Adaptive Receiver Structures for Fiber Communication Systems Employing Polarization-Division Multiplexing

Turgut Oktem, Alper T. Erdogan, *Member, IEEE*, and Alper Demir, *Member, IEEE*

***Abstract*—Polarization-division multiplexing (PDM) is emerging as a promising technique for increasing data rates without increasing symbol rates. However, the distortion effects of the fiber transmission medium poses severe barriers for the implementation of this technological alternative. Especially, due to the fiber-induced polarization fluctuation orthogonally transmitted PDM signals are mixed at the receiver input. Therefore, a receiver compensation structure needs to be implemented to recover the original orthogonal transmitted components from their mixtures at the end of the fiber channel. This is in fact the focus of this article where a receiver algorithm is based on a recently proposed blind source separation scheme exploiting magnitude boundedness of digital communication signals. Through the use of this scheme, new receiver algorithms for recovering the original polarization signals in an adaptive manner are proposed. The key feature of these algorithms is that they can achieve high separation performance while maintaining the algorithmic complexity in a fairly low level that is suitable for implementation in optical fiber communication receivers. The performance of these algorithms are illustrated through some simulation examples.**

***Index Terms*—Blind signal processing, PDM, PMD, fiber optic.**

I. Introduction

CENTRAL to the research efforts in optical fiber communications is to increase data rates to meet the growing demand for high bandwidth applications such as video on demand, teleconferencing, etc. There are different technologic alternatives to be utilized for this purpose. The development of corresponding methods along with the investigation of their efficiency and the practicality is an area of active research.

The most direct approach is to increase the data rate by increasing the symbol rate, and therefore, the physical transmission bandwidth. However, the increasing frequency-selective behavior of the fiber as a function of bandwidth, which is mainly due to the polarization mode dispersion [1]–[5] (PMD), is a major obstacle. There are different receiver compensation alternatives, optical [6]–[11] or electrical [12]–[20], to counteract the intersymbol interference (ISI) caused by the frequency selective behavior. However, a major limiting factor in optical fiber communication systems is the limited availability of computational resources per symbol, which clearly decreases with the growing symbol rate.

Another approach is to squeeze more information bits into given bandwidth by using more complex modulation schemes than ON–OFF signalling, such as the use of multilevel constellations potentially in connection with some channel precoding schemes (e.g., multitone scheme) [21], [22]. Pursuing these alternatives would have clear implications on both transmitter and receiver complexities. The constraint on the real-time implementation of corresponding DSP blocks would be a clear bottleneck.

The utilization of both orthogonal polarizations at the same wavelength as multiplexed transmission paths in a fiber, i.e., polarization-division multiplexing (PDM) [23], is a promising method for increasing the information capacity of a fiber. PDM would simply enable a factor of two increase in bandwidth efficiency subject to the availability of intelligent coding/multiplexing schemes and efficient receiver algorithms. Despite its promising features, there are practical problems in the implementation of the PDM method. The main issue is the mixing of the transmitted symbols in both space (polarization) and time dimensions due to the random fiber-induced polarization fluctuations and PMD. For relatively low symbol rates, which is the subject of this paper, the mixing in time, i.e., ISI due to PMD, can be ignored. However, the mixing of the two polarization sequences due to the random fluctuations of the polarization states in a fiber would still be an important issue. Therefore, intelligent receiver methods are needed to separate the original polarization signals from their mixtures. Due to the nonstationary behavior of the fiber, the separation needs to be done in an adaptive manner.

Due to the promise of doubling bandwidth efficiency, PDM has received considerable attention in the literature. PDM together with WDM was employed to enable terabit/s transmission in a single strand of fiber in 1996 [23]. PDM with solitons was used to double fiber information capacity in ultra-long-haul transmission in [24]. In both [23], [24], an RF pilot tone or a dither signal is added to one of the polarizations and the two polarization channels are demultiplexed using an optical polarization beam splitter and an active optical polarization controller that is electronically controlled based on monitoring the RF

Manuscript received December 25, 2008; revised May 22, 2009. First published July 17, 2009; current version published October 12, 2009. This work was supported in part by the Turkish Academy of Sciences GEBIP program and in part by two TUBITAK (Scientific and Technological Research Council of Turkey) Career Awards (104E073 and 104E057).

T. Oktem is with the Mobile Communications Department of EURECOM, Valbonne 06560, France (e-mail: Turgut-Mustafa.Oktem@eurecom.fr, oktem@eurecom.fr).

A. T. Erdogan and A. Demir are with the Electrical and Electronics Engineering Department, Koc University, Istanbul 34450, Turkey (e-mail: alperdogan@ku.edu.tr; aldemir@ku.edu.tr).

Digital Object Identifier 10.1109/JLT.2009.2027912

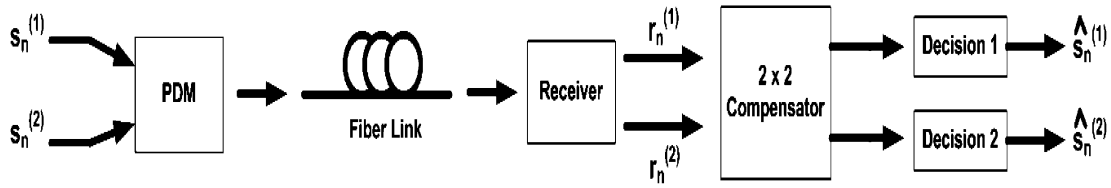


Fig. 1. Overall fiber link and compensator for PDM transmission.

pilot tone/dither signal in the receiver. In [25], slightly different amounts of optical power is launched into the two polarization channels, and the detected power difference between the two channels is used as the feedback signal to control the dynamic optical polarization controller. With this technique based on channel power imbalance, one does not need to modify the existing receiver electronics. The control electronics needed does not need to work at the symbol rate and can be much slower. In [26], a control signal for the dynamic optical polarization controller is obtained based on computing auto- or cross-correlations in the receiver that quantify the amount of crosstalk between the two polarization components. We note here that, in the PDM technique, two independent data sets are transmitted on each of the two orthogonal polarizations using ON–OFF keying or other modulation formats such as (differential) phase-shift keying [(D)PSK]. There is also a class of technique, which use polarization as another dimension in forming a signal constellation, such as the one described in [27] that employs a quaternary signalling scheme based on the total power transmitted in the two polarizations, and the one in [28] that uses joint phase-polarization-shift keying in a direct-detection quaternary scheme. The advantage of both of these techniques is that a dynamically controlled optical polarization controller is not needed and a quaternary receiver with a simple structure can be used. A simplified receiver structure in this case comes at the cost of using a suboptimal (in terms of BER for given power) signal constellation scheme. However, if PDM is used along with coherent receiver structures [29], [30], one can also avoid using dynamically controlled optical polarization controllers and implement adaptive algorithms purely in electronics in order to compensate for the fiber-induced polarization fluctuations. Moreover, with coherent detection, one can also use optimal, multibit signaling schemes such as (D)QPSK in order to further increase bandwidth efficiency and information capacity of the fiber. In devising adaptive algorithms in order to compensate for the polarization fluctuations, the use of training sequences is impractical and bandwidth inefficient in most cases. In this paper, we consider the use of *unsupervised*, or *blind*, adaptive algorithms that do not require any kind of pilot/dither signals and that can be implemented purely in electronics in order to separate the two polarization components and compensate for the fiber-induced polarization fluctuations.

The separation of original sources from their mixtures is an active research problem under blind source separation (BSS) and independent component analysis fields. Applicability of a given BSS algorithm to the PDM separation problem depends on its performance, robustness, data efficiency, and computational complexity. In fact, the most challenging issue is to develop an algorithm, which has sufficiently low complexity to work in high symbol rates of fiber links while achieving a desired level of separation performance. For this reason, we propose the use of the BSS approach for magnitude-bounded (MB-BSS) signals proposed in [31] and its extensions for fiber communication applications. MB-BSS approach provides a suitable framework for the development of low-complexity blind structures, and in this article, we demonstrate that PDM receivers designed based on MB-BSS have the desired low-complexity and high-performance separation features.

The organization of the article is as follows: In Section II, we introduce the low data rate channel model and the simulation model we use in this article. Section III poses the polarization demixing problem as a BSS problem. In Section IV, the MB-BSS algorithm is introduced as an unsupervised adaptive solution for the polarization mixing problem. We provide extensions to this algorithm to reduce implementation complexity and describe simplified hardware for compensator system design. Furthermore, in Section V, we provide simulation results to illustrate the performance of the resulting receiver structures. Finally, Section VI gives the conclusion.

II. Low Symbol Rate Channel Model

In the PDM scheme, two independent signals are transmitted over two orthogonal polarizations. Fig. 1 shows the overall block diagram corresponding to PDM-based fiber link. The PDM multiplexed signals at the transmitter travel along the fiber and reach the receiver. These signals are demodulated by a coherent receiver structure (see for example [30]) in order to extract in-phase and quadrature-phase components of two orthogonal polarizations of the received signal. However, due to the fiber-induced polarization fluctuation and the PMD of the fiber channel, these signals are distorted and mixed versions of the original multiplexed transmit signals. For this reason, a fiber optic link using PDM transmission scheme can be modeled as a 2×2 multiple-input multiple-output (MIMO) channel, where we have two input signals, orthogonal polarization signals sent by the transmitter, and two output signals, the mixtures of the original transmitted signals at the receiver. As a result, the compensator for this type of a channel must also be a 2×2 system that tries to extract two independent original polarization signals from their mixtures. Finally, the compensator outputs are sent to two separate decision devices.

We build a simulation model for the channel to enable the numerical experimentation of the overall fiber communication link in a computer environment. The simulation model consists of three parts. First part is the transmit filter used for pulse shaping. Second part is the fiber channel with PMD, and third part is the receive filter matched to the transmit filter. First, we obtain frequency-domain transfer functions of all components

and combine them to obtain the overall transfer function. Then, a discrete-time channel impulse response is extracted by fitting a finite-impulse response (FIR) filter to the overall frequency spectrum with a least squares approximation. As we use PDM, all of these parts are 2 by 2 systems.

At the transmitter, square-root raised cosine pulse filters are employed as transmit filters. At the receiver, matched filters to maximize SNR are used. The aim in using such transmit–receive filter pairs is to achieve bandwidth limitation of the signal and send the signal in a suitable form to resist ISI. What remains is the mathematical modeling of the PMD of the fiber channel. PMD is mathematically modeled as a concatenation of polarization-maintaining fibers with varying group delays and rotations of the principal axes [5], [32]–[34]. Therefore, fiber frequency response $\mathbf{U}(\boldsymbol{\omega})$ is given by

$$\mathbf{U}(\boldsymbol{\omega}) = \prod_{i=1}^{M} \mathbf{D_i}(\boldsymbol{\omega}) S_i \tag{1}$$

where

$$\mathbf{D_i}(\boldsymbol{\omega}) = \begin{pmatrix} e^{j\omega\frac{\Delta\tau_i}{2}} & 0 \\ 0 & e^{-j\omega\frac{\Delta\tau_i}{2}} \end{pmatrix}$$

and

$$\mathbf{S_i} = \begin{pmatrix} \cos(\alpha_i)e^{j\phi_i} & \sin(\alpha_i)e^{j\phi_i} \\ -\sin(\alpha_i) & \cos(\alpha_i) \end{pmatrix}.$$

. Where $\Delta\tau_i$ represents the group delay induced by the ith section (each section has equal length). $\mathbf{S_i}$ is the scattering matrix and gives a frequency-independent coordinate transformation of the principal axes and α_i and ϕ_i, respectively, denote the random polarization and phase angles and are randomly generated with uniform distributions over $(0, 2\pi)$.

For the scattering matrix $\mathbf{S_i}$, there are more alternatives. Some authors use slightly different scattering matrices [35]. Total product of the matrices in (1) results in a frequency-dependent unitary matrix $\mathbf{U}(\boldsymbol{\omega})$ where ω represents the angular frequency deviation from the carrier angular frequency ω_0. Therefore, we have a baseband equivalent transfer function. Based on the scattering matrices used, $\mathbf{U}(\boldsymbol{\omega})$ may be a general unitary matrix or a structured unitary matrix in the following form [4], [36]:

$$\mathbf{U}(\boldsymbol{\omega}) = \begin{pmatrix} U_1(\omega) & U_2(\omega) \\ -U_2^*(\omega) & U_1^*(\omega) \end{pmatrix} \quad \omega \in [\mathbf{0}, \mathbf{2}\pi) \tag{2}$$

satisfying the unitaryness constraint ($|U_1(\omega)|^2 + |U_2(\omega)|^2 = 1$ for all $\omega \in [0, 2\pi)$).

As a result, overall channel PMD transfer function is a paraunitary matrix (i.e.,unitary at each frequency). Polarization-dependent loss may distort the unitary structure; however, it is assumed to be negligible. For all the simulations presented in this article, we used a fiber cable of length of 100 km having a PMD parameter (D_p) equal to 1 ps/$\sqrt{\text{km}}$, resulting in a total mean DGD of 10 ps.

A. Discrete-Time Equivalent Channel

Since the overall channel for the fiber link is bandlimited due to transmit and receive filters, we can model the effective

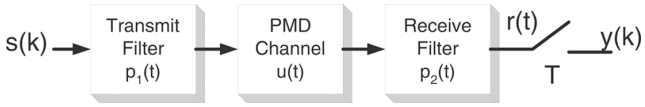


Fig. 2. Channel simulation model.

channel within the transmission band in the sampled domain, as shown in Fig. 2. The impulse response of the corresponding system can be obtained by fitting an FIR impulse response spectrum to the transfer function in the transmit band with a least squares approximation. Although the overall channel is a 2×2 MIMO system, we describe the fitting procedure for one of its four scalar components.

1) The transmitted discrete-time sequence is represented by $\{s(k), k \in \mathcal{Z}\}$. Typically, either real-valued BPSK or complex-valued QPSK-QAM constellations are used for transmission symbols.
2) Let $p_1(t)$ and $p_2(t)$ be the impulse responses of the transmit and the receive filters, respectively, where we choose $p_2(t) = p_1^*(-t)$, i.e., as the matched filter to the transmit pulse shaping filter. Let $u(t)$ represent the impulse response corresponding to the fiber link, i.e., the inverse continuous time Fourier transform (CTFT) of $U(w)$. Then, the overall channel impulse response would be equivalent to

$$c(t) = p_1(t) \otimes u(t) \otimes p_2(t) = p(t) \otimes u(t) \tag{3}$$

where $p(t) = p_1(t) \otimes p_2(t)$.
3) As stated earlier, the combined receive transmit filter impulse response $p(t)$ is selected as the raised cosine spectrum [$p_1(t)$ is a square-root raised cosine filter and $p_2(t)$ is time reversed and conjugated version of $p_1(t)$] whose CTFT is bandlimited to $(1/T)$

$$P(f) = 0, \quad \text{for} \quad |f| > \frac{1}{T} \tag{4}$$

where T is the symbol period. Therefore, the CTFT of the combined channel $c(t)$ is also bandlimited to $(1/T)$

$$C(f) = P(f)U(f) = 0, \quad \text{for} \quad |f| > \frac{1}{T}. \tag{5}$$

4) As $C(f)$ is bandlimited to $(1/T)$, the output of the receive filter $r(t)$, which is given by

$$r(t) = \sum_{k=-\infty}^{\infty} s(k)c(t - kT) \tag{6}$$

would also be bandlimited to $(1/T)$. Therefore, sampling $r(t)$ at a rate of $(2/T)$ (Nyquist sampling rate) would be sufficient

$$\begin{aligned} z(n) &= r\left(n\frac{T}{2}\right) = \sum_{k=-\infty}^{\infty} s(k)c\left(n\frac{T}{2} - kT\right) \\ &= \sum_{k=-\infty}^{\infty} s(k)c\left((n-2k)\frac{T}{2}\right) \\ &= d(k) \otimes g(k) \end{aligned} \tag{7}$$

where $d(k)$ is the upsampled version (with an upsampling rate of 2) of $s(k)$, i.e.,

$$d(k) = \begin{cases} s(k/2) & k \quad \text{is even} \\ 0 & k \quad \text{is odd} \end{cases} \tag{8}$$

and $g(k)$ are the sampled version of the continuous time impulse response $c(t)$ with a sampling rate of $(2/T)$, i.e.,

$$g(k) = c\left(k\frac{T}{2}\right) \quad k \in Z. \tag{9}$$

Since $C(f)$ is bandlimited to $(1/T)$, the discrete-time Fourier transform (DTFT) of g_k obeys

$$G(w) = \sum_{k=-\infty}^{\infty} g(k)e^{-jwk} = C\left(\frac{w}{T\pi}\right) \quad \text{for } -\pi \le \omega \le \pi. \tag{10}$$

Now a causal FIR filter approximation is fit to this frequency spectrum

$$\sum_{k=0}^{L-1} \hat{g}(k)e^{-jwk} \approx C\left(\frac{w}{T\pi}\right), \quad \text{for } -\pi \le \omega \le \pi \tag{11}$$

where L is the filter length, which is chosen to be an odd number. The parameters $\hat{g}(k)$ can be found by applying least squares to the system of equations obtained by evaluating aforementioned equation at different frequencies. Assuming that a symbol-rate sampling is used at the receiver, and we represent the symbol-spaced receive filter output samples with $o(k)$, we have

$$y(n) = r(nT) = s(k) \otimes h(k) \tag{12}$$

where the overall discrete-time equivalent channel $h(k)$ is given by

$$h(k) = \hat{g}(2k) \quad k = 0, 1, \ldots, \frac{L-1}{2}. \tag{13}$$

Following the steps described earlier, all four scalar components of the 2×2 MIMO discrete-time equivalent impulse response of the overall channel can be obtained. In other words, if $h_{ij}(k)$ represent the impulse response of the channel between the input j and output i obtained by the procedure above, we can write the overall discrete-time equivalent impulse response as

$$\mathbf{H}(k) = \begin{bmatrix} h_{11}(k) & h_{12}(k) \\ h_{21}(k) & h_{22}(k) \end{bmatrix}, k = 0, \ldots, N-1. \tag{14}$$

The resulting discrete-time impulse response could be used to measure the effective ISI of the overall channel (before compensation) as a function of symbol rate. For that purpose, we introduce the signal to ISI energy ratio (SISIR) measure defined as

$$\text{SISIR} = \frac{\rho^2}{\sum_{k=0}^{N-1} \|\mathbf{H}(k)\|_F^2 - \rho^2} \tag{15}$$

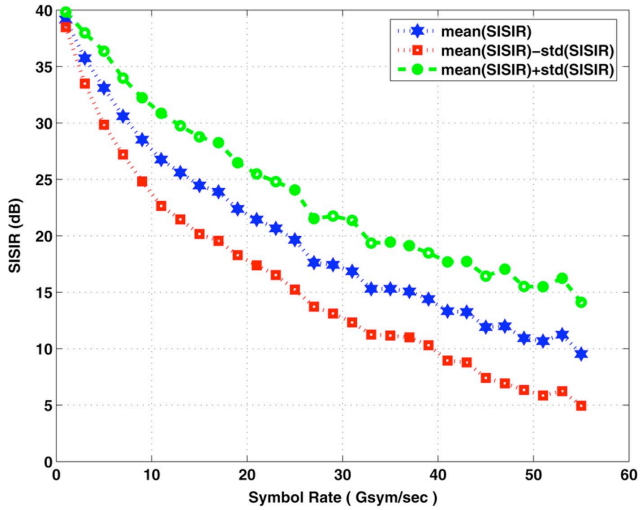


Fig. 3. Mean and standard deviation variation around its mean of SISIR (dB) for data rates between 1 Gsymbol/s and 55 Gsymbols/s (for a fiber with a total mean DGD of 10 ps).

where ρ is the Frobenius norm[1] of the main tap (i.e., the tap with the maximum Frobenius norm)

$$\rho = \max_{k \in \{0,1,\ldots,N-1\}} \|\mathbf{H}(k)\|_F. \tag{16}$$

In Fig. 3, for a fiber with a total mean DGD of 10 ps (which is the fiber link assumed in simulations), we plot SISIR as a function of symbol rate. Due to the random nature of PMD, SISIR is also random. Therefore, we plot its mean and its standard deviation variation around its mean values.

The graph indicates that increasing data rates result in higher ISI (lower SISIR) as expected. This is because when we increase the data rate, the dispersion-induced pulse broadening becomes longer than symbol periods due to the decreased symbol durations.

In the rest of the article, we would assume that the symbol rates are moderately low (e.g., less than 10 Gsymbols/s) such that the ISI is not the limiting factor. In such a scenario, main limiting factor will be the crosstalk between the two PDM channels induced by the fiber-induced polarization fluctuation, and in the following sections, we introduce various algorithms to solve this problem. The high symbol-rate case, where both ISI and the polarization crosstalk need to be handled, will be the subject of a separate article.

III. Blind Separation of Polarization Channels for Low Symbol Rates

Blind processing is used when no training signal is available. Although the training-based (i.e., supervised) adaptive filtering is currently being used in several communication applications, it has serious drawbacks. First of all, the useful available bandwidth of the channel is wasted by training sequences. This reduces the overall capacity of the channel drastically. Furthermore, there are some cases where the use of training sequences

[1]Frobenius norm of an $m \times n$ matrix $\mathbf{A}$ is defined as $\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|A_{ij}|^2}$.

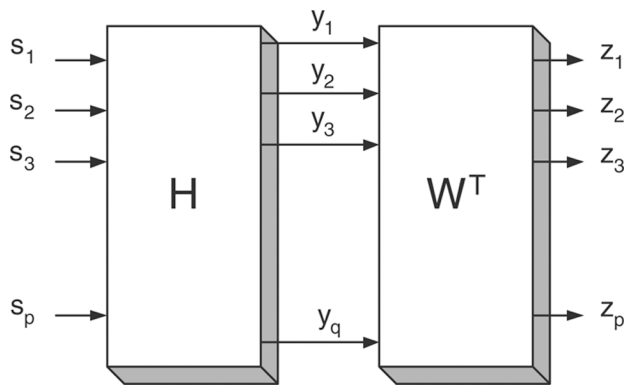


Fig. 4. BSS setup.

is not practical or very difficult to establish: The same communication link may be shared by several users and inclusion of a new user requires interruption in the service for other users. Furthermore, locating and synchronizing the training data location may be a nontrivial task especially for high-rate signals. Finally, due to the nonstationary nature of the fiber link, the receiver algorithm needs to track the variations without using training information. As a result, unsupervised or blind adaptation of receivers is more preferable, which is indeed our approach in this work. Blind separation is the process of extracting original input signals from their mixtures without any information about the mixing matrix and with only some information about the statistics of the input data. A general setup for the separation problem is shown in Fig. 4, where

- $s_1, \ldots, s_p$ are independent source signals and $y_1, \ldots, y_q$ are the mixture signals. In the PDM application, p and q are both equal to 2, i.e., there are two source signals, $s_1(k)$ and $s_2(k)$, corresponding to two sequences transmitted at the orthogonal polarizations and two mixture signals, $y_1(k)$ and $y_2(k)$, corresponding to the orthogonal polarization signals at the receiver.
- $\mathbf{H}$ is the unknown mixing channel. In the PDM application, there are two input polarization signals and two output polarization signals. Therefore, $\mathbf{H}$ is a 2×2 system. For the low symbol rate case assumed in this article, the mixing system $\mathbf{H}$ is memoryless, i.e., the mixtures at any time instants are functions of only the sources at the same time instant, which is referred as the *instantaneous* BSS problem. This refers to the case where there is no mixing in time and therefore no ISI. As noted before, the instantaneous BSS scenario is applicable only if the symbol rate is low such that the frequency response of the channel within the transmission bandwidth can be assumed to be flat. Referencing Fig. 3, for data rates up to 10 Gsymbols/s, the channel can be considered practically ISI free as SISIR is considerably high at those rates (for the assumed fiber link with a total mean DGD of 10 ps). In this case, the outputs y_1 and y_2 of $\mathbf{H}$ can be written as

$$\underbrace{\begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix}}_{\mathbf{y(k)}} = \mathbf{H} \underbrace{\begin{bmatrix} s_1(k) \\ s_2(k) \end{bmatrix}}_{\mathbf{s(k)}}. \tag{17}$$

- The purpose of BSS is to extract the source signals from the observation sequences $\mathbf{y(k)}$ using a linear system with transfer matrix $\mathbf{W}^T$, i.e.

$$\mathbf{z(k)} = \mathbf{W}^T \mathbf{y(k)} \tag{18}$$

where $\mathbf{z(k)} = [\, z_1(k) \quad z_2(k) \quad \ldots \quad z_p(k) \,]^T$ contains the estimates of the source signals. $\mathbf{W}$ is obtained adaptively from the time samples (realizations) of $\mathbf{y(k)}$. No *a priori* knowledge of $\mathbf{H}$ and no training sequences are assumed.
- For a typical BSS algorithm, $\mathbf{W}$ is decomposed into two operators $\mathbf{W} = \mathbf{W}_{\text{pre}}\mathbf{\Theta}$ where
  - $\mathbf{W}_{\text{pre}}$ is used to whiten the channel outputs. The output of this section is two uncorrelated signals.
  - $\mathbf{\Theta}$ is the unitary separator, which is used to convert uncorrelated signals generated by the whitening block into independent original polarization signals.

  Fortunately, in the fiber PDM, application mixing channel is (nearly) unitary due to our standing assumptions about the fiber model. Therefore, whitening phase can be skipped, which brings a considerable relief in terms of complexity for the BSS algorithm implementation. Therefore, we can choose $\mathbf{W}_{\text{pre}} = \mathbf{I}$ and consequently $\mathbf{W} = \mathbf{\Theta}$.

  Although the elements of $\mathbf{y(k)}$ are uncorrelated, they are not necessarily independent. Therefore, our goal is to obtain a unitary matrix $\mathbf{\Theta}$, which will convert the uncorrelated $\mathbf{y(k)}$ vector into an independent vector. For such a choice of $\mathbf{\Theta}$, we would obtain

$$\mathbf{z(k)} = \mathbf{\Theta}^T \mathbf{y(k)} = \mathbf{\Phi P s(k)} \tag{19}$$

where $\mathbf{P}$ is a permutation matrix and $\mathbf{\Phi}$ is a diagonal matrix with unit-magnitude complex entries given as

$$\mathbf{\Phi} = \begin{bmatrix} e^{j\phi_1} & 0 \\ 0 & e^{j\phi_2} \end{bmatrix}. \tag{20}$$

  Here matrices $\mathbf{P}$ and $\mathbf{\Phi}$ represent the unavoidable permutation (ordering of polarizations) and phase ambiguities. These ambiguities can be easily resolved via exploitation of some side information after convergence. The algorithm introduced in the next section restricts the phase ambiguity to a set of finite values (e.g., $\{0, (\pi/2), \pi, -(\pi/2)\}$ for 4-QPSK signaling), which could be easily resolved and compensated.
- Obtaining $\mathbf{\Theta}$ to convert the uncorrelated vector $\mathbf{y}$ into an independent vector $\mathbf{z}$ is typically posed as an optimization problem in the form

$$\begin{aligned} &\text{maximize/minimize} \quad \mathcal{J}(\mathbf{z(k)}) \\ &\text{subject to} \quad \mathbf{\Theta}^H \mathbf{\Theta} = \mathbf{I}. \end{aligned}$$

  Here $\mathcal{J}$ is the objective/cost function, which reflects the level of independence of $\mathbf{z}$ as a function of $\mathbf{\Theta}$, to be optimized. Different choices for $\mathcal{J}$ leads to development of different algorithms with varying performance and complexities. In the next section, we introduce a low-complexity BSS approach suitable for implementation in high symbol rate optical fiber communication systems.

IV. MB-BSS Algorithm for Polarization Separation

The approach that we follow in this article exploits the bounded magnitude property of digital communication sources, in addition to the independence of polarization components; hence, it is named BSS for bounded magnitude sources (MB-BSS) [31].

The MB-BSS approach is based on the optimization setting

$$\begin{aligned} &\text{minimize} \quad \sup\|\mathcal{R}e\{\mathbf{z}(\mathbf{k})\}\|_\infty \\ &\text{subject to} \quad \mathbf{\Theta}^H\mathbf{\Theta} = \mathbf{I} \end{aligned}$$

which corresponds to minimizing the maximum real component of the output over all output components and over the ensemble. This setting does not have a closed form solution, so the solution is obtained through a simple search process. Since the objective function is convex but nonsmooth, the search is performed through the use of subgradient directions. The following sections describe the original algorithm proposed in [31] and its low-complexity variations.

A. Windowed Version of the MB-BSS Algorithm

In order to obtain a practical algorithm for the proposed optimization setting, which can be used in real-time applications, the evaluation of the maximum value of the infinity norm of $\mathbf{z}(\mathbf{k})$ should be limited to a finite window of time samples of z, i.e., $\{\mathbf{z}(\mathbf{k}) : k \in \{0, 1, \ldots, \Omega-1\}\}$. Here, we make a certain ergodicity assumption that these time samples reflect (or approximate) the ensemble behavior in terms of the infinity norm of z. The basic subgradient search algorithm corresponding to the optimization setting above, which is proposed in [31] can be described as follows:

If we define

$$\mathbf{Y} = [\,\mathbf{y(0)} \quad \mathbf{y(1)} \quad \ldots \quad \mathbf{y(\Omega-1)}\,]$$

as the matrix of input values in the window of interest, then for a given $\mathbf{\Theta}$, the corresponding outputs can be placed in a matrix $\mathbf{Z}$

$$\mathbf{Z} = [\,\mathbf{z(0)} \quad \mathbf{z(1)} \quad \ldots \quad \mathbf{z(\Omega-1)}\,] = \mathbf{\Theta}^T \quad \mathbf{Y}.$$

Then, update of the algorithm is given as

$$\begin{aligned} \underline{\mathbf{\Theta}}^{(i+1)} &= \mathbf{\Theta}^{(i)} - \mu^{(i)}\text{sign}\left(\mathcal{R}e\left\{\mathbf{Z}^{(i)}_{m^{(i)},n^{(i)}}\right\}\right)\bar{\mathbf{Y}}_{:,n^{(i)}}\mathbf{e}^T_{m^{(i)}} \\ \mathbf{\Theta}^{(i+1)} &= \mathcal{P}_{\mathcal{U}}\left\{\underline{\mathbf{\Theta}}^{(i+1)}\right\} \end{aligned} \tag{21}$$

where $\mathbf{\Theta}^{(i)}$ is the value of $\mathbf{\Theta}$ at the ith iteration, $\mathbf{Z}^{(i)}$ is the output matrix calculated based on $\mathbf{\Theta}^{(i)}$, $(m^{(i)}, n^{(i)})$ is the index for the maximum real component magnitude entry of $\mathbf{Z}^{(i)}$, $\mu^{(i)}$ is the step size at the ith iteration, $\underline{\mathbf{\Theta}}^{(i+1)}$ is the unprojected version of the updated $\mathbf{\Theta}$, $\mathcal{P}_{\mathcal{U}}$ is the projection operator to the unitary matrix set, $\bar{\mathbf{Y}}_{:,n^{(i)}}$ is the nth column of $\mathbf{Y}^{(i)}$ [which is $\mathbf{y(n)}$] with its elements complex conjugated, and $\mathbf{e}_{m^{(i)}}$ is the mth standard basis vector.

After each subgradient update, $\mathbf{\Theta}$ matrix is potentially moved away from the set of unitary matrices. Therefore, the updated $\mathbf{\Theta}$ matrix needs to be projected back onto the unitary matrix set to realize the unitary constraint. The technique we use for $\mathcal{P}_{\mathcal{U}}$ is minimum-(Frobenious)-distance orthogonalization. With this technique, we try to find the projection operator, which projects $\underline{\mathbf{\Theta}}^{(i+1)}$ to the minimum Frobenius-distance unitary matrix. The projection is obtained by (see e.g., [37])

$$\begin{aligned} \mathbf{\Theta}^{(i+1)} &= \mathcal{P}_{\mathcal{U}}\left\{\underline{\mathbf{\Theta}}^{(i+1)}\right\} \\ &= \mathcal{P}_{\mathcal{U}}\left\{\mathbf{U}^{(i+1)}\mathbf{\Sigma}^{(\mathbf{i+1})}\mathbf{V}^{(\mathbf{i+1})^{\mathbf{H}}}\right\} \\ &= \mathbf{U}^{(i+1)}\mathbf{V}^{(i+1)^H} \end{aligned} \tag{22}$$

where $\mathbf{U}^{(i+1)}\mathbf{\Sigma}^{\mathbf{i+1}}\mathbf{V}^{(\mathbf{i+1})^H}$ is the singular value decomposition (SVD) of $\underline{\mathbf{\Theta}}^{(i+1)}$.

B. Simplified Version of the MB-BSS Algorithm

In order to reduce the implementation complexity of the windowed MB-BSS algorithm, we propose the following simplification: The original windowed version of the MB-BSS algorithm requires storage to keep track of the maximum value and the corresponding subgradient components. The need for such storage could be eliminated by replacing window-based structure with a sample-by-sample update structure. According to this method, the current maximum magnitude real output of the separator is compared against a target value, which could be calculated from the constellation structure. For example, if 4-QAM type constellation is used for both polarizations, the target level for the maximum absolute real component is equal to 1 assuming that the output is normalized to have a variance of 2. The corresponding algorithm can be summarized as follows:

- Assume that the received signal vector is $\mathbf{y}_i$ (channel output) at the instant. Then, the separator output $\mathbf{z}_i$ is given as

$$\mathbf{z}_i = \mathbf{\Theta}^T \quad \mathbf{y_i} = \begin{bmatrix} a+jb \\ c+jd \end{bmatrix}. \tag{23}$$

- Let $K = \max\{|a|, |c|\}$ and let m be the index of this maximum. For example, if $|a| > |c|$ then $m = 1$, otherwise $m = 2$. The simplified MB-BSS algorithm compares K with 1. If $K < 1$, no operation is performed. If $K > 1$, then the separator is updated with the subgradient given by (21).

$$\underline{\mathbf{\Theta}}^{(i+1)} = \mathbf{\Theta}^{(i)} - \mu^{(i)}V^{(i)}\bar{\mathbf{y}}_i\mathbf{e}^T_{m^{(i)}} \tag{24}$$

$$\mathbf{\Theta}^{(i+1)} = \mathcal{P}_{\mathcal{U}}\left\{\underline{\mathbf{\Theta}}^{(i+1)}\right\} \tag{25}$$

where

$$V^{(i)} = \begin{cases} \text{sign}(a), & \text{if } m^{(i)} = 1 \\ \text{sign}(c), & \text{if } m^{(i)} = 2 \end{cases}. \tag{26}$$

After the separator update, the next step is unitary projection indicated by (25).

The second complicated step in the original algorithm is the projection to the set of unitary matrices. This operation requires that SVD is performed for each update or at least for a group of updates. This would necessitate the use of a complicated DSP unit with hard computational restrictions for the real-time implementation.

In order to eliminate the need for the projection to unitary matrices, one can use a parametrization approach to represent the

unitary matrices. According to this approach, the unitary matrix is parameterized with several unconstrained real variables; therefore, the original constrained problem is converted to an unconstrained one. The unconstrained version could potentially enable a low-complexity hardware implementation.

Any general 2×2 unitary matrix $\mathbf{W}$ can be parameterized as follows:

$$\mathbf{W} = e^{j\phi} \overbrace{\begin{bmatrix} \cos(\alpha)e^{jx} & \sin(\alpha)e^{jy} \\ -\sin(\alpha)e^{-jy} & \cos(\alpha)e^{-jx} \end{bmatrix}}^{\mathbf{U}} \tag{27}$$

where the exponential represents a global phase factor and $\mathbf{U}$ is a 2×2 unitary matrix with determinant $+1$. Therefore, according to (27), any general 2×2 unitary matrix can be parameterized using four real parameters (ϕ, α, x, and y).

The derivation of the parametric MB-BSS algorithm includes partial derivatives with respect to the four real variables aforementioned. We will use the chain rule to obtain the updates with respect to these four real parameters. First, subgradient is obtained with respect to the unitary separator. After the subgradient is obtained, the parameterized unitary separator is differentiated with respect to the four real parameters. We will only show the derivation for the variable α since we perform the same operations for the remaining parameters. The gradient with respect to α is derived through the following steps:

1) The cost function can be first written in a more convenient form in terms of the vectorized version of $\mathbf{\Theta}$

$$\begin{aligned} f(\mathbf{\Theta}) &= \|\mathcal{R}e\{\mathrm{vec}(\mathbf{z})\}\|_\infty \\ &= \|\mathcal{R}e\{\underbrace{(\mathbf{z}^T \otimes \mathbf{I})}_{\mathbf{A}} \underbrace{\mathrm{vec}(\mathbf{\Theta}^{\mathbf{T}})}_{\mathbf{b}}\}\|_\infty \\ &= \|\mathcal{R}e\{\mathbf{Ab}\}\|_\infty \end{aligned} \tag{28}$$

where "vec" denotes the vectorization operation of the matrix and $\otimes$ denotes the Kronecker product.

2) We can further process the resulting expression for the cost function to write it in terms of vectors of real variables

$$\begin{aligned} &\|\mathcal{R}e\{\mathbf{Ab}\}\|_\infty \\ &\quad = \|\mathcal{R}e\{(\mathbf{A}_R + j\mathbf{A}_I)(\mathbf{b}_R + j\mathbf{b}_I)\}\|_\infty \\ &\quad = \|\mathbf{A}_R\mathbf{b}_R - \mathbf{A}_I\mathbf{b}_I\|_\infty \\ &\quad = \left\| \underbrace{[\mathbf{A}_R \quad -\mathbf{A}_I]}_{\mathbf{K}} \underbrace{\begin{bmatrix} \mathbf{b}_R \\ \mathbf{b}_I \end{bmatrix}}_{\mathbf{r}} \right\|_\infty \\ &\quad = \|\underbrace{\mathbf{Kr}}_{\mathbf{s}}\|_\infty = f(\mathbf{r}), \end{aligned} \tag{29}$$

where the subscript "R" denotes the real part of the component while "I" denotes the imaginary part.

3) A subgradient of the cost function with respect to $\mathbf{r}$ can be written as

$$\underline{\mathbf{g}} = \mathrm{sign}(s_i)\mathbf{K}_{i,:}^T \quad \text{where} |s_i| = \|\mathbf{s}\|_\infty. \tag{30}$$

4) [2] Assuming that the cost function is differentiable with respect to α, which is true except at a discrete set of points, we apply the chain rule to obtain derivative with respect to α

$$\frac{\partial f(\mathbf{r})}{\partial \alpha} = \underline{\mathbf{g}}^T \cdot \frac{\partial \mathbf{r}}{\partial \alpha}. \tag{31}$$

In order to obtain $(\partial \mathbf{r})/(\partial \alpha)$, first of all, the unitary separator matrix is vectorized to obtain $\mathbf{b}$. Then, real and imaginary parts of the vector $\mathbf{b}$ is placed on top of each other to construct the vector $\mathbf{r}$. These operations were all explained in preceding steps. If we skip the details, then the partial derivative vector is given by

$$\frac{\partial \mathbf{r}}{\partial \alpha} = \begin{bmatrix} -\sin(\alpha)\cos(\phi + x) \\ \cos(\alpha)\cos(\phi + y) \\ -\cos(\alpha)\cos(\phi - y) \\ -\sin(\alpha)\cos(\phi - x) \\ -\sin(\alpha)\sin(\phi + x) \\ \cos(\alpha)\sin(\phi + y) \\ -\cos(\alpha)\sin(\phi - y) \\ -\sin(\alpha)\sin(\phi - x) \end{bmatrix}. \tag{32}$$

5) Final step is the update of the variable achieved as

$$\alpha^{(i+1)} = \alpha^{(i)} - \mu^{(i)} \underline{\mathbf{g}}^T \cdot \frac{\partial \mathbf{r}}{\partial \alpha} \tag{33}$$

where $\underline{\mathbf{g}}$ is given in (30) and $(\partial \mathbf{r})/(\partial \alpha)$ in (32).

The elimination of the window structure together with parametrization described earlier enables the implementation of the compensation using a low-complexity hardware structure, eliminating the need for an embedded CPU. Simplified version of the algorithm removes the need for memory, unitary projection operation, and other complex operations that can only be carried out by DSPs. Therefore, the application of the simplified algorithm with the parametric algorithm updates makes the all-hardware implementation possible. The resulting algorithm would update four real parameters based on the result of the comparison of maximum absolute real component with the threshold value. If an update is required, the four real parameters are updated with the received signal vector $\mathbf{y}_i$. Otherwise no operation is carried out. In the next section, we present simplified hardware implementation corresponding to the parametrization-based algorithm presented earlier.

C. Hardware Implementation of the Simplified MB-BSS

The simplified MB-BSS algorithm is suitable for low-complexity hardware implementation as it does not contain computationally involved operations such as SVD or matrix inversion. Fig. 5 illustrates a sample implementation for the simplified MB-BSS algorithm. The overall structure consists of two major functional sections.

1) "Threshold Decision" Unit: This block has two inputs, which are the real components of the separator outputs. This block determines whether an update is necessary or not by threshold comparison. It has two outputs denoted by signZ and Data Select. If an update is not necessary, signZ is set to 0. Otherwise signZ indicates the sign of the maximum magnitude real output of the separator (which is $V^{(i)}$ given by (26)), and Data Select is the index of this maximum (which is $m^{(i)}$).

[2] $\mathbf{K}_{i,:}$ refers to the ith row of $\mathbf{K}$.

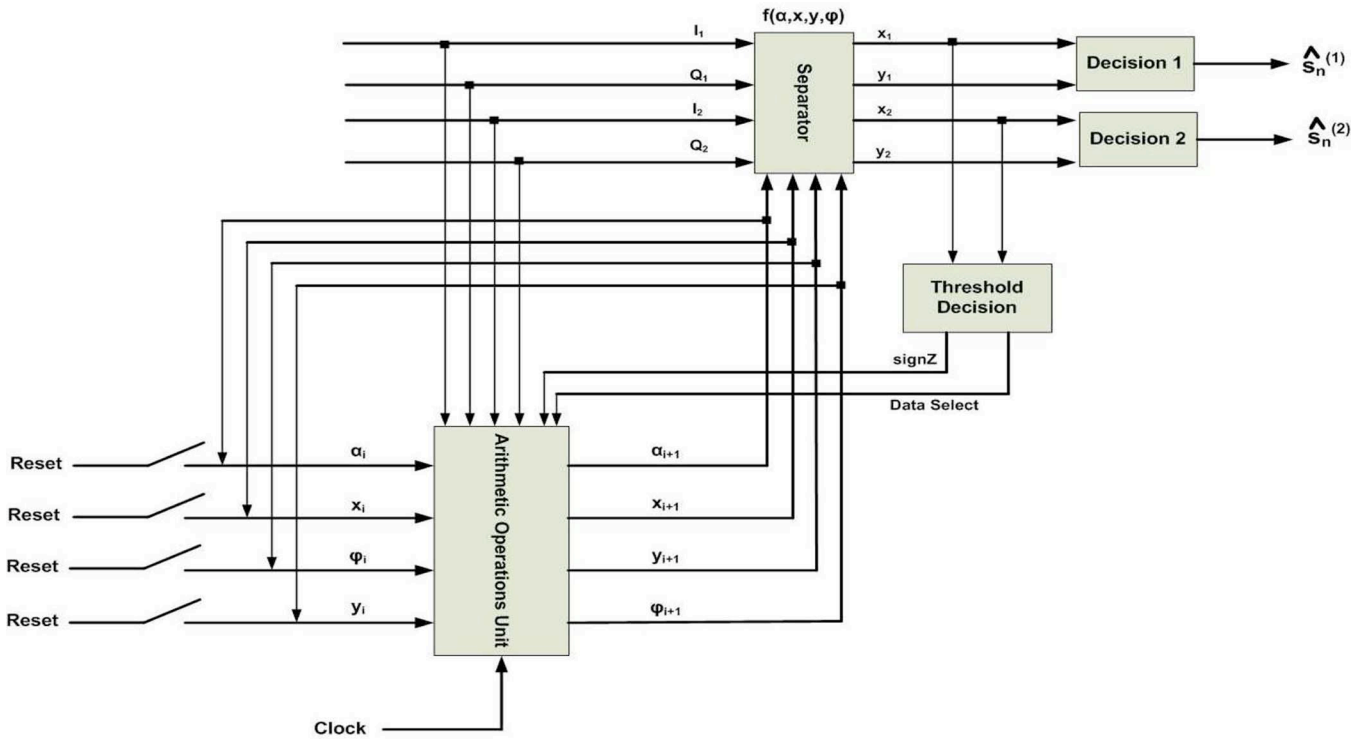


Fig. 5. Simplified hardware implementation of the overall simplified parametric MB-BSS system.

2) "Arithmetic Operations" Unit: This block is used to perform the algorithm updates on the parameters for the unitary separator based on the control signals generated by the threshold decision unit.

As this sample implementation suggests, the proposed adaptive receiver structure could be implemented using a relatively low-complexity structure. The next section presents the performance of the proposed algorithms.

V. Simulation Results

In this section, we present the simulation performance of both windowed MB-BSS algorithm and its simplified version. In the simulations, we assume the following:

1) Each polarization uses 4-QPSK signalling.
2) The fiber link has a total mean DGD level of 10 ps. The symbol rate is assumed to be 5 Gsymbols/s. Therefore, according to Fig. 3, SISIR is typically more than 30 dB, which implies that the effects of ISI due to PMD is negligible. In other words, the performance of the receiver would be limited by the noise and the polarization separation performance.
3) The output of the fiber channel is assumed to be corrupted by amplifier spontaneous emission (ASE) noise caused by erbium-doped fiber amplifiers (EDFAs), which is modeled as additive white Gaussian noise.
4) The receiver A–D converters (ADC) are assumed to have a precision of 5 bits.

A. Simulation Results for the Windowed MB-BSS Algorithm

We first present simulation results for the windowed version of the MB-BSS algorithm. In these simulations, the SNR (where noise component is due to ASE only) level is chosen as 20 dB. The receiver employs a moving window version of the algorithm described earlier, where each window has a length of 30 samples. The simulation results for the windowed MB-BSS algorithm are shown in Fig. 6, where (a) is the constellation plot for one of the polarization components at the compensator input, (b) is the constellation plot for one of the polarizations at the compensator output after convergence. Fig. 6(c) shows the signal to interference noise power ratio (SINR) as a function of iterations. According to this figure, SINR quickly and successfully converges to the vicinity of 20 dB limit in about 100 iterations. This would correspond to a time span of 3000 symbols, which would be equivalent to 0.6 μs. Since random variations in fiber polarization state occur in millisecond time scales, the algorithm would be successful in both acquiring the initial separation level and tracking the channel variation due to the fiber-induced polarization fluctuation.

In order to illustrate the behavior of the moving window algorithm over many different realizations of the fiber channel, we repeated the simulation for 10 000 different realizations of the fiber channel for an SNR level of 20 dB. In each simulation run, we recorded the final SINR value achieved at the output of the separator. Then, we obtained the empirical distribution [probability density function (pdf)] of the SINRs based on these 10 000 values. In order to illustrate the effect of the symbol rate, we performed these experiments for both 5 Gsymbols/s and 20 Gsymbols/s. The empirical probability distribution functions obtained in these simulations are shown in Fig. 7. As can be seen from these figures, for 5 Gsymbols/s symbol rate, the output SINR is closely distributed around 20 dB. This indicates that the algorithm successfully recovers orthogonal polarization signals.

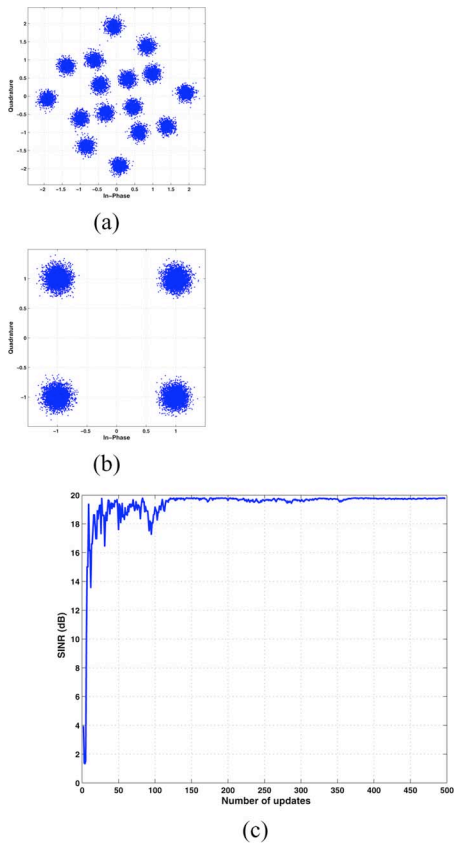


Fig. 6. MB-BSS simulation result for $\mathrm{SNR} = 20$ dB for moving window approach. (a) The corrupted channel output; (b) The separator output (recovered sources); (c) SINR convergence curve.

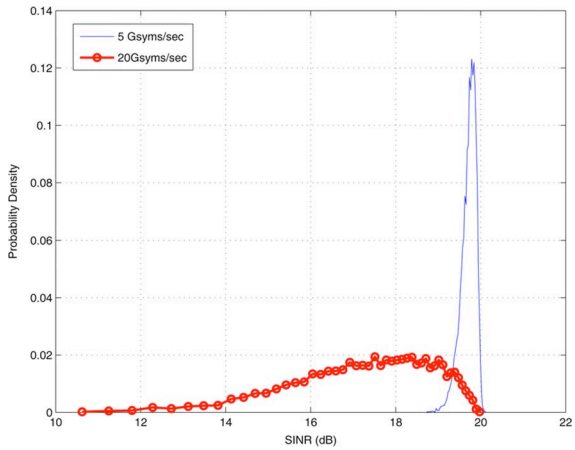


Fig. 7. Windowed algorithm: Output SINR distributions (after the convergence of the algorithm) for 5 Gsymbols/s and 20 Gsymbols/s. (ASE $\mathrm{SNR} = 20$ dB).

In 20 Gsymbols/s symbol rate, since the ISI component of the channel becomes strong, the SINR level is distributed around 12–20 dB. The algorithm is still successful in separating polarizations; however, the ISI component caused by the time dispersion degrades the SINR level. This is not a surprising result, as we noted earlier that the algorithm introduced in this article does not handle the time dispersion effects due to PMD.

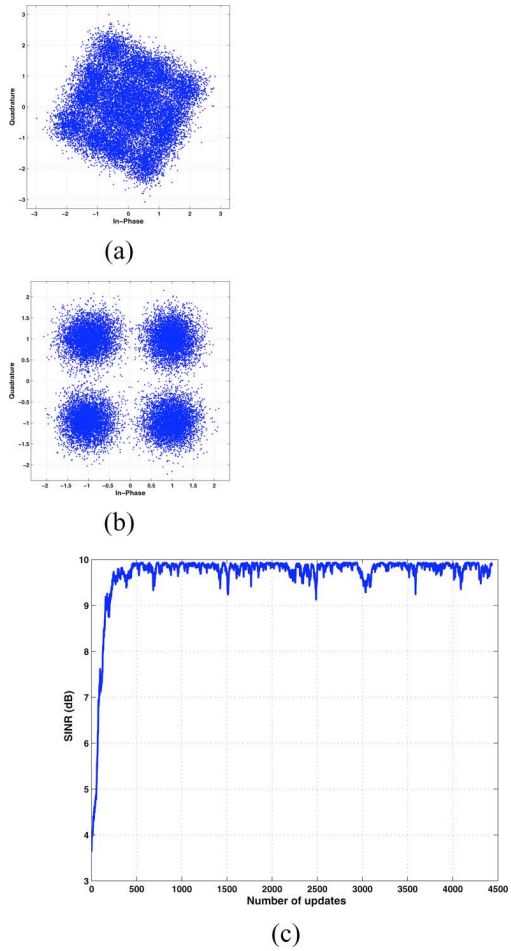


Fig. 8. Simplified parametric MB-BSS simulation result for $\mathrm{SNR} = 10$ dB. (a) The corrupted channel output; (b) The separator output (recovered sources); (c) SINR convergence curve.

B. Simulation Results for the Simplified Parametric MB-BSS Algorithm

In order to illustrate the performance of the simplified MB-BSS algorithm, we repeated the earlier simulation setup with the new algorithm. However, the SNR (due to ASE) level is decreased to 10 dB, which is even more challenging than the previous example. Fig. 8 shows the simulation results, where (a) and (b) are again the constellation plots at the input and output of the receiver compensator, respectively, for one of the polarization components. The convergence of SINR as a function of iterations is displayed in Fig. 8. According to this figure approximately 760 updates are sufficient for the convergence. This data duration corresponds to 0.152 μs at 5 Gsymbols/s. We can, therefore, conclude that simplified parametric MB-BSS has also fast convergence behavior with sufficient performance level. Over the various channel and noise scenarios that we experimented, typically both the windowed algorithm and the simplified parametric algorithm achieve very similar SINR levels, i.e., there is no loss in using the simplified algorithm.

In order to evaluate the performance of the simplified algorithm on various realizations of the fiber channel, we repeated the aforementioned experiment for 10 000 times. We calculated the empirical pdfs for SINRs for both 5 Gsymbols/s

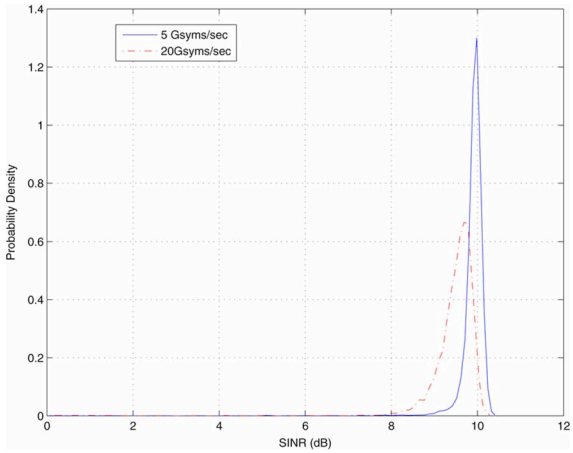


Fig. 9. Simplified parametric algorithm: Output SINR distributions (after the convergence of the algorithm) for 5 Gsymbols/s and 20 Gsymbols/s. (ASE $\mathrm{SNR} = 10$ dB).

and 20 Gsymbols/s, which are shown in Fig. 9. Similar to the windowed version, the SINR values are at the close vicinity of the 10 dB level, especially for 5 Gsymbols/s case. The performance degradation in 20 Gsymbols/s is not high due to the fact that the ISI level in this case is mostly below the noise level corresponding to 10 dB SNR level for the channel.

To conclude, based on the simulation results, for both windowed and simplified parametric algorithms, both of the algorithms are successful in achieving separation of polarization components.

VI. Conclusion

In this work, we proposed a framework for the development of adaptive receivers for fiber optic systems employing the PDM scheme. The use of the PDM scheme promises potential increase of the data transmission rate if the receiver successfully counteracts the space mixing caused by the fiber-induced polarization fluctuation phenomenon. The MB-BSS-based framework targets to achieve this goal, where the proposed algorithms could lead to increase in data rates due to the fact that near perfect polarization separation can be achieved using an unsupervised adaptive structure. The simulation results illustrate that the proposed approach can achieve near optimal performance levels. Due to its fast convergence speed, in microsecond time scales, the tracking of the nonstationary fiber channel is possible.

References

[1] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single mode fibers," *Electron. Lett.*, vol. 22, no. 19, pp. 1029–1030, 1986.

[2] A. O. Lima, I. T. Lima, T. Adali, and C. R. Menyuk, "A novel polarization diversity receiver for PMD mitigation," *IEEE Photon. Technol. Lett.*, vol. 14, no. 4, pp. 465–467, Apr. 2002.

[3] Y. Mochida, N. Yamaguchi, and G. Ishikawa, "Technology-oriented review and vision of 40-gb/s-based optical transport networks," *J. Lightw. Technol.*, vol. 20, no. 12, pp. 2272–2281, Dec. 2002.

[4] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. Natl. Acad. Sci.*, vol. 97, pp. 4541–4550, 2000.

[5] A. Demir and A. T. Erdogan, "Emulation and inversion of polarization mode dispersion: A lossless system and reduced-order modeling perspective," *J. Lightw. Technol.*, vol. 26, no. 17, pp. 3071–3089, Sep. 2008.

[6] C. R. Doerr, S. Chandrasekhar, P. J. Winzer, A. R. Chraplyvy, A. H. Gnauck, L. W. Stulz, R. Pafchek, and E. Burrows, "Simple multichannel optical equalizer mitigating intersymbol interference for 40-gb/s nonreturn-to-zero signals," *J. Lightw. Technol.*, vol. 22, no. 1, pp. 249–257, Jan. 2004.

[7] T. Hirooka, M. Nakazawa, F. Futami, and S. Watanabe, "A new adaptive equalization scheme for a 160-gb/s transmitted signal using time-domain optical fourier transformation," *IEEE Photon. Technol. Lett.*, vol. 16, no. 10, pp. 2371–2373, Oct. 2004.

[8] T. Hirooka and M. Nakazawa, "Optical adaptive equalization of high-speed signals using time-domain optical fourier transformation," *J. Lightw. Technol.*, vol. 24, no. 7, pp. 2530–2540, Jul. 2006.

[9] F. Buchali and H. Bulow, "Adaptive PMD compensation by electrical and optical techniques," *J. Lightw. Technol.*, vol. 22, no. 44, pp. 1116–1126, Apr. 2004.

[10] M. Bohn, W. Rosenkranz, and P. M. Krummrich, "Adaptive distortion compensation with integrated optical finite impulse response filters in high bitrate optical communication systems," *IEEE J. Sel. Top. Quantum Electron.*, vol. 10, no. 2, pp. 273–280, Mar./Apr. 2004.

[11] M. Secondini, E. Forestieri, and G. Prati, "Plc optical equalizer for chromatic and polarization-mode dispersion compensation based on MSE control," *IEEE Photon. Technol. Lett.*, vol. 16, no. 4, pp. 1173–1175, Apr. 2004.

[12] A. J. Weiss, "On the performance of electrical equalization in optical fiber transmission systems," *IEEE Photon. Technol. Lett.*, vol. 15, no. 9, pp. 1225–1227, Sep. 2003.

[13] V. Curri, R. Gaudino, A. Napoli, and P. Poggiolini, "Electronic equalization for advanced modulation formats in dispersion-limited systems," *IEEE Photon. Technol. Lett.*, vol. 16, no. 13, pp. 2556–2558, Nov. 2004.

[14] O. E. Agazzi, M. R. Hueda, H. S. Carrer, and D. E. Crivelli, "Maximum-likelihood sequence estimation in dispersive optical channels," *J. Lightw. Technol.*, vol. 23, no. 2, pp. 749–763, Feb. 2005.

[15] J. Winters and R. Gitlin, "Electrical signal processing techniques in long-haul fiber-optic systems," *IEEE Trans. Commun.*, vol. 38, no. 9, pp. 1439–1453, Sep. 1990.

[16] H. F. Haunstein, W. Sauer-Greff, A. Dittrich, K. Sticht, and R. Urbansky, "Principles for electronic equalization of polarization-mode dispersion," *J. Lightw. Technol.*, vol. 22, no. 22, pp. 1169–1182, Apr. 2004.

[17] H. Wu, J. A. Tierno, P. Pepeljugoski, J. Schaub, S. Gowda, J. A. Kash, and A. Hajimiri, "Integrated transversal equalizers in high-speed fiber-optic systems," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2131–2137, Dec. 2003.

[18] P. M. Watts, V. Mikhailov, S. Savory, P. Bayvel, M. Glick, M. Lobel, B. C. an Peter Kirkpatrick, S. Shang, and R. I. Killey, "Performance of single-mode fiber links using electronic feed-forward and decision feedback equalizers," *IEEE Photon. Technol. Lett.*, vol. 17, no. 10, pp. 2206–2208, Oct. 2005.

[19] F. F. Dai, S. Wei, and R. Jaeger, "Integrated blind electronic equalizer for fiber dispersion compensation," in *Proc. ISCAS 2005*, May , vol. 6, pp. 5750–5753.

[20] A. T. Erdogan, A. Demir, and T. M. Oktem, "Automatic PMD compensation by unsupervised polarization diversity combining coherent receivers," *J. Lightw. Technol.*, vol. 26, no. 13, pp. 1823–1834, Jul. 2008.

[21] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. I. Unification and optimal designs," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1988–2006, Jul. 1999.

[22] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2000.

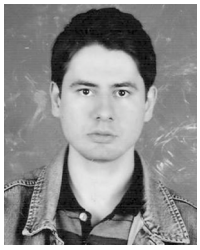
[23] A. Chraplyvy, A. Gnauck, R. W. Tkach, J. L. Zyskind, J. W. Sulhoff, A. J. Lucero, Y. Sun, R. M. Jopson, F. Forghieri, R. M. Derosier, C. Wolf, and A. McCormick, "1-Tb/s transmission experiment," *IEEE Photon. Technol. Lett.*, vol. 8, no. 9, pp. 1264–1266, Sep. 1996.

[24] S. G. Evangelides, L. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," *J. Lightw. Technol.*, vol. 10, no. 1, pp. 28–35, Jan. 1992.

[25] X. S. Yao, L.-S. Yan, B. Zhang, A. Willner, and J. Jiang, "All-optic scheme for automatic polarization division multiplexing," *Opt. Exp.*, vol. 15, no. 12, pp. 7407–7414, 2007.

[26] D. S. R. Noe, S. Hinz, and F. Wust, "Crosstalk detection schemes for polarization division multiplex transmission," *J. Lightw. Technol.*, vol. 19, no. 10, pp. 1469–1475, Oct. 2001.

[27] M. I. Hayee, M. C. Cardakli, A. Sahin, and A. E. Willner, "Doubling of bandwidth utilization using two orthogonal polarizations and power unbalancing in a polarization-division-multiplexing scheme," *IEEE Photon. Technol. Lett.*, vol. 13, no. 8, pp. 881–883, Aug. 2001.
[28] Y. Han and G. Li, "Experimental demonstration of direct-detection quaternary differential polarisation-phase-shift keying with electrical multilevel decision," *Electron. Lett.*, vol. 42, no. 2, pp. 109–111, 2006.
[29] G. P. Agrawal, *Fiber-Optic Communication Systems*. New York: Wiley-Interscience, 2002.
[30] D. E. Crivelli, H. S. Carrer, and M. R. Hueda, "Adaptive digital equalization in the presence of chromatic dispersion, PMD, and phase noise in coherent fiber optic systems," in *Proc. IEEE Commun. Soc. Globecom*, 2004, pp. 2545–2551.
[31] A. T. Erdogan, "A simple geometric blind source separation method for bounded magnitude sources," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 438–449, Feb. 2006.
[32] W. Weiershausen, R. Leppla, F. Kuppers, and H. Scholl, "Polarization-mode dispersion in fibre transmission: Theoretical approach, impact on systems, and suppression of signal-degradation effects," in *Proc. ECOC1999. 25th Eur. Conf. Opt. Commun.*, Nice, France, Sep. , vol. 2, pp. 130–132.
[33] T. Merker, N. Hahnenkamp, and P. Meissner, "Comparison of PMD-compensation techniques at 10 gbit/s using an optical first-order compensator and electrical tranversal filter," *Opt. Commun.*, vol. 182, pp. 135–141, 2000.
[34] H. Sunnerud, C. Xie, M. Karlsson, R. Samuelsson, and P. A. Andrekson, "A comparison between different PMD compensation techniques," *J. Lightw. Technol.*, vol. 20, no. 3, pp. 368–378, Mar. 2002.
[35] D. Marcuse, C. Menyuk, and P. K. A. Wai, "Application of the manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.
[36] G. D. VanWiggeren and R. Roy, "Transmission of linearly polarized light through a single-mode fiber with random fluctuations of birefringence," *Appl. Opt.*, vol. 38, pp. 3888–3892, 1999.
[37] J. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.



Turgut Oktem was born in Ankara, Turkey, in 1981. He received the B.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2004, the M.S. Degree from the Electrical and Computer Engineering Department, Koc University, Istanbul, Turkey, in November 2006. Currently, he is working toward the Ph.D. degree in the Mobile Communications Department of Eurecom Institute, France, since December 2007.

In September 2004, he joined the Electrical and Computer Engineering Department of Koc University as a Teaching and Research Assistant. His recent research interests include mobile communications, digital signal processing, convex optimization, and fiber optic communications.

Alper T. Erdogan (M'00) was born in Ankara, Turkey, in 1971. He received the B.S. degree from the Middle East Technical University, Ankara, Turkey, in 1993, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1995 and 1999, respectively.

He was a Principal Research Engineer in Globespan-Virata Corporation (formerly Excess Bandwidth and Virata Corporations) from September 1999 to November 2001. He joined Electrical and Electronics Engineering Department, Koc University, Istanbul, Turkey, in January 2002, where he is currently an Associate Professor. His current research interests include wireless, fiber and wireline communications, adaptive signal processing, optimization, system theory and control, and information theory.

Dr. Erdogan is the recipient of TUBITAK Career Award in 2005, Werner Von Siemens Excellence Award in 2007, and TUBA GEBIP Outstanding Young Scientist Award in 2008. He currently serves as the Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING.

Alper Demir (S'93–M'96) received the B.S. degree in electrical engineering from Bilkent University, Ankara, Turkey, in 1991, and the M.S. and Ph.D. degrees in electrical engineering and computer sciences from the University of California, Berkeley, in 1994 and 1997, respectively.

In the summer of 1995, he was at Motorola, in the summer of 1996 at Cadence Design Systems, from 1997 to 2000 at Bell Laboratories Research, from 2000 to 2002 at CeLight (start-up in optical communications), and in the summer of 2002 and in August 2005 at the Research Laboratory for Electronics at the Massachusetts Institute of Technology. Since 2002, he has been with the Department of Electrical and Electronics Engineering, Koc University, Istanbul, Turkey. His recent research interests include computational prototyping of electronic and opto-electronic systems, numerical modeling and analysis, stochastic dynamical systems and their applications, and noise in nonlinear electronic, optical, and communication systems. The work he did at Bell Laboratories and CeLight is the subject of six patents. He has coauthored two books in the areas of nonlinear-noise analysis and analog-design methodologies and published around 40 articles in journals and conferences.

Dr. Demir is the recipient of several best paper awards: the 2002 Best of International Conference on Computer-Aided Design (ICCAD) Award: 20 years of excellence in CAD, the 2003 IEEE/Association for Computing Machinery William J. McCalla ICCAD Best Paper Award, and the 2004 IEEE Circuits and Systems Society Guillemin–Cauer Best Paper Award. In 1991, he was the recipient of the Regents Fellowship and the Eugene-Mona Fay Gee Scholarship from the University of California, Berkeley, and was selected to be an Honorary Fellow of the Scientific and Technological Research Council of Turkey (TUBITAK). In 2003, he was selected by the Turkish Academy of Sciences to receive the Distinguished Young Scientist Award, in 2005, he won a TUBITAK Career Award, and in 2007, he received the TUBITAK Young Scientist Award.