IEICE Communications Express, Vol.13, No.4, 106–109

ComEX

LETTER

Geometrically shaped four-dimensional modulation formats for 5–7 bits/symbol designed by autoencoder

Akira Naka [1, a)] **and Mamoru Komatsu** [1]

Abstract We design geometrically shaped four-dimensional (4D) modulation formats for 5–7 bits/4D symbol with an autoencoder, which is one of the deep learning applications, using pre-training with 4D formats based on 2-ary amplitude 8-ary phase-shift keying. The numerical evaluation shows that the obtained 4D modulation formats have 0.5-0.7dB signal-to-noise sensitivity gains in terms of normalized generalized mutual information (NGMI) as well as bit error ratio (BER) performances. Furthermore, we show the NGMI and BER performances of these 4D modulation formats are almost equivalent to those of probabilistic amplitude shaping with enumerative sphere shaping with a block length of 20.

Keywords: geometric shaping, 4D modulation, autoencoder, enumerative sphere shaping, 2-ary amplitude 8-ary phase-shift keying

Classification: Fiber-optic transmission for communications

1. Introduction

Multi-dimensional (MD) modulation formats have been proposed to provide variable capacity with higher signal-to-noise (SNR) receiver sensitivity for optical fiber transmission systems [1, 2, 3, 4, 5]. In such MD formats, some components of polarization modes, time-slot modes and/or frequency (wavelength) modes in single-mode fiber as well as in-phase and quadrature (IQ) modulation components are comprehensively treated as a single symbol.

On the other hand, signal shaping has gained attention due to a better receiver sensitivity in terms of SNR toward to Shannon's channel capacity, either via probabilistic shaping (PS) or geometric shaping (GS). PS uses a non-uniform distribution on equidistant constellation points of standard quadrature amplitude modulation (QAM) constellation with a distribution matcher (DM) which converts a uniformly distributed binary data block into a distributed amplitude data block at the transmitter [6, 7, 8]. Increasing the block length improves the SNR receiver sensitivity, and with infinite block length, the Shannon limit is reached in principle. However, as the block length increases, the computational load to be performed in real time during system operation increases. Furthermore, the longer the block length, the greater the influence of burst errors that could occur during DM inverse calculations and countermeasures are required [9, 10].

On the contrary, the GS optimizes the position of constellation points [11, 12]. The GS relies only on properly placing the constellation points and straightforward modifications of the demapper. That is why the GS is a promising format to be an alternative to the PS that uses a non-uniform distribution on equidistant constellation points of standard QAM constellation.

Recently, the extension of GS from two dimensions to multi-dimension has been reported for optical fiber transmission systems [13, 14]. An autoencoder (AE), a deep learning technique, has been attempted to perform complicated computations to optimize the positions of many constellation points for GS [14, 15]. The layered configuration for the AE shown in [15] allows us to optimize not only a position of constellation points but also labeling of binary bit to each point simultaneously. Moreover, the configuration can be directly applied to multi-dimensional modulation formats without modification, simply changing the parameter of dimensional number.

This paper describes three kinds of four-dimensional (4D) modulation formats for 5–7 bits/4D symbol designed by the AE with pre-training technique using the 4D modulation formats of set partitioned 2-ary amplitude 8-ary phase-shift keying (set partitioned 4D-2A8PSK) proposed in [3]. Normalized generalized mutual information (NGMI) and bit error ratio (BER) performances of the obtained GS-4D formats are evaluated under additive white Gaussian noise (AWGN) and compared with the performances of the set partitioned 4D-2A8PSK formats. The performances are also compared with the performances of probabilistic amplitude shaping (PAS), which is a practical scheme to realize PS [6]. In this paper, the amplitude of the PAS is non-uniformly shaped based on enumerative sphere shaping (ESS), which is a lexicographical indexing algorithm to realize sphere shaping that can minimize the rate loss at any block length of distributed amplitude sequence [8].

2. Autoencoder configuration

Figure 1 shows a communication system over AWGN as an AE to obtain 4D modulation formats. Its configuration is exactly equivalent to the one described in [14], configured with reference to [15]. In the mapper section, a m-th length binary bit vector $\boldsymbol{b}$ is encoded as a one-hot vector, i.e., 2^m-dimensional vector, the one element of which is equal to one and zero otherwise. The one-hot vector is mapped to a constellation point $\boldsymbol{x}$ consisting of n_{ch} channels after a rec-

[1] Dept. of Electrical and Electronic Systems Engineering, Ibaraki University, 4-12-1 Naka-Narusawa, Hitachi-shi, Ibaraki 316-8511, Japan

[a)] akira.naka.dr@vc.ibaraki.ac.jp

DOI: 10.23919/comex.2023XBL0179
Received December 20, 2023
Accepted December 28, 2023
Publicized February 16, 2024
Copyedited April 1, 2024

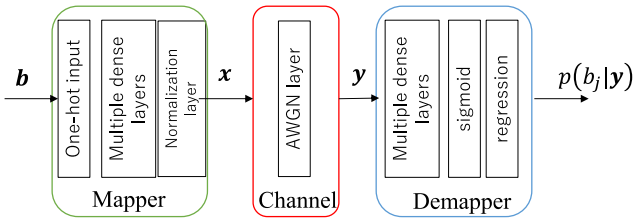


Fig. 1 An autoencoder block diagram.

tified linear unit sandwiched between two fully-connected layers, described as multiple dense layers in Fig. 1, followed by a normalization layer. Note that the number of n_{ch} corresponds to the number of the MD modulation components.

Amount of AWGN is adjusted to SNR value and added to each channel of $\boldsymbol{x}$ in the channel section. In the demapper section, conditional probabilities $p(b_j|\boldsymbol{y}), (j = 1, \ldots, m)$ are obtained at a sigmoid function layer after several multiple dense layers followed by regression layer to calculate the loss of the AE learning. Several iterations of mini-batch training with a size proportional to 2^m are performed using the Adam optimization algorithm to minimize the total binary categorical cross entropy calculated from $\boldsymbol{b}$ and a set of $p(b_j|\boldsymbol{y})$. Note that log-likelihood ratio (LLR) of each binary bit is obtained before sigmoid function layer, which leads to optimization of not only the position of constellation points but also labeling of binary bit to each point [14, 15].

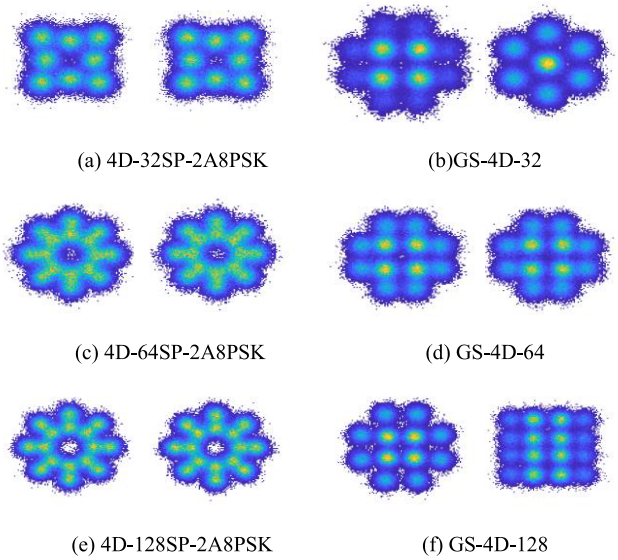
In our calculations, n_{ch} is set to 4, which corresponds to 4D formats. Namely, each channel of $\boldsymbol{x}$ is interpreted as the real or imaginary part of IQ modes, polarization modes or time-slot modes to form the constellation. Further, m is set to 5, 6 or 7, which corresponds to 32, 64 or 128 constellation points, respectively. That is, each transmission capacity is respectively 5, 6 or 7 bits/4D symbol, i.e., 2.5, 3.0 or **3.5** bits/2D symbol, respectively.

To accelerate the AE learning process, we respectively pre-trained the mapper and demapper sections, by setting a constellation of 4D modulation format at a constellation point **x**, before performing the AE learning to obtain the 4D formats with the configuration shown in Fig. 1. We select the set-partitioned 4D-2A8PSK formats for 5-7 bits/4D symbol, namely, 4D-32SP-2A8PSK, 4D-64SP-2A8PSK and 4D-128SP-2A8PSK, since each of which has a set of symmetrical projections about the center point in two 2-dimensional projection, and relatively has good SNR sensitivity [3]. In fact, when we tried several types of constellations including a set-partitioned 4D-16QAM [2], we found that the constellations obtained by the AE learning were not necessarily the same due to convergence of suboptimal conditions. Among them, the constellation obtained with the set-partitioned 4D-2A8PSK formats had the best NGMI and BER performance.

3. Results and discussion

3.1 Obtained 4D constellation

Figures 2(a), 2(c) and 2(e) respectively show the set-partitioned 4D-2A8PSK constellation diagrams in 2 x 2D projection for 5-7 bits/4D symbol used for the pre-training, whereas Figs. 2(b), 2(d) and 2(f) respectively show the optimized 4D modulation constellation diagrams obtained by



(a) 4D-32SP-2A8PSK (b)GS-4D-32

(c) 4D-64SP-2A8PSK (d) GS-4D-64

(e) 4D-128SP-2A8PSK (f) GS-4D-128

Fig. 2 Constellation diagrams of 4D modulation formats in 2×2D projection for a received NGMI ≈ 0.95. The three projections on the left are before the AE learning, and the three on the right are after the AE learning. Communication capacities are respectively 5 bits/4D symbol for (a) and (b), 6 bits/4D symbol for (c) and (d), and 7 bits/4D symbol for (e) and (f).

the AE learning under the conditions of (b) $m = 5$ at SNR of $E_s/N_o = 7.0$ dB, (d) $m = 6$ at $E_s/N_o = 9.0$ dB and (e) $m = 7$ at $E_s/N_o = 11.0$ dB, where E_s/N_o is SNR per 2D symbol.

We find that each constellation point on each set-partitioned 4D-2A8PSK has equal probability as illustrated in Figs. 2(a), 2(c) and 2(e). Note that 4D-32SP-2A8PSK uses only half of the constellation points of 4D-64SP-2A8PSK and 4D-128SP-2A8PSK due to the set-partitioning.

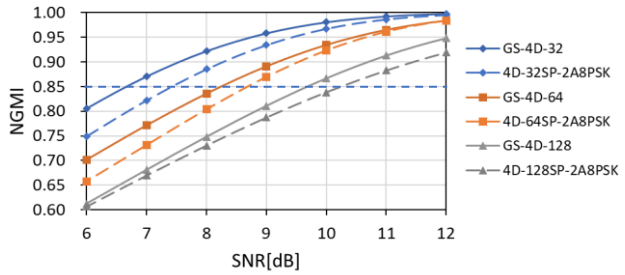
In the case of any transmission capacity of 5-7 bits/4D-symbol, the constellation points significantly change from the ones of the original 2A8PSK formats through the AE learning, so that GS-4D modulation formats are obtained at each capacity, as illustrated in Figs. 2(b), 2(d) and 2(f). Each of the obtained constellations has some constellation points with high probability near the center point in each 2D projection. This allows us to achieve good SNR sensitivity as the non-uniform distribution in the PS does.

The two-dimensionally projected constellations in Figs. 2(b) and 2(f) are asymmetrical to each other, same as the ones shown in [13], but the constellation points are slightly different. This difference is probably because the pre-training performed in [13] uses some set-partitioned 4D-16QAM constellations instead of the set-partitioned 4D-2A8PSK as well as the learning is performed on the constraint of orthant symmetry.

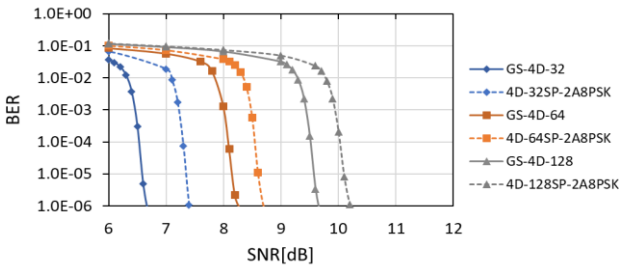
Furthermore, the constellation in Fig. 2(d) is notably similar to the 4D-64SP-12QAM constellation shown in [4]. However, the constellation of Fig. 2(d) does not have equidistant constellation points unlike the 4D 64SP-12QAM, so geometric shaping is expected to have better SNR sensitivity.

3.2 NGMI and BER of GS-4D formats compared to those of set-partitioned 4D-2A8PSK formats

Figures 3(a) and (b) respectively show NGMI and BER performance of the obtained GS-4D modulation formats

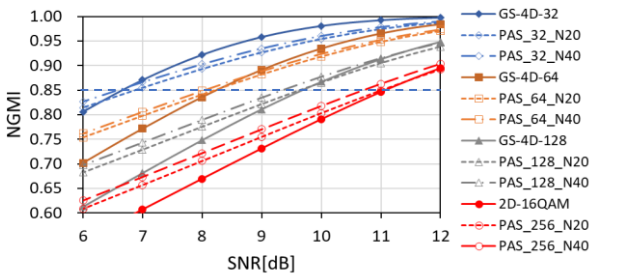


(a) NGMI

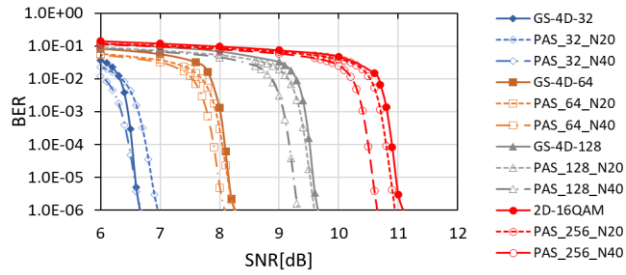


(b) BER

Fig. 3 NGMI and BER performances of the GS-4D modulation formats and the set-partitioned 4D-2A8PSK formats under the number of constellation points 2^m = 32, 64 or 128 in 4D symbol.



(a) NGMI



(b) BER

Fig. 4 NGMI and BER performances of the GS-4D modulation formats and the PAS formats with transmission capacity equivalent to constellation points of 2^m = 32, 64, 128 in 4D symbol, as well as the performances of the PAS format and the 2D-16QM with the capacity equivalent to constellation points of 2^m = 256 in 4D symbol. Amplitude sequence of the PAS is shaped by ESS with a block length of 20 or 40.

compared to the set-partitioned 4D-2A8PSK formats under m = 5, 6 and 7 or the number of constellation points $\mathbf{2^m = 32}$, 64 or 128 using a system configured independently of the AE shown in Fig. 1. BER performances are numerically evaluated by placing a low-density parity-check code (LDPC) encoder and decoder, which is defined by Digital Video Broadcasting–Satellite–Second Generation at a code rate of 4/5, outside the mapper and the demapper.

The LDPC used here can correct all bit errors when NGMI equals to or larger than approximately 0.85 with a code rate of 4/5. Under any calculation conditions, the SNR value when it intersects the black dotted line drawn at NGMI=0.85 in Fig. 3(a) matches the SNR value of the water fall in the BER characteristics in Fig. 3(b).

These results, shown in Figs. 3(a) and (b), show that the NGMI and BER performances of the GS-4D formats are improved by 0.5 to 0.7 dB through the AE learning over the SNR receiver sensitivity of 2A8PSK used for the pre-learning.

3.3 NGMI and BER of GS-4D formats compared to those of PAS formats shaped by ESS

This section shows that the NGMI and BER performances of the obtained GS-4D modulation format in comparison with PAS shaped by ESS [8], in which a block of uniformly distributed input bit sequence is assigned to a block of distributed amplitude sequence in lexicographical order.

The amplitude sequence is limited to the total energy below the maximum total energy in ESS. The block length of amplitude sequence and the total energy is chosen to allow setting the desired information rate. For the amplitude sequence of the block length n, the number of input binary bits $k = \lfloor \log_2 N \rfloor$ is determined by the number N of ways to order the amplitude sequence, where $\lfloor\rfloor$ represents the floor function.

The information rate of the PAS on equidistant constellations of 2D 64QAM after error correction by LDPC is given by

$$R = 2 \cdot \left\{ \frac{k}{n} + 1 - 3(1 - R_c) \right\} \tag{1}$$

where R_c is code rate of LDPC [7]. Here, the maximum total energy values are selected at n = 20 or 40, and R_c = 4/5 for ESS to set k/n = 0.6, 0.8, 1.0 respectively, so that the information rate are 2.0, 2.4, or 2.8 bits/2D symbol after the error correction at R_c = 4/5. These information rates correspond to the transmission capacity of 5, 6 or 7 bits/4D symbol before the error correction, or the number of constellation points 2^m = 32, 64 or 128. Moreover, another energy value is additionally selected to be k/n = 1.2 having information rate of 4 bits/2D symbol before the error correction, corresponding to 2D-16QAM for comparison.

To be noted, we deliberately select the block lengths n of 20 or 40, because each of them allows us to set four kinds of information rate, namely 2.0, 2.4, 2.8 and 3.2 bits/2D symbol. In ESS, the number of input binary bits k or the corresponding information rate R cannot be directly set but is determined by the maximum energy and the block length n, through N, the number of ways to order the amplitude sequence. Therefore, we actually calculate all possible k and R for each maximum energy at each block lengths n, under the condition of $k \leq 53$ which is limited by our calculation environment. In fact, the only n that realizes each of the rate under the above condition is 15 for ESS at R_c = 4/5, except for 20 and 40.

Figures 4(a) and (b) show that the NGMI and BER performances of the GS-4D modulation formats with 2^m = 32, 64 or 128 are almost equivalent to or slightly worse than

those of ESS under the condition of $n = 20$, while GS-4D format with $2^m = 32$ slightly better than that of ESS. When n is 40, each rate loss of ESS becomes smaller than when n is 20, so each ESS performance is improved by about 0.2dB to 0.4dB to outperform each GS-4D format. Noted that the performance of GS-4D-32 is almost equivalent to PAS_32_N40. We assume that this result is due to the performance degradation of PAS_32_N40 formats. That is, as m becomes smaller, the amplitude distribution becomes steeper and deviates from ideal shaping for our numerical evaluation.

These results are quantitatively consistent with that of 2D-16QAM, as the NGMI and BER performances of 2D-16QAM are also equivalent to those of ESS under the condition of $n = 20$. These quantitative analyses show GS-4D formats are equivalent to 2D-16QAM in terms of SNR sensitivity although the transmission capacities are different. Therefore, it can be considered that GS-4D modulation formats for 5 to 7 bits/4D symbol are optimized by the AE learning.

4. Conclusion

We design GS-4D modulation formats for 5–7 bits/4D symbol with the AE learning, using pre-training with set-partitioned 4D-2A8PSK formats. The obtained GS-4D modulation formats respectively have some constellation points with high probability near the center point in each 2D projection to achieve higher SNR sensitivity. The obtained NGMI and BER performances of the GS-4D modulation formats are better than those of the set-partitioned 4D-2A8PSK formats and are almost equivalent to those of 2D PAS format shaped by the ESS with the amplitude sequence block length of 20 at the code rate of 4/5 of a practical LDPC.

References

[1] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *J. Lightw. Technol.*, vol. 27, no. 22, pp. 5115–5126, 2009. DOI: 10.1109/JLT.2009.2029064

[2] J. Renaudier, A. Voicila, O. Bertran-Pardo, O. Rival, M. Karlsson, G. Charlet, and S. Bigo, "Comparison of set-partitioned two-polarization 16QAM formats with PDM-QPSK and PDM-8QAM for optical transmission systems with error-correction coding," ECOC'12, We1. C.5, 2012. DOI: 10.1364/ECEOC.2012.We.1.C.5

[3] K. Kojima, T. Yoshida, T. Koike-Akino, D.S. Millar, K. Parsons, M. Pajovic, and V. Arlunno, "Nonlinearity-tolerant four-dimensional 2A8PSK family for 5–7 Bits/Symbol spectral efficiency," *J. Lightw. Technol.*, vol. 35, no. 8, pp. 1383–1391, 2017. DOI: 10.1109/JLT.2017.2662942

[4] T. Nakamura, E. Le Taillandier de Gabory, H. Noguchi, W. Maeda, J. Abe, and K. Fukuchi, "Long haul transmission of four-dimensional 64SP-12QAM signal based on 16QAM Constellation for Longer Distance at same spectral efficiency as PM-8QAM," ECOC'15, pp. 1–5, 2015. DOI: 10.1109/ECOC.2015.7341874

[5] M. Nakamura, F. Hamaoka, A. Matsushita, K. Horikoshi, H. Yamazaki, M. Nagatani, A. Sano, A. Hirano, and Y. Miyamoto, "Coded eight-dimensional QAM technique using iterative soft-output decoding and its demonstration in high baud-rate transmission," *J. Lightw. Technol.*, vol. 35, no. 8, pp. 1369–1375, 2017. DOI: 10.1109/JLT.2017.2669919

[6] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4651–4665, 2015. DOI: 10.1109/TCOMM.2015.2494016

[7] J. Cho, P.J. Winzer, "Probabilistic constellation shaping for optical fiber communications," *J. Lightw. Technol.*, vol. 37, no. 6, pp. 1590–1607, 2019. DOI: 10.1109/JLT. 2019.2898855

[8] Y.C. Gültekin, W.J. van Houtum, A.G.C. Koppelaar, F.M.J. Willems, and W.J. van Houtum, "Enumerative sphere shaping for wireless communications with short packets," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1098–1112. 2020. DOI: 10.1109/TWC.2019.2951139

[9] P. Skvortcov, T. Koike-Akino, D.S. Millar, K. Kojima, and K. Parsons, "Dual coding concatenation for burst-error correction in probabilistic amplitude shaping," *J. Lightw. Technol.*, vol. 40, no. 16, pp. 5502–5513, Aug. 15, 2022. DOI: 10.1109/JLT.2022.3178675

[10] A. Naka, "Burst error characteristics in probabilistic constellation shaping," *IEICE Commun. Express*, vol. 10, no. 10, pp. 775–779, 2021. DOI: comex.2021XBL0117

[11] L. Galdin, A. Edwards, W. Yi , E. Sillekens, Y. Wakayama, T. Gerard, W.S. Pelouch, S. Barnes, T. Tsuritani, R.I. Killey, D. Lavery, and Polina Bayvel, "Optical fibre capacity optimisation via continuous bandwidth amplification and geometric shaping," *IEEE Photon. Technol. Lett.*, vol. 32, no. 17, pp. 1021–1024, 2020. DOI: 10.1109/LPT.2020.3007591

[12] S. Zhang and F. Yaman, "Design and comparison of advanced modulation formats based on generalized mutual information," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 416–423, 2018. DOI: 10.1109/JLT.2017.2779753

[13] B. Chen, Y. Lei, G. Liga, Z. Liang, W. Ling, X. Xue, and A. Alvarado, "Geometrically-shaped multi-dimensional modulation formats in coherent optical transmission systems," *J. Lightw. Technol.*, vol. 41, no. 3, pp. 897–910, 2023. DOI: 10.1109/JLT.2022.3204101

[14] A. Naka and M. Komatsu, "Geometrically shaped multidimensional modulation formats designed by deep learning," *IEICE Commun. Express*, vol. 12, no. 4, pp. 139–144, 2023. DOI: 10.1587/comex.2022XBL0176

[15] S. Cammerer, F.A. Aoudia, S. Dörner, M. Stark, J. Hoydis, and S. Brink, "Trainable communication systems: Concepts and prototype," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5489–5503, 2020. DOI: 10.1109/TCOMM.2020.3002915