Optics Communications 409 (2018) 7–12

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Invited paper

Constellation design with geometric and probabilistic shaping

Shaoliang Zhang *, Fatih Yaman

NEC Laboratories America, Inc., Princeton, NJ 08540, USA

ARTICLE INFO

Keywords:
Generalized mutual information
Geometric shaping
Multi-dimensional constellation
Pairwise optimization
Probabilistic shaping

ABSTRACT

A systematic study, including theory, simulation and experiments, is carried out to review the generalized pairwise optimization algorithm for designing optimized constellation. In order to verify its effectiveness, the algorithm is applied in three testing cases: 2-dimensional 8 quadrature amplitude modulation (QAM), 4-dimensional set-partitioning QAM, and probabilistic-shaped (PS) 32QAM. The results suggest that geometric shaping can work together with PS to further bridge the gap toward the Shannon limit.

1. Introduction

Advanced modulation formats have been received more and more attention due to the limited received signal-to-noise ratio (SNR) and fiber nonlinearity. Thanks to the high-speed digital coherent receiver and digital-to-analog converters (DAC), well-established multi-level quadrature amplitude modulation (QAM) formats in the wireless communication, such as 8QAM and 16QAM, are applied in the coherent optical communication to deliver beyond 100 Gbps per wavelength data rate [1]. Despite their simple implementation complexity, the recent study shows that these modulation formats are operating far away from the Shannon limit [2]. In other words, the current optical systems are not fully utilizing the capacity provided by the received SNRs.

The pursuit of capacity-approaching modulation formats is under intensive research to close the gap toward the Shannon limit. Although there are many varieties of modulation formats proposed in the literature, they can be generally divided into geometric shaping (GS) and probabilistic shaping (PS) categories. Using matrix rotation of an 8-dimensional (8-dim) bi-orthogonal format, 8-dim 2QAM has been demonstrated to outperforming its two-dimensional (2-dim) counterparts [3], such as dual-polarization binary phase-shifted keying (BPSK) and dual-polarization quadrature phase-shifted keying (QPSK). The recently proposed four-dimensional (4-dim) 8QAM [4] and two-amplitude (2A) 8PSK [5] can beat their 2-dim star-8QAM counterparts by ~0.5 dB at the forward error correction (FEC) limit. Moreover, six-dimensional 16QAM has been proposed to outperform star-8QAM by increased Euclidean distance [6]; however, its performance advantage at the FEC limit comes at the cost of a computationally expensive iterative decoding scheme. The tailoring of modulation formats has been found to become critical in accomplishing the goal to approach the fundamental capacity limit.

Although the shaping gain increases with constellation size, it is capped at 1.53 dB [7]. GS-based constellation, such as amplitude-phase-shifted keying (APSK) and 32QAM, have been demonstrated over trans-Atlantic distance at > 8 b/s/Hz SE over C-band only [8]. To further reduce the gap toward the Shannon limit, PS-based 64QAM have been demonstrated to outperforming 32QAM over 6600 km [9,10] at the expense of increased complexity non-uniform constellation shaping [11]. In [12] GS-optimized 32QAM is designed in a nonlinear optical channel and demonstrated to having higher capacity limits than 32QAM over 11 185 km transmission.

In this paper, we review the generalized pairwise optimization algorithm [13] to design GS multi-dimension constellation together with PS based on two objective functions: minimize the bit error rate (BER) or maximize the modulation capacity at the given SNR and bits mapping. The design procedure with uniform and non-uniform signaling is presented in details and the designed optimized constellation is examined in both analysis and experiments to demonstrate the effectiveness of the proposed algorithm.

The paper is structured as follows: in Section 2 the generalized pairwise optimization algorithm is reviewed together with two objective functions. In Section 3, three design cases are presented using the proposed algorithm: 2-dim 8QAM, 4-dim SP32-QAM. Section 4 summarizes the work.

* Corresponding author.
E-mail address: szhang@nec-labs.com (S. Zhang).

http://dx.doi.org/10.1016/j.optcom.2017.08.063
Received 17 May 2017; Received in revised form 18 August 2017; Accepted 28 August 2017
Available online 12 September 2017
0030-4018/

2. Pairwise optimization algorithm

Optimal 2-dim constellation was designed based on pairwise optimization algorithm at a non-uniform distribution. Two essential steps are proposed in the original paper [14]:

1. Given the target SNR, symbol error rate (SER) is first minimized by optimizing the geometric positions of a pair of symbols $(\mathbf{s}_i, \mathbf{s}_j)$ at a time. The optimal 2-dim constellation is derived by repeating the process after looping through all the $M \times (M-1)/2$ pairs for any M-QAM constellation.
2. Optimum bits mapping is exhaustively searched through layers to minimize the BER of the derived optimal geometric constellation.

In this paper, the aforementioned two steps are consolidated into one simple objective function: *minimize the BER or* $-1 \times$ GMI *given the bits mapping and SNR*. Both objective functions can be calculated using Monte Carlo (MC) simulation. However, instead of using time-consuming MC simulation for BER computation, analytic BER equations are derived here to facilitate the optimization process.

2.1. Analytic BER of N-dimensional constellation

Considering an N-dimensional (N-dim) constellation with M non-equiprobable symbols $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M\}$ at probability mass function (PMF) $p_i, i \in \{1, 2, \dots, M\}$ to encode $m = -\sum_{i=1}^{M} p_i \log_2 p_i$ bits, their bits mappings are represented by $\beta_i, i \in \{1, 2, \dots, M\}$. Note that $\mathbf{s}_i$'s are $N \times 1$ vectors which are denoted in lower case bold, and scalar variables are denoted in lower case normal font. The PMF can be equiprobable, i.e., $\{p_i = \frac{1}{M}, \forall i\}$, to represent uniform constellation signaling. The Hamming distance $h(\beta_i, \beta_j)$ is defined as the number of bits encoded between symbols $\mathbf{s}_i$ and $\mathbf{s}_j$ are different.

The analytic SER between any pair of symbols $(\mathbf{s}_i, \mathbf{s}_j)$ in an additive white Gaussian noise (AWGN) channel can be given as [14]

$$P_s(\epsilon_{ij}) = Q\left(\frac{\|\mathbf{s}_i - \mathbf{s}_j\|}{\sqrt{4\sigma_n^2/N}} + \frac{\sqrt{4\sigma_n^2/N}\ln\frac{p_j}{p_i}}{2\|\mathbf{s}_i - \mathbf{s}_j\|}\right), \tag{1}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-y^2/2}dy$ is the Gaussian Q-function, $\|\cdot\|$ is the norm operation of a vector, and σ_n^2 is the noise variance accumulated in N-dim space. Note that $P_s(\epsilon_{ij})$ is the probability that $\mathbf{s}_i$ is decoded as the sent symbol given that $\mathbf{s}_j$ is sent.

As a result, the SER can be tightly upper bounded by

$$\begin{aligned} P_s &= \sum_{j=1}^{M} P(\epsilon|\mathbf{s}_j)P(\mathbf{s}_j) \\ &\le \sum_{j=1}^{M}\sum_{i\neq j} P(\epsilon_{ij})P(\mathbf{s}_j). \end{aligned} \tag{2}$$

The Hamming distance $h(\beta_i, \beta_j)$ between each symbol pair is taken into account when converting the SER Eq. (2) to BER:

$$P_b \le \sum_{j=1}^{M}\sum_{i\neq j} h(\beta_i, \beta_j)P(\epsilon_{ij})P(\mathbf{s}_j). \tag{3}$$

The derived BER upper bound from Eq. (3) is compared in Fig. 1 with the ones from MC simulation for both uniform and non-uniform 16QAM. The non-uniform 16QAM is produced by adjusting the PMF p_i's governed by Maxwell–Boltzmann distribution depending on the symbol power:

$$\begin{aligned} p_i &= K(\lambda)\exp(-\lambda\|\mathbf{s}_i\|^2), \\ \text{where} \quad K(\lambda) &= \frac{1}{\sum_{i=1}^{M}\exp(-\lambda\|\mathbf{s}_i\|^2)}. \end{aligned} \tag{4}$$

As shown in Fig. 1, the analytic BER of Eq. (3) follows closely with the ones from MC simulation for both PMF scenarios up to the BER level of 2×10^{-2}, which is close to the state-of-the-art soft-decision FEC threshold. Moreover, the performance difference between two formats remains at high BER regime despite they drift away from the MC curves. The density plots of these two formats are plotted in the right insets of Fig. 1. Due to the high accuracy of the analytic BER equations derived here, the pairwise optimization process can take advantage of fast computation of analytic BER to design the optimal geometric shaping within a few seconds.

2.2. GMI calculation of N-dim constellation

In a bit-interleaved coded modulation (BICM) system with a suboptimally mismatched bit-wise decoder, generalized mutual information (GMI) is regarded as an effective approach to compare its post-FEC capacity of a signal constellation χ encoding m bits per symbol [15]. Its channel capacity can be represented by BICM-GMI:

$$C_{\text{GMI}} = H(\chi) - \sum_{i=1}^{m} H(B_i|Y), \tag{5}$$

where $H(\chi)$ denotes the entropy of signal χ associated with a PMF $p(a_i), a_i \in \chi$, given by $H(\chi) = -\sum_{i=1}^{M} p(x_i) \times \log_2 p(x_i)$. The conditional entropy $H(B_i|Y)$ computes uncertainty between the random binary variables $B_i \in \{0, 1\}$ at ith position and the channel output Y. Note that GMI derived here is not the upper bound of the BICM capacity under bit-wise decoder, and its upper bound analysis is beyond the scope of this paper. Despite that GMI can be numerically derived as proofed in [15], MC simulation is used in our paper to estimate the GMI by generating N_s random symbols from constellation χ following a specified PMF $p(a_i)$. The received complex symbols are given by $y = x+n$, where n is complex AWGN with zero mean and variance σ_n^2. At the given SNR $\mathbb{E}\left[\|x_i\|^2\right]/\sigma_n^2$, its BICM-GMI capacity can be computed as

$$\begin{aligned} C^{BICM} &= -\mathbb{E}\left[\log_2 p(a_i)\right] \\ &\quad + \sum_{i=1}^{m}\mathbb{E}\left[\log_2 \frac{\sum_{a\in\chi}\sum_{b=0}^{1} p\left(y|a_i^b\right) p\left(a_i^b\right)\mathbb{I}(x_i^b)}{\sum_{a_j\in\chi} p(y|a_j)p(a_j)}\right]. \end{aligned} \tag{6}$$

Here, $\mathbb{E}[\cdot]$ and $p(y|a_j)$ represent the expectation operation and conditional probability of the received channel outputs given the input symbol a_j, a_i^b and x_i^b denote the transmitted symbols with bit $b \in \{0, 1\}$ at ith position, $\mathbb{I}(x_i^b)$ is an indicator function that equals to 1 when $x_i^b = b$ otherwise 0. In the MC simulation, Eq. (6) can be approximated by

$$C_{\text{GMI}} = H(\chi) + \sum_{i=1}^{\log_2 M}\mathbb{E}\left\{\log_2\left[\frac{b_i(j)\cdot L_1(j) + (1-b_i(j))\cdot L_0(j)}{P_y(j)}\right]\right\}, \tag{7}$$

where $b_i(j) \in \{0, 1\}$ denotes the transmitter bits for the jth symbol at ith bit and $P_y(j)$ is the probability of the received jth symbol $y(j)$. Note that the $\mathbb{E}[\cdot]$ in Eq. (7) is performed with respect to the symbol index j. L_0 and L_1 stand for the bit likelihood (L) for the received jth symbol $y(j)$, which can be computed based on the following equation:

$$\begin{aligned} L_{b_i=0,1} &= \sum_{x^{b_i=0,1}} p(y|x^{b_i=0,1})p(x^{b_i=0,1}) \\ &= \sum_{x^{b_i=0,1}} \left(2\pi\sigma^2\right)^{-n/2}\exp\left(-\frac{\|y - x^{b_i=0,1}\|^2}{2\sigma^2}\right)p(x^{b_i=0,1}), \end{aligned} \tag{8}$$

where $\sigma^2 = \sigma_n^2/N$ is the noise variance in each dimension. Note that σ^2 can be estimated from the recovered constellation in the experiments.

GMI is shown to be affected by both bits mapping and geometric shapes of the constellation. It is well-known that non-Gray-mapped constellation, such as 8QAM, 32QAM and other multi-dimensional constellation, suffers from a BICM-capacity loss because of smaller GMI than their mutual information $I(X;Y)$. Iterative decoding and demapping between LDPC decoder and bits demapper has to be used to overcome the capacity loss in these non-Gray-mapped constellation. To deal with the issue, it has been proposed in [12] to apply C_{GMI} as the

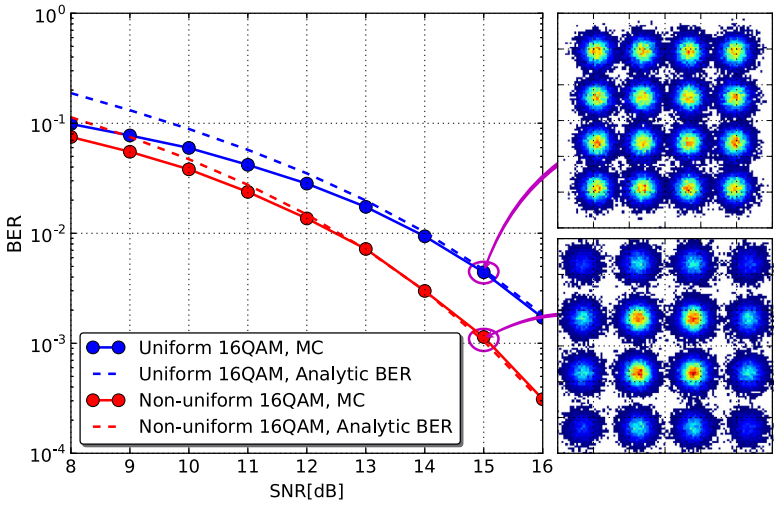

Fig. 1. The performance comparison between analytic BER from Eq. (3) and MC simulation of uniform and non-uniform 16QAM constellation.

objective function in the pairwise optimization algorithm to design GMI-optimized constellation without resorting to high-complexity iterative decoding. The designed constellation based on C_{GMI} objective function will be presented in the Section 3.

2.3. Generalized pairwise optimization

The generalized pairwise optimization algorithm can be summarized as following. The objective function at the specified SNR is to

$$\text{minimize} \quad P_b \quad \text{Eq. (3)} \quad \text{or} \quad -1 \times C_{\mathrm{GMI}} \quad \text{Eq. (7)}, \tag{9}$$

$$\text{subject to} \quad \mathbb{E}\left[\mathbf{s}_i\right] = \mathbf{0}, \quad \text{and} \quad \mathbb{E}[\|\mathbf{s}_i\|^2] = M. \tag{10}$$

Taking any pair $(\mathbf{s}_i, \mathbf{s}_j)$ of M constellation symbols, both the zero mean and average power constraints Eq. (10) can be rewritten as [14]

$$p_i \mathbf{s}_i = -\mathbf{b} - p_j \mathbf{s}_j, \tag{11}$$

$$\|p_j \mathbf{s}_j + \mathbf{b}\|^2 + \frac{p_j}{p_i^2}\|\mathbf{s}_j\|^2 = M - d, \tag{12}$$

where $\mathbf{b} = \sum_{k=1,k\neq i,k\neq j}^{M} p_k \mathbf{s}_k$, and $d = \sum_{k=1,k\neq i,k\neq j}^{M} p_k \|\mathbf{s}_k\|^2$. In the other words, the minimization of the objective function Eq. (9) is simplified into finding $\mathbf{s}_j$ on a hypersphere with center and radius given by the other $M-2$ constellation points. The details about the derivation of center and radius can be found in [14]. Note Eq. (11) defines the relationship between $\mathbf{s}_i$ and $\mathbf{s}_j$, and Eq. (12) reduces the optimization dimension space of $\mathbf{s}_j$ to $N-1$ instead of N. In principle, more constraints can be further added in the pairwise optimization algorithm, such as constant symbol power, $\|\mathbf{s}_i\| = 1, \forall i \in 1, 2, \ldots, M$ [16]. Here we make the pairwise optimization algorithm more general to cover all the potential local or global minimums. The pairwise optimization process need to loop through all $M \times (M-1)/2$ pairs before converging to the optimal state.

3. Optimized multi-dimensional constellation

3.1. Optimized 2-dim 8QAM constellation

The impact of initial constellation on the design of pairwise optimization is investigated in Fig. 2 by using four varieties of 2-dim 8QAM formats: 8PSK, star-8QAM, SP-8QAM and circular (circ)-8QAM at the target 12 dB SNR. Compared the constellation at initialization with the ones after pairwise optimization algorithm, the resultant optimized constellation is significantly different than the initial constellation in Fig. 2(a) and (b). In contrast, little change is observed to the constellation of SP-8QAM and circ-8QAM due to the near-local minimum of the initial states. After pairwise optimization algorithm, the optimized constellation share similar feature in Fig. 2(a), (b) and (d) that center symbol has almost zero power to maximize the Euclidean distance. However, their constellation shapes are not exactly the same to indicate the presence of many local minimum. The principle behind pairwise algorithm is to balance the trade-off between the number of the nearest neighbors and the distance between symbols with different Hamming distance $h(\beta_i, \beta_j)$. A similar methodology has been applied in [17] to optimize star-8QAM and circ-8QAM formats. The estimated Q-factor is plotted in Fig. 3 against the iteration steps at SNR = 12 dB. Here each iteration denotes the optimization of one pair of constellation symbols. It clearly demonstrates how the optimization progresses for each initial constellation to the steady state in less than 100 steps, and there is ∼ 1.4 dB and ∼ 0.5 dB Q-factor improvement, respectively, over 8PSK and star-8QAM. The optimized star-8QAM has the similar performance as the optimized circ-8QAM.

Extensive MC simulation is tried out to locate the global optimum for the optimized 8QAM constellation by changing the initial constellation and optimization seeds. The derived optimized 2-dimensional (opt-2D) 8QAM is plotted in the inset of Fig. 4 with left-MSB bits mapping. Its GMI is compared against other multi-dim 8QAM-equivalent formats proposed in [4–6]. Despite that the derived opt-2D cannot be confirmed to be global optimum, it achieves the best GMI performance among all the proposed 8QAM-equivalent formats. The performance advantage of the designed constellation proves the effectiveness of the pairwise optimization in geometric shaping. It is worth mentioning that simulated-annealing type optimization procedure can be applied here to reduce the probability to be trapped in the local minima.

3.2. Optimized 4-dim SP32 constellation

Generally beyond 2-dim constellation does not have Gray bits mapping due to too many neighbors in N-dim space. As a result, the bits mapping and geometric shaping of N-dim constellation are more critical at $N > 2$. Take SP32 constellation (see Fig. 5(a)) for example, it is derived by partitioning dual-polarization SP-8QAM into two subsets of 4QAM with correlation between two polarization to encode 5 bits/symbol in four dimensions (X-polarization and Y-polarization) [18]. The advantage of SP32 constellation is to reduce the SNR requirement by > 2 dB at the expense of sacrificing 1 bits/symbol compared to dual-polarization SP-8QAM. However, due to the non-Gray-mapped characteristic, Fig. 5 shows that the SNR performance gap between SP32 and SP-8QAM is reduced to about 1 dB at 5 dB Q-factor.

To design optimized SP32 (opt-SP32) constellation, SP32 is fed into the generalized pairwise optimization algorithm as the initial

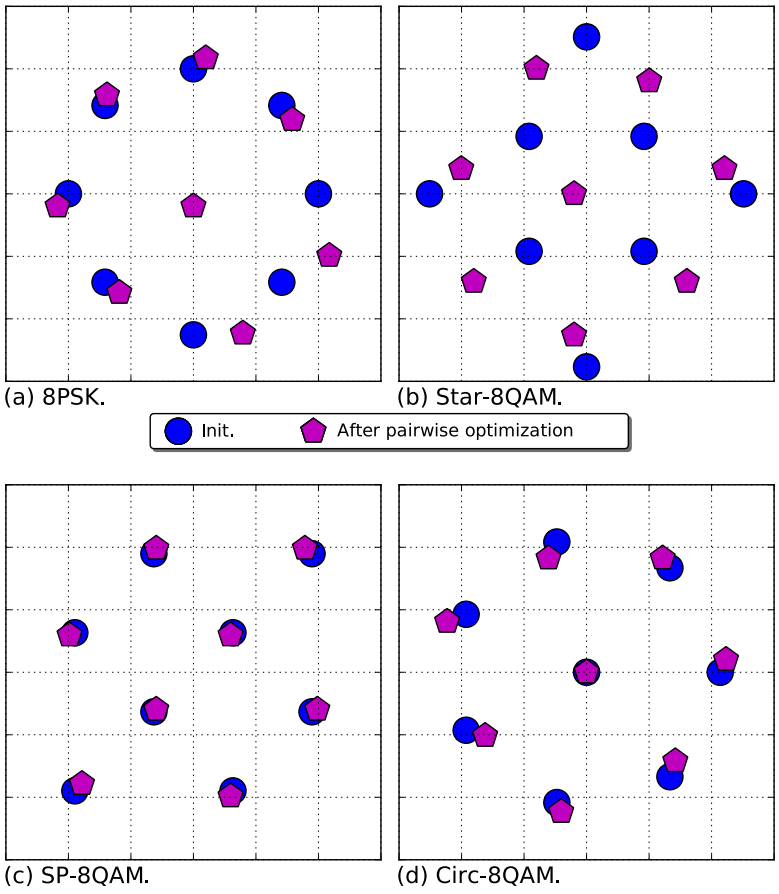

Fig. 2. The constellation at the initial state : (a) 8PSK, (b) Star-8QAM, (c) SP-8QAM, (d) Circ-8QAM, and after pairwise optimization algorithm at 12 dB SNR.

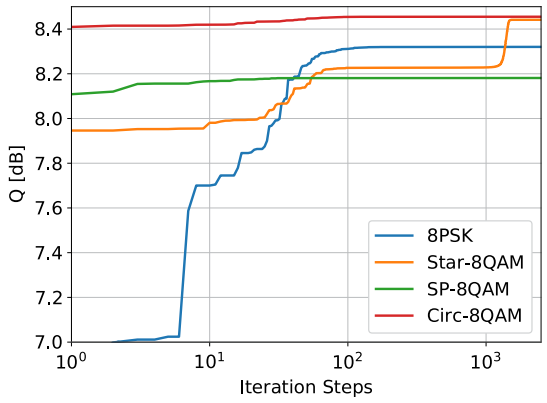

Fig. 3. Q versus iteration.

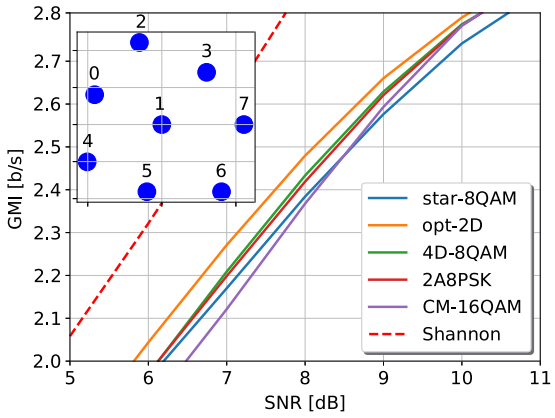

Fig. 4. The analytic GMI versus SNR for various 8QAM-equivalent formats. Inset shows the constellation of opt-2D 8QAM with left-MSB bits mapping.

constellation and bits mapping. Despite there are in total $32\times31/2 = 496$ pairs to loop through, the optimization speed is still fast thanks to the analytic BER given in Eq. (3). The optimization procedure is performed at 9 dB SNR and converged within 1000 steps. Due to the free movement of the constellation points, the resulting opt-SP32 constellation has non-overlapped 32 points in 4-dim space, which is projected to 32 distinct points in each of the 2-dim X-polarization and Y-polarization plane. Therefore, it poses challenge to both transmitter and receiver on generation and recovery. The K-nearest neighboring (KNN) [19] points in each polarization are grouped together if they are separated further than the minimum Euclidean distance. In this manner, the 2-dim projection of the opt-SP32 constellation has only manageable number of points in each polarization to reduce the implementation penalty. The recovered opt-SP32 constellation is plotted in Fig. 5(b) from 34 Gbaud back-to-back experiments. The KNN clustering algorithm is checked to degrade the performance by < 0.2 dB Q-factor at high SNR regime and cause trivial penalty at low SNRs. Note that $K \in \{1, 2, 3, 4\}$ is optimized depending on the allowable performance loss.

As shown in Fig. 5, the derived opt-SP32 outperforms SP32 by ~0.5 dB SNR at 6 dB Q-factor in both theory and experiment. The advantage of opt-SP32 over SP32 remains up to 9.8 dB Q-factor due to the optimization process involving the trade-off between bits mapping and Euclidean distance. Note that 2-dim blind digital signal processing

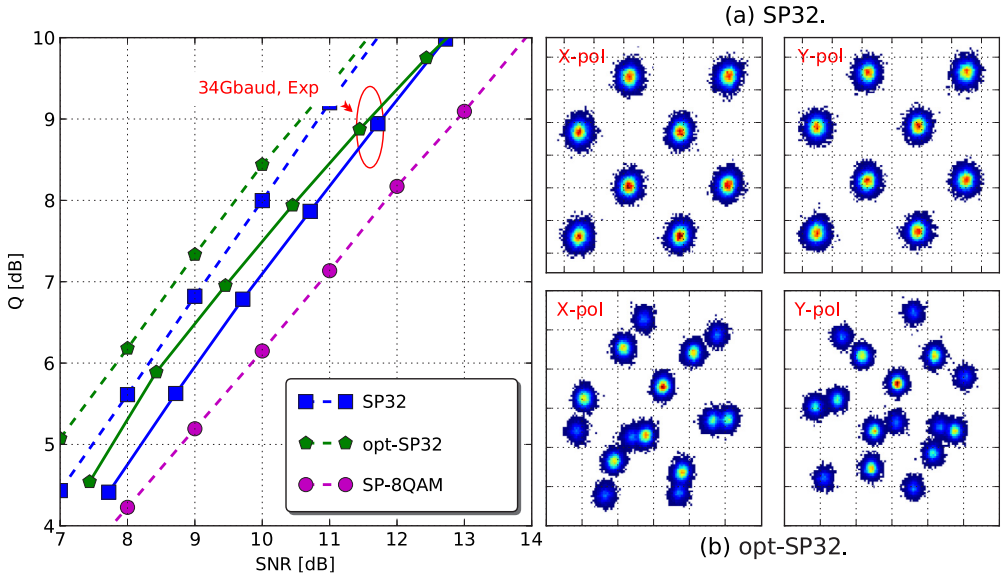

Fig. 5. Performance comparison between SP32, opt-SP32 and SP-8QAM. Solid line: 34Gbaud experiment; dashed line: theory. Constellation of (a) SP32 and (b) Opt-SP32 recovered in the experiments.

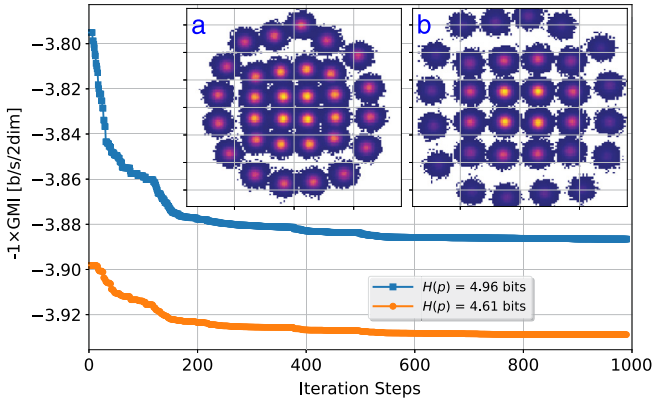

Fig. 6. The GMI trace in the pairwise optimization algorithm at SNR = 12 dB at two $H(p)$. (a) $H(p) = 4.96$ bits; (b) $H(p) = 4.61$ bits.

(DSP) algorithm is used to recover both modulation formats, which causes more implementation penalty to opt-SP32 at < 6 dB Q-factor due to the DSP convergence by distinguishing the closely-spaced points in X-polarization and Y-polarization plane. It is expected to achieve better implementation performance when switching the DSP algorithm into 4-dim space due to the larger symbol separation [6].

3.3. Optimized PS-32QAM constellation

In the previous two cases, the constellation is designed based on minimization of analytic BER. Here we show another example using C_{GMI} as the objective function for optimizing PS-32QAM constellation. The initial non-uniform PS-32QAM constellation is selected to give entropy $H(p) = 4.96$ bits/symbol and $H(p) = 4.61$ bits/symbol by adjusting the PMF p based on the Maxwell–Boltzmann distribution Eq. (4). Based on the GMI trace as a function of optimization steps shown in Fig. 6, the optimization procedure can be converged within 500 steps to demonstrate the fast convergence despite using MC simulation for computing GMI based on Eq. (7). The optimized PS-32QAM constellation could improve the GMI of PS-32QAM by ∼0.1 bit/symbol at $H(p) = 4.96$ bits/symbol. The resulting constellation is plotted in the inset(a) of Fig. 6. Whereas the optimized PS-32QAM only shows trivial ∼0.03 bit/symbol improvement over PS-32QAM at $H(p) = 4.61$ bits/symbol due to the limited gap between the Shannon limit and PS-32QAM at the given SNR. The resulting optimized PS-32QAM constellation is shown in Fig. 6(b). The similar approach using both GS-56APSK together with PS is also reported in [20] to demonstrate 70.4 Tb/s transmission over 7400 km.

Fig. 7 compares the GMI of the optimized PS-32QAM improvement over PS-32QAM and uniform 32QAM. At the target GMI between 3 b/s and 4 b/s, the optimized PS-32QAM is capable of achieving >0.5 dB and ∼0.3 dB SNR gain over 32QAM and PS-32QAM, respectively, at $H(p) = 4.96$ bits/symbol. Note that the PMF of optimized PS-32QAM remains the same as PS-32QAM without introducing additional complexity to non-uniform constellation shaping. However, the improvement is diminished to ∼0.15 dB SNR at $H(p) = 4.61$ bits/symbol. This result suggests that GS brings little gain to PS-QAM format in a linear AWGN channel. The benefits of joint GS and PS in a nonlinear optical channel remains to be studied in the future.

3.4. Optimized-constellation in the presence of fiber nonlinearity

The aforementioned optimization procedure is carried out in the linear AWGN channel, such that the resultant constellation is not guaranteed to outperform the standard M-QAM formats after transmission due to the fiber nonlinearity (NL). Modulation-dependent NL has been confirmed in both analysis and experiments to suggest that large signal moments leads to larger NL [10]. Thanks to the accuracy of the time-domain NL model, this modulation-dependent NL can be analyzed at the given transmission link. An optimized 32QAM with a quasi-Gray bits mapping based on GMI is designed over an optical nonlinear channel by taking into account the fiber NL dependency on the moments of the signal constellation [12,21]. Both analytic and experimental results demonstrate ∼0.5 dB SNR improvement over 32QAM in a BICM system between 3 and 4 b/s GMI. The NL tolerance of opt32 is experimentally shown to be similar as 32QAM after 4290 km [21] and 11 185 km [12] transmission.

4. Conclusion

A generalized pairwise optimization algorithm is reviewed to design optimized constellation based on minimization of BER or maximization of the GMI at the given SNR and bits mapping. The impact of initial constellation and convergence speed are investigated to demonstrate its

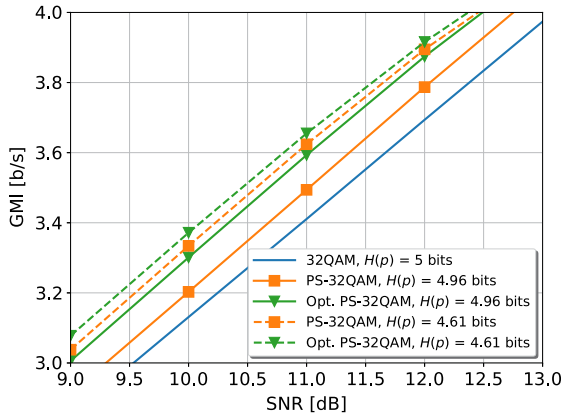

Fig. 7. The analytic GMI versus SNR for uniform 32QAM, PS-32QAM and optimized PS-32QAM.

feasibility to design optimized multi-dim constellation. This approach has been applied to generate 4-dim SP32 and optimized PS-32QAM constellation at $H(p) = 4.96$ bits/symbol to achieve ∼0.5 dB SNR improvement over their counterparts in both simulation and experiments.

References

[1] S. Okamoto, K. Yonenaga, K. Horikoshi, M. Yoshida, Y. Miyamoto, M. Tomizawa, T. Okamoto, H. Noguchi, J. ichi Abe, J. Matsui, H. Nakashima, Y. Akiyama, T. Hoshida, H. Onaka, K. Sugihara, S. Kametani, K. Kubo, T. Sugihara, 400 Gbit/s/ch field demonstration of modulation format adaptation based on pilot-aided OSNR estimation using real-time DSP, IEICE Transactions on Communications advpub (2017). http://dx.doi.org/10.1587/transcom.2017OBI0003.

[2] S. Zhang, Capacity-approaching transmission based on GMI-optimized modulation formats, in: Proc. SPIE, 10130, 2017. http://dx.doi.org/10.1117/12.2252950. 101300L–101300L-6.

[3] A.D. Shiner, M. Reimer, A. Borowiec, S.O. Gharan, J. Gaudette, P. Mehta, D. Charlton, K. Roberts, M. O'Sullivan, Demonstration of an 8-dimensional modulation format with reduced inter-channel nonlinearities in a polarization multiplexed coherent system, Opt. Express 22 (17) (2014) 20366–20374.

[4] M. Reimer, S.O. Gharan, A.D. Shiner, M.O. Sullivan, (2016) Nonlinear tolerant optical modulation formats at high spectral efficiency, US Patent App. 14/744,395 (Dec. 22 2016).

[5] K. Kojima, D.S. Millar, T. Koike-Akino, K. Parsons, Constant modulus 4D optimized constellation alternative for DP-8QAM, in: ECOC, 2014, p. P.3.25. http://dx.doi.org/10.1109/ECOC.2014.6964188.

[6] H. Zhang, C. Davidson, H. Batshon, M. Mazurczyk, M.A. Bolshtyansky, D. Foursa, A. Pilipetskii, DP-16QAM based coded modulation transmission in C+L band system at transoceanic distance, in: Optical Fiber Communication Conference, 2016, p. W1I.2.

[7] F.R. Kschischang, S. Pasupathy, Optimal nonuniform signaling for Gaussian channels, IEEE Trans. Inf. Theory 39 (3) (1993) 913–929.

[8] S. Zhang, F. Yaman, Y.K. Huang, J.D. Downie, D. Zou, W.A. Wood, A. Zakharian, R. Khrapko, S. Mishra, V. Nazarov, J. Hurley, I.B. Djordjevic, Capacity-approaching transmission over 6375 km using hybrid quasi-single-mode fiber spans, J. Lightwave Technol. (ISSN: 0733-8724) 35 (3) (2017) 481–487. http://dx.doi.org/10.1109/JLT.2016.2631151.

[9] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, W. Idler, Rate adaptation and reach increase by probabilistically shaped 64-QAM: an experimental demonstration, J. Lightwave Technol. 34 (7) (2016) 1599–1609.

[10] 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, J. Renaudier, Advanced C+L-band transoceanic transmission systems based on probabilistically shaped PDM-64QAM, J. Lightwave Technol. (ISSN: 0733-8724) 35 (7) (2017) 1291–1299. http://dx.doi.org/10.1109/JLT.2017.2657329.

[11] P. Schulte, G. Bcherer, Constant composition distribution matching, IEEE Trans. Inf. Theory (ISSN: 0018-9448) 62 (1) (2016) 430–434. http://dx.doi.org/10.1109/TIT.2015.2499181.

[12] S. Zhang, F. Yaman, Y.-K. Huang, J.D. Downie, X. Sun, A. Zakharian, R. Khrapko, W. Wood, I.B. Djordjevic, E. Mateo, Y. Inada, 50.962 Tb/s over 11 185 km Bi-Directional C+L transmission using optimized 32QAM, in: Conference on Lasers and Electro–Optics, Optical Society of America, 2017, p. JTh5A.9.

[13] S. Zhang, F. Yaman, E. Mateo, T. Inoue, K. Nakamura, Y. Inada, A generalized pairwise optimization for designing multi-dimensional modulation formats, in: Optical Fiber Communication Conference, Optical Society of America, 2017, p. W4A.6. http://dx.doi.org/10.1364/OFC.2017.W4A.6.

[14] B. Moore, G. Takahara, F. Alajaji, Pairwise optimization of modulation constellations for non-uniform sources, IEEE Can. J. Elect. Comput. Eng. 34 (4) (2009) 167–177. http://dx.doi.org/10.1109/CJECE.2009.5599424.

[15] A. Alvarado, E. Agrell, D. Lavery, R. Maher, P. Bayvel, Replacing the soft-decision FEC limit paradigm in the design of optical communication systems∗, J. Lightwave Technol. 34 (2) (2016) 707–721.

[16] M.A. Reimer, S.O. Gharan, A. Shiner, M. O'Sullivan, Optimized 4 and 8 dimensional modulation formats for variable capacity in optical networks, in: Optical Fiber Communication Conference, 2016, p. M3A.4.

[17] S. Zhang, K. Nakamura, F. Yaman, E. Mateo, T. Inoue, Y. Inada, Optimized BICM-8QAM formats based on generalized mutual information, in: 2015 European Conference on Optical Communication (ECOC), 2015, pp. 1–3. http://dx.doi.org/10.1109/ECOC.2015.7341639.

[18] J.K. Fischer, S. Alreesh, R. Elschner, F. Frey, M. Nölle, C. Schmidt-Langhorst, C. Schubert, Bandwidth-variable transceivers based on four-dimensional modulation formats, J. Lightwave Technol. 32 (16) (2014) 2886–2895. http://dx.doi.org/10.1016/0003-4916(63)90068-X.

[19] G. James, D. Witten, T. Hastie, R. Tibshirani, An introduction to statistical learning, Springer, 2013.

[20] J.-X. Cai, H.G. Batshon, M. Mazurczyk, O.V. Sinkin, D. Wang, M. Paskov, W. Patterson, C. Davidson, P. Corbett, G. wolter, T. Hammon, M.A. Bolshtyansky, D. Foursa, A. Pilipetskii, 70.4 Tb/s capacity over 7600 km in C + L band using coded modulation with hybrid constellation shaping and nonlinearity compensation, in: Optical Fiber Communication Conference Postdeadline Papers, 2017, p. Th5B.2.

[21] S. Zhang, K. Nakamura, F. Yaman, E. Mateo, T. Inoue, Y. Inada, Design and performance evaluation of a GMI-optimized 32QAM, in: 2017 European Conference on Optical Communication (ECOC), 2017, p. Tu.1.D.4.