Likelihood-Based Selection Radius Directed Equalizer With Time-Multiplexed Pilot Symbols for Probabilistically Shaped QAM

Gabriele Di Rosa, *Graduate Student Member, IEEE*, and André Richter, *Senior Member, IEEE*

***Abstract*—Probabilistically shaped (PS) quadrature amplitude modulation (QAM) recently established itself as the solution to adopt for state-of-the-art coherent transponders. This in turn has drawn attention to the rework of conventional digital signal processing (DSP) algorithms, which are not optimal for QAM-PS. In this context, we propose a novel pilot-aided equalization technique based on the ubiquitous radius directed equalizer (RDE). Our solution employs both payload and time-multiplexed pilot symbols for the adaptive filters update. The payload symbols are treated on the basis of the likelihood of their correct blind assignment. We show that this approach is at the same time able to reduce the global pilot overhead (POH) required by the DSP chain and to provide greatly improved performance in the tracking ability of the equalizer in the event of fast state of polarization rotations. We describe a simulation environment in which we model accurately static and dynamic polarization-related impairments. We test our algorithm over different shaped modulation formats and in the presence of transceiver impairments and demonstrate its superior performance with respect to standard feed-forward implementations in all the scenarios considered.**

***Index Terms*—DSP, probabilistic shaping, channel equalization, polarization demultiplexing, data-aided DSP.**

I. Introduction

DIGITAL coherent transceivers and multilevel complex modulation formats have been leading the progress of optical fiber communications for more than a decade, allowing for a continuous increase of the achievable channel capacity. To obtain this result, efficient digital signal processing (DSP) algorithms for high order quadrature amplitude modulation (QAM) have been developed. Recently, with the actual systems approaching the Shannon limit, probabilistically shaped (PS) modulation established itself as the solution to adopt to further increase the spectral efficiency [1] and provide at the same time extremely flexible rate adaptivity [2]. Commercial coherent transponders exploiting this technology are already a reality and increasing research attention is being drawn to the rework of standard DSP algorithms that have been developed to work optimally for unshaped QAM constellations [3]–[7]. In particular, channel equalization proved to be a particularly challenging operation, with most pioneering works relying on fully data-aided solutions [2] and more practical implementations relying on periodically transmitted training sequences [8].

While this last approach offers reliable operation and modulation format independence, it comes at the cost of a reduced data throughput and limited dynamic channel tracking ability, at least without incurring in considerably higher pilot overhead (POH). In this context, we present a novel pilot-aided algorithm for channel equalization based on the ubiquitous radius directed equalizer (RDE) [9]. Our solution takes advantage of single periodically transmitted quadrature phase shift keying (QPSK) pilot symbols, which are commonly introduced to perform carrier phase recovery (CPR) or at least to mitigate for cycle slips [4], [10]. The reuse of these pilot symbols for equalization has the advantage of reducing the global POH of the DSP. At the same time, our approach exploits the unknown payload symbols. However, these are not simply blindly assigned to a constellation level to update the adaptive filter as for the standard RDE (STD-RDE). In our likelihood-based selection (LBS) RDE, the filter taps are updated upon reception of a sample only if the likelihood of its amplitude to be correctly assigned to a constellation level is sufficiently high. This operation was already proposed in our previous work in the context of fully blind equalization [11].

In this paper, we show that our proposed solution is able to enhance the performance of equalizers that rely only on constant amplitude pilots for the filters update, as the one described in [10], over static and dynamic channels, while preserving the reduced POH if compared to training sequence based equalization. In particular, severe requirements on the equalizer are posed by fast state of polarization (SOP) rotations, which can arise from temperature fluctuations, mechanical vibrations and electrical transients. In this context, the most demanding scenario is represented by overground links, for which it has been reported in laboratory experiments [12] and extensive field trials [13] that optical ground wire cables can be subject to SOP rotation of several Mrad/s as a consequence of lightnings striking in the vicinity of the transmission line. In this scenario, the LBS-RDE shows an impressive reduction in the signal to noise ratio (SNR) penalty needed to preserve error-free transmission upon occurrence of fast SOP rotations.

Manuscript received January 8, 2021; revised April 1, 2021 and June 18, 2021; accepted July 14, 2021. Date of publication July 20, 2021; date of current version October 4, 2021. This work was supported by the EU under H2020-MSCA-ITN-2018, under Grant 814276 (WON), and by the German Ministry of Education and Research under Grant 16KIS0993 (OptiCON). *(Corresponding author: Gabriele Di Rosa)*

The authors are with VPIphotonics GmbH, Berlin 10587, Germany (e-mail: gabriele.di.rosa@vpiphotonics.com; andre.richter@vpiphotonics.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2021.3098220.

Digital Object Identifier 10.1109/JLT.2021.3098220

For the sake of comparison, we also implemented a solution in which the payload symbols are directly processed following the STD-RDE update rule and show that our proposed solution outperforms it in all tested conditions. The different performance of the algorithms are assessed through a channel model consisting of multiple birefringent fiber sections and abrupt fast linear polarization rotation for shaped constellations based on 16, 64 and 256-QAM with different pilot rates. We considered cases characterized by moderate (≈ 12 ps) or high (≈ 60 ps) differential group delay (DGD). Furthermore, the impact of transceiver IQ amplitude and phase imbalances and IQ-skew is assessed. We demonstrate that also in the presence of these impairments our proposed solution outperforms the investigated alternatives for realistic transceiver imbalances.

Finally, we compare the proposed LBS-RDE to a different feedback equalizer structure based on the decision-directed least mean square (DD-LMS) algorithm and we point out the conditions in which our feed-forward solution is beneficial.

The rest of the paper is organized as follows: In Section II, conventional equalization schemes are reviewed and our proposed improved algorithm is described. In Section III, the transmission system used to test the equalizer is detailed and, in Section IV, the performance of our proposed algorithm is assessed and compared to conventional approaches. Section V provides the conclusions of this work.

II. Algorithm Description

In this section, we describe in detail the proposed algorithm and the ones we implemented for comparison. We start by introducing the standard constant modulus algorithm (CMA) and the STD-RDE in Section II-A. We then describe the RDE probability aware (PA) implementation [3] in Section II-B. In Section II-C, the LBS concept is introduced and its benefits with respect to the update rules of the other equalizers are discussed. In Section II-D, the implementation of the standard amplitude-based equalizers is described, while the proposed LBS-RDE is detailed in II-E. Finally, in Section II-F, the DD-LMS equalizer is introduced and its differences compared to the other algorithms considered are discussed.

A. Constant Modulus and Radius Directed Equalizer

CMA and STD-RDE are standard solutions for time-domain polarization demultiplexing and dynamic channel equalization when employing QPSK and multilevel QAM signals, respectively. The equalizer can be implemented as four N tap finite impulse response (FIR) filters organized in a butterfly structure [14]. The output of the equalizer in the two orthogonal polarizations can be obtained sample by sample as:

$$\begin{aligned} x_{out}[n] &= \mathbf{h}_{xx}^{H}[n]\mathbf{x}_{in}[n] + \mathbf{h}_{xy}^{H}[n]\mathbf{y}_{in}[n] \\ y_{out}[n] &= \mathbf{h}_{yx}^{H}[n]\mathbf{x}_{in}[n] + \mathbf{h}_{yy}^{H}[n]\mathbf{y}_{in}[n], \end{aligned} \tag{1}$$

where $\mathbf{h}_{xx}$, $\mathbf{h}_{xy}$, $\mathbf{h}_{yx}$, $\mathbf{h}_{yy}$ are vectors containing the N tap weights and $\mathbf{x}_{in}[n]$, $\mathbf{y}_{in}[n]$ are N input samples with index ranging from $n-N+1$ to n. The dynamic optimization of the FIR filters taps is performed by means of a stochastic gradient descent algorithm [14] for which the weights are updated as:

$$\begin{aligned} \mathbf{h}_{xx}[n+1] &= \mathbf{h}_{xx}[n] + \mu\epsilon_x[n]x_{out}^{*}[n]\mathbf{x}_{in}[n] \\ \mathbf{h}_{xy}[n+1] &= \mathbf{h}_{xy}[n] + \mu\epsilon_x[n]x_{out}^{*}[n]\mathbf{y}_{in}[n] \\ \mathbf{h}_{yx}[n+1] &= \mathbf{h}_{yx}[n] + \mu\epsilon_y[n]y_{out}^{*}[n]\mathbf{x}_{in}[n] \\ \mathbf{h}_{yy}[n+1] &= \mathbf{h}_{yy}[n] + \mu\epsilon_y[n]y_{out}^{*}[n]\mathbf{y}_{in}[n], \end{aligned} \tag{2}$$

where μ is the adaptation speed and $\epsilon_{x,y}$ are the error functions for X and Y polarization. For a MQAM constellation with K possible amplitude rings these error functions can be calculated as [15]:

$$\begin{aligned} \epsilon_x[n] &= R_k^2[n] - |x_{out}[n]|^2 \\ \epsilon_y[n] &= R_k^2[n] - |y_{out}[n]|^2. \end{aligned} \tag{3}$$

For the STD-RDE $R_k[n]$ is the constellation radius that is closest to the received n-th sample and depends on the transmitted MQAM constellation points $A_m,\ m = 1, \ldots, M$ and the possible amplitude levels that they can assume $R_k,\ k = 1, \ldots, K$. On the contrary, for the CMA it takes a single constant value $R_k[n] = R_0 = \sqrt{E[|A_m|^4]/E[|A_m|^2]}$. Although the CMA can be used for multilevel constellations, in this situation the STD-RDE offers the advantage of an optimal error criterion. This is obtained through an additional stage prior to the error calculation, in which the received sample has to be assigned to the closest constellation amplitude radius based on the Euclidean distance between the received sample amplitude A and the transmitted levels R_k.

B. Probability Aware Decision Regions

The Euclidean distance metric has been shown to be particularly suboptimal to assign the received samples amplitude in two conditions: i) when the constellation amplitude levels have highly uneven transmission probabilities, as for MQAM-PS, and ii) for low SNR at the receiver, which is a common requirement nowadays in high capacity links characterized by the extensive use of powerful forward error correction (FEC) codes. Because of this, optimized decision regions have been proposed [3] and applied to different stages of the DSP chain at the receiver side [3], [4], [7], [11]. In presence of additive white Gaussian noise (AWGN) with variance σ^2 it is in fact possible to calculate the likelihood of a received symbol with amplitude A to belong to the ring R_k as a Rician distribution:

$$f(A \mid R_k) = \frac{A}{\sigma^2} \exp\left[\frac{-(A^2 + R_k^2)}{2\sigma^2}\right] I_0\left(\frac{AR_k}{\sigma^2}\right), \tag{4}$$

where $I_0(\cdot)$ is the 0^{th} order modified Bessel function. At this point we can use the knowledge of the constellation rings transmission probability distribution functions (PDF) $f(R_k)$. By means of this quantity, we can obtain the optimum discrimination by maximizing the probability that the received symbol with amplitude A belongs to the k-th constellation ring R_k among all possible K amplitude levels:

$$\max_{k \in 1,\ldots,K} \left(f(R_k \mid A)\right) \iff \max_{k} \left(f(R_k) f(A \mid R_k)\right), \tag{5}$$

where the logical equality follows from Bayes theorem. From (4) and (5) it follows that we have to find k such that:

$$\max_{k\in 1,\ldots,K}\left(f(R_k)\frac{A}{\sigma^2}\exp\left[\frac{-(A^2+R_k^2)}{2\sigma^2}\right]I_0\left(\frac{AR_k}{\sigma^2}\right)\right). \quad (6)$$

Nevertheless, because of the extremely rapid growth of the modified Bessel function, from a numerical point of view to preserve accuracy and avoid overflow it is more convenient to calculate $\hat{I}_n(x) = I_n(x)\exp(-|x|)$ and to maximize over the natural logarithm of the expression. We have then the equivalent condition:

$$\max_{k\in 1,\ldots,K}\left(\ln\left(f(R_k)\right) - \frac{(A-R_k)^2}{2\sigma^2} + \hat{I}_0\left(\frac{AR_k}{\sigma^2}\right)\right). \quad (7)$$

Moreover, it is possible to calculate the average values of the Rician distributions and set them as optimized radii $\hat{R}_k$ over which the equalizer error function is calculated:

$$\hat{R}_k = \sigma\sqrt{\frac{\pi}{2}}\mathrm{L}_{1/2}\left(-\frac{R_k^2}{2\sigma^2}\right), \quad (8)$$

where $\mathrm{L}_{1/2}(x)$ is a Laguerre polynomial and can be expressed as:

$$\mathrm{L}_{1/2}(x) = \exp\left(\frac{x}{2}\right)\left[(1-x)\,I_0\left(-\frac{x}{2}\right) - xI_1\left(-\frac{x}{2}\right)\right]. \quad (9)$$

Also in this case the modified Bessel functions require the normalization previously introduced. A numerically accurate calculation of the $\hat{R}_k$ values can be performed as:

$$\hat{R}_k = \sigma\sqrt{\frac{\pi}{2}}\exp\left(\ln\left[(1+R_\sigma)\,\hat{I}_0\left(\frac{R_\sigma}{2}\right) + R_\sigma\hat{I}_1\left(\frac{R_\sigma}{2}\right)\right]\right), \quad (10)$$

where $R_\sigma = R_k^2/2\sigma^2$. For sufficiently high SNR, the Rician distribution tends to degenerate into a Gaussian distribution and $\hat{R}_k$ approaches R_k. In Fig 1(a-c) this behavior is easy to observe for a 16QAM based PS constellation with an entropy of 3 bit/symbol. With SNR $= 5$ dB the means of the Rician distributions $\hat{R}_k$ clearly differentiate from the transmitted amplitude levels R_k (Fig. 1(a)), while already for SNR $= 10$ dB this difference is greatly reduced (Fig. 1(b)). Nevertheless, also for this higher SNR the strongly uneven amplitude levels distribution leads to a clear variation in the optimum decision regions. It is only for an extremely high SNR $= 20$ dB that both, average values and decision boundaries, converge to the conventional ones (Fig. 1(c)). It is clear then that in most realistic working conditions probability aware (PA) decisions are necessary to be implemented to approach the optimum solution when employing MQAM-PS constellations.

Finally, in Fig. 1(d) the same analysis is carried out for unshaped 16QAM. This case is investigated since also unshaped MQAM signals do not possess equiprobable amplitude distributions, due to the variable number of symbols belonging to the different amplitude rings. For this simulation, we set SNR $=$ 14.7 dB in order to achieve the same normalized generalized mutual information (NGMI) as for the case of Fig. 1(b). This allows us to fairly compare modulation formats with different entropy. In fact, by comparing Fig. 1(b) and (d) we observe

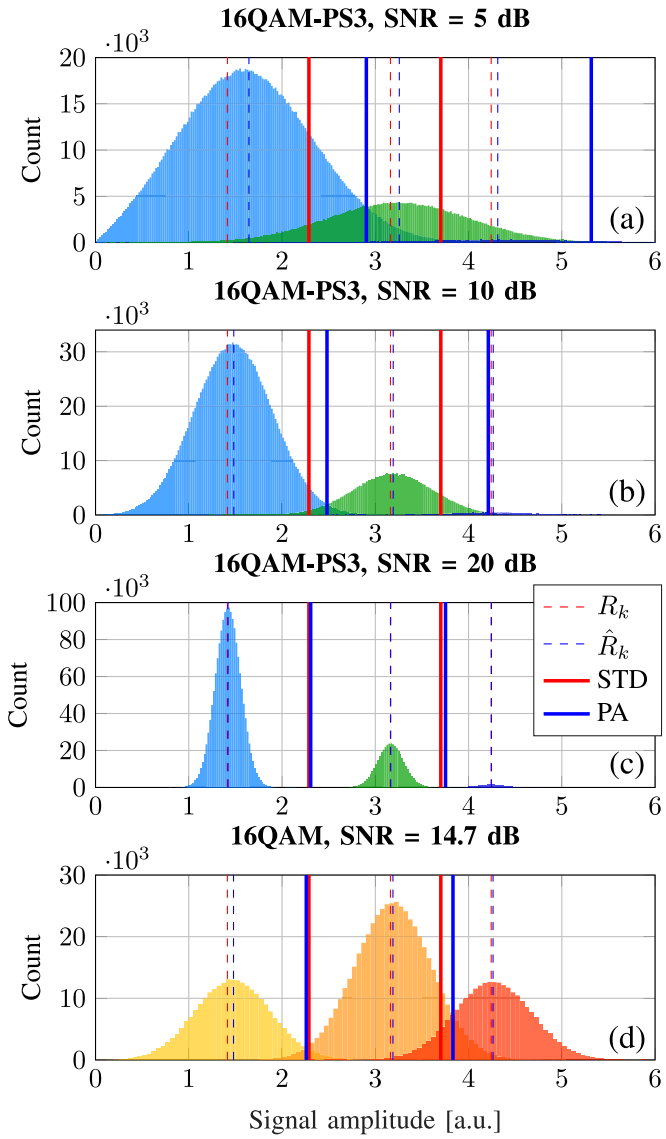


Fig. 1. Radii (dashed lines) and decision regions thresholds (solid lines) for standard Euclidean distance based (STD) and probability aware (PA) decisions for a MQAM-PS constellation with an entropy of 3 bit/symbol (16QAM-PS3) with SNR $= 5$ dB (a), 10 dB (b) and 20 dB (c) and for unshaped 16QAM with SNR $= 14.7$ dB (d).

that the optimized radii, which are dependent only on the noise variance, have very similar values. On the contrary, we observe that the optimized decision thresholds are much closer to the standard ones in the unshaped case. This behavior is due to the amplitude transmission probabilities being much more uneven in the 16QAM-PS case. We conclude that the optimized decision regions lead to improved amplitude assignments also for unshaped MQAM, but with a reduced impact with respect to MQAM-PS.

C. *Likelihood Based Selection: Concept and Benefits*

The concepts introduced in the previous section have first been applied to a fully blind RDE algorithm by simply replacing the conventional radii and decision thresholds with the optimized

ones just introduced [3]. In our previous work, we exploited further the knowledge of the received signal amplitude distribution by introducing the LBS concept [11]: In the STD-RDE all the received samples are always assigned to the closest radius and used for the filter taps update. On the contrary, for the LBS-RDE we consider the relative likelihood of any received amplitude level to be correctly assigned to the determined radius. This likelihood, which we denote as α, can be obtained as:

$$\alpha(A) = \max_{k\in 1,\dots,K} \frac{f(R_k)f(A \mid R_k)}{\sum_{k=1}^{K} f(R_k)f(A \mid R_k)}. \tag{11}$$

To calculate with good accuracy this quantity we can once again repeat the process performed to derive (7) and calculate the single terms of the equation as:

$$f(R_k)f(A \mid R_k) = \exp\left[\ln\left(f(R_k)\right) + \ln\left(\frac{A}{\sigma^2}\right) + \right.$$
$$\left. - \frac{(A-R_k)^2}{2\sigma^2} + \hat{I}_0\left(\frac{AR_k}{\sigma^2}\right)\right], \; k = 1,\dots,K. \tag{12}$$

The likelihood function $\alpha(A)$ provides information about how much the received sample can be trusted to lead to a correct update. In general, we can calculate it for each received sample and adapt consequently the update speed, but the rather complex calculations involved could make it unfeasible or excessively resource hungry for a real-time application. To simplify the implementation we propose to define a threshold value α_{th} and to perform the filters update only for received samples with amplitude $A[n]$ such that $\alpha(A[n]) > \alpha_{th}$. This can be efficiently implemented by optimizing a priori α_{th} and storing the redefined decision regions as a function of signal entropy and SNR in a look-up table, similarly to what has been already proposed in related works [3], [4]. In Fig. 2 an example for a 64QAM based PS constellations with entropy $H = 4$ bit/symbol is shown for two different SNR values. Here we set as an example $\alpha_{th} = 0.8$ in both cases and highlight the resulting decision regions. First, we can notice that the regions are not anymore continuous as for the STD-RDE. On the contrary, received amplitudes that fall outside the highlighted areas are discarded for the filters update. To include this difference in the filters update rule (2), we can simply redefine the filter adaptation speed as $\hat{\mu} = \mu u(\alpha(A) - \alpha_{th})$, where $u(x)$ is the Heaviside step function. Moreover, we can appreciate how the transmission probability and the accumulated noise shape the likelihood function: The inner rings, characterized by an higher number of occurrences, show more prominent peaks in the likelihood to be correctly assigned also for the lower SNR value shown, where the increased noise variance leads in general to a strong decrease of $\alpha(A)$. One last factor to take into account is the spacing between neighboring amplitude levels: it is in fact clear from both figures that the amplitudes located in the middle of the signal range, where the levels are more closely spaced, present much lower likelihood values. To understand the benefit of PA decisions and the LBS concerning the adaptive filters update, we calculate the amplitude assignment error rate, which represent the probability to perform a wrong update of the filters taps. In Fig. 3 this quantity is shown for the same SNR values as in

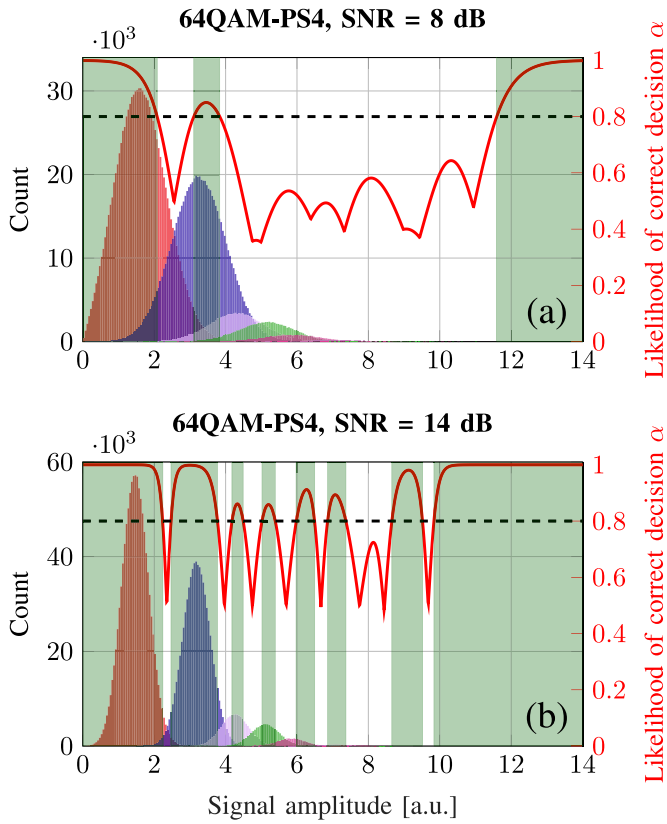


Fig. 2. Signal amplitude distribution and likelihood of correct assignment function for a 64QAM PS constellation with entropy $H = 4$ bit/symbol (64QAM-PS4) with SNR = 8 dB (a) and 14 dB (b).

Fig. 2. The results for the LBS are evaluated versus the likelihood threshold α_{th} and are compared to standard Euclidean distance based and PA assignments. At the same time, the probability to update the filter upon reception of a symbol for the LBS is plotted versus α_{th}. We can observe that PA decisions reduce to less than half the amount of erroneous assignments with respect to the conventional approach for both SNR values tested, going from $\approx 4.1\%$ to $\approx 1.6\%$ for SNR $= 14$ dB and from $\approx 23.2\%$ to $\approx 7.7\%$ for SNR $= 8$ dB. Nevertheless, we can notice that beyond a given α_{th} the LBS starts to discard unreliable symbols for the filters update and the error rate drops consequently. In particular, it is possible to observe in both plots a region in which the number of assignment errors drops rapidly below measurable values with our finite simulated sequence. On the other hand, this range is reached only below a given update rate, around 40% for the lower SNR $= 8$ dB (Fig. 3(a)) and $\approx 80\%$ when the SNR is increased to 14 dB. Here a trade-off is needed, since an excessively reduced update rate worsens the equalizer tracking speed and may discard sufficiently good samples. In fact, in the limit case in which the error rate drops below measurable levels, the algorithm is discarding all samples except the ones with the smallest amplitudes. These samples have higher likelihood since they have the highest tranmission probability and they do not have a neighboring ring nearby, but are also more sensitive to noise, which strongly impairs the equalization performance. Moreover, the equalizer has to operate on a signal impaired not

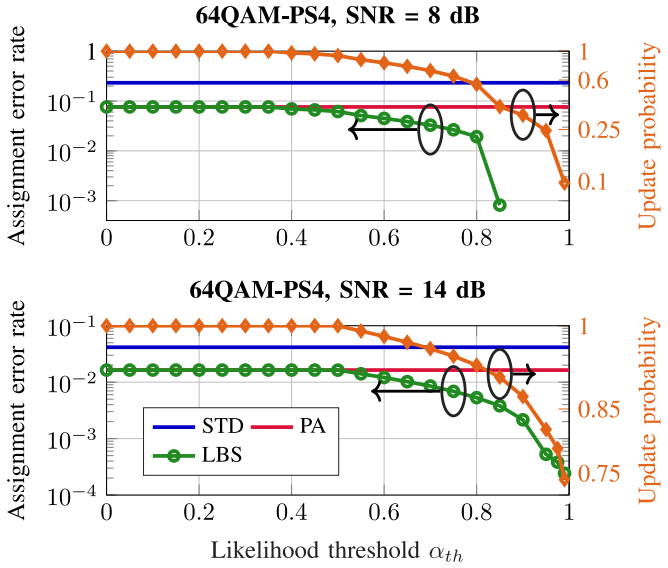


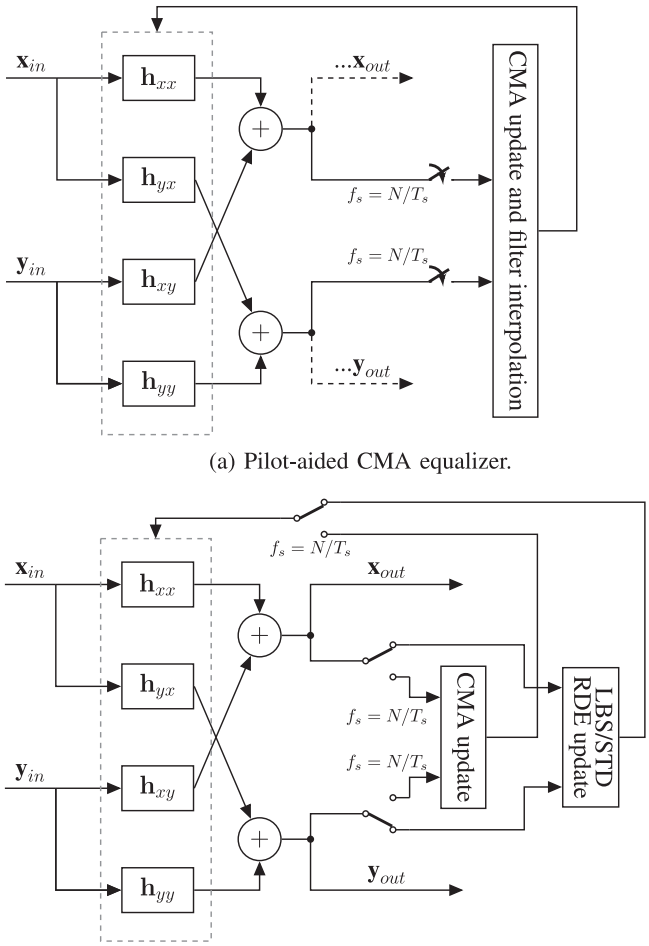
Fig. 3. Assignment error rate for Euclidean distance based (STD), probability aware (PA) and likelihood based selection (LBS) decisions. (a): SNR = 8 dB, (b): SNR = 14 dB.

only by AWGN noise but also by polarization related effects, such as polarization mode dispersion (PMD) and dynamic SOP rotations, and other deterministic impairments. This makes the optimization of α_{th} for given requirements a non-trivial task in which both, an excessively sporadic or biased update and a large assignment error rate have to be avoided.

D. Pilot-Aided Channel Equalization

The solutions presented in Section II-B and II-C have been applied in the context of channel equalization to enhance the performance of a fully blind STD-RDE [3], [11]. Blind operation benefits from the absence of pilot symbols by achieving higher data throughput in the same conditions, but does not guarantee the robustness and operative range of pilot-based equalizers. In particular, these algorithm can ultimately fail in scenarios characterized by large PMD, noise and/or dynamic SOP rotations.

Because of this, several pilot-aided solutions have been proposed, which commonly rely on transmitting periodically a short known data sequence with good spectral properties as for example constant amplitude zero autocorrelation (CAZAC) sequences [16]. This choice provides important benefits such as avoiding iterative processes for the equalizer to converge and modulation format independence, but at the same time introduces a POH that rapidly increases with the channel impulse response length and the required tracking speed [17]. A different approach is to use single time-multiplexed QPSK pilot symbols and to perform equalization by means of the CMA over these known data points and interpolate the filters taps between them [10]. This solution allows to drastically reduce the global POH of the DSP chain since the pilots can be shared with the data-aided carrier phase recovery (CPR) for a sufficiently high POH, or can support blind CPR schemes by the mitigation of cycle slips, for which POH$< 1\%$ have been shown to be effective [4].



(a) Pilot-aided CMA equalizer.

(b) Proposed pilot-aided LBS-RDE and STD-RDE.

Fig. 4. Schematics of the implemented equalizers. The switches change state for one symbol duration T_s every N received symbols, where N is the spacing between two consecutive pilots.

In this work, we implement this solution and compare the pilot-aided CMA approach just discussed to our proposed equalizer, which will be now detailed in Section II-E. The block diagram for the pilot-aided CMA equalizer implementation is shown in Fig. 4(a). The switches in the schematic change state for one symbol period T_s every N received symbols, where $N = 1/\text{POH}$ is the spacing between two consecutive pilots. It is finally necessary to remark that in Fig. 4(a) only the filters update process is graphically described. A realistic hardware implementation for the filter interpolation and the subsequent filtering operation is instead not discussed.

E. Proposed Pilot-Aided LBS-RDE

The pilot-aided CMA equalizer introduced in Section II-D has two main drawbacks: (i) the algorithm has an evident penalty for low POH, due to the filter linear interpolation that under this condition deviates consistently from the ideal equalizer response and (ii) the seldom filters update reduces the equalizer tracking speed over dynamic channels characterized by rapid SOP rotations. To try to eliminate these weaknesses, we propose to avoid discarding the payload symbols between consecutive

pilots for the filters update but instead to operate blindly over them. For this purpose, we perform the blind update by means of the LBS update rule, which we introduced in Section II-C, or simply by using the STD-RDE update. In Fig. 4(b) the block diagram of the proposed equalizer is shown.

Although the convergence limits of blind equalizers have been already pointed out, we find that the problem can be solved by using a sufficient number of interleaved pilot symbols, as it will be shown in Section IV. In particular, we propose the LBS update as the optimum solution while we implement the STD-RDE update only for the sake of performance comparison. We have shown in Fig. 3 that the LBS update is able to strongly reduce the number of blind assignment errors after convergence while using a considerable fraction of the payload for the filters update, thus providing a combination of accurate equalization and good tracking speed. Moreover, the LBS is crucial also during the convergence of the equalizer. In fact, the ability to discard most of the wrong assignments while being supported by the pilot symbols allows to achieve convergence also in extremely noisy conditions and with a reduced POH as it will be demonstrated in Section IV.

Since all the equalizers implemented for this work are pilot-aided, for simplicity we will from now on refer to the three tested equalizers as constant modulus equalizer (CME), LBS-RDE and STD-RDE. The CME is the implementation shown in Fig. 4(a), which relies only on pilot symbols for the filters update and on interpolation. LBS-RDE and STD-RDE are instead as from Fig. 4(b) and exploit pilot and payload symbols.

F. Decision-Directed Least Mean Square Equalizer

Until now we focused only on amplitude-based feed-forward solutions based on the CMA and RDE algorithms. A different approach that can operate blindly on the payload data and guarantees an optimal error criterion is to use the DD-LMS equalizer. In this case, the error function used for the filters update becomes [18]:

$$\begin{aligned} \epsilon_x[n] &= (d_x[n] - \exp(-j\phi)x_{out}[n])\exp(j\phi) \\ \epsilon_y[n] &= (d_y[n] - \exp(-j\phi)y_{out}[n])\exp(j\phi), \end{aligned} \tag{13}$$

where ϕ is the estimated phase and $d_{x,y}$ is the hard symbol decision performed on the equalizer output after phase rotation. This solution is commonly employed after a pre-convergence stage because of its improved steady-state performance compared to amplitude-based adaptation. On the other hand, this approach requires a feedback loop to perform CPR and optionally frequency offset estimation (FOE) to calculate the phase-dependent error function before updating the adaptive filters. Compared to feed-forward solutions based on CMA and RDE, this inherent loop structure results in the impossibility to decouple equalization and phase estimation and introduces a larger feedback delay for the filters update. The delay depends on the specific DSP implementation considered and harms the equalizer tracking ability over dynamic channels [18]. In the most critical cases, a fast SOP rotation can lead to positive feedback between the degradation of phase estimation and equalization, failing to recover the signal [19].

In Section IV-D, we compare the performance of the proposed LBS-RDE and the DD-LMS equalizer with different loop delays. In this way, we are able to study the trade-off between improved steady-state equalization performance and increased filter update delay in different system scenarios. The DD-LMS equalizer implemented includes a CME pre-convergence stage and in-loop phase estimation using a pilot-based maximum-likelihood CPR. This is implemented as a moving average filter whose averaging length is optimized to maximize the NGMI and with linear phase interpolation performed between phase estimates.

III. Investigated Transmission System

In this section, we describe in detail the simulation setup that we designed to assess the performance of the proposed equalizer. We focus in Section III-A on the generation of the MQAM-PS signals and the assessment of the system performance. Then, we describe in Section III-B the optical system with particular attention to the channel model used to reproduce accurately polarization-related impairments caused by PMD, extremely fast SOP transients, and their mutual interaction. Finally, in Section III-C, we provide an overview of the DSP chain used at the receiver before the performance estimation. The investigated transmission system was simulated in VPIphotonics Design Suite 11.0 and is depicted in Fig. 5.

A. Signal Generation and Performance Metric for MQAM-PS

The MQAM-PS signals are generated by probabilistic amplitude shaping (PAS). The I and Q components of the transmitted data frame, whose length is set to 2^{17} symbols, are randomly drawn from a Maxwell-Boltzmann distribution and combined to generate the transmitted complex sequence. This distribution of the amplitude levels has been shown to be optimal for the AWGN channel [20] and is described as:

$$f(A_i) = \frac{\exp\left(-\lambda A_i^2\right)}{\sum\limits_{A_i \in A_{I,Q}} \exp\left(-\lambda A_i^2\right)}, \tag{14}$$

where $A_{I,Q}$ is the set of amplitude values that the I and Q components of the constellation can assume and λ is a factor that is optimized in order to achieve the desired signal entropy $H = -\sum_{A_i \in A_{I,Q}} f(A_i)\log_2[f(A_i)]$.

For our analysis, we generate shaped signals based on 16QAM with $H = 3$ bit/symbol (16QAM-PS3), 64QAM with $H = 4$ bit/symbol (64QAM-PS4), 5 bit/symbol (64QAM-PS5) and 256QAM with $H = 6$ bit/symbol (256QAM-PS6).

Finally, to assess the performance of the MQAM-PS signals we calculate the generalized mutual information (GMI) from which we obtain the NGMI as in [21]:

$$\text{NGMI} = 1 - [H - \text{GMI}] \,/\, \log_2(M). \tag{15}$$

This figure of merit is chosen since it has been shown to provide the most accurate post-FEC bit error rate (BER) prediction irrespective of the modulation format [22].

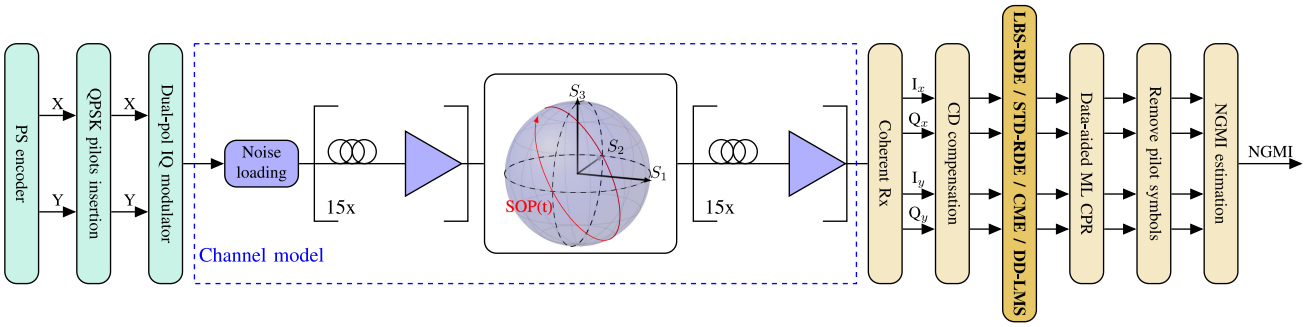


Fig. 5. Transmission system investigated to evaluate the performance of the implemented equalizers.

B. Optical System and Channel Model

The polarization multiplexed MQAM-PS signals are generated at a symbol rate of 32 Gbaud and are first periodically interleaved with QPSK pilots every N symbols. The signal is then pulse shaped by means of a raised cosine filter with 0.1 roll-off factor before being modulated around $f_c = 193.1$ THz. The pilot rate is set to either 1/32 or 1/128 in order to simulate conditions characterized by a relatively high POH $= 3.125\%$, which can be at the same time used to perform data-aided CPR, and a lower POH $< 1\%$. The second case is suitable for conditions in which higher throughput is desired and the pilots act only as a support for cycle slips mitigation for the CPR. The amplitude of the pilot symbols A_p is optimized to maximize the NGMI obtained with the CME over a static linear channel for the different modulation formats used: While in fact higher A_p can improve the performance of the equalizer, at the same time the average energy of the transmitted signal is increased, resulting in a non-trivial trade-off. Also for the LBS-RDE, STD-RDE and DD-LMS equalizer the same A_p is used to allow for a fair comparison.

After the transmitter, the signal is noise loaded to set a given SNR before entering the fiber model. Propagation takes place over 30 spans of 100 km of standard single mode fiber (SSMF) with dispersion $D = 16$ ps/(nm $\cdot$ km). Polarization effects in the fiber are modeled through the established coarse step method, which approximates continuous variations of birefringence by series of many short polarization sections with constant birefringence and random SOP rotation at the sections boundaries. The length of the scattering section is set in our setup to 1 km. At this point we define two deterministic birefringence profiles with different average PMD for which we estimate the average accumulated DGD over the signal bandwidth by considering the Jones Matrix of the channel. The two profiles provide respectively DGD $= 12$ ps and DGD $= 60$ ps. In this way, we are able to simulate conditions characterized by moderate and large DGD, while performing a fair comparison among the different equalizers in repeatable conditions. Moreover, we introduce in the middle of the transmission line, after 15 fiber spans, a lumped linear SOP rotation. With this additional block we want to emulate the impact of events such as lightning strikes in the vicinity of the fiber on the signal SOP as it has been reported from laboratory experiments [12] and field trials [13]. The dynamic SOP rotation is implemented by transforming the signal according to the Jones matrix:

$$\mathbf{J}(t) = \begin{bmatrix} \cos(\omega_r t) & \sin(\omega_r t) \\ -\sin(\omega_r t) & \cos(\omega_r t) \end{bmatrix}, \tag{16}$$

where ω_r is the SOP rotation rate in rad/s. It must be remarked that fiber nonlinearities are not simulated and inline amplification is assumed noiseless. These choices are made in order to be able to set a specific SNR and to assess the performance of the equalizer without being affected by additional impairments arising from the nonlinear behavior of the channel, such as uncompensated nonlinear phase noise.

Finally, at the receiver after photo-detection the I and Q components pass through low-pass filters with cutoff frequency $f_{-3\,\mathrm{dB}} = 32$ GHz and are digitized at 2 sample/symbol before entering the DSP chain.

C. Digital Signal Processing

The simple DSP chain has the purpose to support the equalizers under test, which are the focus of this work. After digital to analog conversion, chromatic dispersion (CD) is fully compensated. Then the signal enters the equalization stage, which can consist of the CME, the proposed LBS-RDE, the STD-RDE, or the DD-LMS equalizer. The number of taps of the adaptive filters is set to 25, which proved to be sufficient in all tested conditions. At the output of the equalizer the signal is down-sampled to 1 sample/symbol before passing through a fully data-aided maximum-likelihood CPR, which is implemented as a moving average filter with a window length of 51 symbols. Finally, the pilot symbols are removed from the stream before calculating the NGMI over the received samples.

IV. Performance Assessment

In this section, we first evaluate the performance of the proposed LBS-RDE versus the equivalent implementation using the STD-RDE update rule and the CME described in Section II-D. In Section IV-A, the comparison is performed for transmission over a static channel while in Section IV-B the signal is also subject to a local fast SOP rotation in the middle of the link, as illustrated in Fig. 5. In Section IV-C, the robustness of the tested equalizers in the presence of transceiver impairments is investigated. Finally,

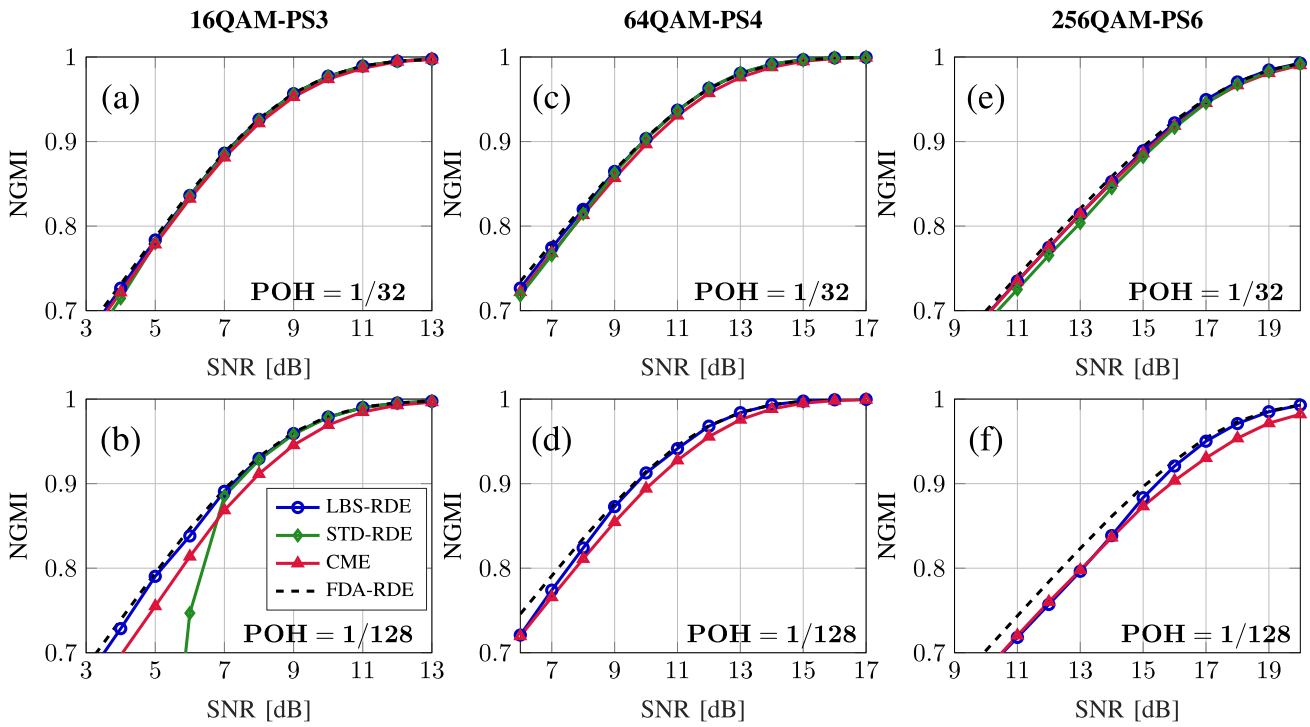


Fig. 6. Performance in static conditions: NGMI versus SNR for the equalizers under test for the different shaped modulation format tested and POH of 1/32 and 1/128.

in Section IV-D, the performance of the proposed LBS-RDE is compared to the one of the feedback-based DD-LMS equalizer.

A. *Performance in Static Conditions*

The first analysis that we performed concerns transmission over a static channel with AWGN and fiber birefringence only. We have DGD $= 60$ ps and sweep the SNR in order to observe the range $0.7 \leq \text{NGMI} < 1$. We then compare the results obtained with the performance of a fully data-aided implementation of the conventional RDE (FDA-RDE). This represents a limit best-case condition in which the algorithm has the knowledge of the whole transmitted sequence and never performs assignment errors. All equalizers are insured to reach convergence by iterating over 20 repetitions of the transmitted sequence and the performance are calculated only over the last block of 2^{17} symbols. The adaptation speed μ is separately optimized for the different algorithms for each tested SNR value. The same operation is performed for the LBS-RDE also for the additional parameter α_{th}.

The results of this analysis are shown in Fig. 6, where we can observe the NGMI for the different modulation formats tested and for a pilot overhead of $1/32$ and $1/128$ versus the SNR. From Fig. 6(a,b) we notice that for the analyzed 16QAM based PS signal all the equalizers perform equally well with the higher POH $= 1/32$ while the situation drastically changes for POH $= 1/128$. Here in fact, only the LBS-RDE is able to approach the best-case performance of the FDA equalizer over the whole SNR range. The STD-RDE, on the contrary, fails for SNR < 7 dB but approaches the performance of the LBS-RDE with higher SNR values. This is due to the seldom pilot-aided update of the filters taps, which results into poor convergence when misled by the large number of blind assignment errors characterizing this extremely noisy scenario. Finally, looking at the CME we clearly notice a performance penalty over most of the analyzed range. This arises from the interpolation operation, which does not represent accurately the time-domain behavior of the signal when the interpolation step becomes larger because of the reduced pilot rate.

Moving to Fig. 6(c,d) we see results for the 64QAM-PS constellation. Again, when transmitting the pilots every 32 symbols we observe the same behavior as for 16QAM-PS3. On the contrary, some differences arise for POH $= 1/128$: In this case the STD-RDE performance is in fact not plotted since the algorithm does not provide a reliable convergence over most of the range, making it unsuitable for this application. Coming to the LBS-RDE and the CME results, we observe again a performance gap over most of the SNR range, with the LBS-RDE outperforming the CME and approaching the ideal FDA-RDE performance. However, focusing on the lowest SNR simulated we notice a deterioration of the performance of the LBS-RDE, which approaches the one of the CME. In this condition the strong noise applied over the higher order 64QAM-PS constellation, which is characterized by reduced symbol spacing, make blind adaptation challenging, even when using the LBS concept. Nevertheless, our proposed solution still outperforms the CME for NGMI ≥ 0.8.

Finally, we move to Fig. 6(e, f) and our results for the 256QAM-PS constellation. As expected, to achieve a performance improvement through blind equalization becomes more challenging with the increasing modulation order and the resulting reduction in symbol spacing. We observe that the STD-RDE

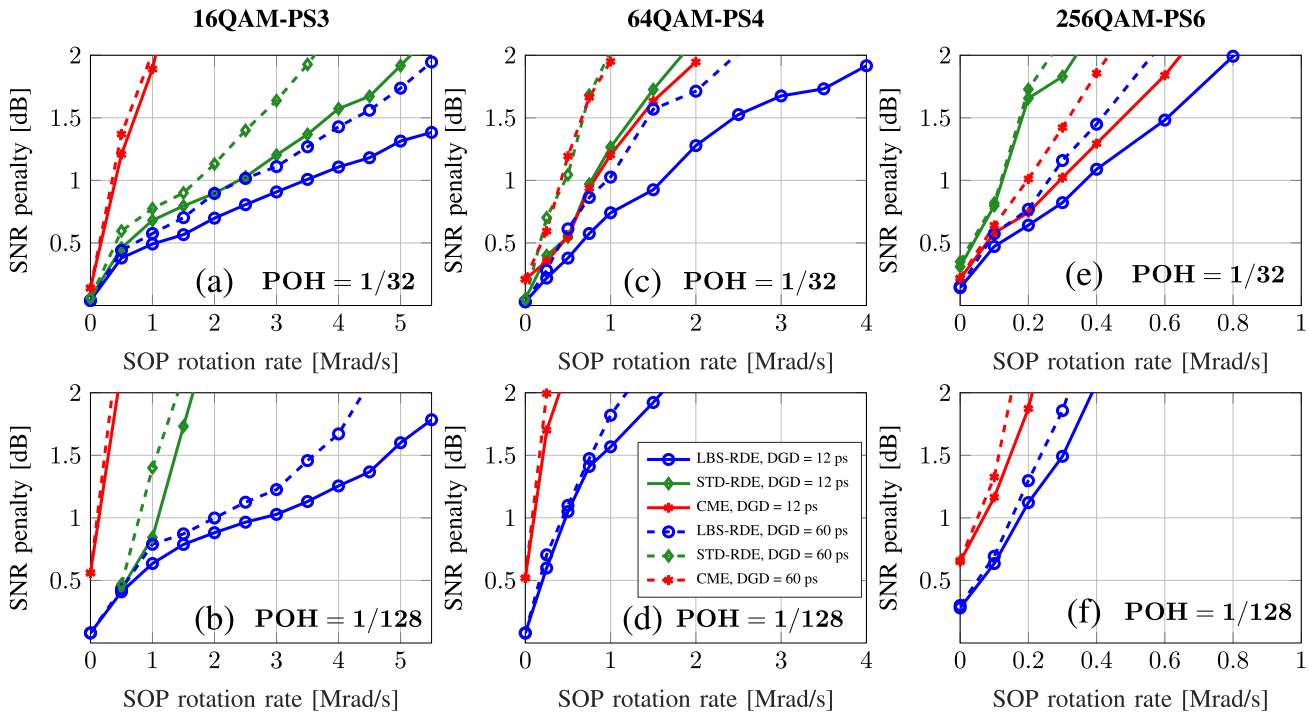


Fig. 7. Performance in dynamic SOP conditions: SNR penalty with respect to ideal fully data-aided equalization over a static channel to preserve NGMI > 0.88. The reference SNR values for the different cases are: (a) 6.84 dB, (b) 6.71 dB, (c) 9.35 dB, (d) 9.10 dB, (e) 14.61 dB, (f) 14.52 dB.

is not suitable for POH $= 1/128$ and also shows a penalty with respect to the CME for POH $= 1/32$. On the other side, the LBS-RDE keeps outperforming the CME for the lowest pilot overhead but on a reduced range with respect to lower cardinality constellations. We observe a visible gain in fact only for NGMI ≥ 0.9.

We conclude that the LBS-RDE is able to significantly improve the performance of a purely pilot-based equalizer when a low overhead is desired in most cases for 16/64QAM based constellations and for 256QAM when also a limited FEC overhead is used.

B. Performance in Dynamic SOP Conditions

We now proceed to analyze the performance of the equalizers over the channel model depicted in Fig. 5, in which at the middle of the transmission line the signal SOP is subject to a rapid linear rotation at a given rate ω_r. To simulate this transient behavior, first the signal is transmitted over a static channel until the equalizers reach convergence. At this point, the signal is passed through the same channel but with the inclusion of the SOP rotation stage and the NGMI is calculated over this last data frame. For the performance evaluation we calculate the SNR penalty that is needed to preserve NGMI > 0.88 compared to the ideal FDA-RDE results in static conditions versus the SOP rotation rate ω_r. This NGMI threshold allows for a post-FEC BER of 10^{-15} by using a low-density parity-check code with code rate (R_c) 0.86 concatenated with an outer staircase code with 6.25% overhead, thus obtaining a global $R_c = 0.81$ [22]. The study is performed for DGD $= 12$ ps and DGD $= 60$ ps. The adaptation speed μ is separately optimized for the different equalizers and SNR values. This optimization guarantees that with a given μ and received SNR the algorithm is able to provide NGMI > 0.88 in static conditions while at the same time maximizing the performance when subject to the SOP rotation.

In Fig. 7 the results of this analysis are shown. We start from Fig. 7(a), where we observe that both solutions exploiting the payload for the filters update vastly outperform the CME. For fast dynamic conditions, in fact, even an update rate of 1/32 proves to be not sufficient to preserve the performance. We see that the CME has already an SNR penalty of 2 dB for a maximum tolerated $\omega_r = 1$ Mrad/s. For what concerns the other two solutions, once again the LBS-RDE outperforms the STD-RDE in all conditions. For DGD $= 12$ ps in particular, the LBS-RDE has an SNR penalty < 1.5 dB for rotation rates up to 5.5 Mrad/s, a value that is larger than the fastest transient observed in field trials [13]. We also observe the impact of the interaction between SOP rotation and larger fiber birefringence: all tested algorithm show in fact evident penalties at the same ω_r when larger PMD is present.

Moving to Fig. 7(b) we observe that reducing the overhead has an extreme impact on the CME and the STD-RDE, more than halving the maximum allowed ω_r at an SNR penalty of 2 dB. For this lower POH the LBS-RDE is also subject to a performance drop, though it is much smaller compared to the other two solutions. This is achieved thanks to the improved blind usage of the payload symbols, which is particularly effective in eliminating errors for smaller constellations such as 16QAM and at the same time is able to preserve a large fraction of the received samples (> 85 % in the analyzed case) for the filters update.

For 64QAM-PS4 again the LBS-RDE clearly outperforms the other solutions, although the increased base modulation order makes resistance to very fast SOP rotations > 4 Mrad/s not achievable in the tested range of SNR also for a pilot overhead of 1/32 (Fig. 7(c)). Moreover, additional penalties arise for DGD $=$ 60 ps. The other two algorithms tested have instead opposite behaviors with respect to the previously analyzed situation with 16QAM-PS: the CME performance is in fact improved thanks to the increased constellation energy. This led to an higher optimum amplitude of the pilots, which direct better the equalization. On the contrary, the deterioration of the blind assignment error rate for the STD-RDE makes it perform poorly, bringing no extra benefit to the CME alone.

Looking then at Fig. 7(d)), we see that in these conditions also the LBS-RDE is not able to stand SOP rotations faster than 1 Mrad/s with the reduced POH. However, it still performs extremely well compared to the CME, both in static and dynamic conditions. The STD-RDE is not shown since it is not able to converge even in static conditions for the SNR range analyzed here, as pointed out already in Section IV-A.

In Fig. 7(e,f) we find finally the results for the highest order constellation that we tested: 256QAM-PS6. In general the maximum tolerated SOP rotation rate is strongly reduced for this larger base constellation, being ≈ 800 krad/s even for the LBS-RDE with the highest POH tested for an SNR penalty of 2 dB and with DGD $= 12$ ps. The STD-RDE performance again deteriorates, becoming evidently worse than the CME also for a pilot overhead of 1/32. On the contrary, the LBS-RDE remains clearly the best performing algorithm, although the performance improvement with respect to the CME is reduced when compared to the previous cases analyzed.

C. Performance in Presence of Transceiver IQ-Imbalances

In a real coherent transceiver, we find amplitude and phase imbalances arising from the amplitude mismatch and the quadrature error between the I and Q branches. At the same time, timing differences between the two tributaries result in IQ-skew. Imperfect calibration, device aging and temperature variations can lead to the presence of residual impairments. When left uncompensated, these have a detrimental impact on the DSP chain, in particular for high order modulation formats.

Amplitude and phase imbalance generated at the receiver side can be monitored and mitigated before CD compensation by means of the Gram-Schmidt orthogonalization procedure (GSOP) [23]. At the same time, receiver IQ-skew can be handled together with channel equalization also in the presence of large accumulated CD by modifying the 2x2 equalizer into a 4x2 structure [24]. On the contrary, transmitter impairments can only be compensated after FOE and CPR [23]. This implies that transmitter impairments will strongly affect the signal at the equalization stage also after convergence of the equalizer. The result is a deviation from the AWGN channel assumption, over which our LBS update rule is defined.

To test the impact of these impairments on equalization, we include now in the simulations amplitude imbalance, phase imbalance and IQ-skew at the transmitter side. Assuming linear operation for the modulator, we can describe the baseband signal in the polarization p at the output of the transmitter as:

$$u_{tx,p}(t) = a_{Ip} \exp\left[j\phi_{Ip}\right] I_p(t + \tau_{Ip}) + \\ + j a_{Qp} \exp\left[j\phi_{Qp}\right] Q_p(t + \tau_{Qp}), \tag{17}$$

where $I_p(t), Q_p(t)$ are the ideal transmitted IQ sequences for polarization p. Following from this definition, we can write the amplitude imbalance ξ, the phase imbalance Φ and the IQ-skew $\Delta\tau$ as:

$$\xi = 20 \log_{10} \left(\frac{a_{Qp}}{a_{Ip}} \right), \tag{18}$$

$$\Phi = \phi_{Qp} - \phi_{Ip}, \tag{19}$$

$$\Delta\tau = \tau_{Qp} - \tau_{Ip}. \tag{20}$$

We now test the performance of LBS-RDE, STD-RDE and CME versus the different impairments for 64QAM-PS4 with POH $=$ 1/32 and DGD $= 60$ ps. Simulations are performed in static SOP conditions and for a rotation rate of 1 Mrad/s. In these two cases, we set the SNR respectively to 9.5 dB and 11.5 dB. These values are chosen in order to obtain for all tested algorithms NGMI > 0.88 in absence of transceiver impairments, as obtained from the results of Fig. 6 and Fig. 7. The impairments are applied with the same magnitude for the X and Y polarizations. We verified that the conclusions obtained are qualitatively representative also for the more general case of uneven impairments in the two branches. The results of this analysis are presented in Fig. 8.

In Fig. 8(a-b), the NGMI achieved by means of the different equalizers versus amplitude imbalance is shown: without SOP rotation (Fig. 8(a)) the three equalizers show a similar performance penalty. On the contrary, with 1 Mrad/s SOP rotation (Fig. 8(b)) the STD-RDE shows a clear performance drop for $\xi > 0.6$ dB while the other two solutions present a smoother decrease of the NGMI. At this point, it is important to notice that the performance gap between the LBS-RDE and CME is reduced for strong amplitude imbalance values. This behavior is due to the reduced effectiveness of the LBS update rule when operating significantly outside the AWGN assumption. Nevertheless, the LBS-RDE outperforms the CME over the whole tested range. A clear reduction in performance gain is observed only for large $\xi > 1.5$ dB.

The same analysis is carried out in presence of phase imbalance. Results are shown in Fig. 8(c-d), and similar conclusions as for the amplitude imbalance case are drawn. In static conditions we observe a comparable behavior of the equalizers, and with the additional SOP rotation we have similar results as in Fig. 8(b). A strong performance drop of the STD-RDE for $\Phi > 6^\circ$ is observed, together with a reduced performance gap between LBS-RDE and CME for larger phase imbalances.

Finally, in Fig. 8(e-f), the impact of transmitter IQ-skew is assessed. The qualitative behavior of the equalizers is similar to the other two imbalances previously considered, but in this case we observe a region in which the CME outperforms the STD-RDE in static conditions. Similarly, the CME outperforms the LBS-RDE upon occurrence of the tested 1 Mrad/s SOP rotation,

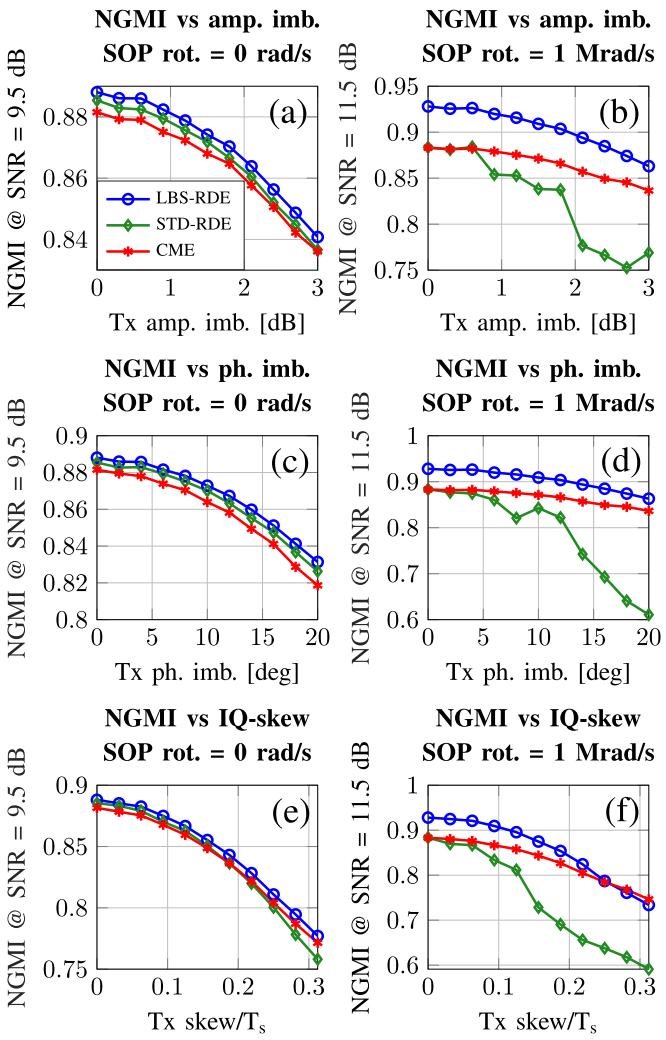


Fig. 8. Performance in the presence of transmitter impairments: NGMI versus amplitude imbalance (a-b), phase imbalance (c-d) and the ratio between IQ-skew and symbol period (e-f).

showing the limits of blind operation of the equalizer with strong transceiver impairments. However, this happens only for a very large $\Delta\tau > 25\%$ of the symbol period.

We conclude that a strong deviation from the AWGN assumption due to transceiver imperfections has an impact on the blind operation of the equalizer, also with the improved LBS update rule. Nevertheless, the impairments values for which this becomes predominant are much larger than what is usually found in standard applications. This demonstrates the effectiveness of our approach also in the presence of residual transceiver impairments.

D. Performance Comparison With Feedback Equalization

After having assessed that the LBS-RDE is the best performing equalizer among the tested feed-forward solutions, we now perform a comparison with the feedback-based DD-LMS equalizer. As it was introduced in Section II-F, the DD-LMS update rule provides the advantage of improved steady-state performance, which results in increased tracking speed and lower output error [25], but at the cost of increased filter update delay and coupling to phase estimation. It is interesting to analyze this trade-off to understand in which conditions the improved feed-forward equalization based on the LBS-RDE outperforms the feedback-based approach. However, to perform this study, some changes in the equalizers implementation previously described have to be made to capture critical limitations:

1) *Parallel block processing:* Until now we considered a sample-by-sample update of the adaptive filters. This best-case condition was simulated to assess the maximum theoretical performance of the tested update algorithms but does not represent a realistic condition. In practical cases, the desired throughput sets a minimum requirement on the equalizer parallelism (P). To emulate this timing limitation, we now process in parallel a block of P samples with filters having the same coefficients, which are then adapted once per block by summing the single $P/2$ updates obtained as in (2). While performing this operation, we want to stress that we are not aiming to simulate a realistic parallel hardware implementation. Our goal is only to include limitations arising from hardware delays in our investigated scenario while keeping the possibility to update the filter over every symbol to obtain best-case achievable performance, consistently with the previous analyses.
2) *Filter update delay:* Until now we assumed instantaneous updates for the adaptive filters. In reality, the operations needed to calculate the filter update take several clock cycles in a circuit implementation. Combined with the parallel block processing, this results in the input signal block $\#N$ being equalized with the filter adapted over the received symbols in block $\#(N - D - 1)$, where D represents the propagation delay inside the feedback loop in clock cycles [18].

In the following analysis, a parallelism $P = 512$ is set to consider efficient implementations satisfying the requirements of current high baud rate systems [18]. For what concerns the update delay, a specific value for D is impossible to be set without referring to a particular hardware implementation. Nevertheless, the delay is larger for feedback equalizers, since they require additional phase estimation (and optionally FOE as in [19]) compared to feed-forward equalization. In this context, the CPR implementation plays a crucial role: While a simple coarse phase estimation may provide lower latency, more accurate algorithms can require a larger amount of sequential operations. In the following, we set as a reference $D = 0$ for the LBS-RDE and $D = 1, 6, 11, 16$ for the DD-LMS equalizer. This allows us to understand the limitations arising from an additional update delay and to observe the conditions in which feed-forward equalization is to be preferred without focusing on a single DSP implementation. To perform this analysis, we simulate static and dynamic SOP conditions using the setup depicted in Fig. 5 and optimize the equalizers parameter with the same procedure performed in Section IV-A and IV-B. The linewidth of transmitter and local oscillator lasers is set to 100 kHz and POH $= 1/32$ is chosen. This overhead allows for an effective, yet realistic, pilot-aided CPR in the DD-LMS feedback. Moreover, it proved

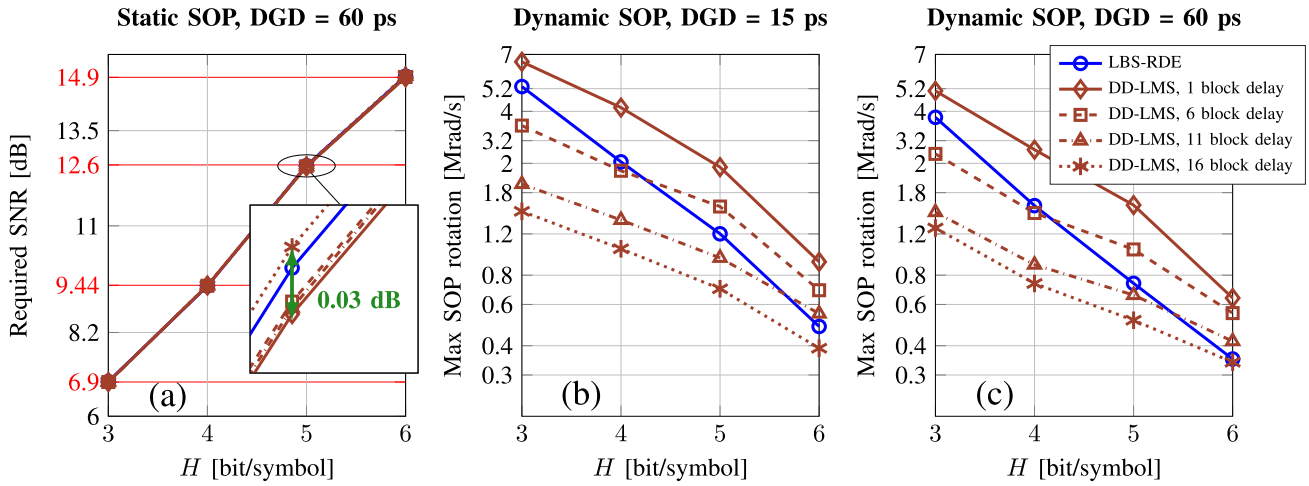


Fig. 9. Performance comparison of LBS-RDE and DD-LMS equalizer. (a): Required SNR to achieve NGMI > 0.88 for variable constellation entropy and DGD $= 60$ ps. (b-c): Maximum SOP rotation tolerated for 1.5 dB SNR penalty to preserve NGMI > 0.88 with DGD $= 15$ ps (b) and 60 ps (c).

to give the best performance for the LBS-RDE with dynamic SOP, making it a case of particular interest for this study. The signal entropy is swept from 3 to 6 bit/symbol, with the base MQAM constellation chosen as the smallest for which it is possible to achieve the desired entropy. The results obtained from this study are reported in Fig. 9. In Fig. 9(a), the SNR values required to achieve NGMI > 0.88 with DGD $= 60$ ps and static SOP versus the signal entropy are shown. From the plots, we notice that the LBS-RDE and the DD-LMS equalizer show extremely similar performance for all simulated entropy values. This demonstrates that in the analyzed scenario both update criteria approach similar output error values in static conditions. Moreover, as expected, the introduction of a delay in the filter update does not have any impact on the performance for static SOP conditions.

More insights can be gained from Fig. 9(b-c). We now consider a dynamic channel model, and find the maximum tolerated SOP rotation for LBS-RDE and DD-LMS equalizer with different feedback delays for a 1.5 dB SNR penalty to the reference SNR values shown in red in Fig. 9(a). We immediately notice that the DD-LMS is the best performing algorithm if an almost instantaneous delay of one block is assumed for the carrier recovery. Nevertheless, the situation changes drastically when longer delays are assumed. While we see that the maximum tolerated SOP rotation by the LBS-RDE drops faster for increasing entropy than for the DD-LMS equalizer, we also observe that the additional loop delay introduces a large penalty on the whole entropy range. Already for a delay of 6 blocks, the DD-LMS equalizer is outperformed by the LBS-RDE for $H \leq 4$ bit/symbol, while a larger delay of 11 blocks makes the LBS-RDE overcome the performance of the feedback-based equalizer with the 64QAM-PS5 signal. Finally, only for an even larger delay of 16 blocks, the DD-LMS equalizer shows worse performance than the LBS-RDE also for the 256QAM-PS6 constellation. These results confirm what was already observed in Fig. 7, with the LBS-RDE being particularly beneficial over fast-varying SOP for smaller constellations with $H \leq 4$ bit/symbol.

We conclude that, while the LBS-RDE was found to be clearly the best solution among feed-forward algorithms, the choice of the DD-LMS or the LBS-RDE to maximize the equalizer performance over a dynamic channel is tightly connected to the actual system scenario and DSP implementation. Moreover, in the analyzed scenario, the LBS-RDE is to be preferred for operation over a sufficiently slowly varying channel if decoupling of equalization and phase estimation is desired. We have shown, in fact, that the two equalizers present identical performance in static conditions.

V. Conclusions

We proposed a novel feed-forward algorithm for polarization demultiplexing and channel equalization of MQAM-PS formats that exploits at the same time periodically transmitted pilot symbols and the unknown payload symbols to enhance the performance of a purely pilot-aided solution. In our new approach, the payload data are taken effectively into account for the filters update by considering the statistics of their amplitude distribution to select only the ones that present a sufficiently high likelihood to lead to a correct filters update. This method was tested over a channel model in which AWGN, fiber birefringence, and dynamic SOP rotations are included. Additionally, the robustness of the proposed approach with respect to transceiver impairments has been verified. We assessed the performance of our proposed algorithm compared to a conventional solution that adapts the equalizer filters only over pilot symbols and to an equivalent one in which the statistics of the received signal are ignored. The observed result is that our LBS-RDE outperforms in all simulated conditions these other two implementations. In particular, in the presence of rapid SOP rotation, our solution shows strongly enhanced performance in all tested scenarios. Moreover, a considerable gain is obtained also for a static channel when a low pilot overhead ($< 1\%$ in our tested scenarios) is used. Finally, the proposed LBS-RDE has been compared also to a feedback equalization scheme based on the DD-LMS algorithm. Although the improved steady-state

performance of the DD-LMS equalizer provides better performance over a dynamic channel in ideal conditions, we have shown that the proposed feed-forward LBS-RDE is beneficial in the presence of considerable additional feedback circuit delay. This behavior is particularly evident already for the shorter delays tested for 16QAM-PS and 64QAM-PS. Moreover, in the analyzed scenario, the two algorithms have proven to perform identically in static conditions. Our results suggest that the LBS-RDE is the best choice for operation over a sufficiently slowly varying channel if decoupling of equalization and phase estimation is desired.

REFERENCES

[1] T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, "On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel," *J. Lightw. Technol.*, vol. 34, no. 21, pp. 5063–5073, 2016.

[2] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *J. Lightw. Technol.*, vol. 34, no. 7, pp. 1599–1609, 2016.

[3] S. Dris *et al.*, "Blind polarization demultiplexing and equalization of probabilistically shaped QAM," in *Proc. Opt. Fiber Commun. Conf.*, 2019, pp. 1–3.

[4] G. D. Rosa and A. Richter, "Low complexity blind carrier phase recovery for probabilistically shaped QAM," *IEEE Photon. Technol. Lett.*, vol. 32, no. 17, pp. 1109–1112, Sep. 2020.

[5] F. A. Barbosa *et al.*, "Phase and frequency recovery algorithms for probabilistically shaped transmission," *J. Lightw. Technol.*, vol. 38, no. 7, pp. 1827–1835, Apr. 2019.

[6] Q. Yan *et al.*, "Blind carrier frequency offset estimation in coherent optical communication systems with probabilistically shaped M-QAM," *J. Lightw. Technol.*, vol. 37, no. 23, pp. 5856–5866, 2019.

[7] Q. Zhang and C. Shu, "Viterbi and viterbi algorithm based phase recovery for probabilistically shaped signals," *J. Lightw. Technol.*, vol. 39, no. 5, pp. 1364–1370, Mar. 2021.

[8] Y. Zhu *et al.*,, "Spectrally-efficient single-carrier 400 G transmission enabled by probabilistic shaping," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, 2017, pp. 1–3.

[9] M. J. Ready and R. P. Gooch, "Blind equalization based on radius directed adaptation," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 1990, pp. 1699–1702.

[10] M. P. Yankov, E. P. daSilva, F. Da Ros, and D. Zibar, "Experimental analysis of pilot-based equalization for probabilistically shaped WDM systems with 256QAM/1024QAM," in *Proc. Opt. Fiber Commun. Conf.*, 2017, pp. W 2A-48.

[11] G. D. Rosa and A. Richter, "Blind radius directed equalizer with likelihood-based selection for probabilistically shaped and high order QAM," in *Proc. Eur. Conf. Opt. Commun.*, 2020, pp. 1–4.

[12] P. M. Krummrich, D. Ronnenberg, W. Schairer, D. Wienold, F. Jenau, and M. Herrmann, "Demanding response time requirements on coherent receivers due to fast polarization rotations caused by lightning events," *Opt. Exp.*, vol. 24, no. 11, pp. 12442–12457, 2016.

[13] D. Charlton *et al.*,, "Field measurements of SOP transients in OPGW, with time and location correlation to lightning strikes," *Opt. Exp.*, vol. 25, no. 9, pp. 9689–9696, 2017.

[14] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Sep./Oct. 2010.

[15] M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats," *J. Lightw. Technol.*, vol. 35, no. 5, pp. 1125–1141, 2017.

[16] M. Nölle, F. Frey, R. Elschner, C. Schmidt-Langhorst, J. K. Fischer, and C. Schubert, "Investigation of CAZAC sequences for data-aided channel estimation considering nonlinear optical transmission," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, 2015, pp. 1–3.

[17] M. Kuschnerov *et al.*, "Data-aided versus blind single-carrier coherent receivers," *IEEE Photon. J.*, vol. 2, no. 3, pp. 387–403, Jun. 2010.

[18] W. Yi *et al.*,, "Performance of momentum-based frequency-domain MIMO equalizer in the presence of feedback delay," *Opt. Exp.*, vol. 28, no. 13, pp. 19133–19143, 2020.

[19] J. Li *et al.*,, "Real-time fast polarization tracking based on polarization phase locking least mean square algorithm," *Opt. Exp.*, vol. 27, no. 16, pp. 22 116–22126, 2019.

[20] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 913–929, May 1993.

[21] J. Cho, L. Schmalen, and P. J. Winzer, "Normalized generalized mutual information as a forward error correction threshold for probabilistically shaped QAM," in *Proc. Eur. Conf. Opt. Commun.*, 2017, pp. 1–3.

[22] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems, "Achievable information rates for fiber optics: Applications and computations," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 424–439, 2018.

[23] J. Liang, Y. Fan, Z. Tao, X. Su, and H. Nakashima, "Transceiver imbalances compensation and monitoring by receiver DSP," *J. Lightw. Technol.*, early access, 2019, doi: 10.1109/JLT.2019.2944830.

[24] R. Rios-Müller, J. Renaudier, and G. Charlet, "Blind receiver skew compensation and estimation for long-haul non-dispersion managed systems using adaptive equalizer," *J. Lightw. Technol.*, vol. 33, no. 7, pp. 1315–1318, 2015.

[25] I. Fatadin, D. Ives, and S. J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," *J. Lightw. Technol.*, vol. 27, no. 15, pp. 3042–3049, 2009.

Gabriele Di Rosa (Student Member, IEEE) received the B.Sc. degree in electronic engineering (*cum laude*) from Politecnico di Torino, Turin, Italy, in 2016 and the double M.Sc. degree in electronic engineering from Politecnico di Torino *(cum laude)* and in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2019. While working toward the M.Sc. degree in 2019, he was a Product Development Engineer with Finisar AB, Järfälla, Sweden. He is currently enrolled with the Technical University of Berlin, Berlin, Germany and is working toward the Ph.D. degree with VPIphotonics GmbH, Berlin, Germany, as an early-stage Researcher of the Marie Skłodowska-Curie European Training Network WON. His research interests include ultra-wideband fiber-optic coherent communications, digital signal processing, advanced modulation techniques, and transceiver optimization.

André Richter (Senior Member, IEEE) received the M.Sc. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 1995, and the Dipl.-Ing. and Dr.-Ing. degrees from the Technical University of Berlin, Berlin, Germany, in 1997 and 2002, respectively, for novel work in modeling long haul fiber optical communications. In 1998, he was a Research Fellow with the University of Maryland, Baltimore County, Baltimore, MD, USA. In 1997, he joined the VPIphotonics team and serves there since then with his interests in modeling and design aspects of optical communications, fiber-optics, and integrated photonics, and in developing professional software solutions addressing these fields. He invoked an industry training program that served more than 1000 engineers and contributed to many co-funded research programs. He led international teams in application engineering, product management, research & development before being appointed as a General Manager in 2013. He coauthored more than 120 publications, including two book chapters. He is a Senior Member of OSA, and Member of VDE/ITG and SPIE, and has served in several technical program committees of OFC, ACP, and Summer Topicals.