Non-Linear Phase Noise Mitigation over Systems using Constellation Shaping

Non-Linear Phase Noise Mitigation over Systems using Constellation Shaping

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

***Abstract*—This paper presents a modified soft-decoding strategy, which improves performance in the presence of strong phase noise. This can substantially increase the reach of systems that are severely affected by phase noise, generated by fiber non-linear Kerr effect. This strategy is applied to two different experimental scenarios employing constellation shaping, which is known to generate strong non-linear phase noise. In the first experiment, we show that the strategy significantly improves the performance of probabilistically-shaped (PS) 64-QAM over low-dispersion fibers. In the second experiment, the strategy is used to optimize the position of the points of a 32-QAM constellation (geometrical shaping). This optimized constellation is then compared to standard 32-QAM and PS 64-QAM over standard single-mode fiber. Also in this case, the modified strategy is able to give significant reach gains.**

***Index Terms*—Optical communications, non-linear phase noise, probabilistic constellation shaping, geometric constellation shaping.**

I. Introduction

RECENTLY, constellation shaping has been applied to optical communications as a method to further increase transceiver sensitivity [1], [2]. Different techniques are enclosed under the "constellation shaping" definition, and they can be divided into two main groups: probabilistic constellation shaping (PCS) and geometric constellation shaping (GCS). Both of these methods change the shape of standard quadrature amplitude modulation (QAM) constellations, with the purpose of optimizing the constellation to a specific data rate and channel. Consequently, this improves the sensitivity of the transceiver, allowing to reach longer distances at high data rates. Moreover, PCS can also continuously adapt the data-rate, allowing greater flexibility compared to standard QAM constellations [3].

The coherent optical channel is well approximated by the additive white Gaussian noise (AWGN) channel [4], for which the optimal transmit constellation is the Gaussian constellation. Since it requires an infinite set of transmitted symbols, such a constellation cannot be realized in practice. Both PCS and GCS try to approximate a Gaussian constellation using a finite set of points. For instance, the Maxwell-Boltzmann probability distribution [5] was found to be optimal in terms of maximization of the mutual information (MI), when PCS is applied to a standard QAM constellation over an AWGN channel.

While the AWGN approximation is valid for a wide range of long-haul optical communications scenarios, there are situations where this approximation is not accurate. For example, single-span systems [6], transmitters using low symbol rates [7] or transmission over low-dispersion fiber [2] are known to generate non-AWGN non-linear interference (NLI). In those cases, a significant portion of NLI is in the form of phase and polarization rotation noise (PPRN) [8], [9], which changes the statistics of the transmission channel. This effect is stronger when using constellation shaping, since Gaussian-like constellations inherently generate more NLI [8], [10]. Therefore, the Gaussian constellation is not optimal anymore over such channel.

The main effect of PPRN on the constellation is similar to laser phase noise (PN), and it is called non-linear phase noise (NLPN). This effect can be further divided into two main categories, depending on its memory (length of its auto-correlation function) [11]. Long-autocorrelation NLPN can be almost fully compensated for by standard carrier phase recovery CPE algorithms, which are already employed at the receiver to compensate for laser phase noise. On the other hand, short-autocorrelation NLPN cannot be compensated for by the CPE, and thus it impairs the detection of the received signal. Therefore, this short-autocorrelation NLPN is the main effect that makes the channel statistics differ from AWGN.

A lot of work has been recently done to improve transmission performance in the non-linear optical channel. Most of it focused on the optimization of symbol probabilities (for PCS) or locations (for GCS) in order to *generate less NLI* [6], [12]–[15]. However, the real-world gain of these techniques was found to be small. Part of the reason lies on the fact that the minimization of the power of the total NLI does not allow distinguishing between short-autocorrelation and long-autocorrelation NLPN.

In [16]–[20], the authors used a radically different approach. Instead of preventing NLI generation, they focused on its compensation at the receiver. If the compensation is successful, then the channel becomes AWGN, allowing the use of powerful AWGN-optimized constellations. In [16], the authors designed a phase-tracking algorithm based on the assumption that phase noise is a first-order Wiener process, and each sample of phase noise is distributed according to the Tikhonov (or Von Mises) distribution. A similar approach was followed in [20], where the authors used a Kalman adaptive filter to track and compensate the time-varying intersymbol

D. Pilori and G. Bosco are with DET, Politecnico di Torino, 10129 Torino, Italy. E-mail: dario.pilori@polito.it. A. Nespola is with LINKS foundation, 10138 Torino, Italy. F. Forghieri is with CISCO Photonics S.r.l., 20871 Vimercate (MB), Italy.

This work is an extended version of [22].

interference (ISI) effect of NLI. Both of these methods were found to be effective, and they are not tailored to a specific scenario. However, they require a change in the standard digital signal processing (DSP) chain of a coherent receiver. Moreover, the hardware implementation complexity may be too high to be practically realized. Indeed, to the best of our knowledge, no real-time implementation has been presented.

This work is based on the latter approach, and proposes a modification of the soft-decoding strategy at the receiver, in order to take into account the presence of NLPN, which alters the channel statistics. A similar method has already been successfully applied to wireless communications in [21], assuming a Tikhonov-distributed memoryless phase noise. In [17] another approach was adopted, where a multi-dimensional Gaussian distribution was assumed. In this work, we use the same channel model as [21], applying it to the non-linear optical channel, with a particular focus on systems using constellation shaping.

This paper is organized as follows. In Section II, the theoretical channel model is described, providing the expressions used by the receiver to decode the symbols. Afterwards, the effectiveness of this method is investigated in two different experimental scenarios. In the first scenario, illustrated in Section III, the modified decoding metric is applied to a system using both standard and PCS QAM constellations, propagating over low-dispersion fiber. These results have been partially published in [22]. In the second scenario, shown in Section IV, the method is used to design a geometrically-shaped constellation that is tailored to a channel with NLPN. Conclusions are drawn in Section V.

II. Theoretical Model

Modern coherent receivers for long-haul optical transmission employ pragmatic soft-decision forward error correction (FEC) decoding, also called bit-interleaved coded modulation (BICM) [23], [24]. These receivers perform constellation demapping and binary FEC decoding in two separate steps. The de-mapper calculates bit-wise log-likelihood ratios (LLRs) (also called *L*-values) after estimating channel statistics, usually under the AWGN assumption. Since standard CPE algorithms (such as blind phase search (BPS)) are able to remove most of the long-autocorrelation NLPN [11], residual phase noise after CPE can be assumed *memoryless*. Any residual memory can be suppressed using a large symbol interleaver [25]. With this key approximation, LLRs can be computed by taking into account the presence of this residual phase noise, potentially improving decoding performance, without substantially changing the receiver structure.

In this Section, we derive the expressions of the LLRs which take into account NLPN. To perform this, a simple channel model, based on BICM, is introduced. Detailed analyses on this scheme are out of the scope of this work, therefore we refer the interested reader to [24], [25] (and references therein).

A. Information-theoretical channel model

A block diagram of the channel model used in this section is shown in Fig. 1. The binary FEC encoder encodes information bits to channel bits. A block of m channel bits $\boldsymbol{b}_k = \begin{bmatrix} b_{1,k} & b_{1,k} & \dots & b_{m,k} \end{bmatrix}$ is mapped into a complex symbol a_k drawn from a constellation $\mathcal{C}$ with cardinality $|\mathcal{C}| = 2^m$. In these equations, k represents the time index.

A sequence of complex symbols is then transmitted over the "optical channel" block, which is composed by the optical transmitter, the channel and the receiver (including adaptive equalization and phase recovery), thereby obtaining a received symbol y_k. Assuming that the optical channel is memoryless, i.e. the CPE is able to compensate for most of the long-autocorrelation phase noise, it can be fully described by its conditional probability $p(y|a)$, where y and a are scalars. Moreover, assuming that the channel is stationary, the time index k can be dropped without any loss of generality.

The demapper then takes the received symbol y and generates a $m \times 1$ vector $\boldsymbol{l}$ of bit-wise LLRs. They are real numbers that represent the reliability of the estimation of the binary representation $\boldsymbol{b}_k$. Large positive (or small negative) values of l mean that the estimation of that bit is reliable, while values of l close to zero mean a less reliable estimation. LLRs are used by the binary soft FEC decoder to estimate the information bits.

LLRs are calculated using a deterministic function of y, which depends on the channel transition probability $p(y|a)$

$$l_i = \log \frac{P(b_i = 1|y)}{P(b_i = 0|y)} = \log \frac{\sum_{a_1 \in \mathcal{C}_i^1} p(y|a_1) P(a_1)}{\sum_{a_0 \in \mathcal{C}_i^0} p(y|a_0) P(a_0)} \tag{1}$$

In this equation, $i = 1, \dots, m$ is the bit index, and $\mathcal{C}_i^b$ is the set of constellation symbols whose i-th bit is equal to b. $P(a)$ is the a-priori probability to transmit symbol a. If no PCS is applied, then all transmitted symbols are equiprobable, i.e. $P(a) = 2^{-m}$ for all the symbols.

In the next sub-sections, expressions of (1) are provided for both the AWGN channel and the AWGN channel with phase noise.

B. AWGN conditional probability

In an AWGN channel, the received sample y_k of Fig. 1 is expressed as:

$$y_k = a_k + n_k \tag{2}$$

where n_k is an additive white Gaussian noise sample, which is a combination of amplified spontaneous emission (ASE) noise and NLI, with variance σ_n^2. The conditional probability $p(y|a)$ is the circularly symmetric Gaussian probability density function (PDF), centered around a

$$p(y|a) = \frac{1}{\pi \sigma_n^2} \exp\left(-\frac{|y-a|^2}{\sigma_n^2} \right) \tag{3}$$

This expression is then substituted into (1) to calculate the LLRs. In this paper, this detection technique will be briefly called "AWGN".

C. Conditional Probability with NLPN

The presence of phase noise modifies (2) into

$$y_k = a_k e^{j\phi_k} + n_k \tag{4}$$

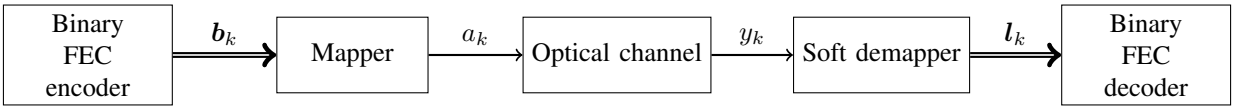


Fig. 1. Information-theoretical system model. The optical channel, which includes transmitter and receiver DSP – assumed memoryless – is fully described by the probability density function (PDF) $p(y|a)$.

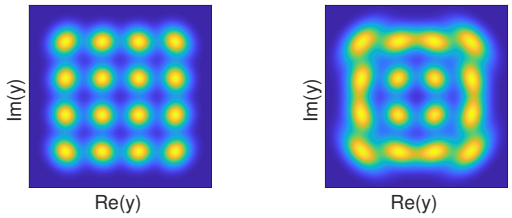


Fig. 2. Probability of the complex-valued received sample y computed using (7) for 16-QAM with $\sigma_n^2 = 0.25$, $\kappa_\phi = 205$ (a) and $\kappa_\phi = 33$ (b).

where ϕ_k represents residual phase noise. As stated at the beginning of this Section, ϕ_k is assumed to be a memoryless process, i.e. ϕ_k is statistically independent from ϕ_n for $k \neq n$.

Then, an appropriate probability distribution must be chosen for ϕ. Since it is a phase, it is bounded between $-\pi$ and π, therefore distributions with unbounded support (like the Gaussian distribution) are not suitable. In the literature, a well known distribution for random phases is the Tikhonov (or von Mises) probability distribution [26]:

$$p(\phi) = \frac{e^{\kappa_\phi \cos\phi}}{2\pi I_0(\kappa_\phi)} \qquad \phi \in (-\pi, \pi] \tag{5}$$

where the parameter κ_ϕ is called concentration, and $I_0(.)$ is the modified Bessel function of the first kind. This distribution is an approximation of a Gaussian distribution wrapped between $-\pi$ and π. Moreover, for large values of concentration (i.e. small phase noise), the Tikhonov distribution resembles a zero-mean Gaussian distribution with variance $\sigma_\phi^2 \approx 1/\kappa_\phi$.

With those assumptions, the channel conditional probability can be expressed as [21]

$$p(y|a) \approx \sqrt{\frac{\kappa_\phi}{8\pi^3}} \frac{e^{-\kappa_\phi}}{\sigma_n^2} \exp\left(-\frac{|y|^2 + |a|^2}{2\sigma_n^2} + \left|\frac{ya^*}{\sigma_n^2} + \kappa_\phi\right|\right) \tag{6}$$

where the modified Bessel function of the first kind has been approximated as $I_0(x) \approx e^x/\sqrt{2\pi x}$ (valid for large values of x). We found that this approximation is accurate for the values of κ_ϕ encountered in the experiments.

This expression can then be then substituted in (1) to evaluate LLRs for this channel. The receiver just needs to estimate, other than AWGN variance σ_n^2 (which is also required in standard AWGN receivers), an additional parameter κ_ϕ. In this paper, this detection technique will be called "PN-aware".

Fig. 2 shows two examples of the PDF of the received signal $p(y)$. The constellation is 16-QAM, with an AWGN variance $\sigma_n^2 = 0.25$ and two different concentrations of phase noise: $\kappa_\phi = 205$ (a) and $\kappa_\phi = 33$ (b). The probability of y has been calculated from (6) using the law of total probability

$$p(y) = \sum_{a_i \in \mathcal{C}} P(a_i) p(y|a_i) \tag{7}$$

Smaller values of concentration (i.e. larger phase noise) change the shape of the PDF, especially in the outer points, which become similar to "rain drops". On the other hand, large values of concentration make the PDF similar to the AWGN PDF (3).

D. Performance metrics

In an optical communication system, the most important performance metric is the post-FEC bit error ratio (BER). For high-capacity long-haul transmissions, it must usually be below 10^{-15}. To perform a reliable post-FEC BER measurement, a very large number of codewords needs to be transmitted, which is extremely time-consuming, even for real-time field programmable gate array (FPGA) receivers.

To avoid performing this operation, a popular approach is the estimation of the so-called achievable information rate (AIR) metrics, which allow to infer post-FEC BER performance without performing the actual decoding [27]. For BICM receivers, a widely used metric is the generalized mutual information (GMI), which can be easily evaluated from the LLRs using Monte-Carlo integration [24, eq. (35)]:

$$\mathrm{GMI} \approx H(\mathcal{C}) - \frac{1}{N} \min_{s \geq 0} \sum_{i=1}^{m} \sum_{k=1}^{N} \log_2\left(1 + e^{s(-1)^{b_{i,k}} l_{i,k}}\right) \tag{8}$$

where $H(\mathcal{C}) = -\sum_{a_i \in \mathcal{C}} P(a_i) \log_2 P(a_i)$ is the entropy of the transmitted constellation, N the number of symbols and s an optimization parameter. The optimization over such parameter is mandatory whenever the exact channel law is unknown [24].

To compare standard QAM modulation formats with PCS QAM, assuming the use of the probability amplitude shaping (PAS) architecture introduced in [28], the normalized GMI (NGMI) [29] is often used. The NGMI of a standard 2^m-QAM constellation is

$$\mathrm{NGMI} = \frac{\mathrm{GMI}}{m} \tag{9}$$

while the NGMI of a PS 2^m-QAM constellation with entropy $H(\mathcal{C})$ is

$$\mathrm{NGMI} = 1 - \frac{H(\mathcal{C}) - \mathrm{GMI}}{m} \tag{10}$$

By fixing the same NGMI threshold, different constellations, with and without PCS, can be fairly compared [29].

III. Experimental Demonstration over Low-Dispersion Fibers

Low-dispersion fibers – e.g. G.653 or G.655 – are known to generate strong short-autocorrelation NLPN; this effect can be explained by the time-domain pulse-collision theory [8]. Therefore, the modified decoding metric of (6) can be beneficial for signals transmitted over those legacy fibers,

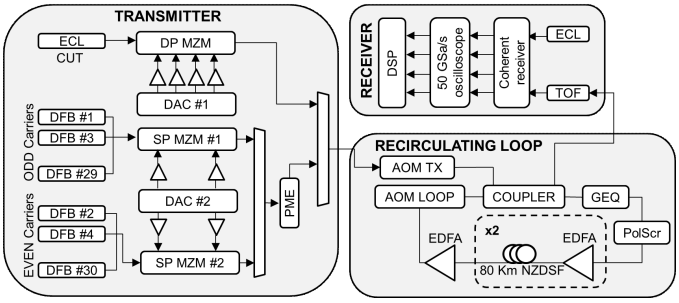


Fig. 3. Experimental setup for the NZDSF experiment.

TABLE I
GMI THRESHOLDS FOR AN NGMI THRESHOLD OF 0.9

Constellation	Entropy (bit/symbol)	GMI threshold (bit/symbol)	Data rate (Gbit/s)
16-QAM	4	3.6	106.6
32-QAM	5	4.5	133.3
PS 64-QAM-1	$13/3 \approx 4.33$	$56/15 \approx 3.73$	106.6
PS 64-QAM-2	$31/6 \approx 5.17$	$137/30 \approx 4.57$	133.3

which are still widely installed and deployed in several countries, like Brasil [30] and Japan [31]. In order to assess the gain that can be achieved by using the modified metric (6), we ran an experiment over G.655 non-zero dispersion-shifted fiber (NZDSF), comparing the performance of PS 64-QAM constellations with (uniform) 16-QAM and 32-QAM at the same net data rate. The experimental setup and the transmitted constellations are the same as in [2, Sec. III.G]. Differently from [2], the use of bit-wise decoding is assumed and consequently the GMI is used as a performance metric.

A. Experimental setup

The experimental setup is shown in Fig. 3. Two four-channel digital-to-analog converters (DACs) generate the channel under test (CUT) and 30 interfering wavelength division multiplexing (WDM) channels at 16 GBaud around 1558 nm. As modulation format, four polarization-multiplexed (PM) constellations have been tested: 16-QAM, 32-QAM and two different PS 64-QAM (-1 and -2). The parameters of the four constellations are summarized in Table I. Using these parameters, 16-QAM and PS 64-QAM-1, 32-QAM and PS 64-QAM-2 achieve the same net (i.e. post-FEC) data-rate, assuming a soft-decision FEC with an overhead of 20%. The PS 64-QAM constellations have been obtained with the exponential probability mass function [2, Eq. (1)], whose parameters have been chosen to obtain the entropy shown in Table I.

An integrated coherent receiver, connected to a four-channel 50 Gs/s real-time oscilloscope, detected the CUT. Receiver offline DSP was made of a 24-tap fractionally-spaced least-mean-squares (LMS) adaptive equalizer, followed by a pilot-aided BPS-maximum likelihood (ML) phase recovery [2], [32]. BPS used 18 test phases, and pilot symbols (1%) were used only to aid phase unwrapping. The memory of the moving average was optimized for every received waveform.

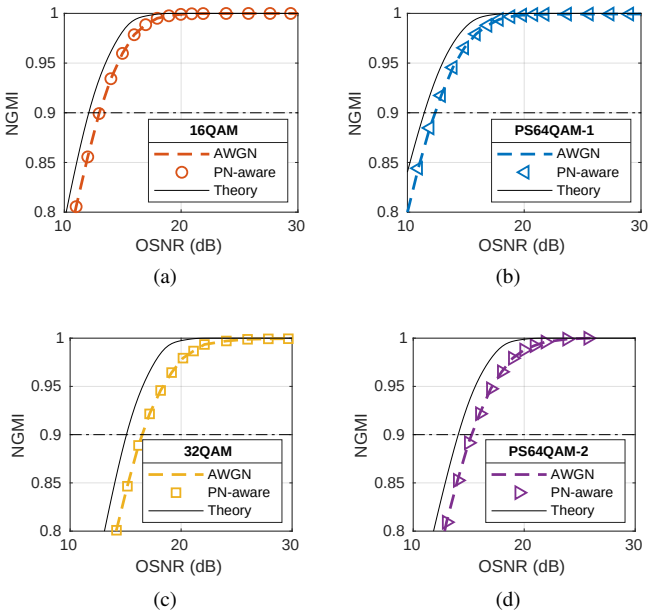


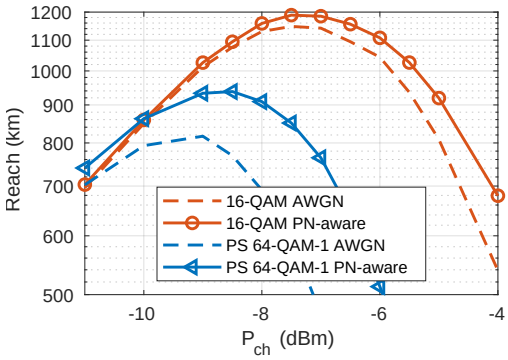
Fig. 4. Optical back-to-back results, in terms of NGMI, for the four considered constellations as a function of the OSNR (normalized to the symbol rate).

The LLRs are calculated by the decoder using (1), using both AWGN (3) and PN-aware (6) detection. Afterwards, the GMI is evaluated from the LLRs using Monte-Carlo integration [24]. We chose an NGMI threshold of 0.9, which gives different GMI thresholds for all constellations. These thresholds are shown in Table I. In these experiments, the optimal values of concentration κ_ϕ and variance of amplitude noise σ_n^2 were found, for each waveform, by maximizing the GMI.

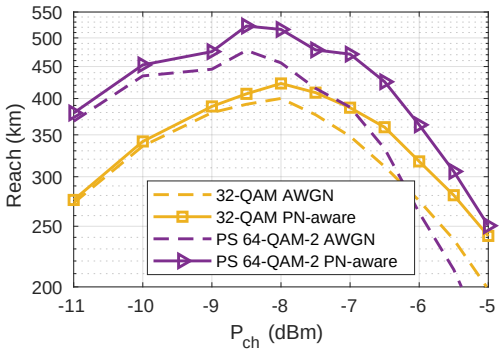
B. Back-to-back results

The performance of the four constellations in optical back-to-back is shown in Fig. 4, in terms of NGMI as a function of the optical signal-to-noise ratio (OSNR), where the noise has been calculated over a bandwidth equal to the symbol rate. Dashed lines have been obtained using the AWGN decoder, whilst markers have been calculated with the PN-aware decoding metric. Results are compared with the theoretical curves (solid black lines). Since no other sources of phase noise are present, the proposed algorithm gives no improvement in back-to-back. This suggests that the phase noise introduced by the transmitter and receiver lasers (< 100 kHz linewidth) is almost fully compensated for by the CPE. This hypothesis is further justified by the fact that the measured non-circularity index (NCI) [11, Sec. IV-A] was always lower or equal than 0.04 dB at the NGMI threshold for all constellations.

At the NGMI threshold of 0.9, shown as a black dashed-dotted line in Fig. 4, back-to-back penalties with respect to the theoretical performance are approximately 0.9 dB for 16-QAM and PS 64-QAM-1 and 1.4 dB for 32-QAM and PS 64-QAM-2.

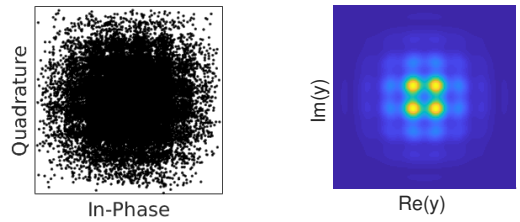


(a)



(b)

Fig. 5. Propagation results over NZDSF. In each plot, the constellations achieve the same net (post-FEC) data-rate: 106.6 Gbit/s (a) and 133.3 Gbit/s (b). The NGMI threshold is equal to 0.9 for all formats.



(a) (b)

Fig. 6. Recovered constellation of PS 64-QAM-1 at its maximum reach over NZDSF (a). Its PDF, estimated using (7) with $\sigma_n^2 = 0.35$ and $\kappa_\phi = 72.6$, is shown in (b).

TABLE II
FIBER PARAMETERS

Fiber	α (dB/km)	β_2 (ps^2/km)	γ (1/(W km))
SMF (G.652)	0.20	-21.27	1.3
NZDSF (G.655)	0.23	3.38	2

C. Propagation results

The WDM signal was transmitted over a recirculating fiber loop, shown in Fig. 3, made of two 80-km spans of NZDSF fiber with EDFA amplification. The main parameters of the fiber are summarized in Table II. The performance results are shown in Fig. 5 as maximum reach (at the NGMI threshold of 0.9) for different per-channel launch power values. From Fig. 5a it can be noticed that the PS constellation has a shorter

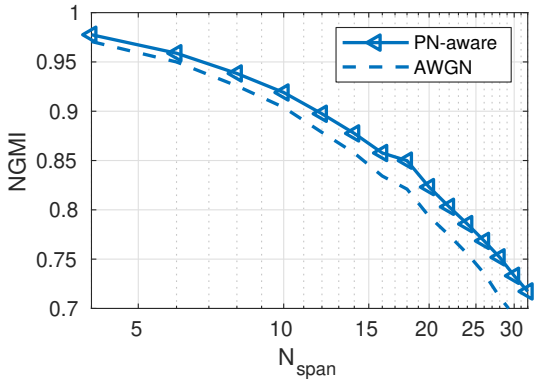


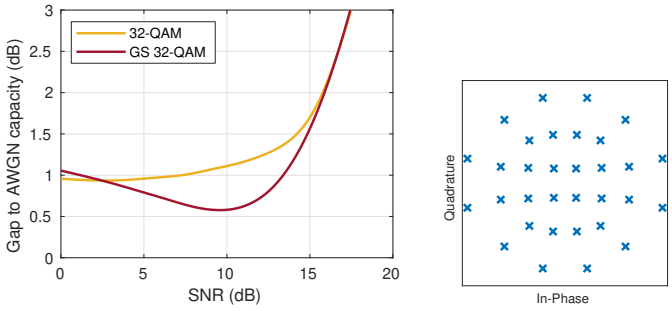
Fig. 7. NGMI as a function of the number of spans, measured at the optimal launch power, for PS 64-QAM-1 over 80-km spans of NZDSF.

reach than 16-QAM. This contrasts with the back-to-back results, which showed that PS 64-QAM constellations have a better sensitivity with respect to 16-QAM and 32-QAM. In that scenario, NLPN is so strong that the sensitivity gain of PS is overcome, even using PN-aware decoding. Moreover, there may be an additional penalty, due to a suboptimal performance of the BPS phase recovery algorithm, which has been shown to have a penalty when applied to PS constellations [33]. Nevertheless, PN-aware decoding is able to substantially increase its reach by 14.1% compared to AWGN decoding, while the reach increase with 16-QAM is only 4%. A constellation diagram of PS 64-QAM-1 at its maximum reach is shown in Fig. 6, together with the corresponding PDF, estimated using (7). In (b), the difference between 32-QAM and PS 64-QAM-2 is smaller, and the reach increase due to PN-aware decoding is equal to 5.6% and 9.9%, respectively. This difference is caused also by the absence of an exact Gray mapping on 32-QAM, which reduces its sensitivity with respect to 64-QAM.

The higher gain obtained by applying the modified metric to PS constellations is related to the fact that a larger amount of NLPN is generated during propagation for Gaussian-like constellations. For the same reason, the gain of the modified detection strategy increases with the propagation distance, as shown in Fig. 7, where the NGMI at the optimum launch power is shown for the PS 64-QAM-1 constellation as a function of the number of spans. PN-aware decoding gives a better NGMI over AWGN decoding at every distance. The gain is slightly increasing with distance, due to the generation of additional NLPN.

D. Discussion and conclusion

As expected from the theory, propagation of PS constellations over NZDSF generates strong short-autocorrelation NLPN, which cannot be fully compensated by receiver CPE. This is particularly evident when PCS constellations are compared, at the same net data rate, with lower-cardinality uniform QAM constellations [2]. The use of the PN-aware decoding (6) is able to partially mitigate this effect, without increasing significantly the receiver complexity. Nevertheless, it is not able to fully compensate for NLPN. To achieve this effect,



(a) (b)

Fig. 8. Optimization of 32-QAM. Gap to capacity assuming a BICM-AWGN channel (a), and scatter diagram of the optimized constellation (b).

TABLE III
GMI THRESHOLDS FOR AN NGMI THRESHOLD OF 0.86

Constellation	Entropy (bit/symbol)	GMI threshold (bit/symbol)	Data rate (Gbit/s)
32-QAM	5.0	4.30	128
PS 64-QAM-3	5.2	4.36	128

more powerful (and, possibily, complex) DSP schemes need to be designed.

IV. EXPERIMENTAL DEMONSTRATION WITH GEOMETRICAL SHAPING

GCS is a constellation shaping technique that optimizes the position (and the bit mapping) of constellation points, according to a given metric. The metric can be arbitrarily chosen; therefore, by choosing an appropriate metric that takes into account Kerr effects, a maximum-reach gain can be obtained.

GCS has been widely applied to the non-linear fiber-optic channel [6], [12]–[15]. However, as discussed in the introduction, these methods are mostly focused on generating less NLI. In this example, following [21], we optimize the constellation using the PN-aware metric (6). In this case, assuming that the variance of NLPN does not significantly change by "moving" constellation points, this operation improves the performance of the PN-aware decoding described in Section II. In particular, we used the simulated annealing algorithm [34] to optimize constellation points, given the additive noise variance and the Tikhonov PN concentration. More details on this procedure can be found in [21, Sec. II].

A. Optimization of 32-QAM

As an example, we ran the optimization algorithm over a 32-QAM constellation, obtaining the results shown in Fig. 8. The constellation has been optimized with the simulated annealing algorithm over a channel with a fixed SNR of 13 dB (considering only additive noise) and a Tikhonov phase noise with concentration $\kappa_\phi = 637.1$. These two values were heuristically chosen among a large set of SNRs and concentrations, and the resulting constellation was found to have a good back-to-back and propagation performance.

A BICM receiver is assumed, i.e. the adopted performance metric was the GMI, and the LLRs were calculated using the PN-aware metric of (6). With such scheme, 32-QAM has a penalty with respect to square QAM constellations (e.g. 64-QAM), due to the absence of an exact Gray bit-to-symbol labeling. The resulting constellation is shown in Fig. 8b, and its performance over an AWGN-BICM channel is shown in Fig. 8a. To aid visualization, results have been quoted as gap to the AWGN capacity, which is defined as

$$\mathrm{SNR}_{\mathrm{dB}} - 10\log_{10}\left(2^{\mathrm{GMI}} - 1\right) \tag{11}$$

where the SNR is the ratio between the variance of the signal and the variance of the additive noise in (3).

The largest SNR gain is ~ 0.5 dB, which is obtained for an SNR of approximately 10 dB. Unfortunately, as it will be shown later, this SNR is too low for the considered NGMI thresholds. Nevertheless, this constellation gives a non-negligible shaping gain across a wide range of SNR. We remark that the gain of GCS, albeit smaller than PCS, is obtained without requiring a distribution matcher and a proper coded modulation scheme.

B. Experimental setup

The effectiveness of the PN-optimized GS 32-QAM was experimentally tested. The experimental setup is identical to Fig. 3, where the loop fiber was replaced with 4×80-km spans of G.652 SMF, whose parameters are summarized in Table II. The constellations under test were three: 32-QAM (unshaped), GS 32-QAM and PS 64-QAM-3, with entropy 5.2 bit/symb. This value was chosen to obtain the same net data rate of 32-QAM (both shaped and unshaped) with a $4/5$-rate FEC, corresponding to a 25% overhead. An NGMI threshold of 0.86 was used for all the constellations under test, which is lower than the one used in Sec. III, since the FEC overhead is higher. The corresponding GMI thresholds are reported in Table III. The theoretical SNR gains, measured at the threshold with respect to 32-QAM, are 0.37 dB and 1 dB for GS 32-QAM and PS 64-QAM-3, respectively. As in Sec. III, PS 64-QAM-3 has been obtained with the exponential probability mass function.

C. Results and discussion

The experimental results are shown in Fig. 9 and Fig. 10. Fig. 9 shows the back-to-back results, expressed as NGMI as a function of the OSNR, normalized to the symbol rate. The NGMI threshold (0.86) is shown as a black dashed line. Like in Fig. 4, AWGN and PN-aware decoders give the same result in back-to-back. Therefore, only the results with the AWGN decoder are shown. As expected from theory (Fig. 8), GS 32-QAM has a slight advantage compared to 32-QAM, while PS 64-QAM-3 has the largest gain, thanks to the higher number of constellation points.

Propagation results, expressed as maximum reach (at the NGMI threshold) for different per-channel optical launch powers, are shown in Fig. 10. Solid lines (and markers) have been obtained with PN-aware decoding, while dashed lines have been obtained with AWGN decoding. Comparing the three constellations, there is a noticeable reduction of the optimal launch power with PS 64-QAM-3. This means that

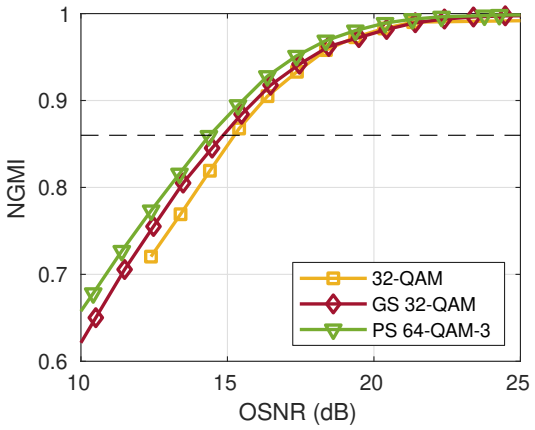


Fig. 9. Back-to-back measurements of GS 32-QAM, compared to 32-QAM and PS 64-QAM-3. All constellations have a net data rate of 128 Gbit/s.

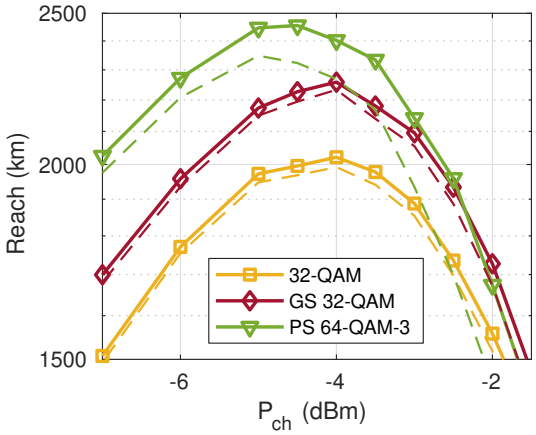


Fig. 10. Experimental results of GS 32-QAM, compared to 32-QAM and PS 64-QAM-3, over 80-km spans of SMF. Solid lines represent the results obtained with the PN-aware decoder, while dashed lines represent the results obtained with the AWGN decoder. All constellations have a net data rate of 128 Gbit/s.

NLPN is affecting this constellation, reducing its maximum reach. This effect can be also seen by looking at the dashed lines (AWGN decoding), where the gain of PN-aware decoding over AWGN decoding is noticeably larger with PS 64-QAM-3.

Maximum-reach gains with respect to 32-QAM were $+11.7\%$ and $+21.0\%$ for GS 32-QAM and PS 64-QAM-3, respectively. These gains can be converted into a sensitivity gain using [35, (29)], assuming that all NLI is generated according to the GN-model. A reach gain of $+11.7\%$ corresponds to a sensitivity gain of 0.48 dB, while a gain of $+21.0\%$ corresponds to 0.83 dB. Comparing these gains with the theoretical gains over a pure AWGN channel (0.37 dB and 1 dB for GS 32-QAM and PS 64-QAM-3, respectively), the PS constellation has a penalty, while the GS constellation has a *gain*. This is due to its higher resilience to NLPN, which was not taken into account in the AWGN results of Fig. 8.

V. Conclusion

In this paper, we presented a modified soft-decoding strategy that improves the performance in the presence of short-autocorrelation non-linear phase noise. We tested this strategy over two experimental scenarios. In the first scenario, the metric improved the performance of PS 64-QAM over low-dispersion fibers. In the second scenario, the modified strategy was used to optimize a 32-QAM constellation, which reduced the amount of NLPN over standard single-mode fiber.

Acknowledgments

The authors would like to thank G. Montorsi and P. Poggiolini for useful discussions. This work has been partially supported by CISCO Photonics and the PhotoNext center at Politecnico di Torino.

References

[1] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *Journal of Lightwave Technology*, vol. 34, no. 7, pp. 1599–1609, 4 2016.

[2] D. Pilori, L. Bertignono, A. Nespola, F. Forghieri, and G. Bosco, "Comparison of probabilistically shaped 64qam with lower cardinality uniform constellations in long-haul optical systems," *Journal of Lightwave Technology*, vol. 36, no. 2, pp. 501–509, 1 2018.

[3] G. Bosco, "Advanced modulation techniques for flexible optical transceivers: the rate/reach tradeoff," *Journal of Lightwave Technology*, vol. 37, no. 1, pp. 36–49, jan 2019.

[4] 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, 2 2014.

[5] F. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for gaussian channels," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 913–929, 5 1993.

[6] O. Geller, R. Dar, M. Feder, and M. Shtaif, "A shaping algorithm for mitigating inter-channel nonlinear phase-noise in nonlinear fiber systems," *Journal of Lightwave Technology*, vol. 34, no. 16, pp. 3884–3889, 8 2016.

[7] D. Pilori, F. Forghieri, and G. Bosco, "Residual non-linear phase noise in probabilistically shaped 64-QAM optical links," in *Optical Fiber Communication Conference (OFC)*. OSA, 2018.

[8] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications," *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 593–607, 1 2016.

[9] R. Dar and P. Winzer, "Nonlinear interference mitigation: methods and potential gain," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 903–930, 2 2017.

[10] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Optics Express*, vol. 22, no. 13, p. 16335, 6 2014.

[11] 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, 2 2017.

[12] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "On shaping gain in the nonlinear fiber-optic channel," in *IEEE International Symposium on Information Theory (ISIT)*. IEEE, 6 2014.

[13] J. Cho, S. Chandrasekhar, R. Dar, and P. J. Winzer, "Low-complexity shaping for enhanced nonlinearity tolerance," in *European Conference on Optical Communications (ECOC)*, 2016.

[14] E. Sillekens, D. Semrau, G. Liga, N. A. Shevchenko, Z. Li, A. Alvarado, P. Bayvel, R. I. Killey, and D. Lavery, "A simple nonlinearity-tailored probabilistic shaping distribution for square QAM," in *Optical Fiber Communication Conference (OFC)*. OSA, 2018.

[15] E. Sillekens, D. Semrau, D. Lavery, P. Bayvel, and R. I. Killey, "Experimental demonstration of geometrically-shaped constellations tailored to the nonlinear fibre channel," in *European Conference on Optical Communications (ECOC)*, 2018.

[16] M. P. Yankov, T. Fehenberger, L. Barletta, and N. Hanik, "Low-complexity tracking of laser and nonlinear phase noise in WDM optical fiber systems," *Journal of Lightwave Technology*, vol. 33, no. 23, pp. 4975–4984, 12 2015.

[17] T. A. Eriksson, T. Fehenberger, P. A. Andrekson, M. Karlsson, N. Hanik, and E. Agrell, "Impact of 4D channel distribution on the achievable rates in coherent optical communication experiments," *Journal of Lightwave Technology*, vol. 34, no. 9, pp. 2256–2266, 5 2016.

[18] M. P. Yankov, F. D. Ros, E. P. da Silva, T. Fehenberger, L. Barletta, D. Zibar, L. K. Oxenløwe, M. Galili, and S. Forchhammer, "Non-linear phase noise compensation in experimental WDM systems with 256QAM," *Journal of Lightwave Technology*, vol. 35, no. 8, pp. 1438–1443, 4 2017.

[19] M. P. Yankov, L. Barletta, and D. Zibar, "Phase noise compensation for nonlinearity-tolerant digital subcarrier systems with high-order QAM," *IEEE Photonics Journal*, vol. 9, no. 5, pp. 1–12, 10 2017.

[20] O. Golani, M. Feder, and M. Shtaif, "Kalman-MLSE equalization for NLIN mitigation," *Journal of Lightwave Technology*, vol. 36, no. 12, pp. 2541–2550, 6 2018.

[21] F. Kayhan and G. Montorsi, "Constellation design for memoryless phase noise channels," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2874–2883, 5 2014.

[22] D. Pilori, A. Nespola, P. Poggiolini, F. Forghieri, and G. Bosco, "Low-complexity non-linear phase noise mitigation using a modified soft-decoding strategy," in *Optical Fiber Communication Conference (OFC)*, vol. M1I.2. OSA, 2019.

[23] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, 5 1998.

[24] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems, "Achievable information rates for fiber optics: Applications and computations," *Journal of Lightwave Technology*, vol. 36, no. 2, pp. 424–439, 1 2018.

[25] E. Agrell and M. Secondini, "Information-theoretic tools for optical communications engineers," in *IEEE Photonics Conference (IPC)*. IEEE, 9 2018.

[26] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein, "On the selection of a two-dimensional signal constellation in the presence of phase jitter and gaussian noise," *Bell System Technical Journal*, vol. 52, no. 6, pp. 927–965, 7 1973.

[27] 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, 1 2016.

[28] 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, 12 2015.

[29] J. Cho, L. Schmalen, and P. J. Winzer, "Normalized generalized mutual information as a forward error correction threshold for probabilistically shaped QAM," in *European Conference on Optical Communication (ECOC)*. IEEE, 9 2017.

[30] B. Clesca, P. Perrier, H. Fevrier, D. il Chang, S. Burtsev, H. de Pedro, and W. Pelouch, "Field deployment of advanced photonic technologies for ultra-high bit rate and ultra-long reach terrestrial WDM transmission in brazil," in *Asia Communications and Photonics Conference (ACP)*. OSA, 2014.

[31] M. Fukui, "Influence of dispersion fluctuation of installed dispersion-shifted fiber on four-wave mixing induced degradation in WDM transmission system," in *European Conference on Optical Communications (ECOC)*. IEE, 1997.

[32] 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, 7 2010.

[33] D. A. A. Mello, F. A. Barbosa, and J. D. Reis, "Interplay of probabilistic shaping and the blind phase search algorithm," *Journal of Lightwave Technology*, vol. 36, no. 22, pp. 5096–5105, 2018.

[34] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 5 1983.

[35] V. Curri, A. Carena, A. Arduino, G. Bosco, P. Poggiolini, A. Nespola, and F. Forghieri, "Design strategies and merit of system parameters for uniform uncompensated links supporting nyquist-WDM transmission," *Journal of Lightwave Technology*, vol. 33, no. 18, pp. 3921–3932, 9 2015.