Low-complexity geometric shaping

N.B. When citing this work, cite the original published paper.

(article starts on next page)

Low-Complexity Geometric Shaping

Ali Mirani [iD], Erik Agrell [iD], *Fellow, IEEE,* and Magnus Karlsson [iD], *Senior Member, IEEE, Fellow, OSA*

Abstract—**Approaching Shannon's capacity via geometric shaping has usually been regarded as challenging due to modulation and demodulation complexity, requiring look-up tables to store the constellation points and constellation bit labeling. To overcome these challenges, in this paper, we study lattice-based geometrically shaped modulation formats in multidimensional Euclidean space. We describe and evaluate fast and low complexity modulation and demodulation algorithms that make these modulation formats practical, even with extremely high constellation sizes with more than 10^{28} points. The uncoded bit error rate performance of these constellations is compared with the conventional quadrature amplitude modulate (QAM) formats in the additive white Gaussian noise and nonlinear fiber channels. At a spectral efficiency of 2 bits/sym/polarization, compared with 4-QAM format, transmission reach improvement of more than 38% is shown at the hard-decision forward error correction threshold of 2.26×10^{-4}.**

Index Terms—**Optical communication, coherent receiver, multidimensional modulation format, geometric shaping, lattice.**

I. INTRODUCTION

MUCH effort have been put on different aspects of optical links to overcome the need for higher information throughput. Among them, the optimization of modulation formats has been an important research area [1], [2]. One way to increase the transmission rate is to use higher-order modulation formats, i.e., transmitting more information bits in each channel use. However, higher-order modulation formats require more signal power to achieve a certain performance which can be a problem because fiber links are limited in transmitting power by fiber nonlinearities [3, chap. 8]. Another method to improve the transmission rate in optical communication is to use constellation shaping. Shaping can be performed in two different ways, which can also be combined with each other, known as *probabilistic shaping* [4] and *geometric shaping* [5]. Probabilistic shaping uses a nonuniform distribution over a regular, low-dimensional constellation such as QAM, whereas geometric shaping refers to a nonrectangular constellation, often multidimensional, with a uniform distribution. Recently, probabilistic shaping has been extensively studied and shown to provide rate adaptability and energy efficiency gains [6], [7]. However, its need for a distribution matcher increases the system complexity. Also geometric shaping has been studied, although less, but used to design nonlinear tolerant and multidimensional constellations [8], [9], [10].

A. Mirani and M. Karlsson are with the Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-41296 Gothenburg, Sweden (email: mirani@chalmers.se; magnus.karlsson@chalmers.se).

E. Agrell is with the Department of Electrical Engineering, Chalmers University of Technology, SE-41296 Gothenburg, Sweden (email: agrell@chalmers.se).

Exploiting multidimensional modulation formats brings about two independent performance gains known as *coding gain* and *shaping gain*. Coding gain comes from the fact that constellation points can be packed more densely in multidimensional space, and the shaping gain is achieved when the constellation boundary is closer to a hypersphere. The maximum shaping gain that can be obtained by a spherical boundary in high dimensional space is 1.53 dB compared to a cubic boundary [11].

The electromagnetic field has 4 degrees of freedom [12] and combined with coherent detection schemes, receivers with improved sensitivity can be utilized to compensate for channel distortions, i.e., polarization drift and nonlinear effects [13]. Furthermore, taking advantage of wavelengths, time slots, and spatial modes/cores in a fiber link can increase the available degrees of freedom and increase data transmission rates. Typically, these dimensions have been modulated independently in fiber optic communications, but there are studies that show that optimizing formats in many dimensions will improve the performance of the system.

Modern research on multidimensional modulation formats in optical communication systems dates back to 2009 when the possibility of optimizing the constellations in the 4-dimensional space formed by the optical field was demonstrated [14]. High-dimensional modulation formats were then optimized by sphere packing to find the most power-efficient and spectral-efficient modulation formats [15]. Later on, there were even more research on employing higher dimensions such as 8 and 24. In [16], 8-dimensional lattice modulation formats were simulated based on spherical cut [17] with 128 and 256 constellation points. Also, by combining pulse-position modulation with polarization-switched quadrature phase-shift keying (QPSK), an 8-dimensional modulation format was experimentally investigated in [18] showing significant improvement over polarization multiplexed QPSK. In [19] and [20], 8- and 24-dimensional modulation formats using temporally adjacent time slots were used to transmit data over short-reach intensity modulated links. In [21], multidimensional modulation formats in coherent optical systems were demonstrated by spherical-cut constellations and block-coded modulation. The advantages of combining multidimensional shaping with non-binary coded modulation were shown in [22]. An optimized 8-dimensional modulation format based on polarizations and time slots was compared with polarization multiplexed binary phase-shift keying (BPSK) to show its improved nonlinear performance in [23]. In [24], multidimensional constellations based on lattices, so called *Voronoi constellation (VC)*, were investigated and were suggested as coded modulation schemes with higher shaping gains than generalized cross constellations indicated in [11]. Also, the densest lattices for the sphere packing problem in different dimensions with fast and low-

complexity modulation and demodulation algorithms were studied in [25] and [26]. In [27], to achieve high spectral efficiency (SE), digital backpropagation was combined with trellis shaping where bit-interleaved coded modulation and a convolutional shaping code were used for lower and higher bit levels, respectively. The complexity of this method is mainly due to the Viterbi algorithm decoder in the transmitter [28].

In this paper, we evaluate the performance of coherent fiber links with ultra-large Voronoi constellations with more than 10^{28} points, which has never been done before to our knowledge. The low-complexity algorithms are numerically compared with maximum likelihood (ML) detection and shown to perform almost equivalently at high SE or high signal to noise ratio (SNR). Low-complexity natural binary and quasi-Gray constellation bit labelings are used to map bits to constellation points and their performances are compared. The asymptotic power efficiency of lattices for different dimensions and SEs, as well as the optimum choice of Voronoi boundary are studied. The uncoded bit error rate (BER) performance of these VCs are then examined in the additive white Gaussian noise (AWGN) and nonlinear fiber-optic channels.

The remainder of the paper is organized as follows. In section II, we describe the lattice theory, lattice properties, and how to use them for finite constellations. Fast modulation and demodulation algorithms for lattice-based VCs are then presented in section III. In section IV, figures of merit for comparing modulation formats at different dimensions are defined and numerical simulations are applied to demonstrate these concepts. We finally give the performance of the VCs in section V in both the AWGN and the nonlinear fiber-optic channels. The conclusion of this paper is presented in section VI.

II. PRELIMINARIES

In this section, we introduce the basics of lattice theory, lattice properties, constellations and their construction based on lattices.

In general, lattices are periodic structures of points in N-dimensional Euclidean space. Mathematically, lattices can be described and constructed using a set of linearly independent vectors $\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_l \in \mathbb{R}^N$ in N-dimensional space where $\mathbb{R}$ is the set of real numbers and $l \leq N$. Linear combination of these vectors with integer coefficients creates the lattice points which are then defined by

$$\Lambda = \left\{ z_1\mathbf{g}_1 + \cdots + z_l\mathbf{g}_l \,\middle|\, z_i \in \mathbb{Z}, i \in \{1, \cdots, l\} \right\}. \quad (1)$$

In this paper, we only consider full-rank lattices, i.e., $l = N$. In order to describe lattices in a more compact way, the square *generator matrix* $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_N] \in \mathbb{R}^{N \times N}$ is defined, where $\mathbf{g}_1, \ldots, \mathbf{g}_N$ are column vectors. Thus, $\det(\mathbf{G}) \neq 0$ and $\Lambda = \left\{ \mathbf{G}\mathbf{z} \mid \mathbf{z} \in \mathbb{Z}^N \right\}$. It should be noted that the generator matrix is not unique and different generator matrices can create the same lattice points.

In communication systems, information bits are mapped to symbols that are sent every T_s seconds and modulate a carrier which are then transmitted through a channel. Symbols are selected from a set of M elements, the *constellation*,
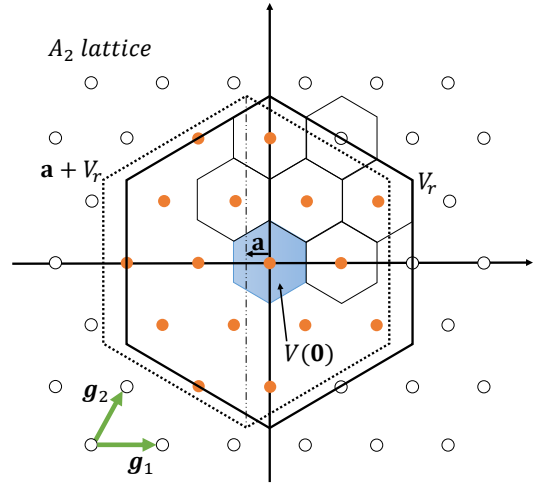

Fig. 1: Properties for A_2 lattice. The vectors $\mathbf{g}_1 = [1\ 0]^T$ and $\mathbf{g}_2 = [1/2\ \sqrt{3}/2]^T$ show the basis vectors of the lattice. Circular points show the lattice points, small hexagons are the Voronoi regions around the lattice points and $V(\mathbf{0})$ is the Voronoi region of the lattice. The filled points are the selected constellation points with a Voronoi cut. The boundary V_r indicates the scaled Voronoi region with $r = 4$, containing $r^2 = 16$ points and $\mathbf{a}$ is the shift vector.

$\mathcal{C} = \{\mathbf{c}_0, \mathbf{c}_1, ..., \mathbf{c}_{M-1}\}$. Constellations can be described in an N-dimensional space and therefore each symbol ($\mathbf{c}_i$) is represented by an N-dimensional vector. In this work, an equal probability of generating each symbol from the constellation is considered. Therefore, the entropy of this source or average information of each symbol is $m = \log_2 M$ *bits*, and the information rate of the source is equal to $R_b = \frac{m}{T_s}$ *bits per second*.

The average energy per symbol is defined as

$$E_s = \frac{1}{M} \sum_{i=0}^{M-1} \|\mathbf{c}_i\|^2 = m \cdot E_b, \quad (2)$$

where E_b is the average energy per bit. The average power is calculated as $P = E_s/T_s$.

In order to use the (infinite) set of points in a lattice in communications, we need to select a finite set of points as a constellation. In other words, we need to cut a constellation from the lattice. There are different ways of doing this, e.g., spherical or cubic cuts [21], [29], but what is interesting for us in this paper is the *Voronoi cut*. The Voronoi region around a lattice point $\mathbf{x}$, known as $V(\mathbf{x})$, is defined as the set of points in $\mathbb{R}^N$ which are closer to point $\mathbf{x}$ than any other points of the lattice, i.e.,

$$V(\mathbf{x}) = \left\{ \mathbf{y} \in \mathbb{R}^N \,\middle|\, \|\mathbf{y} - \mathbf{x}\| \leq \|\mathbf{y} - \mathbf{w}\|, \forall \mathbf{w} \in \Lambda \right\}. \quad (3)$$

It should be noted that the all-zero vector, $\mathbf{0}$, is always a lattice point and the Voronoi region $V(\mathbf{0})$ is called the Voronoi region of the lattice.

A VC is generally defined as the set of (possibly shifted) lattice points that lie inside the Voronoi region of a sparser lattice [24]. In this work, we focus on the special case where the

sparser lattice is a scaled version of Λ, for which simple and efficient modulation and demodulation algorithms are available [25]. Denoting the scale factor by the integer r, we define V_r as the Voronoi region of the lattice $r\Lambda = \{r\mathbf{x} \mid \mathbf{x} \in \Lambda\}$. The VC is now defined as $\mathcal{C}_\Lambda(r, \mathbf{a}) = \{\mathbf{x} - \mathbf{a} \mid \mathbf{x} \in \Lambda \cap (\mathbf{a} + V_r)\}$, where $\mathbf{a}$ is a *shift vector* suitably chosen to yield zero-mean constellations and at the same time resolve potential ties when lattice points lie on the boundary of V_r [25]. The number of points in the VC is $M = r^N$.

The concepts of basis vectors, infinite lattice, Voronoi regions, scaled and shifted Voronoi regions, and finite constellation points are all exemplified in Fig. 1 for the two-dimensional hexagonal lattice A_2.

The hexagonal lattice is the densest lattice in two-dimensional space. Finding the densest structure in a given dimension is known as a sphere packing problem, and is notoriously difficult. Lattices that are used in this paper are the cubic lattice ($\mathbb{Z}^N$), the hexagonal lattice (A_N, known to be densest in $N = 2$ and $N = 3$), the checkerboard lattice (D_N, densest in $N = 4$), the Gosset lattice (E_8, densest in $N = 8$), and the Leech lattice (Λ_{24}, densest in $N = 24$). For a more in-depth discussion of these lattices and their properties, see [30, chap. 4].

III. VORONOI CONSTELLATION DESIGN

In this section, we design VCs of various dimensions N and sizes M. This is not done by tabulating all constellation points, as in traditional geometric shaping, but by algebraic descriptions of the constellations. Specifically, a Voronoi constellation $\mathcal{C}_\Lambda(r, \mathbf{a})$ is fully determined by the generator matrix $\mathbf{G}$, the scale factor r, and the shift vector $\mathbf{a}$.

A. Modulation

After introducing the infinite lattices and creating finite constellations using Voronoi cuts, the question that remains is how to select a constellation point with low computational complexity to transmit the information bits [25].

Suppose that the constellation based on lattice Λ is $\mathcal{C}_\Lambda(r, \mathbf{a}) = \{\mathbf{c}_0, \mathbf{c}_1, \cdots, \mathbf{c}_{M-1}\}$ with $M = r^N$ constellation points. Using $m = \log_2 M$ bits, one of the constellation points can be uniquely selected. To avoid storage problems, we want to avoid storing constellation points in a look-up table, but generate each required constellation point based on its index each time it is needed. Therefore, m bits are mapped to a constellation index $K \in \{0, 1, \cdots, M - 1\}$. Then, following Algorithm 1 [25], a VC point for the index K can be selected.

In the first line of the algorithm, K is represented in base r using $K = k_{N-1}r^{N-1} + \cdots + k_1 r + k_0$. Then, in the subsequent lines, a point in the lattice is selected and finally mapped to the region where the constellation is considered based on $\mathcal{C}_\Lambda(r, \mathbf{a})$. In the fourth line, the closest point algorithm (CPA) is used to find the closest lattice point to $\mathbf{w}$. The CPA is described in [26] and we will discuss it briefly in section III-C.

B. Demodulation

After selecting a constellation point ($\mathbf{c}$) and transmitting it over the channel, in the receiver side, the received samples ($\mathbf{y} \in \mathbb{R}^N$) need to be mapped back to a probable VC point to find its index and the transmitted bits. This procedure is described in Algorithm 2 [25].

First, the received point is shifted from the finite constellation coordinates to the infinite lattice coordinates. Then, the closest lattice point is found by CPA and the index for this lattice point is calculated. Since this point can be outside the constellation area, it is moved back to a point inside the constellation region by the modulo operation (this procedure is shown in Fig. 2b by arrows). Finally, the base of the index is changed and bits can be recovered in the end.

The demodulation algorithm is almost the reverse procedure of the modulation algorithm and their core part is the CPA. Therefore, the complexity of the modulation and the demodulation algorithms are similar. This is one of the advantages of using the lattice structures as the constellation in transmission systems.

C. Closest point algorithm

The CPA which is used in the modulation and demodulation procedure is the main part of these algorithms and it provides an optimum way to find the closest point of an infinite lattice to a point in $\mathbb{R}^N$. The CPA was described in [26] for different lattices such as $\mathbb{Z}^N, A_N, D_N$, and E_8. To implement CPA for Λ_{24}, we use the information in [31] and [32] and the construction based on extended Golay code and D_{24} lattice. Usually, a more complicated lattice can be constructed using a union of cosets (shifted lattices) of other simpler lattices. Therefore, finding the closest lattice point in these complicated lattices will be simplified to applying CPA to the simpler cosets and finally comparing their results to find the closest point. For instance, A_2 can be viewed as two rectangular lattices (scaled $\mathbb{Z}^2$ lattices) and E_8 is the union of two cosets of D_8. Consequently, the CPA is applied over these cosets and their results are compared to find the closest point. The CPA also finds the closest lattice point in an infinite lattice

---

Algorithm 1: Modulation

Result: $K \rightarrow \mathbf{c}$
1: $K \xrightarrow{\text{changing base}} (k_{N-1} \cdots k_1 k_0)_r, \quad 0 \le k_i \le r - 1$
2: $\mathbf{x} = \sum_{i=1}^{N} k_i \mathbf{g}_i$
3: $\mathbf{w} = (\mathbf{x} - \mathbf{a})/r$
4: $\boldsymbol{\lambda} = \text{CPA}(\mathbf{w}), \quad \boldsymbol{\lambda} \in \Lambda$
5: $\mathbf{c} = \mathbf{x} - r\boldsymbol{\lambda} - \mathbf{a}, \quad \mathbf{c} \in \mathcal{C}_\Lambda(r, \mathbf{a})$

Algorithm 2: Demodulation

Result: $\mathbf{y} \rightarrow K$
1: $\mathbf{w} = \mathbf{y} + \mathbf{a}$
2: $\boldsymbol{\lambda} = \text{CPA}(\mathbf{w}), \quad \boldsymbol{\lambda} \in \Lambda$
3: $\mathbf{k}' = \mathbf{G}^{-1}\boldsymbol{\lambda}, \quad \mathbf{k}' = [k'_{N-1}, \cdots, k'_1, k'_0]^T$
4: $\mathbf{k} = \mathbf{k}' \pmod{r}, \quad \mathbf{k} = [k_{N-1}, \cdots, k_1, k_0]^T$
5: $(k_{N-1} \cdots k_1 k_0)_r \xrightarrow{\text{changing base}} K$

(a) Maximum likelihood (ML) detection.
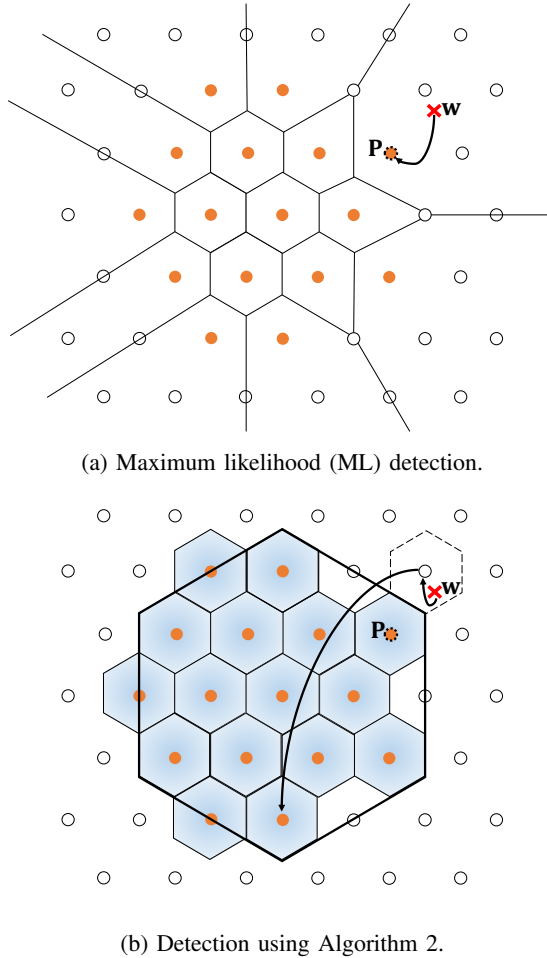

(b) Detection using Algorithm 2.

Fig. 2: Detection methods in the receiver.

to a point in $\mathbb{R}^N$ with few comparisons which depends on the number of dimensions regardless of the infinite number of lattice points. The CPA independence from the number of lattice points is one of the most important advantages of using lattice-based geometrically shaped constellations compared to other shaping methods.

D. The constellation shift vector

In order to find the best shift vector to minimize the constellation energy, an iterative algorithm was suggested in [25] which can converge in a few iterations, $\mathbf{a}_{n+1} = (1/M) \sum_{\mathbf{x} \in \mathcal{C}_\Lambda(r, \mathbf{a}_n)} \mathbf{x}$. However, if M is very large, as for some constellations in this paper (see Table II), then the sum cannot be exactly calculated, and if the sum is replaced by a Monte-Carlo approximation, then the algorithm is not guaranteed to converge.

In Fig. 3, we show that when the number of constellation points increases, we can ignore the optimization of finding the best shift vector, i.e., for large constellations, choosing any shift vector will give almost the same constellation energy as optimizing it. To show this concept, we define $\mu := \mathbb{E}_{\mathbf{a}_0} \left[|E_{s,0} - E_{s,\text{opt}}| / E_{s,\text{opt}} \right]$ where $\mathbb{E}_{\mathbf{a}_0}$ stands for expectation with respect to choosing $\mathbf{a}_0$ uniformly in $V(\mathbf{0})$, $E_{s,0}$ is the constellation energy using a random $\mathbf{a}_0$, and $E_{s,\text{opt}}$ is
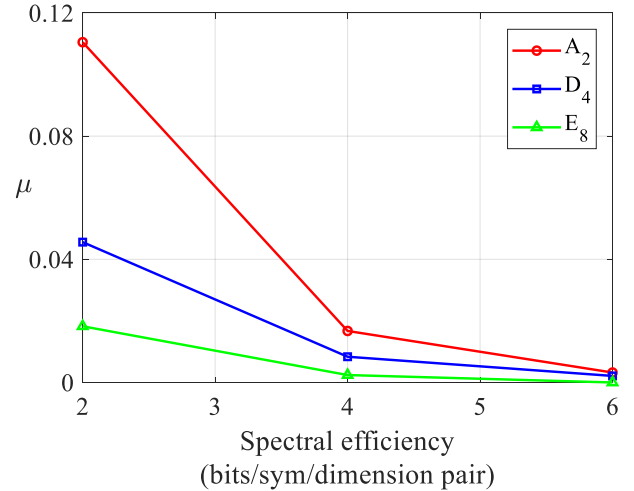

Fig. 3: Effect of random selection of the shift vector on the constellation energy.

the constellation energy optimized by the iterative algorithm in [25] initialized with $\mathbf{a}_0$. Figure 3 shows that μ decreases as the number of constellation points increases. Therefore, in constellations taken from the Leech lattice and E_8 lattice with SE of 8 bits/sym/dimension pair, where generating all constellation points is not feasible, we choose a random shift vector $\mathbf{a}$ in $V(\mathbf{0})$. In this case, the effect of constellation average offset is negligible on the constellation energy, as shown in Fig. 3. However, for other constellations in this paper, the proposed algorithm in [25] with $\mathbf{a}_0 = \mathbf{0}$ is applied and will always result in a zero-mean constellation.

IV. FIGURES OF MERIT

In this section, parameters to study and compare different multidimensional modulation formats are introduced.

1) Spectral efficiency: For a constellation with M separate symbols in N dimensions, the SE is calculated as

$$SE = \frac{\log_2 M}{N/2} = 2 \log_2 r \tag{4}$$

and has the unit of *bits per symbol per dimension pair*. In fiber-optic communication, usually, the polarization of light is considered as a pair of dimensions because each polarization can support two dimensions of in-phase and quadrature components. If sinc pulse shaping is used, the unit of SE can also be considered as *bits per second per Hertz* which shows how efficient the frequency spectrum is being used [33].

Based on Eq. (4), the SE is related to the scaling factor (r) of the Voronoi region. This shows that any SEs can be achieved by changing r for every lattice Λ. In this work, to simplify bit mapping, we focus on scaling factors that are powers of 2, i.e., $r = 2, 4, 8, 16$.

2) Symbol and bit error rate: For equiprobable constellations with ML estimation, using pairwise error probability and union bound, the symbol error rate (SER) is bounded by

$$\text{SER} \leq \frac{1}{M} \sum_{i=0}^{M-1} \sum_{j \neq i} \frac{1}{2} \text{erfc} \left(\frac{d_{ij}}{2\sqrt{N_0}} \right) \quad (5)$$

where d_{ij} is the Euclidean distance between $\mathbf{c}_i$ and $\mathbf{c}_j$. For high SNR, the terms with $d_{ij} = d_{min}$ dominate, and for lattices, all constellation points have at least one neighbor at distance d_{min} [34, chap. 5]. Therefore, the SER is bounded by

$$\frac{1}{2} \text{erfc} \left(\frac{d_{min}}{2\sqrt{N_0}} \right) \leq \text{SER} \overset{\sim}{\leq} \frac{\bar{\tau}}{2} \text{erfc} \left(\frac{d_{min}}{2\sqrt{N_0}} \right) \quad (6)$$

which $\overset{\sim}{\leq}$ means an approximate upper bound that approaches the true value as N_0 goes to zero. Also, $\bar{\tau} = (1/M) \sum_{i=0}^{M-1} \tau_i$ which τ_i is the number of neighbours at the minimum distance from symbol $\mathbf{c}_i$. Equation (6) shows that two important factors for SER of lattices are d_{min} and average of number of closest neighbors which for high-dimensional lattices can be very high [34, chap. 5].

The BER depends on the bit labeling of the symbols as well. For dense packing high dimensional lattices, there is no Gray mapping solution to optimize the BER performance. In this paper, instead of looking for the best bit mapping, we looked at two fast and low-complexity methods which we call natural binary and quasi-Gray constellation bit labeling. Natural binary labeling is the direct transformation of the symbol index (K) to its base-2 representation. Quasi-Gray labeling means transforming the symbol indexes in each dimension (k_i) to a binary Gray code.

3) Sensitivity penalty: For every modulation format, d_{min} is the minimum Euclidean distance between two symbols of the constellation. For high SNR and the AWGN channel, the probability of symbol error can be approximated using the union bound and its dominant terms which are proportional to $\text{erfc}(d_{min}/(2\sqrt{N_0}))$ where $\text{erfc}(x) = (2/\sqrt{\pi}) \int_x^\infty \exp(-t^2) dt$ is the complementary error function and $N_0/2$ is the variance of the Gaussian noise in each dimension. Therefore, the SER is a monotonically decreasing function of

$$\frac{d_{min}^2}{4N_0} = \frac{P}{R_b N_0} \gamma = \frac{E_b}{N_0} \gamma \quad (7)$$

where $\gamma = d_{min}^2/(4E_b)$ is defined as the asymptotic power efficiency because for a given SER, the required power is proportional to $1/\gamma$. The parameter γ describes the constellation geometry and is often given in dB. If different modulation formats are compared at the same power and bit rate, γ shows the power gain for high SNR with respect to BPSK, QPSK and dual polarization QPSK modulation formats because for these constellations $\gamma = 0$ dB. The sensitivity penalty is also defined as $1/\gamma$ and it shows the performance penalty compared to BPSK, QPSK and dual polarization QPSK for high SNR [14]. The sensitivity penalty comparison of different VCs with respect to QAM are shown in Fig. 4.
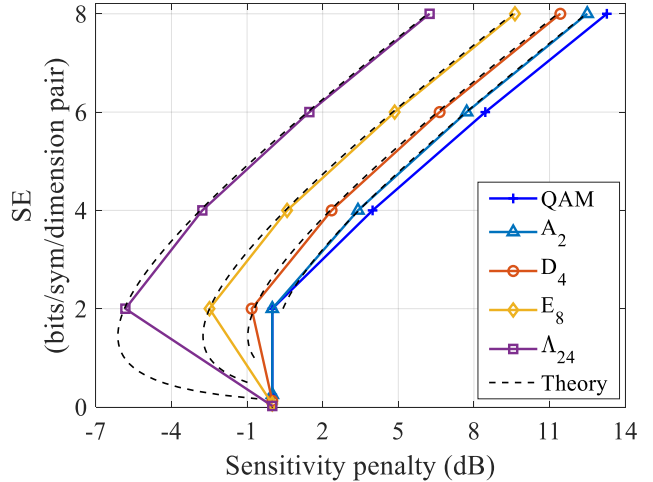

Fig. 4: The SE vs. sensitivity penalty for lattice base multidimensional modulation formats. The dashed lines are calculated using Eq. (23) in [14].

TABLE I: Shaping and coding gains of multidimensional lattices [24], [34, chap. 5]

Lattice	$\mathbb{Z}^N$	A_2	D_4	E_8	Λ_{24}
γ_c (dB)	0	0.62	1.51	3.01	6.02
γ_s (dB)	0	0.17	0.37	0.65	1.03

For multidimensional modulation formats, as the SE increases, the asymptotic power efficiency (γ) becomes proportional to the product of two important gains over the cubic lattice called the shaping gain (γ_s) and coding gain (γ_c) [11]. Shaping gain shows how spherical the boundary of the constellation is, and in this work, we use a Voronoi cut. The coding gain indicates the density of the lattice points. These gains for different lattices are shown in Table I and their mathematical definitions were presented in [11].

V. RESULTS

In this section, we present the simulation results applying the VCs in the AWGN and nonlinear fiber channel. The results are shown as the performance of these modulations in terms of uncoded BER and SER.

A. Additive white Gaussian noise channel

In this part, the bit labeling methods, detection algorithms in the receiver, and the performance of the VCs with Algorithms 1 and 2 are compared over the AWGN channel with respect to the equivalent QAM formats with ML detection.

To compare the natural binary and quasi-Gray constellation bit labeling, we simulate their BER performance over the AWGN channel and use them to encode and decode the transmitted and received VC symbols. The simulation setup includes a transmitter which generates multidimensional symbols and a multidimensional AWGN channel $\mathcal{N}(0, \frac{N_0}{2} \mathbf{I}_{N \times N})$. In the receiver, we use Algorithm 2 to estimate the VC
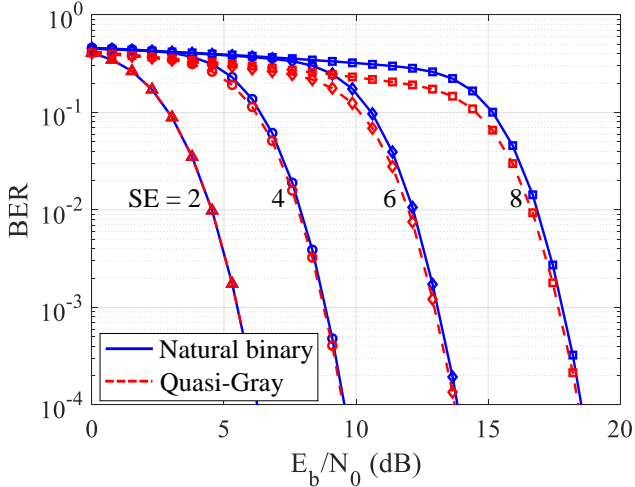
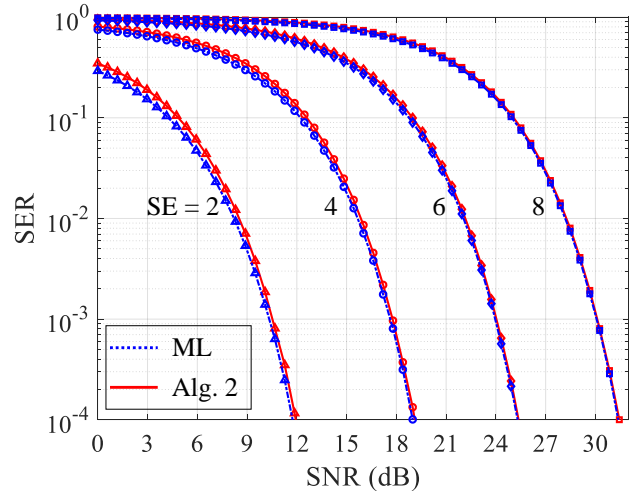Fig. 5: Natural binary vs. quasi-Gray bit labeling for Leech lattice at different SEs over AWGN channel.

Fig. 6: The ML estimation vs. Algorithm 2 performance for QAM formats at different SEs over AWGN channel.

symbols. At a $SE = 2$ bits/sym/dimension pair, these two methods create the same symbol labeling for the constellation points. However, for higher SEs, quasi-Gray performs better as shown in Fig. 5 for the Leech lattice (Λ_{24}).

In the receiver, an alternative detection scheme for VCs is the ML estimation which is the optimum detection rule for constellations with uniform probability distribution of symbols. The ML estimation compares the received information with all possible constellation points and finds the one which has the minimum Euclidean distance from the received point. The ML estimation requires a lookup table the size of the constellation cardinality. Therefore, for our VCs, ML detection can be extremely time consuming for the constellation sizes shown in Table II. However, the fast and low-complexity performance of Algorithm 2 comes at the expense of suboptimal behaviour, which is illustrated in Fig. 2 and 6.

In Fig. 2a, the decision regions for the ML estimation are shown. If point $\mathbf{P}$ is transmitted and $\mathbf{w}$ is received, since it is still in the decision region of point $\mathbf{P}$, it will be detected correctly. However, in Fig. 2b, using Algorithm 2, the received point will be mapped to an incorrect constellation point. This situation only happens when the received vector $\mathbf{w}$ is outside the Voronoi region of the transmitted vector ($V(\mathbf{P})$) and also outside the scaled Voronoi region (V_r). Therefore, when the number of constellation points increases, the ratio between the boundary constellation points and the total number of points decreases, and these errors will be less likely. Also, in high SNR regimes, the received points are more probable to remain
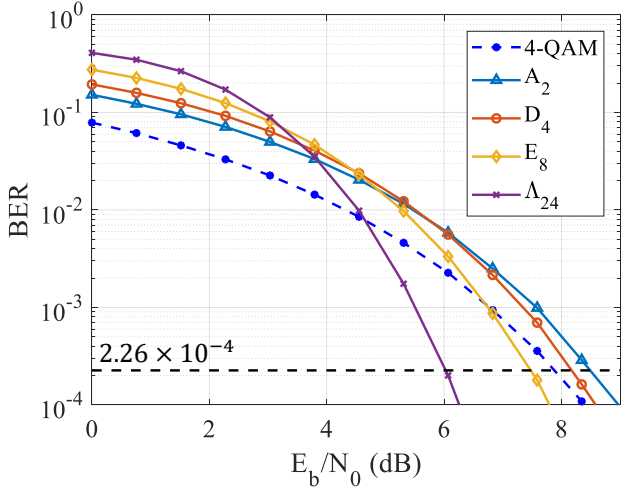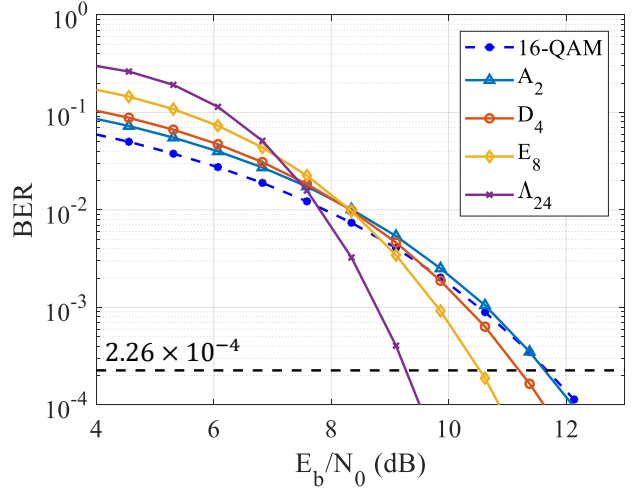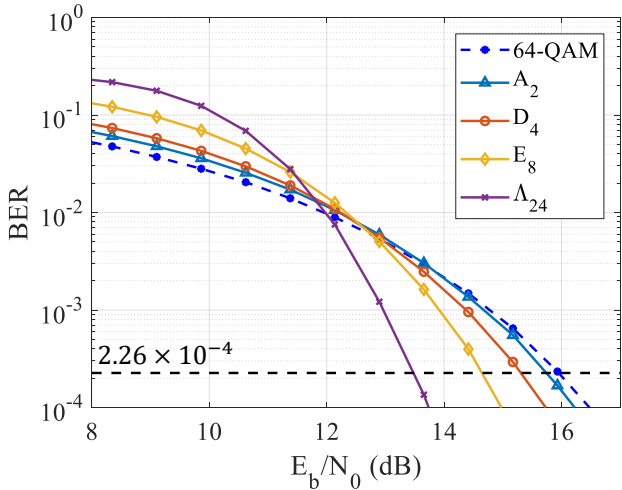
in the Voronoi region of their transmitted points therefore Algorithm 2 is able to detect them correctly similar to the ML estimation.

In Fig. 6, the ML estimation and Algorithm 2 are compared for QAM formats in different SEs over AWGN channel. Interestingly, as the SNR (defined as $E_s/(N \cdot N_0/2)$) or the number of the constellation points increases, the gap between the SER performance of Algorithm 2 and the ML estimation decreases and they approach each other. Therefore, it can be concluded that, for finite VCs, Algorithm 2 is close to optimum for constellations with high cardinality, or in the limit of high SNR.

Finally, in Fig. 7, the BER performance of VCs in different dimensions are compared with the equivalent QAM format at the same SE. Except for the A_2 lattice in Fig. 7a, all VCs will eventually perform better than the QAM format as the E_b/N_0 increases. The reason for this behavior is the asymptotic power efficiency that was discussed in section IV. However, for low E_b/N_0, QAM formats outperform VCs because of Gray labeling and smaller $\bar{\tau}$. In Fig. 4, at $SE = 2$ bits/sym/dimension pair, the A_2 and the QAM format have the same sensitivity penalty, however, in Fig. 7a, 4-QAM always performs better than A_2 in all E_b/N_0. This is because of 4-QAM Gray labeling and the average number of closest neighbors which for the A_2 lattice with 4 constellation points is $\bar{\tau} = 2.5$ and for the 4-QAM is $\bar{\tau} = 2$.

B. Nonlinear optical channel

An important characteristic of the VCs is their performance in the nonlinear fiber-optic channel. In order to investigate this problem, a nonlinear fiber-optic channel has been simulated using the split-step Fourier method (SSFM). In an experimental setup, multidimensional modulation formats can be realized in different ways. Among the possible dimensions in fiber, in this work, we focus on in-phase, quadrature, polarization, and wavelength. Using different combinations of dimensions, e.g.,

TABLE II: Constellation size (M) in different dimensions and SEs

SE	2	4	6	8
A_2	4	16	64	256
D_4	16	256	4,096	65,536
E_8	256	65,536	16,777,216	4,294,967,296
Λ_{24}	16,777,216	$\approx 2.8 \times 10^{14}$	$\approx 4.7 \times 10^{21}$	$\approx 7.9 \times 10^{28}$

(a) SE = 2 bits/sym/dimension pair

(b) SE = 4 bits/sym/dimension pair

(c) SE = 6 bits/sym/dimension pair
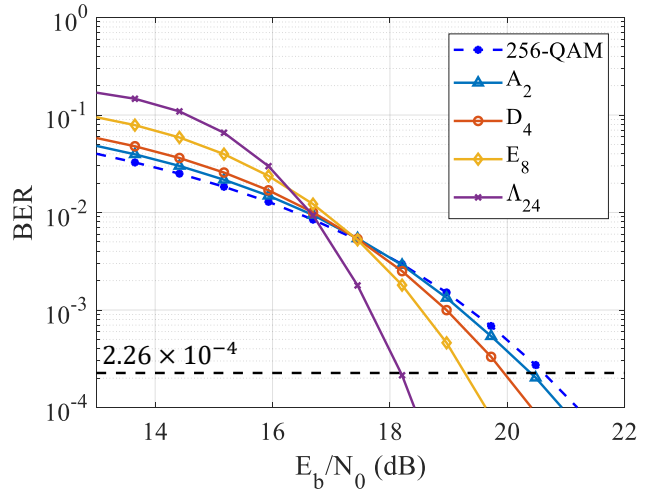
(d) SE = 8 bits/sym/dimension pair

Fig. 7: The BER vs. E_b/N_0 for multidimensional modulation formats over AWGN channel

time slots, wavelengths, spatial modes, in practice will give different performance results because the coupling and cross-talk between the dimensions in a nonlinear channel might differ. In the wavelength-division multiplexing (WDM) system considered in this paper, multiple synchronized coherent receivers are needed to detect each channel. Another design could possibly include time slots as additional dimensions which will reduce the number of coherent receivers and the nonlinear interaction of channels.

The simulation setup consists of a transmitter side including a symbol generator that produces multidimensional symbols according to the selected VC. These symbols are then parallelized into multiple groups of 4-dimensional symbols because each wavelength can carry up to 4 dimensions in a single-mode fiber. At this stage, each group of 4-dimensional symbols is combined with pilot symbols that have QPSK modulation format and they are used for pilot-aided digital signal processing in the receiver. Finally, in the transmitter, the combined payloads and pilots are upsampled and filtered with a root-

raised-cosine (RRC) filter. Then, each group is shifted to its corresponding wavelength and they are added together to form a WDM system. Before the channel, the total optical power of the signal is set and then using the Manakov equations [35] and SSFM [36, chap. 2], the signal is propagated over the nonlinear dispersive fiber-optic channel. The signal is amplified using an erbium-doped fiber amplifier (EDFA) at the end of each span in the fiber link. In our simulations, uniform QAM constellations over multiple wavelengths and polarizations are considered as benchmarks to compare the transmission performance with VCs at the same SE.

On the receiver side, first, digital dispersion compensation is applied for the whole propagation link distance. Nonlinearity-compensation methods, such as digital backpropagation, are not performed in the receiver; however, they can improve the performance of the constellations by compensating the fiber impairments at the cost of increased complexity. Then, each wavelength channel is filtered out using an RRC filter and downsampled. Using the QPSK pilots, polarization de-
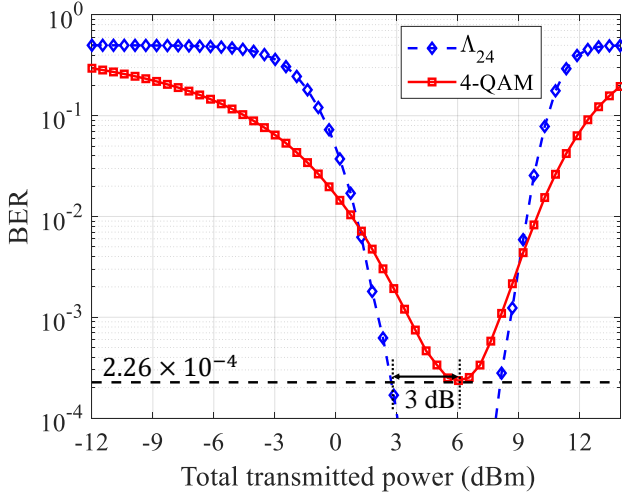
Fig. 8: The BER vs. total transmitted optical power after 80 spans of 80 km fibers over a WDM system with 6 wavelengths.
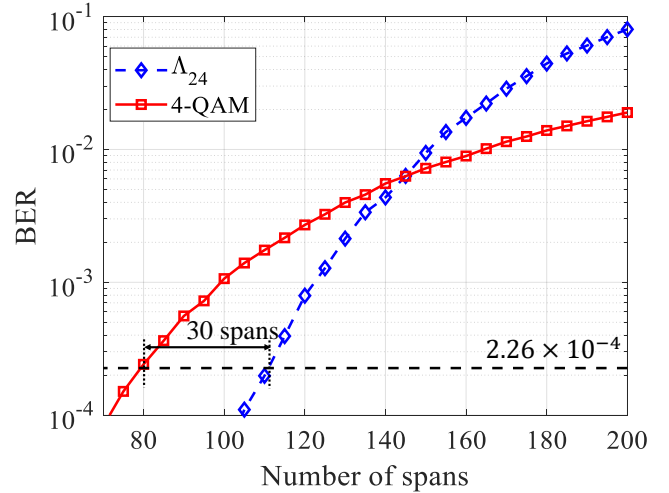

Fig. 9: The BER vs. transmission distance at the optimum optical power over nonlinear fiber-optic channel.

TABLE III: Nonlinear simulation setup parameters

Parameter	Value
Symbol rate	28 GBaud
RRC roll-off factor	0.2
WDM spacing	50 GHz
Fiber nonlinear coefficient	1.3 $W^{-1}km^{-1}$
Fiber dispersion	17 ps/nm/km
Fiber attenuation	0.2 dB/km
Span length	80 km
EDFA noise figure	5 dB
SSFM step size	0.5 km
Oversampling factor	32
Pilot overhead	1.56%

multiplexing and phase tracking are applied to the payload symbols. After that, the processed payloads are combined together to form the multidimensional symbols again and using Algorithm 2 and ML estimation, it is decided what the transmitted symbol has been for VCs and QAM format, respectively. In the end, BER is calculated based on the transmitted and the received bits. The parameters of the nonlinear simulation setup are listed in Table III.

In Fig. 8, the BER performance of the Leech lattice is compared with the 4-QAM at the same SE. The total propagation distance is 80 spans (6400 km) using 6 wavelengths to transmit the 24-dimensional signals. At a BER of 2.26×10^{-4}, which corresponds to 10^{-15} after error correction using a Reed-Solomon (544, 514) (KP4) code [37], the achieved transmitted power gain is more than 3 dB.

In Fig. 9, the BER performance is shown at the optimum power in different transmission distances. According to our simulations, the optimum power almost remains constant at different transmission distances. In this representation, transmission distance improvement of more than 30 spans or 38% is shown at a BER equal to 2.26×10^{-4}.

Other shaping methods, e.g., probabilistic amplitude shaping, offer around 7 to 15% reach improvement at a given achievable information rate, compared with uniform QAM [7], [38], [39]. However, some caution is advised when comparing

reach gains obtained using different performance metrics. For example, the SER and BER benefit from both coding gain and shaping gain, whereas information rates typically reflect shaping gain only. Nevertheless, it seems clear that our geometric shaping method can be considered as a competitive shaping approach with low complexity to increase the performance of optical communication systems.

Future work may include a more detailed comparison with probabilistic shaping, as well as trellis shaping [27], [28]. Also, combinations of VC and hard-decision forward error correction codes can be investigated in low latency and low power consumption applications, e.g., video conferencing or instrument control links that are latency sensitive.

VI. CONCLUSION

We have analyzed Voronoi constellations as a geometric shaping method for optical communication systems. The complexity of modulation and demodulation are low compared with other geometric schemes because modulation and demodulation are carried out algorithmically and the constellation need not be stored in a table, paving the way for more or less unlimited cardinalities. Reach enhancements as high as 38% is found in nonlinear fiber link simulations. We conclude that lattice-based geometric shaping can provide higher gains than what has been reported using probabilistic shaping, at much shorter block lengths and moderate complexity.

VII. ACKNOWLEDGMENT

REFERENCES

[1] P. J. Winzer, "High-spectral-efficiency optical modulation formats," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3824–3835, 2012.

[2] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: the previous 20 and the next 20 years," *Optics Express*, vol. 26, no. 18, pp. 24 190–24 239, 2018.

[3] M. Seimetz, *High-Order Modulation for Optical Fiber Transmission*. Springer, 2009, vol. 143.

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

[5] G. D. Forney, Jr., R. G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. Qureshi, "Efficient modulation for band-limited channels," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 5, pp. 632–647, 1984.

[6] J. Cho and P. J. Winzer, "Probabilistic constellation shaping for optical fiber communications," *Journal of Lightwave Technology*, vol. 37, no. 6, pp. 1590–1607, 2019.

[7] T. Fehenberger, 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, 2016.

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

[9] R. T. Jones, T. A. Eriksson, M. P. Yankov, and D. Zibar, "Deep learning of geometric constellation shaping including fiber nonlinearities," in *European Conference on Optical Communication (ECOC)*, 2018.

[10] A. A. El-Rahman and J. Cartledge, "Multidimensional geometric shaping for QAM constellations," in *European Conference on Optical Communication (ECOC)*, 2017.

[11] G. D. Forney, Jr. and L.-F. Wei, "Multidimensional constellations-Part I: Introduction, figures of merit, and generalized cross constellations," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 877–892, 1989.

[12] M. Karlsson and E. Agrell, "Four-dimensional optimized constellations for coherent optical transmission systems," in *European Conference on Optical Communication (ECOC)*, 2010.

[13] E. Ip, A. P. T. Lau, D. J. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Optics Express*, vol. 16, no. 2, pp. 753–791, 2008.

[14] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *Journal of Lightwave Technology*, vol. 27, no. 22, pp. 5115–5126, 2009.

[15] M. Karlsson and E. Agrell, "Power-efficient modulation schemes," in *Impact of Nonlinearities on Fiber Optic Communications*. Springer, 2011, pp. 219–252.

[16] T. Koike-Akino, D. S. Millar, K. Kojima, and K. Parsons, "Eight-dimensional modulation for coherent optical communications," in *European Conference on Optical Communication (ECOC)*, 2013.

[17] T. Koike-Akino and V. Tarokh, "Sphere packing optimization and EXIT chart analysis for multi-dimensional QAM signaling," in *IEEE International Conference on Communications (ICC)*, 2009.

[18] T. A. Eriksson, P. Johannisson, E. Agrell, P. A. Andrekson, and M. Karlsson, "Biorthogonal modulation in 8 dimensions experimentally implemented as 2PPM-PS-QPSK," in *Optical Fiber Communication Conference (OFC)*, 2014.

[19] X. Lu, A. Tatarczak, V. Lyubopytov, and I. Tafur Monroy, "Optimized eight-dimensional lattice modulation format for IM-DD 56 Gb/s optical interconnections using 850 nm VCSELs," *Journal of Lightwave Technology*, vol. 35, no. 8, pp. 1407–1414, 2017.

[20] X. Lu, V. S. Lyubopytov, and I. Tafur Monroy, "24-dimensional modulation formats for 100 Gbit/s IM-DD transmission systems using 850 nm single-mode VCSEL," in *European Conference on Optical Communication (ECOC)*, 2017.

[21] D. S. Millar, T. Koike-Akino, S. Ö. Arık, K. Kojima, K. Parsons, T. Yoshida, and T. Sugihara, "High-dimensional modulation for coherent optical communications systems," *Optics Express*, vol. 22, no. 7, pp. 8798–8812, 2014.

[22] T. Matsumine, T. Koike-Akino, D. S. Millar, K. Kojima, and K. Parsons, "Short lattice-based shaping approach exploiting non-binary coded modulation," in *European Conference on Optical Communication (ECOC)*, 2019.

[23] A. Shiner, M. Reimer, A. Borowiec, S. O. Gharan, J. Gaudette, P. Mehta, D. Charlton, K. Roberts, and M. O'Sullivan, "Demonstration of an 8-dimensional modulation format with reduced inter-channel nonlinearities in a polarization multiplexed coherent system," *Optics Express*, vol. 22, no. 17, pp. 20 366–20 374, 2014.

[24] G. D. Forney, Jr., "Multidimensional constellations-Part II. Voronoi constellations," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 941–958, 1989.

[25] J. Conway and N. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Transactions on Information Theory*, vol. 29, no. 6, pp. 820–824, 1983.

[26] ——, "Fast quantizing and decoding and algorithms for lattice quantizers and codes," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 227–232, 1982.

[27] B. P. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," *Journal of Lightwave Technology*, vol. 30, no. 13, pp. 2047–2053, 2012.

[28] G. D. Forney, Jr., "Trellis shaping," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 281–300, 1992.

[29] M. Karlsson and E. Agrell, "Spectrally efficient four-dimensional modulation," in *Optical Fiber Communications Conference (OFC)*, 2012.

[30] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. Springer, 1999.

[31] J. H. Conway and N. J. Sloane, "On the Voronoi regions of certain lattices," *SIAM Journal on Algebraic Discrete Methods*, vol. 5, no. 3, pp. 294–305, 1984.

[32] J.-P. Adoul and M. Barth, "Nearest neighbor algorithm for spherical codes from the Leech lattice," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 1188–1202, 1988.

[33] M. Karlsson and E. Agrell, "Multidimensional Modulation and Coding in Optical Transport," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 876–884, 2016.

[34] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: with Wireless Applications*. Springer, 1999.

[35] S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Nonlinear propagation in multimode and multicore fibers: Generalization of the Manakov equations," *Journal of Lightwave Technology*, vol. 31, no. 3, pp. 398–406, 2012.

[36] G. P. Agrawal, *Nonlinear Fiber Optics*, 5th ed. Springer, 2013.

[37] E. Agrell and M. Secondini, "Information-theoretic tools for optical communications engineers," in *IEEE Photonics Conference (IPC)*, 2018.

[38] C. Pan and F. R. Kschischang, "Probabilistic 16-QAM shaping in WDM systems," *Journal of Lightwave Technology*, vol. 34, no. 18, pp. 4285–4292, 2016.

[39] T. Fehenberger, G. Böcherer, A. Alvarado, and N. Hanik, "LDPC coded modulation with probabilistic shaping for optical fiber systems," in *Optical Fiber Communication Conference (OFC)*, 2015.