Two-Stage Dimension-Wise Coded Modulation for Four-Dimensional Hurwitz-Integer Constellations

Sebastian Stern[1], Felix Frey[1,2], Johannes K. Fischer[2], Robert F.H. Fischer[1]
[1]Institute of Communications Engineering, Ulm University, Ulm, Germany
[2]Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Berlin, Germany
Email: {sebastian.stern, felix.frey, robert.fischer}@uni-ulm.de, johannes.fischer@hhi.fraunhofer.de

***Abstract*—In state-of-the-art fiber-optical systems and upcoming wireless standards, transmission over both planes of polarization of electromagnetic waves is an important approach to increase the spectral efficiency. To this end, four-dimensional modulation schemes are suited. In this paper, coded modulation for signal constellations over the Hurwitz integers, an isomorphic representation of the four-dimensional checkerboard lattice, is studied. In particular, a two-stage coding strategy is proposed where the components of the four-dimensional signal are preprocessed jointly before conventional coded-modulation techniques like bit-interleaved coded modulation or multistage decoding are applied individually per dimension. The theoretical capacities as well as numerical results from simulations with LDPC codes show that the proposed approach enables a performance gain over the straightforward application of ASK constellations per dimension or QAM constellations per polarization.**

I. Introduction

In order to satisfy the growing demand for ultra-high data-rate transmission, modulation techniques that enable a very high bandwidth efficiency are required. A straightforward but effective strategy is to use both (orthogonal) planes of polarization of electromagnetic waves. In *fiber-optical transmission*, this approach has already been used for quite some time. Recently, in the field of wireless communications, so-called *dual-polarized antennas* for the radiation of both horizontally- and vertically-polarized waves have been proposed [5] and become more and more popular, cf., e.g., [11], [13].

Taking advantage of radio-frequency modulation in both planes of polarization, four orthogonal components are available in equivalent baseband domain. This can be modeled as a quaternion-valued (QV) transmission [8], [13], [18], i.e., with signal points from the set of quaternions [4]. Then, the system performance may not only benefit from a doubled spectral efficiency, but also from QV modulation formats. In particular, instead of choosing the signal set from the four-dimensional (4D) integer lattice $\mathbb{Z}^4$—represented by the QV set of *Lipschitz integers* [4]—it is more advantageous to take a subset of the checkerboard lattice $\boldsymbol{D}_4$—represented by the QV set of *Hurwitz integers* [3], [4]: the densest 4D packing [3] reduces the constellation's variance and hence the transmit power.

In fiber-optical transmission, constellations based on the Lipschitz integers are well studied and known as polarization-multiplexed quadrature-amplitude modulation (PM-QAM). In addition, constellations from the lattice $\boldsymbol{D}_4$ have been proposed, e.g., [10] gives an overview on 4D modulation in optical communication. Nevertheless, the central issue of the latter is the combination with channel coding schemes, i.e., the development of a strategy for coded modulation (CM). Particularly, bit-interleaved coded modulation (BICM) [2] is not suited, cf. [1], as—due to an increased number of nearest neighbors [3]—a Gray labeling does not exist. A straightforward capacity-approaching CM strategy in combination with state-of-the-art channel codes is still an open problem.

To overcome this problem, in this paper, we propose a novel CM approach for signal sets based on Hurwitz integers. Following the concept of capacity-achieving multistage decoding (MSD) [16], the encoding and hence the corresponding decoding are split into two main stages: The first stage handles the 4D (QV) signal as a whole. Based thereon, in the second main stage, it is split into its four components where conventional CM schemes are applied—either BICM with Gray labeling in the style of multistage bit-wise (MSBW) receivers [17], or multilevel coding (MLC) by continuing the philosophy of MSD. For both cases, based on the capacities of the *equivalent channels* [16], we explain how to choose the rates of the component codes to maximize the performance over the additive white Gaussian noise (AWGN) channel. On that basis, numerical simulations with low-density parity-check (LDPC) codes show the benefits of the proposed strategies.

The paper is structured as follows: Sec. II reviews 4D lattices and related signal constellations. Two-stage dimension-wise CM for Hurwitz-integer constellations is proposed in Sec. III and related numerical results are provided in Sec. IV. The paper closes with a summary and outlook in Sec. V.

II. Four-Dimensional Lattices and Constellations

In digital transmission, it is common practice to draw the data symbols from a regular grid, in particular a *lattice* [3]. In the following, we will briefly review the most important 4D lattices and related signal constellations.

A. Four-Dimensional Lattices

We consider 4D lattices of the form

$$\boldsymbol{\Lambda}(\boldsymbol{G}) = \left\{ \boldsymbol{v} = \boldsymbol{G}\boldsymbol{u} \mid \boldsymbol{u} \in \mathbb{Z}^4 \right\} , \tag{1}$$

where $\boldsymbol{G} \in \mathbb{R}^{4\times 4}$ denotes the *generator matrix* of the lattice that produces lattice points $\boldsymbol{v} = [v_1, v_2, v_3, v_4]^{\mathsf{T}} \in \mathbb{R}^4$.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—FI 982/13-1 & FI 982/14-1.

If 2D lattices are considered—especially in case of complex baseband processing of (single-polarized) radio-frequency signals—it may be convenient to employ (scalar) complex numbers instead of 2D vectors. Transferring this concept to the 4D (dual-polarized) case, the set of *quaternions* [3], [4]

$$\mathbb{H} = \{v = v^{(1)} + v^{(2)}\mathrm{i} + v^{(3)}\mathrm{j} + v^{(4)}\mathrm{k} \mid v^{(1,2,3,4)} \in \mathbb{R}\} \quad (2)$$

with the imaginary units i, j, and k (where $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$) enables an equivalent scalar representation of the lattice vectors according to (1). Noteworthy, this QV representation is well suited to describe a rotation and/or crosstalk between the polarization planes, cf., e.g., [8], [18].

1) Lipschitz Integers (Lattice $\mathbb{Z}^4$): The simplest 4D lattice is the integer lattice $\mathbb{Z}^4$ with the 4×4 identity generator matrix $\boldsymbol{G} = \boldsymbol{I}$. It is isomorphic to the QV set of *Lipschitz integers* [4]

$$\mathcal{L} = \{v = v^{(1)} + v^{(2)}\mathrm{i} + v^{(3)}\mathrm{j} + v^{(4)}\mathrm{k} \mid v^{(1,2,3,4)} \in \mathbb{Z}\}\ , \quad (3)$$

i.e., all quaternions that only consist of integer components. All elements of $\mathcal{L}$ (or $\mathbb{Z}^4$) have eight nearest neighbors at the minimum (squared) distance $d^2_{\min} = 1$.

2) Hurwitz Integers (Lattice $\boldsymbol{D}_4$): The *checkerboard* or *Schläfli lattice* $\boldsymbol{D}_4$ is the *densest packing* in four dimensions, particularly achieving a packing gain of 1.51 dB w.r.t. $\mathbb{Z}^4$ [3], [7]. One possible generator matrix thereof reads [3]

$$\boldsymbol{G} = \begin{bmatrix} 1 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 1/2 \\ 0 & 0 & 0 & 1/2 \end{bmatrix}\ . \quad (4)$$

The lattice $\boldsymbol{D}_4$ according to (4) is isomorphically expressed by the QV set of *Hurwitz integers* [3], [4]

$$\begin{aligned}\mathcal{H} &= \{v = v^{(1)} + v^{(2)}\mathrm{i} + v^{(3)}\mathrm{j} + v^{(4)}\mathrm{k} \mid \\ &\qquad [v^{(1)}, v^{(2)}, v^{(3)}, v^{(4)}]^{\mathsf{T}} \in \mathbb{Z}^4 \cup (\mathbb{Z}+1/2)^4\} \quad (5)\\ &= \mathcal{L} \cup (\mathcal{L} + (1+\mathrm{i}+\mathrm{j}+\mathrm{k})/2)\ .\end{aligned}$$

The Hurwitz integers (and the lattice $\boldsymbol{D}_4$) can be split into two disjunct sets: i) the set of Lipschitz integers and ii) the set of Lipschitz integer shifted by $1/2$ in each component. In comparison to $\mathcal{L}$ (or $\mathbb{Z}^4$), the number of nearest neighbors is tripled to 24 [3], but the related minimum (squared) distance remains $d^2_{\min} = 1$. Graphically, in four dimensions, there is enough space to place additional lattice points in between the (Lipschitz) integers without decreasing the minimum distance.

B. Four-Dimensional Signal Constellations

In digital transmission, the set of data symbols $\mathcal{A}$ (signal constellation) can be chosen to be a (zero-mean) subset of a lattice—the so-called *signal-point lattice* [7]—with cardinality $M = |\mathcal{A}|$ and variance σ_a^2. In the 4D case, both abovementioned types of lattices are suited and discussed below.

1) Lipschitz Constellations: Following the concept of (1D) amplitude-shift keying (ASK) and (2D) square QAM constellations, signal constellations based on the Lipschitz integers as the signal-point lattice are defined by [13]

$$\mathcal{A}_{\mathcal{L}} = \mathcal{A}_{\mathrm{d}} + \mathcal{A}_{\mathrm{d}}\,\mathrm{i} + \mathcal{A}_{\mathrm{d}}\,\mathrm{j} + \mathcal{A}_{\mathrm{d}}\,\mathrm{k}\ , \quad (6)$$

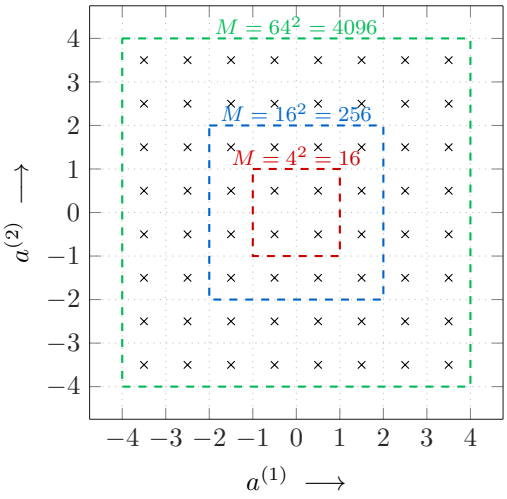

Fig. 1. 2D projection (first two components) of the Lipschitz constellations $\mathcal{A}_{\mathcal{L}}$ with $M = 2^4 = 4^2 = 16$ (red boundaries; $b = 4$ bit), $M = 4^4 = 16^2 = 256$ (blue boundaries; $b = 8$ bit), and $M = 8^4 = 64^2 = 4096$ (green boundaries; $b = 12$ bit).

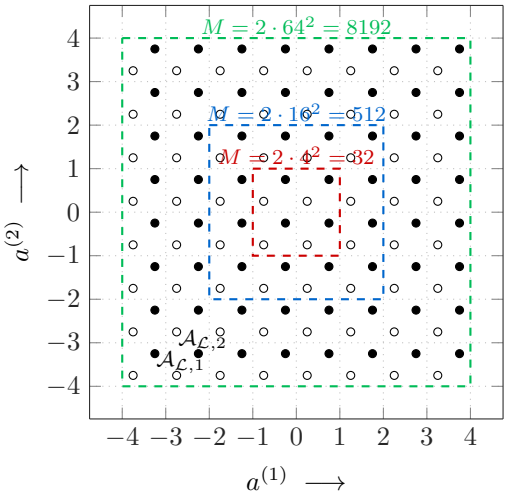

Fig. 2. 2D projection (first two components) of the Hurwitz constellations $\mathcal{A}_{\mathcal{H}}$ with $M = 2{\cdot}2^4 = 2{\cdot}4^2 = 32$ (red boundaries; $b = 5$ bit), $M = 2{\cdot}4^4 = 2{\cdot}16^2 = 512$ (blue boundaries; $b = 9$ bit), and $M = 2{\cdot}8^4 = 2{\cdot}64^2 = 8192$ (green boundaries; $b = 13$ bit). Subset $\mathcal{A}_{\mathcal{L},1}$: circles filled white; subset $\mathcal{A}_{\mathcal{L},2}$: circles filled black.

with the zero-mean M_{d}-ary independent components

$$\mathcal{A}_{\mathrm{d}} = \{m - (M_{\mathrm{d}}-1)/2 \mid m \in \{0, \ldots, M_{\mathrm{d}}-1\}\}\ . \quad (7)$$

In Fig. 1, where the 2D projections of some Lipschitz-based constellations are shown, we see that the data symbols are located within a hypercube. If non-scaled versions thereof are chosen (i.e., $d^2_{\min} = 1$), the variance per component reads $\sigma^2_{a,\mathrm{d}} = (M_{\mathrm{d}}^2 - 1)/12$. In total, we have $\sigma_a^2 = 4\,\sigma^2_{a,\mathrm{d}}$. It is convenient to choose $M_{\mathrm{d}} = 2^{b_{\mathrm{d}}}$, where $b_{\mathrm{d}} \in \mathbb{N}\backslash 0$ is the number of bits to be transmitted per dimension. Then, the total cardinality reads $M = M_{\mathrm{d}}^4 = 2^{4b_{\mathrm{d}}} = 16, 256, 4096, \ldots$, and the total number of bits per symbol $b = 4\,b_{\mathrm{d}} = 4, 8, 12, \ldots$. These parameters are summarized in Table I. Noteworthy, since M_{d} is even, the offset $o_{\mathcal{L}} = (1+\mathrm{i}+\mathrm{j}+\mathrm{k})/2$ to the Lipschitz integers $\mathcal{L}$ is required to be zero-mean, i.e., $\mathcal{A}_{\mathcal{L}} \subset \mathcal{L} + o_{\mathcal{L}}$.

TABLE I
VARIANCE AND NUMBER OF BITS OF (ZERO-MEAN) LIPSCHITZ AND HURWITZ CONSTELLATIONS IN TOTAL AND PER DIMENSION.

Type	M	M_d	σ_a^2	$\sigma_{a,\mathrm{d}}^2$	b	b_d
$\mathcal{A}_\mathcal{L}$	16	2	1	0.25	4	1
$\mathcal{A}_\mathcal{H}$	32	2.3784	1.25	0.3125	5	1.25
$\mathcal{A}_\mathcal{L}$	256	4	5	1.25	8	2
$\mathcal{A}_\mathcal{H}$	512	4.7568	5.25	1.3125	9	2.25
$\mathcal{A}_\mathcal{L}$	4096	8	21	5.25	12	3
$\mathcal{A}_\mathcal{H}$	8192	9.5137	21.25	5.3125	13	3.25

2) Hurwitz Constellations: If the Hurwitz integers are taken as the signal-point lattice we can take advantage of the abovementioned packing gain. Since they consist of two Lipschitz-based subsets, the straight-forward construction [13]

$$\mathcal{A}_\mathcal{H} = \mathcal{A}_{\mathcal{L},1} \cup \mathcal{A}_{\mathcal{L},2} \tag{8}$$

can be applied, where we have the two Lipschitz-based subsets

$$\mathcal{A}_{\mathcal{L},1} = \mathcal{A}_\mathcal{L} - (1+\mathrm{i}+\mathrm{j}+\mathrm{k})/4 \tag{9}$$

$$\mathcal{A}_{\mathcal{L},2} = \mathcal{A}_\mathcal{L} + (1+\mathrm{i}+\mathrm{j}+\mathrm{k})/4\,. \tag{10}$$

In Fig. 2, the 2D projections of some Hurwitz constellations are shown. We see that two Lipschitz-based signal sets (9) and (10) with the offset $\pm o_\mathcal{H} = \pm(1+\mathrm{i}+\mathrm{j}+\mathrm{k})/4$ to $\mathcal{A}_\mathcal{L}$ are actually present; however, the minimum (squared) distance of the resulting Hurwitz constellation (8) is still $d_\mathrm{min}^2 = 1$. In particular, we have $\mathcal{A}_\mathcal{H} \subset \mathcal{H} + o_\mathcal{H}$, where the offset ensures that the resulting constellation is zero-mean.

In comparison to the Lipschitz constellations, the number of signal points is doubled within the same boundary regions (cf. Fig. 1 vs. Fig. 2). It is quite obvious that, given Lipschitz-based subsets with 16, 256, or 4096 signal points, we obtain the cardinalities $M = 32$, $M = 512$, or $M = 8192$, cf. Table I. As a consequence, one additional bit can be transmitted (in total), or one quarter bit per dimension. Due to the same boundaries, the constellations' variances remain nearly the same in comparison to the subsets. As can be seen from Table I, only a slight increase of 0.0625 per dimension or 0.25 in total is present, caused by $o_\mathcal{H}$ [13]. Nevertheless, due to 24 nearest neighbors, a Gray labeling is not possible and a straightforward application of BICM is not promising.

In optical transmission, these constellations are also known as M-ary set-partitioned QAM (M-SP-QAM) [10], which result from the *extension*[1] of two independent $\sqrt{M/2}$-ary QAM constellations via the addition or subtraction of the 4D offset (two *set partitions*). In contrast to the QV representation, they are usually represented as (shifted) subsets of the lattice $\boldsymbol{D}_4$.

C. Capacities over the AWGN Channel

In order to assess the system performance in case of coded transmission, the coded-modulation (or constellation-constrained) capacity over the AWGN channel C^CM is a relevant quantity. It can be obtained by means of numerical integration [14], e.g., via Gauss-Hermite quadratures [12].

[1]Moreover, another type of $\boldsymbol{D}_4$-based constellations can be obtained by a technique called *reduction* [10], which, however, prevents independent modulation per partition and dimension (by analogy with *rectangular* vs. *non-rectangular* QAM). In this paper, we restrict to extended constellations.

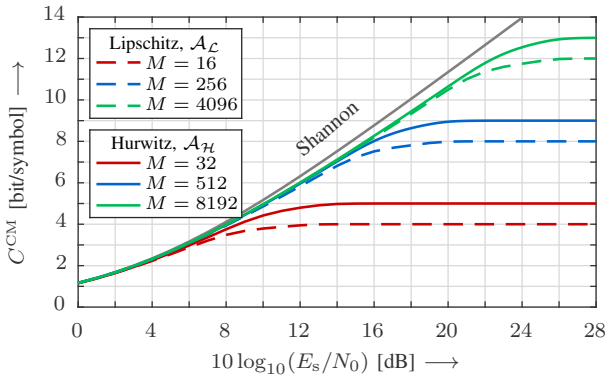

Fig. 3. Coded-modulation capacities C^CM in bit/symbol vs. signal-to-noise ratio over the AWGN channel for Lipschitz (dashed) and Hurwitz (solid) constellations. The Shannon limit is given as gray solid line.

In Fig. 3, the capacities of the above Lipschitz and Hurwitz constellations are shown over the signal-to-noise ratio (SNR) in dB. It is represented as

$$\frac{E_\mathrm{s}}{N_0} = \frac{\sigma_{a,\mathrm{d}}^2}{2\,\sigma_{n,\mathrm{d}}^2} = 2\,\frac{\sigma_a^2}{\sigma_n^2}\,, \tag{11}$$

where $\sigma_{n,\mathrm{d}}^2$ denotes the variance of real-valued zero-mean white Gaussian noise within each dimension and σ_n^2 the total variance of the 4D noise representation (with independent components). We clearly see that, in the high-SNR regime, one additional bit is achieved by applying the Hurwitz instead of the Lipschitz constellations. Besides, if a target rate close to a Lipschitz-constellation's maximum capacity is desired, e.g., 7 bit with $M = 256$, it is more advantageous to choose the related Hurwitz constellation ($M = 512$) since the power efficiency may be increased (here: SNR gain of about 0.5 dB). A similar gain may also be achieved by choosing the next (hypercube) Lipschitz constellation, however, with the cost of a tremendously higher processing and decoding effort ($M = 4096$ vs. $M = 512$ signal points).

III. CODED MODULATION AND LABELING

In this section, we discuss how channel coding can be efficiently combined with 4D modulation based on Hurwitz constellations. To that end, we will propose a straightforward two-stage dimension-wise (TD) CM scheme for Hurwitz constellations and discuss how to determine the rates of the component codes via the respective level capacities of the *equivalent channels* [16].

A. Two-Stage Dimension-Wise Coded Modulation

The proposed CM scheme is motivated by the construction of the Hurwitz constellation according to (8) via two interlaced Lipschitz subsets. In the following, the TD scheme is tailored jointly with the binary labeling rule $\mathcal{M} : \mathfrak{c} \mapsto a$, which maps the address[2] vector $\mathfrak{c} = [\mathfrak{c}_1, \ldots, \mathfrak{c}_b] \in \mathbb{F}_2^b$ to the data

[2]Symbols over the finite field $\mathbb{F}_2$ are typeset in Fraktur font, e.g., $\mathfrak{q}$, $\mathfrak{c}$.

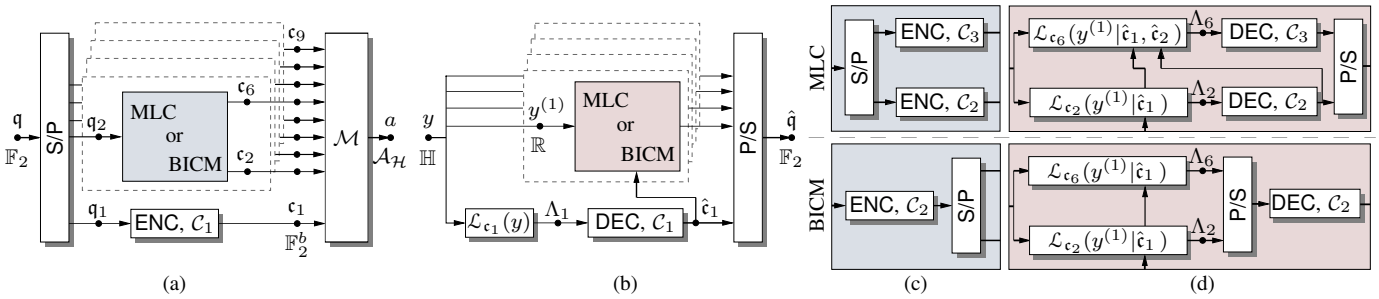

Fig. 4. Two-stage dimension-wise coded modulation transmitter (a) and receiver (b) for the four-dimensional Hurwitz constellation $\mathcal{A}_\mathcal{H}$ with $M = 512$ (i.e., $b = 9$ binary address levels $\mathfrak{c}_l$). The lowest bit level with $l = 1$ is protected by a low-rate code $\mathcal{C}_1$ and selects via the mapping $\mathcal{M}$ between the two subsets $\mathcal{A}_{\mathcal{L},1}$ and $\mathcal{A}_{\mathcal{L},2}$. A *dimension-wise* encoding is performed for the remaining levels, i.e., with four independent stages each following either the MLC or BICM (c,d) paradigm and resembling a one-dimensional ASK constellation $\mathcal{A}_\mathrm{d}$ with $M_\mathrm{d} = 4$ and $b_\mathrm{d} = 2$. On the receive side, the lowest level is decoded first requiring a single, *quaternion*-valued (4D) computational step on the receive symbol $y \in \mathbb{H}$. In a second step, each MLC/BICM stage operates independently on the *real*-valued component $y^{(\cdot)}$ and calculates its likelihood-values Λ_l after the lowest level is already decoded.

symbols $a \in \mathcal{A}_\mathcal{H}$. The corresponding inverse function is given by $\mathcal{M}^{-1}(a)$ and the function $\mathcal{M}_l^{-1}(a)$ returns the l^th bit of the symbol a, i.e., $\mathcal{M}_l^{-1}\left(\mathcal{M}([\mathfrak{c}_1, \ldots, \mathfrak{c}_b])\right) = \mathfrak{c}_l$.

The additional bit obtained by packing the two Lipschitz subsets selects between the two alphabets $\mathcal{A}_{\mathcal{L},1}$ and $\mathcal{A}_{\mathcal{L},2}$, i.e., it encodes the sign of the 4D offset $o_\mathcal{H}$. As a result, the set $\mathcal{A}_\mathcal{H}$ is divided into two subsets with the same minimum intra-subset Euclidean distance. This is in accordance with the well-known fact that in 4D set partitioning (SP) the intra-subset distance only increases in every second partitioning step [15].

Due to the isomorphism between $\mathcal{L}$ and the lattice $\mathbb{Z}^4$, each Lipschitz subset $\mathcal{A}_{\mathcal{L},1}$ and $\mathcal{A}_{\mathcal{L},2}$ can be interpreted as the Cartesian product of four M_d-ary ASK constellations $\mathcal{A}_\mathrm{d}$, cf. (6). As a consequence, Gray labeling in each dimension is possible, i.e., binary labels of adjacent ASK signal points only differ in a single bit position. This allows the straightforward (independent) application of BICM in each of the four dimensions. Alternatively, the four M_d-ary ASK constellations can be labeled according to the SP strategy and MLC with MSD can be applied per dimension. The mapping rule $\mathcal{M}_\mathrm{d} : \{\mathfrak{o}, \mathfrak{1}\}^{b_\mathrm{d}} \mapsto a^{(\chi)}$ maps tuples of b_d address bits to the χ^th quaternion component $a^{(\chi)} \in \mathcal{A}_\mathrm{d} \pm o_\mathrm{d}$ with $\chi = 1, 2, 3, 4$. The dimension-wise offset $o_\mathrm{d} = 1/4$ is either subtracted ($\mathfrak{c}_1 = \mathfrak{o}$) or added ($\mathfrak{c}_1 = \mathfrak{1}$) to the one-dimensional ASK constellation $\mathcal{A}_\mathrm{d}$ to select between $\mathcal{A}_{\mathcal{L},1}$ and $\mathcal{A}_{\mathcal{L},2}$.

The corresponding structure of the TD CM scheme is shown in Fig. 4 by example of $\mathcal{A}_\mathcal{H}$ with $M = 512$ signal points. The source bits $\mathfrak{q}$ are drawn from the binary field $\mathbb{F}_2$ and demultiplexed into 5 levels (i.e, one for the offset and four for each dimension) with block lengths equivalent to the dimension of the successive codes. The lowest source bit stream $\mathfrak{q}_1$ is encoded (ENC) with the error-correcting code $\mathcal{C}_1$ of rate $R_{\mathrm{c},1}$ to obtain the lowest address bit level $\mathfrak{c}_1$. The remaining four source bit streams $\mathfrak{q}_2, \ldots, \mathfrak{q}_5$ correspond to the four one-dimensional M_d-ary ASK constellations $\mathcal{A}_\mathrm{d}$. Each bit stream is encoded independently according to either the MLC or BICM paradigm[3], cf. Fig. 4 (c). Tuples of $b_\mathrm{d} = 2$ address bits[4] are then mapped ($\mathcal{M}_\mathrm{d}$) to an M_d-ary ASK constellation $\mathcal{A}_\mathrm{d}$ in each quaternion component $a^{(\chi)}$. SP labeling is used in conjunction with MLC and Gray labeling is used with BICM.

On the receive side, the lowest bit level is decoded first. The calculation of the log likelihood ratio (LLR) reads

$$\Lambda_1 = \mathcal{L}_{\mathfrak{c}_1}(y) = \log\left(\frac{\sum_{a\in\mathcal{A}_{\mathcal{L},1}} \mathsf{f}_{Y|A}(y \mid a)}{\sum_{a\in\mathcal{A}_{\mathcal{L},2}} \mathsf{f}_{Y|A}(y \mid a)}\right), \qquad (12)$$

where $\mathsf{f}_{Y|A}(y \mid a)$ is the QV conditional probability density function (PDF) which in case of the AWGN channel can be isomorphically expressed as a multivariate normal distribution with zero mean and separate components over the four dimensions. The decoder (DEC) delivers the estimate $\hat{\mathfrak{c}}_1$ which is used for the LLR calculation of the remaining levels $l \geq 2$ according to[5]

$$\Lambda_l = \mathcal{L}_{\mathfrak{c}_l}(y \mid \hat{\mathfrak{c}}_1) = \log\left(\frac{\sum_{a\in\mathcal{A}_l^{(\mathfrak{o})}} \mathsf{f}_{Y|A}(y \mid a)}{\sum_{a\in\mathcal{A}_l^{(\mathfrak{1})}} \mathsf{f}_{Y|A}(y \mid a)}\right), \qquad (13)$$

where $\mathcal{A}_l^{(\mathfrak{u})} = \{\, a \in \mathcal{A}_\mathcal{H} \mid \mathcal{M}_l^{-1}(a) = \mathfrak{u} \cap \mathcal{M}_1^{-1}(a) = \hat{\mathfrak{c}}_1 \,\}$ and $\mathfrak{u} \in \{\mathfrak{o}, \mathfrak{1}\}$. Due to independent components concerning signal modulation and noise the calculation can independently be performed on the distribution of each *real-valued* component $y^{(\chi)} \in \mathbb{R}$ via

$$\Lambda_l = \log\left(\frac{\sum_{a^{(\chi)}\in\mathcal{A}_{\mathrm{d},l}^{(\mathfrak{o})}} \mathsf{f}_{Y^{(\chi)}|A^{(\chi)}}(y^{(\chi)} \mid a^{(\chi)})}{\sum_{a^{(\chi)}\in\mathcal{A}_{\mathrm{d},l}^{(\mathfrak{1})}} \mathsf{f}_{Y^{(\chi)}|A^{(\chi)}}(y^{(\chi)} \mid a^{(\chi)})}\right), \qquad (14)$$

where $\mathcal{A}_{\mathrm{d},l}^{(\mathfrak{u})} = \{\, a^{(\chi)} \in \mathcal{A}_\mathrm{d} \pm \hat{o}_\mathrm{d} \mid \mathcal{M}_{\mathrm{d},l_\mathrm{d}}^{-1}(a^{(\chi)}) = \mathfrak{u} \,\}$ and l_d is the bit-level index in dimension χ. The offset $\hat{o}_\mathrm{d} = 1/4$ is either subtracted ($\hat{\mathfrak{c}}_1 = \mathfrak{o}$) from or added ($\hat{\mathfrak{c}}_1 = \mathfrak{1}$) to $\mathcal{A}_\mathrm{d}$. The PDF per component reduces to the one-dimensional normal distribution. After parallel decoding of the upper stages the estimate $\hat{\mathfrak{q}}$ is obtained. Hence, the calculation in (12) is in fact the only QV (i.e., 4D) operation in the TD receiver structure.

[3]For BICM—due to the successive demultiplexing—the codeword length $N_{\mathrm{c},2}$ in each stage must satisfy the relation $N_{\mathrm{c},2} = b_\mathrm{d} N_{\mathrm{c},1}$.

[4]The ordering of bit levels l is done here by the common convention in MLC that the bit-level capacities C_l increase with l, cf. the next subsection.

[5]Here, we only explicitly state the LLR calculation for the BICM receiver. The extension to MSD is, however, straightforward.

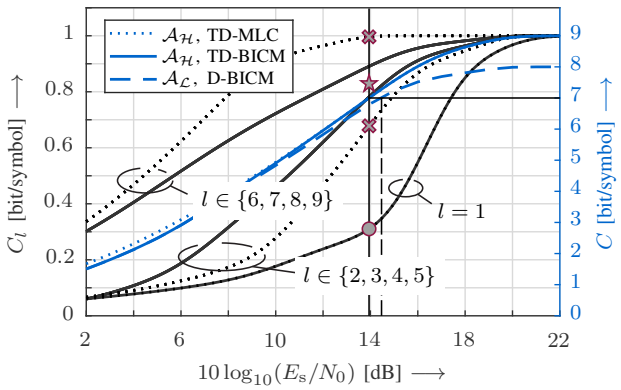

Fig. 5. Bit-level capacities C_l (left axis, black) for two-stage dimension-wise (TD) CM of the Hurwitz constellation $\mathcal{A}_\mathcal{H}$ with cardinality $M = 512$. MLC-MSD with SP (dotted) and BICM with Gray labeling (solid) is done in each one-dimensional 4ASK stage. Total capacities C (right axis, blue) are contrasted to dimension-wise (D)-BICM (dashed) using the Lipschitz constellation $\mathcal{A}_\mathcal{L}$ with $M = 256$. Target rate is $R_\mathrm{m} = 7$ bit/symbol; required SNR for TD-MLC/BICM is 13.94 dB and for D-BICM 14.41 dB. Markers indicate rates of the component codes for BICM (star) and MLC-MSD (crosses) while the lowest level has the same rate in both cases (circle).

B. Bit-Level Capacities

According to the chain rule of information theory [16], the coded-modulation capacity C^CM, shown in Fig. 3, can be decomposed into the mutual information (MI) of its *equivalent channels* l via[6]

$$C^\mathrm{CM} = I(Y;A) = I(Y;\mathfrak{C}_1,\mathfrak{C}_2,\ldots,\mathfrak{C}_b) \tag{15}$$

$$= \sum_{l=1}^{b} I(Y;\mathfrak{C}_l \mid \{1,2,\ldots,l-1\}) \,, \tag{16}$$

where $I(Y;\mathfrak{C}_l \mid \mathcal{I}) = I(Y;\mathfrak{C}_l \mid \{\mathfrak{C}_i \mid i \in \mathcal{I}\})$ gives the MI of the l^th bit given the bits in the set $\mathcal{I}$.

1) Lipschitz Constellation: Transferring this concept to the *dimension-wise* (D) decoding of Lipschitz constellations we have for MLC with MSD

$$C^\mathrm{D\text{–}MLC} = \sum_{n=1}^{4} \sum_{l\in\mathcal{D}_n} \underbrace{I(Y;\mathfrak{C}_l \mid \mathcal{I}\cap\mathcal{D}_n)}_{C_l^\mathrm{D\text{–}MLC}} = \sum_{l=1}^{b} C_l^\mathrm{D\text{–}MLC} \,, \tag{17}$$

where $\mathcal{D} = \{\mathcal{D}_1,\ldots,\mathcal{D}_4\}$ partitions the bit-level indices into sets of equal size $|\mathcal{D}_n| = b_\mathrm{d}$, i.e., each $\mathcal{D}_n$ collects all indices l that are associated with the n^th dimension. The coded-modulation capacity can be achieved if the code rate $R_{\mathrm{c},l}$ in level l equals the corresponding level capacity $C_l^\mathrm{D\text{–}MLC}$.

Similarly, the achievable rate of the BICM scheme can be given as the sum of the unconditioned MIs in each dimension, particularly

$$C^\mathrm{D\text{–}BICM} = \sum_{n=1}^{4} \sum_{l\in\mathcal{D}_n} \underbrace{I(Y;\mathfrak{C}_l)}_{C_l^\mathrm{D\text{–}BICM}} = \sum_{l=1}^{b} C_l^\mathrm{D\text{–}BICM} \,, \tag{18}$$

[6]In the context of information-theoretic elements, upper case letters indicate random variables, i.e., here A, Y, $\mathfrak{C}$.

TABLE II
CODE RATES FOR BICM AND MLC FOR DIMENSION-WISE (D) CM OVER LIPSCHITZ CONSTELLATIONS OR TWO-STAGE DIMENSION-WISE (TD) CM OVER HURWITZ CONSTELLATIONS ACCORDING TO THE CAPACITY RULE.

Scenario	$\mathcal{A}$	M	CM	$R_{\mathrm{c},1}$	$R_{\mathrm{c},2}$	$R_{\mathrm{c},3}$	$R_{\mathrm{c},4}$
$R_\mathrm{m} = 3.5$	$\mathcal{A}_\mathcal{L}$	16	–	–	0.8750	–	–
	$\mathcal{A}_\mathcal{H}$	32	TD	0.4909	0.7523	–	–
$R_\mathrm{m} = 7$	$\mathcal{A}_\mathcal{L}$	256	D-BICM	–	0.8750	–	–
			D-MLC	–	0.7507	0.9993	–
	$\mathcal{A}_\mathcal{H}$	512	TD-BICM	0.3103	0.8362	–	–
			TD-MLC		0.6763	0.9962	–
$R_\mathrm{m} = 10.5$	$\mathcal{A}_\mathcal{L}$	4096	D-BICM	–	0.8750	–	–
			D-MLC	–	0.6302	0.9948	1
	$\mathcal{A}_\mathcal{H}$	8192	TD-BICM	0.1734	0.8606	–	–
			TD-MLC		0.5911	0.9906	1

and the rate of the code $\mathcal{C}_n$ in each dimension is $R_{\mathrm{c},n} = \sum_{l\in\mathcal{D}_n} C_l^\mathrm{D\text{–}BICM}/b_\mathrm{d}$. The rate loss of (18) in comparison to (17) depends on the actual labeling. Gray labeling is known to perform well in the high-SNR regime.

2) Hurwitz Constellation: For the TD CM scheme applied to the Hurwitz constellations and under the premise that the lowest level encodes the 4D offset between both Lipschitz subsets, we have

$$C^\mathrm{TD\text{–}MLC} = \underbrace{I(Y;\mathfrak{C}_1)}_{C_1^\mathrm{TD\text{–}MLC}} + \sum_{n=1}^{4} \sum_{l\in\mathcal{D}_n} \underbrace{I(Y;\mathfrak{C}_l \mid (\mathcal{I}\cap\mathcal{D}_n)\cup\mathfrak{C}_1)}_{C_l^\mathrm{TD\text{–}MLC}} \,, \tag{19}$$

and similarly for BICM applied in each dimension

$$C^\mathrm{TD\text{–}BICM} = \underbrace{I(Y;\mathfrak{C}_1)}_{C_1^\mathrm{TD\text{–}BICM}} + \sum_{n=1}^{4} \sum_{l\in\mathcal{D}_n} \underbrace{I(Y;\mathfrak{C}_l \mid \mathfrak{C}_1)}_{C_l^\mathrm{TD\text{–}BICM}} \,. \tag{20}$$

The bit-level capacities for the Hurwitz-based approach with $M = 512$ are shown in Fig. 5. Both variants, i.e., TD-MLC with SP labeling and TD-BICM with Gray labeling, perform well at the target rate of $R_\mathrm{m} = 7$. The SNR gain w.r.t. the Lipschitz approach is about 0.5 dB. Note that TD-BICM approaches the performance of D-BICM in the low-SNR regime. The rates of the component codes can be deduced from the level capacities, cf. the markers in Fig. 5. In Table II, they are listed together with the ones for some other cardinalities and target rates.

IV. NUMERICAL RESULTS

Results obtained from numerical simulations are provided to complement our theoretical considerations. To this end, LDPC codes, particularly the subclass of irregular repeat-accumulate codes [9], have been constructed. The left (non-staircase) part of the parity-check matrix has randomly been chosen according to a predefined degree distribution (90 % of the columns have weight 3 and 10 % weight 4). In case of BICM, the block interleaver from the DVB-S2 standard [6, Sec. 5.3.3] is applied. Codes of length $N_\mathrm{c} = 64800$ as in DVB-S2 are employed; the average over 10^5 codewords is taken. The large code length ensures a sufficient number of redundancy bits even in case of very high code rates above 0.99. In practice,

these code rates can alternatively be realized by algebraic codes. To enable a fair comparison, we display the results over $E_\mathrm{b}/N_0 = 2\,\sigma_a^2/(R_\mathrm{m}\,\sigma_n^2)$, where R_m is the abovementioned CM rate, i.e., the number of information bits per QV symbol.

We consider the scenarios listed in Table II. The proposed Hurwitz-based CM strategies are contrasted to the Lipschitz-based ones, i.e., ASK per component, for the same CM rates. This is equivalent to neglecting the first stage in Fig. 4, i.e., the first level is unused and thus the 4D offset deactivated.

In Fig. 6, the bit error ratio (BER) is depicted. Given the first scenario with target rate $R_\mathrm{m} = 3.5$, we see an SNR gain of about 0.8 dB for the TD Hurwitz approach with $M = 32$ over the dimension-wise Lipschitz one with $M = 16$. This complies with the capacities (cf. Fig. 3); the required SNRs for this target rate are additionally shown. Due to the additional offset bit, the code rate in the second (i.e., 1D) stage $R_{\mathrm{c},2}$ can be lowered, improving the total performance. In this stage, CM is straightforward as only one bit per dimension is present.

Going over to the second scenario with $R_\mathrm{m} = 7$, both Hurwitz-based approaches ($M = 512$) enable a gain of about 0.5 dB over the respective Lipschitz ones ($M = 256$). This is again in accordance with the SNR gap w.r.t. the theoretical capacities. In both cases, MLC/MSD performs slightly better than BICM even though the total capacities in Fig. 5 are nearly the same. The degradation is present since the LLR-values are not completely independent—as actually required for BICM. However, the effect can be reduced if channel code and interleaver are optimized jointly. In contrast, the shorter code length in the first stage of the TD-BICM strategy does not play a significant role if the codes are already very long. Moreover, compared to the difference between $\mathcal{A}_\mathcal{L}$ and $\mathcal{A}_\mathcal{H}$, the gap induced by BICM instead of MLC is rather negligible.

Considering the third scenario with $R_\mathrm{m} = 10.5$, the same conclusions can be drawn. However, in that case, the SNR gain of the Hurwitz approaches reduces to roughly 0.2 dB in accordance with the theoretical capacities. Nevertheless, given the 8192-ary Hurwitz constellation, a gain of up to 1 dB over the 4096-ary Lipschitz one may be achieved if the target rate is increased to more than 11 bit, cf. Fig. 3.

V. Summary and Outlook

CM strategies for Hurwitz constellations, i.e., signal sets with elements drawn from the checkerboard lattice $\boldsymbol{D}_4$, have been discussed. More specifically, applying a 4D preprocessing in a first stage, state-of-the-art CM approaches like BICM or MLC can be applied in a second stage. The related capacities over the AWGN channel as well as numerical simulations have revealed that—when compared with conventional integer constellations like ASK or QAM ones—significant performance gains are possible. Thereby, the low-complexity approach of BICM performs almost as well as the MSD of MLC which is optimal in an information-theoretic point of view.

Future work deals with the implementation and assessment of the proposed coding strategies in fiber-optical systems where not only noise but also non-linearities may be present, or the application in dual-polarized wireless fading channels.

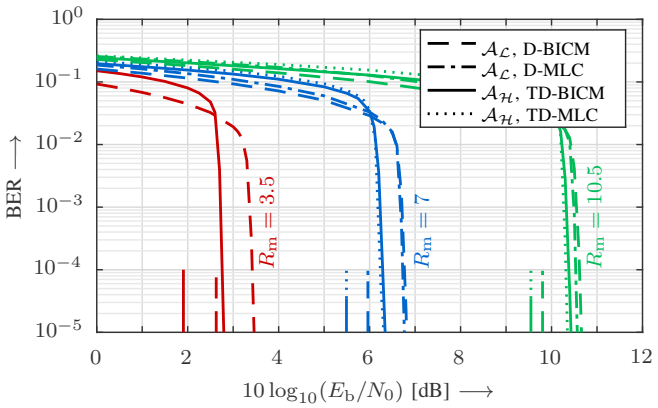

Fig. 6. Coded BER vs. E_b/N_0 in dB over the AWGN channel using dimension-wise (D) BICM or MLC for Lipschitz and two-stage dimension-wise (TD) CM for Hurwitz constellation. Parameters according to Table II. The respective CM capacity limits are shown as vertical lines.

References

[1] A. Alvarado, E. Agrell. Four-Dimensional Coded Modulation with Bit-wise Decoders for Future Optical Communications. *J. Lightwave Technol.*, vol. 33, no. 10, May 2015.

[2] G. Caire, G. Taricco, E. Biglieri. Bit-Interleaved Coded Modulation. *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 927–946, May 1998.

[3] J.H. Conway, N.J.A. Sloane. *Sphere Packings, Lattices and Groups*. Third Edition, Springer, 1999.

[4] J.H. Conway, D.A. Smith. *On Quaternions and Octonions: Their Geometry, Arithmetic, and Symmetry*. Taylor & Francis, 2003.

[5] Y.H. Cui, R.L. Li, H.Z. Fu. A Broadband Dual-Polarized Planar Antenna for 2G/3G/LTE Base Stations. *IEEE Trans. Antennas Propag.*, vol. 62, no. 9, pp. 4836–4840, Sep. 2014.

[6] European Telecommunications Standards Institute. DVB-S2. *ETSI Standard EN 302 307 V1.4.1: Digital Video Broadcasting (DVB)*, 2014.

[7] R.F.H. Fischer. *Precoding and Signal Shaping for Digital Transmission*. Wiley-IEEE Press, 2002.

[8] O.M. Isaeva, V.A. Sarytchev. Quaternion Presentations Polarization State. *2nd Topical Symp. on Combined Optical-Microwave Earth and Atmosphere Sensing*, pp. 195–196, Apr. 1995.

[9] H. Jin, A. Khandekar, R. McEliece. Irregular Repeat-Accumulate Codes. *2nd Int. Symp. on Turbo Codes and Related Topics*, Sep. 2000.

[10] M. Karlsson, E. Agrell. Multidimensional Optimized Optical Modulation Formats. *Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks*. John Wiley & Sons, 2015.

[11] W. Liu. Channel Equalization and Beamforming for Quaternion-Valued Wireless Communication Systems. *J. of the Franklin Institute*, vol. 354, no. 18, pp. 8721–8733, Dec. 2017.

[12] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P Flannery. *Numerical Recipes in C*. Second Edition, Cambridge University Press, 1987.

[13] S. Stern, R.F.H. Fischer. Quaternion-Valued Multi-User MIMO Transmission via Dual-Polarized Antennas and QLLL Reduction. *25th Int. Conf. on Telecommunications*, pp. 63–69, June 2018.

[14] G. Ungerböck. Channel Coding with Multilevel/Phase Signals. *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[15] G. Ungerböck. Trellis-Coded Modulation with Redundant Signal Sets Part II: State of the Art. *IEEE Commun. Mag.*, vol. 25, no. 2, pp. 12–21, Feb. 1987.

[16] U. Wachsmann, R.F.H. Fischer, J.B. Huber. Multilevel Codes: Theoretical Concepts and Practical Design Rules. *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361–1391, July 1999.

[17] F. Wäckerle, R.F.H. Fischer. Multistage Bit-Wise Receivers for 4D Modulation Formats in Optical Communications. *10th Int. ITG Conf. on Systems, Communications and Coding*, Feb. 2015.

[18] B.J. Wysocki, T.A. Wysocki, J. Seberry. Modeling Dual Polarization Wireless Fading Channels using Quaternions. *IST Workshop on Sensor Networks and Symp. on Trends in Communications*, pp. 68–71, June 2006.