Power-Efficient Voronoi Constellations for Fiber-Optic Communication Systems

Citation for the original published paper (version of record):
Li, S., Mirani, A., Karlsson, M. et al (2023). Power-Efficient Voronoi Constellations for Fiber-Optic Communication Systems. Journal of Lightwave Technology, 41(5): 1298-1308.
http://dx.doi.org/10.1109/JLT.2022.3222423

Power-Efficient Voronoi Constellations for Fiber-Optic Communication Systems

Shen Li, Ali Mirani, Magnus Karlsson, *Senior Member, IEEE, Fellow, OSA,* and Erik Agrell, *Fellow, IEEE*

***Abstract*—Voronoi constellations (VCs) are considered as an effective geometric shaping method due to their high power efficiencies and low complexity. In this paper, the performance of 16- and 32-dimensional VCs with a variety of spectral efficiencies transmitted in the nonlinear fiber channel are investigated. Both single-channel and wavelength-division multiplexing systems are considered for the transmission of the VCs, as well as different signal processing schemes, including chromatic dispersion compensation and digital backpropagation. Multiple performance metrics including the uncoded bit error rate, mutual information (MI), and generalized mutual information (GMI) of VCs are evaluated. Compared with quadrature amplitude modulation (QAM) formats, the VCs provide 1.0–2.4 dB launch power gains, up to 0.50 bits/symbol/dimension-pair MI gains, up to around 30% potential reach increase at the same MI, and up to 0.30 bits/symbol/dimension-pair GMI gains in a limited launch power range. The observed performance gains over QAM are found higher than in the back-to-back case. Moreover, a general GMI estimation method for very large constellations using importance sampling is proposed for the first time.**

***Index Terms*—Fiber-optic communication, generalized mutual information, geometric shaping, information rates, lattices, multidimensional modulation formats, Voronoi constellations.**

I. Introduction

GEOMETRIC shaping is a way to gain power efficiency by adjusting the position of constellation points with respect to uniform quadrature amplitude modulation (QAM). In recent years, great interests have been shown in performing shaping in fiber-optical communications, partially because no theoretical limits have been found for the maximum shaping gain of the nonlinear fiber channel yet, which could potentially be more than the ultimate 1.53 dB shaping gain in the linear additive white Gaussian noise (AWGN) channel. Actually, there have been indications showing that geometrically shaped modulation formats might achieve gains higher than 1.53 dB [1]. On the other hand, some works show that probabilistically shaped constellations introduce a modulation-dependent nonlinear interference for fiber transmission, which reduces the shaping gain achieved in the AWGN channel [2], [3].

Coherent fiber communication inherently consists of four dimensions: two orthogonal quadratures in two orthogonal polarizations. More dimensions can be realized utilizing time slots, wavelengths, and spatial dimensions, and performing shaping over higher dimensions jointly is expected to achieve larger performance gains. A comprehensive review of several geometrically-shaped modulation formats can be found in [4, Table I]. Much work has been devoted to design geometrically shaped 4-dimensional constellations [2], [5]–[8]. Among the work in the literature, usually look-up tables storing all coordinates of constellation points are needed, which makes the detection complexity and storage requirement increase exponentially when extending to higher dimensions.

Voronoi constellations (VCs) based on lattices, inherently performing a joint shaping of multiple dimensions, can be a good trade-off between shaping gain and complexity. VCs have a shaping lattice providing the shaping gain and a coding lattice providing the coding gain, which were first proposed by Conway and Sloane in [9], together with their low-complexity encoding and decoding algorithms, and then generalized by Forney [10]. No look-up tables are needed, neither no dramatic complexity increase in high dimensions. VCs were used for shaping in some wireless network applications [11]–[16]. For the AWGN channel, uncoded BER gains of VCs over QAM were reported in [17], [18], and mutual information (MI) gains of VCs were demonstrated in [19].

VCs were first studied for fiber communications in [18], where Mirani *et al.* reported significant uncoded bit error rate (BER) gains of VCs over QAM transmitted in a wavelength-division multiplexing (WDM) system. Later in [20], power gains of VCs were demonstrated in experiments for a 80 km single-channel transmission.

In this paper, different from the work in [18] and [20], we investigate the performance of another type of VC with a cubic coding lattice in both single-channel and WDM simulations. The considered VCs have different numbers of dimensions, shaping lattices, and spectral efficiencies. Several important performance metrics for both hard- and soft-decision forward error correction (FEC) are evaluated for the VCs, including the uncoded BER, MI, and generalized mutual information (GMI), among which the MI and GMI performances for VCs in the nonlinear fiber channel are demonstrated for the first time to our knowledge. Consistently with [1], the observed power, MI, and GMI gains of VCs over QAM formats in fiber transmission are found higher than in the AWGN channel. The data rates and transmission distances of the studied systems are comparable to 400ZR [21] and the upcoming 800 Gbps and 1.2 Tbps standards. The target systems of VCs could be data center interconnects, e.g., campus and metro data centers, which require rather high throughput with limited complexity for short distances (usually less than 100 km), or metro

This research was funded in part by the Swedish Research Council (VR) under grants no. 2017-03702 and no. 2021-03709 and the Knut and Alice Wallenberg Foundation under grant no. 2018.0090.

S. Li and E. Agrell are with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden. e-mail: shenl@chalmers.se.

A. Mirani and M. Karlsson are with the Department of Microtechnology and Nanoscience, Chalmers University of Technology, 412 96 Gothenburg, Sweden.

optical links with up to hundreds of kilometers. Moreover, a general GMI estimation method for very large constellations is proposed, using importance sampling, extending the MI estimation method in [19].

Notation: Bold lowercase symbols denote row vectors and bold uppercase symbols denote random vectors or matrices. The sets of integer, real, complex, and natural numbers are denoted by $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{N}$, respectively. Other sets are denoted by calligraphic symbols. Rounding a vector to its nearest integer vector is denoted by $\lfloor\cdot\rceil$, in which ties are broken arbitrarily.

II. Basics of VCs

VCs are structured multidimensional lattice-based constellations. An n-dimensional lattice is spanned by n linearly independent basis vectors which are the rows of its $n \times n$ generator matrix $\boldsymbol{G}$. All linear combinations of these basis vectors with integer coefficients form the lattice, i.e,

$$\Lambda \triangleq \{\boldsymbol{u}\boldsymbol{G}:\ \boldsymbol{u} \in \mathbb{Z}^n\}. \tag{1}$$

From the definition, a lattice must contain the all-zero point $\mathbf{0}$. The *Voronoi region* of a lattice Λ, denoted as $\Omega(\Lambda)$, contains all points[1] in $\mathbb{R}^n$ having the all-zero point $\mathbf{0}$ as their closest lattice point in Λ.

A VC has a *coding lattice* Λ and a *shaping lattice* $\Lambda_{\rm s}$, and $\Lambda_{\rm s}$ is a sublattice of Λ, i.e., $\Lambda_{\rm s} \subset \Lambda$. Forney defined a VC based on the lattice partition $\Lambda/\Lambda_{\rm s}$ as a set of translated lattice points of Λ enclosed by the Voronoi region of $\Lambda_{\rm s}$ [10], i.e.,

$$\Gamma \triangleq (\Lambda - \boldsymbol{a}) \cap \Omega(\Lambda_{\rm s}), \tag{2}$$

where $\boldsymbol{a} \in \mathbb{R}^n$ is the *offset vector* (see Fig. 1 for an example VC). The offset vector avoids lattice points falling on the boundary of $\Omega(\Lambda_{\rm s})$, and can be optimized to minimize the average symbol energy of the VC. The number of points in a VC is $M \triangleq |\det \boldsymbol{G}_{\rm s}|/|\det \boldsymbol{G}|$, where $\boldsymbol{G}_{\rm s}$ and $\boldsymbol{G}$ are the generator matrices of $\Lambda_{\rm s}$ and Λ, respectively. The length of the bit labels is $m = \log_2 M$ bits and the *spectral efficiency* [22]–[24] of a VC per dimension-pair is defined as

$$\beta = 2m/n \text{ [bits/symbol/dimension-pair]}. \tag{3}$$

A dimension-pair could refer to any two real dimensions for a multidimensional AWGN channel, and for fiber-optic channels, it could be an in-phase and quadrature (I/Q) pair of a single polarization component. The average symbol energy of a VC is

$$E_{\rm s} = \frac{1}{M}\sum_{\boldsymbol{x}\in\Gamma} \|\boldsymbol{x}\|^2. \tag{4}$$

The VCs based on the lattice partition $\mathbb{Z}^n/\Lambda_{\rm s}$ are the focus of this paper due to the following reasons. First, the decoding process is simpler for such VCs than the classical VCs considered in [18]. Second, the cubic coding lattice enables a "pseudo-Gray" labeling scheme where most of the constellation points differ by only one bit from their adjacent constellation points. Third, the analysis of the MI and log-likelihood ratio becomes possible with a cubic coding lattice, which has never been studied for such large constellations. Fourth, though the cubic coding lattice does not provide any coding gain, the coding gain is provided by FEC coding in the considered optical systems, which partly serves the same purpose as maximizing the minimum distance between constellation points.

The encoding and decoding algorithms (mapping between integers coordinates and constellation points) of VCs based on the lattice partition $\mathbb{Z}^n/\Lambda_{\rm s}$ are summarized in [19, Alg. 1, Alg. 2] and [25]. As seen from the algorithms, the decoder only comprises linear operations. Though it is not a maximum-likelihood decoder, it has a much lower complexity, and there is evidence showing that the performance gap to the maximum-likelihood decoder becomes negligible when M or the signal-to-noise ratio (SNR) increases [18]. The mapping between information bits and integer coordinates is according to the binary reflected Gray code in order to minimize the BER [17]. In this paper, some commonly-used multidimensional lattices are scaled and adopted as the shaping lattices, including the 2-dimensional checkerboard lattice D_2, 4-dimensional checkerboard lattice D_4, 8-dimensional Gosset lattice E_8, 16-dimensional Barnes–Wall lattice Λ_{16} [26, Ch. 4] and a 32-dimensional lattice L_{32} [19, Table I]. The VCs considered in this paper together with their parameters are listed in Table I. The offset vector $\boldsymbol{a}$ is optimized using an iterative algorithm proposed in [9] to minimize the average symbol energy of VCs with small M values, such as D_2^5 and D_4^{17}. For larger VCs, the average symbol energy cannot be minimized via Monte Carlo simulations. We validated that choosing a random $\boldsymbol{a} \in \Omega(\Lambda)$ does not affect the performance significantly for large VCs [18]. The considered VCs only achieve the shaping gains from $\Lambda_{\rm s}$, and no coding gain is provided due to the cubic coding lattice being used. The asymptotic shaping gains (ASGs) [22] determined by $\Lambda_{\rm s}$ of these VCs are also listed in Table I. The ASGs indicate the maximum shaping gains these VCs can achieve over QAM in the linear AWGN channel asymptotically, but these values do not limit their performance gains (e.g., SNR gains over QAM at the same BER) at a finite SNR. Higher ASGs might imply higher performance gains in the nonlinear fiber channel.

Example 1: A simple example is D_2^5 having a shaping lattice which does not provide any shaping gain but is good for illustration. The generator matrices of $4D_2$ and $\mathbb{Z}^n$ are

$$\boldsymbol{G}_{\rm s} = \begin{pmatrix} 8 & 0 \\ 4 & 4 \end{pmatrix},\ \boldsymbol{G} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{5}$$

and the offset vector $\boldsymbol{a} = (-1/2, 0)$. Fig. 1 illustrates the mapping rule between bits and constellation points for this example.

III. GMI estimation

The amount of information per symbol that a certain channel can transmit with an arbitrarily small error probability, using any (optimal or suboptimal) encoder/decoder pair, is known as an achievable information rate (AIR). The maximum AIRs over a certain channel indicating the fundamental limits of

[1]Arbitrary points in $\mathbb{R}^n$ are denoted as "points" in this paper. To avoid ambiguity, "lattice point" is used when a point also belongs to a lattice. Later throughout the paper, "constellation points" refers to the points in VCs.

TABLE I: The VCs and their parameters considered in this paper.

Name	$\Lambda/\Lambda_{\mathrm{s}}$	M	m	β	$\boldsymbol{a}$	ASG [dB]
D_2^5	$\mathbb{Z}^2/4D_2$	32	5	5	$(-0.5, 0)$	0
D_4^{17}	$\mathbb{Z}^4/16D_4$	131072	17	8.5	$(0.500, 0, -0.168, 0.334)$	0.37
E_8^{24}	$\mathbb{Z}^8/8E_8$	16777216	24	6	$\in \Omega(\mathbb{Z}^8)$	0.65
Λ_{16}^{60}	$\mathbb{Z}^{16}/8\Lambda_{16}$	$\approx 1.2 \times 10^{18}$	60	7.5	$\in \Omega(\mathbb{Z}^{16})$	0.86
Λ_{16}^{76}	$\mathbb{Z}^{16}/16\Lambda_{16}$	$\approx 7.6 \times 10^{22}$	76	9.5	$\in \Omega(\mathbb{Z}^{16})$	0.86
Λ_{16}^{92}	$\mathbb{Z}^{16}/32\Lambda_{16}$	$\approx 5.0 \times 10^{27}$	92	11.5	$\in \Omega(\mathbb{Z}^{16})$	0.86
L_{32}^{123}	$\mathbb{Z}^{32}/8L_{32}$	$\approx 1.1 \times 10^{37}$	123	7.6875	$\in \Omega(\mathbb{Z}^{32})$	0.94
L_{32}^{155}	$\mathbb{Z}^{32}/16L_{32}$	$\approx 4.6 \times 10^{46}$	155	9.6875	$\in \Omega(\mathbb{Z}^{32})$	0.94
L_{32}^{187}	$\mathbb{Z}^{32}/32L_{32}$	$\approx 2.0 \times 10^{56}$	187	11.6875	$\in \Omega(\mathbb{Z}^{32})$	0.94

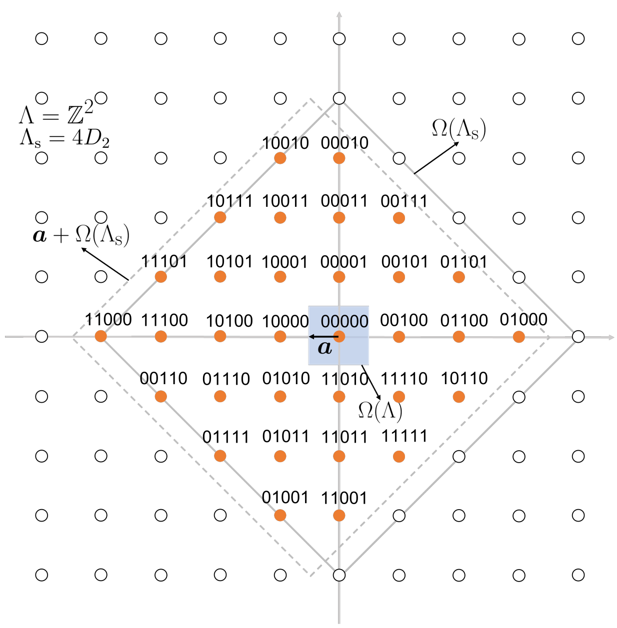


Fig. 1: An illustration of the VC in Example 1.

a coded modulation scheme, are the MI and GMI, of which the former applies to nonbinary codes or multilevel codes, and the latter applies to binary interleaved coded modulation (BICM) [27]. The MI and GMI of VCs transmitted in the nonlinear fiber channel are definitely worth investigation. However, their calculation usually requires full enumeration of all constellation points, which is not applicable to high-dimensional VCs with high cardinalities. Thus, a method based on importance sampling [28, Ch. 9] was proposed to solve this challenge in [19]. In this section, the same idea is extended to the estimation of GMI for very large constellations in general cases, and then specifically applied to VCs.

A. GMI estimation method based on importance sampling

We consider a modulation format $\mathcal{X}$ consisting of M equally probable symbols with distinct binary labels of length $m = \log_2 M$. For a memoryless discrete channel with the conditional probability $f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})$, where $\boldsymbol{x} \in \mathcal{X}$ denotes the transmitted symbol and $\boldsymbol{y} \in \mathbb{R}^n$ denotes the received noisy symbol, the GMI can be written as [27, Eq. (15)]

$$\mathrm{GMI} \triangleq \frac{1}{M} \sum_{k=1}^{m} \sum_{b\in\{0,1\}} \sum_{\boldsymbol{x}\in\mathcal{X}_k^b} \int_{\mathbb{C}^n} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) \log_2 \frac{f_{k,b}(\boldsymbol{y})}{\frac{1}{2}\left(f_{k,0}(\boldsymbol{y}) + f_{k,1}(\boldsymbol{y})\right)} d\boldsymbol{y}, \tag{6}$$

where $\mathcal{X}_k^b \subset \mathcal{X}$ is the set of constellation points with a bit b at position k in their m-bit binary label, and

$$f_{k,b}(\boldsymbol{y}) = \frac{2}{M} \sum_{\boldsymbol{x}\in\mathcal{X}_k^b} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}). \tag{7}$$

If N_{s} samples are uniformly and independently drawn from $\mathcal{X}$ and transmitted through the simulated channel $f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})$, then the GMI in (6) can be approximated using these N_{s} channel realization pairs $(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)})$, where the superscripts represent the time index. Let the set $\mathcal{Z}_k^b$ denote all the time samples of input symbols $\boldsymbol{x}^{(i)}$ that have a bit b at position k. Then (6) can be approximated by

$$\mathrm{GMI} \approx \frac{1}{N_{\mathrm{s}}} \sum_{k=1}^{m} \sum_{b\in\{0,1\}} \sum_{i\in\mathcal{Z}_k^b} \log_2 \frac{f_{k,b}(\boldsymbol{y}^{(i)})}{\frac{1}{2}\left(f_{k,0}(\boldsymbol{y}^{(i)}) + f_{k,1}(\boldsymbol{y}^{(i)})\right)}. \tag{8}$$

It can be noted that enumerating all constellation points in $\mathcal{X}$ to calculate (7) is practically impossible when M it too large. Extending the idea from [19], instead of enumerating every point in $\mathcal{X}$, we consider a much smaller importance set $\mathcal{I}(\boldsymbol{y})$, such that the contribution to the sum $f_{k,b}(\boldsymbol{y})$ from constellation points in $\mathcal{X} - \mathcal{I}(\boldsymbol{y})$ is negligible. Then $f_{k,b}(\boldsymbol{y})$ can be approximated by enumerating all constellation points in $\mathcal{I}^{k,b}(\boldsymbol{y}) = \mathcal{I}(\boldsymbol{y}) \cap \mathcal{X}_k^b$. The importance set $\mathcal{I}(\boldsymbol{y})$ can be further divided into D disjoint subsets $\mathcal{I}_d(\boldsymbol{y})$ for $d = 1, \ldots, D$. From each subset, K_d random samples $\boldsymbol{x}_{d,j}$ for $j = 1, \ldots, K_d$ are drawn uniformly if $K_d < |\mathcal{I}_d(\boldsymbol{y})|$, unless $|\mathcal{I}_d(\boldsymbol{y})|$ is small enough so that enumerating all constellation points in it is more computationally efficient, then $\boldsymbol{x}_{d,j}$ for $j = 1, \ldots, K_d$ are all constellation points in $\mathcal{I}_d(\boldsymbol{y})$, where $K_d = |\mathcal{I}_d(\boldsymbol{y})|$. If we call the subscript j "sample indices", and let the set $\mathcal{T}_d^{k,b}(\boldsymbol{y}) \subseteq \{1, \ldots, K_d\}$ denote all the sample indices of the random samples (or all constellation points in small subsets

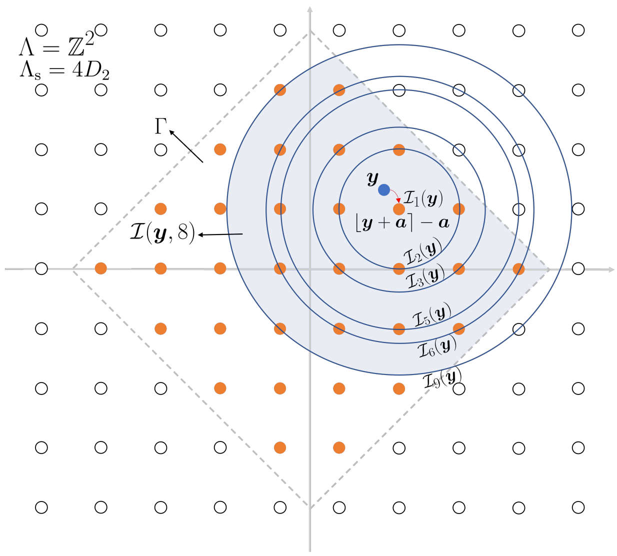


Fig. 2: The importance region for D_2^5 in Example 2. The constellation points inside the gray shaded region (including those falling on the boundary) form the importance region for GMI estimation. The constellation points falling on each "shell" form the subsets $\mathcal{I}_d(\boldsymbol{y})$.

$\mathcal{I}_d(\boldsymbol{y})$) $\boldsymbol{x}_{d,j}$ that have a binary label b at position k, then (7) can be approximated as

$$f_{k,b}(\boldsymbol{y}) \approx \frac{2}{M} \sum_{\boldsymbol{x}\in\mathcal{I}^{k,b}(\boldsymbol{y})} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) \tag{9}$$

$$\approx \frac{2}{M} \sum_{d=1}^{D} \frac{|\mathcal{I}_d(\boldsymbol{y})|}{K_d} \sum_{j\in\mathcal{T}_d^{k,b}(\boldsymbol{y})} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}_{d,j}). \tag{10}$$

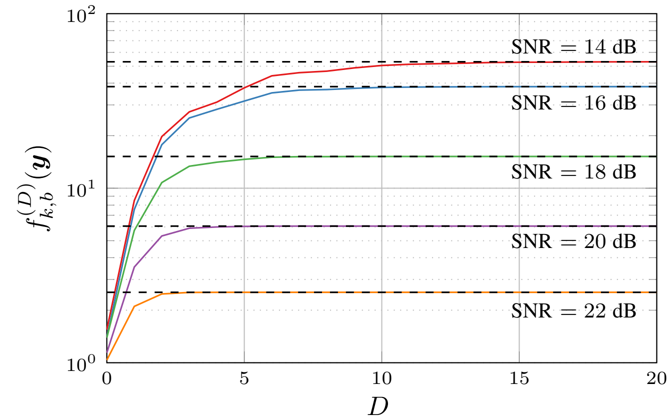
As D and K_d increase, (10) should converge to (7). So far, combining (8) and (10), the expressions for estimating the GMI of very large constellations are obtained. These expressions are applicable to any analytical channel (or an analytical auxiliary channel if the real channel is not analytically known), and any structured modulation formats as long as the importance set can be well defined. Generally, the important set depends on the modulation format $\mathcal{X}$ and the channel law $f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})$. If the constellation points in the important set cannot be enumerated, nor sampled, then it is infeasible to estimate the GMI by the proposed method.

B. GMI estimation for the VCs

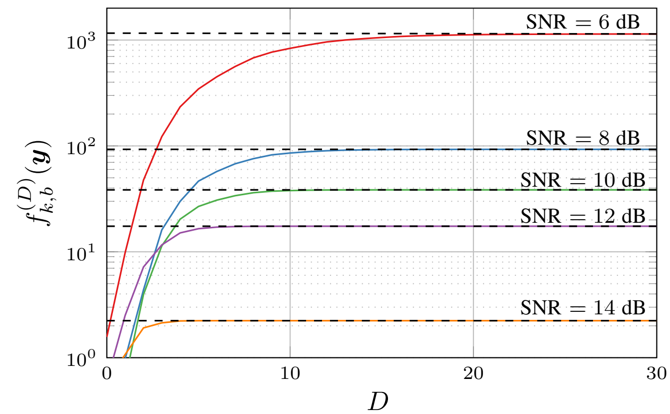
Since an analytical channel law of the nonlinear fiber channel is not known, an auxiliary channel law $q_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})$ is usually used to replace $f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})$ in (7), generating a lower bound on the GMI for the fiber channel. A common and reasonable choice of the auxiliary channel is the Gaussian channel, which has the conditional distribution

$$q_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{(2\pi\sigma^2/n)^{n/2}} \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{x}\|^2}{2\sigma^2/n}\right), \tag{11}$$

where σ^2 is the total noise power for n real dimensions. Then the SNR is defined as E_s/σ^2.



(a) Example 2: D_4^{17}



(b) Example 3: E_8^{24}

Fig. 3: The estimated value $f_{k,b}^{(D)}(\boldsymbol{y})$ as a function of D in Examples 2 and 3. The black dashed lines are the corresponding benchmark values $f_{k,b}(\boldsymbol{y})$ calculated using (7). Accurate estimation is validated by the convergence of all curves to the benchmark values.

Upon receiving a noisy symbol $\boldsymbol{y}$, according to (11), the constellation points that are close to $\boldsymbol{y}$ have the most impact on the GMI estimation. The nice structure of the cubic coding lattice makes these constellation points easy to enumerate in a "Euclidean ball" centered at $\lfloor \boldsymbol{y}+\boldsymbol{a} \rceil$ with radius R

$$\mathcal{B}(\boldsymbol{y}, R^2) \triangleq \{\boldsymbol{x} : \|\boldsymbol{x}+\boldsymbol{a} - \lfloor \boldsymbol{y}+\boldsymbol{a} \rceil\|^2 \leq R^2,\ \boldsymbol{x}+\boldsymbol{a} \in \mathbb{Z}^n\}, \tag{12}$$

where the squared radius $R^2 \in \mathbb{N}$. The ball can be further divided into R^2+1 "shells" defined as

$$\mathcal{S}(\boldsymbol{y}, r^2) \triangleq \{\boldsymbol{x} : \|\boldsymbol{x}+\boldsymbol{a} - \lfloor \boldsymbol{y}+\boldsymbol{a} \rceil\|^2 = r^2,\ \boldsymbol{x}+\boldsymbol{a} \in \mathbb{Z}^n\}, \tag{13}$$

for $r^2 = 0, 1, \ldots, R^2$.

The importance set in this specific application is not only a function of $\boldsymbol{y}$, but also depends on its squared radius R^2, and can be defined as

$$\mathcal{I}(\boldsymbol{y}, R^2) = \mathcal{B}(\boldsymbol{y}, R^2) \cap \Gamma, \tag{14}$$

which consists of $D = R^2+1$ disjoint subsets $\mathcal{I}_d(\boldsymbol{y}) = \mathcal{S}(\boldsymbol{y}, d-1) \cap \Gamma$ for $d = 1, \ldots, D$.

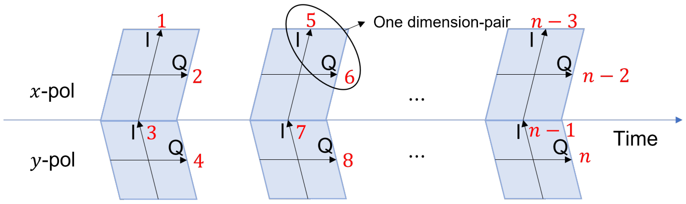


Fig. 4: The mapping of VCs to one wavelength. Integers from 1 to n represent the information carried by the n dimensions of VCs. All spectral efficiencies given in this paper are normalized to a single I/Q plane (2 dimensions).

In analogy with (8) and (10), the lower bound of the GMI can be approximated as

$$\widetilde{\text{GMI}} \approx \frac{1}{N_s}\sum_{k=1}^{m}\sum_{b\in\{0,1\}}\sum_{i\in\mathcal{Z}_k^b}\log_2\frac{\tilde{f}_{k,b}(\boldsymbol{y}^{(i)})}{\frac{1}{2}\left(\tilde{f}_{k,0}(\boldsymbol{y}^{(i)})+\tilde{f}_{k,1}(\boldsymbol{y}^{(i)})\right)} \tag{15}$$

where

$$\tilde{f}_{k,b}(\boldsymbol{y}) \approx \frac{2}{M}\sum_{d=1}^{D}\frac{|\mathcal{I}_d(\boldsymbol{y})|}{K_d}\sum_{j\in\mathcal{T}_d^{k,b}(\boldsymbol{y})} q_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}_{d,j}). \tag{16}$$

In our simulations, we found that the subsums in $\mathcal{I}_d^{(k,b)}(\boldsymbol{y})$ for all $d=1,\ldots,D$ in (16) can be approximated very well with $K_d=\min\{|\mathcal{I}_d(\boldsymbol{y})|,10^5\}$ samples. With this number of samples in each subset, we increase D from 1, until the condition

$$\max_{i=1,\ldots,N_s}\left(\frac{\tilde{f}_{k,b}^{(D+1)}(\boldsymbol{y}^{(i)})-\tilde{f}_{k,b}^{(D)}(\boldsymbol{y}^{(i)})}{\tilde{f}_{k,b}^{(D)}(\boldsymbol{y}^{(i)})}\right)<0.5\% \tag{17}$$

is met for all $k=1,\ldots,m$ and $b\in\{0,1\}$, where the superscripts of the estimated $\tilde{f}_{k,b}(\boldsymbol{y}^{(i)})$ denote the number of subsets being used. Then we believe that with no less than D subsets, $\tilde{f}_{k,b}(\boldsymbol{y}^{(i)})$ converges to the exact value (7). Finally, accurate approximation of $\tilde{g}_{k,b}(\boldsymbol{y}^{(i)})$ for all $i=1,\ldots,N_s$ yields accurate estimation of $\widetilde{\text{GMI}}$.

Example 2: We again take D_2^5 as an example due to the simplicity for illustration. Fig. 2 shows the importance region $\mathcal{I}(\boldsymbol{y},8)$ and its subsets when $D=9$. Given a noisy $\boldsymbol{y}$, the importance region is centered at $\lfloor\boldsymbol{y}+\boldsymbol{a}\rceil-\boldsymbol{a}$, which itself forms the first subset $\mathcal{I}_1(\boldsymbol{y})$. In this example, $\mathcal{I}_4(\boldsymbol{y})=\mathcal{I}_7(\boldsymbol{y})=\mathcal{I}_8(\boldsymbol{y})=\varnothing$.

Example 3: We transmit D_4^{17} over the Gaussian channel. For random k and $\boldsymbol{y}$, Fig. 3a shows the convergence of the estimated $f_{k,b}^{(D)}(\boldsymbol{y})$ using (10) to the benchmark values $f_{k,b}(\boldsymbol{y})$ computed using (7) for different SNRs. In this example, $K_d=|\mathcal{I}_d(\boldsymbol{y})|$ for all $d=1,\ldots,D$.

Example 4: We transmit E_8^{24} over the Gaussian channel. For random k and $\boldsymbol{y}$, Fig. 3b shows the convergence of the estimated $f_{k,b}^{(D)}(\boldsymbol{y})$ using (10) to the benchmark values $f_{k,b}(\boldsymbol{y})$ computed using (7) in the medium SNR range. In this example, $K_d=10^5$ uniform samples from $\mathcal{I}_d(\boldsymbol{y})$ are used for subsets with $d>8$ and $K_d=|\mathcal{I}_d(\boldsymbol{y})|$ for subsets with $d\le 8$.

TABLE II: The simulation parameters.

Parameters	Values
Symbol rate	45 Gbaud
Pulse shape	RRC, rolloff 0.1
Fiber attenuation	0.2 dB/km
Fiber nonlinear coefficient	1.27 $\text{W}^{-1}\text{km}^{-1}$
Dispersion parameter	17 pm/nm/km
EDFA noise factor	5 dB
Span length	80 or 100 km
Center frequency	1550 nm
Channel spacing	50 GHz
Oversampling factor in SSFM	32
Oversampling factor in DBP (if any)	32
Step size in SSFM	1 km
Pilot overhead	2%
Number of symbol time slots transmitting VC symbols	8×10^4

IV. Performance Analysis

In this section, we study the BER, MI, and GMI performance of 16-dimensional and 32-dimensional VCs in Table II over the nonlinear fiber channel, since their ASGs are high, which implies large potential shaping gains in the nonlinear fiber channel. The benchmarks for comparison are the QAM formats with the same spectral efficiencies. For transmission of an n-dimensional VC, we consider (a) a single-channel system in which $n/4$ time slots are needed to transmit one VC symbol, and (b) a WDM system with $n/4$ wavelengths in which each wavelength carries one VC symbol in $n/4$ time slots. The mapping of VCs to one wavelength is depicted in Fig. 4. For the WDM case, many other different ways of assigning VCs to physical dimensions are possible. Nevertheless, our simulations show no big performance difference using these different ways. The experimental results of comparing different physical realizations for another type of VC draw the same conclusion [29]. Therefore, we choose the mapping in Fig. 4 because this scheme does not need many coherent receivers to decode one VC symbol.

VCs can have non-integer spectral efficiencies from the definition of (3), as β is not guaranteed to be an integer. The performance of VCs should be compared with QAM constellations at the same spectral efficiencies. Thus, for QAM transmission, in order to have a non-integer spectral efficiency, two two-dimensional QAM formats with different cardinalities M_1 and M_2 can be assigned to different I/Q pairs, resulting in a spectral efficiency of

$$\beta_{\text{QAM}}=\frac{d_1\log_2 M_1+d_2\log_2 M_2}{d_1+d_2}\ \text{[bits/symbol/dimension-pair]},$$

where d_1 and d_2 are the number of I/Q pairs in Fig. 4 transmitting M_1-QAM and M_2-QAM, respectively, and $d_1+d_2=n/2$. The QAM constellation having the same spectral efficiency as Λ_{16}^{76} has parameters $d_1=4$, $d_2=4$, $M_1=1024$, and $M_2=512$; for the QAM corresponding to L_{32}^{155}, we should set $d_1=11$, $d_2=5$, $M_1=1024$, and $M_2=512$. The two constituent QAM constellations are scaled to the same minimum distance, which maximizes the minimum distance of the resulting hybrid-QAM constellation for a given n-

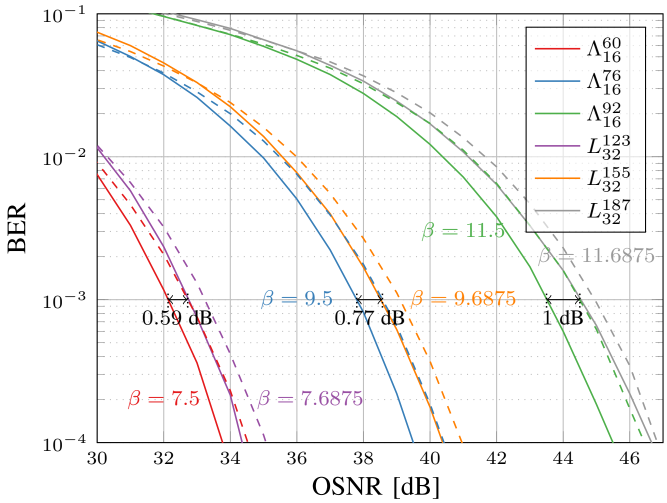
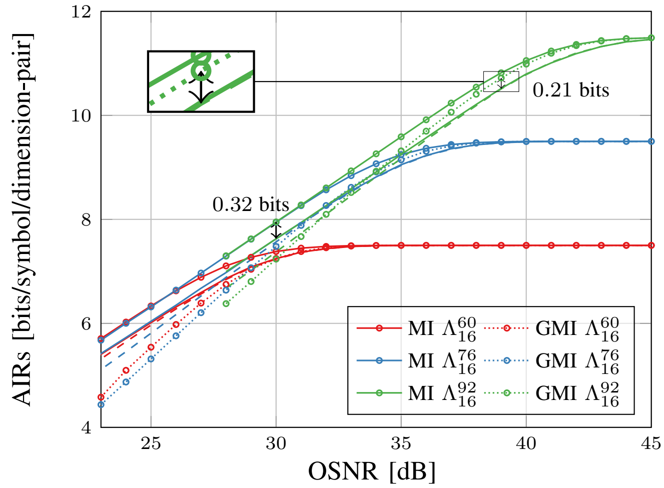
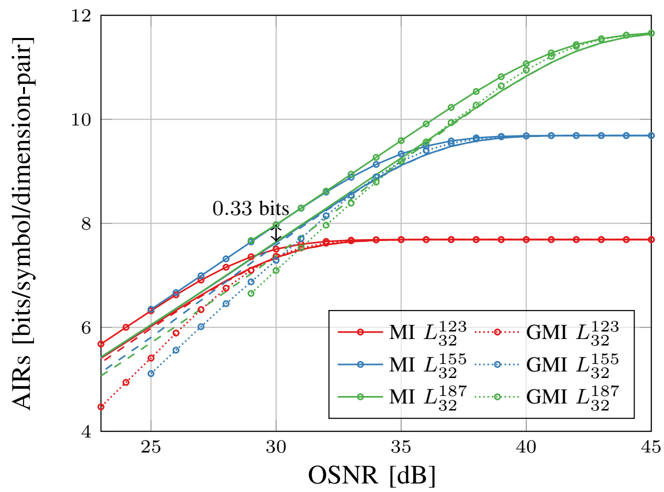


Fig. 5: The BER as a function of the OSNR for VCs compared with QAM constellations in the B2B scenario. Dashed lines represent QAM constellations. Lines with the same colors represent the same spectral efficiencies.



(a) 16-dimensional constellations



(b) 32-dimensional constellations

Fig. 6: The maximum AIRs as a function of the OSNR for (a) 16-dimensional and (b) 32-dimensional VCs in the B2B scenario. Solid and dashed lines without markers represent the MI and GMI of QAM, respectively. Lines with the same colors represent the same spectral efficiencies β (see Table I).

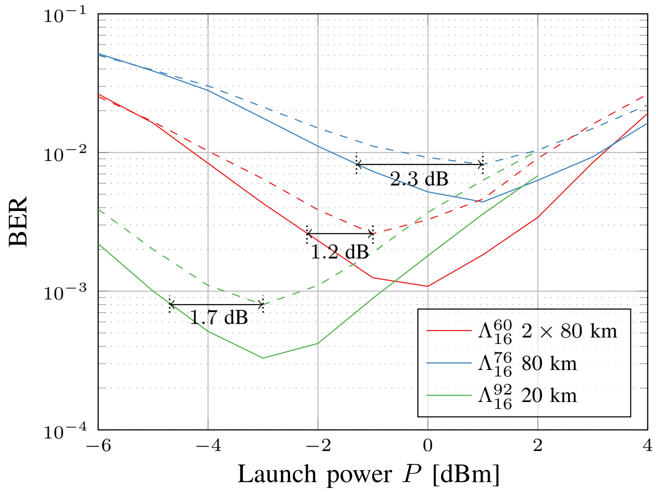
dimensional symbol energy E_s [30, Ch. 4.3]. For modeling the fiber channel, the Manakov equation [31] is adopted and the split-step Fourier method (SSFM) [32] is used to simulate the channel, which is sufficiently accurate to capture the nonlinearities in a real fiber. In the digital signal processing (DSP) chain at the receiver side, a one-time chromatic dispersion (CD) compensation is performed. Alternatively, a full-field digital backpropagation (DBP) can be implemented to increase the transmission distance of VCs with high spectral efficiencies. Also, we use a simple pilot-aided DSP algorithm, in which a certain percentage of 4-QAM pilot symbols are added in front of the transmitted symbols in each polarization to further compensate for the phase rotation in each channel due to fiber nonlinearities. The QAM formats are Gray labeled and detected with the maximum likelihood decoder. The simulation parameters are listed in Table II.

In experiments, other impairments such as phase noise, frequency offset, transceiver impairments, etc., may need to be compensated using a more sophisticated DSP chain. In general, standard DSP algorithms designed for QAM work for VCs as well without modification. In [20], a pilot-based DSP algorithm described in [33] has been used for transmission of VCs in experiments.

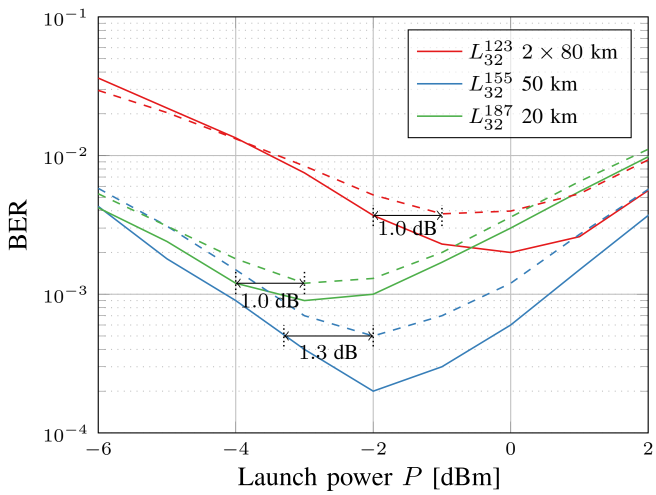
A. *Back-to-back case*

A back-to-back (B2B) scenario reflects the performance of VCs in the absence of fiber nonlinearities. The BER performance of the considered VCs depends on the spectral efficiencies, number of dimensions, and the labeling scheme. Fig. 5 presents the BER performance of the considered VCs compared with QAM in the B2B scenario. The optical signal-to-noise ratio (OSNR) is calculated assuming that it is measured in a reference optical bandwidth of 12.5 GHz (0.1-nm wavelength). VCs with higher spectral efficiencies have larger OSNR gains. Up to 1 dB power gains are observed at a BER of 10^{-3} for Λ_{16}^{76}. The 32-dimensional VCs do not show higher gains than 16-dimensional VCs, since the bit labeling scheme is more efficient in 16-dimensional VCs. To compare the labeling performance for these very large VCs, one could calculate the Gray penalty [34], [35] using Algorithm 5 in [17]. For example, Λ_{16}^{76} has a Gray penalty of approximately 1.33, which is smaller than that of L_{32}^{155} (1.40).

Fig. 6 shows the estimated MI and GMI as a function of OSNR in the B2B scenario. The MI is estimated using the method in [19, Sec. V-B] and parameters are chosen as suggested therein. The GMI is estimated using the method and parameters proposed in section III-B. The results show that VCs outperform QAMs throughout the whole OSNR range in terms of MI. Up to 0.33 bits/symbol/dimension-pair MI gains over QAMs can be observed. However, as for the GMI, VCs only outperform QAM in a limited OSNR range, and VCs with higher spectral efficiencies tend to have larger GMI gains. A 0.21 bits/symbol/dimension-pair GMI gain is observed for Λ_{16}^{92}

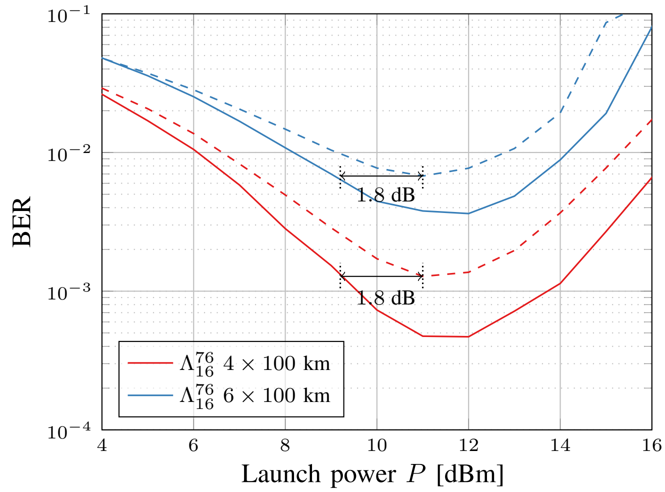


(a) 16-dimensional constellations

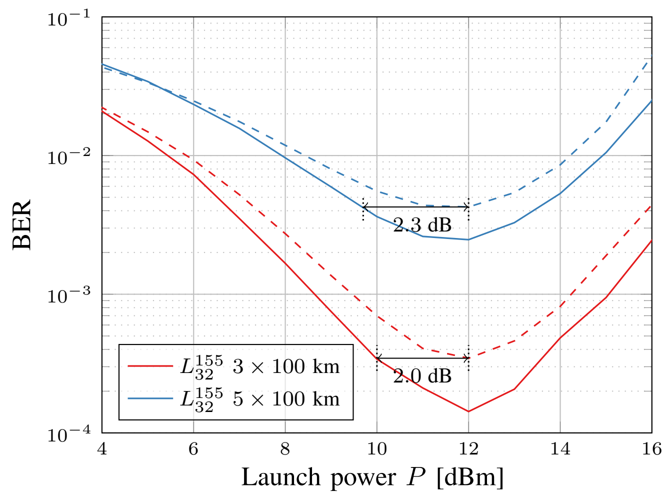


(b) 32-dimensional constellations

Fig. 7: The BER as a function of the launch power for VCs in the single-channel case with electronic CD compensation. Dashed lines with the same colors represent QAM having the same spectral efficiencies β as VCs.



(a) 16-dimensional constellations



(b) 32-dimensional constellations

Fig. 8: The BER as a function of the launch power for VCs in the single-channel case with a small step size DBP. Dashed lines with the same colors represent QAM having the same spectral efficiencies β as VCs.

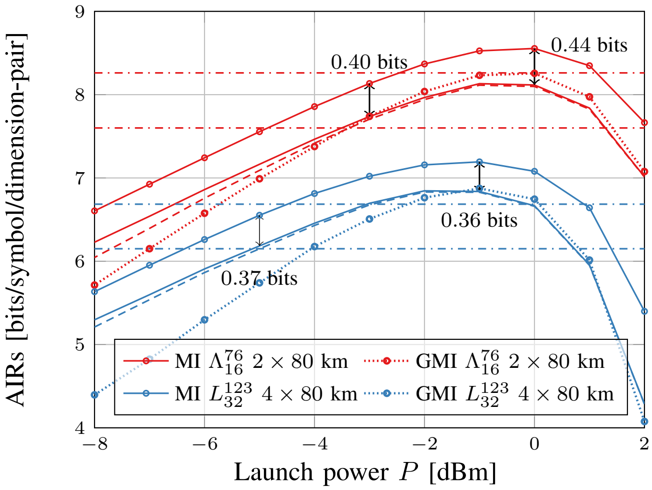
when the GMI is close to β. Compared with 16-dimensional VCs, the 32-dimensional VCs have almost no GMI gains due to the worse labeling performance.

B. Single-channel case

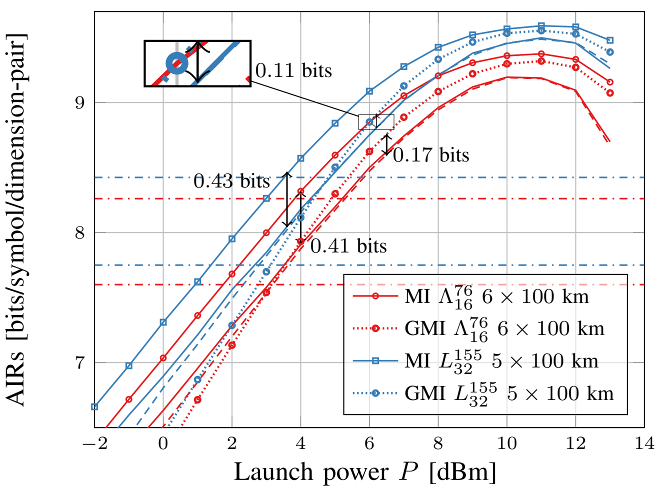
Fig. 7 shows the BER performance of the considered VCs over a nonlinear fiber channel. Only CD compensation is performed at the receiver. The BER is presented between 10^{-4} and 10^{-2}, which covers most commonly-used hard-decision forward error correction limits for fiber communications [36]. The simulated transmission distances are short, since the spectral efficiencies are high. It shows that the considered VCs reduce the minimum BER over QAM by 25%–58%, and show launch power gains ranging from 1.0–2.3 dB over QAM at the minimum BER achieved by QAM, larger than the OSNR improvements realized in the B2B case, which are less than 1 dB.

The considered VCs are shown to be suitable for short-distance transmission. Despite this, we also study their performance when a full-field DBP with a rather small step size (1 km) is used to support the high SNR needed by these VCs for longer transmission distances. As an example, Λ_{16}^{76} and L_{32}^{155} still maintain good launch power gains as shown in Fig. 8 even when most of the nonlinearities are compensated in the DBP. This extreme case shows the advantage of transmitting VCs over long distances if high complexity is allowed in the DSP.

Fig. 9 shows the maximum AIR performance for the VCs with and without DBP. With only CD compensation, 0.36–0.44 bits/symbol/dimension-pair MI gains of VCs over QAM can be observed throughout the whole launch power range. However, small GMI gains are only observed for Λ_{16}^{76} in a limited launch power range. With a small step size DBP, VCs achieve similar MI gains as without DBP. The GMI gains tend to be somewhat higher when the SNR is higher. If we assume an FEC code with 20% overhead, which reduces the



(a) CD used in DSP

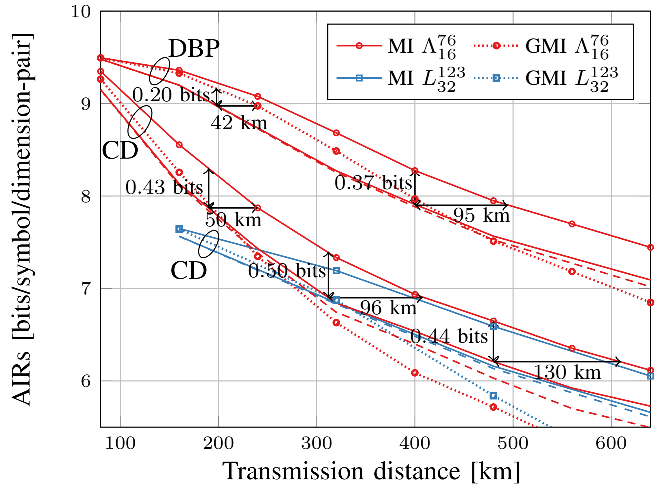


(b) Small step size DBP used in DSP

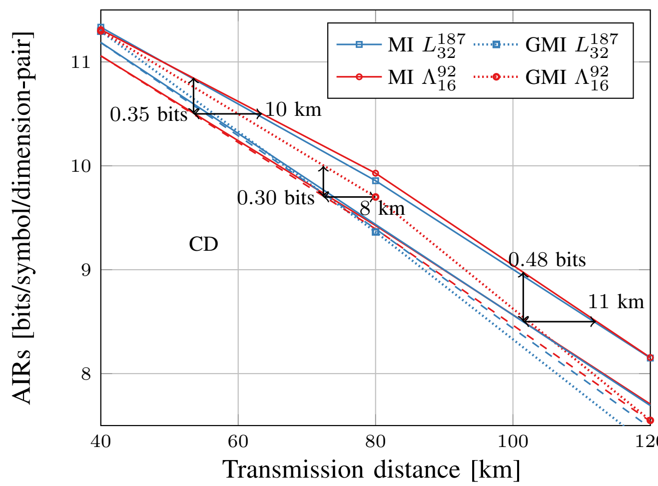
Fig. 9: The maximum AIRs as a function of the launch power of VCs in the single-channel case. In (a), a one-time electronic CD compensation is performed. In (b), a full-field 1-km step size DBP is used. Solid and dashed lines without markers represent the MI and GMI of QAM, respectively. Lines with the same colors have the same spectral efficiencies β. For each VC, two horizontal dash-dotted lines are drawn at 80% and 87% of β, corresponding to an FEC overhead of 25% and 15%, respectively.

spectral efficiency to 83% of the uncoded spectral efficiency β, 0.17 and 0.11 bits/symbol/dimension-pair GMI gains are observed for Λ_{16}^{76} and L_{32}^{155}, respectively. VCs only outperform QAM when the GMI is above 80% and 85% of β for Λ_{16}^{76} and L_{32}^{155}, respectively. This means that there exists a BICM scheme with an FEC overhead smaller than around 25% and 18% for Λ_{16}^{76} and L_{32}^{155}, respectively, in which using VCs as modulation formats achieves better performance than QAM.

Fig. 10 shows the maximum AIRs as a function of the transmission distance, in which the launch power is optimized for each distance. For medium distance range, VCs can achieve up to 0.50 bits/symbol/dimension-pair MI gains over QAM at the same transmission distance. Without any compensation of nonlinearities, at around 310 km, Λ_{16}^{76} and L_{32}^{123} increase the reach by 96 km achieving the same MI. No significant



(a) Medium distance range



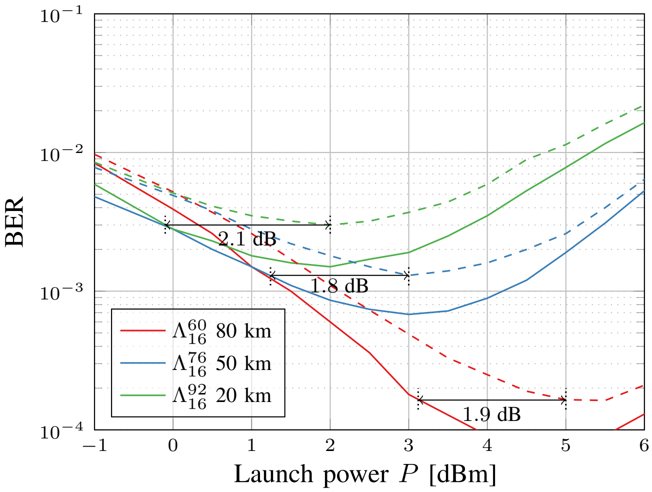
(b) Short distance range

Fig. 10: The maximum AIRs as a function of the transmission distance for the VCs in the single-channel case, with a span length of 80 km. Solid and dashed lines without markers represent the MI and GMI of QAM, respectively. Lines with the same colors have the same spectral efficiencies β.

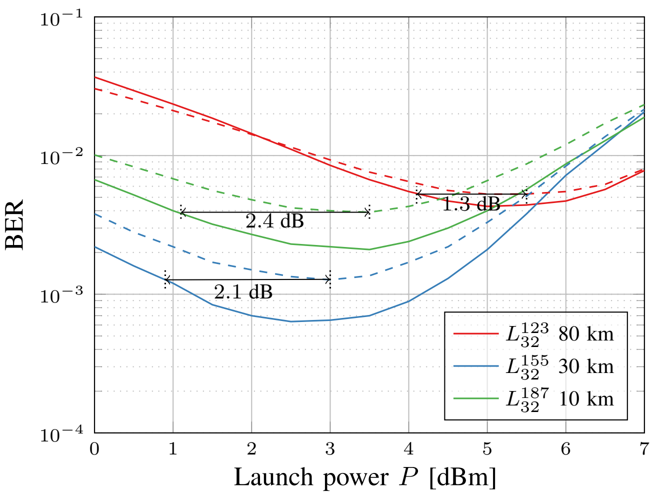
gains are shown in terms of GMI. With DBP at the receiver (10-km step size), Λ_{16}^{76} realizes larger MI and GMI gains, or reach increase at the same MI and GMI, compared with the case without DBP. For example, at the same GMI value, Λ_{16}^{76} increases the transmission distance of QAM from 198 to 240 km. For short transmission distances (40–120 km) without DBP, VCs maintain large MI gains and reach increase at the same MI, and realize up to 0.30 bits/symbol/dimension-pair GMI gains. The maximum observed MI and GMI gains over QAM are up to 0.50 and 0.30 bits/symbol/dimension-pair, respectively, which are larger than the maximum observed MI and GMI gains in the B2B case (0.33 and 0.21 bits/symbol/dimension-pair), respectively.

C. WDM case

The BER performance of VCs compared with QAM in WDM systems without and with a small step size DBP are

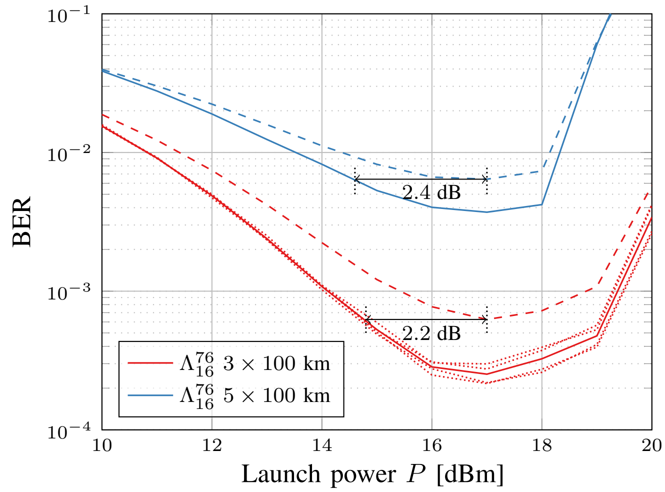


(a) 16-dimensional constellations

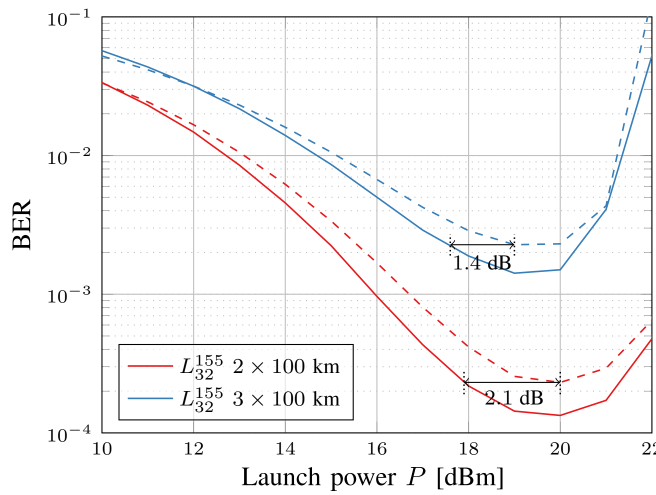


(b) 32-dimensional constellations

Fig. 11: The BER as a function of the launch power for VCs in single-span WDM transmission with CD compensation in DSP. Dashed lines with the same colors represent QAM having the same spectral efficiencies β as VCs.



(a) 16-dimensional constellations



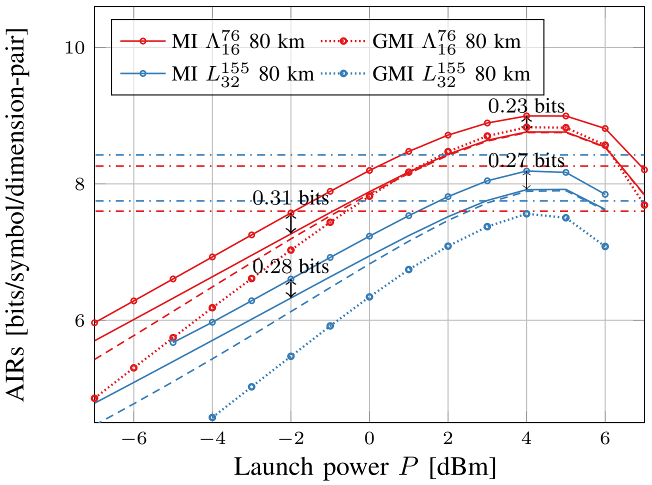
(b) 32-dimensional constellations

Fig. 12: The BER as a function of the launch power for VCs in the WDM case with a small step size (1 km) DBP. Dashed lines with the same colors represent QAM having the same spectral efficiencies β as VCs. In (a), the two dotted red curves above the solid Λ_{16}^{76} curve represent the BERs of the two center channels, and the two below are from the two side channels.

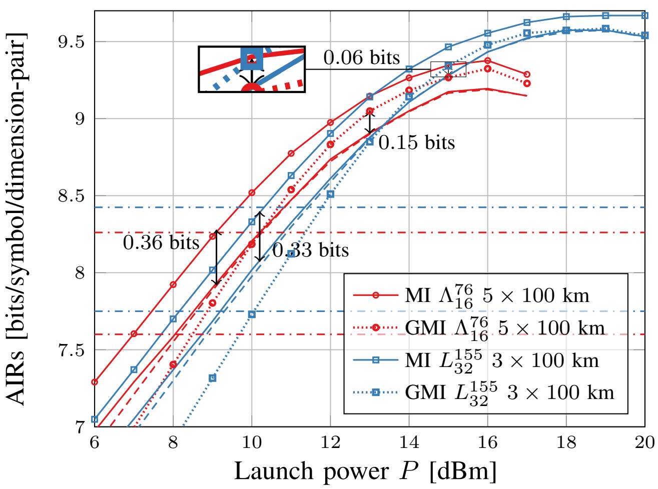
shown in Fig. 11 and Fig. 12, respectively. The shown BER is the average BER over all wavelengths, not from one of the two center channels as usual, which however is shown to not affect the performance analysis significantly, since there is no big difference among the BER of each single channel; see the dotted curves in Fig. 12a. Both with and without DBP, the considered VCs reduce the minimum BER of QAM by 18%–50%, and show 1.3–2.4 dB launch power gains over QAM at the minimum BER achieved by QAM.

Fig. 13 reports some MI and GMI results as a function of the launch power. Around 0.23–0.36 bits/symbol/dimension-pair MI gains over QAM are achieved by the VCs for the whole studied launch power range, whereas the GMI gains are subject to the SNR. An extreme example is L_{32}^{155} transmitted at 80 km without DBP, whose GMI performance is much worse than for QAM. On the other hand, for Λ_{16}^{76} transmitted at 80 km with DBP, up to 0.15 bits/symbol/dimension-pair GMI gains are observed in a limited launch power range.

Fig. 14 illustrates the maximum AIRs as a function of the transmission distance. For the whole studied medium distance range, without DBP, VCs achieve up to 0.29 bits/symbol/dimension-pair MI gains and allow for larger potential reach than QAM at the same MI, whereas the GMI gains are negligible. With a 10-km step size DBP at the receiver, the MI gains and reach increase are larger than without DBP. For short distances, VCs achieve higher maximum AIR gains than for medium distances, and increase the transmission distance by 10%–20% at the same MI. Overall, the AIR gains in WDM are marginally smaller than in the corresponding single-channel case. This might be due to that the fiber nonlinearities, especially the cross-phase modulation, might have a slightly stronger impact on VCs than QAM formats. However, the maximum observed MI gain of VCs over QAM is still found higher than in the B2B case

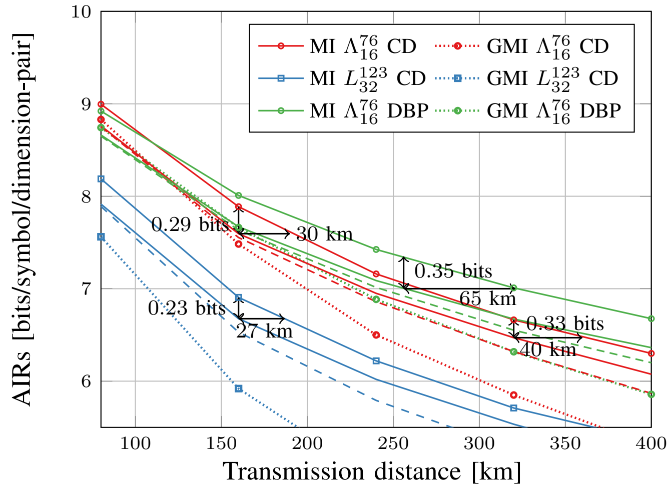


(a) CD used in DSP

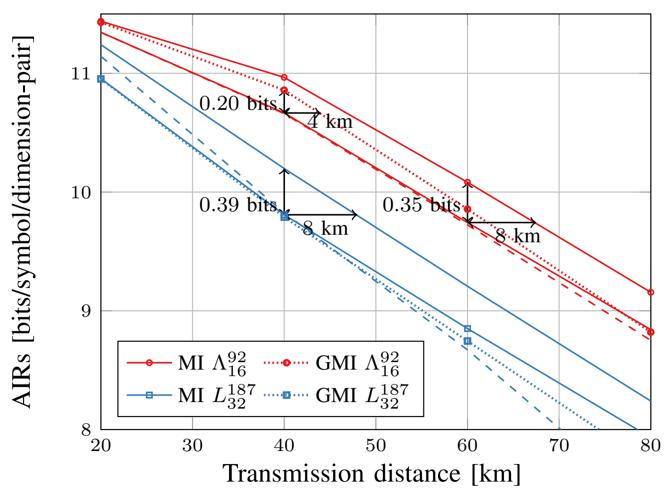


(b) Small step size DBP used in DSP

Fig. 13: The maximum AIRs as a function of the launch power of VCs in the WDM case (a) without and (b) with a 1-km step size DBP. Solid and dashed lines without markers represent the MIs and GMIs of QAM, respectively. Lines with the same colors represent the same spectral efficiencies β. For each VC, two horizontal dash-dotted lines are drawn at 80% and 87% of β, corresponding to an FEC overhead of 25% and 15%, respectively.



(a) Medium distance range



(b) Short distance range

Fig. 14: The maximum AIRs as a function of the transmission distance for VCs in the WDM case, with a span length of 80 km. Solid and dashed lines without markers represent the MI and GMI of QAM, respectively. Lines with the same colors represent the same spectral efficiencies β.

($0.39 > 0.33$ bits/symbol/dimension-pair), which is consistent with the observation in the single-channel case.

V. Conclusion

We have simulated transmission of 16- and 32-dimensional Voronoi constellations with a cubic coding lattice over the nonlinear fiber channel and studied their performance in both single-channel and WDM systems. The BER, MI, and GMI performance of VCs are compared with QAM constellations at the same spectral efficiencies. Extended from the MI estimation method from our previous work [19], a GMI estimation method for very large constellations is proposed and applied to the considered VCs. The MI and GMI performance of VCs in fiber-optic communications are first demonstrated to our knowledge. The launch power gains of VCs over QAM imply that they can achieve better performance than QAM in systems with hard-decision FEC decoding. However, in systems with BICM and soft-decision decoding, where the GMI is the right predictor for the post-FEC BER, the performance gains of VCs might be limited. However, their good MI gains over QAM imply that designing a multilevel code specifically for the considered VCs might achieve these potential rate gains and reach increases for fiber communications, which remains as future work.

References

[1] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "On shaping gain in the nonlinear fiber-optic channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2014, pp. 2794–2798.

[2] T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, "On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel," *J. Lightw. Technol.*, vol. 34, no. 21, pp. 5063–5073, 2016.

[3] M. N. Tehrani, M. Torbatian, H. Sun, P. Mertz, and K.-T. Wu, "A novel nonlinearity tolerant super-Gaussian distribution for probabilistically shaped modulation," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Rome, Italy, 2018.

[4] 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," *Journal of Lightwave Technology*, pp. 1–14, 2022.

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

[6] B. Chen, C. Okonkwo, H. Hafermann, and A. Alvarado, "Polarization-ring-switching for nonlinearity-tolerant geometrically shaped four-dimensional formats maximizing generalized mutual information," *J. Lightw. Technol.*, vol. 37, no. 14, pp. 3579–3591, 2019.

[7] R.-J. Essiambre, R. Ryf, M. Kodialam, B. Chen, M. Mazur, J. I. Bonetti, R. Veronese, H. Huang, A. Gupta, F. A. Aoudia, E. C. Burrows, D. F. Grosz, L. Palmieri, M. Sellathurai, X. Chen, N. K. Fontaine, and H. Chen, "Increased reach of long-haul transmission using a constant-power 4D format designed using neural networks," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Brussels, Belgium, Dec. 2020.

[8] 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.

[9] J. H. Conway and N. J. A. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 6, pp. 820–824, 1983.

[10] G. D. Forney, Jr., "Multidimensional constellations—part II: Voronoi constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 941–958, 1989.

[11] N. Sommer, M. Feder, and O. Shalvi, "Shaping methods for low-density lattice codes," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Taormina, Italy, Oct. 2009.

[12] B. M. Kurkoski, J. Dauwels, and H.-A. Loeliger, "Power-constrained communications using LDLC lattices," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, July 2009.

[13] N. S. Ferdinand, B. M. Kurkoski, B. Aazhang, and M. Latva-aho, "Shaping low-density lattice codes using Voronoi integers," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Hobart, Australia, Nov. 2014.

[14] N. S. Ferdinand, B. M. Kurkoski, M. Nokleby, and B. Aazhang, "Low-dimensional shaping for high-dimensional lattice codes," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7405–7418, 2016.

[15] N. di Pietro and J. J. Boutros, "Leech constellations of construction-A lattices," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4622–4631, 2017.

[16] C. Feng, D. Silva, and F. R. Kschischang, "An algebraic approach to physical-layer network coding," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7576–7596, 2013.

[17] S. Li, A. Mirani, M. Karlsson, and E. Agrell, "Designing Voronoi constellations to minimize bit error rate," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Melbourne, Australia, July 2021.

[18] A. Mirani, E. Agrell, and M. Karlsson, "Low-complexity geometric shaping," *J. Lightw. Technol.*, vol. 39, no. 2, pp. 363–371, 2021.

[19] S. Li, A. Mirani, M. Karlsson, and E. Agrell, "Low-complexity Voronoi shaping for the Gaussian channel," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 865–873, 2022.

[20] A. Mirani, K. Vijayan, Z. He, S. Li, E. Agrell, J. Schröder, P. Andrekson, and M. Karlsson, "Experimental demonstration of 8-dimensional Voronoi constellations with 65,536 and 16,777,216 symbols," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Bordeaux, France, Sep. 2021.

[21] Optical Internetworking Forum, "OIF-400ZR-01.0 – implementation agreement 400ZR," https://www.oiforum.com/wp-content/uploads/OIF-400ZR-01.0_reduced2.pdf, March 2020.

[22] G. D. Forney, Jr. and L.-F. Wei, "Multidimensional constellations—part I: Introduction, figures of merit, and generalized cross constellations," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 6, pp. 877–892, 1989.

[23] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 913–929, 1993.

[24] M. Karlsson and E. Agrell, "Spectrally efficient four-dimensional modulation," in *Proc. Opt. Fiber Commun. Conf. (OFC)*, Los Angeles, CA, 2012.

[25] B. M. Kurkoski, "Encoding and indexing of lattice codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6320–6332, 2018.

[26] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. New York, NY: Springer, 1999.

[27] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems, "Achievable information rates for fiber optics: Applications and computations," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 424–439, 2018.

[28] A. B. Owen, "Monte Carlo theory, methods and examples," 2013. [Online]. Available: https://statweb.stanford.edu/~owen/mc

[29] A. Mirani, K. Vijayan, S. Li, Z. He, J. Schröder, P. Andrekson, E. Agrell, and M. Karlsson, "Comparison of physical realizations of multidimensional Voronoi constellations in single mode fibers," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Basel, Switzerland, Sept. 2022.

[30] G. L. Stüber, *Principles of Mobile Communication*, 4th ed. Springer, 2017.

[31] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringent fibers," *Opt. Lett.*, vol. 16, no. 16, pp. 1231–1233, Aug 1991.

[32] R. H. Hardin and F. Tappert, "Application of the split-step Fourier method to the numerical solution of nonlinear and variable coefficient wave equations," *SIAM Review Chronicle*, vol. 15, no. 2, p. 423, Apr. 1973.

[33] J. Schröder and M. Mazur, "QAMPy a DSP chain for optical communications," DOI: 10.5281/zenodo.1195720, 2018.

[34] M. K. Simon and J. G. Smith, "Hexagonal multiple phase-and-amplitude-shift-keyed signal sets," *IEEE Trans. Commun.*, vol. 21, no. 10, pp. 1108–1115, 1973.

[35] J. G. Smith, "Odd-bit quadrature amplitude-shift keying," *IEEE Trans. Commun.*, vol. 23, no. 3, pp. 385–389, 1975.

[36] E. Agrell and M. Secondini, "Information-theoretic tools for optical communications engineers," in *IEEE Photon. Conf. (IPC)*, Reston, VA, USA, Sept. 2018.