# Block versus Adaptive MIMO Equalization for Coherent PolMux QAM Transmission Systems

Mehrez Selmi[(1)], Philippe Ciblat[(1)], Yves Jaouën[(1)], Christophe Gosset[(1)]

[(1)] Institut Telecom/Telecom ParisTech, CNRS UMR 5141,46 rue Barrault 75634 Paris, France
✉ mehrez.selmi@telecom-paristech.fr

***Abstract*** *We propose a block equalization algorithm using optimal step size. The algorithm shows fast convergence, low steady-state error and good tracking capacity in comparison to standard equalizers operating on a sample-by-sample basis.*

## Introduction

Phase modulated formats, such as QPSK, combined with polarization multiplexing (PolMux) have shown the capability to reach 100 Gb/s or above using coherent detection associated with digital signal processing (DSP). This DSP deals at least with the compensation of the Inter-Symbol Interference (ISI) genererated by polarization mode dispersion (PMD) and residual chromatic dispersion (CD). M-ary quadrature amplitude modulation (M-QAM) formats appear as promising candidates to reach 400 Gb/s, with the current state of the art of ADC and FPGA circuits[1]. In Fig. 1, we describe the DSP structure of a PolMux coherent receiver for which our goal is to improve the block in gray color corresponding to MIMO equalization, i.e., PMD/residual CD mitigation. Usually the ISI is mitigated by means

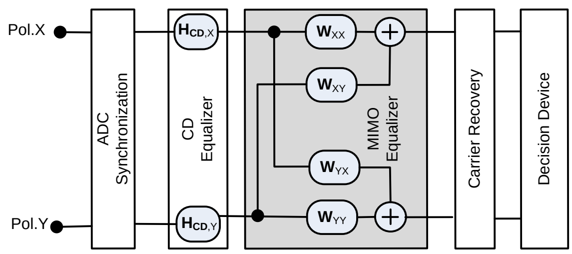


**Fig. 1**: DSP structure for PolMux coherent receiver.

of an adaptive equalizer based on the stochastic gradient descent algorithm for sake of simplicity[2,3] using either the constant modulus (CM), the multi-modulus (MM) or the decision-directed (DD) criterion. Moreover, in the state-of-the-art, the stochastic gradient descent algorithm is carried out with a constant step size during all the tracking sequence. The choice of this step size is connected to the trade-off between the convergence speed and the steady state performance. In order to overcome the problem of suboptimal step size, we proposed the stochastic Pseudo-Newton algorithm based on the computation of the inverse of the Hessian matrix[4]. The main advantage of an adaptive approach is its ability to track propagation channel variation. However in optical communications, the propagation channel may vary quite slowly compared to the symbol period[5,6], i.e., the channel can be assumed to be constant over a large observation window. Therefore it is worthy treating the data block-by-block and not sample-by-sample.

The proposed algorithm is a block one (i.e., operating block-by-block) instead of being adaptive (i.e., operating sample-by-sample). The main advantage of this new approach is to improve the statistics estimation and so the algorithm behavior. In the sequel of this paper, we introduce our blockwise algorithm which relies on the (non-stochastic) gradient-descent algorithm associated with the CM criterion and an optimal step size implementation. Derivations and simulations are done using 16-QAM modulation format.

## Proposed algorithm

Let $x(n)$ be the transmitted sequence of QAM modulated signal and $y_a(t)$ the continuous-time signal associated with one polarization at the receiver. As it is usual to oversample the received signal at twice the baud rate, we focus on $y(n) = y_a(nT_s/2)$ where $T_s$ is the symbol period. In order to "work" at the symbol rate, we stack two consecutive received samples into a bivariate process as follows:

$$\mathbf{y}(n) = [y_a(nT_s), y_a(nT_s + T_s/2)]^{\mathrm{T}} \qquad (1)$$

where $(.)^{\mathrm{T}}$ stands for the transposition. If the received signal is only disturbed by linear operations (such as CD and $1^{st}$ order PMD), $\mathbf{y}(n)$ is filtered version of $x(n)$ that takes the following form

$$\mathbf{y}(n) = \sum_{k=0}^{K} \mathbf{h}(k)x(n-k) + \mathbf{b}(n) \qquad (2)$$

where $\mathbf{h}(k)$ is the $k$-th component of the filter, and $\mathbf{b}(n)$ is the additive noise indenpendent of the data. In order to compensate for the channel response, we introduce a $T_s/2$-fractionally spaced equalizer (FSE). Let $\mathbf{w}(n)$ be the $n$-th component of the FSE assumed to be a FIR of length $L$, we have

$$z(n) = \sum_{k=0}^{L} \mathbf{w}(k)^* \mathbf{y}(n-k) = \mathbf{W}^{\mathrm{H}} \mathbf{Y}_L(n) \quad (3)$$

where $(.)^{\mathrm{H}}$ stands for conjugate transposition, $\mathbf{W} = [\mathbf{w}(0), \mathbf{w}(1), \cdots, \mathbf{w}(L-1)]^{\mathrm{T}}$ and $\mathbf{Y}_L(n) = [\mathbf{y}(n)^{\mathrm{T}}, \mathbf{y}(n-1)^{\mathrm{T}}, \cdots, \mathbf{y}(n-L+1)^{\mathrm{T}}]^{\mathrm{T}}$. We now would like to exhibit the filter $\mathbf{W}$ enabling us to have $z(n)$ close to $x(n)$. To do that, it is relevant to use the CM criterion defined as the minimization of the following cost function[7]

$$J(\mathbf{W}) = \mathbb{E}[J_n(\mathbf{W})] \quad (4)$$

with $J_n(\mathbf{W}) = (|z(n)|^2 - R)^2$ and $R = \mathbb{E}[|x(n)|^4]/\mathbb{E}[|x(n)|^2]$. In order to estimate the mathematical expectation, we propose to minimize the following estimated cost function

$$\hat{J}_N(\mathbf{W}) = \frac{1}{N} \sum_{n=0}^{N-1} J_n(\mathbf{W}) \quad (5)$$

where $N$ is the number of available bivariate samples $\mathbf{y}(n)$. Our purpose is now to find the minimum of $\mathbf{W} \mapsto \hat{J}_N(\mathbf{W})$. To do that, we suggest to use the (non-stochastic) gradient descent algorithm with optimal step size. If $\mathbf{W}_i$ is the estimated equalizer at the $i$-th iteration (note that the data block is the same for each iteration), we have the following update relation[10,11]

$$\mathbf{W}_{i+1} = \mathbf{W}_i - \mu_i \mathbf{\Delta}_i \quad (6)$$

where $\mathbf{\Delta}_i = \partial \hat{J}_N(\mathbf{W}) / \partial \mathbf{W}^*{}_{|\mathbf{W}_i}$ writing as

$$\mathbf{\Delta}_i = \frac{1}{N} \sum_{n=0}^{N-1} (|z(n)|^2 - R) z(n)^* \mathbf{Y}_L(n) \quad (7)$$

and where $(.)^*$ stands for complex conjugation.
In order to find the optimal step size $\mu_i$ at the $i$-th iteration, we proceed into the minimization the estimate cost function with respect to $\mu_i$, i.e.,

$$\mu_i = \arg\min_{\mu} \hat{J}_N(\mathbf{W}_i - \mu \mathbf{\Delta}_i). \quad (8)$$

The derivative of $\mu \mapsto \hat{J}_N(\mathbf{W}_i - \mu \mathbf{\Delta}_i)$ is the following 3-rd degree polynomial.

$$P_i(\mu) = p_{3,i}\mu^3 + p_{2,i}\mu^2 + p_{1,i}\mu + p_{0,i} \quad (9)$$

where

$$p_{3,i} = \frac{1}{N} \sum_{n=0}^{N-1} a_n^2, \quad p_{2,i} = \frac{1}{N} \sum_{n=0}^{N-1} a_n b_n,$$

$$p_{1,i} = \frac{1}{N} \sum_{n=0}^{N-1} (2a_n c_n + b_n^2), \quad p_{0,i} = \frac{1}{N} \sum_{n=0}^{N-1} b_n c_n$$

with $a_n = |z(n)|^2$, $b_n = -2\Re(z(n)\delta_{n,i}^*)$, $c_n = (|z(n)|^2 - R)$ and $\delta_{n,i} = \mathbf{\Delta}_i^{\mathrm{H}} \mathbf{Y}_L(n)$. To minimize Eq. (8), we derive in closed-form the roots of polynomial $P_i(.)$ and we select the real-valued root prodiving the minimum value of $\mu \mapsto \hat{J}_N(\mathbf{W}_i - \mu \mathbf{\Delta}_i)$. In Fig. 2, we summarize the architecture of the proposed equalizer.

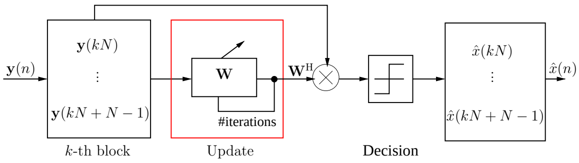


**Fig. 2**: Structure of the proposed blockwise equalizer.

Note that similar approach can be done with DD criterion. In such a case, we have

$$\mathbf{\Delta}_{i,\mathrm{DD}} = \frac{1}{N} \sum_{n=0}^{N-1} (z(n) - \hat{x}(n))^* \mathbf{Y}_L(n) \quad (10)$$

where $\hat{x}(n)$ is the current decision on the symbol $x(n)$, and

$$\mu_{i,\mathrm{DD}} = \frac{\sum_{n=0}^{N-1} \Re\{\delta_{n,i,\mathrm{DD}}^*\}(z(n) - \hat{x}(n))}{\sum_{n=0}^{N-1} 2|\delta_{n,i,\mathrm{DD}}^*|^2} \quad (11)$$

where $\delta_{n,i,\mathrm{DD}} = \mathbf{\Delta}_{i,\mathrm{DD}}^{\mathrm{H}} \mathbf{Y}_L(n)$.
Our algorithm can be adapted in order to treat the two polarizations jointly. Moreover our approach can be straightforwardly applied for the MM algorithm (MMA)[3]. We limit our study to the CM criterion in this paper.

## Simulations results

A 112Gbit/s transmission was achieved by multiplexing both polarizations with 16-QAM modulated signals which corresponds to 14Gbaud transmission per polarization. The transmit shaping filter is a square root raised cosine filter with a roll-off factor equal to 1. A matched filter to the shaping filter is applied at the receiver side. The received electrical signal is sampled at a rate of 2 samples per symbol. A 5-th order Bessel filter with a 3dB bandwidth equal to 80% of the symbol rate was used as anti-aliasing filter. The performance of the algorithm is evaluated using 1000 Monte Carlo trials.
Except otherwise stated, we considered a transmission line that exhibits a 1000ps/nm of CD, a $\pi/4$ polarisation rotation and a 50ps DGD[8]. The OSNR was set to 20dB and no phase noise was considered. The equalizer length $L$ is fixed to 6. This equalizer $\mathbf{W}$ is estimated using the proposed method and applied to both polarizations.
The speed of convergence of the algorithm for different block sizes is depicted in Fig. 3. We observe that the convergence is obtained after

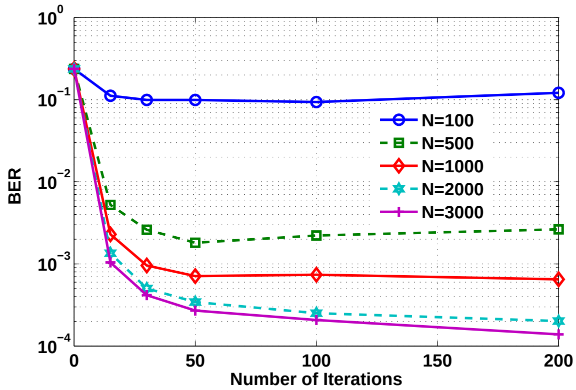


**Fig. 3**: BER vs. #iterations for different block sizes $N$. (OSNR=20dB, DGD=50ps, polarization rotation=$\frac{\pi}{4}$)

nearly 50 iterations. Secondly, we can also remark that lower BER (below the FEC limit $\sim 2.10^{-3}$) is reached as soon as the block size is larger than 1000.

We now would like to compare our algorithm with the standard CM algorithm (CMA) and the Pseudo Newton (PN) CMA[4] operating adaptively, i.e., sample-by-sample. In Fig. 4, we analyse the BER for different sizes of observation window (the observation window per polarization is denoted by $N$). We remind that the proposed

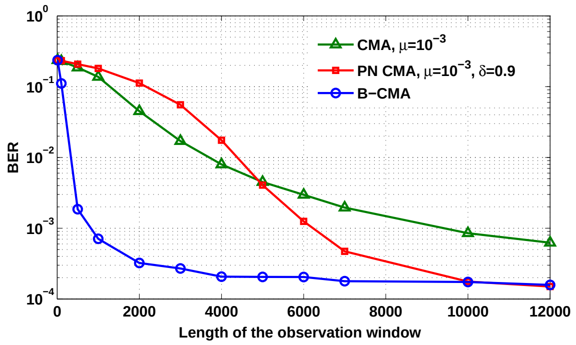


**Fig. 4**: BER vs. the length of the observation window. (#iterations= 50 for the Block CMA, OSNR=20dB, DGD=50ps, polarization rotation= $\frac{\pi}{4}$)

algorithm (called Block (B) CMA) evaluates the cost function $\hat{J}_N(.)$ along the entire observation window (which is then associated with one block) and then iterates to find the minimum of this cost function (50 iterations have been considered here to obtain Fig. 4) whereas the adaptive based algorithm iterates at each sample $n$ with a new cost function $J_n(.)$ until the end of the observation window. The B-CMA improves significantly the convergence speed with regard to the observation window size. The steady-states however are similar for all the approaches.

As noticed by some authors[5,6], the propagation channel has a coherence time around a few milliseconds. In addition, experimental measurements[9] have shown that the maximum state of polarization (SOP) variation is about 26rad/s which corresponds a coherence time equal to 3.8ms (if the coherence time is defined as the maximum delay for which the normalized mean square error between both channel impulse response is less than 1%). As we consider a 14Gbaud transmission per polarization, the symbol period is equal to 71ps and the block size with $N$=1000 has a duration of 71ns. As a consequence, we have checked that the channel is stationary or slowly-varying throughout one block.

Let us now comment on the tracking ability of the proposed blockwise algorithm. By observing Fig. 4, we remarked that the blockwise algorithm is able to estimate the channel with a few thousands of samples. As a consequence, if the channel is dramatically modified inside one block, the blockwise algorithm will be able to find fastly the new channel value with the next block while the adaptive equalizer will find this new channel value after at least a few tens of thousands iterations and thus samples. Surprisingly the blockwise approach is more adapted to channel variation than the adaptive one. Actually, the adaptive one may be interesting when the channel is varying oftenly (i.e., inside each block for example) but the measurments done in installed fibers in[5,9] show that it is not the case in optical communications.

## Conclusion

We proposed a block blind equalizer using an optimal step size. Simulations showed faster convergence than CMA and PN-CMA algorithms, and excellent tracking capabilities. Avoiding operating on a sample-by-sample basis relaxes the constraints on real-time implementation, therefore our proposal is an attractive solution for QAM coherent systems using polarization multiplexing.

## Acknowledgement

This work has been funded by the project “Réseaux du futur“ located at Institut Telecom and supported by the European Network of Excellence EURO-FOS.

## References

1 A.H. Gnauck et al., OFC, PDPB8 (2009).
2 S.J. Savory, Opt. Express, **16** (2008).
3 H. Louchet et al., ECOC, Tu.1.E.6 (2008).
4 M. Selmi et al., OFC, OThM3 (2010).
5 H. Bulow et al. OFC, OWE4 (1999).
6 E. Ip et al. Opt. Express, **16** (2008).
7 D.N. Godard, IEEE TCOM, **28** (1980).
8 I. Fatadin et al., JLT, **19** (2009).
9 S. Salaun et al., ECOC, P1.04 (2009).
10 L. Mazet et al., SPAWC, (1997).
11 V. Zarzoso et al., IEEE TCOM, **56** (2008).