Probabilistic Shaping over Multi-Dimensional Constellations for Optical Fiber Transmissions: Trade-offs and Insights

Jingtian Liu
Télécom Paris, Institut Polytechnique de Paris
19 place Marguerite Perey, 91120 Palaiseau, FRANCE
jingtian.liu@telecom-paris.fr

Élie Awwad
Télécom Paris, Institut Polytechnique de Paris
19 place Marguerite Perey, 91120 Palaiseau, FRANCE
elie.awwad@telecom-paris.fr

Yves Jaouën
Télécom Paris, Institut Polytechnique de Paris
19 place Marguerite Perey, 91120 Palaiseau, FRANCE
yves.jaouen@telecom-paris.fr

***Abstract*—In this work, we combine probabilistic constellation shaping (PCS) using a Maxwell Boltzmann distribution with multi-dimensional modulations with the purpose of increasing performance gains in both the linear and non-linear regimes of optical fiber transmission systems. We particularly study the case of polarization division multiplexed long-haul coherent systems. By applying set partitioning and energy constraints on four-dimensional (4D) modulations followed by probabilistic shaping, we give insights into constructing novel modulations with enhanced performance compared to conventional PCS constellations in which shaping is performed separately over each real dimension.**

***Index Terms*—Probabilistic constellation shaping, multi-dimensional constellations, coherent transmission systems, modulation and coding schemes, non-linear Kerr effect.**

I. Introduction

Extensive studies have shown that high spectral efficiency is achieved by using polarization division multiplexing with M-ary quadrature amplitude modulation (PDM M-QAM) formats and forward error correction (FEC) codes over optical fiber transmission systems [1]–[3]. However, this comes at the cost of a reduced transmission distance because of the accumulation of both amplified spontaneous emission (ASE) noise and Kerr-induced nonlinear distortions [4]. Constellation shaping is one way of approaching the Shannon limit over an additive white Gaussian noise (AWGN) channel. In [5], Kschischang and Pasupathy showed that the probability distribution that maximizes the achievable information rate over an AWGN channel, for a fixed average energy constraint, is the Maxwell-Boltzmann (MB) distribution. If a constellation symbol $\mathbf{r}$ with energy $\|\mathbf{r}\|^2$, it will be chosen with a probability proportional to $\exp\left(-\lambda\|\mathbf{r}\|^2\right)$ where λ is a non-negative parameter governing the trade-off between the bit rate and the average energy.

To apply an MB distribution, probabilistic constellation shaping (PCS) based on constant composition distribution matching (CCDM) [6] was suggested as a competitive candidate to implement PCS over rectangular or square M-QAM, showing near-optimal linear shaping gains [7]. However, the increased peak-to-average power ratio (PAPR) and standardized moments of PCS schemes will trigger more non-linear Kerr-induced distortions compared to unshaped uniform modulations.

More recently, several versions of PCS and distribution matching (DM) were suggested to either further enhance the performance compared to CCDM (in terms of rate loss or robustness to channel effects) or to make the shaped signal better suited to the conventional DSP chain while reducing the implementation complexity. For instance, truncated PCS based on PDM-64QAM was designed by suppressing the maximum-energy symbols to make the scheme better suited to the adaptive multiple-input multiple-output (MIMO) equalizer in [8]. In [9], a multi-dimensional DM was applied to distribute the probability of amplitudes over multiple dimensions and enhance the performance compared to separate 1D distribution matching of amplitudes over the available dimensions. Both techniques offered a better performance than regular PCS constellation shaping.

In this work, we propose to apply PCS over a 4-dimensional set-partitioned modulation defined according to energy constraints to target both enhanced spectral efficiency and better tolerance to non-linear effects. The obtained simulation results inform us that a trade-off must be found between the achievable linear and non-linear gains in order to benefit from this new construction.

II. PCS modulation design over 4D

We start building the modulation scheme with a 4D modulation consisting of two square M-QAM symbols, one for each polarization tributary of the PDM signal. For the sake of low complexity, we can restrict the M-QAM constellations to rectangular or square QAMs defined over $\mathbb{Z}^2$. To increase the minimum squared Euclidean distance (MSED) of these QAM constellations, we apply set partitioning to the constellation.

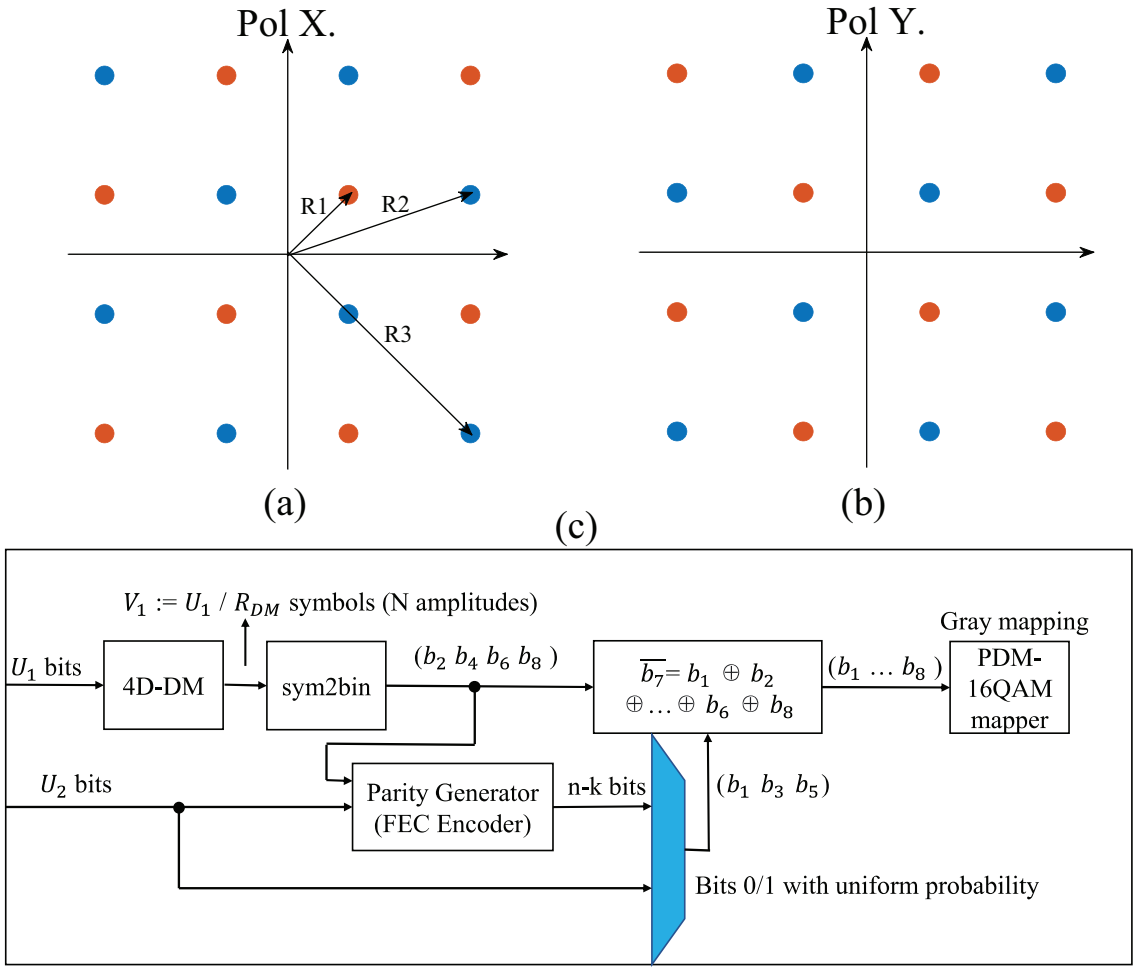

Fig. 1. (a,b) 2D-projections of the 4D 128SP-16QAM in Pol X. (a) and Pol. Y (b). The three arrows represent the three radii values. (c) Block diagram of the PCS 4D CCDM transmitter with the FEC encoder

In this paper, we illustrate the construction with a square 16-QAM ($M = 16$) and define the set-partitioned modulation as follows: a 4D symbol is obtained by combining 2D symbols from two different subsets, shown in different colors in Fig. 1(a-b), and assigning them to two orthogonal polarization tributaries. Each subset contains $8^2 = 64$ 4D symbols. All in all, we get 128 combinations of 4D symbols and the MSED is two-fold increased compared to the one of PDM-16QAM. We call this new constellation 128SP-16QAM which carries $\log_2(128) = 7$ bits per 4D symbol. In the general case of a set-partitioned square M-QAM with a single level of partitioning, the number of bits per 4D symbol will be $2\log_2(M) - 1$.

From this 128SP-16QAM constellation, we can select a subset of symbols by applying constraints either on the desired 4D energy levels or on the desired combinations of 1D energy levels. In the former case, we can reduce the peak-to-average power ratio (PAPR) by removing high energy symbols. In the latter case, we can remove selected 1D-energy combinations to reduce the number of neighbors at the MSED. For instance, following the first rule, we can define a constellation with 88 4D symbols by removing the 4D symbols containing a pair of 2D symbols laying both on the radius R_3 and the ones with a 2D symbol on the radius R_2 and another 2D symbol on the radius R_3 (where R_1, R_2 and R_3 are shown on the 2D projection in Fig.1(a)). This leads to an entropy of $\log_2(88) \approx 6.46$ bits/4D symbol. We call this constellation 88SP-16QAM. On the other hand, following the second rule, we can remove all symbols of an intermediate 4D energy level or alternatively all symbols corresponding to a given combination of 1D amplitudes.

As a next step in the design, we apply PCS over the 4D constellation using a Maxwell-Boltzmann distribution such as the probability of occurrence of the energy values decreases from the lowest to the highest 4D energy. For the distribution matching of the conventional PCS scheme, refered to as PCS 2D in this paper, we choose constant composition distribution matching (CCDM) [6] that involves a low-density-parity check (LDPC) encoder that encodes k input bits to n output bits. For a regular PCS 2D scheme with a square 16-QAM in each polarization, distribution matching can be implemented in parallel over each of the four real dimensions. For each real dimension, the output of the DM has 2 amplitude levels

$\{1,3\}$ to which we associate two probabilities from the desired probability distribution as done in [6]. In [10], an implementation of a PCS transmitter for a single quadrature is described including a 1D distribution matcher and a forward error correction (FEC) encoder. The total net rate of the PCS 2D CCDM transmitter is $R_T = 4\left(U_1 + U_2\right)/V_1 = 4R_{DM} + 8R_C - 4$ bits per 4D symbol where U_1 and U_2 are bit streams at the input of each 1D transmitter, R_{DM} is the rate of the 1D distribution matcher (each 1D DM encodes U_1 bits into $V_1 = U_1/R_{DM}$ symbols), and $R_C = k/n$ is the rate of the embedded FEC encoder that encodes the bits coming from the U_2 bits and the Gray-mapped V_1 symbols. The formula for the total net rate does not take into account the rate loss introduced by the finite-length distribution matching, which for a practical implementation slightly modifies the length of U_1.

Now, we need to design a constant composition distribution matcher and a bit-to-symbol mapping for the 4D shaped 128SP-16QAM constellation and its variants (subsets of this constellation). To achieve this, we consider a bit mapping based on a Gray-mapped PDM-16QAM in order to avoid the use of a look-up-table. We separate the 8 bits of a PDM-16QAM symbol into 4 sign bits (b_1, b_3, b_5, b_7) and four amplitude bits (b_2, b_4, b_6, b_8). In Fig. 1(c), we show the structure of the proposed PCS 4D transmitter that is inspired by the transmitter in [10]. The major difference between the two transmitters is that the initial PCS transmitter for polarization division multiplexed systems consists in 4 parallel 1D PCS transmitters whereas our proposed PCS 4D transmitter performs distribution matching directly over the 4 dimensions. The U_1 bits generate $V_1 = U_1/R_{DM,4D}$ amplitude combinations through a 4D distribution matcher (4D-DM) with rate $R_{DM,4D}$ bits per 4D symbol. Each of these V_1 amplitude combinations are then Gray-demapped into four amplitude bits, which generates $4V_1$ bits in total. For the different variants of the PCS 4D 128SP-16QAM constellations, the output of the DM can be one out of $N \leq 2^4 = 16$ combinations of 1D amplitudes $\{1,3\}$ with the desired probability distribution. In particular, the 128SP-16QAM constellation has $N = 16$ combinations of 1D amplitudes and the 88SP-16QAM has $N = 11$ combinations. The $4V_1$ bits are combined with U_2 uniformly distributed bits as the input of a (k,n) LDPC encoder. Then, the generated $n-k$ parity bits are combined with the U_2 bits to form the sequence of sign bits (b_1, b_3, b_5). Finally, b_7 is generated as $\overline{b_7} = b_1 \bigoplus b_2 \bigoplus \ldots b_8$ where $\bigoplus$ is the XOR operator, and the eight bits are mapped to a PDM-16QAM symbol. From the described structure, we find that $k = U_2 + 4V_1$, and $n - k + U_2 = 3V_1$ to have enough sign bits for the $V1$ amplitude combinations. Hence, our transmitter receives $U_1 = \lfloor nR_{DM,4D}/7 \rfloor$ and $U_2 = \lfloor k - 4n/7 \rfloor$ bits at its input where $\lfloor x \rfloor$ is the greatest integer less than or equal to x. The total net rate of this 4D transmitter is $R_T = (U_1+U_2)/V_1 = R_{DM,4D}+7R_C-4$ bits per 4D symbol. As a last note, for the 4D 128SP-16QAM (without applying additional constraints), as we suppress none of the amplitude combinations, we can use the same 1D distribution matching as for the conventional PCS 2D.

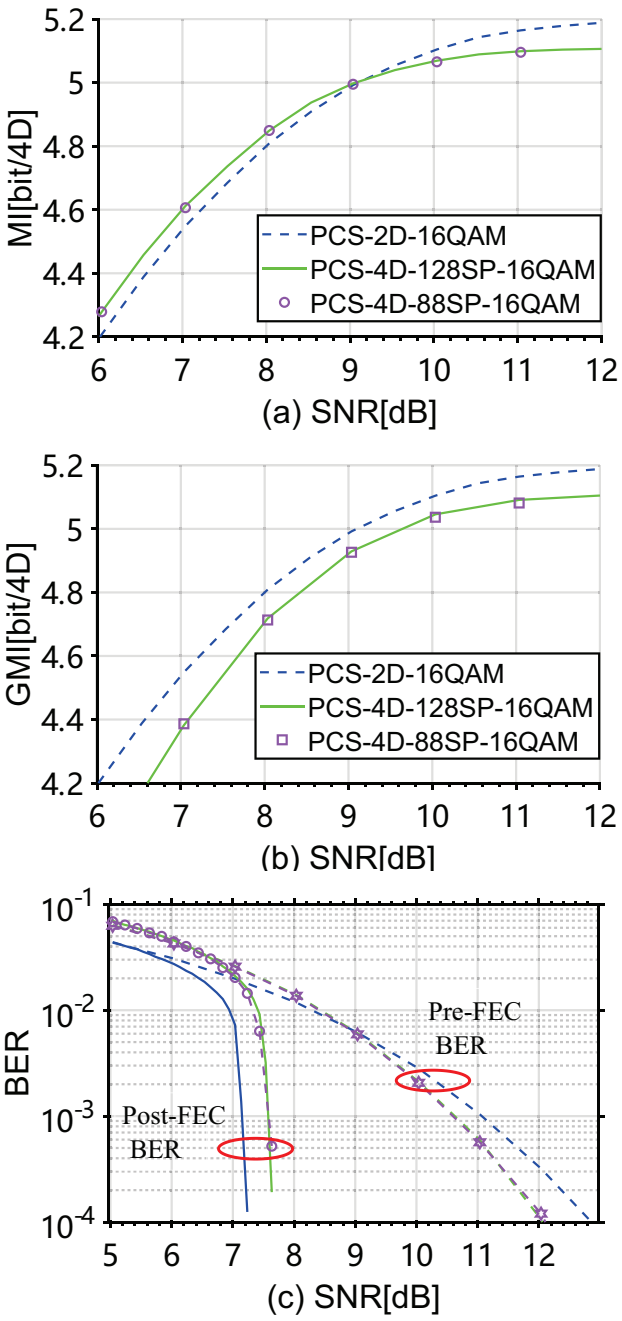

Fig. 2. Performance over AWGN channel: (a) MI versus SNR for PCS schemes optimized for a net rate of 4.4 bits/4D with $R_C = 0.9$, (b) GMI vs SNR, (c) Pre-FEC BER and post-FEC BER vs SNR (PCS-2D-16QAM in blue, PCS-4D-128SP-16QAM in green, PCS-4D-88SP-16QAM in purple).

III. Performance over an AWGN channel

In Fig. 2, we show respectively in (a), the mutual information (MI) defined and computed as Eq. (8) in [11] that we adapt to 4D modulations, and in (b), the generalized mutual information (GMI) that takes into account the bit mapping, computed as Eq. (10) in [11] and that we also adapt to 4D modulations. These information rates are plotted versus the symbol-wise signal-to-noise ratio (SNR) for two PCS 4D schemes, the

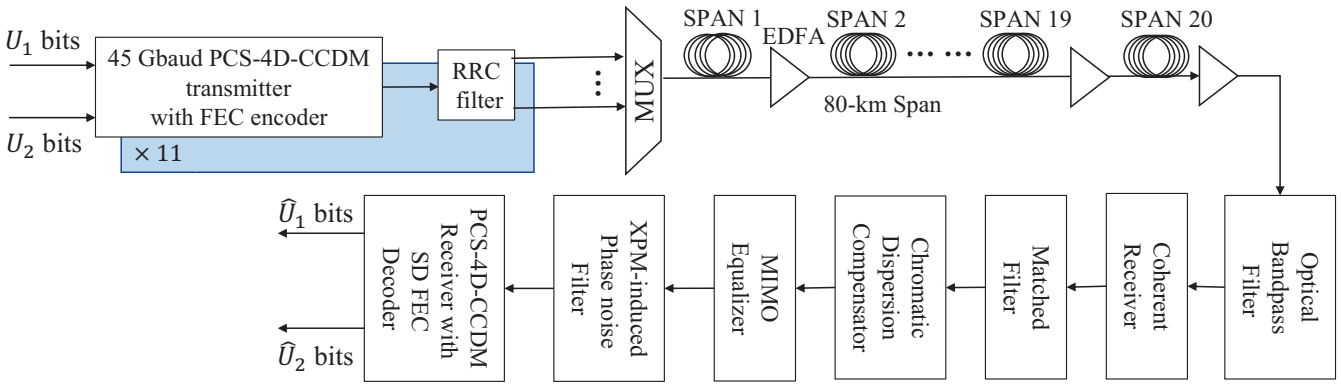

Fig. 3. Schematic of the simulation setup with 4D PAS transmitter and receiver

128SP-16QAM (with five 4D energy levels) and the 88SP-16QAM (with three 4D energy levels) defined in section I. The performance is measured over an additive White Gaussian noise (AWGN) channel. We compare the two 4D schemes to a conventional PDM PCS-2D-16QAM (blue curve). The shaping for the PCS-4D schemes was optimized by targeting a maximization of their generalized mutual information (GMI) at an SNR around 7 dB. The optimization is done by testing different combinations of R_C and $R_{DM,4D}$. The results in Fig. 2 are given for PCS 4D transmitters using an LDPC code of length $n = 64800$ bits from the DVB-S2 standard with a rate $R_C = 0.9$. For the PCS 4D schemes, the highest achievable GMI is 4.4 bits/4D-symbol for an SNR of 7 dB and it is obtained with a DM rate $R_{DM,4D} = 2.1$ bits/4D. The CCDM block length of PCS 4D is hence $V_1 = 9257$ symbols ($U_1 = 19360$ bits if the rate loss is zero; in practice for $n = 64800$, $U_1 = 19469$), $U_2 = 21291$ bits and the rate loss per 1D amplitude is 3.2×10^{-3} bit. The rate loss is computed as Eq. (21) in [12]. For the conventional PCS 2D transmitter, we use the following parameters: $U_2 = 25920$ bits, $V_1 = 32400$ symbols (hence $U_1 = 9720$ bits if the rate loss is zero; in practice for $n = 64800$, $U_1 = 9732$), $R_{DM} = 0.3$ bits/1D and the rate loss is 4×10^{-4} bit per 1D amplitude.

At an SNR of 7 dB, we notice that the mutual information is indeed the highest for the PCS 4D schemes. However, if we look at their generalized mutual information, we see that the GMI deteriorates compared to the conventional scheme. On one side, these behaviors can be explained by the increased MSED of the two PCS 4D schemes with a set-partitioned PDM-QAM compared to a conventional PCS 2D over a non-partitioned QAM. On the other side, the 4D modulations have a larger number of symbol pairs at MSED which cannot fit a Gray mapping for all the 4D neighbors. This sub-optimal bit mapping explains the degraded GMI.

In Fig. 2 (c), we validate the observations on the achievable information rates through bit-error-rate (BER) curves. To decode the FEC code, we use belief propagation (BP) decoding at the receiver side with 50 decoding iterations. The pre-FEC BER curves show the MSED advantage at high SNR and the post-FEC curves show the 0.4 dB penalty due to the sub-optimal bit mapping. This penalty can be suppressed by using joint iterative demapping and decoding or non-binary LDPC codes at the cost of an increased complexity.

IV. Performance over a multi-span non-linear fiber link

After measuring the performance over an AWGN channel, we now evaluate the performance of the three schemes over a multi-span fiber link through numerical simulations, as shown in Fig.3. The transmitted signal consists of 11 WDM channels, each modulated with one of the three schemes at 45 GBaud. The net rate per channel is 198 Gbits/s. The channel spacing is 50 GHz and the central channel is located at 1550 nm. All channels are pulse shaped with a root-raised-cosine filter (RRC) with a roll-off factor of 0.1. The tested fibers are standard single mode fibers (SSMF) with a chromatic dispersion coefficient $D = 17$ ps/nm/km, a non-linearity coefficient $\gamma = 1.32$ $(\mathrm{W} \cdot \mathrm{km})^{-1}$, and a fiber attenuation parameter $\alpha = 0.2$ dB/km at 1550 nm. Polarization mode dispersion (PMD) is added to the simulated link with a PMD coefficient equal to 0.04 ps/$\sqrt{\text{km}}$. The link configuration is a homogeneous multi-span link of 20×80 km. We add an extra noise loading of -13 dBm at the receiver side to get a faster evaluation of low BER values in the dip of the BER curve. Each span is followed by an erbium-doped fiber amplifier (EDFA) with a noise figure of 5 dB. At the receiver side, the central channel is optically filtered and coherently detected. The digital signal processing (DSP) consists of matched filtering, chromatic dispersion compensation, genie-aided linear MIMO channel equalization to remove PMD and polarization crosstalk, and finally, a genie-aided XPM-induced phase noise filter (XPM stands for cross-phase modulation) with an averaging window of 64 symbols as equations (2) and (3) in [13]. We introduce no laser phase noise and no frequency offset between the signal carrier and the local oscillator. We choose a genie-aided DSP approach (perfect knowledge of the linear MIMO channel and optimal XPM-phase noise correction) to measure the intrinsic performance of the 4D-shaped

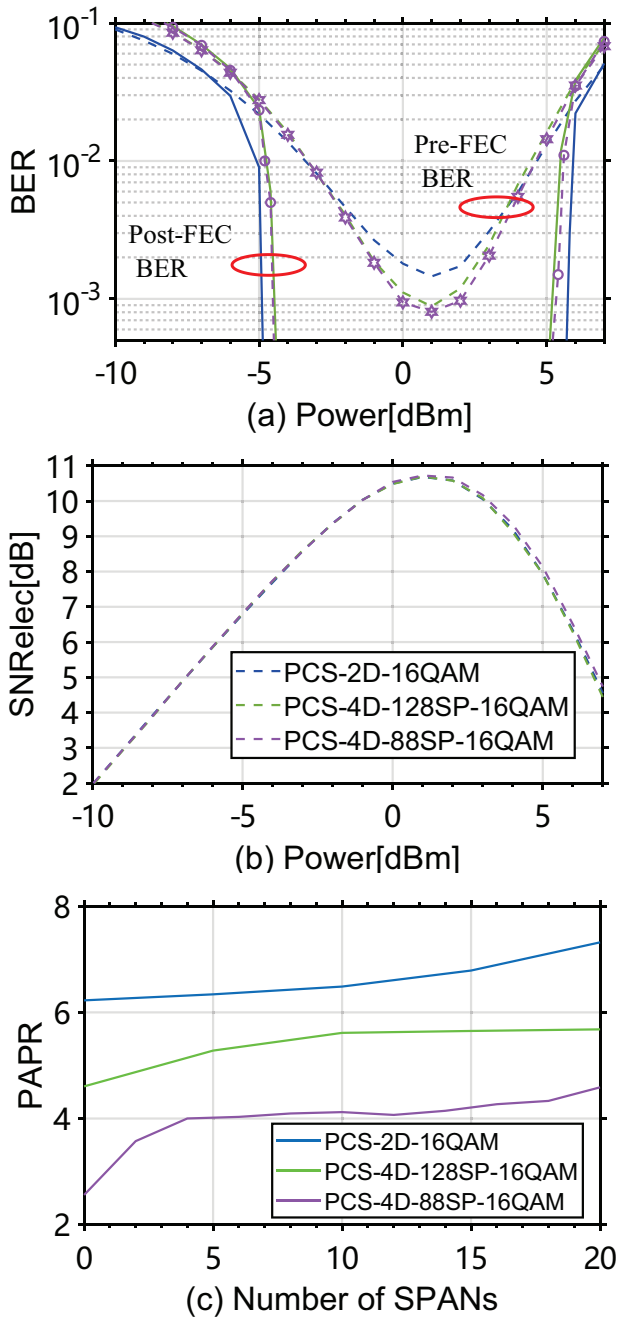

Fig. 4. Performance over a multi-span link (same PCS schemes as in Fig. 2): (a) BER vs launched power/channel. (b) Electrical SNR vs launched power/channel. (c) PAPR vs number of spans.

constellations independently from the potential penalties that a sub-optimal DSP chain might add. The equalized symbols are then passed to the belief-propagation decoder for soft-decision FEC decoding. We do not perform iterations between the demapper and the decoder in order to avoid an increase of the complexity of the receiver.

In Fig. 4(a), we plot the BER curves, and in Fig. 4(b), the electric SNR curves as a function of the launched optical power. Electric SNR is the SNR measured from the noisy constellation points at the input of the decision circuit. As the electric SNR curves are almost the same with slightly better performance for the PCS 4D 88SP-16QAM, the BER enhancement of the PCS 4D schemes, for BER values lower than 10^{-2}, stems from the increased MSED that brings performance gains in the linear regime as well as from the reduced PAPR compared to PCS 2D which brings gains in the non-linear regime. Indeed, in Fig. 4 (c), we plot the PAPR evolution versus the propagation distance for the three PCS schemes at the optimal launch power. We clearly see that the lowest PAPR is obtained with the PCS 4D 88SP-16QAM modulation.

V. Discussion and Conclusion

Through numerical simulations, we illustrated that both linear and non-linear gains, namely an increase in MSED and PAPR reduction, can be achieved with probabilistic shaping of set-partitioned multi-dimensional modulations. For instance, for the suggested PCS-4D 88SP-16QAM, we chose to apply PCS over three 4D energy levels whereas, for the initial 128SP-16QAM constellation, five 4D energy levels are used. While the use of fewer energy levels in the constellation is sub-optimal for a linear AWGN channel, the removal of the two highest energies slightly increased the performance in the non-linear regime. However, the observed gains are measured at BER levels lower than 10^{-2} for a receiver implementing sequential demapping and decoding. To achieve gains at higher BERs, a joint demapper-decoder is required at the cost of higher complexity. Although the observed gains with the two PCS-4D gains are not significant, the described design rule paves the way to further investigations that may lead to multi-dimensional PCS schemes with higher gains. Additional refinements can be brought to the proposed construction to further enhance the trade-off between the linear and non-linear gains, such as the optimization of the selection procedure from the set-partitioned modulation and the search for a better bit mapping.

Acknowledgments

This work was funded by Huawei Technologies France. We would like to thank Hartmut Hafermann for fruitful discussions.

References

[1] T. Pfau, S. Hoffmann, and R. Noé, "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, 2009.

[2] K. Onohara, T. Sugihara, Y. Konishi, Y. Miyata, T. Inoue, S. Kametani, K. Sugihara, K. Kubo, H. Yoshida, and T. Mizuochi, "Soft-Decision-Based Forward Error Correction for 100 Gb/s Transport Systems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1258–1267, 2010.

[3] K. Kikuchi, "Fundamentals of Coherent Optical Fiber Communications," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 157–179, 2016.

[4] F. Frey, L. Molle, R. Emmerich, C. Schubert, J. K. Fischer, and R. F. H. Fischer, "Single-step Perturbation-based Nonlinearity Compensation of Intra- and Inter-Subcarrier Nonlinear Interference," in *2017 European Conference on Optical Communication (ECOC)*, 2017, pp. 1–3.

[5] F. R. Kschischang *et al.*, "Optimal nonuniform signaling for Gaussian channels," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 913–929, 1993.

[6] P. Schulte *et al.*, "Constant composition distribution matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, 2015.
[7] A. Alvarado, T. Fehenberger, B. Chen, and F. M. Willems, "Achievable information rates for fiber optics: Applications and computations," *Journal of Lightwave Technology*, vol. 36, no. 2, pp. 424–439, 2017.
[8] M. Kong *et al.*, "800-Gb/s/carrier WDM Coherent Transmission Over 2000 km Based on Truncated PS-64QAM Utilizing MIMO Volterra Equalizer," *Journal of Lightwave Technology*, vol. 40, no. 9, pp. 2830–2839, 2022.
[9] M. Fu *et al.*, "Multi-Dimensional Distribution Matching With Bit-Level Shaping for Probabilistically Shaped High Order Modulation Formats," *Journal of Lightwave Technology*, vol. 40, no. 9, pp. 2870–2879, 2022.
[10] L. Schmalen, "Probabilistic constellation shaping: Challenges and opportunities for forward error correction," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, 2018, pp. 1–3.
[11] G. Böcherer, "Probabilistic signal shaping for bit-metric decoding," in *2014 IEEE International Symposium on Information Theory*, 2014, pp. 431–435.
[12] Y. C. Gültekin, W. Van Houtum, and F. Willems, "On constellation shaping for short block lengths," in *Proc. Symp. Inf. Theory Signal Process. Benelux*, 2018, pp. 1–11.
[13] T. Fehenberger *et al.*, "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)*. IEEE, 2015, pp. 1–3.