Direct Detection of Bipolar Pulse Amplitude Modulation

Marco Secondini *Senior Member, IEEE*, Enrico Forestieri *Senior Member, IEEE*

Abstract—**Pulse amplitude modulation (PAM) is a widely employed digital modulation format. PAM formats are generally classified as bipolar PAM (BPAM) if using both positive and negative amplitude levels, and unipolar PAM (UPAM) if using only non-negative amplitude levels. While BPAM formats are in principle more energy efficient, they are not compatible with conventional direct detection (DD) schemes, so that UPAM formats are usually preferred for short-reach optical communications. In this work, we propose a novel DD scheme that employs oversampling at two samples per symbol to extract both the amplitude and phase information from the received optical signal, enabling the detection of BPAM signals. The proposed scheme uses a single photodetector, has no special requirements in terms of bandwidth, and requires only some minimal additional processing compared to a conventional scheme. A theoretical analysis, confirmed by numerical simulations, shows that, in the presence of optical amplifier noise, the proposed BPAM/DD transmission technique provides large optical signal-to-noise ratio gains compared to a more conventional UPAM/DD transmission and a good tolerance to group velocity dispersion and non-ideal filtering.**

Index Terms—**Optical fiber communication, pulse amplitude modulation, , direct detection, optical interconnections, modulation formats.**

I. INTRODUCTION

According to the terminology used in optical communications, optical systems are commonly divided into two main categories: *coherent detection* (CD) systems and *direct detection* (DD) systems. The former are characterized by the use of a local oscillator laser, whose light is mixed with the received optical signal before photodetection,[1] whereas the latter do not require a local oscillator and employ a single photodiode to detect the intensity of the received optical signal.

With respect to CD, DD is simpler, cheaper, and has no issues related to phase and polarization stability, being in fact insensitive to phase and polarization rotations. For these reasons, DD systems were the only ones practically used in optical communications until about a decade ago and are, still today, the preferred choice for short-reach links and optical interconnects. However, DD has also some relevant limitations compared to CD, so that the latter is now preferred for long-haul and submarine links [2, Sec. 2,6,7,12–14].

A first limitation of DD systems is their sensitivity to the receiver thermal noise, which is instead typically neg-

M. Secondini and E. Forestieri are with the Institute of Communication, Information and Perception Technologies, Scuola Superiore Sant'Anna, Pisa, Italy, and with the National Laboratory of Photonic Networks, CNIT, Pisa, Italy (email: marco.secondini@sssup.it; forestieri@sssup.it).
This work was supported in part by Ericsson.

[1]In wireless systems, "coherent detection" is commonly used with a different meaning, as it refers to a detection scheme that assumes perfect knowledge of the carrier phase (see also [1, Footnote 10] for a discussion on the different terminologies adopted in the optical and wireless communities).

ligible in CD systems thanks to the inherent amplification mechanism provided by the local oscillator. This limitation is, however, almost irrelevant for *optically amplified* systems, whose performance is dominated by amplified spontaneous emission (ASE) noise. In fact, due to the greater technological complexity of CD, the advent of optical amplifiers in the early nineties has practically stopped the research on CD systems for nearly two decades [3, Sec. 1.1.2].

A second limitation of DD systems is their ability to detect only the intensity of the optical signal, apparently losing the information contained in its phase, whereas CD systems are able to detect both amplitude and phase of the signal (or, equivalently, the two quadrature components). This information loss puts some constraints on the modulation formats that can be detected by DD and makes harder, or even impossible, the compensation of some transmission impairments. Usually, in DD systems, the baseband modulating signal is constrained to be real and positive, so that it can be directly mapped to the field intensity and recovered by the DD receiver: these systems are referred to as *intensity-modulation/direct-detection* (IM/DD) systems.

The aforementioned constraint makes IM/DD systems less efficient than CD systems. For instance, in a generic *pulse amplitude modulation* (PAM) alphabet, the M amplitude levels are selected to optimize the system performance under some specific constraints. IM/DD systems, due to the non-negativity constraint on the amplitude levels set by the square-law detector, employ a *unipolar* PAM (UPAM) alphabet $\mathcal{A} = \{A_1, A_2, \ldots, A_M\}$, with $A_M > A_{M-1} > \ldots > A_1 \geq 0$ [4]. On the other hand, CD systems, which are not constrained by the square-law detector, usually resort to a more efficient *bipolar* PAM (BPAM) alphabet $\mathcal{A} = \{\pm A_1, \ldots, \pm A_{M'}\}$, with $M' = M/2$ and $A_{M'} > A_{M'-1} > \ldots > A_1 > 0$ [5, Sec. 5.2], [6, Sec. 4.3.1]. In general, BPAM can achieve the same minimum distance between the amplitude levels as UPAM with a significantly lower energy per bit—up to $6\,\mathrm{dB}$ lower for large M. This advantage can be exploited by CD systems, but not by IM/DD systems, making the former much more efficient than the latter in terms of required *signal-to-noise ratio* (SNR), even in the presence of optical amplifiers.

There have been various attempts to overcome also this second limitation of DD systems and combine the advantages of amplitude and phase modulation with the simplicity of DD [7]. Some important results in this sense are discrete multitone PAM [8], carrierless amplitude and phase modulation [9], single-sideband (SSB) modulation combined with orthogonal frequency-division multiplexing (OFDM) [10], [11] and the Kramers–Kronig (KK) receiver [12]. Typically, the possibility of detecting both signal amplitude and phase comes at the

cost of inserting a strong optical carrier (hence frustrating any benefit in terms of required SNR) and/or increasing the modulator/demodulator complexity and the receiver bandwidth.

A fundamental question that arises when considering DD is if and how much it reduces the capacity of the *additive white Gaussian noise* (AWGN) channel with respect to the classical CD case studied by Shannon in his seminal work [13]. The capacity of the DD channel with AWGN has been investigated in the optical literature in two main scenarios [14, Ch. 11], depending on whether the AWGN is added *before* the photodetector (e.g., optical noise generated by optical amplifiers) [15]–[17] or *after* the photodetector (e.g., thermal noise) [18]–[20]. The first case has been studied also in the context of phase-noise channels [21], [22], whereas a mixed scenario, with both optical and thermal noise, has been investigated in [23]. Here, we are specifically interested in the optical-noise scenario. Formally, the problem can be reduced to comparing the capacity of two band-limited *waveform* channels: the AWGN channel $y(t) = x(t) + w(t)$—where $x(t)$ and $y(t)$ are, respectively, the input and output complex waveforms with (two-sided) bandwidth B, and $w(t)$ is the AWGN noise—and the corresponding DD channel, $y'(t) = |x(t) + w(t)|^2$. A similar problem has been widely investigated in the past, but with specific reference to the *discrete-time* versions of the same channels, $y_k = x_k + w_k$ and $y'_k = |x_k + w_k|^2$, where x_k, y_k, w_k, and y'_k are the samples of $x(t)$, $y(t)$, $w(t)$, and $y'(t)$ at time $t = k/B$. In this case, the loss of the phase information nearly halves the available degrees of freedom, approximately reducing the capacity by a factor of two [15], [16], [21], [22], [24]. However, a key issue that needs to be considered when referring to waveform channels—especially in the presence of a nonlinear element such as the square-law detector—is if the considered discrete-time channels are correct equivalent representations of the original waveform channels [25]. In fact, while sampling the output of the band-limited AWGN channel at rate B provides a sufficient statistic—indeed, this is the approach used by Shannon to derive his famous capacity result on this channel [13]—the output of the DD channel requires a doubled sampling rate to obtain a sufficient statistic. This is understood by noting that the square-law detector transforms a complex signal of bandwidth B into a real signal of bandwidth $2B$, so that the phase information is hopefully not completely lost, but simply transferred at higher frequency. This intuition is the basis for the demonstration of the following recent result on the capacity of the DD channel: with a sampling rate of $2B$, the apparently lost phase information is almost fully recovered, so that the capacity of the DD channel is at most 1 bit (per channel use) less than the capacity of the AWGN channel [17].

In this work, inspired by the key result on the capacity of the DD channel [17], we propose a simple DD scheme that allows to extract both the amplitude and phase information by employing oversampling: while the samples taken in the middle of each symbol time provide information about the symbol amplitude, the additional samples taken at intermediate times provide information about the phase difference between consecutive symbols. The proposed scheme enables the detection of BPAM signals, with a significant advantage in terms of required SNR and tolerance to group velocity
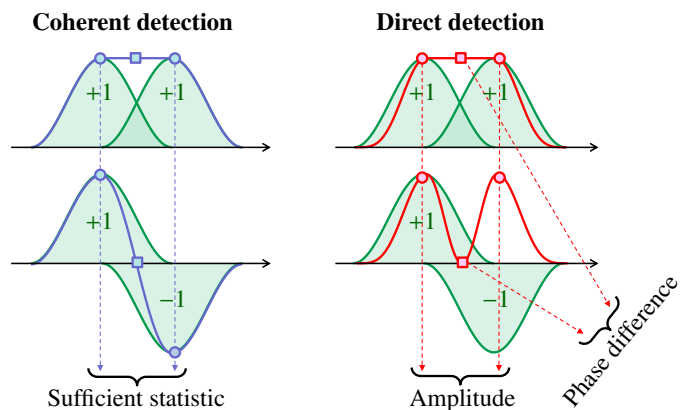

Fig. 1. Basic principle for detection of amplitude and phase by oversampling.

dispersion (GVD) compared to conventional IM/DD systems. The performance of BPAM/DD systems and their superiority compared to IM/DD systems is theoretically and numerically investigated, both in an ideal and in a more realistic scenario. A preliminary experimental demonstration of these results has been presented in [26].

The work is organized as follows. Section II introduces the basic principle of operation of the proposed detection scheme. Section III describes the modulation and direct detection of UPAM and BPAM signals, compares the two formats in terms of energy efficiency, and introduces a semianalytical method for the evaluation of performance (whose details are provided in the Appendix). Section IV validates the semianalytical method and provides some numerical results on the performance of UPAM and BPAM and their tolerance to GVD and non-ideal filtering. The practical implementation of the system and the comparison with other DD techniques are discussed in Section V. The conclusions are finally drawn in Section VI.

II. DIRECT DETECTION OF AN AMPLITUDE- AND PHASE-MODULATED SIGNAL

Before delving into a detailed description of the proposed transmission technique, we briefly illustrate the basic principle on which it is based. Fig. 1 shows two consecutive pulses (green filled lines) amplitude-modulated by symbols with either the same (upper row) or opposite (lower row) phases. *Coherent detection* of the pulses, as illustrated in the left column of Fig. 1, produces a photocurrent that is proportional to the *complex envelope* (blue line) of the optical signal. In this case, the *even* samples (circles), taken in the middle of each symbol time, are a *sufficient statistic* to determine both the amplitude and the phase of the transmitted symbols. On the other hand, *direct detection* of the pulses, as illustrated in the right column of Fig. 1, produces a photocurrent that is proportional to the *intensity* (red line) of the optical signal, so that the even samples provide information only about the amplitude of the symbols (always "1" in this example), but not about their phase. However, the *odd* samples (squares), which depend on the *constructive* or *destructive* interference between the two pulses, supply the missing piece of information necessary to determine the phase difference. In the example of Fig. 1, the
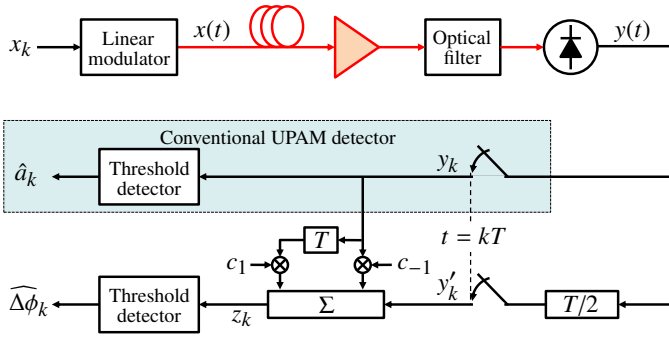
Fig. 2. Lowpass equivalent representation of the system.

odd sample has a high or low value when the symbols have the same or opposite phases, respectively. This mechanism can be extended to symbols with arbitrary amplitudes and phases, allowing the detection of both amplitude and (differential) phase modulation.

Based on the above, we consider a generic system with linear modulation of the optical field [5, eq. (2.107)] and direct detection, corrupted by amplifier noise. The lowpass equivalent representation of such a system is depicted in Fig. 2. The choice of an ideal linear modulator is due to its amenability to analytical developments, but a Mach–Zehnder modulator (MZM) can be used in practical realizations, as explained in Section V. The complex envelope of the transmitted optical field is

$$x(t) = \sum_{i=-\infty}^{\infty} x_i p(t - iT) \qquad (1)$$

where $p(t)$ is the supporting pulse with Fourier transform $P(f)$, T the symbol time, and x_i the i-th symbol, drawn from a given modulation alphabet. The signal, after propagating through the optical fiber with transfer function $H_f(f)$, is corrupted by ASE noise with power spectral density N_0, filtered by an optical filter with low-pass equivalent transfer function $H_o(f)$, and photodetected. Defining the overall channel transfer function as $H(f) = P(f)H_f(f)H_o(f)$ with corresponding impulse response $h(t) = \mathcal{F}^{-1}\{H(f)\}$, and assuming an ideal noiseless photodetector with infinite bandwidth and unit responsivity, the photodetected signal can be expressed as

$$y(t) = \left| \sum_{i=-\infty}^{\infty} x_i h(t - iT) + w(t) \right|^2 \qquad (2)$$

where $w(t)$ is the filtered ASE noise, with power spectral density $N_0|H_o(f)|^2$. The photodetected signal is then sampled by a couple of parallel symbol-time samplers, offset by $T/2$, to obtain two sequences of samples (corresponding to the even and odd samples when sampling at twice the symbol rate). Denoting by $h_k \triangleq h(kT/2)$ the channel coefficients and $w_k \triangleq w(kT/2)$ the noise samples, we have

$$y_k = y(kT) = \left| \sum_{i=-\infty}^{\infty} h_{2(k-i)} x_i + w_{2k} \right|^2 \qquad (3)$$

$$y_k' = y(kT - T/2) = \left| \sum_{i=-\infty}^{\infty} h_{2(k-i)-1} x_i + w_{2k-1} \right|^2 \qquad (4)$$

In order to simplify the analysis, we assume that all the coefficients h_k vanish for $|k| > 1$.[2] This simplified assumption is adopted only to illustrate the working principle and to design the detection strategy, but will be removed in the numerical results shown in Section IV. With this assumption, letting

$$n_k = 2\Re\{h_0 x_k w_{2k}^*\} + |w_{2k}|^2 \qquad (5)$$

$$n_k' = 2\Re\{(h_{-1} x_k + h_1 x_{k-1}) w_{2k-1}^*\} + |w_{2k-1}|^2 \qquad (6)$$

the received samples (3), (4) can be rewritten as

$$y_k = |h_0 x_k + w_{2k}|^2 = |h_0 x_k|^2 + n_k \qquad (7)$$

$$y_k' = |h_{-1} x_k + h_1 x_{k-1} + w_{2k-1}|^2$$
$$= |h_{-1} x_k|^2 + |h_1 x_{k-1}|^2$$
$$+ 2|h_{-1} h_1 x_k x_{k-1}| \cos(\Delta\phi_k + \theta) + n_k' \qquad (8)$$

where $\Delta\phi_k$ is the phase difference between the symbols x_{k-1} and x_k, θ the phase difference between the coefficients h_1 and h_{-1}, and (5), (6) are the noise terms affecting the even and odd samples, respectively. Eventually, the receiver in Fig. 2 computes the auxiliary decision variable

$$z_k = y_k' - c_{-1} y_k - c_1 y_{k-1} \qquad (9)$$

where the real coefficients $c_{\pm 1} = |h_{\pm 1}/h_0|^2$ depend on the overall channel response. By using (7) and (8), the auxiliary decision variable (9) can be expressed as

$$z_k = 2|h_{-1} h_1 x_k x_{k-1}| \cos(\Delta\phi_k + \theta) + \eta_k \qquad (10)$$

where $\eta_k = n_k' - c_{-1} n_k - c_1 n_{k-1}$ is the noise term.

Conventional IM/DD systems employ only the even samples (3) obtained from the upper branch of the receiver in Fig. 2. As shown in (7), these samples depend only on the squared modulus $|x_k|^2$ of the corresponding modulation symbols. In this case, any information about the symbol phase (or even polarity) is completely lost, so that only alphabets with real non-negative symbols are usually considered. On the other hand, as shown in (10), the auxiliary decision variable z_k, obtained from the lower branch of the receiver, depends also on $\Delta\phi_k$. Thus, once the amplitude information has been recovered, it can be employed to determine also the phase difference between consecutive symbols, provided that $-\theta \leq \Delta\phi_k \leq \pi - \theta$ ($\theta = 0$ in the case of a real or symmetric response, as it usually occurs in practical cases) and that the symbol amplitudes are large enough (which excludes the null symbol from the alphabet). Thus, the proposed DD scheme gives the possibility to exploit more energy-efficient modulation formats—for instance by properly distributing the modulation symbols over the whole real axis or in the upper complex half-plane—hopefully getting closer to the capacity bound derived in [17]. This approach is pursued with the BPAM format in the next section.

---

[2]This is a more restrictive assumption with respect to the Nyquist criterion, which ensures that only the even coefficients h_{2m}, with $|m| > 0$, vanish. As an example, two possible choices for the modulation pulse $p(t)$ that satisfy this condition when combined with the corresponding optical matched filter in the absence of fiber dispersion (back-to-back), are i) when $p(t)$ is an ideal rectangular pulse and ii) when its Fourier transform $P(f)$ is a root-raised cosine function with unit roll-off. In both cases, it can be shown that all the coefficients vanish but for $h_0 = 1$ and $h_{\pm 1} = 0.5$.

III. Unipolar and Bipolar PAM With Direct Detection

As shown in the previous section, the DD scheme of Fig. 2 allows the detection of both amplitude and (differential) phase values. A possible application, discussed in this section, is the adoption of BPAM alphabets even in DD systems, exploiting their superior energy efficiency compared to UPAM alphabets. In the following, we describe the modulation format and detection strategy of the proposed DD/BPAM system and compare them to those employed in conventional DD/UPAM systems, also highlighting the expected advantage of the former over the latter.

A. Bipolar PAM

The typical PAM constellation considered in many textbooks is a *bipolar* constellation with an even number M of equispaced levels symmetrically arranged around zero, so that $x_k \in \mathcal{A} = \{\pm A, \pm 3A, \dots, \pm(M-1)A\}$ [5, Sec. 5.2], [6, Sec. 4.3.1]. The spacing between adjacent levels is $2A$ and the amplitude A is selected to obtain the desired average energy per bit E_b. This constellation maximizes the minimum distance between levels for a given E_b, minimizing the error probability with coherent detection. The same constellation, however, turns out to be non-optimal when combined with the DD strategy in Fig. 2. On the other hand, a nearly optimum performance is obtained by the slightly modified alphabet

$$\mathcal{A}_{\mathrm{BPAM}} = \{\pm A, \pm 2A, \dots, \pm M'A\} \qquad (11)$$

where the spacing between the two innermost levels is still $2A$ while the spacing between any other two adjacent levels is A. An intuitive explanation and a numerical validation of this fact are reported in Sec. III-C and Sec. IV, respectively. Hereafter, we refer to (11) simply as the BPAM alphabet (levels).

Due to the inherent differential nature of the DD scheme in Fig. 2, the bit-to-symbol mapping requires a slight modification with respect to the conventional case. The symbol transmitted at discrete time k is expressed as $x_k = a_k e^{j\phi_k}$, where $a_k \in \{A, \dots, M'A\}$ is the (positive) amplitude, $\phi_k = \phi_{k-1} + \Delta\phi_k$ is the phase, and $\Delta\phi_k = \{0, \pi\}$ is the phase difference between consecutive symbols. A total of $m = \log_2 M$ bits are encoded on each transmitted symbol, with $m - 1 = \log_2(M')$ bits mapped to the M' positive amplitude levels according to a Gray map, and one bit encoded on the two phase difference values.

At the receiver side, each transmitted amplitude a_k and phase difference $\Delta\phi_k$ are recovered by processing the corresponding even sample (3) and auxiliary decision variable (9), respectively. Each even sample depends only on a_k, which is detected as in a conventional PAM by means of a (multiple) threshold detector

$$\widehat{a}_k = \begin{cases} A_1, & y_k \le \gamma_1 \\ A_2, & \gamma_1 < y_k \le \gamma_2 \\ \vdots & \vdots \\ A_{M'} & y_k > \gamma_{M'-1} \end{cases} \qquad (12)$$

where $\gamma_1, \dots, \gamma_{M'-1}$ are the optimized thresholds. On the other hand, the sign of the auxiliary decision variable depends

only on $\Delta\phi_k$, which is detected by a second threshold detector according to

$$\widehat{\Delta\phi}_k = \begin{cases} 0, & z_k \le \beta \\ \pi, & z_k > \beta \end{cases} \qquad (13)$$

where β is an optimized threshold. In general, the optimal thresholds are not located at the intermediate points between the signal levels. Due to the square-law detection, both the ISI and noise-noise beating terms are not zero-mean and affect the position of the optimal thresholds, with a complicated dependence on the SNR and filter bandwidth. Therefore, in our simulations, all the thresholds have been numerically optimized.[3] Eventually, information bits are recovered from $\widehat{a}_k$ and $\widehat{\Delta\phi}_k$ according to the encoding rule specified above. Equations (3), (9), (12), and (13) show that decisions are made symbol by symbol, with no decision feedback, so that no error propagation takes place at the receiver. Note that the detection strategy (12)–(13) is not optimal in a maximum a posteriori probability sense, but it is selected for its low implementation complexity and good performance.

B. Unipolar PAM

Conventional DD systems, due to the square-law detection, usually employ UPAM alphabets. A typical choice is selecting the *amplitude* levels of the optical signal as

$$\mathcal{A}_{\mathrm{UPAM-I}} = \{0, A, A\sqrt{2}, \dots, A\sqrt{M-1}\} \qquad (14)$$

so that, after square-law detection, the corresponding *intensity* levels are equally spaced starting from zero. The UPAM-I alphabet in (14) asymptotically guarantees the minimum error probability when the performance is dominated by electronics noise, modeled as AWGN that is added to the signal after the square-law detection.

On the contrary, when the performance is dominated by optical noise—AWGN that is added to the signal before the square-law detection—the best performance is obtained by equally spacing the *amplitude* levels according to

$$\mathcal{A}_{\mathrm{UPAM-A}} = \{0, A, 2A, \dots, (M-1)A\} \qquad (15)$$

so that the minimum distance between the amplitude levels is maximized [27]. The practical implementation of (14) and (15) is discussed in Section V. However, the significant advantage of (15) with respect to (14) may reduce or disappear in the presence of transceiver imperfections [27] or, as we will show in Section IV, transmission impairments.

In both the UPAM alphabets, a Gray map is used to encode $m = \log_2 M$ bits on each symbol. At the receiver side, only the upper branch of the detection scheme depicted in Fig. 2 is required to recover the transmitted symbols. The photodetected signal is sampled at symbol rate, and each sample is compared with $M - 1$ threshold levels to determine which symbol has been transmitted.

---

[3]It can be shown that in the absence of ISI and with matched filtering, the thresholds in (12) are given by $\gamma_n = (n + 1/2)^2 A^2 E_p^2$, $n = 1, 2, \dots, M'$, where E_p is the energy of the supporting pulse $p(t)$ in (1). However, even in this simplified case we did not succeed in obtaining an analytical expression for the optimum threshold in (13).
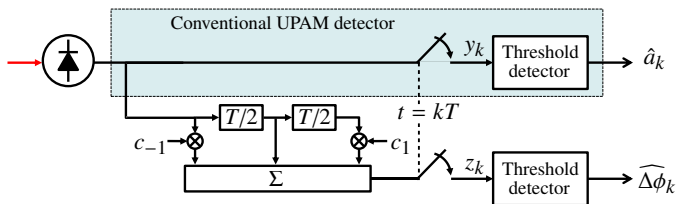
Fig. 3. BPAM receiver equivalent to the one in Fig. 2.

Fig. 4. BER vs the ratio between the amplitude levels of a BPAM-4 signal for $E_b/N_0 = 13\,\mathrm{dB}$ (symbols: simulations; solid lines: KLSE method).

C. Energy Efficiency

In order to get an idea of the performance gain that we can expect by employing the proposed BPAM system, we consider the ideal conditions $h_0 = 1$, $h_{\pm 1} = 0.5$ (see Footnote 2) and follow a simple heuristic approach. Since optical amplifiers introduce AWGN in the *optical domain*, we assume that the performance is mainly determined by the minimum distance between the different values taken by the *optical signal* at the sampling times [5, Sec. 4.3.6]. Thus, we compare the mean energy per bit that is required by the three alphabets to obtain a prescribed minimum distance in the optical domain. While this simple approach would be rigorously justified only for the optimal CD strategy, we conjecture that it provides a reasonable indication of the performance *difference* between the various formats even for the proposed DD strategy, also providing some insight on the BPAM format and its optimization.

For the even samples, relevant for amplitude detection and, hence, both for UPAM and BPAM systems, the minimum distance can be expressed as

$$d_e = \min_{m \neq n} |A_m - A_n| \qquad (16)$$

while for the odd samples, relevant for phase detection and, hence, only for BPAM systems, the minimum distance can be expressed as

$$d_o = \min |A_m| \qquad (17)$$

First of all, we note that in these ideal conditions the BPAM alphabet in (11) achieves the same minimum distance on both samples, $d_o = d_e = A$, guaranteeing the best performance trade-off between amplitude and phase detection. This is why we prefer this alphabet over the classical PAM $\{\pm A, \pm 3A, \ldots\}$ that, on the other hand, would give $d_o = d_e/2 = A$, favoring amplitude detection over phase detection.

Then, we consider the UPAM-I and UPAM-A alphabets in (14) and (15), for which the minimum distances are $d_e = A(\sqrt{M-1} - \sqrt{M-2})$ and $d_e = A$, respectively. Eventually, by choosing the parameter A such that to have the same minimum distance d for each format, the corresponding required average energy per bit turns out to be

$$E_b = \begin{cases} \dfrac{(M/2+1)(M+1)}{6\log_2 M} d^2, & \text{BPAM} \\[2ex] \dfrac{(M-1)d^2}{2\log_2 M(\sqrt{M-1} - \sqrt{M-2})^2}, & \text{UPAM-I} \\[2ex] \dfrac{(M-1)(2M-1)}{6\log_2 M} d^2, & \text{UPAM-A} \end{cases} \qquad (18)$$
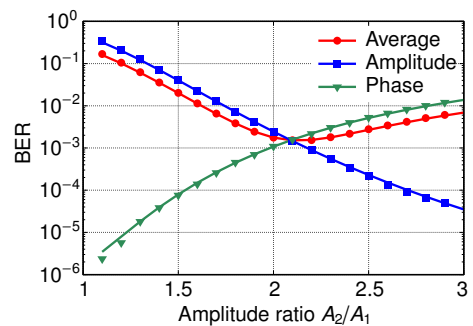
By comparing the corresponding E_b in (18), we see that BPAM has a potential advantage over UPAM-A of approximately 1.4 dB, 3.6 dB, and 4.8 dB for $M = 4, 8, 16$, respectively, asymptotically approaching 6 dB for $M \to \infty$. The potential advantage of BPAM over UPAM-I is even larger: 7.7 dB, 10.8 dB, and 12.3 dB for $M = 4, 8, 16$, respectively, asymptotically approaching 13.8 dB for $M \to \infty$.

In the next section, these potential gains, obtained by the heuristic approach described above, will be compared to the actual results obtained by numerical simulations. We defer to a future work the derivation of analytical expressions for the error probability in the BPAM ideal case, while very accurate analytical approximations for the UPAM-I and UPAM-A formats can be found in [27]. Observing that the receiver part in Fig. 2 is equivalent to the receiver in Fig. 3, in the general case the error probability can be computed by a straightforward extension of the Karhunen–Loéve series expansion (KLSE) method described in [28]. Indeed, it suffices incorporating the transfer function of the (analog) 3-tap equalizer into the post-detection filter and deal with two parallel detectors. The KLSE method is outlined in the Appendix and validated in the next section.

IV. NUMERICAL RESULTS

In this section, we first validate the theoretical predictions of Section III-C and the accuracy of the semianalytical KLSE method described in the Appendix by a comparison with Monte Carlo simulations and direct error counting. Then, we investigate the performance of the proposed BPAM/DD system and compare it to that of UPAM/DD systems. This is done first in the ideal scenario considered in Section II, and then in a more realistic scenario. In the ideal scenario, $P(f)$ has a root-raised-cosine shape with unit roll-off, the optical filter is matched to the pulse shape $H_o(f) = P(f)$, the photodetector has no band limitations, and both the optical signal and the noise are in a single polarization. The coefficients $c_{\pm 1}$ are set to the exact value indicated after (9). A large gain optical amplifier is considered, so that the performance is dominated by the ASE noise and depends on the E_b/N_0 ratio, while the receiver noise is negligible. The realistic scenario, on the other hand, is described later in greater detail and includes non-ideal pulse shaping, non-matched optical filtering, band-limited postdetection filtering, channel dispersion, and mismatched values of $c_{\pm 1}$. No adaptive equalization is used in any
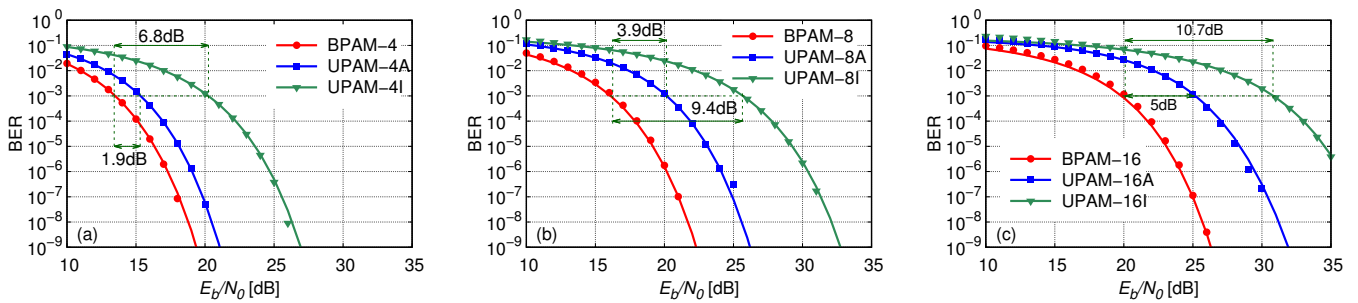
Fig. 5. Performance of BPAM, UPAM-A, and UPAM-I with (a) 4 levels (b) 8 levels, and (c) 16 levels (symbols: simulations; solid lines: KLSE method).
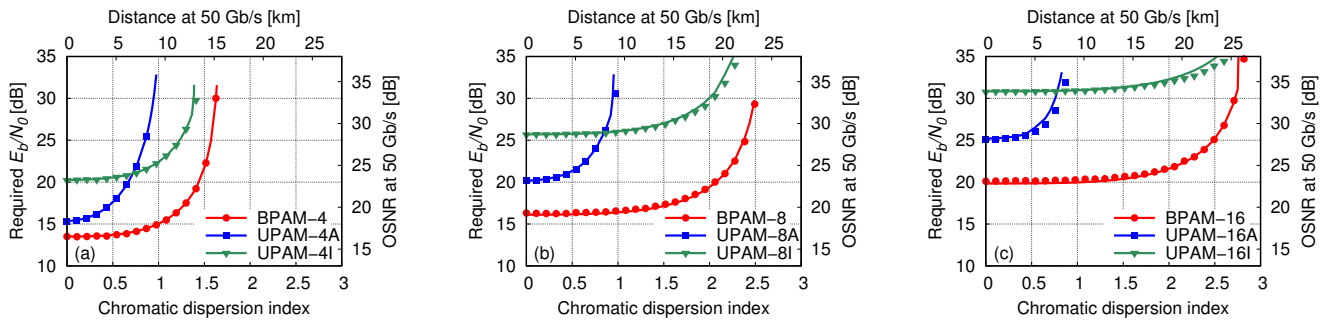

Fig. 6. GVD tolerance of BPAM, UPAM-A, and UPAM-I with (a) 4 levels (b) 8 levels, and (c) 16 levels (symbols: simulations; solid lines: KLSE method).

scenario. In Monte Carlo simulations, the BER is estimated by direct error couting, averaging over up to 10^8 random symbols to ensure a good accuracy down to BER values of the order of 10^{-6}.

First of all, we verify if the amplitude levels considered in (11) are optimal for the BPAM-4 format in the back-to-back case (without fiber dispersion). Given the mean energy per bit E_b, the BPAM-4 alphabet is fully defined by the ratio between the two positive amplitude levels A_2/A_1. Fig. 4 reports the BER as a function of the amplitude ratio A_2/A_1 for a fixed E_b/N_0 ratio of 13 dB. The results obtained by the KLSE method (solid line) and by Monte Carlo simulations (symbols) are in excellent agreement, but for some statistical fluctuations of the latter at low BER values. Besides the average BER, the BER on the even and odd samples (for amplitude and phase detection, respectively) are also reported. The minimum average BER is obtained around the optimal ratio $A_2/A_1 \approx 2.15$. This value is close to the theoretical value $A_2/A_1 = 2$ that has been heuristically predicted in Sec. III-C. However, since the optimal ratio slightly changes for different pulse or filter shapes and in the presence of GVD, we decided, for the sake of simplicity, to consider a fixed ratio $A_2/A_1 = 2$ in all results of the paper, corresponding to the BPAM alphabet defined in (11).

Fig. 5 shows the BER as a function of the E_b/N_0 ratio for the BPAM, UPAM-A, and UPAM-I formats. The 4-level, 8-level, and 16-level formats are considered in cases (a), (b), and (c), respectively. The agreement between the KLSE method and Monte Carlo simulations is confirmed in all the cases. As expected, when the dominant source of noise is the ASE, UPAM-A performs better than UPAM-I [27]. More importantly, the proposed BPAM format has an even better performance. For instance, the E_b/N_0 ratio required by

BPAM-4 at a BER of 10^{-3} is about 1.9 dB and 6.8 dB lower than that required by UPAM-4A and UPAM-4I, respectively. The gains respectively increase to 3.9 dB and 9.4 dB for the 8-level formats, and to 5 dB and 10.7 dB for the 16-level formats. Thus, despite the use of a non-conventional and non-optimal detection strategy based on the interference between consecutive pulses, the theoretical advantage of the BPAM alphabet over the UPAM ones predicted by (18) effectively translates into a relevant performance gain. Note that the actual gains gets closer to the theoretical ones predicted by (18) for lower BER values.

Another important characteristic that we want to test is the robustness of the format to GVD. Fig. 6 shows the E_b/N_0 ratio required by the three formats to achieve a BER of 10^{-3} as a function of the chromatic dispersion index $\gamma = -2\beta_2 R_b^2 L$, a dimensionless parameter that measures the accumulated GVD normalized to the bit-rate [29]. The 4-level, 8-level, and 16-level formats are considered in cases (a), (b), and (c), respectively. The agreement between the KLSE method and Monte Carlo simulations is confirmed also in this scenario. As a useful reference, the upper horizontal and right vertical axes report, respectively, the corresponding transmission distance and the required optical SNR (OSNR) for a 50 Gb/s transmission over a standard fiber with a GVD parameter $\beta_2 = -21.7\,\mathrm{ps}^2/\mathrm{km}$. From this figure, we note two important things. First, UPAM-A, though very convenient compared to UPAM-I in back-to-back, is much more sensitive to GVD and loses its initial advantage for $\gamma \approx 0.8$ (i.e., after about 8 km at 50 Gb/s). This, combined with its lower resilience to transceiver imperfections [27], makes UPAM-A a rather inconvenient choice, so that UPAM-I is usually preferred. On the other hand, BPAM has both the best back-to-back performance (as already shown in Fig. 5) and a good

tolerance to GVD. Thus, in the presence of GVD, BPAM keeps its large advantage compared to UPAM-I, and increases its advantage compared to UPAM-A. As an example, in back-to-back configuration, BPAM-16 requires about the same E_b/N_0 ratio (20 dB) as UPAM-4I and UPAM-8A. Moreover, when allocating 5 dB more ($E_b/N_0 = 25\,\mathrm{dB}$) for GVD tolerance, BPAM has twice and three times their reaches, respectively.

Finally, we investigate the performance of the proposed strategy in a more realistic scenario—i.e., considering non-ideal pulse shaping, a non-matched optical filter, and a limited electrical bandwidth. In particular, we consider a rectangular nonreturn-to-zero signal filtered by a Gaussian lowpass filter with 3-dB bandwidth $R_b/3$, a fourth-order super-Gaussian optical bandpass filter, a Gaussian electrical lowpass filter, and include ASE noise on both polarizations. The contour plots in Fig. 7(a)–(c) show the E_b/N_0 ratio required in a back-to-back case by BPAM-4, UPAM-4A, and UPAM-4I (at a BER of 10^{-3}) as a function of the optical and electrical 3-dB bandwidths, B_o and B_e, normalized to the bit-rate R_b. The parameters $c_{\pm 1}$ in (9) have been held fixed to the (ideal) common value $c_{\pm 1} = 0.25$ for all bandwidth values.

Even in this more realistic scenario, BPAM-4 has an advantage of about 2.2 dB and 7.3 dB over UPAM-4A and UPAM-4I, respectively, when considering their respective optimal configurations. Indeed, BPAM-4 has a wider optimal electrical bandwidth ($\sim 1.9R_b$) and is slightly more sensitive to deviations of the optical bandwidth from the optimal value compared to UPAM-4A and UPAM-4I. This peculiar behavior is due to the fact that the detection of the odd samples is based on the interference of adjacent pulses in the optical domain, which is favored by narrowing the optical bandwidth but not the electrical bandwidth. Moreover, the fixed choice $c_{\pm 1} = 0.25$ is nearly optimal only around the best combination of optical and electrical bandwidth. Nevertheless, even reducing the electrical bandwidth to more realistic values and considering a reasonable range of values for the optical bandwidth (e.g., $B_e = 0.6R_b$ and $0.65R_b \leq B_o \leq 0.75R_b$), the penalty experienced by BPAM-4 is less than 0.5 dB and its advantage over the UPAM-4 formats only slightly reduced.

The combined effect of non-ideal pulse shaping and filtering, GVD, and imperfect knowledge of the coefficients $c_{\pm 1}$ is eventually investigated in Fig. 8, which shows the contour plots of the E_b/N_0 required by BPAM-4 to achieve a BER of 10^{-3} with a chromatic dispersion index $\gamma = 0.9$ (corresponding to a standard fiber of about 8.3 km and a bit-rate of 50 Gb/s) and three different fixed values of $c_{\pm 1}$. All other parameters are the same as in Fig. 7. Two important observations can be made. First, the choice of the parameters $c_{\pm 1}$ in (9) is not at all critical: an increase of 20% and 40% of their value causes, respectively, a negligible change in the required E_b/N_0 in Fig. 8(b) and only a small (less than 0.4 dB in the interesting area) increase in Fig. 8(c). Second, the 1 dB penalty caused by GVD in this realistic scenario (measured by comparing Fig. 8(a) with Fig. 7(a) around the optimal configuration points) is similar to the penalty predicted by Fig. 6(a) in the ideal matched filtering scenario, though the absolute values of E_b/N_0 clearly differ.

V. DISCUSSION

In Sections II and III, in order to introduce the theoretical foundations of the BPAM/DD technique, we have considered an ideal implementation of the proposed system. In a practical implementation, the ideal modulator can be replaced by a conventional push–pull MZM, whose optical field transfer function can be written as [30]

$$T(\Delta V) = \cos\left[\frac{\pi(\Delta V - V_{\mathrm{bias}})}{2V_\pi}\right] \qquad (19)$$

where ΔV is the (zero-mean) voltage difference applied at the two arms of the MZM, V_{bias} the constant bias voltage, and V_π the *halfwave voltage* of the modulator. In order to obtain the BPAM modulation in (11), which takes both positive and negative values symmetrically distributed around zero, the MZM is biased at the *null point*, $V_{\mathrm{bias}} = V_\pi$. When a small driving voltage $\Delta V(t) \ll V_\pi$ is applied, (19) ensures an approximately *linear modulation* of the optical *amplitude*

$$T(\Delta V) = \sin\left(\frac{\pi \Delta V}{2V_\pi}\right) \approx \frac{\pi \Delta V}{2V_\pi} \qquad (20)$$

This is also the typical configuration employed to generate various zero-mean formats, such as BPSK or (on each arm of a nested MZM structure) QPSK and quadrature amplitude modulation. Often, to minimize the insertion loss of the modulator, the full characteristic of the MZM is exploited, $-V_\pi \leq \Delta V(t) \leq V_\pi$. In this case, the MZM nonlinearity in (20) causes a compression of the outer levels with respect to the inner ones. If needed, this compression can be compensated for by properly predistorting the modulation levels of the driving signal $\Delta V(t)$. For BPAM, however, such a predistortion may be unnecessary, since the nonlinear compression induced by the MZM is useful to obtain the correct level distribution in (11). This is particularly evident in BPAM-4, which can be directly obtained by driving the MZM with four voltage levels evenly spaced over the whole transmission characteristic

$$\Delta V = \left\{\pm \frac{V_\pi}{3}, \pm V_\pi\right\} \Rightarrow T(\Delta V) = \left\{\pm \frac{1}{2}, \pm 1\right\} \qquad (21)$$

The same MZM can be used also for UPAM-I and UPAM-A modulation, but with different settings. For UPAM-I, the bias is set at the so-called *quadrature point*, $V_{\mathrm{bias}} = V_\pi/2$, corresponding to 50% power transmission. When a small driving voltage is applied, this ensures an approximately *linear modulation* of the optical *intensity* around its average value

$$|T(\Delta v)|^2 = \cos^2\left(\frac{\pi \Delta V}{2V_\pi} - \pi/4\right) \approx \frac{1}{2} + \frac{\pi \Delta V}{2V_\pi} \qquad (22)$$

On the other hand, to maximize the extinction ratio, the driving voltage should range between the minimum and maximum transmission points, $-V_\pi/2 \leq \Delta V(t) \leq V_\pi/2$, while a proper predistortion should be applied to compensate for the MZM nonlinearity. Finally, UPAM-A requires a linear modulation of the optical amplitude as BPAM, but only positive amplitude levels as UPAM-I. This requires using the same portion of the MZM characteristic as UPAM-I, but with modulation levels asymmetrically distributed around the bias point—a somewhat
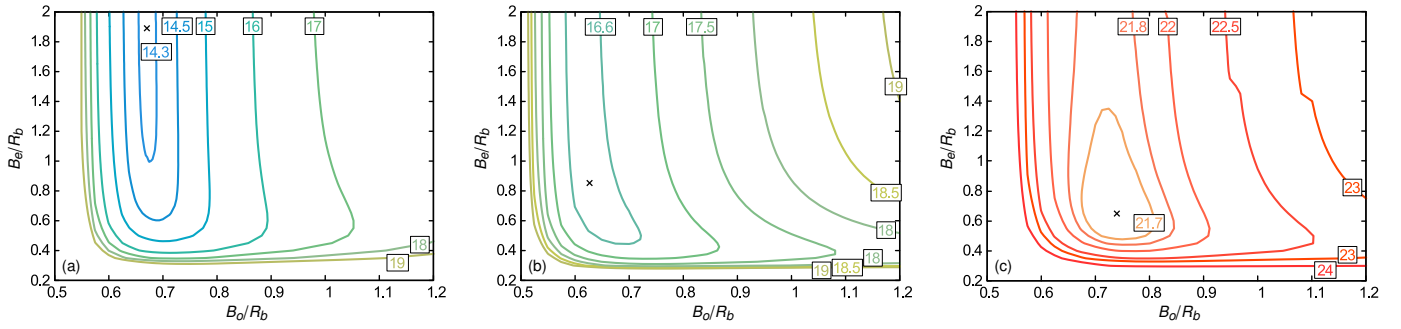
Fig. 7. Contour plot of the SNR per bit required by (a) BPAM-4, (b) UPAM-4A, and (c) UPAM-4I as a function of the optical and electrical filter bandwidths. The optimal configurations are denoted by the "×" symbols and require about 14.3 dB, 16.5 dB, and 21.6 dB, respectively.
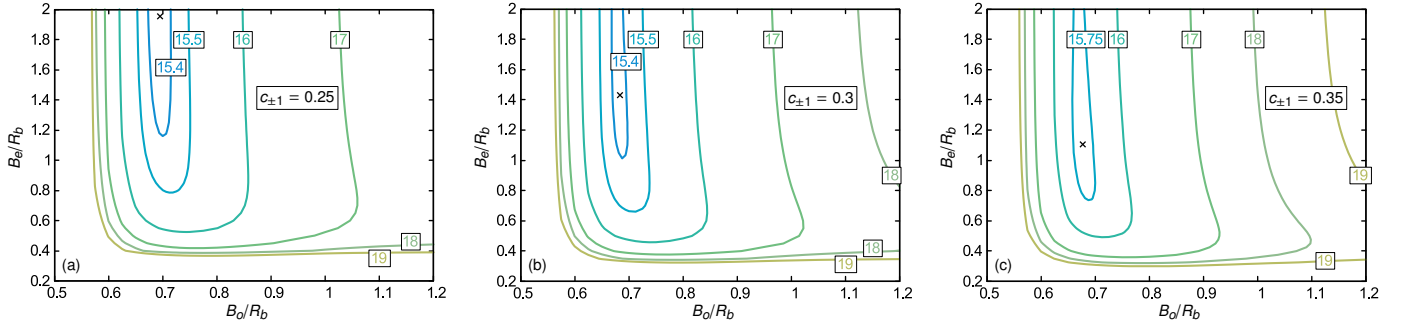

Fig. 8. Contour plot of the SNR per bit required by BPAM-4 in the presence of GVD (chromatic dispersion index $\gamma = 0.9$) as a function of the optical and electrical filter bandwidths, and for three different values of the parameters $c_{\pm 1}$. The optimal configurations are denoted by the "×" symbols and require about (a) 15.37 dB, (b) 15.38 dB, and (c) 15.72 dB.

more complicated configuration that makes UPAM-A harder to generate than BPAM and UPAM-I.

The optimal pulse shaping and filtering considered in Section II to obtain an ISI-free operation are not essential to implement a practical system and obtain relevant gains compared to conventional IM/DD systems. In fact, Fig. 6, Fig. 7, and Fig. 8 consider more realistic scenarios and include the impact of ISI due to propagation and/or filtering effects. The advantage of BPAM over both UPAM formats is confirmed also in these conditions.

Apparently, BPAM detection requires knowledge of the channel response to determine the coefficients $c_{\pm 1}$ in (9). This is also not a critical issue, as it is sufficient to select reasonable values for $c_{\pm 1}$ to obtain a good performance, as shown for instance in Fig. 7(a) and in Fig. 8(a)–(c). In practice, there are two possible choices: to select some fixed prescribed values for $c_{\pm 1}$ (e.g., the ideal values $c_{\pm 1} = 0.25$ or some other values based on reasonable assumptions about the system components) or to optimize $c_{\pm 1}$ to obtain the best performance. In the latter case, Fig. 3 shows that the auxiliary decision variable in (9) can be seen as the output of a 3-tap linear equalizer with tap coefficients $\{c_{-1}, 1, c_1\}$. The coefficients $c_{\pm 1}$ can thus be efficiently optimized (or even adaptively controlled, if desired) by using one of the several methods available for adaptive equalization [5, Chap. 8], [6, Chap. 11].

With respect to a conventional DD scheme, which can operate with one sample per symbol, the detector in Fig. 2 requires two samples per symbol. In a practical implementation, this requires either two parallel symbol-time samplers operating

with a relative delay of half a symbol time, or a single sampler operating at twice the sampling rate. This is the price to pay to obtain the large sensitivity gain shown in Section IV. However, as explained below, the increase of the sampling rate does not entail a corresponding increase of the required bandwidth. Moreover, the use of a fractionally-spaced (at two samples per symbol) feed-forward (or more complex) equalizer is often included in the implementation of conventional IM/DD systems to compensate for the transmitter and receiver bandwidth limitations [7]. In this case, the oversampling and additional processing in (9) required by BPAM come at no extra cost with respect to conventional PAM, though a separate equalization of the even and odd samples should be performed.

A preliminary experimental demonstration of BPAM, based on the practical implementation described above, is provided in [26]. The results show that, even employing standard devices and in the presence of a significant implementation penalty, the advantage of BPAM over UPAM is almost fully preserved.

The proposed system uses oversampling at two samples per symbol to extract some additional phase information from the photodetected signal. This information is, in part, contained at high frequencies after the square-law detector. Therefore, one may be tempted to conclude that BPAM requires twice the bandwidth of conventional IM/DD systems. This is not true. In fact, from a theoretical standpoint, the bandwidth increase caused by square-law detection takes place after photodetection, hence *not at optical level*. Thus, as it can be clearly appreciated from the behavior of the contour lines in Fig. 7(a)–(c) with respect to the horizontal axis B_o/R_b,

BPAM does not require more optical bandwidth than a conventional IM/DD system. Moreover, even if the total signal bandwidth is theoretically doubled after square-law detection, the receiver bandwidth need not be doubled to extract the desired information. This, too, can be verified in Fig. 7(a)–(c): while the optimal receiver bandwidth is indeed nearly doubled, the penalty experienced by the BPAM format when operating at a much lower bandwidth is very small.

In this work, we have compared the proposed BPAM/DD solution with more conventional IM/DD solutions based on UPAM-I and UPAM-A formats. There are two main reasons for that: the first reason is that UPAM-4I (simply known as *PAM4* in the optical literature) is a widely deployed standard for optical transmission [31]—UPAM-A being just a minor variant of UPAM-I. The other reason is that BPAM requires only a minor system modification with respect to UPAM-I and UPAM-A and is, therefore, directly comparable with them. Many other transmission techniques have been proposed in the last years to improve the performance of DD systems. Two relevant examples are SSB-OFDM [10], [11] and the KK receiver [12]. Even if all these techniques share the same idea of extracting phase information from a DD receiver, they are based on different approaches and have quite different characteristics. In fact, both SSB-OFDM and KK require the insertion of a strong optical carrier and significant additional complexity to extract the full phase information as in CD. On the other hand, the solution proposed in this work is based on the transmission of a carrierless BPAM signal, requires only minimal additional processing, and extract only partial phase information to allow the use of more energy efficient (zero-mean) constellations. In conclusions, these formats have different scope and applications. On one hand, BPAM/DD offers improved sensitivity compared to conventional IM/DD with only minimal additional complexity. On the other hand, KK and SSB-OFDM enables some useful features of CD (e.g., higher spectral efficiency and digital compensation of propagation impairments) at the cost of a higher required optical power and complexity.

VI. CONCLUSION

We have proposed a novel DD scheme that employs oversampling at two samples per symbol to detect an optical BPAM signal. With respect to a conventional UPAM/DD system, the proposed BPAM/DD system provides high OSNR gains and GVD robustness at the cost of minimal additional processing at the receiver, while still using a single modulator and a single photodetector. In an ideal optical-noise-limited regime, BPAM-4, BPAM-8, and BPAM-16 respectively provide OSNR gains of 6.8 dB, 9.4 dB, and 10.7 dB compared to UPAM-I formats with the same number of equally spaced intensity levels. Similar gains are observed also in a more realistic scenario, including GVD, non-ideal pulse shaping, and practical transceiver bandwidth limitations. The gains are still relevant, though lower (1.9 dB, 3.9 dB, and 5 dB, respectively), compared to UPAM-A formats with equally spaced amplitude levels. UPAM-A formats are, however, much more sensitive to GVD and other transceiver imperfections, being in fact rarely used in practice.

The proposed transmission technique is particularly appealing for optically-amplified DD links, in which the detrimental impact of signal-noise interaction limits the system performance and makes UPAM formats with more than four amplitude levels impractical. By contrast, the proposed BPAM/DD technique, thanks to its lower OSNR requirements and good GVD tolerance, allows the use of higher-level formats, increasing the system reach and decreasing the bandwidth requirements. For instance, for a fixed bit rate of 50 Gb/s, BPAM-16 requires an OSNR of about 23 dB in back-to-back, nearly the same as UPAM-4I and UPAM-8A. Moreover, by allocating 5 dB of OSNR margin for GVD tolerance, BPAM has twice and three times their reaches, respectively.

A preliminary experimental demonstration of the novel optical transmission technique introduced in this work can be found in [26]. We defer to a future work a more extensive experimental investigation and the analysis of other important issues, such as the performance in thermal-noise- and shot-noise-limited regimes, the use of more efficient constellations (e.g., with more than two phase values), and the potential application in a wireless context.

APPENDIX

In this appendix we briefly explain how to apply the KLSE method to the BPAM case. However, mutatis mutandis, the procedure also holds for both the UPAM cases.

In order to account for the ISI due to $n - 1$ M-ary symbols, we need a de Bruijn sequence $\{x_k\}_{k=1}^N$ of length $N = M^n$ [32] (see also [33] for a method to generate a quaternary sequence, easily extendable to sequences with M a power of 2). As already explained, the symbols are written as $x_k = a_k e^{j\phi_k}$, with $a_k \in \{A_1, A_2, \ldots, A_{M'}\}$, $M' = M/2$, and $\phi_k = \phi_{k-1} + \Delta\phi_k$, $\Delta\phi_k \in \{0, \pi\}$ being the phase difference between consecutive symbols. Denoting by $\mathcal{E}_k^{(1)}$ and $\mathcal{E}_k^{(2)}$ the error events associated with the k-th amplitude a_k and phase ϕ_k, respectively, assuming equiprobable symbols and a Gray map on the amplitude levels, the average probability of a bit error P_b can be written as

$$P_b \simeq \frac{1}{N \log_2 M} \sum_{k=1}^N \left(P\big(\mathcal{E}_k^{(1)}\big) + P\big(\mathcal{E}_k^{(2)}\big) \right) \qquad (23)$$

where $P\big(\mathcal{E}_k^{(1)}\big)$ and $P\big(\mathcal{E}_k^{(2)}\big)$ are, respectively, the symbol error probabilities in the upper and lower branch in Fig. 3. Taking into account (12) and (13), these probabilities can be evaluated through the pdfs of the samples y_k and z_k as

$$P\big(\mathcal{E}_k^{(1)}\big) = \begin{cases} P(y_k > \gamma_1) & \text{if } a_k = A_1 \\ P(y_k < \gamma_{i-1}) \\ \quad + P(y_k > \gamma_i) & \text{if } a_k = A_i, \ 2 \le i \le M' - 2 \\ P(y_k < \gamma_{M'-1}) & \text{if } a_k = A_{M'-1} \end{cases}$$
$$(24)$$

$$P\big(\mathcal{E}_k^{(2)}\big) = \begin{cases} P(z_k > \beta) & \text{if } \Delta\phi_k = 0 \\ P(z_k < \beta) & \text{if } \Delta\phi_k = \pi \end{cases} \qquad (25)$$

By choosing the sampling times such that they occur at a maximum eye opening, the minimum average P_e is obtained by minimizing (23) over the thresholds in (24) and (25). Those

appearing in (24) and (25) are the probabilities that a random variable with a generalized chi-square distribution exceeds (or does not) a given threshold. It can be shown that such probabilities may be approximated with very high accuracy by the so-called *saddlepoint approximation* [34], so that (note that the same expressions hold by replacing y_k with z_k)

$$P(y_k < \gamma_{\text{th}}) \simeq \frac{\exp\left[\Phi_{y_k}(s_0^-)\right]}{\sqrt{2\pi\Phi_{y_k}''(s_0^-)}} \qquad (26)$$

$$P(y_k > \gamma_{\text{th}}) \simeq \frac{\exp\left[\Phi_{y_k}(s_0^+)\right]}{\sqrt{2\pi\Phi_{y_k}''(s_0^+)}} \qquad (27)$$

where Φ_{y_k}'' denotes the second order derivative of the phase function $\Phi_{y_k}(s)$ defined as

$$\Phi_{y_k}(s) \triangleq \ln\left[\Psi_{y_k}(s)\frac{\exp(-y_{\text{th}}s)}{|s|}\right], \qquad s \in \mathbb{R} \qquad (28)$$

$\Psi_{y_k}(s) = E\{\exp(y_k s)\}$ being the moment generating function (MGF) of y_k. The values s_0^+ and s_0^- are the positive and negative, respectively, saddle points on the real s-axis of $\exp[\Phi_{y_k}(s)]$, and may be evaluated as the roots of the equation $\Phi_{y_k}'(s) = 0$. Hence, for computing the error probability, one simply needs to know the MGFs of the samples y_k and z_k. These MGFs can be computed by using one of the methods reviewed in [28], paying attention to incorporate the transfer function of the equalizer in the lower branch of Fig. 3 into the postdetection filter for evaluating the MGF of the samples z_k.

ACKNOWLEDGMENT

REFERENCES

[1] G. Colavolpe, T. Foggi, E. Forestieri, and G. Prati, "Robust multilevel coherent optical systems with linear processing at the receiver," *J. Lightw. Technol.*, vol. 27, no. 13, pp. 2357–2369, July 2009.

[2] E. Agrell, M. Karlsson, A. R. Chraplyvy, D. J. Richardson, P. M. Krummrich, P. Winzer, K. Roberts, J. K. Fischer, S. J. Savory, B. J. Eggleton, M. Secondini, F. R. Kschischang, A. Lord, J. Prat, I. Tomkos, J. E. Bowers, S. Srinivasan, M. Brandt-Pearce, and N. Gisin, "Roadmap of optical communications," *J. Optics*, vol. 18, no. 6, p. 063002, 2016.

[3] G. P. Agrawal, *Fiber-optic communications systems*, 3rd ed. Wiley, 2002.

[4] J. M. Kahn and K.-P. Ho, "Spectral efficiency limits and modulation/detection techniques for DWDM systems," *IEEE J. Select. Topics Quantum Electron.*, vol. 10, no. 2, pp. 259–272, 2004.

[5] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: with Wireless Applications*. Plenum Publishers, 1999.

[6] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, Aug. 2000.

[7] K. Zhong, X. Zhou, T. Gui, L. Tao, Y. Gao, W. Chen, J. Man, L. Zeng, A. P. T. Lau, and C. Lu, "Experimental study of PAM-4, CAP-16, and DMT for 100 Gb/s short reach optical transmission systems," *Opt. Exp.*, vol. 23, no. 2, pp. 1176–1189, 2015.

[8] S. C. J. Lee, S. Randel, F. Breyer, and A. M. J. Koonen, "PAM-DMT for intensity-modulated and direct-detection optical communication systems," *IEEE Photonics Technology Letters*, vol. 21, no. 23, pp. 1749–1751, 2009.

[9] L. Tao, Y. Wang, Y. Gao, A. P. T. Lau, N. Chi, and C. Lu, "Experimental demonstration of 10 Gb/s multi-level carrier-less amplitude and phase modulation for short range optical communication systems," *Opt. Exp.*, vol. 21, no. 5, pp. 6459–6465, 2013.

[10] J. Lowery and J. Armstrong, "Orthogonal-frequency-division multiplexing for dispersion compensation of long-haul optical systems," *Opt. Express*, vol. 14, no. 6, pp. 2079–2084, Mar 2006.

[11] B. J. C. Schmidt, A. J. Lowery, and J. Armstrong, "Experimental demonstrations of electronic dispersion compensation for long-haul transmission using direct-detection optical ofdm," *J. Lightwave Technol.*, vol. 26, no. 11, pp. 196–203, Jan 2008.

[12] A. Mecozzi, C. Antonelli, and M. Shtaif, "Kramers–kronig coherent receiver," *Optica*, vol. 3, no. 11, pp. 1220–1227, Nov 2016.

[13] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3/4, pp. 379–423/623–656, July/Oct. 1948.

[14] K.-P. Ho, *Phase-modulated optical communication systems*. Springer, 2005.

[15] A. Mecozzi and M. Shtaif, "On the capacity of intensity modulated systems using optical amplifiers," *IEEE Photonics Technology Letters*, vol. 13, no. 9, pp. 1029–1031, Sep. 2001.

[16] K.-P. Ho, "Exact evaluation of the capacity for intensity-modulated direct-detection channels with optical amplifier noises," *IEEE Photon. Technol. Lett.*, vol. 17, no. 4, pp. 858–860, Apr. 2005.

[17] A. Mecozzi and M. Shtaif, "Information capacity of direct detection optical transmission systems," *J. Lightw. Technol.*, vol. 36, no. 3, pp. 689–694, Feb. 2017.

[18] S. Hranilovic and F. R. Kschischang, "Capacity bounds for power-and band-limited optical intensity channels corrupted by Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 50, no. 5, pp. 784–795, May 2004.

[19] A. Lapidoth, S. M. Moser, and M. A. Wigger, "On the capacity of free-space optical intensity channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.

[20] A. A. Farid and S. Hranilovic, "Capacity bounds for wireless optical intensity channels with Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 56, no. 12, pp. 6066–6077, Dec. 2010.

[21] A. Lapidoth, "On phase noise channels at high SNR," in *Proc. IEEE Inf. Theory Workshop.*, Bangalore, India, Oct. 2002.

[22] M. Katz and S. Shamai (Shitz), "On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2257–2270, Oct. 2004.

[23] K. Keykhosravi, E. Agrell, M. Secondini, and M. Karlsson, "When to use optical amplification in noncoherent transmission: an information-theoretic approach," *IEEE Trans. Commun.*, 2020, dOI: 10.1109/TCOMM.2020.2968890.

[24] N. M. Blachman, "A comparison of the informational capacities of amplitude-and phase-modulation communication systems," *Proceedings of the IRE*, vol. 41, no. 6, pp. 748–759, 1953.

[25] M. Secondini, "Information capacity of optical channels," in *Optical Fiber Telecommunications VII*, A. Willner, Ed. Elsevier, 2020, pp. 867–920.

[26] M. Secondini, E. Forestieri, F. Fresi, L. Potì, T. Catuogno, and F. Cavaliere, "Bipolar pulse amplitude modulation with direct detection," in *Proc. European Conf. on Optical Commun.*, Dublin, Ireland, Sep. 2019.

[27] F. Fresi, E. Forestieri, M. Secondini, L. Potì, and F. Cavaliere, "PAM-4 performance optimization in presence of optical amplification," in *Proc. European Conf. on Optical Commun.*, Sep. 2018.

[28] E. Forestieri and M. Secondini, "On the error probability evaluation in lightwave systems with optical amplification," *J. Lightw. Technol.*, vol. 27, no. 6, pp. 706–717, Mar. 2009.

[29] A. F. Elrefaie, R. E. Wagner, D. Atlas, and D. Daut, "Chromatic dispersion limitations in coherent lightwave transmission systems," *J. Lightw. Technol.*, vol. 6, no. 5, pp. 704–709, 1988.

[30] J. Conradi, *Optical fiber telecommunications IV B*. New York: Academic, 2002, ch. Bandwidth-efficient modulation formats for digital fiber transmission systems, pp. 862–901.

[31] *IEEE Standard for Ethernet*, Std. IEEE Std 802.3bs, 2017.

[32] S. W. Golomb, *Shift register sequences*. San Francisco: Holden-Day, 1967.

[33] D. V. Den Borne, E. Gottwald, G. D. Khoe, and H. D. Waardt, "Bit pattern dependence in optical DQPSK modulation," *Electron. Lett.*, vol. 43, no. 22, 2007.

[34] C. W. Helstrom, "Approximate evaluation of detection probabilities in radar and optical communications," *IEEE Transactions on Aerospace Electronic Systems*, vol. 14, pp. 630–640, Jul. 1978.