Rate-Adaptive Coded Modulation for Fiber-Optic Communications

Lotfollah Beygi, Erik Agrell, *Senior Member, IEEE*, Joseph M. Kahn, *Fellow, IEEE*, and Magnus Karlsson, *Member, IEEE, Fellow, OSA*

***Abstract*—Rate-adaptive optical transceivers can play an important role in exploiting the available resources in dynamic optical networks, in which different links yield different signal qualities. We study rate-adaptive joint coding and modulation, often called coded modulation (CM), addressing non-dispersion-managed (non-DM) links, exploiting recent advances in channel modeling of these links. We introduce a four-dimensional CM scheme, which shows a better tradeoff between digital signal processing complexity and transparent reach than existing methods. We construct a rate-adaptive CM scheme combining a single low-density parity-check code with a family of three signal constellations and using probabilistic signal shaping. We evaluate the performance of the proposed CM scheme for single-channel transmission through long-haul non-DM fiber-optic systems with electronic chromatic-dispersion compensation. The numerical results demonstrate improvement of spectral efficiency over a wide range of transparent reaches, an improvement over 1 dB compared to existing methods.**

***Index Terms*—Fiber-optic communications, four-dimensional set partitioning, non-dispersion managed links, nonlinear channel model, probabilistic shaping, rate-adaptive coded modulation.**

I. Introduction

THE tremendous growth in the demand for high data rates in optical networks makes efficient use of available bandwidth indispensable [1]–[3]. The spectral efficiency, i.e., the number of information bits sent in each polarization per symbol period, of these channels can be improved by joint design of modern coding and advanced modulation formats, so-called coded modulation (CM). Forward error correction (FEC) [4] has already become a vital part of optical transport network standards and has evolved in several generations [5]. CM schemes are known [6]–[8] to be superior to conventional FEC techniques with independent FEC and modulation designs, in the sense of requiring less signal power for the same amount of redundancy and the same bit-error ratio (BER). Also, the joint design provides more freedom in the trade-off between digital signal processing complexity and transparent reach.

Manuscript received June 19, 2013; revised August 14, 2013; accepted August 20, 2013. Date of publication October 15, 2013; date of current version December 18, 2013. This work was published in part at Globecom 2012. The research was supported by the Swedish Foundation for Strategic Research (SSF) under grant RE07-0026 and the Swedish Governmental Agency for Innovation Systems (VINNOVA/CELTIC) under Grant 2010-01238. The calculations were performed in part on resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE.

L. Beygi was with Chalmers University of Technology, Sweden, and is now with Qamcom Research & Technology AB (e-mail: beygi@qamcom.se).

E. Agrell and M. Karlsson are with Chalmers University of Technology, Sweden (e-mail: agrell@chalmers.se; magnus.karlsson@chalmers.se).

J. M. Kahn is with Stanford University, USA (e-mail: jmk@ee.stanford.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2013.2285672

In general, fiber-optic channels are non-Gaussian due to the interplay of noise, dispersion, and nonlinearity. In contrast to additive white Gaussian noise (AWGN) channels, there is no standard framework for quantifying fundamental limits [9] as well as designing capacity-approaching schemes for such channels. Recently, however significant advances [10], [11] have been introduced in channel modeling of non-dispersion-managed (non-DM) fiber-optic links with sufficiently high symbol rates and sufficiently weak nonlinearity, often called pseudolinear regime, where the dispersion length is much smaller than the nonlinear length [1], [12]. The new Gaussian noise-like model introduced for pseudolinear regime makes it possible to adapt available CM techniques from AWGN channels to these channels.

Three main categories of CM schemes, namely trellis-coded modulation (TCM) [7], multilevel coded modulation (MLCM) [13], and bit-interleaved coded modulation (BICM) [14] have been studied for fiber-optic links operating in the pseudolinear [1], [12] and nonlinear [15, Ch. 4], [16] regimes. TCM was first proposed in [17] for fiber-optic systems with an 8-point cubic (three-dimensional) polarization-shift keying constellation. Later, the simplest 4- and 16-state TCM schemes were applied to 8-point phase shift keying (PSK) and differential PSK in [18]. The concatenation of TCM with different outer codes, Reed–Solomon (RS) and Bose–Chaudhuri–Hocquenghem codes, was studied in [19]. MLCM was proposed in [20] for a memoryless nonlinear fiber-optic channel with RS component codes. Two MLCM schemes were introduced in [21] and [22] with staircase codes and nonbinary low-density parity-check (LDPC) codes, respectively.

A comprehensive study of BICM was provided in [23] for fiber-optic communications with different modulation formats. Moreover, multidimensional BICM was studied in [24], [25]. Furthermore, BICM has been applied to polarization-multiplexed (PM) iterative polar modulation in [26]. CM schemes constructed by nonbinary component codes such as moderate-length nonbinary LDPC codes were proposed for fiber-optic communication in [27], [28].

A dynamic or heterogeneous structure of optically switched mesh networks demands adaptive transceivers to operate with different signal qualities. In other words, the required error protection provided by a CM scheme is varying with the uncoded link performance. Therefore, a CM scheme with the possibility of adapting the data transmission scheme to the channel state information (CSI), a so-called rate-adaptive scheme, is needed in these networks. To this end, a rate-adaptive CM scheme was proposed in [29] using three nonbinary LDPC codes with different rates together with three 4-, 8-, and 16-ary constellations. In [28],

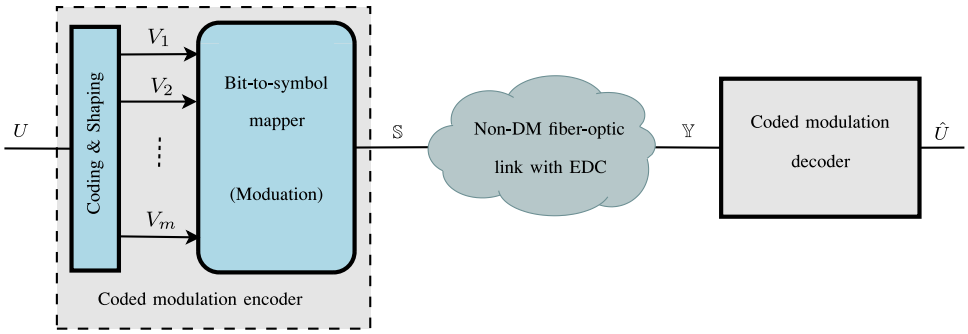

Fig. 1. A non-DM fiber link including a CM encoder at the transmitter, a non-DM fiber-optic channel with N spans, each consisting of an SMF and an EDFA, and the CM decoder and EDC at the receiver (U is the input information bit sequence and $\hat{U}$ is the decoded bit sequence).

the authors designed a rate-adaptive scheme with six nonbinary LDPC codes to provide a transmission bit rate between 100 Gb/s and 300 Gb/s in steps of 26.67 Gb/s at a fixed symbol rate. In a more practical scenario, a rate-adaptive scheme [30] is proposed exploiting six combinations of binary LDPC and RS codes together with three modulation formats. This scheme was capable of operating within 2.9 and 3.9 dB from the AWGN capacity in long and short non-DM single-channel fiber-optic links, respectively, showing 50% increase in transparent reach compared to the rate-adaptive system introduced in [31] with hard-decision FEC.

The aim of this paper is to introduce a low-complexity rate-adaptive CM scheme based on the recently introduced channel model for non-DM fiber-optic links [11], [32]. To this end, we introduce a new four-dimensional (4D) CM scheme to reduce the complexity of the nonbinary LDPC CM introduced in [28]. More precisely, we change the bit-to-symbol mapper using a new constellation labeling inspired by the polar coding approach [33] to reduce the order of the Galois field (GF), and hence the complexity, of the exploited nonbinary LDPC code.

A distinct contribution of the new CM scheme is in providing a flexible 4D structure, using a new 4D mapper and a probabilistic shaping method based on the shell-mapping algorithm [34]. This flexibility is used to obtain a rate-adaptive scheme with a single LDPC code. Simulation results are provided for a 4D nonbinary LDPC CM scheme with probabilistic shaping over a non-DM PM single-channel fiber-optic link. According to the numerical results, the proposed scheme can operate within 2.7 dB from the AWGN capacity for bit rates between 178 Gb/s and 343 Gb/s, showing 1 dB performance improvement compared to [30]. Finally, the performance of the proposed system is compared with other rate-adaptive schemes in the literature as well as AWGN bounds.

II. System Model

The system model including the transmitter, the non-DM fiber-optic link, and the receiver is depicted in Fig. 1.

A. Transmitter

As shown in this figure, the CM unit encodes the information bit sequence U to a sequence of 4D symbols, each of which is a pair of two-dimensional (2D) standard quadrature amplitude modulation (QAM) symbols in two polarizations. The symbol period is T. This encoding is represented by the matrix $\mathbb{S}$. Then, the coded symbols are sent through the non-DM fiber-optic link after performing pulse shaping. The code rate R is defined as the ratio of the spectral efficiencies of the coded system to the uncoded system. Moreover, the system redundancy overhead is defined as $\mathrm{OH} = 1/R - 1$. If we represent the 4D symbol by $(S_{\mathrm{xi}}, S_{\mathrm{xq}}, S_{\mathrm{yi}}, S_{\mathrm{yq}})$ at a specific time instant, its energy is computed as $E_s = S_{\mathrm{xi}}^2 + S_{\mathrm{xq}}^2 + S_{\mathrm{yi}}^2 + S_{\mathrm{yq}}^2$. The energies in polarization x and y are represented by $S_{\mathrm{xi}}^2 + S_{\mathrm{xq}}^2$ and $S_{\mathrm{yi}}^2 + S_{\mathrm{yq}}^2$, respectively. The energies of the four available dimensions are given by S_{xi}^2, S_{xq}^2, S_{yi}^2, and S_{yq}^2. It is assumed that the signals in polarizations x and y have the same average transmitted power P.

B. Non-Dispersion-Managed Fiber-Optic Link

The non-DM fiber-optic link has N spans, each consisting of a single mode fiber (SMF) and an erbium-doped fiber amplifier (EDFA) with single-wavelength data transmission. Electronic chromatic-dispersion compensation (EDC) is used at the receiver. Moreover, we assume that each EDFA compensates for the attenuation in each fiber span of length L and adds an amplified spontaneous emission (ASE) noise. This noise is modeled as a circular white complex Gaussian vector with variance $\sigma_{\mathrm{ASE}}^2 = GF_n h\nu/(2T)$ in each polarization [35, Eq. 8.1.15], where G is the required gain to compensate for the attenuation in a span, $F_n = 2n_{\mathrm{sp}}(1 - G^{-1})$ is the noise figure, in which n_{sp} is the spontaneous emission factor, and $h\nu$ is the photon energy. The Gaussian noise model introduced in [11] is used for the calculation of prior information of the CM decoder. According to this model, the received signal $\mathbb{Y}$ in a PM fiber-optic channel with EDC (as shown in Fig. 1) is represented by

$$\mathbb{Y} = \zeta \mathbb{S} + \mathbb{Z} \tag{1}$$

where $\mathbb{Z}$ represents the PM complex zero-mean circularly symmetric AWGN in each polarization, and

$$|\zeta|^2 \approx 1 - 3N^{1+\epsilon}\gamma^2\alpha^{-2}\tanh\left(\frac{\alpha}{4}L_{\mathrm{D}}\right)P^2 \tag{2}$$

in which γ is the nonlinear coefficient, α is the attenuation coefficient, β_2 is the dispersion coefficient, and $L_{\mathrm{D}} = T^2/(|\beta_2|)$. Here, the linear growth of the nonlinear noise with N,

introduced in [11], has been changed to $N^{1+\epsilon}$, where

$$\epsilon = \frac{3}{10}\log\left(1+\frac{6}{\alpha L \operatorname{asinh}\left(\frac{\pi^2}{2\alpha L_{\mathrm{D}}}\right)}\right).$$

As discussed in [10], this takes into account the partially correlated rather than entirely uncorrelated accumulation of nonlinear noises from different spans. The variance of the AWGN noise in each polarization is given by $\sigma^2_{Z_{\mathrm{x}}} = \mathbb{E}\{\mathbb{Z}^H\mathbb{Z}\}/2 = N\sigma^2_{\mathrm{ASE}} + \sigma^2_{\mathrm{NL}}$, where

$$\sigma^2_{\mathrm{NL}} = (1-|\zeta|^2)P \approx 3N^{1+\epsilon}\gamma^2\alpha^{-2}\tanh\left(\frac{\alpha}{4}L_{\mathrm{D}}\right)P^3 \quad (3)$$

where P is the transmitted power per polarization. In contrast to conventional AWGN channels with a noise variance independent from the input power, the variance of this equivalent AWGN is proportional to the cube of the transmit power.

C. Receiver

The coded modulation decoder extracts the sequence $\hat{U}$ from the received signal $\mathbb{Y}$ after EDC. According to (1), it needs the variance of the AWGN Z in each polarization to compute the a posteriori probabilities of the coded symbols[1]. We assume perfect clock and carrier synchronization, and perfect compensation of chromatic dispersion, polarization-mode dispersion and polarization rotation. The signal-to-noise ratio (SNR) is defined as $|\zeta|^2 P/\sigma^2_{Z_{\mathrm{x}}}$. The optimum power for each transparent reach is computed by $\partial\mathrm{SNR}/\partial P = 0$ analytically. Our numerical evaluation using the SSFM to determine the power that minimizes the SER is in good agreement with the analytical results.

There is a minimum SNR γ (in dB) to obtain a BER of 10^{-15} at the output of the CM decoder, which is usually computed by numerical simulations. The gap $\Delta\gamma$ between γ and the minimum SNR obtained using the Shannon formula for an AWGN channel with the spectral efficiency η is a useful measure to compare different CM schemes[2]. This gap, referred to as gap from AWGN capacity [30], can be expressed as $\Delta\gamma = \gamma - 10\log_{10}(2^\eta - 1)$ dB.

III. Design of the Coded Modulation Scheme

The transmitter of the coded modulation scheme can be represented as a mapper transforming the sequence of information bits to a sequence of symbols from a 4D constellation. As shown in Fig. 1, the scheme maps m bit sequences $V_1, V_2, \ldots, V_m$ to a 4D symbol sequence $\mathbb{S}$. The symbol sequence $\mathbb{S}$ is transmitted through a non-DM fiber-optic channel and received as the distorted symbol sequence $\mathbb{Y}$ after the EDC. The additive noise $\mathbb{Z}$ represents the added linear ASE noise and nonlinear noise-like interference. The channel capacity of a discrete-time memoryless channel is

$$\max_{p(V_1,\ldots,V_m)} I(V_1,\ldots,V_m;\mathbb{Y}),$$

[1]An alternative analytical result was introduced in [10] for the same model as $|\zeta|^2 \approx 1 - 32/(27\pi)N^{1+\epsilon}\gamma^2\alpha^{-1}L_{\mathrm{D}}^{-1}\operatorname{asinh}(\pi^2/(2\alpha L_{\mathrm{D}}))P^2$, which is in a good agreement with (2) numerically.

[2]This AWGN capacity, although popular as a benchmark, does not represent the capacity of the nonlinear fiber-optic channel [9], [36].

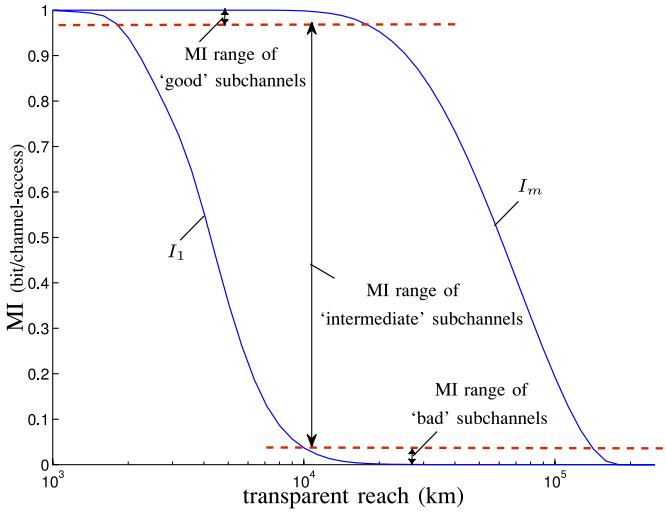

Fig. 2. A schematic example of bitwise conditional MIs together with the range of MIs for the 'good', 'intermediate', and 'bad' groups.

where $I(A;B)$ denotes the mutual information (MI) between A and B [37, Eq. (7.1)]. The maximum is taken over all possible input vector distributions $p(V_1,\ldots,V_m)$.

A. Probabilistic Signal Shaping

Probabilistic shaping changes the uniform distribution of the equivalent binary channels inputs $V_1,\ldots,V_m$ such that the distribution of the generated 1D symbols (elements of the matrix $\mathbb{S}$) from these bits better approximates a Gaussian distribution. In other words, instead of 4D symbols with uniformly distributed 1D elements, the 1D symbols (elements) close to the origin (with small amplitudes) are sent more often than 1D symbols far from the origin of the constellation (with large amplitudes). For a system without probabilistic shaping, the input bits are equally likely or uniformly distributed. Thus, we use the MI between the channel input and output $I = I(V_1,\ldots,V_m;\mathbb{Y})$ with uniformly-distributed V_is, i.e., $p(V_i) = 0.5, i = 1,\ldots,m$ for a system without probabilistic shaping.

B. Information-Theoretic Design Framework

The MI I can be decomposed [13] as $I = \sum_{i=1}^m I_i$, where $I_i = I(V_i;\mathbb{Y}|V_1,\ldots,V_{i-1})$ is the conditional MI of the subchannel i, provided that the transmitted bits of the subchannels $1,\ldots,i-1$ are given. Since according to (1), a non-DM fiber-optic link with EDC can be approximately modeled as a memoryless discrete-time AWGN channel [11], the MIs of the binary subchannels can be calculated by a numerical method [37, Ch. 9]. We exploit the efficient numerical method introduced in [38, Appendix] to evaluate the MIs using Gauss–Hermite quadratures for AWGN channels. This design framework, based on equivalent parallel binary subchannels I_is [13], can be used to analyze the design of different CM schemes. For a non-DM fiber optic link, as shown in Fig. 2, the subchannels may have different MIs for different transparent reaches (or SNRs). Hence, to approach the channel MI, an unequal error-protecting technique [13] needs to be applied over the m binary subchannels.

To this end, one may exploit several binary component codes, similar to [39], to devise a capacity-achieving CM scheme, or a single nonbinary code as in [28], which is the approach taken in this paper. We propose a new mapper to reduce the number of binary codes or the order of the corresponding GF for mitigating the complexity of the schemes introduced in [28], [39].

C. Polar Code Technique

If we consider a mapper with an arbitrary number of dimensions, the goal is to design this unit such that the equivalent binary subchannels are categorized in two groups, namely 'bad' and 'good' subchannels, with MIs 0 and 1, respectively, known as polar coding. Unfortunately, this ideal grouping can be reached only in an asymptotic case for very large dimensions, at the cost of a high decoding complexity. For a 4D mapper, we need to add a third group, namely 'intermediate' subchannels, with MIs between 0 and 1, as illustrated in Fig. 2. We consider two thresholds, shown with horizontal dashed lines in Fig. 2, to categorize the subchannels into the three groups. Our criterion in the design of the 4D mapper is to minimize the number of 'intermediate' subchannels. We notice here that for a given transparent reach, or SNR, the sum of the MIs of the subchannels yields the total MI between the channel input and output, independent from the mapper employed [13]. Thus, reducing the number of 'intermediate' subchannels leads to increasing the number of bad and good subchannels, which decreases the system complexity. In general, for a one-dimensional (1D) constellation with a large number of symbols, this discrete optimization of labeling to obtain the minimum number of 'intermediate' subchannels is very complicated, and even more so for 4D constellations. Therefore, we solely performed the optimization over the 1D labelings including the binary reflected Gray code, the natural binary code, and the folded binary code [38, Ch. 1]. We found that the natural labeling provides the minimum number of 'intermediate' subchannels in the region of interest, i.e., moderate SNRs, with PM 16- and 64-QAM constellations.

IV. Bit-to-Symbol Mapper

In this section, we introduce a channel-aware 4D bit-to-symbol mapper, often called constellation labeling, which reduces the number of 'intermediate' subchannels for non-DM fiber-optic links and provides a suitable structure to add probabilistic shaping. We consider the 4D constellation as the Cartesian product of two square QAM constellations or equivalently four 1D PAM constellations denoted by $\mathcal{A}\times\mathcal{A}\times\mathcal{A}\times\mathcal{A}$, where $\mathcal{A}$ is a PAM constellation.

A. Four-Dimensional Mapper Without Probabilistic Shaping

Without loss of generality, we proceed by describing the 4D mapper using the PM 16-QAM constellation for the sake of simplicity. For a 4-PAM constellation $\mathcal{A}=\{-3,-1,1,3\}$ with a natural labeling $\{00,01,10,11\}$, the set of four symbols is split into two sets $\mathcal{A}_0=\{-3,1\}$ and $\mathcal{A}_1=\{-1,3\}$, where the first bit (the least significant bit) of the binary labeling is 0 and 1 for symbols in $\mathcal{A}_0$ and $\mathcal{A}_1$, respectively. This partitioning of the signal set based on a specific bit in the binary labeling is called set partitioning with respect to the corresponding bit. Here, the 1D constellation $\mathcal{A}$ with a minimum Euclidean distance (MED) of d_0 is set partitioned into two subsets $\mathcal{A}_0$ and $\mathcal{A}_1$ with MEDs of $2d_0$.

As mentioned above, the 4D constellation is represented as the Cartesian product of four 1D 4-PAM constellations. We use this property and the method proposed in [40], which is a generalized version of the technique in [41] for an arbitrary dimension, to develop a 4D set partitioning based on the set partitioning of its constituent 4-PAM constellations. As seen in Table II(a), each step (one labeling bit) in the set partitioning of the 4-PAM constellation $\mathcal{A}$ is used in four steps (four labeling bits) of set partitioning of the 4D constellation $\mathcal{A}\times\mathcal{A}\times\mathcal{A}\times\mathcal{A}$. Using the partitioning of the 1D constellation, the 4D constellation is split into 16 subsets as shown in Table II(a). We define the inter-MED between 4D sets as the MED between the 4D symbols of the two sets. As seen in Table II(a), the bit v_1 partitions the set $\mathcal{A}\times\mathcal{A}\times\mathcal{A}\times\mathcal{A}$ into two sets with an inter-MED of d_0 (v_i, $0<i\leq m$, represents a bit from the sequence V_i at a specific time instant). Provided that v_1 is known, v_2 splits the corresponding subset into two sets with inter-MED of $\sqrt{2}d_0$. In an analogous way, set partitioning with the bits v_3 and v_4 result inter-MEDs of $\sqrt{2}d_0$ and $2d_0$, respectively, as illustrated in Fig. 3. Each of the subsets $\mathcal{A}_0$ and $\mathcal{A}_1$ can be further set partitioned into subsets $\mathcal{A}_{00}$, $\mathcal{A}_{01}$, $\mathcal{A}_{10}$, and $\mathcal{A}_{11}$ and so on (the same notation as in [13], [42]). The 4D set partitioning can be analogously continued for the labeling bits v_5, v_6, v_7, and v_8.

Since the first step of the set partitioning of the 4-PAM constellation yields two sets $\mathcal{A}_0$ and $\mathcal{A}_1$ with the same average energy of 5, the distributions of the first four bits of the 4D labeling given in Table II(a) have no effect on the average energy of the corresponding 4D constellation. In other words, after performing set partitioning by the labeling bits $v_1,\ldots,v_4$, we are left with 16 subsets with the same average 4D symbol energy. Each subset has 16 4D symbols, as shown in Table II(b) for the first subset. Moreover, as shown in Fig. 3, the inter-MED of the subsets resulting from the 4D set partitioning in each step for a given bit shows a nondecreasing behavior from the least to the most significant bit. Since the corresponding neighboring multiplicities can increase from v_1 to v_8, the inter-MED is not a good measure to categorize the channels into 'good', 'intermediate', and 'bad'. Hence, we use the MIs of binary subchannels for this purpose. As mentioned in the previous section the natural labeling provides the minimum number of 'intermediate' subchannels for the 4D mapper among the labelings introduced in Section III-C. This conclusion was observed by a numerical brute-force search.

Example: The MIs of binary subchannels $v_1,\ldots,v_8$ are plotted in Fig. 4 for different transparent reaches of a non-DM fiber-optic link. The AWGN model (1) with the system parameters given in Table I is used to compute these MIs numerically [38, Appendix]. For each transparent reach NL, the SNR is calculated as $|\zeta|^2P/\sigma^2_{Z_{\mathrm{x}}}$ with ζ and $\sigma^2_{Z_{\mathrm{x}}}$ given in Section II-B and the optimum transmit power P given in Section II-C. As seen in this figure, one may exploit the 4D mapper to categorize the

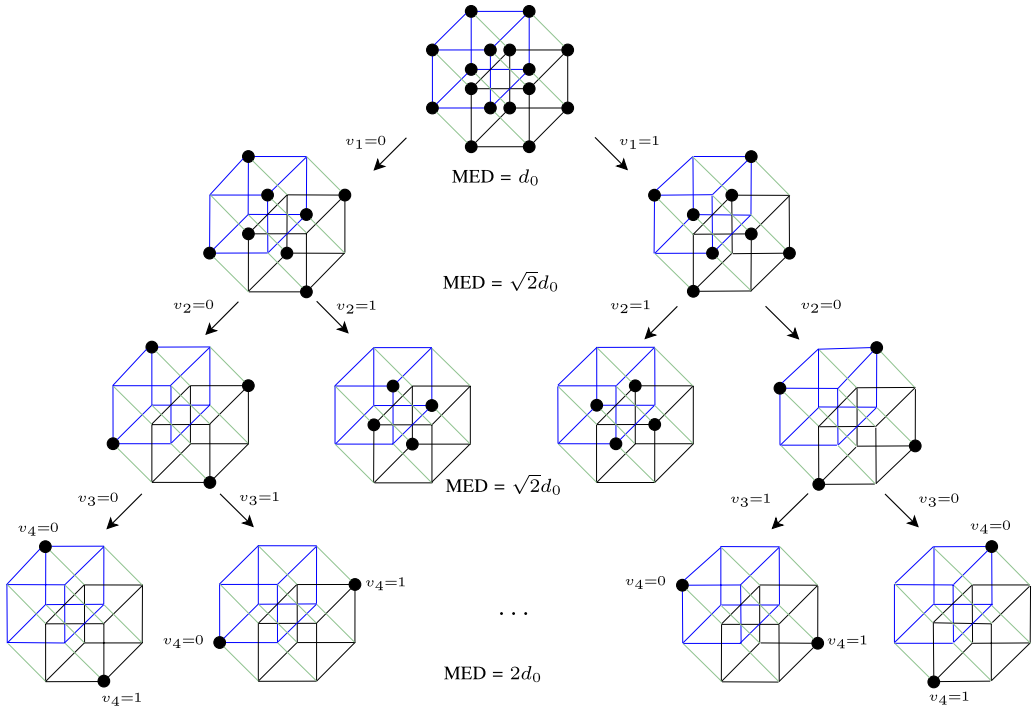

Fig. 3. First four steps of the 4D set partitioning of PM 16-QAM. Black circles represent the subsets given in Table II(a). The first four bits of the binary labeling for this constellation are represented by $v_4 v_3 v_2 v_1$.

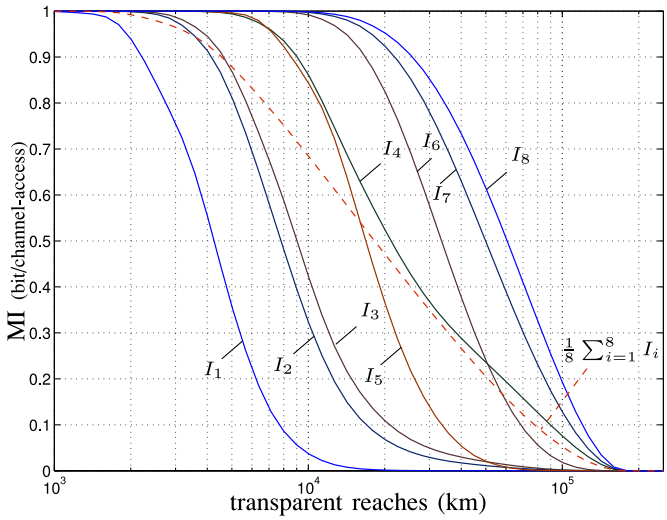

Fig. 4. The bitwise conditional MIs for a 4D mapper with PM 16-QAM. The fiber-optic link is implemented with the parameters given in Table I.

TABLE I
SYSTEM PARAMETER VALUES

Symbol rate R_s	32 Gbaud
Nonlinear coefficient γ	1.4 $\text{W}^{-1}\text{km}^{-1}$
Attenuation coefficient α	0.2 dB/km
Dispersion coefficient D	17 ps/nm/km
Optical center wavelength λ	1550 nm
EDFA noise figure F_n	5 dB
Span length L	80 km

binary subchannels into three types, namely 'bad', 'intermediate', and 'good' subchannels, at a specific transparent reach. As an example, for transparent reaches less than 20000 km in Fig. 4, we solely have 'intermediate' and 'good' subchannels, while for transparent reaches greater than 20000 km, the binary subchannel are categorized in 'intermediate' and 'bad' subchannels only.

B. Four-Dimensional Mapper With Probabilistic Shaping

Here, we describe how the 4D set partitioning, constructed based on a 1D constellation with the natural labeling, can be modified to devise this 4D mapper, which accounts for probabilistic shaping. Indeed, we manipulate the introduced 4D labeling in the previous section such that the average transmit power can be reduced by changing the distributions of zeros and ones at the input of the 'good' subchannels.

The second set partitioning of the 4-PAM constellation generates subsets with energies of 1 and 9. Therefore, the average energy of the 4D constellation can be reduced by manipulating the distributions of the second set of four bits. As will be discussed in Section V-B, we change the distribution of zeros and ones solely in 'good' subchannels from uniform to a nonuniform distribution, to obtain binary streams with higher prior probability for zeros. For the selected PM 16-QAM constellation, we can consider $v_5, \ldots, v_8$ as 'good' subchannels, which are nonuniformly shaped by the binary shaping algorithm introduced in Section V-B. Now, the labeling of 4D symbols inside each subset of $\mathcal{A} \times \mathcal{A} \times \mathcal{A} \times \mathcal{A}$ needs to be modified to account for signal shaping. Intuitively, the labels of the 4D symbols in each

TABLE II
THE 4D BINARY LABELING OF THE PM16-QAM CONSTELLATION ($\mathcal{A}_0 = \{-3, 1\}$ AND $\mathcal{A}_1 = \{-1, 3\}$). (a) THE 'BAD' AND 'INTERMEDIATE' SUBCHANNELS. (b) THE 'GOOD' SUBCHANNELS DEVISED TO ACCOUNT FOR PROBABILISTIC SHAPING IN SUBSET $\mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0$

(a)

v_1	v_2	v_3	v_4	subsets of $\mathcal{A} \times \mathcal{A} \times \mathcal{A} \times \mathcal{A}$
0	0	0	0	$\mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0$
0	0	0	1	$\mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_1$
0	0	1	0	$\mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_1$
0	0	1	1	$\mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_0 \times \mathcal{A}_0$
0	1	0	0	$\mathcal{A}_1 \times \mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_0$
0	1	0	1	$\mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_0 \times \mathcal{A}_1$
0	1	1	0	$\mathcal{A}_1 \times \mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_1$
0	1	1	1	$\mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_0$
1	0	0	0	$\mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_1$
1	0	0	1	$\mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_0$
1	0	1	0	$\mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_0$
1	0	1	1	$\mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_0 \times \mathcal{A}_1$
1	1	0	0	$\mathcal{A}_1 \times \mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_1$
1	1	0	1	$\mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_0 \times \mathcal{A}_0$
1	1	1	0	$\mathcal{A}_1 \times \mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0$
1	1	1	1	$\mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_1 \times \mathcal{A}_1$

(b)

v_5	v_6	v_7	v_8	weight	4D symbols in $\mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0$	E_s
0	0	0	0	0	(1, 1, 1, 1)	4
0	0	0	1	1	(1, 1, 1, −3)	12
0	0	1	0	1	(1, 1, −3, 1)	12
0	1	0	0	1	(1, −3, 1, 1)	12
1	0	0	0	1	(−3, 1, 1, 1)	12
0	0	1	1	2	(1, 1, −3, −3)	20
0	1	0	1	2	(1, −3, 1, −3)	20
0	1	1	0	2	(1, −3, −3, 1)	20
1	0	0	1	2	(−3, 1, 1, −3)	20
1	0	1	0	2	(−3, 1, −3, 1)	20
1	1	0	0	2	(−3, −3, 1, 1)	20
0	1	1	1	3	(1, −3, −3, −3)	28
1	0	1	1	3	(−3, 1, −3, −3)	28
1	1	0	1	3	(−3, −3, 1, −3)	28
1	1	1	0	3	(−3, −3, −3, 1)	28
1	1	1	1	4	(−3, −3, −3, −3)	36

subset can be obtained by assigning binary labels with small Hamming weights to 4D symbols with small symbol energies as outlined in the following steps.

1) Create a table of all 2^4 binary labels of length 4. Sort the table from the lowest to highest Hamming weights. Order labels with the same weight lexicographically, considering the bit on the right as the least significant bit.
2) Sort the 4D symbols of each subset from lowest to highest energy. Symbols with the same energy are ordered based on their energies in polarization x. If the symbols have the same 4D and 2D symbol energies, they are ordered based on their 1D energies, starting from S_{xi}^2 to S_{yq}^2 (as introduced in Section II).
3) Assign the table generated in 1) as the indices of the table generated in 2), in the same order.

This approach yields the labeling bits $v_5, \ldots, v_8$ ('good' subchannels) shown in Table II(b) for the first subset $\mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0$. We observe that labels with low weights correspond to symbol with low energies, as intended. This approach is illustrated in Fig. 5 for a subset consisting of 4D symbols with energies of 4, 12, 20, 28, and 36.

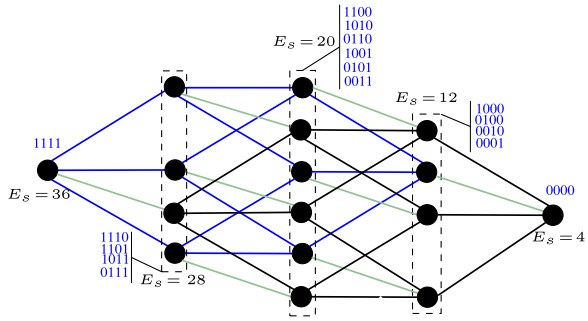

Fig. 5. The labeling bits $v_5\, v_6\, v_7\, v_8$ of 'good' subchannels shown in Table II(b) for the first subset $\mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0 \times \mathcal{A}_0$ of PM 16-QAM. This labeling is suitable for probabilistic signal shaping.

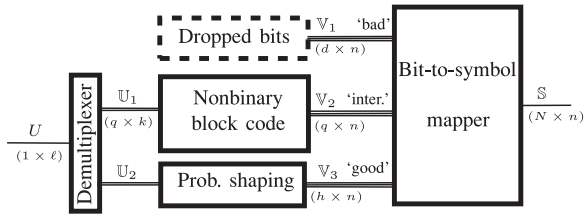

Fig. 6. The encoder for the proposed CM scheme, consisting of a block code, a 4D mapper, and a probabilistic shaping unit.

V. POLAR CODED MODULATION

Since the 4D mapper introduced in Section IV is inspired by the polar code concept, we call the devised CM scheme 'polar CM'. As seen in Fig. 6, the 4D bit-to-symbol mapper described in the previous section performs the role of a polar code [33] in our scheme. It divides the bit positions of the binary labeling of constellation symbols into 'bad' ($\mathbb{V}_1$), 'intermediate' ($\mathbb{V}_2$), and 'good' ($\mathbb{V}_3$) subchannels. Since the MI of 'bad' channels is close to zero, the input of these channels is frozen to zero or one. The matrix $\mathbb{V}_1$ is an all-zero matrix of size $d \times n$ representing the dropped bits, which do not carry any information. The sequence of information bits U of length ℓ is split into two groups. The first group $\mathbb{U}_1$ is a matrix of size $q \times k$, which is encoded by a block code. The block encoder generates a matrix $\mathbb{V}_2$ of size $q \times n$ by adding $(n-k)q$ bits or $n-k$ symbols from the GF of order 2^q, denoted by GF(2^q), as redundancy. Finally, the second group $\mathbb{U}_2$ consists of h row vectors with lengths r_i, $0 < i \leq h$. It is encoded by a probabilistic shaping unit to generate a matrix $\mathbb{V}_3$ of size $h \times n$. The 4D mapper unit maps the $(d+q+h) \times n$ binary matrix $\mathbb{V} = \{\mathbb{V}_1^T, \mathbb{V}_2^T, \mathbb{V}_3^T\}^T$ at the time instant i to the $4 \times n$ real matrix $\mathbb{S}$. Each column of $\mathbb{S}$ is a 4D symbol taken from a constellation of size 2^{d+q+h}.

A 2^q-ary nonbinary LDPC is used as the block code in the simulations. However, one may instead consider a binary LDPC code provided that iterative decoding is exploited between the log-likelihood ratio (LLR) calculator and LDPC decoder. The decoding of the binary subchannels with MI close to 1 is performed after the detection of 'intermediate' subchannels (i.e.,

coded by the nonbinary LDPC code). This is because the MI of these subchannels are conditioned on the input bits of the 'intermediate' subchannels, as discussed in Section III-C.

A. LDPC Coding and Decoding

Since the 'intermediate' subchannels are dependent (the detector of 'intermediate' subchannel i needs the transmitted bits on subchannels $1, \ldots, i-1$), they should be decoded jointly for optimal performance, provided that a nonbinary encoder is used in the encoder. Independent bit-wise decoding would give rise to performance degradation of the CM scheme. Hence, the nonbinary LDPC code performs on a vector of input bits of 'intermediate' subchannels. The main shortcoming of nonbinary LDPC codes is their decoding complexity, which increases with the order of the GF employed [43]. Since the error protection using the nonbinary LDPC code is solely performed over 'intermediate' subchannels, the proposed 4D mapper reduces the number of 'intermediate' subchannels and consequently the required order of the GF for the nonbinary LDPC code. Hence, the complexity of the CM scheme compared to [28] is reduced without performance degradation. The introduced channel model in Section II for the non-DM fiber-optic channel simplifies the computation of the a posteriori information required for the decoder of the LDPC code. In other words, according to this model the additive noise is white, hence the a posteriori probabilities of the received symbols are computed using the noise variance given in Section II and no iterative equalization [44, Ch. 7] is needed.

B. Probabilistic Shaping

Shell-mapping [34], [45], [46] and trellis shaping [47] are two well-known algorithms for performing probabilistic shaping over a constellation with uniformly distributed symbols. Since the shell-mapping algorithm and the LDPC code can be implemented jointly with lower complexity than trellis shaping, we exploit shell-mapping in our CM scheme. Indeed, the new 4D mapper introduced in Section IV allows us to exploit a binary shaping encoder over 'good' subchannels. Interestingly, the major part of the shaping gain is obtained by applying probabilistic shaping over the 'good' (uncoded) subchannels $\mathbb{V}_3$. This approach not only allows us to exploit a hard-decision decoder for the shaping unit but also avoids using concatenated (or joint) signal shaping and channel coding schemes, resulting a scheme with low complexity.

As shown in Fig. 6, the exploited binary shell-mapping is simply described as h parallel binary encoders, where the encoder i maps the input vector of length r_i to an output vector of length n. The shaping encoder i can be simply described as a look-up table consisting of 2^{r_i} binary vectors of length n. To construct this table, all 2^n binary vectors of length n are first sorted from lowest to highest Hamming weights. Vectors with the same weight are ordered lexicographically. Then, the last 2^{n-r_i} vectors are discarded. Therefore, it is readily seen that $r_i \leq n$ and we define $R_i = r_i/n$ as the shaping rate of the 'good' subchannel i.

To find the rates of the shaping encoders, we first use the approach introduced in [13] to force symbol S to get a (discrete) Gaussian distribution in each dimension. For example, the distribution of the in-phase component of the x polarization is considered as

$$P_{S_{\mathrm{xi}}}(a) = e^{-\lambda|a|^2} \left(\sum_{b \in \mathcal{A}} e^{-\lambda|b|^2} \right)^{-1}, \tag{4}$$

where the parameter λ controls the trade-off between the average energy of the 1D constellation and its entropy $H(\mathcal{A})$. Then, we maximize the MI $I(\mathbb{Y}; \mathbb{V})$ by a numerical optimization over the parameter λ. Now, we can compute the nonuniform distribution of 'good' subchannels and consequently the required rate of shaping encoders for the 'good' subchannels. The signal shaping reduces the required SNR or transmit power at the cost of shaping redundancy (or constellation expansion).

VI. Complexity Analysis

The encoder and decoder of the component codes together with the LLR calculation from soft (distorted) received symbols represent the main part of the DSP complexity of a CM scheme. For CM schemes with binary codes and a Gray-labeled constellation, the LLRs of the subchannels can be computed using the computationally efficient max-log approximation [48, Ch. 7]. Finding the closest among the 4D constellation symbols to the received (distorted) symbol requires approximately four times the computational complexity of finding the closest symbol in the constituent 1D constellation, neglecting the three additions which one may be needed to compute the 4D MED from four 1D MEDs [41]. Hence, the LLR vector for a 4D CM scheme can be computed with a very low complexity. This complexity analysis implies that one may compare the complexity of the receivers for CM schemes with different dimensions by taking into account solely the complexity of the component code decoders per dimension.

The complexity of LDPC and RS codes have been well-studied in the literature. The computational complexity required per iteration of the fast Fourier transform sum-product algorithm in decoding of a 2^q-ary regular nonbinary LDPC code designed over GF(2^q) is in the order of $O(J\rho q 2^q)$ [48, Ch. 14], where J and ρ are the number and weight of the rows of the parity-check matrix of the nonbinary LDPC code, respectively. For RS codes, the complexity is in the order of $O(q^2 2^q)$ [49]. Moreover, the number of iterations required for the convergence of LDPC iterative decoding also influence the complexity of the decoder of these codes.

VII. Numerical Results

In this section, we provide numerical results for the achievable rates of the proposed CM scheme as well as the AWGN bounds [10], [11] for different transparent reaches of a single-channel non-DM PM fiber-optic system. Although the LDPC codes with large girth employed are capable of having no error floor on the performance curve down to BERs around 10^{-12} as used in [50], we have evaluated the performance

TABLE III
THE RATES OF SHAPING ENCODERS FOR THE RATE-ADAPTIVE CM SCHEME IN TABLE V

Reach	R_1	R_2	R_3	R_4	R_5	R_6
17×80	0.97	0.95	0.95	0.94	0.94	0.91
31×80	0.88	0.84	0.84	0.83	0.75	-
44×80	0.70	0.70	0.70	0.70	-	-
55×80	0.83	0.75	0.63	-	-	-
90×80	0.83	-	-	-	-	-

The encoder output length is 64 bits.

of the proposed CM scheme with a small-girth LDPC code together with an RS outer code. Moreover, we choose the 'intermediate' subchannels such that the LDPC code rate satisfies $R_{\rm c} \leq I(Y; V_{d+1}, \ldots, V_{d+q}|V_1, \ldots, V_d)$ for different SNRs. The CM scheme is simulated with a (3,15)-regular quasi-cyclic nonbinary LDPC (1920,1546) over GF(2^6) to bring the uncoded BER down to BERs around 3×10^{-5}. Then, we meet BER $\leq 10^{-15}$ using a (1022,1004) RS outer code over GF(2^{10}). We use a shortening technique to match the length of the RS code and the LDPC input block length [51, Ch. 5]. The system parameters are given in Table I. The Fourier transform sum-product algorithm is used for the decoding of the nonbinary LDPC code.

The probabilistic shaping encoder output length is fixed to 64 bits. The optimized rates of the shaping encoders for different 'good' subchannels based on the method introduced in Section V-B are given in Table III. The total rate of the CM scheme is

$$R = \frac{R_{\rm RS}}{d+q+h}\left(\sum_{i=1}^{h} R_i + qR_{\rm c}\right) \tag{5}$$

where $R_{\rm RS}$ is the rate of outer RS code. The total FEC overheads of the systems for different transparent reaches are given in Tables IV and V. The split-step Fourier method (SSFM) [52, Eq. 2.4.10] is used to simulate a fiber-optic channel based on the Manakov equation with an adaptive segment length [53] of $\Delta_i = (\kappa L_{\rm N} L_{\rm D}^2)^{1/3}$, where i is the segment index, $\kappa = 10^{-4}$, and $L_{\rm N} = 1/(\gamma P_{i-1})$ is the nonlinear length of segment $i-1$ [52, p. 55] with the input power P_{i-1}. In the simulations, the receiver is assumed to have perfect knowledge of the polarization state. The optical filters are assumed to be unity gain with double-sided bandwidth equal to the sampling frequency used, which is usually greater than the signal bandwidth. We consider a root-raised-cosine pulse [54, p. 675] with an excess bandwidth of 0.2 and a truncation length of 16 symbols.

Fig. 7 shows the information bit rate per two polarizations versus the transparent reach for a non-DM fiber-optic link. The gap between the bit rate achieved using the proposed CM scheme without probabilistic shaping and the AWGN capacity is around 50 Gbps. As seen for transparent reaches smaller than 3000 km, probabilistic shaping decreases the gap to around 40 Gbps (20% reduction). In this figure, the AWGN capacity is plotted based on the Shannon formula, $2\log_2(1+\text{SNR})$, for an AWGN channel for different transparent reaches. To this end, the SNR was estimated in two ways: analytically using the variance $\sigma^2_{Z_{\rm x}}$ introduced for the Gaussian noise model and empirically by simulating at the output of the equalizer using a Monte-Carlo simulation in a similar way as in [30]. Fig. 7 indicates a good agreement between the analytical and empirical approaches.

The maximum uncoded symbol error rate of a hard-decision demodulator, denoted by $\text{SER}_{\rm Th}$, for obtaining an information BER of 10^{-15} at the output of the CM decoder for different transparent reaches are also given in Tables IV and V. As mentioned earlier, we use a semi-analytical approach to compute this FEC threshold. In fact, the required SNR to get a BER around 2×10^{-5} at the output of the proposed CM scheme is computed by Monte-Carlo simulations. Then, an interleaver is considered to make the errors independent at the input of the outer RS decoder. Finally, the BER at the output of the exploited outer RS decoder with no decoding failures is obtained by [55, Eq. (16)–(19)] assuming negligible probability of undetected errors. For example, a CM scheme with 17 spans (see first row of Table V) and 16.47% overhead has a SER FEC threshold ($\text{SER}_{\rm Th}$) of 0.08.

VIII. RATE ADAPTATION

To improve the utilization of optical networks with a dynamic or heterogeneous structure, we use the same approach as [30] to adapt the CM scheme to the CSI estimated by the receiver and reported back to the transmitter of the fiber-optic system. Similarly, we consider two choices for the CSI: (i) SNR, which is the symbol SNR estimated after polarization tracking and EDC, and (ii) the inner LDPC code performance, which is computed by a syndrome-based error estimator. The CSI is usually quantized to an integer value and sent to the transmitter using a reliable feedback channel.

The proposed 'polar CM' scheme provides a flexible structure to implement an adaptive-rate CM scheme with a single LDPC code. More precisely, the number of bits in the different ('good', 'intermediate', and 'bad') groups introduced in the 'polar CM' scheme are adjusted according to the CSI, as shown in Tables IV and V (these tables are discussed in Section VII). In contrast to [30], this approach exploits a simple circuitry to provide a rate-adaptive CM scheme. Since the mapper is solely a simple look-up table, the rate adaptation is straightforward to implement. As an example, in a single channel scenario for a short link with a transparent reach smaller than 1500 km (or SNR $>$ 18 dB), PM 64-QAM with no dropped bits is a suitable scheme, while for a long link with a transparent reach of 9000 km (SNR $\simeq$ 9 dB), PM 16-QAM with two dropped bits satisfies the required BER of 10^{-15}.

Interestingly, this rate adaptation can be seen as a proper selection of 4D constellations extracted from well-known lattices [6]. Tables IV and V indicate how the number of dropped bits (DB) and uncoded bits (UB) in the 4D mapper needs to be changed to support spectral efficiencies (SE) from 2.37 to 5.32 per polarization. As seen, for high SNRs, the PM 64-QAM constellation extracted from the 4D cubic lattice is used, while by decreasing the SNR, simply by changing the mapper, we obtain the 2048-ary 4D constellation extracted from the so-called D_4 lattice, which is the best 4D packing lattice [56], [57]. The number of coded bits, i.e., 'intermediate' subchannels, is fixed to six and these bits were coded by a nonbinary LDPC code.

TABLE IV
THE CM SCHEME WITHOUT PROBABILISTIC SHAPING FOR DIFFERENT TRANSPARENT REACHES

Reach (km)	bit rate (Gbps)		CM parameters					$\Delta\gamma$ (dB)	SNR (dB)	SER_{Th}
	CM	AWGN bound	Constellation	DB	UB	SE	OH%			
13×80	341	411	PM 64-QAM	0	6	5.32	12.78	3.35	19.26	0.07
20×80	309	368	PM 64-QAM	1	5	4.83	24.25	2.84	17.22	0.19
27×80	278	338	PM 64-QAM	2	4	4.34	38.32	2.96	15.80	0.29
39×80	246	301	PM 64-QAM	3	3	3.85	55.98	2.73	14.00	0.42
53×80	215	274	PM 16-QAM	0	2	3.36	19.21	2.98	12.64	0.08
81×80	183	232	PM 16-QAM	1	1	2.86	39.65	2.58	10.56	0.18
113×80	152	203	PM 16-QAM	2	0	2.37	68.56	2.80	9.01	0.29

System parameters are given in Table I, BER $\leq 10^{-15}$, RS(1022,1004) outer code over GF(2^{10}) and (1920,1546) LDPC inner code over GF(2^6).

TABLE V
THE CM SCHEME WITH PROBABILISTIC SHAPING FOR DIFFERENT TRANSPARENT REACHES

Reach (km)	bit rate (Gbps)		CM parameters					$\Delta\gamma$ (dB)	SNR (dB)	SER_{Th}
	CM	AWGN bound	Constellation	DB	UB	SE	OH%			
17×80	330	398	PM 64-QAM	0	6	5.15	16.47	2.67	18.68	0.08
31×80	282	326	PM 64-QAM	1	5	4.41	36.15	2.13	15.19	0.19
44×80	241	293	PM 64-QAM	2	4	3.76	59.81	2.60	13.57	0.29
55×80	221	267	PM 64-QAM	3	3	3.46	73.65	2.33	12.32	0.42
90×80	178	223	PM 16-QAM	1	1	2.78	43.89	2.41	10.10	0.18

System parameters are given in Table I, BER $\leq 10^{-15}$, RS(1022,1004) outer code over GF(2^{10}) and (1920,1546) LDPC inner code over GF(2^6).
The rates of shaping encoders are given in Table III.

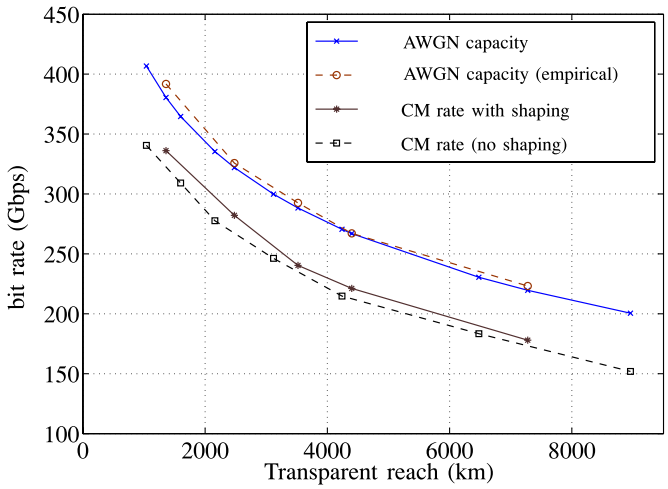

Fig. 7. Bit rates per two polarizations achieved by the CM scheme versus the transparent reach of a non-DM link with EDC, with and without probabilistic shaping.

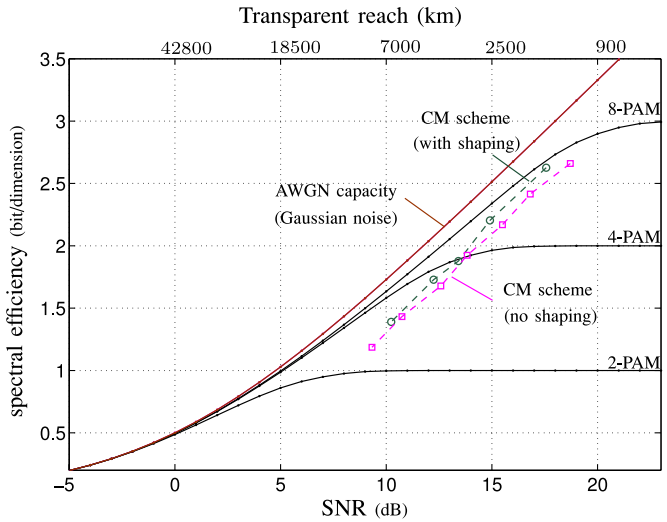

Fig. 8. The spectral efficiency per dimension versus the transparent reach and the SNR for a non-DM link with EDC. The CM scheme curves are based on the results given in Tables V and IV and the spectral efficiency for the Gaussian noise model is computed by $(1/2)\log_2(1+\text{SNR})$, where $\text{SNR} = |\zeta|^2 P/\sigma^2_{Z_x}$.

In Fig. 8, the AWGN capacity (spectral efficiency) per dimension is illustrated versus the transparent reach as well as the SNR (using the results provided in Sections II-B and II-C) for the rate-adaptive CM scheme over a non-DM fiber-optic link with the parameters given in Table I. The spectral efficiencies of the system with standard constellations, 4-PAM and 8-PAM, with and without probabilistic shaping, are also plotted in this figure. The results show that the rate-adaptive CM scheme using a single nonbinary code with probabilistic shaping can achieve $\Delta\gamma < 2.7$ dB for transparent reaches from 17×80 to 112×80 km.

The transmission bit rates for different transparent reaches are given in Table VI for four rate-adaptive schemes, including our proposed CM scheme. Since we have not considered the band-pass filtering of the signal propagated through many cascaded reconfigurable optical add–drop multiplexers (ROADMs) over long transparent reaches as considered in [30], we solely compare the gap from the AWGN capacity rather than the transparent reach. As shown in Table VI, the rate-adaptive CM scheme

TABLE VI
PERFORMANCE (Gb/s, km) OF RATE-ADAPTIVE SCHEMES WITH DIFFERENT BIT RATES PER TWO POLARIZATIONS AND TRANSPARENT REACHES

[29], BP	[30], EDC	[21], EDC	This work, EDC
(200,6400)	(50,11000)	(540,2000)	(150,9000)
(300,3900)	(100,6000)	(662,1000)	(220,4400)
(400,2100)	(200,1920)	(748,500)	(330,1440)

introduced in [29] can achieve information rates even higher than the AWGN capacity computed for a scheme with EDC. Moreover, the results introduced in [21] with distributed Raman amplification show significantly larger transparent reaches, however the gap from the AWGN capacity ($\Delta\gamma$) is larger than with our proposed scheme.

The results shown in Table V indicate larger than 1 dB performance improvement (or reduction in $\Delta\gamma$) compared to the previous results presented in [30]. However, according to the complexity analysis introduced in Section VI, the complexity of the exploited regular nonbinary LDPC code is slightly higher than irregular binary LDPC codes used in [30].

IX. Conclusion

The paper introduced a new 4D LDPC CM scheme for nonlinear fiber-optic channels. The design framework was supported by an information-theoretic approach. The proposed scheme exploits a 4D mapper, inspired by polar coding, to reduce the computational complexity of the nonbinary CM schemes without performance degradation. The 4D scheme provides a flexible structure to adapt the CM scheme for links with different signal qualities in a fiber-optic network.

In contrast to existing rate-adaptive schemes in the literature, the proposed scheme uses a single LDPC code rather than several binary or nonbinary component codes. Furthermore, exploiting a probabilistic shaping based on the shell-mapping algorithm, the system FEC threshold is improved with a reasonable increase in the system complexity.

Acknowledgment

The calculations were performed in part on resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE.

References

[1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[2] B. P. Smith and F. R. Kschischang, "Future prospects for FEC in fiber-optic communications," *IEEE J. Select Topics Quantuum Electron.*, vol. 16, no. 5, pp. 1245–1257, Sep./Oct. 2010.

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

[4] F. Chang, K. Onohara, and T. Mizuochi, "Forward error correction for 100 G transport networks," *IEEE Commun. Mag.*, vol. 48, no. 3, pp. S48–S55, Mar. 2010.

[5] ITU-T G.975.1, "Forward error correction for high bit-rate DWDM submarine systems," Feb. 2004.

[6] G. D. Forney, Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2384–2415, Oct. 1998.

[7] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[8] D. J. Costello, Jr. and G. D. Forney, Jr., "Channel coding: The road to channel capacity," *Proc. IEEE*, vol. 95, no. 6, pp. 1150–1177, Jun. 2007.

[9] E. Agrell. (2012). "On monotonic capacity–cost functions," [Online]. Available: http://arxiv.org/abs/1108.0391

[10] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[11] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3440–3450, Nov. 2012.

[12] A. Mecozzi, C. Clausen, and M. Shtaif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," *IEEE Photon. Technol. Lett.*, vol. 12, no. 4, pp. 392–394, Apr. 2000.

[13] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1361–1391, Jul. 1999.

[14] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873–884, May 1992.

[15] K.-P. Ho, *Phase-Modulated Optical Communication Systems*. New York, NY, USA: Springer-Verlag, 2005.

[16] L. Beygi, E. Agrell, M. Karlsson, and P. Johannisson, "Signal statistics in fiber-optical channels with polarization multiplexing and self-phase modulation," *J. Lightw. Technol.*, vol. 29, no. 16, pp. 2379–2386, Aug. 2011.

[17] S. Benedetto, G. Olmo, and P. Poggiolini, "Trellis coded polarization shift keying modulation for digital optical communications," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 1591–1602, Feb.–Apr. 1995.

[18] H. Zhao, E. Agrell, and M. Karlsson, "Trellis-coded modulation in PSK and DPSK communications," in *Proc. Eur. Conf. Exhibit. Optic. Commun.*, Sep. 2006.

[19] M. Magarini, R.-J. Essiambre, B. E. Basch, A. Ashikhmin, G. Kramer, and A. J. de Lind van Wijngaarden, "Concatenated coded modulation for optical communications systems," *IEEE Photon. Technol. Lett.*, vol. 22, no. 16, pp. 1244–1246, Aug. 2010.

[20] L. Beygi, E. Agrell, P. Johannisson, and M. Karlsson, "A novel multilevel coded modulation scheme for fiber optical channel with nonlinear phase noise," in *Proc. IEEE Global Commun. Conf.*, Dec. 2010.

[21] B. P. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," *J. Lightw. Technol.*, vol. 30, no. 13, pp. 2047–2053, Jul. 2012.

[22] L. Beygi, E. Agrell, and M. Karlsson, "Adaptive coded modulation for nonlinear fiber-optical channels," in *Proc. IEEE Global Commun. Conf.*, Anaheim, Dec. 2012, pp. 331–335.

[23] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3518–3530, Aug. 2009.

[24] H. Batshon, I. Djordjevic, L. Xu, and T. Wang, "Multidimensional LDPC-coded modulation for beyond 400 Gb/s per wavelength transmission," *IEEE Photon. Technol. Lett.*, vol. 21, no. 16, pp. 1139–1141, Aug. 2009.

[25] H. Bülow and E. Masalkina, "Coded modulation in optical communications," in *Proc. Optic. Fiber Commun. Conf.*, Mar. 2011.

[26] I. B. Djordjevic, H. G. Batshon, L. Xu, and T. Wang, "Coded polarization-multiplexed iterative polar modulation (PM-IPM) for beyond 400 Gb/s serial optical transmission," in *Proc. Optic. Fiber Commun. Conf.*, Mar. 2010, Paper OMK2.

[27] I. Djordjevic and B. Vasic, "Nonbinary LDPC codes for optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 17, no. 10, pp. 2224–2226, Oct. 2005.

[28] M. Arabaci, I. B. Djordjevic, L. Xu, and T. Wang, "Nonbinary LDPC-Coded modulation for high-speed optical fiber communication without bandwidth expansion," *IEEE Photon. J.*, vol. 4, no. 3, pp. 728–734, Jun. 2012.

[29] M. Arabaci, I. Djordjevic, T. Schmidt, R. Saunders, and R. Marcocci, "Rate-adaptive nonbinary-LDPC-coded modulation with backpropagation for long-haul optical transport networks," in *Proc. Int. Conf. Transp. Opt. Netw.*, Jul. 2010.

[30] G.-H. Gho and J. M. Kahn, "Rate-adaptive modulation and low-density parity-check coding for optical fiber transmission systems," *J. Opt. Commun. Netw.*, vol. 4, no. 10, pp. 760–768, Oct. 2012.

[31] G.-H. Gho and J. M. Kahn, "Rate-adaptive modulation and coding for optical fiber transmission systems," *J. Lightw. Technol.*, vol. 30, no. 12, pp. 1812–1818, Jun. 2012.

[32] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 742–744, Jun. 2011.

[33] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[34] R. Laroia, N. Farvardin, and S. A. Tretter, "On optimal shaping of multidimensional constellations," *IEEE Trans. Inform. Theory*, vol. 40, no. 4, pp. 1044–1056, Jul. 1994.

[35] G. P. Agrawal, *Fiber-Optic Communication Systems*, 3rd ed. New York, NY, USA: Wiley, 2002.
[36] E. Agrell and M. Karlsson, "Satellite Constellations: Towards the nonlinear channel capacity," in *Proc. IEEE Photon. Conf.*, Burlingame, CA, USA, Sep. 2012, pp. 316–317.
[37] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.
[38] A. Alvarado. (2010). "Towards fully optimized BICM transmissions," Ph.D. dissertation, Chalmers Univ. Technol., Göteborg, Sweden, [Online]. Available:http://publications.lib.chalmers.se/records/fulltext/130824.pdf
[39] J. Hou, P. Siegel, L. Milstein, and H. Pfister, "Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 49, no. 9, pp. 2141–2155, Sep. 2003.
[40] L. Beygi, E. Agrell, and M. Karlsson, "On the dimensionality of multilevel coded modulation in the high SNR regime," *IEEE Commun. Lett.*, vol. 14, no. 11, pp. 1056–1058, Nov. 2010.
[41] L. F. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Trans. Inform. Theory*, vol. 33, no. 4, pp. 483–501, Jul. 1987.
[42] R. Hormis and X. Wang, "Low-complexity coded-modulation for ISI-constrained channels," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1836–1846, Jun. 2009.
[43] D. Declercq and M. Fossorier, "Decoding algorithms for nonbinary LDPC codes over GF(q)," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 633–643, Apr. 2007.
[44] I. B. Djordjevic, W. Ryan, and B. Vasic, *Coding for Optical Channels*. New York, NY, USA: Springer-Verlag, 2010.
[45] F. R. Kschischang and S. Pasupathy, "Optimal shaping properties of the truncated polydisc," *IEEE Trans. Inform. Theory*, vol. 40, no. 3, pp. 892–903, May 1994.
[46] A. K. Khandani and P. Kabal, "Shaping multidimensional signal spaces—Part I: Optimum shaping, shell mapping," *IEEE Trans. Inform. Theory*, vol. 39, no. 6, pp. 1799–1808, Nov. 1993.
[47] G. D. Forney, Jr., "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 281–300, Mar. 1992.
[48] W. E. Ryan and S. Lin, *Channel Codes: Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[49] J. Justesen, "On the complexity of decoding Reed–Solomon codes," *IEEE Trans. Inform. Theory*, vol. 22, no. 2, pp. 237–238, Mar. 1976.
[50] I. Djordjevic, J. Xu, K. Abdel-Ghaffar, and S. Lin, "A class of low-density parity-check codes constructed based on Reed-Solomon codes with two information symbols," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 317–319, Jul. 2003.
[51] S. Lin and D. J. Costello, Jr., *Error Control Coding*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2004.
[52] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. New York, NY, USA: Academic, 2007.
[53] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *J. Lightw. Technol.*, vol. 28, no. 6, pp. 939–951, Mar. 2010.
[54] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York, NY, USA: McGraw-Hill, 2008.
[55] W. J. Ebel and W. H. Tranter, "The performance of Reed-Solomon codes on a bursty-noise channel," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 298–306, Feb.–Apr. 1995.
[56] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. New York, NY, USA: Springer-Verlag, 1998.
[57] M. Karlsson and E. Agrell, "Spectrally efficient four-dimensional modulation," in *Proc. Optic. Fiber Commun. Conf.*, Mar. 2012, Paper OTu2C.1.

Lotfollah Beygi received the Ph.D. degree from the Chalmers University of Technology, Göteborg, Sweden, in 2013. He is currently with Qamcom Research and Technology AB. His main research interests include coded modulation and digital signal processing for fiber-optical and wireless communications.

Erik Agrell (M'99–SM'02) received the Ph.D. degree in information theory in 1997 from the Chalmers University of Technology, Sweden. From 1997 to 1999, he was a Postdoctoral Researcher with the University of California, San Diego and the University of Illinois at Urbana-Champaign. In 1999, he joined the faculty of the Chalmers University of Technology, where he has been a Professor in communication systems since 2009. In 2010, he cofounded the Fiber-Optic Communications Research Center (FORCE) at Chalmers, where he leads the signals and systems research area. His research interests include information theory, coding theory, and digital communications, and his favorite applications are found in optical communications. Dr. Agrell served as the Publications Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY from 1999 to 2002 and is an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS since 2012. He received the 1990 John Ericsson Medal, the 2009 ITW Best Poster Award, the 2011 GlobeCom Best Paper Award, the 2013 CTW Best Poster Award, and the 2013 Chalmers Supervisor of the Year Award.

Joseph M. Kahn (M'90–SM'98–F'00) is a Professor of Electrical Engineering at Stanford University. His research addresses communication and imaging through optical fibers, including modulation, detection, signal processing and spatial multiplexing. He received the A.B. and Ph.D. degrees in Physics from U.C. Berkeley in 1981 and 1986, respectively. From 1987 to 1990, he was at AT&T Bell Laboratories, Crawford Hill Laboratory, in Holmdel, NJ. He was on the Electrical Engineering faculty at U.C. Berkeley from 1990 to 2003. In 2000, he co-founded StrataLight Communications, which was acquired by Opnext, Inc. in 2009. He received the National Science Foundation Presidential Young Investigator Award in 1991. In 2000, he was elected a Fellow of the IEEE.

Magnus Karlsson received the Ph.D degree from the Chalmers University of Technology, Gothenburg, Sweden, in 1994. Since 1995, he has been with the Photonics Laboratory, Chalmers University of Technology, first as an Assistant Professor and since 2003 as a Professor in photonics. He has authored or coauthored more than 240 scientific journal and conference contributions in the areas of nonlinear optics and fiber optic transmission, and co-founded the Chalmers fiber-optic communication research center FORCE in 2010. He has served in the technical committee for the Optical Fiber Communication Conference (OFC), and currently serves in the technical program committees for the European Conference of Optical Communication (ECOC) and the Asia Communications and Photonics Conference (ACP). He has been an Associate Editor for *Optics Express* since 2010. He received the CELTIC excellence award in 2011, the best paper award at GlobeCom 2011, and was appointed Fellow of the Optical Society of America in 2012.