Comparison of Probabilistically Shaped 64QAM with Lower-Cardinality Uniform Constellations in Long-Haul Optical Systems

Comparison of Probabilistically Shaped 64QAM with Lower-Cardinality Uniform Constellations in Long-Haul Optical Systems

Dario Pilori, *Student Member, OSA,* Luca Bertignono, *Student Member, OSA,* Antonino Nespola, Fabrizio Forghieri and Gabriella Bosco, *Fellow, OSA*

(Top Scored)

***Abstract*—In this paper, we compare the performance of probabilistically-shaped 64QAM with uniform 16QAM and 32QAM modulation formats at the same *net* data rate in long-haul coherent optical communications systems. Experimental results at 16 GBaud are shown, with off-line post-processing of the data performed using either an ideal or a realistic carrier phase estimation (CPE) scheme. We show that the choice of CPE algorithm at the receiver is crucial, since, as predicted by current models, most of the additional nonlinear noise introduced by the shaping is non-linear phase noise (NLPN). Thanks to the use of probabilistic shaping (PS), maximum reach gains ranging from 15.5% and 34% are obtained over pure silica-core fiber (PSCF), where the NLPN can be efficiently compensated for by standard CPE algorithms, whilst over non-zero dispersion-shifted fiber (NZDSF) the gain of PS is drastically reduced, due to residual short-correlated NLPN.**

***Index Terms*—Probabilistic shaping, optical fibers, coherent optical communications, quadrature amplitude modulation**

I. Introduction

RECENTLY, the use of Probabilistic Shaping (PS) techniques applied to high-order modulation formats has been proposed in coherent optical communication systems, as a means of increasing both the sensitivity and the flexibility of the transceivers [1]–[7]. The PS technique allows to approach the optimum Gaussian constellation in an Additive White Gaussian Noise (AWGN) channel by changing the a-priori probability of the symbols, thus reducing the shaping loss of standard QAM constellations.

In the past year, several works have been published, reporting maximum-reach gains with respect to standard QAM constellations ranging from 7% [2] to 40% [3]. In fact, the potential gain can be significantly different depending on several factors, such as reference modulation format, target transmission rate and the system scenario. In many cases, the comparison has been carried out between schemes with different asymptotic Mutual Information (MI) and considering Forward Error Correction (FEC) schemes with different overheads, making it difficult to assess the real gains achieved by constellation shaping.

Moreover, while performance of probabilistically-shaped constellations have been extensively studied over pure AWGN channels (*e.g.* [8], [9]), its performance over the non-linear fiber-optic channel is still under active investigation [6], [10], [11]. One of the most important differences between an AWGN channel and the coherent fiber-optic channel is the presence of Non-Linear Phase Noise (NLPN), which is constellation-dependent and may reduce the AWGN benefit of probabilistic shaping. Since NLPN is (at least partially) compensated by standard Carrier Phase Estimator (CPE) algorithms, it is not trivial to predict performance of PS constellations with standard Non-Linear Interference Noise (NLIN) models, such as the EGN-model [12]. Therefore, the amount of residual (after-CPE) NLPN strongly depends on the choice of CPE algorithm and on the kind of optical fiber.

Consequently, this work expands the results presented in [13], reporting a thorough experimental investigation of the maximum-reach gain that can be achieved by switching from standard QAM constellations to higher cardinality probabilistically-shaped constellations, keeping the same net (*i.e.* post-FEC) data rate. While in [13] we compared their performances over a single kind of fiber (PSCF) and with an ideal CPE algorithm, this work extends it over a different kind of fiber with a more realistic CPE to create conditions where the impact of NLPN may be different. This gain was measured using standard Digital Signal Processing (DSP) algorithms (either data-aided or decision-directed), without any specific NLIN compensation technique (*e.g.* DBP). In particular, we compared uniformly-shaped 16 and 32QAM with PS-64QAM, either at the same constellation entropy (which requires different FEC) or at the same FEC rate (which implies different entropy). In all of these cases, in order to have a fair comparison, we kept constant the net data rate. Afterwards, the experimental results are compared with two different NLIN models to find the more suitable model to accurately predict performance of PS constellations.

This paper is structured as follows. In section II we describe in detail the probabilistic shaping scheme and the adopted PS-64QAM constellations, showing their theoretical advantage in a pure AWGN channel. Then, in section III, we present the experimental setup and its results. In section IV we conclude the paper, summarizing the key results.

D. Pilori, L.Bertignono and G. Bosco are with Department of Electronics and Telecommunications (DET), Politecnico di Torino, 10129 Torino, Italy. E-Mail: dario.pilori@polito.it.

A. Nespola is with Istituto Superiore Mario Boella, 10129 Torino, Italy.

F. Forghieri is with Cisco Photonics Italy S.r.l., 20871 Vimercate (MB), Italy.

II. Basics of Probabilistic Shaping

The basic idea of probabilistic shaping is the transmission of standard QAM constellations points with different probabilities. This is done because QAM constellations are not optimal in an AWGN channel [14], since the optimal (*i.e.* capacity-achieving) constellation is the Gaussian constellation itself, which is not practical to implement in realistic communications systems.

In particular, uniform QAM constellations have, in the high-SNR regime, an asymptotic gap of $\pi e/6 \approx 1.53$ dB from the Gaussian constellation [14, Fig. 1], which represents the maximum SNR gain that can be achieved by constellation shaping. In this section, we will review the basic concepts of probabilistic shaping that have been applied to our analysis.

A. Performance metric

Instead of the popular pre-FEC BER, for this work we used the Mutual Information (MI) as performance metric [15]. Since in this work we did not consider any particular FEC implementation, we used the metric that gives the performance of the optimal receiver, the AWGN-channel MI. The MI has been calculated over an AWGN auxiliary channel with a Monte-Carlo method [16].

B. Constellation probability

It can be proven that the optimal probability distribution for a standard QAM constellation in AWGN is the so-called Maxwell-Boltzmann distribution [8], in which each constellation point a_i is transmitted with probability $\mathcal{P}(a_i) \propto \exp(-\lambda |a_i|^2)$. An increase of λ increases the amount of "shaping" applied to the constellation, reducing its entropy $\mathbb{H}[A] = -\sum_i \mathcal{P}(a_i) \log_2 \mathcal{P}(a_i)$ and increasing its efficiency in a specific SNR region [17].

However, the Maxwell-Boltzmann distribution, although optimal, presents some implementation issues with large values of λ. For instance, let us consider a PS-64QAM constellation with entropy $\mathbb{H}[A] = 4$ bit/symbol, which is obtained with $\lambda = 0.1334$ assuming a minimum distance between adjacent symbols $d_{\text{min}} = 2$. In this case, the ratio between the probability of the innermost and the outermost points of the constellation is $\exp[\lambda(98-2)] \approx 3.7 \times 10^5$. This means that one point in the inner-most ring of 64QAM is $\sim 3.7 \times 10^5$ times more likely to be transmitted than a point in the outer-most ring, and the number of symbols required to accurately approximate the desired distribution is extremely high. Therefore, since our DAC limits the length of the transmit sequence to 2^{16} symbols at 16 GBaud, it is very difficult to construct a transmit sequence with the given Maxwell-Boltzmann distribution.

To overcome this limitation we tried to slightly modify the Maxwell-Boltzmann distribution in order to reduce the ratio between the probabilities of inner and outer constellation points without compromising its optimality, significantly. A good compromise was obtained by removing the square in the modulus of the constellation, *i.e.*

$$\mathcal{P}(a_i) \propto e^{-\lambda |a_i|} \tag{1}$$

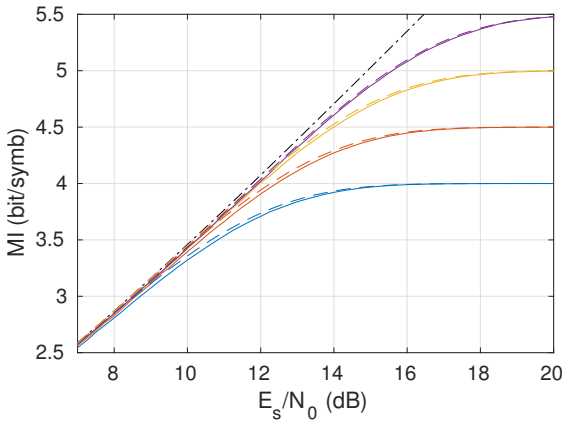


Fig. 1. Comparison between PS-64QAM constellations with entropy $\mathbb{H}[A] = 4, 4.5, 5, 5.5$ bit/symbol over an AWGN channel with Maxwell-Boltzmann distribution (dashed lines) and the distribution we employed (1) (solid lines). Channel capacity is indicated as black dashed-dotted line.

Considering the same example (PS-64QAM with $\mathbb{H}[A] = 4$ bit/symbol), with this distribution the outermost points of the constellations are just ~ 1700 more likely to be transmitted than innermost points. Therefore, it is now feasible to accurately generate a transmit sequence with this probability distribution that can fit in a DAC's memory.

Obviously, this distribution is not optimal anymore. The performance difference, in terms of MI over an AWGN channel for different values of SNR, is shown in Fig. 1. The maximum difference between the Maxwell-Boltzmann distribution and (1) happens with $\mathbb{H}[A] = 4$ bit/symbol, where it reaches $\Delta\text{MI} = 0.046$ bit/symbol for $E_s/N_0 = 8$ dB, which is a quite low value of SNR. At a reasonable target MI of 3.33 bit/symbol (corresponding to the operating point of an ideal FEC with rate $r = 5/6$), the SNR penalty is 0.17 dB.

However, as it will be shown later, this distribution still allows significant gains in maximum reach. Moreover, considering again PS-64QAM with $\mathbb{H}[A] = 4$ bit/symb, the Peak-to-Average Power Ratio (PAPR) of the Maxwell-Boltzmann distribution is 11.17 dB, while for the distribution of (1) is 10.96 dB, which is a 0.21-dB improvement that can mitigate the performance loss with a realistic DAC.

While a theoretical comparison between the Maxwell-Boltzmann distribution and other probability distributions is out of the scope of this paper, further investigations are needed to address this issue, since almost all current high-speed DACs have a strongly limited memory.

C. Probabilistic Amplitude Shaping (PAS)

While there are several methods to implement probabilistic constellation shaping, we chose to use the Constant Composition Distribution Matching (CCDM) algorithm [18], which allows a fixed-to-fixed length matching between random input bits to a probabilistically-shaped output constellation. Since the CCDM algorithm must be applied before FEC encoding, it is important that the FEC encoder does not change symbol probabilities. In [9] the authors propose a scheme called

TABLE I
PARAMETERS OF PS-64QAM CONSTELLATIONS

Comparison with	$\mathbb{H}[A]$ (bit/symb)	FEC rate	SNR gain (dB)	MI th. (bit/symb)
16QAM	4	8/9	0.85	3.6
16QAM	$13/3 \approx 4.33$	5/6	1.29	3.6
32QAM	5	31/36	0.80	4.5
32QAM	$31/6 \approx 5.17$	5/6	1.02	4.5

Probabilistic Amplitude Shaping (PAS) to achieve this result; the interested reader can refer to [3], [9] for details on PAS.

The net transmission rate with this scheme in bit/symbol is

$$R = \mathbb{H}[A] - (1-r)m \tag{2}$$

where m is the cardinality (in bit) of the constellation (*e.g.* $m = 6$ for 64QAM) and r the FEC code rate.

This equation is different than the same equation for uniform QAM constellations:

$$R = m_{\mathrm{U}} r_{\mathrm{U}} \tag{3}$$

where r_{U} is FEC code rate and m_{U} the cardinality of the uniform constellation.

It is trivial to see that, given the differences between (2) and (3), it is not easy to find adequate conditions to fairly compare uniform-shaped constellations and probabilistically-shaped constellations. Therefore, to perform a fair comparison, for this work we decided to compare one PS constellation with uniform lower-cardinality constellations at the same *net* data rate (*i.e.* R). By imposing equality between (2) and (3), there are two different conditions over which the net data rate is the same:

1) Same-entropy condition: We can force the entropy of the probabilistically-shaped constellation to be equal to the cardinality of the uniformly-shaped constellation (which corresponds to its entropy): $\mathbb{H}[A] = m_{\mathrm{U}}$. In this case, the PS constellation requires a FEC rate lower than the FEC rate used for the uniform constellation

$$r = 1 - \frac{m_{\mathrm{U}}}{m}(1 - r_{\mathrm{U}}) \tag{4}$$

2) Same-FEC-rate condition: By imposing the use of the same FEC $r = r_{\mathrm{U}}$, the entropy of the PS constellation is larger than the cardinality of the uniform constellation

$$\mathbb{H}[A] = m - r(m - m_{\mathrm{U}}) \tag{5}$$

D. Employed PS constellations

In this work, we compared PS-64QAM with 16QAM and 32QAM, at the same net data rate, in the two conditions mentioned before. The FEC code rate for uniform constellations was $5/6$, which corresponds to the standard 20% overhead. The parameters of the four PS-64QAM constellations are detailed in table I.

A plot of the AWGN performances of those four constellations is shown in Fig. 2. Considering the two PS-64QAM constellations with the same entropy as 16 and 32QAM, for large values of SNR they saturate at their entropy, similarly to standard QAM constellations. However, for medium values of SNR, they are closer to Shannon's limit than standard QAM constellation, exhibiting a sensitivity gain of approximately 1 dB.

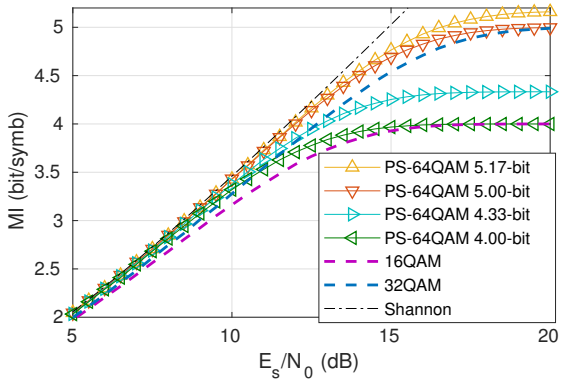


Fig. 2. Performance over an AWGN channels of the four PS-64QAM constellations detailed in table I. These constellations are compared with 16QAM and 32QAM (dashed lines), and channel capacity (dash-dot line).

To perform comparisons at the same net data rate, the constellations have to be compared at the same mutual information. Assuming an infinite block-length ideal FEC with 20% overhead, the minimum MI to achieve error-free transmission is $10/3 \approx 3.33$ bit/symb for 16QAM and $25/6 \approx 4.17$ bit/symb for 32QAM. To take into account the implementation penalties of realistic FEC, in this paper we compared those constellations at larger values of MI, namely 3.6 and 4.5 bit/symbol for respectively 16QAM and 32QAM, corresponding to a FEC implementation penalty of approximately 1.3 dB. Note that the exact MI threshold depends on the specific FEC scheme.

The SNR gain at the given MI thresholds of PS-64QAM compared with 16 and 32QAM is detailed in the last column of table I.

III. PROPAGATION PERFORMANCE

A. Introduction

In the previous section, the performance of PS constellations was evaluated in an ideal AWGN channel, which is a good approximation [19] of the optical channel with coherent detection.

However, more sophisticated models of the impact of fiber Kerr non-linearities on optical communications [12], [20], [21] have found a dependence on the modulation format and its statistics on the generation of NLIN. Moreover, recent investigations have found that NLIN can be subdivided in two main categories: additive short-correlated Gaussian-noise-like interference and long-correlated Phase and Polarization-Rotation Noise (PPRN) [22]–[24]. While the first category of NLIN is very well approximated as AWGN, the second interference is quite different than AWGN. Additionally, NLPN can be (at least partially) compensated using standard phase-recovery algorithms. Therefore, it is quite difficult to predict the performance of PS in realistic long-haul optical

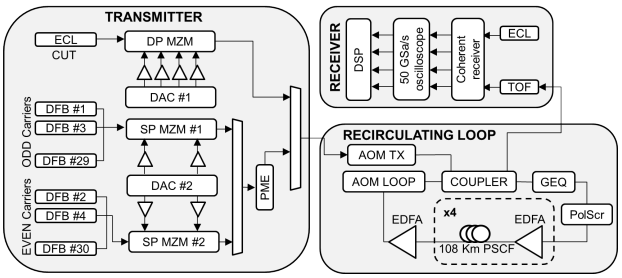


Fig. 3. Experimental setup.

TABLE II
FIBER PARAMETERS

Fiber	α (dB/km)	D_λ (ps/nm km)	γ (1/W km)	L_{span} (km)
PSCF	0.17	20.17	0.75	108
NZDSF	0.23	2.65	2	80

communications scenarios using standard models (such as the EGN model [12] or NLIN wizard [25]).

In [24], the authors performed a detailed analysis (using the theoretical model of [21]) of PPRN in fully-loaded C-band WDM systems, showing its dependence on main system parameters. In particular, they showed that PPRN increases with the increase of number of points of the constellation, and showed the maximum SNR gain obtainable with ideal PPRN mitigation. In [23], the authors performed split-step simulations to measure the magnitude and auto-correlation of NLPN. They also showed the maximum-reach gain obtainable with an ideal CPE, which approximately corresponds to the EGN model with QPSK correction factor for all modulation formats. This remarkable result has been experimentally validated in [26].

In this section, we evaluate the maximum-reach gain that can be obtained by switching from uniform-shaped 16 and 32QAM constellations to PS-64QAM at the same net data rate. Instead of focusing on the non-linear modeling (or the information-theoretical) implications, the goal of this section is finding the gain that can be achieved by *standard* receivers in *realistic* systems scenarios.

To understand the impact of main system parameters on the amount of residual NLPN, the comparison will be performed over two different kinds of optical fiber (PSCF and NZDSF), and two different CPE algorithms. The results will be then compared with the GN [19] and EGN [12] models predictions.

B. Experimental Setup

A scheme of the experimental setup is shown in Fig. 3. At the transmitter, 31 16-GBaud WDM channels are generated with 25 GHz spacing. The central Channel Under Test (CUT) is generated with a < 100 kHz External Cavity Laser (ECL), modulated with a dual-polarization lithium-niobate Mach-Zehnder Modulator (MZM) driven by four 64 GS/s DACs. The 30 interfering channels are generated with Distributed Feedback Lasers (DFBs), modulated by two single-polarization MZMs followed by a Polarization Multiplexing Emulator (PME), then added to the CUT. The 16 GBaud electrical signals were digitally pre-filtered to overcome components bandwidth limitations and shaped with a 15%-rolloff Root Raised Cosine (RRC) filter.

The WDM comb is then transmitted using a recirculating loop, made either by 4×108-km spans of Pure-Silica Core Fiber (PSCF) or 2×80-km spans of Non-Zero Dispersion-Shifted Fiber (NZDSF). The span length of both fibers has been chosen to have approximately the same span loss (18.4 dB). The main parameters of the fibers used in this experiment are reported in table II.

At the end of each span an Erbium-Doped Fiber Amplifier (EDFA) with a noise factor of 5.2 dB recovers span loss. An additional EDFA is used to compensate the extra losses of the Acousto-Optic Modulators (AOMs) and the Programmable Optical Filter (GEQ) used to compensate for non-flat EDFA gain. A Polarization Scrambler is adopted to randomize the state of polarization at the beginning of each loop.

At the receiver, a Tunable Optical Filter (TOF) selects the central CUT, which is mixed with a < 100 kHz ECL in an integrated coherent receiver. The four electrical signals at the output of the coherent receiver are sampled and digitized by a 50 GS/s real-time oscilloscope, then offline processed.

C. Receiver DSP

At the receiver, the signal is re-sampled at 2 samples per symbol and compensated for chromatic dispersion; the frequency offset is coarsely detected and removed using the standard periodogram method. Then, it is equalized with a 30-tap 4×4 real-valued MIMO LMS-based adaptive equalizer [27], with a modification to remove phase noise in the error computation [28].

After adaptive equalization, the CPE estimates and compensates for phase noise, which is a combination of laser phase noise and non-linear phase-noise. Therefore, the choice of the CPE algorithm strongly influences the performance after propagation. In this work, we implemented two different phase recovery schemes: fully-data-aided (genie-aided) CPE and a combination of Blind Phase Search (BPS) and Maximum Likelihood (ML) with pilot symbols.

1) Ideal CPE: This scheme, also called Phase Noise Receiver [29], tries to obtain the performance of a "perfect" blind CPE. Phase error is estimated using the transmitted data, then it is averaged with a moving average to avoid compensation of short-correlated noise (like ASE noise). The choice of the moving average length is crucial: a too small length may compensate part of ASE noise, while a too large length may not compensate for all phase noise. Therefore, we measured the Non-Circularity Index (NCI) of the constellation for different moving average lengths. The NCI, described in details in [23], is defined as the ratio between the variances of the radial and the tangent part of noise affecting the constellation. We then chose the moving average length as the value that makes the NCI as close as possible to 0 dB, *i.e.* the constellation points are as close as possible to circles. Since this value depends on the OSNR, we performed this operation for every received waveform.

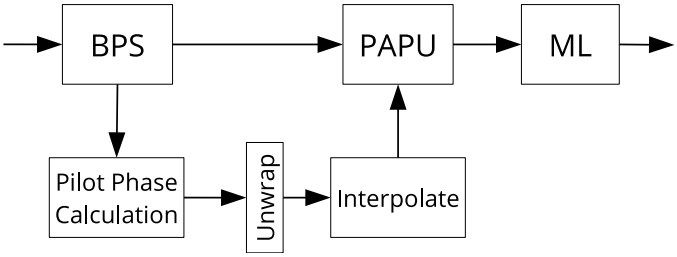


Fig. 4. Pilot-aided BPS-ML Carrier Phase Estimator (CPE). First, a coarse Blind Phase Search (BPS) [30] estimates carrier phase in the range $[-\pi/4, +\pi/4)$. Pilot blocks are extracted from the received signal, and phase between each block is unwrapped and linearly interpolated. A Pilot-Aided Phase Unwrapper (PAPU) unwraps the output phase of the BPS [33]. Afterwards, a Maximum Likelihood (ML) phase estimator obtains a fine estimation of the carrier phase [31].

2) BPS-ML: This scheme is a realistic pilot-symbol-based feed-forward CPE algorithm. It is based on a coarse BPS [30] algorithm followed by a fine ML phase estimator [31]. While these algorithms are completely blind, BPS gives an estimation of the phase in a quadrant, therefore it needs a phase unwrapping algorithm to get the "actual" phase. For high OSNRs, phase unwrapping can be performed blindly, but for our target OSNRs this is not possible. Therefore, following [32] [33], we periodically inserted short blocks of pilot symbols to perform a correct phase unwrapping procedure. Since the inserted pilot symbols were used *only* for phase unwrapping, the pilot overhead depends only on the amount of phase noise, and it can be kept low. Additionally, the choice of pilot overhead does not influence the performance of the receiver without cycle slips, but changes only the cycle slip probability. A high-level scheme of this CPE is shown in Fig. 4.

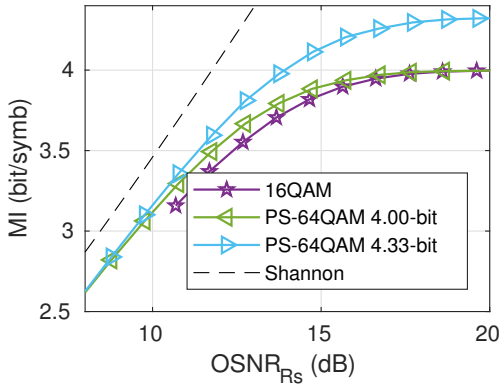
For this work, we used BPS with 18 test angles in the range $[-\pi/4, +\pi/4)$, and a 10-symbol pilot block every 502 data symbols, which approximately correspond to a 2% pilot overhead. This pilot overhead is larger than needed but, as previously stated, a larger pilot overhead influences only the cycle slip probability. As for the ideal CPE, we optimized the moving average length for every received waveform. Since this scheme is not genie-aided, we chose the length that maximizes the overall MI.

D. Optical back-to-back

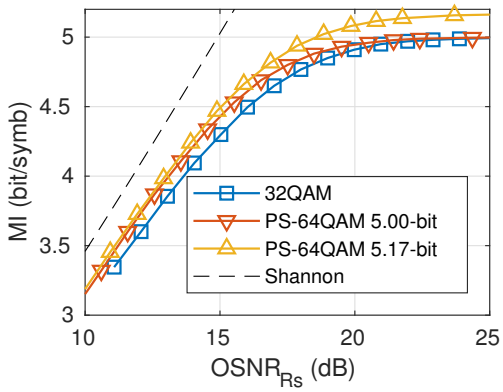
In order to characterize the setup and measure the penalty introduced by the transmitter and receiver, we performed a set of measurements with an optical back-to-back setup, shown in Fig. 5. The Optical Signal-to-Noise Ratio (OSNR) has been normalized to the symbol rate (16 GBaud). Comparing these results with the AWGN performances (Fig. 2), uniform constellations have ~ 0.9 dB of penalty, while PS-64QAM, due to its larger cardinality, has a penalty of ~ 1.05 dB.

E. Propagation over PSCF

Results after propagation over PSCF (see Fig. 3) are shown in Fig. 6, as maximum reach (expressed in kilometers) for



(a)



(b)

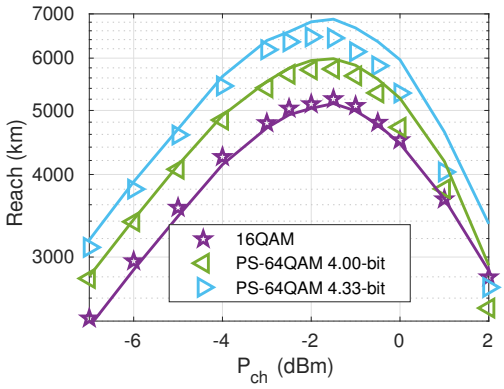
Fig. 5. Optical back-to-back results. The OSNR has been normalized to the symbol rate (16 GBaud). Average penalty with respect to theory (Fig. 2) is ~ 0.9 dB. These results have been used to calibrate the NLIN models described in section III-F.

different launch powers per WDM channel (P_{ch}). The maximum reach values have been calculated at the target MIs described in section II-D, namely 3.6 bit/symb (Fig. 6a) and 4.5 bit/symb (Fig. 6b). To obtain intermediate results, the MI has been linearly interpolated between different recirculating loops (4×108 km).

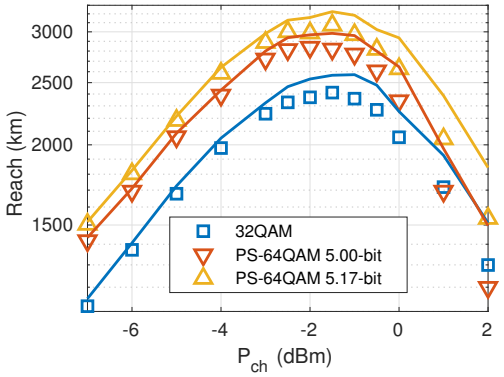
In the low-power region, where performance is ASE-noise limited, max-reach results achieved with the ideal CPE (solid lines) and BPS-ML CPE (markers) are almost identical. This suggests that, with laser phase noise only, the BPS-ML CPE can reach the performance of a perfect CPE.

Considering the ideal-CPE curves, the PS gain, which is the difference (in maximum reach) between uniform constellations and PS constellations, is the same both in the linear region and at the optimal power (maximum-reach region). This suggests that an ideal CPE is able to fully compensate the additional NLPN caused by "Gaussian-like" constellations. Using this ideal CPE, PS gains of PS-64QAM with respect to 16QAM are $+17.1\%$ and $+34.4\%$, respectively for 4 and 4.33 bit/symb of entropy, and $+15.5\%$ and $+24.9\%$ of PS-64QAM over 32QAM.

Looking at the results with BPS-ML CPE, they exhibit some penalty in the presence of NLPN, which translates into a slight reduction of maximum reach (and optimal power). 16QAM,



(a)



(b)

Fig. 6. Maximum-reach for different launch power P_{ch} over 108-km spans of PSCF. Maximum reach has been calculated at a target MI of 3.6 bit/symb (a) and 4.5 bit/symb (b). Solid lines are obtained with an ideal (genie-aided) CPE, while markers have been obtained with a realistic CPE (BPS-ML).

as predicted by the theoretical models, has less constellation-dependent NLIN, therefore keeps approximately the same performance with both CPEs. With BPS-ML, PS gains become $+11.3\%$ and $+24.4\%$ with respect to 16QAM and $+17.8\%$ and $+27.5\%$ with respect to 32QAM.

F. Comparison with NLIN models

In order to compare the experimental results with the NLIN models, we calculated the MI at the optimal launch power at the end of each loop, collecting the results in Fig. 7. Since the NLIN models output an equivalent Signal-to-Noise Ratio, we converted it to MI using the back-to-back measurements of Fig. 5. With this procedure, NLIN models predictions automatically take into account transceiver penalties.

We used two NLIN models: the well-known incoherent GN-model [19] (dashed lines), and the (coherent) EGN-model [12] (solid lines) with the correction factor for constant-amplitude (*e.g.* QPSK) modulation. Since the EGN model is computationally intensive to calculate, we used an analytical approximation [34], which, in this scenario, gives results quite close to the predictions of the complete EGN model.

From the results of Fig. 7, it can be seen that the EGN with QPSK correction factor predicts with great accuracy the performance of all the modulation formats with the ideal CPE, except for very short distances. However, performance reduces with the realistic BPS-ML CPE, becoming very close to the predictions of the incoherent GN-model. As shown in the previous section (Fig. 6), the impact of NLIN is weakly constellation-dependent with BPS-ML. In fact, the incoherent GN-model (iGN) slightly underestimates the performance of 16QAM, and slightly overestimates the performance of (for instance) PS-64QAM 4.33-bit. Nevertheless, the error (in maximum reach) at the target MIs is very small. For instance, the iGN-model underestimates the maximum reach of 16QAM by $\sim 3.2\%$, while it overestimates the maximum reach of PS-64QAM 4.33-bit by $\sim 2.5\%$. This result is consistent with past analyses with uniform-shaped constellations [35].

G. Propagation over NZDSF

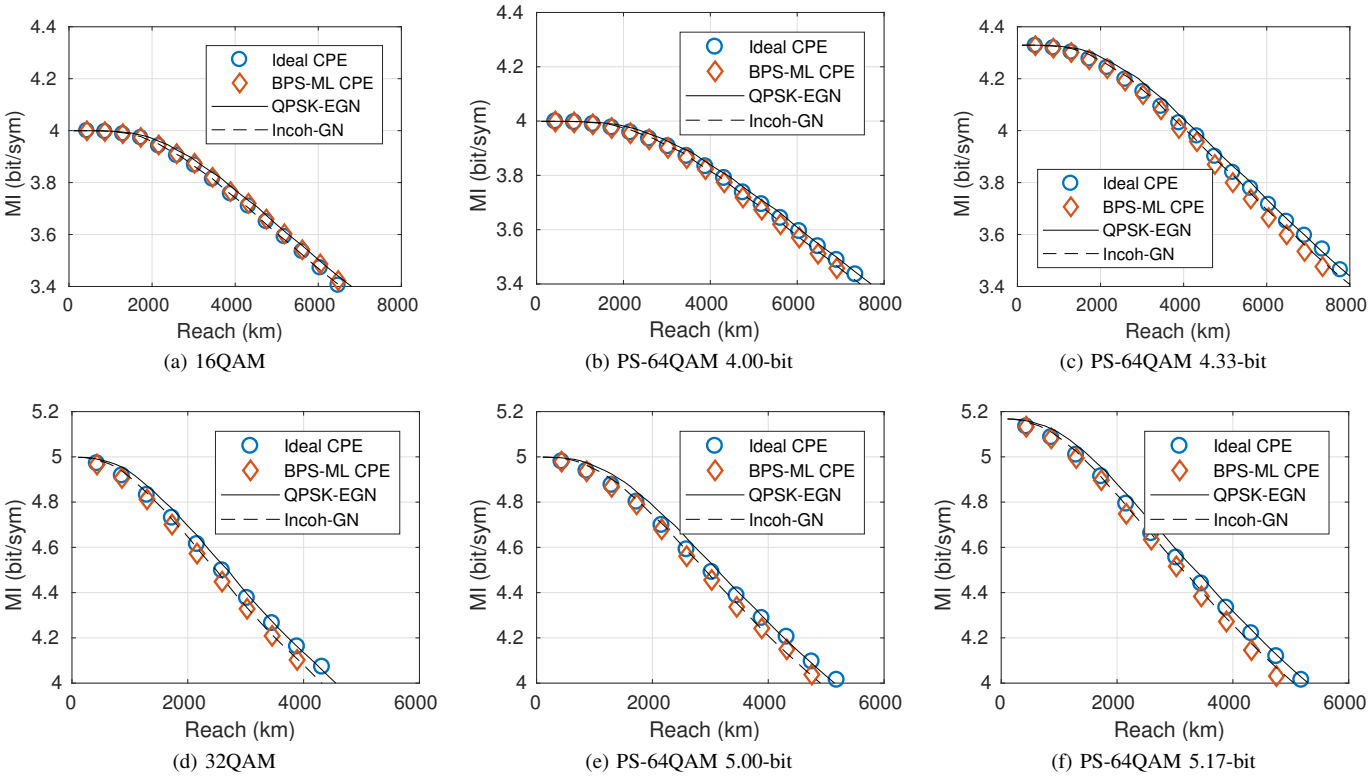
In the previous section, we showed that impact of residual NLPN after a BPS-ML CPE on PS-64QAM constellations over PSCF, is approximately the same as in lower-cardinality uniform constellations. Therefore, propagation performance can be accurately predicted by the GN-model using the back-to-back sensitivities.

However, this result holds only when the memory of NLPN is long enough to be almost completely compensated by the CPE algorithm, and this hypothesis holds only for high-dispersion optical fibers (such as PSCF). To test this hypothesis, we measured the performance of the same PS constellations considered over NZDSF. While this may seem an unrealistic scenario, NZDSF is still widely installed in long-haul networks of some countries, such as Italy [36, p. 37].

The experimental setup is the same as Fig. 3, where we replaced the 4×108-km spans of PSCF with 2×80-km spans of NZDSF, whose parameters are summarized in table II. The span length was reduced to 80 km to keep the same total span loss (approximately 18.4 dB).

Results are shown in Fig. 8 in terms of maximum reach for different launch powers per WDM channel (P_{ch}). As opposed to PSCF, with NZDSF constellation-dependent NLIN is strong, heavily reducing PS gain. In the case of 16QAM, PS gain is negative, *i.e.* 16QAM performs better than PS-64QAM. The gain of the ideal CPE with respect to BPS-ML is grater than the result of PSCF and more constellation dependent, which suggests that the additional NLIN is mostly short-correlated NLPN. Due to its short correlation, the estimation of NLPN by the CPE, especially the BPS-ML, is more difficult.

To validate this hypothesis, we set up a numerical simulation using the Split-Step Fourier method [37] to measure the correlation of NLPN without ASE noise nor laser phase noise. To perform a fair comparison with the experimental results, simulations have been carried out with the same number of WDM channels (31 at 16 GBaud with 25 GHz spacing). For PSCF, we simulated propagation over 70×108 km with $P_{\text{ch}} = -1.7$ dBm, while for NZDSF we propagated over 24×80 km with $P_{\text{ch}} = -6.5$ dBm. We then measured, following [23], the auto-correlation (in symbols) of the phase component of NLIN, which is shown in Fig. 9. Over PSCF, the auto-correlation of NLPN is much longer than NZDSF. For instance, the normalized auto-correlation reaches 0.1 after



(a) 16QAM (b) PS-64QAM 4.00-bit (c) PS-64QAM 4.33-bit

(d) 32QAM (e) PS-64QAM 5.00-bit (f) PS-64QAM 5.17-bit

Fig. 7. Maximum reach at the optimal power measured at the end of each loop of 4×108 km of PSCF (markers), using two different phase recovery schemes: ideal (circles) and BPS-ML (diamonds). Moving average length of each CPE has been optimized in every point. Results are compared with the incoherent GN-model [19] (dashed lines) and EGN model with correction factor for QPSK [34] (solid lines).

950 samples with PSCF and 45 samples with NZDSF, which explains why the CPE is less effective over NZDSF.

While a theoretical analysis of NLPN with probabilistic shaping is out of the scope of this work, we can conclude that the incoherent GN-model well predicts PS performance in scenarios where constellation-dependent NLPN has a sufficiently long auto-correlation to be compensated with standard CPE algorithms. This is the case in the PSCF propagation experiments. On the contrary, on low-dispersion fibers (*e.g.* NZDSF), the auto-correlation is significantly shorter, strongly reducing the gain of probabilistic shaping over standard QAM constellations.

We remark that those results have been obtained with a relatively low symbol rate (16 GBaud). A possible mitigation of this effect would be an increase of symbol rate. While it increases the overall NLIN [38], it also increases the auto-correlation length of NLPN [24], potentially increasing the gain of probabilistic shaping over NZDSF.

IV. Conclusion

In this paper, we experimentally compared PS-64QAM with uniform 16QAM and 32QAM at the same *net* data rate in long-haul optical communications systems. Using the PAS algorithm, we imposed a modified Maxwell-Boltzmann probability distribution (1) which allowed to transmit shorter PRBS sequences at the expense of a slight back-to-back penalty.

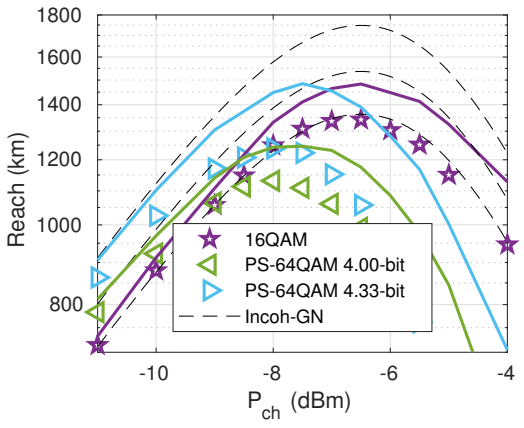
After propagation over PSCF and NZDSF, we found that the choice of CPE algorithm at the receiver is crucial, since, as predicted by current models, most of the additional NLIN introduced by PS is Non-Linear Phase Noise (NLPN).

Over PSCF, using both realistic and ideal CPE algorithms, we found that residual NLIN is weakly constellation-dependent due to its long auto-correlation. Therefore, the gain of PS-64QAM over 16 and 32QAM matches the gains measured in back-to-back. In particular, with a realistic CPE (such as BPS-ML with pilot tones), we found that the maximum reach of all the considered modulation formats is very well approximated by the incoherent GN-model. The gains (in maximum reach) were ranging between 15.5% and 34%. Moreover, the incoherent accumulation of NLIN allows the inclusion of probabilistic shaping in optical networking tools [39], [40], enabling the evaluation of its network benefits [41].

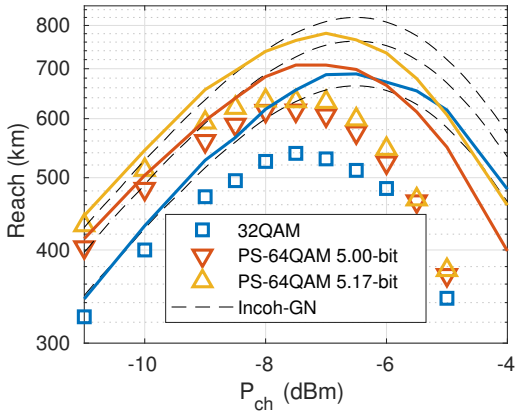
However, using low-dispersion fibers, such as NZDSF, we found that the performance of PS becomes worse than expected from the back-to-back results. This is due to the short auto-correlation of NLPN, that cannot be compensated by standard CPE algorithms. We remark that those results have been obtained at a (relatively low) 16 GBaud with a modified Maxwell-Boltzmann distribution. Future work should address a comparison at a higher symbol rate with the correct Maxwell-Boltzmann distribution, which may lead to different conclusions.

Acknowledgment

This work was supported by Cisco Photonics with an SRA contract. The authors would like to thank A. Bisplinghoff, C.



(a)



(b)

Fig. 8. Maximum-reach for different launch power P_{ch} over 80-km spans of NZDSF. Maximum reach has been calculated with a target MI of 3.6 bit/symb (a) and 4.5 bit/symb (b). Solid lines are obtained with an ideal (genie-aided) CPE, while markers have been obtained with a realistic CPE. Results are compared with incoherent GN-model predictions (dashed lines).

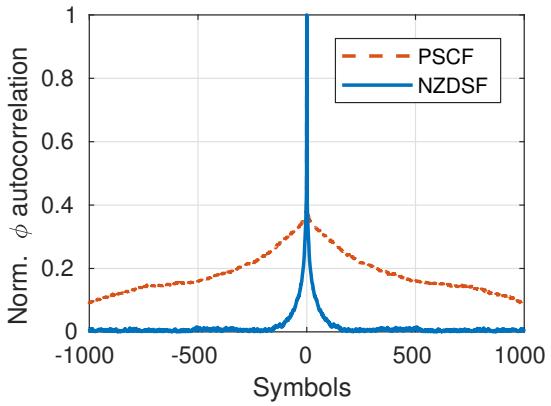


Fig. 9. Auto-correlation of phase-component of NLIN of PS-64QAM 4.00-bit over PSCF and NZDSF in maximum-reach conditions. These results have been obtained with numerical simulations without ASE noise and laser phase noise.

Fludger and P. Poggiolini for useful discussions. We would also like to thank the OFC 2017 attendees that gave their feedback at the poster W2A.57.

REFERENCES

[1] M. P. Yankov, D. Zibar, K. J. Larsen, L. P. B. Christensen, and S. Forchhammer, "Constellation shaping for fiber-optic channels with QAM and high spectral efficiency," *IEEE Photonics Technology Letters*, vol. 26, no. 23, pp. 2407–2410, Dec 2014.

[2] C. Pan and F. R. Kschischang, "Probabilistic 16-QAM shaping in WDM systems," *Journal of Lightwave Technology*, vol. 34, no. 18, pp. 4285–4292, Sept 2016.

[3] F. Buchali, F. Steiner, G. Bcherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *Journal of Lightwave Technology*, vol. 34, no. 7, pp. 1599–1609, April 2016.

[4] S. Chandrasekhar, B. Li, J. Cho, X. Chen, E. Burrows, G. Raybon, and P. Winzer, "High-spectral-efficiency transmission of PDM 256-QAM with parallel probabilistic shaping at record rate-reach trade-offs," in *ECOC 2016 - Post Deadline Paper; 42nd European Conference on Optical Communication*, Sept 2016, pp. 1–3.

[5] J. Cho, X. Chen, S. Chandrasekhar, G. Raybon, R. Dar, L. Schmalen, E. Burrows, A. Adamiecki, S. Corteselli, Y. Pan, D. Correa, B. McKay, S. Zsigmond, P. Winzer, and S. Grubb, "Trans-atlantic field trial using probabilistically shaped 64-QAM at high spectral efficiencies and single-carrier real-time 250-Gb/s 16-QAM," in *Optical Fiber Communication Conference Postdeadline Papers*. Optical Society of America, 2017, p. Th5B.3. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2017-Th5B.3

[6] A. Ghazisaeidi, I. F. de Jauregui Ruiz, R. Rios-Mller, L. Schmalen, P. Tran, P. Brindel, A. C. Meseguer, Q. Hu, F. Buchali, G. Charlet, and J. Renaudier, "Advanced C+L-band transoceanic transmission systems based on probabilistically shaped PDM-64QAM," *Journal of Lightwave Technology*, vol. 35, no. 7, pp. 1291–1299, April 2017.

[7] W. Idler, F. Buchali, L. Schmalen, E. Lach, R.-P. Braun, G. Böcherer, P. Schulte, and F. Steiner, "Field trial of a 1 Tb/s super-channel network using probabilistically shaped constellations," *J. Lightwave Technol.*, vol. 35, no. 8, pp. 1399–1406, Apr 2017. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-35-8-1399

[8] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for gaussian channels," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 913–929, May 1993.

[9] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 4651–4665, Dec 2015.

[10] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "On shaping gain in the nonlinear fiber-optic channel," in *2014 IEEE International Symposium on Information Theory*, June 2014, pp. 2794–2798.

[11] T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, "On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel," *J. Lightwave Technol.*, vol. 34, no. 21, pp. 5063–5073, Nov 2016. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-34-21-5063

[12] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, Jun 2014. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-22-13-16335

[13] L. Bertignono, D. Pilori, A. Nespola, F. Forghieri, and G. Bosco, "Experimental comparison of PM-16QAM and PM-32QAM with probabilistically shaped PM-64QAM," in *Optical Fiber Communication Conference*. Optical Society of America, 2017, p. M3C.2. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2017-M3C.2

[14] G. D. Forney and G. Ungerboeck, "Modulation and coding for linear gaussian channels," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2384–2415, Oct 1998.

[15] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft-decision FEC limit paradigm in the design of optical communication systems," *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 707–721, Jan 2016.

[16] A. Alvarado, "Information rates and post-FEC BER prediction in optical fiber communications," in *Optical Fiber Communication Conference*. Optical Society of America, 2017, p. Th3F.3. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2017-Th3F.3

[17] D. Pilori, F. Forghieri, and G. Bosco, "Maximization of the achievable mutual information using probabilistically shaped squared-qam constellations," in *Optical Fiber Communication Conference*. Optical Society of America, 2017, p. W2A.57. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2017-W2A.57

[18] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, Jan 2016.

[19] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, Feb 2014.

[20] A. Mecozzi and R.-J. Essiambre, "Nonlinear shannon limit in pseudolinear coherent systems," *J. Lightwave Technol.*, vol. 30, no. 12, pp. 2011–2024, Jun 2012. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-30-12-2011

[21] O. Golani, R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Modeling the bit-error-rate performance of nonlinear fiber-optic systems," *J. Lightwave Technol.*, vol. 34, no. 15, pp. 3482–3489, Aug 2016. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-34-15-3482

[22] M. Secondini and E. Forestieri, "Analytical fiber-optic channel model in the presence of cross-phase modulation," *IEEE Photonics Technology Letters*, vol. 24, no. 22, pp. 2016–2019, Nov 2012.

[23] P. Poggiolini and Y. Jiang, "Recent advances in the modeling of the impact of nonlinear fiber propagation effects on uncompensated coherent transmission systems," *Journal of Lightwave Technology*, vol. 35, no. 3, pp. 458–480, Feb 2017.

[24] R. Dar and P. J. Winzer, "Nonlinear interference mitigation: Methods and potential gain," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 903–930, Feb 2017.

[25] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Accumulation of nonlinear interference noise in fiber-optic systems," *Opt. Express*, vol. 22, no. 12, pp. 14 199–14 211, Jun 2014. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-22-12-14199

[26] A. Nespola, L. Bertignono, G. Bosco, A. Carena, P. Poggiolini, and F. Forghieri, "Independence of the impact of inter-channel non-linear effects on modulation format and system implications," in *ECOC 2016; 42nd European Conference on Optical Communication*, Sept 2016, pp. 1–3.

[27] M. Paskov, D. Lavery, and S. J. Savory, "Blind equalization of receiver in-phase/quadrature skew in the presence of nyquist filtering," *IEEE Photonics Technology Letters*, vol. 25, no. 24, pp. 2446–2449, Dec 2013.

[28] S. Randel, P. J. Winzer, M. Montoliu, and R. Ryf, "Complexity analysis of adaptive frequency-domain equalization for MIMO-SDM transmission," in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*, Sept 2013, pp. 1–3.

[29] T. Fehenberger, M. P. Yankov, L. Barletta, and N. Hanik, "Compensation of XPM interference by blind tracking of the nonlinear phase in WDM systems with QAM input," in *2015 European Conference on Optical Communication (ECOC)*, Sept 2015, pp. 1–3.

[30] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, April 2009.

[31] X. Zhou, "An improved feed-forward carrier recovery algorithm for coherent receivers with M-QAM modulation format," *IEEE Photonics Technology Letters*, vol. 22, no. 14, pp. 1051–1053, July 2010.

[32] M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, T. Pfau, M. Pepe, M. Bertolini, and G. Gavioli, "Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers," *IEEE Photonics Technology Letters*, vol. 24, no. 9, pp. 739–741, May 2012.

[33] H. Cheng, Y. Li, M. Yu, J. Zang, J. Wu, and J. Lin, "Experimental demonstration of pilot-symbols-aided cycle slip mitigation for QPSK modulation format," in *OFC 2014*, March 2014, pp. 1–3.

[34] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A simple and effective closed-form GN model correction formula accounting for signal non-gaussian distribution," *Journal of Lightwave Technology*, vol. 33, no. 2, pp. 459–473, Jan 2015.

[35] A. Carena, G. Bosco, V. Curri, P. Poggiolini, and F. Forghieri, "Impact of the transmitted signal initial dispersion transient on the accuracy of the gn-model of non-linear propagation," in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*, Sept 2013, pp. 1–3.

[36] Idealist, "D1.1 elastic optical network architecture: Reference scenario, cost and planning," Tech. Rep., 2015. [Online]. Available: http://www.ict-idealist.eu/index.php/wp1/finish/23-wp1/465-d1-1-elastic-optical-network-architecture-reference-scenario-cost-and-planning

[37] D. Pilori, M. Cantono, A. Carena, and V. Curri, "FFSS: The fast fiber simulator software," in *2017 19th International Conference on Transparent Optical Networks (ICTON)*, July 2017, pp. 1–4.

[38] P. Poggiolini, A. Nespola, Y. Jiang, G. Bosco, A. Carena, L. Bertignono, S. M. Bilal, S. Abrate, and F. Forghieri, "Analytical and experimental results on system maximum reach increase through symbol rate optimization," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1872–1885, April 2016.

[39] M. Cantono, R. Gaudino, and V. Curri, "Potentialities and criticalities of flexible-rate transponders in DWDM networks: A statistical approach," *J. Opt. Commun. Netw.*, vol. 8, no. 7, pp. A76–A85, Jul 2016. [Online]. Available: http://jocn.osa.org/abstract.cfm?URI=jocn-8-7-A76

[40] V. Curri, M. Cantono, and R. Gaudino, "Elastic all-optical networks: A new paradigm enabled by the physical layer. how to optimize network performances?" *J. Lightwave Technol.*, vol. 35, no. 6, pp. 1211–1221, Mar 2017. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-35-6-1211

[41] A. Ferrari, M. Cantono, U. Waheed, A. Ahmad, and V. Curri, "Networking benefits of advanced DSP techniques and hybrid fiber amplification," in *2017 19th International Conference on Transparent Optical Networks (ICTON)*, July 2017, pp. 1–4.