) and $H_0^{-1}(\mathbf{W}) = \delta \mathbf{Id}$ with the identity matrix $\mathbf{Id}$ and a fixed positive number δ. In Table 1, we summarize the value of the gradient and $f(n)$ for CMA and DD algorithms. The computational load of the gradient (resp. Pseudo-Newton) approach is $\mathcal{O}(L)$ (resp. $\mathcal{O}(L^2)$). In the case

	CMA	**DD**
$\nabla J_n(\mathbf{W})$	$(\lvert z(n)\rvert^2 - R) z(n)^* \mathbf{y}(n)$	$(z(n) - \hat{x}(n))^* \mathbf{y}(n)$
$f(n)$	$2\lvert z(n)\rvert^2$	1

Table 1. Gradient and $f(n)$ for CMA and DD algorithms

of PolMux channels, we have to consider a MIMO equalizer (to compensate simultaneously for the polarization cross-talk, CD and PMD) since there are two outputs: one per polarization. Then we have $z_1(n) = \mathbf{W}_{11}^* \mathbf{y}_1(n) + \mathbf{W}_{12}^* \mathbf{y}_2(n)$ and $z_2(n) = \mathbf{W}_{21}^* \mathbf{y}_1(n) + \mathbf{W}_{22}^* \mathbf{y}_2(n)$ where index 1 (resp. index 2) corresponds to polarization X (resp. Y). Update equations for filters $\mathbf{W}_{ij}$ are the same as these described above.

3. Simulations and results

A 112Gbit/s transmission is achieved by multiplexing both polarizations with 16-QAM modulated signals which corresponds to 14Gbaud transmission per polarization. In order to reduce the wide spectrum of the QAM signal, we used a square root raised cosine filter with a roll-off factor equal to 1, each at the transmitter and the receiver side. The received electrical signal is digitized using a 6bit resolution analog-to-digital converters at a rate of 2 samples per symbol. A 5^{th} order Bessel electrical filter with a -3dB bandwidth 80% of the Baud rate was included as the anti-aliasing filter [4]. To evaluate the speed of convergence of the equalizers, we considered first a single polarization transmission (56Gbaud), Fig.1 shows that Pseudo-Newton (PN) based algorithms converge faster and ensure lower BER at the steady state. The OSNR penalties versus CD is depicted in Fig.2, the system can tolerate up to 1500ps/nm

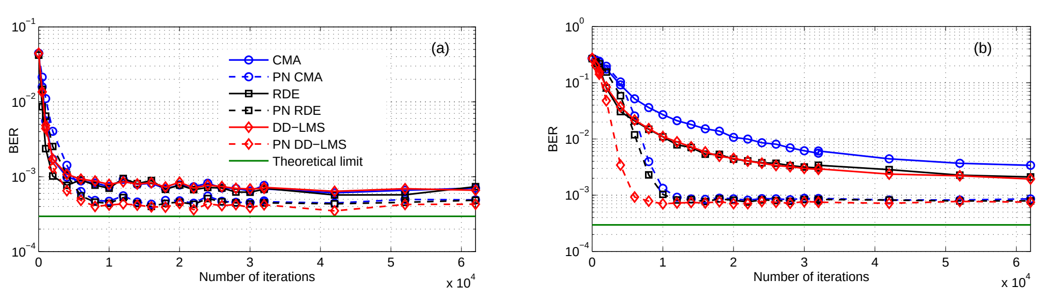

Fig. 1. convergence of the equalizers for OSNR=18dB, 6 taps FSE, $\mu = 10^{-3}$, $\delta = 0.9$: (a) CD=500ps/nm, (b) CD=1500ps/nm.

of CD using 6 taps (and up to 2250ps/nm using 8 taps).

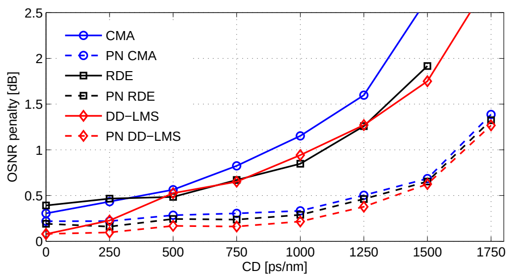

Fig. 2. BER vs. CD in PolMux 16-QAM, 6-taps RDE and CMA equalizers, ($\theta = \frac{\pi}{4}$, DGD=50ps and $\Delta\nu$=1.4MHz)

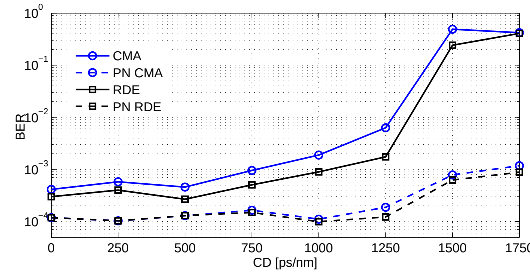

Fig. 3. BER vs. CD in PolMux 16-QAM, 6-taps RDE and CMA equalizers, ($\theta = \frac{\pi}{4}$, DGD=50ps and $\Delta\nu$=1.4MHz)

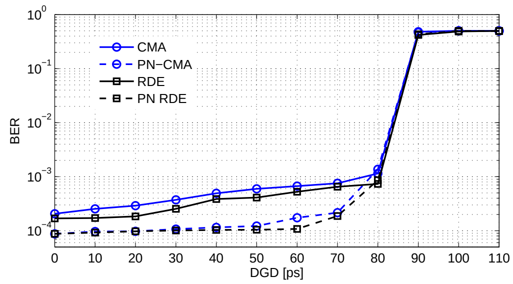

Fig. 4. BER versus DGD in PolMux 16-QAM, 6-taps RDE and CMA equalizers, ($\theta = \frac{\pi}{4}$, CD=500ps/nm and $\Delta\nu$=1.4MHz)

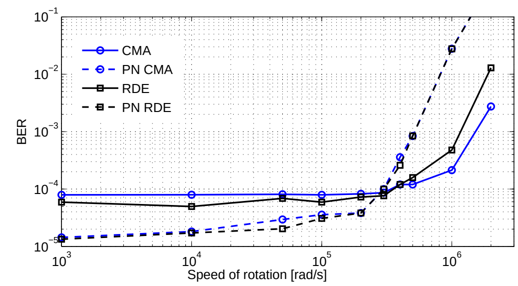

Fig. 5. Dynamic response of the 6-taps equalizers, OSNR=20dB

The performance of the overall system (PolMux) is then investigated. Phase noise (laser linewidth $\Delta\nu$=1.4MHz), CD, DGD and polarization rotation(angle θ) were included and the OSNR was set to 22dB. After equalization, the carrier phase recovery algorithm[4] is applied. As shown in Fig.3, the BER is kept below 10^{-3} up to 1500ps/nm. The constellations of the 16QAM signal in Fig.6 show that the PN based algorithms ensure better steady state. However, PN algorithms do not compensate for bigger values of DGD, all the algorithms were shown to tolerate up to 80ps of DGD (Fig.4). Finally, to evaluate the dynamic behavior of the equalizers, we considered an infinite rotation of the polarization using Jones matrix[6], no phase noise was considered and the OSNR was set to 20dB, gradient based algorithms showed better performance for rotations higher than 200Krad/s.

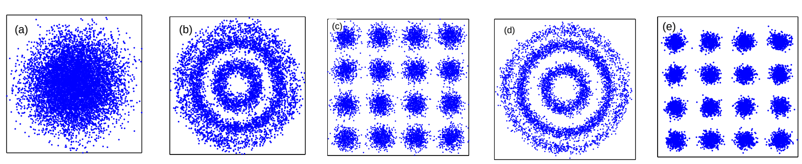

Fig. 6. PolMux 16-QAM constellations ($\theta = \frac{\pi}{4}$,CD=1250ps/nm, DGD=50ps,$\Delta\nu$=1.4MHz) (a)Input of the equalizers (b)-(c) equalized and recovered (RDE)(d-e) equalized and recovered (PN-RDE)

4. Conclusion

ISI suppression and source separation in PolMux 16-QAM coherent optical transmission system using gradient descent and Pseudo-Newton algorithms are compared. We showed that Pseudo-Newton algorithms can compensate up to 1500ps/nm using a 6 taps $T_s/2$ FSE equalizer, and offer better performances that the gradient descent ones when tracking polarization fluctuations slower than 200Krad/s. thus is may remain a candidate for the implementation of 16-QAM coherent systems.

References

1. D.N. Godard, "Self Recovering Equalization and carrier tracking in two dimensional data communication systems," IEEE TCOM, vol. 28, no.11, pp. 1867-1875, 1980.

2. M. Seimetz, "Performance of coherent optical square 16-QAM systems based on IQ-transmitters and homodyne receivers with digital phase estimation," OFC/NFOEC, 2006.
3. H. Louchet K. Kuzmin and A. Richter, "Improved DSP algorithms for coherent 16-QAM transmission," ECOC, 2008.
4. I. Fatadin, D. Ives, and S.J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM coherent optical system," JLT, vol. 27, no. 15 , pp. 3042-3049, 2009.
5. W. J. Zeng, "Adaptive Newton Algorithms for blind equalization using the generalized constant modulus criteria," ICASSP, 2009.
6. S.J. Savory, "Digital filters for coherent optical receivers," Optics Express, vol. 16, no. 2, pp. 804-817, 2008.