Analysis and Experimental Demonstration of Orthant-Symmetric Four-dimensional 7 bit/4D-sym Modulation for Optical Fiber Communication

Bin Chen, Member, IEEE, Alex Alvarado, Senior Member, IEEE, Sjoerd van der Heide, Student Member, IEEE, Menno van den Hout, Student Member, IEEE, Hartmut Hafermann, Senior Member, IEEE, and Chigo Okonkwo, Senior Member, IEEE.

Abstract—We propose a new four-dimensional orthant-symmetric 128-ary modulation format (4D-OS128) with a spectral efficiency of 7 bit/4D-sym. The proposed format fills the gap between polarization-multiplexed 8-ary and 16-ary quadrature-amplitude modulation (PM-8QAM and PM-16QAM). Numerical simulations show that 4D-OS128 outperforms state-of-the-art geometrically-shaped modulation formats by up to 0.65 dB for bit-interleaved coded modulation at the same spectral efficiency. These gains are experimentally demonstrated in a 11×233 Gbit/s wavelength division multiplexing (WDM) transmission system operating at 5.95 bit/4D-sym over 6000 km and 9000 km for both EDFA-only and hybrid amplification scenarios, respectively. A reach increase of 15% is achieved with respect to 128-ary set-partitioning 16QAM.

I. INTRODUCTION

The demand for higher capacity and longer transmission distances in optical fiber communications has been growing for several years. In order to further support the exponential traffic growth, various multiplexing or modulation dimensions such as time, polarization, wavelength and space (multi-mode/multi-core fibers), have been used. In particular, wavelength division multiplexing (WDM) transmission with coherent detection enables high-order modulation formats. These formats have been shown to be promising for next-generation high-speed long-haul transmission systems.

B. Chen is with the School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China. B. Chen is also with Eindhoven University of Technology, 5600 MB, Eindhoven, The Netherlands (e-mail: chen.bin.conan@gmail.com).

A. Alvarado is with the Information and Communication Theory Lab, Signal Processing Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB, Eindhoven, The Netherlands (e-mail: a.alvarado@tue.nl).

S. van der Heide, M. van den Hout and C. Okonkwo are with the Institute for Photonic Integration, Eindhoven University of Technology, Eindhoven 5600 MB, The Netherlands (e-mail: {s.p.v.d.heide, m.v.d.hout, c.m.okonkwo}@tue.nl).

H. Hafermann is with the Optical Communication Technology Lab, Paris Research Center, Huawei Technologies France SASU, 92100 Boulogne-Billancourt, France (e-mail: hartmut.hafermann@huawei.com).

This research is supported in part by Huawei France through the NLCAP project. The work of B. Chen is supported by the National Natural Science Foundation of China (NSFC) under Grant 61701155 and Fundamental Research Funds for the Central Universities under Grant J2019HGBZ0130. S. van der Heide, M. van den Hout and C. Okonkwo are supported in part by the Dutch NWO Gravitation Program: Research centre for Integrated Nanophotonics (Grant Number 024.002.033). The work of A. Alvarado is supported by the Netherlands Organisation for Scientific Research (NWO) via the VIDI Grant ICONIC (project number 15685).

Achieving higher spectral efficiency (SE) by employing high order polarization multiplexed M-ary quadrature amplitude modulation (PM-MQAM) formats comes at the cost of a reduced transmission reach and has been extensively studied [1, 2]. In order to maintain or extend transmission distances in high-speed fiber-optic systems, signal shaping has recently been the focus of considerable attention in the optical communications community.

Shaping methods can be broadly categorized into probabilistic shaping (PS) and geometric shaping (GS), both having distinct advantages and disadvantages. In PS, long coded sequences induce certain nonuniform probability distribution on the constellation points [3]–[9]. In GS, nonrectangular constellations are used with the same probabilities [10]–[13]. For the additive white Gaussian noise (AWGN) channel, it has been shown that PS outperforms GS and more closely approaches Shannon’s channel capacity when the number of constellation points is limited [14] Sec. 4.3], [15]. On the other hand, GS over multiple dimensions can not only reduce the gap to the Shannon capacity [16]–[20], but could also mitigate the nonlinear effects in the optical channel [21]–[26]. Multidimensional (MD) GS therefore offers an interesting approach to achieve shaping gains with low implementation complexity. The key advantage of GS over PS is that GS relies only on the selection of the location of constellation points in a (relatively low) MD space and the design of the corresponding MD detector.

Four-dimensional (4D) modulation formats are typically optimized in the four dimensions consisting of the two quadratures (I/Q) and the two polarization states (X/Y) of the optical field. These formats are often designed to achieve large minimum Euclidean distances [16]–[19]. Conventional polarization multiplexed formats are not true 4D formats, because they are only optimized in I, Q, X, and Y independently. True 4D modulation formats can be obtained by applying Ungerboeck’s set-partitioning (SP) scheme [27] or using sphere packing argument[28]. Set-partitioning PM-16QAM has been investigated to achieve fine granularity as 32-ary set-partitioning QAM (32SP-16QAM) [29], [30], or using sphere packing argument[28]. Set-partitioning PM-16QAM has been investigated to achieve fine granularity as 32-ary set-partitioning QAM (32SP-16QAM) [29], [30], 64-ary set-partitioning QAM (64SP-16QAM) [31] and 128-ary set-partitioning QAM (128SP-16QAM) [32]–[34]. 32SP-

1An excellent summary of MD constellations is given in the online database [28].
amplification, with a data rate of 233 Gb/s per channel. We demonstrate two amplifier configurations, EDFA-only and hybrid, by both numerical simulations and experimentally. We demonstrate the transmission performance of [24], all of them having the same spectral efficiency (7 bit/4D-sym). In addition, we investigate the transmission performance [24], with orthant symmetry constraint has negligible performance loss with 8D formats give interesting performance advantages, larger GMI gains (available for larger constellation sizes and higher dimensionality) are difficult to obtain due to the challenging multi-parameter constellation and labelling optimization. Previous works only solve the 4D or 8D GMI optimization problem for up to SE of 6 bits/4D-sym [25]. In this paper, we propose a novel four-dimensional orthant-symmetric 128-ary modulation format (4D-OS128) to go beyond the 6 bits/4D-sym limit. The format 4D-OS128 has an SE of 7 bit/4D-sym and is obtained via GS by jointly optimizing constellation coordinates and labeling in 4D to maximize GMI. For the design we use the orthant symmetry idea to overcome the challenging multi-parameter optimization, which can be seen as a trade-off between optimization efficiency and performance. We found that the obtained 4D format with the orthant symmetry constraint has negligible performance loss with respect to the one without constraint. The obtained 4D format is compared in terms of linear performance to 128-SP-QAM and a 7 bit modulation in 4D-2A8PSK family (7b4D-2A8PSK) [24], all of them having the same spectral efficiency (7 bit/4D-sym). In addition, we investigate the transmission performance by both numerical simulations and experimentally. We demonstrate two amplifier configurations, EDFA-only and hybrid amplification, with a data rate of 233 Gb/s per channel. We target a GMI lower than the SD-FEC threshold of 5.95 bit/4D-sym [24], which corresponds to a FEC rate of 0.8 (25% FEC OH). For the baseline constellations, distances around 5000 km (EDFA-only) and 8000 km (hybrid amplification) are therefore targeted. Compared at the same bit rate, the proposed 4D-OS128 format achieves a 15% longer transmission reach than 128SP-16QAM and 7b4D-2A8PSK for 11 wavelength division multiplexing (WDM) channels transmission.

This paper is organized as follows. In Sec. II the design methodology and the proposed modulation format are introduced. In Sec. III numerical results are shown for both AWGN and nonlinear optical fiber simulations. The experimental setup of the WDM optical fiber system and the experimental results are described in Sec. IV. Conclusions are drawn in Sec. V.

II. 4D MODULATION FORMAT AND OPTIMIZATION

A. GS Optimization: General Aspects

The optical channel suffers from intersymbol and interpolarization interference. This channel can be modeled by a conditional PDF $P_{Y|X}$, where $X$ and $Y$ are the transmitted and received symbols, respectively. The transmitted MD symbols $X$ are assumed to be MD symbols with $N$ real dimensions (or equivalently, with $N/2$ complex dimensions) drawn uniformly from a discrete constellation $\mathcal{X}$ with cardinality $M = 2^n = |\mathcal{X}|$. The most popular case in fiber optical communications is $N = 4$, which corresponds to coherent optical communications using two polarizations of the light (4 real dimensions). This naturally results in four-dimensional (4D) modulation formats.

In general, the channel law $P_{Y|X}$ is such that the dimensionality of $Y$ is larger than the dimensionality of the transmitted symbols $X$. This comes from the fact that the fiber optical channel introduces memory over multiple symbols, even after dispersion compensation. From now on, however, we consider a channel law $P_{Y|X}$ where the output symbols $Y$ have also $N/2$ complex dimensions. Due to this assumption, we are actually only approximating the true optical channel. As we will explain below, this approximation is well-matched to the fact that typical optical receivers ignore potential memory across 4D symbols.

Throughout this paper, we consider BICM, which is one of the most popular coded modulation schemes. The transmitted symbols $X$ are jointly modulated in 4D space by a set of constellation coordinates and the corresponding labeling strategy. The $i$th constellation point is denoted by $s_i = [s_{i1}, s_{i2}, s_{i3}, s_{i4}] \in \mathbb{R}^4$ with $i = 1, 2, \ldots, M$ in four real dimensions. We use the $M \times 4$ matrix $S = [s_1; s_2; \ldots; s_M]$ to denote the 4D constellation. The $i$th constellation point $s_i$ is labeled by the length-$m$ binary bit sequence $b_i = [b_{i1}, \ldots, b_{im}] \in \{0, 1\}^m$. The binary labeling matrix is denoted by a $M \times m$ matrix $B = [b_1; b_2; \ldots; b_M]$, which contains all unique length-$m$ binary sequences. The 4D constellation and its binary labeling are fully determined by the pair of matrices $(S, B)$.

3One empirical model that properly takes this effect into account for channel capacity calculations is the so-called finite-memory GN model [41]. Another example is the time-domain perturbation models in [42]–[44], where the received symbol depends on previous and future transmitted symbols (potentially across other polarizations and channels).

2Notation convention: Random (row) vectors are denoted by $X$ and its corresponding realization are denoted by $x$. Matrices are denoted by $X$. A semicolon is used to denote vertical concatenation of vectors, e.g., $X = [x_1; x_2] = [x_1^T, x_2^T]^T$, where $[\cdot]^T$ denotes transpose.
In this paper, we are interested in maximizing the GMI. The transmitted bits are assumed to be independent and uniformly distributed, which implies uniform symbols $X$. The receiver assumes a memoryless channel and also uses a bit-metric decoder (i.e., a standard BICM receiver). In this case, the receiver uses a decoding metric $q(y, c)$ proportional to the product of the bit-wise metrics, i.e., the decoding metric is

$$q(y, c) \propto \prod_{k=1}^{m} p_{Y|C_k}(y|c_k)$$

where $C = [C_1, C_2, \ldots, C_m]$ is the random vector representing the transmitted bits, and $I(C_k; Y)$ is the mutual information between the bits and the symbols, and where the notation $\mathbb{G}$ emphaizes the dependency of the GMI on the constellation, binary labeling, and channel law. Furthermore, for any $N$-dimensional channel law, $G$ can be expressed as \[ G(S, B, p_Y|X) = \sum_{k=1}^{m} I(C_k; Y), \] where $I(C_k; Y)$ is the mutual information between the bits and the symbols, and where the notation $G(S, B, p_Y|X)$ emphaizes the dependency of the GMI on the constellation, binary labeling, and channel law. Furthermore, for any $N$-dimensional channel law, $G$ can be expressed as \[ G(S, B, p_Y|X) = m + \frac{1}{M} \sum_{k=1}^{m} \sum_{b \in \{0, 1\}} \sum_{i \in I_k^b} \int_{\mathbb{R}^N} p_Y|X(y|x_i) \log_2 \frac{\sum_{j \in I_k^b} p_Y|X(y|x_j)}{\sum_{j=1}^{M} p_Y|X(y|x_j)} dy. \]

As shown in [4], a GMI optimization requires a jointly optimization of the 4D coordinates and its binary labeling. A GMI-based optimization finds a constellation $S^*$ and labeling $B^*$ for a given channel conditional PDF $p_Y|X$ and energy constraints, i.e.,

$$\{S^*, B^*\} = \arg \max_{S, B; E[||X||^2] \leq \sigma_X^2} G(S, B, p_Y|X),$$

where $\sigma_X^2$ represents the transmitted power, $S^*$ and $B^*$ indicate the optimal constellation and labeling, resp. For the nonlinear fiber optical channel, the constraint $E[||X||^2] \leq \sigma_X^2$ can be omitted as this channel is interference limited, i.e., power is typically not a limitation.

Note that for any given channel $p_Y|X$, the optimization problem in [5] is a single objective function $G$ with multiple parameter and constraints. From previous works [23, 25, 29], it is known that the constellation optimization and GMI calculation for large constellations and/or for constellations with high dimensionality is computationally demanding. Therefore, an unconstrained optimization is very challenging. Unconstrained formats also impose strict requirements for the generation of the signals (i.e., high-resolution digital-to-analog converter (DAC)) as well as complex MD detectors.

### B. Orthonormal Symmetric (OS) Geometric Shaping Optimization

To solve the multi-parameter optimization challenges described above and to reduce the transceiver requirements, we propose to impose an “orthonormal symmetry” constraint to the $N$-dimensional modulation format to be designed. Our proposed approach makes the MD format to be generated from a first-orthant labeled constellation (see Definition 7 below). These concepts are defined in what follows, and are based on $N$-dimensional orthants, defined as the intersection of $N$ mutually orthogonal half-spaces passing through the origin. By selecting the signs of the half-spaces, the $2^N$ orthants available in an $N$-dimensional space can be obtained.

Let $L_q = \{l_1; l_2; \ldots; l_{2q}\}$ with $l_j \in \{0, 1\}$ and $j = 1, 2, \ldots, 2^q$ denote a $2^q \times q$ labeling matrix of order $q$, which contains all unique length-$q$ binary vectors. Let $H_k$ be a $N \times N$ rotation matrix defined as

$$H_k = \begin{bmatrix} (-1)^{l_{k1}} & 0 & \cdots & 0 \\ 0 & (-1)^{l_{k2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (-1)^{l_{kN}} \end{bmatrix},$$

with $k = 1, 2, \ldots, 2^N$ and $l_k = [l_{k1}, l_{k2}, \ldots, l_{kN}]$ are the rows of the labeling matrix of order $N \mathbb{L}_N$.

**Example 1 (Rotation matrices):** For the 1D case ($N = 1$), there are only two orthants (an orthant is a ray), the rotation matrices are $H_1 = 1$ and $H_2 = -1$. For the 2D case ($N = 2$), there are four orthants (an orthant is a quadrant), the rotation matrices are

$$H_1 = \begin{bmatrix} +1 & 0 \\ 0 & +1 \end{bmatrix}, \quad H_2 = \begin{bmatrix} -1 & 0 \\ 0 & +1 \end{bmatrix},$$

$$H_3 = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix}, \quad H_4 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$  

For the $N = 3$ case, there are $2^3 = 8$ orthants (an orthant is an octant), the eight rotation matrices which can be obtained by $[6]$, with $I_j \in \{0, 1\}^3$, and $j = 1, 2, \ldots, 8$. \[ \square \]

We now define a first-orthant and orthant-symmetric (OS) labeled constellations.

**Definition 1 (First-orthant labeled constellation):** The pair of matrices $\{T, \mathbb{L}_{m-N}\}$ is said to be a first-orthant labeled constellation if $T = \{t_1; t_2; \ldots; t_{2^{m-N}}\}$ is a constellation matrix such that $t_j \in \mathbb{R}_+^N, \forall j \in \{1, 2, \ldots, 2^{m-N}\}$ (all entries are nonnegative), and $\mathbb{L}_{m-N}$ is a labeling matrix of order $m-N$. \[ \square \]

**Definition 2 (Orthant-symmetric labeled constellation):** The pair of matrices $\{S, B\}$ is said to be an OS labeled constellation if $S = [S_1; S_2; \ldots; S_{2^N}]$ is a $2^m \times N$ constellation matrix and $B = [B_1; B_2; \ldots; B_{2^N}]$ is a $2^m \times m$ binary labeling matrix, where the constellation matrix $S$ is constructed via $S_k = T B_k$ with $k = 1, 2, \ldots, 2^N$ and $H_k$ is given by [6], and the labeling matrix $B$ is such that $B_l = \mathbb{O}_{l_k}$ with $\mathbb{O}_{l_k}$ is a $m \times N$ matrix, where $\mathbb{O}_{l_k} = [l_{k1}; l_{k2}; \ldots; l_{kN}]$ and $l_k$ are the rows of $L_N$, with $k = 1, 2, \ldots, 2^N$, and $L_{m-N}$ is a labeling matrix of order $m - N$. \[ \square \]

For an $N$-dimensional OS $2^m$-ary constellation in Definition [7] each orthant contains $2^{m-N}$ constellation points. The
first $2^{m-N}$ constellation points correspond to the points in $T = [t_1; t_2; \ldots; t_{M/2^N}]$ (see Definition 1), and are found via $S_1 = TH_1 = T$ (see in Definition 2). The constellation points in $S_k$ with $k \geq 1$ are generated by “folding” the first-orthant points in $S_1$ to other orthant via $S_k = TH_k$. The binary labeling for the proposed $N$-dimensional OS formats is such that the constellation points in a given orthant $S_k$ use the binary labeling $L_{m-N}$, i.e., $m - N$ bits are used within an orthant. The remaining $N$ bits (which define the matrix $O_k$ in Definition 2) are the bits used to select the orthant. To clarify this general definition, we present now two examples. We focus on traditional square $2^m$-ary QAM constellations labeled by the binary-reflect Gray code (in one and two polarizations), which are shown to belong to the class of OS labeled constellations.

**Example 2 (16QAM):** Fig. 1 shows a 16QAM constellation ($N = 2, m = 4$) labeled by the binary-reflect Gray code independently in the first and second dimension. 16QAM is quadrant-symmetric (four quadrants) and consists of the same number of signal points (four points) in each quadrant. The corresponding first-quadrant (orthant) labeled constellation is

$$S_1 = T = \begin{bmatrix} 3 & 3 \\ 1 & 3 \\ 1 & 1 \\ 3 & 1 \end{bmatrix}, \quad B_1 = [O_1, L_2] = \begin{bmatrix} 00 & 00 \\ 00 & 01 \\ 00 & 11 \\ 00 & 10 \end{bmatrix}.$$

The constellation matrix $S$ and labeling matrix $B$ of 16QAM satisfy $S_k = TH_k$ and $B_k = [O_k, L_{m-N}]$ with $k = 1, 2, 3, 4$ in Definition 2 and thus, this format is an OS labeled constellation.

**Example 3 (PM-16QAM):** Consider a polarization-multiplexed QAM constellation (PM-16QAM, $N = 4, m = 8$) labeled by the binary-reflect Gray code, independently in each real dimension. This constellation is obtained as a Cartesian product of two 2D–16QAM, and therefore, PM-16QAM consists of 16 orthants with 16 points in each orthant. The first-orthant labeled constellation $\{S_1, B_1\}$ of PM-16QAM is given by

$$S_1 = T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 3 & 3 & 1 & 1 \\ 3 & 1 & 1 & 1 \end{bmatrix}, \quad B_1 = [O_1, L_4] = \begin{bmatrix} 0000 & 0000 \\ 0000 & 0100 \\ 0000 & 1100 \\ 0000 & 1000 \end{bmatrix}.$$
which consists of $M = 2^m$ points $s_i, i \in \{1, 2, \ldots, 128\}$ labeled by 7 bits $b_i = [b_1, b_2, \ldots, b_7]$. For the 4D-OS128 constellation, each orthant contains $2^{m-N} = 8$ constellation points, and therefore the 8 constellation points in first orthant are considered as the first-orthant labeled constellation $T = [t_1; t_2; \ldots; t_8]$ with labeling matrix $G_3$. The $j$th first-orthant point is denoted by $t_j = [t_{j1}, t_{j2}, t_{j3}, t_{j4}] \in \mathcal{R}_4^+$ with $j = \{1, 2, \ldots, 8\}$. The first orthant is labeled by four binary bits $[b_{j1}, b_{j2}, b_{j3}, b_{j4}] = l_1$. The remaining 3 bits $[b_{j5}, b_{j6}, b_{j7}]$ determine the point $t_j$ in the corresponding orthant.

For the 4D-OS128 constellation, only eight 4D coordinates in the matrix $T$ and the corresponding binary labeling of order 3 $L_3$ need to be optimized in (9). We solve the optimization problem (9) under a circularly symmetric Gaussian noise assumption, i.e., $p_Y|X$ is fully determined by the variance $\sigma_2^2$, which corresponds to the total power per two real dimensions.

We optimized the 4D constellation at the target GMI of 0.85 mbit/4D-sym, where $m$ is the number of bits transmitted per symbol. In other words, the 4D-OS128 modulation was optimized to minimize the SNR requirements for a GMI of $0.85 \times 7 = 5.95$ mbit/4D-sym. The obtained 4D-OS128 modulation format has 128 nonoverlapping points in 4D space. For better visualization, these points can be projected on the two polarizations. This projection results in 20 distinct points in each 2D space, as shown in Fig. 2(a). In order to clearly show the inter-polarization dependency, we use a similar color coding strategy as in [25]: 2D projected symbols in the first and second polarization are valid 4D symbols only if they share the same color. The coordinates of the 8 vectors defining the matrix $T$ are

$$t_j \in \{[\pm t_1, \pm t_1, \pm t_3, \pm t_3], [\pm t_2, \pm t_5, \pm t_3, \pm t_3], [\pm t_5, \pm t_2, \pm t_3, \pm t_3], [\pm t_4, \pm t_4, \pm t_3, \pm t_3],$$

$$[\pm t_2, \pm t_3, \pm t_1, \pm t_1], [\pm t_3, \pm t_3, \pm t_2, \pm t_5], [\pm t_5, \pm t_2, \pm t_3, \pm t_5],$$

$$[\pm t_3, \pm t_2, \pm t_4, \pm t_4],\}$$

where $t_5 > t_4 > t_3 > t_2 > t_1 > 0$. These 8 points are represented in Fig. 2(a) with gray, red, brown, blue, green, magenta, orange and cyan markers in the shadow area, respectively. The points with the same color in other orthants can be obtained $S_k = T H_k$, and therefore, the proposed format is highly symmetric in 16 orthants.

The color coding scheme used in Fig. 2(a) also shows the binary labeling: 4 out of 7 bits $[b_{j1}, b_{j2}, b_{j3}, b_{j4}]$ determine the 16 orthant, and the remaining 3 bits $[b_{j5}, b_{j6}, b_{j7}]$ determine the 8 constellation points in the corresponding orthant. In other words, $[b_{j5}, b_{j6}, b_{j7}]$ determines the color of the transmitted points in Fig. 2(a), while $[b_{j1}, b_{j2}, b_{j3}, b_{j4}]$ determine the coordinate of the point in the same color. Fig. 2(b) shows an example of 4D mapping and demapping with $[b_{j5}, b_{j6}, b_{j7}] = [0, 0, 1]$, which indicates that one of the red 4D points in Fig. 2(b) is selected as the transmitted symbol $x$.

The coordinates of the 4D-OS128 modulation format and the corresponding binary labeling are given in Appendix (see Table II).

D. Comparison with Other Modulation Formats

Here we include two other well-known 4D 128-ary formats for comparisons at the same SE of 7 bit/4D-sym. The first constellation is 128-ary set-partitioning 16QAM (128SP-16QAM) (see Fig. 3(a)), which has been demonstrated for optical communications systems in [22]–[24]. The second one is a 7 bit modulation in the 4D-2A8PSK family (7b4D-2A8PSK) [24] (see Fig. 3(b)), which is currently used in state-of-the-art commercial programmable transponders [49]. The inner ring/outer ring ratio of 0.59 is chosen for 7b4D-2A8PSK to maximize the GMI performance as described in [24].

To better understand the GMI performance and tolerance to fiber nonlinearities, we conduct a comparison in terms of energy per transmitted symbol, peak-to-average power ratio (PAPR), variance of the signal energy with respect to its meanaverage energy $\sigma_2^2 = E[|S|^2] - E[|\mathcal{S}|^2]$, squared
Euclidean distance (SED), and the numbers of pairs of constellation points at minimum squared Euclidean distance (MSED). While PAPR and energy per transmitted symbol can be seen as a rough indication of nonlinearity tolerance, SED and the number of pairs at MSED can be argued to be performance indicators for the AWGN channel. The analysis and discussion below only gives an intuition on the performance of the proposed format. A precise comparison of these modulation formats will be presented in both the linear and nonlinear channels by numerical simulations (Sec. III) and experimental results (Sec. IV).

We assume that all the constellations are normalized to $E_s = 2$ (i.e., unit energy per polarization). Under this assumption, the comparison of energy for each 4D symbol for all three modulation formats are shown Fig. 3. The 7b4D-2A8PSK format has a constant-modulus property, and thus, it can significantly reduce the nonlinear interference noise (NLIN). For 128SP-16QAM, 5 energy levels are visible, while the proposed 4D-OS128 shows only 3.

Table I shows four properties of the four modulation formats under consideration. In the first two columns, we use two performance metrics to compare the modulation-dependent nonlinear interference: PAPR and $\sigma_{\text{S}}^2$, for a given modulation format. Due to the constant-modulus property, both of these two performance metrics are zero for 7b4D-2A8PSK, which is expected to be better than the other two modulation formats in terms of effective SNR. Since PAPR only depends on the few constellation points with largest energy, it cannot reflect the complete nonlinear performance. In contrast, $\sigma_{\text{S}}^2$ is the variation of all the possible transmitted symbols’ energy, and thus smaller $\sigma_{\text{S}}^2$ should in principle result in higher nonlinear noise tolerance. We expect that this smaller variation in terms of energies will give 4D-128SP-16QAM a small nonlinearity tolerance. These predictions will be confirmed in Sec. III.

In addition to the nonlinear noise tolerance property, we study the structure of the formats in terms of MSED, which we denote by $d_s$. We also look at the number of pairs of constellation points at MSED, which we denote as $n$. These two parameters are shown in the last two columns of Table I. A large $d_s$ and small $n$ should in principle result in high MI in the high-SNR regime, as recently proved in [47]. Even though the proposed 4D-OS128 has the smallest $d_s$, it has only 16 pairs at MSED. 4D-128SP-16QAM and 7b4D-2A8PSK have a large number of pairs of constellation points at MSED, which will degrade the AIR performance at medium SNR range. To better understand this, we also study the SED “spectrum” for the three constellations. This is shown as a histograms in Fig. 5.

It has been shown in [25] that GMI does not only depend on $d_s$ and $n$, but also the Hamming distance (HDs) of the binary labels of the constellation points at MSED. This figure also shows a classification of the pairs at a given SED: blue bars for pairs at HD larger than one, and red bars for pairs at HD one. From the SED spectra in Fig. 5, we can see that there are less pairs at lower SED and most of the pairs are at Hamming distance one for the proposed 4D-OS128, which in principle results in a better GMI in medium SNR range.

### III. Simulation Results

#### A. Linear Performance

Considering the 128SP-16QAM and 7b4D-2A8PSK as the baselines, Fig. 6 shows the linear performance in terms of GMI for the proposed 4D-OS128 modulation format. The results in Fig. 6 indicates that 4D-OS128 can provide gains of 0.65 dB at GMI of 5.95 bit/4D-sym. Meanwhile, the proposed 4D-OS128 can provide 0.27 bit/4D-sym gain over 128SP-16QAM at SNR=9.5 dB. Even though 7b4D-2A8PSK performs better than 128SP-16QAM for SNR below 10 dB, there is at least

3The results in [47] hold for 1D constellations only. However, the authors in [47] conjectured that the results holds verbatim for any number of dimensions.
Fig. 5. Histograms of SEDs of three 4D formats: (a) 4D-128SP-16QAM, (b) 7b4D-2A8PSK, and (c) 4D-OS128. The red bars show the number of pairs with Hamming distance of 1 at the SED $d^2$. The MSEDs are 0.8, 0.23, and 0.14, for 4D-128SP-16QAM, 7b4D-2A8PSK, and 4D-OS128, respectively.

a gap of 0.5 dB between 7b4D-2A8PSK and 4D-OS128. In addition, starting with the 4D-OS128 in Fig. 2 as an initial constellation, we further optimize by using (5), which removes the orthant-symmetric constraint.

The optimization process can reach steady state within 2000 steps as shown in the inset (b) of Fig. 6. The optimized 4D geometrically-shaped (4D-GS128) constellation without orthant-symmetric constraint is plotted in the inset (a) of Fig. 6. The red and blue circles represent the symbols transmitted in X and Y polarization respectively. Despite the 0.027 bit/4D-sym improvement provided by the optimization, the result is a constellation where symbols are very close to each other in the 2D space. This rather complex constellation becomes particularly challenging to generate using a high-speed DAC with limited effective number of bits (ENOB). In this paper, we only investigate the performance of the proposed orthant-symmetric 4D-OS128 modulation format.

For verifying GMI results, we use LDPC codes from the DVB-S2 standard with code rates $R \in \{0.83, 0.8\}$ (20% and 25% overhead (OH)) and blocklength $N = 64800$. Fig. 7 shows post-FEC BER of $\text{BER}=4.5 \times 10^{-3}$, between 0.55 dB and 0.65 dB, which is in excellent agreement with the prediction of the GMI. Moreover, we also evaluate the efficiency of the 4D modulations by using low complexity max-log demapper (MaxLog), which significantly reduces the computational com-

---

Note that this process is not guaranteed to find the globally optimum constellation.
complexity by avoiding logarithmic and exponential functions as opposed to the optimum Maximum likelihood (ML) demapper. Note that using a MaxLog approximation for the proposed 4D-OS128 leads to no observable degradation with respect to the ML demapper. A slightly larger penalty is observed for 128SP-16QAM, in all cases using either 20% or 25% OH.

B. Nonlinear Performance: Multi-span WDM transmission

We consider a dual-polarization long haul WDM transmission system with 11 co-propagating channels generated at a symbol rate of 45 GBaud, a WDM spacing of 50 GHz and a root-raised-cosine (RRC) filter roll-off factor of 0.1. Each WDM channel carries $2^{16}$ 4D symbols in two polarizations at the same launch power per channel $P_{ch}$. For the transmission link, a multi-span standard single-mode fiber (SSMF) is used with attenuation $\alpha = 0.21 \text{ dB} \cdot \text{km}^{-1}$, dispersion parameter $D = 16.9 \text{ ps} \cdot \text{nm}^{-1} \cdot \text{km}^{-1}$, and nonlinear coefficient $\gamma = 1.31 \text{ W}^{-1} \cdot \text{km}^{-1}$. Each span consists of an 80 km SSMF through a split-step Fourier solution of the nonlinear Manakov equation with step size 0.1 km and is followed by an erbium-doped fiber amplifier (EDFA) with a noise figure of 5 dB.

At the receiver side, channel selection is firstly applied and then, the signal is downsamplled to 2 samples/symbol. Chromatic dispersion (CD) compensation is performed before applying an RRC matched filter and downsampling to 1 sample/symbol. An ideal phase rotation compensation is performed. Then, log-likelihood ratios (LLRs) are calculated and passed to the soft-decision LDPC decoder. We focus on evaluating the performance of the center WDM channel because it is more affected by the inter-channel crosstalk and non-linear interference.

Fig. 8 shows the effective SNR (after fiber propagation and receiver DSP) for different 4D modulation formats with SE=7 bits/4D-sym over a 5600 km SSMF. As we expected, the proposed 4D-OS128 modulation format achieves the highest GMI compared to 128SP-16QAM without extra SNR penalty, which is consistent with the analysis in Sec. II-D. Therefore, the shaping gain in linear regime can be maintained in nonlinear regime and translates into a reach increase, which will be shown in Fig. 9.

Fig. 9 shows GMI as a function of the transmitted power for different modulation formats with SE=7 bits/4D-sym over a 5600 km SSMF. As we expected, the proposed 4D-OS128 modulation format achieves the highest GMI compared to 128SP-16QAM and 7b4D-2A8PSK. The gains compared to 128SP-16QAM for different launch power are almost constant due to similar effective SNR performance. Meanwhile, the gains compared to 7b4D-2A8PSK is reduced as the launch power increases.

Fig. 10 shows GMI as a function of transmission distance
for different modulation formats with SE=7 bits/4D-sym at optimal launch power. The proposed 4D-OS128 modulation format leads to a 920 km increase in reach relative to the 128SP-16QAM modulation format at GMI of 5.6 bit/4D-sym, and more than 400 km relative to 7b4D-2A8PSK modulation format. For GMI above 6.1 bit/4D-sym, 128SP-16QAM provides the best performance. This can not only be attributed to the larger minimum Euclidean distance for 128SP-16QAM compared to 7b4D-2A8PSK and 4D-OS128. The proposed 4D-OS128 is not designed for higher SNR at shorter transmission distances.

IV. EXPERIMENTAL SETUP AND RESULTS

A. Experimental Transmission Setup

Fig. 10 depicts the experimental transmission setup. The transmitted signal is modulated using either 128SP-16QAM, 7b4D-2A8PSK [24], or 4D-OS128 symbols. Pseudo-random sequences of $2^{16}$ symbols are generated offline, pulse shaped using a RRC filter with 1% roll-off at 41.79 GBd, and uploaded to a 100-GSa/s DAC. The positive differential DAC outputs are connected to the optical multi-format transmitter (OMFT) which consists of an external cavity laser (ECL), a dual-polarization IQ-modulator (DP-IQM), an automatic bias controller (ABC) and RF-amplifiers. The channel under test (CUT), which can be defined at any of the 11 tested C-band channels, is modulated by the OMFT and subsequently amplified. The loading channels are provided by the negative outputs of the DAC and modulated onto the tones provided by 10 ECLs using a DP-IQM. These loading channels are amplified, split into even and odd channels, decorrelated by 10,200 (50 m) and 40,800 symbols (200 m), and multiplexed together with the CUT on a 50-GHz grid using an optical tunable filter (OTF). Bandwidth limitations due to transmitter electronics are initially compensated using an OTF and the residual effects are mitigated digitally as proposed in [59].

The 11-channel 50-GHz-spaced dense wavelength division multiplexing (DWDM) signal is amplified and through an acousto-optic modulator (AOM) enters the recirculating loop which consists of a loop-synchronous polarization scrambler (LSPS), a 75-km span of SSMF, an EDFA, an AOM and an OTF used for gain flattening. The inset of Fig. 11 shows the optical spectrum after 80 circulations, which corresponds to 6000 km of transmission using only EDFA-amplification. Optionally, a hybrid amplification scheme can be used by adding a 750 mW 1445 nm Raman pump in a backward pumping configuration. The output of the recirculating loop is amplified, filtered by a wavelength selective switch (WSS) and digitized by a coherent receiver consisting of a local oscillator (LO), a 90-degree hybrid, four balanced photo-diodes and an 80 GSa/s analog-to-digital converter (ADC). The offline digital signal processing (DSP) includes front-end impairment correction using blind moment estimation, chromatic dispersion compensation, frequency offset estimation and correction between transmitter and LO laser. A widely-linear multiple-input multiple-output (MIMO) equalization [51] with blind phase search (BPS) [52] inside the update loop is employed to correct for phase noise, error counting and GMI evaluation. In the following sections, we evaluate and discuss the results for two configurations as EDFA-only amplification (Sec. IV-B) and a hybrid of EDFA and Raman (Sec. IV-C).

B. Experimental Results: EDFA-only Amplifier

Fig. 12 and Fig. 13 show the 2D projections of the received constellations after optical back-to-back and after transmission over 70 spans respectively. Note that the proposed 4D-OS128 modulation induces nonuniform probability distribution when projected onto 2D, which is similar to probabilistic amplitude shaped 16QAM with very short blocklength of 4. Instead of using four time slots as PS, the 4D-OS128 shapes the constellation using the two quadratures (I/Q) and the two polarization states (X/Y) as four dimensions.

Fig. 14 (a) shows the GMI as a function of transmission distance for the optimal launch power of 9.5 dBm. The GMI in Fig. 14 (a) should be interpreted as the reach that an ideal SD-FEC would achieve. For the considered rate (6 bit/4D-sym), 7b4D-2A8PSK offers approximately the same reach as 128SP-16QAM around 5300 km. The proposed 4D-OS128 reaches 6110 km, which corresponds to a gain of 810 km (15%). Fig. 14 (a) also shows that the average GMI per channel resulting in a 0.26 bit/4D increase for 4D-OS128 with respect to 128SP-16QAM after 6340 km transmission. The GMI vs. launch power for a transmission distance of 6000 km and the three modulation formats under consideration are also shown as inset in Fig. 14. At the optimal launch power, 4D-OS128 outperforms 128SP-16QAM and 7b4D-2A8PSK with a gain of 0.22 bits/4D-sym.

BER performance before and after FEC are shown in Fig. 14 (b). For the experiment, 20 low-density parity-check (LDPC) blocks are constructed, which are then encoded using the DVB-S2 LDPC code with 20% overhead and code length $n = 64800$. An outer hard decision forward error correction (HD-FEC) staircase code with rate 0.9373 [54] that corrects bit errors after LDPC decoding is assumed. The BER threshold is $4.5 \times 10^{-3}$ [54, Fig. 8], which makes 4D-OS128 15%...
Section C: Experimental Results: Hybrid Amplification

In the hybrid amplification scheme, a 750 mW 1445 nm Raman pump is also used in a backward configuration as shown in Fig. 11.
Proposed 4D-OS128

Fig. 15. Experimental results using EDFA-only amplification. Per-channel performance versus GMIs (left) and BERs (right) measured for all 11 channels individually after 6000 km showing GMIs above 5.95 bit/4D \[24\] and BERs below the FEC threshold $4.1 \cdot 10^{-2}$ \[53\].

The GMI vs. launch power for a transmission distance of 9000 km and the three modulation formats under consideration are also shown as inset in Fig. 16 (a). At the optimal launch power of 6.5 dBm, 4D-OS128 maximizes the average GMI per channel resulting in a 0.22 bits/4D-sym increase for with respect to 128SP-16QAM and 7b4D-2A8PSK. Therefore, we use 6.5 dBm as launch power to evaluate the transmission performance. Fig. 16 (a). shows the GMI as a function of transmission distance. We can observe that the relative GMI gains of 0.25 bit/4D-sym and 1100 km (13.5\%) are similar to the EDFA-only case.

BER performance before and after FEC are shown in Fig. 16 (b). Under the assumption of concatenated LDPC and staircase code with rate 0.9373, 4D-OS128 shows a post-FEC reach increase of 1150 km (14\%). By measuring all the 11 WDM channels’ performance, Fig. 17 shows all channels above the GMI threshold of 5.95 bit/4D-sym (0.85 NGMI) enabling 9000 km error-free transmission of net 233 Gbit/s per channel after 25.5% overhead. Comparing to the results of the EDFA-only amplification in Sec. IV-B, utilizing Raman amplifier causes an apparent improvement (50% reach increase) on the transmission performance compared with using the EDFA for the 4D-OS128 modulation.

V. CONCLUSIONS

A new modulation format (4D-OS128) with a spectral efficiency of 7 bits/4D-sym was introduced using the concept of orthant symmetry. The format was designed based on the generalized mutual information, and thus, it finds applications to systems with soft-decision forward error correcting codes and bit-wise decoding. The 4D-OS128 format provides sensitivity gains of up to 0.65 dB after LDPC decoding versus state-of-the-art formats. The experimental results confirm the overall superior receiver sensitivity and reliability of 4D-OS128 versus previously published formats. Transmission reach extensions of more than 15\% is demonstrated. We believe that the proposed format is a good alternative for future high capacity long haul transmission systems, which provides an intermediate solution between PM-8QAM and PM-16QAM. The design of orthant-symmetric constellation for higher dimensions (e.g., 16 dimensions) and larger constellation sizes (e.g., 256-ary formats) is left for further investigation.

APPENDIX A

COORDINATES AND BINARY LABELING FOR 4D-OS128

Table II lists the coordinates of the constellation points and the bit-to-symbol mapping of 4D-OS128. The constellation is assumed to be normalized to $E_s = 2$, i.e., to unit energy per polarization.
TABLE II
Coordinates and binary labeling of the proposed 4D-OS128 format. The coordinates are rounded to four decimal points:
\[ (t_1, t_2, t_3, t_4, t_5) = (0.2875, 0.3834, 0.4730, 1.1501, 1.2460) \]

| Coordinates | Labeling | Coordinates | Labeling |
|-------------|----------|-------------|----------|
| \(+t_3 +t_1 +t_4 +t_1\) | 000011 | \(+t_2 +t_5 +t_3 +t_4\) | 000001 |
| \(+t_3 +t_1 +t_5 +t_1\) | 000011 | \(-t_2 -t_5 +t_3 +t_4\) | 110000 |
| \(+t_3 -t_1 +t_5 +t_1\) | 010011 | \(+t_2 -t_5 +t_3 +t_4\) | 010000 |
| \(+t_3 +t_1 +t_5 +t_5\) | 000111 | \(-t_2 -t_5 +t_3 +t_4\) | 110000 |
| \(+t_3 -t_1 +t_5 +t_5\) | 010111 | \(+t_2 -t_5 +t_3 +t_4\) | 010000 |
| \(+t_1 +t_2 +t_3 +t_5\) | 000100 | \(+t_4 +t_2 +t_5 +t_5\) | 000000 |
| \(+t_1 +t_2 +t_3 +t_4\) | 010100 | \(+t_4 +t_2 +t_3 +t_4\) | 010000 |
| \(+t_1 +t_2 +t_3 +t_5\) | 010111 | \(+t_4 +t_2 +t_3 +t_4\) | 010000 |
| \(+t_1 +t_2 +t_3 +t_5\) | 010111 | \(+t_4 +t_2 +t_3 +t_4\) | 010000 |
| \(+t_1 +t_2 +t_3 +t_5\) | 010111 | \(+t_4 +t_2 +t_3 +t_4\) | 010000 |
| \(+t_1 +t_2 +t_3 +t_5\) | 010111 | \(+t_4 +t_2 +t_3 +t_4\) | 010000 |
| \(+t_1 +t_2 +t_3 +t_5\) | 010111 | \(+t_4 +t_2 +t_3 +t_4\) | 010000 |
| \(+t_1 +t_2 +t_3 +t_5\) | 010111 | \(+t_4 +t_2 +t_3 +t_4\) | 010000 |
| \(+t_1 +t_2 +t_3 +t_5\) | 010111 | \(+t_4 +t_2 +t_3 +t_4\) | 010000 |

ACKNOWLEDGEMENT
Fraunhofer HHI and ID Photonics are gratefully acknowledged for providing their Optical-Multi-Format Transmitter.

REFERENCES
[1] H. Bülow, “Polarization QAM modulation (POL-QAM) for coherent detection schemes,” in 2009 Optical Fiber Communication Conference (OFC), Mar. 2009.
[2] P. J. Winzer, A. H. Gauck, C. R. Doerr, M. Magarini, and L. L. Buhl, “Spectrally efficient long-haul optical networking using 112-Gb/s polarization-multiplexed 16-QAM,” Journal of Lightwave Technology, vol. 28, no. 4, pp. 547–556, Feb. 2010.
[3] F. Buchali, F. Stein, 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, Apr. 2016.
[4] T. Fehrenberger, A. Alvarado, G. Böcherer, and N. Hanik, “On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel,” Journal of Lightwave Technology, vol. 34, no. 21, pp. 5063–5073, Nov. 2016.
[5] G. Böcherer, F. Stein, and P. Schulte, “Fast probabilistic shaping implementation for long-haul fiber-optic communication systems,” in 2017 European Conference on Optical Communication (ECOC), Sep. 2017.
[6] F. Buchali, W. Idler, R. Dirschler, T. Eriksson, and L. Schmalen, “Spectrally efficiently probabilistically shaped square 64QAM to 256QAM,” in 2017 European Conference on Optical Communication (ECOC), Sep. 2017.
[7] R. Maher, K. Croussore, M. Lauermann, R. Going, X. Xu, and J. Rahn, “Constellation shaped 66 GbD DP-1024QAM transceiver with 400 km transmission over standard SMF,” in 2017 European Conference on Optical Communication (ECOC), Sep. 2017.
[8] A. Amari, S. Goossens, Y. C. Güttekin, O. Vassilieva, I. Kim, T. Ikeuchi, C. Okonkwo, F. M. J. Willems, and A. Alvarado, “Introducing enumerative sphere shaping for optical communication systems with short blocklengths,” Journal of Lightwave Technology, vol. 37, no. 23, pp. 5926–5936, Dec. 2019.
[9] S. Goossens, S. van der Heide, M. van den Hout, A. Amari, Y. C. Güttekin, O. Vassilieva, I. Kim, F. M. J. Willems, A. Alvarado, and C. Okonkwo, “First experimental demonstration of probabilistic enumerative sphere shaping in optical fiber communications,” in Opto-Electron. and Commun. Conf. (OECC) and Int. Conf. on Photon. in Switch. and Comput. (PSC), Fukuoka, Japan, July 2019.
[10] Z. Qu and I. B. Djordjevic, “Geometrically shaped 16QAM outperforming probabilistically shaped 16QAM,” in 2017 European Conference on Optical Communication (ECOC), Sep. 2017.
[11] S. Zhang, F. Yaman, E. Mateo, T. Inoue, K. Nakamura, and Y. Inada, “Design and performance evaluation of a GMI-optimized 32QAM,” in 2017 European Conference on Optical Communication (ECOC), Sep. 2017.
[12] B. Chen, C. Okonkwo, H. Hafermann, and A. Alvarado, “Increasing achievable information rates via geometric shaping,” in 2018 European Conference on Optical Communication (ECOC), Sep. 2018.
[13] B. Chen, C. Okonkwo, D. Lavery, and A. Alvarado, “Geometrically-shaped 64-point constellations via achievable information rates,” in 2018 26th International Conference on Transparent Optical Networks (ICTON), July 2018, pp. 1–4.
[14] L. Szczecinski and A. Alvarado, Bit-interleaved coded modulation: fundamentals, analysis and design. John Wiley & Sons, 2015.
[15] F. Stein and G. Boecherer, “Comparison of geometric and probabilistic shaping with application to ATSC 3.0,” in 11th International ITG Conference on Systems, Communications and Coding (SCC2017), Feb. 2017, pp. 1–6.
[16] E. Agrell and M. Karlsson, “Power-efficient modulation formats in coherent transmission systems,” Journal of Lightwave Technology, vol. 27, no. 22, pp. 5115–5126, Nov. 2009.
[17] M. Karlsson and E. Agrell, “Which is the most power-efficient modulation format in optical links?” Opt. Express, vol. 17, no. 13, pp. 10814–10819, June 2009.
[18] T. Koike-Akino, D. S. Millar, K. Kojima, and K. Parsons, “Eight-dimensional modulation for coherent optical communications,” in 2013 European Conference on Optical Communication (ECOC), Sep. 2013.
A. Alvarado and E. Agrell, “Four-dimensional coded modulation with bit-wise decoders for future optical communications,” in Journal of Lightwave Technology, vol. 43, pp. 158–162, 2018.

A. Alvarado, E. Agrell, G. Durisi, and M. Karlsson, “Capacity of a nonlinear optical channel with finite memory,” in IEEE Journal of Lightwave Technology, vol. 32, no. 16, pp. 2862–2876, Aug. 2014.

R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, “Inter-channel nonlinear interference noise in WDM systems: Modeling and mitigation,” in Journal of Lightwave Technology, vol. 33, no. 5, pp. 1044–1053, Mar. 2015.

“Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications,” in Journal of Lightwave Technology, vol. 34, no. 2, pp. 593–607, Jan. 2016.

A. Ghazisaeidi, “A theory of nonlinear interactions between signal and amplified spontaneous emission noise in coherent wavelength division multiplexed systems,” in Journal of Lightwave Technology, vol. 35, no. 23, pp. 5150–5175, Dec. 2017.

G. Caire, G. Taricco, and E. Biglieri, “Bit-interleaved coded modulation,” in IEEE Transaction on Information Theory, vol. 44, no. 3, pp. 927–946, May 1998.

E. Agrell and A. Alvarado, “Optimal alphabets and binary labelings for BICM at low SNR,” in IEEE Transaction on Information Theory, vol. 57, no. 10, pp. 6650–6672, Oct. 2011.

A. Alvarado, F. Brännström, E. Agrell, and T. Koch, “High-SNR asymptotics of mutual information for discrete constellations with applications to BICM,” in IEEE Transaction on Information Theory, vol. 60, no. 2, pp. 1061–1076, Feb. 2014.

A. Alvarado, F. Brännström, and E. Agrell, “High SNR bounds for the BICM capacity,” in IEEE Information Theory Workshop, Paraty, Brazil, Nov. 2011.

“IFINITY T600 transport blade,” [Online]. Available: https://www.fujitsu.com/us/Images/IFINITY-T600-Data-Sheet.pdf

J. Lin, H. Sepehrian, L. A. Rusch, and W. Shi, “Joint digital and optical pre-compensation for 720 Gb/s all-silicon IQ modulator single carrier transmission,” in 2018 European Conference on Optical Communication (ECOC), Sep. 2018.

E. P. da Silva and D. Zibar, “Widely linear equalization for IQ imbalance and skew compensation in optical coherent receivers,” in Journal of Lightwave Technology, vol. 34, no. 15, pp. 3577–3586, Aug. 2016.

T. Pfau, S. Hoffmann, and R. Noé, “Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for m-QAM constellations,” in Journal of Lightwave Technology, vol. 27, no. 8, pp. 989–999, Apr. 2009.

K. Sugihara, Y. Miyata, T. Sugihara, K. Kubo, H. Yoshida, W. Matsumoto, and T. Mizuochi, “A spatially-coupled type LDPC code with an NCG of 12 dB for optical transmission beyond 100 Gb/s,” in 2013 Optical Fiber Communications Conference (OFC), Mar. 2013.

B. P. Smith, A. Farhood, A. Hunt, F. R. Kschischang, and J. Lodge, “Staircase codes: FEC for 100 Gb/s OTN,” in Journal of Lightwave Technology, vol. 30, no. 1, pp. 110–117, Jan. 2012.