

Article

Parallel Distribution Matcher Base on CCDM for Probabilistic Amplitude Shaping in Coherent Optical Fiber Communication

Yao Zhang ¹, Hongxiang Wang ^{1,*}, Yuefeng Ji ¹ and Yu Zhang ²

¹ State Key Laboratory of Information Photonics and Optical Communications, School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

² Academy of Broadcasting Science, NRTA, Beijing 102206, China

* Correspondence: wanghx@bupt.edu.cn

Abstract: As a typical high-order modulation format optimization technology, constellation probability shaping enhances generalized mutual information (GMI) by optimizing the probability distribution of each constellation point of the signal. It can improve the transmission capacity of the same order M Quadrature Amplitude Modulation (QAM) signal under the condition of limited average transmission power, and further narrow the gap with the Shannon limit capacity. The distribution matcher is a key part of constellation probability shaping since it not only ensures the one-to-one mapping of input and output sequences but also realizes the function of probability shaping. The constant composition distribution matcher (CCDM) structure is a widely utilized distribution matcher in the current probability shaping technology. Based on CCDM, a parallel distribution matcher scheme is proposed in this paper. It has a lower rate loss than CCDM for short output lengths (n is less than 100). Block lengths can be reduced by up to 30% with the same rate loss. When the GMI is the same as for the probability shaping (PS) 64QAM signal using CCDM, the OSNR required by the PS-64QAM signal using this scheme can be enhanced by 0.12dB, the block length can be reduced by 40%, and the transmission distance in a standard single-mode fiber can be slightly extended.



Citation: Zhang, Y.; Wang, H.; Ji, Y.; Zhang, Y. Parallel Distribution Matcher Base on CCDM for Probabilistic Amplitude Shaping in Coherent Optical Fiber Communication. *Photonics* **2022**, *9*, 604. <https://doi.org/10.3390/photonics9090604>

Received: 31 May 2022

Accepted: 23 August 2022

Published: 25 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: probabilistic amplitude shaping; constant composition distribution matcher; higher order modulation format signal; constellation shaping; LDPC coding

1. Introduction

The communication transmission rate and data flow in the network continue to expand as the Internet of Everything era dawns [1,2]. Optical fiber communication network is an essential supporting platform in the process of transmitting and exchanging information. Continuously improving its system capacity is the eternal goal of the development of optical communication. The high-order modulation format optical signal carries more bit information per symbol, boosting spectral efficiency and transmission capacity, and is the primary modulation method used in the current high-speed and large-capacity coherent optical communication system. As the probability of the appearance of each symbol in the standard M-QAM modulation format optical signal is the same, the system capacity is difficult to approach the Shannon limit. The gap between the standard M-QAM signal and the Shannon capacity in the additive white Gaussian noise channel is around 1.53 dB [3]. As a typical high-order modulation format optimization technology, constellation shaping enhances mutual information and generalized mutual information by optimizing the distribution of each constellation point of the signal. It includes geometric shaping (GS) and probability shaping (PS), which can improve the transmission capacity of the same order M-QAM signal under the condition of limited average transmission power and further narrow the gap with Shannon limit capacity.

In comparison with geometric shaping, the position of each constellation point of the signal remains the same after probability shaping, but the transmission probability of

different symbols varies. As a result, it is compatible with the existing modulation receiving system and digital signal processing technology, and can be flexibly combined with a range of multiplexing methods and channel coding technology, with low system complexity [3], this paper focuses on constellation probability shaping.

When the source obeys the continuous Gaussian distribution, the largest capacity can be achieved in the AWGN channel with power constraint P , according to Shannon [4]. For the probability shaping technology, the constellation points are dispersed discretely, making it difficult to determine what kind of distribution to use to approach the continuous Gaussian distribution. Kschischang demonstrated in 1993 that the Maxwell–Boltzmann (MB) distribution is the optimum for constellation probability shaping [5]. Subsequently, many probability shaping schemes, such as Gallager’s Scheme [6–8], Trellis Shaping [9], Concatenated Shaping [10], and Bootstrap Scheme [11], have been proposed to realize the shaping and optimization of QAM signals. P. Schulte and G. Böcherer innovatively proposed a probabilistic amplitude shaping (PAS) architecture in 2015 that combined a constant-composition distribution-matcher (CCDM) with forward error correction (FEC) encoding to implement PS [12]. Coding and shaping are decoupled in the PAS scheme thanks to a parallel transmitter design, which substantially simplifies the implementation of encoder and decoder and makes PS techniques practical. Many researchers have proven that probability shaping technology improves transmission distance, spectrum efficiency, and bit error performance in communication systems using the PAS scheme [13–18]. In the present high-order QAM modulation system, it is one of the preferred technologies.

The distribution matcher is a key component of PAS since it not only ensures the one-to-one mapping of input and output sequences but also realizes the function of probability shaping. The CCDM structure [19] is a widely utilized distribution matcher in the current probability shaping technology. It is based on arithmetic encoding, and the rate loss of CCDM can only tend to zero as the output symbol length approaches infinity. Hardware implementation is relatively complex in today’s high-speed optical communication system, and the complexity increases linearly with the length of symbols. Furthermore, arithmetic coding is a highly serial coding method. It divides the interval into intervals to reflect the input and output sequences. While the length of input and output sequences is long, the interval that must be divided grows, the borders between intervals blur, it is difficult to distinguish when mapping, and mistakes are common. It is, therefore, necessary to propose a distribution matcher with better performance even when the output symbol length is short.

A multiset-partition distribution matcher (MPDM) has been proposed to reduce the needed output lengths and can be considered as layered CCDM operations [20,21]. MPDM also adopts the idea of pairwise optimization; it needs to build a Huffman tree structure to index different complementary pairs and uses serial processing for the input data. The coding process is complex and error-prone. Many distributed matcher techniques with a parallel structure have been proposed to address the extremely serial coding approach of CCDM, such as bit-level distribution matcher [22,23], parallel-amplitude distribution matcher [24], and Hierarchical distribution matcher [25]. In addition, widely used enumerative sphere shaping (ESS)[26] and shell mapping (SM) are notable sphere shaping (SpSh) methods, which are different from CCDM. The SpSh considers amplitude sequences in a sphere. ESS orders these sequences lexicographically, while SM orders them based on their energy. Compared with ESS and SM, CCDM has similar latency and is significantly superior in terms of storage complexity [27].

In this work, we propose a parallel distribution matcher based on CCDM to improve the performance of CCDM at short block lengths. Contrary to the constant composition of the CCDM, this DM can output variable composition by using pairwise optimization and parallel structure. The proposed structure has a lower rate loss than CCDM for short output lengths (n less than 100), and the output block lengths can be lowered by up to 30% with the same rate loss. When the value of generalized mutual information is the same as it is when using CCDM, the OSNR tolerance of the PS-64QAM signal using the proposed

approach can be enhanced by 0.12 dB and the block length can be reduced by 40%. At the same time, the transmission distance in the standard single-mode fiber link becomes longer.

2. Fundamentals of Probabilistic Shaping

2.1. Constant Composition Distribution Matcher

The function of constant composition distribution is to convert binary bit sequences $B^k = b_1b_2 \dots b_k$ into amplitude sequences $A^n = a_1a_2 \dots a_n$ in a one-to-one, reversible mapping. The term “constant composition” has two meanings: the length of the input bit sequence and the output amplitude sequence are both fixed, as is the composition of the output amplitude sequence, which means the probability of each amplitude occurring is constant.

Specifically, the input sequence B^k is a uniformly distributed binary sequence of length k . If we define a set of amplitudes $\chi = \{\alpha_1, \alpha_2, \dots, \alpha_i\}$, then the output sequence A^n is of length n and the probability of each amplitude will be $P_A(\alpha_i) = \frac{n_{\alpha_i}}{n}$, indicating that the amplitude α_i will occur n_{α_i} times in this sequence with length n . Use $\Gamma_{P_A}^n$ to denote the set of all sequences of length n with a probability P_A of each amplitude. For example, for a 4ASK signal, $\chi = \{1, 3\}$; when $n = 4$, $P(1) = 3/4$, $P(3) = 1/4$, $\Gamma_{P_A}^n = \{1113, 1131, 1311, 3111\}$, the CCDDM needs to map the amplitude sequence in $\Gamma_{P_A}^n$ with a sequence of binary bits of length k , $k = \log_2 \Gamma_{P_A}^n = 2$, such as the following mapping: $00 \Leftrightarrow 1113, 01 \Leftrightarrow 1131, 10 \Leftrightarrow 1311, 11 \Leftrightarrow 3111$.

The key steps for constructing a CCDDM are as follows: specify the probability $P_A(\alpha_i)$ of the amplitude α_i and the length n of the output amplitude sequence; the number of input bits for CCDDM of a set $\Gamma_{P_A}^n$ is given by (1). $\lfloor \cdot \rfloor$ is rounded down. Since CCDDM is an invertible mapping, the size of the codebook C_{ccddm} should be 2^k ; select a subset of $\Gamma_{P_A}^n$ as the codebook and establish the mapping relation f_{ccddm} to finish one-to-one mapping of the input and output sequences: $\{0, 1\}^k \xrightarrow{f_{ccddm}} C_{ccddm} \subseteq \Gamma_{P_A}^n$.

$$k = \lfloor \log_2 \Gamma_{P_A}^n \rfloor = \left\lfloor \log_2 \frac{n!}{\prod_{i=1}^i n_{\alpha_i}!} \right\rfloor, n_{\alpha_i} = n \cdot P_A(\alpha_i), \tag{1}$$

Any finite length distribution matcher exists a rate loss problem. The rates of CCDDM are $R = k/n$ (bit/symbol), which means that each symbol in the output amplitude sequence carries k bits of information. Under the corresponding probability distribution, the information that each symbol is theoretically able to carry is $H(A)$, with a rate loss of (2).

$$R_{loss} = H(A) - R = H(A) - \frac{k}{n}, \tag{2}$$

Similar to the above 4ASK signal, $P(1) = 3/4$, $P(3) = 1/4$, $H(A) = 0.8113$ bit/symbol; when $n = 20$, $\lfloor \log_2 \Gamma_{P_A}^n \rfloor = 13$, $R_{loss} = 0.8113 - 0.65 = 0.1613$; when n reaches 10,000, $R = 8106/10,000 = 0.8106$ bit/symbol. So, CCDDM is an asymptotically optimal mapping scheme when the output sequence length $n \rightarrow \infty$, $R \rightarrow H(A)$. When n is relatively small, CCDDM has a large rate of loss.

2.2. Probabilistic Amplitude Shaping (PAS) Scheme

The data are modulated to the two polarizations of the optical carrier in the polarization multiplexing coherent optical transmitter, and the in-phase and quadrature components of the optical carrier are modulated separately in each polarization; so, the constellation of two-dimensional square m^2 -QAM signal can be expressed as the Cartesian product of the one-dimensional m -ASK signal. Use the variable X to represent the constellation point of the m -ASK signal, and use x to represent the specific value. Take $x = \pm 1, \pm 3, \dots, \pm(m - 1)$; when the gray mapping is utilized, each constellation point can be stated as the product of magnitude and sign and is independent—that is, $X = A * S$, where

$A = 1, 3, \dots, (m-1)$ and $S = 1$ or -1 . Since each constellation point is symmetrical about the origin, the probabilities of the symmetrical constellation points are equal, $P_X(x) = P_X(-x)$, $P_S(1) = P_S(-1) = 1/2$, the probability shaping of the two-dimensional square m^2 -QAM signal can be converted into the probability shaping of the one-dimensional m -ASK signal, which can be further converted into probabilistic shaping of the positive amplitude of the m -ASK signal, hence the name probabilistic amplitude shaping.

Figure 1a shows the principle diagram of the probability amplitude shaping using a constant composition distribution matcher and low-density parity check code. Taking the 64QAM modulation format signal as an example, the I channel or the Q channel is an 8ASK signal; the generating process of the bit sequence corresponding to the PS-64QAM signal is shown in Figure 1b.

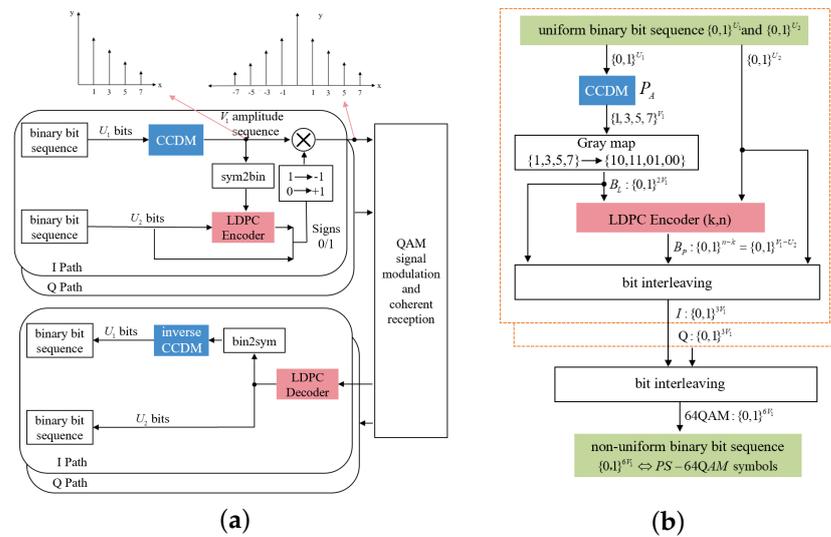


Figure 1. (a) The principle diagram of probability amplitude shaping using a CCDM and LDPC. (b) Bit sequence generation process of PS-64QAM signal.

A sequence of binary data of length U_1 at the input is fed into CCDM to generate a sequence of amplitudes of length V_1 , the amplitudes are picked from $\{1, 3, 5, 7\}$, and the number of occurrences of each amplitude are equal to the target probability. The rate of CCDM is $R_{DM} = U_1/V_1$. Then, apply the gray map on the V_1 amplitude sequences, where the mapping rule is $\{1, 3, 5, 7\} \Leftrightarrow \{10, 11, 01, 00\}$, the mapped binary bit sequence is B_L , and the length is $m * V_1$ ($m = 2$). Next, the binary information sequences of length U_2 at the input and B_L are concatenated and encoded together in LDPC to create check bits B_p of length $n - k = V_1 - U_2$, the encoding rate determined by (3). Finally, U_2 and B_p are utilized as the sign bit together and B_L is used as the amplitude bit, which are, respectively, assigned to different positions of the codeword by the interleaver. The interleaved bitstream can be used as data information for the I channel or the Q channel to modulate the QAM signal. At the receiving end, the input binary bit sequence can be recovered by performing the reverse process. In the following sections, the mutual information and bit error rate performance of PS-QAM signals will be analyzed based on this probabilistic amplitude shaping scheme.

$$R_c = k/n = (m * V_1 + U_2)/(V_1 - U_2 + k) = (m * V_1 + U_2)/(m + 1) * V_1, \quad (3)$$

3. Parallel Distribution Matching Based on CCDM

When the length of the output block (the length of the output amplitude sequence) n is small, CCDM suffers from rate loss. In this section, a paired optimization method is adopted to solve this issue, and a parallel distribution matcher structure based on CCDM is proposed.

3.1. Paired Optimization Principle

When the length n of CCDM output amplitude sequence is small, the probability distribution of each amplitude must be quantized—that is, $P_A \rightarrow P_{A'}$, to ensure that the number of occurrences of each amplitude is an integer and the quantization should try to satisfy that P_A and $P_{A'}$ are as close as possible.

When the output amplitude sequence of the CCDM is $A^n = a_1 a_2 \dots a_n$, the length is n , the amplitude setting is $\chi = \{\alpha_1, \alpha_2, \dots, \alpha_i\}$, each amplitude value in the sequence A^n is taken from χ , the number of occurrences of the amplitude α_i is n_{α_i} , and $n_{\alpha_i} = n \cdot P_{A'}(\alpha_i)$; the set $C = \{n_{\alpha_1}, n_{\alpha_2}, \dots, n_{\alpha_i}\}$ is called $P_{A'}$ composition with a probability distribution a , and the size of the set C can be determined using the permutation and combination formula as shown in (4), which is equivalent to $\Gamma_{P_A}^n$.

$$M(C) = \frac{n!}{n_{\alpha_1}! n_{\alpha_2}! \dots n_{\alpha_i}!} \tag{4}$$

$C_{typ} = \{n_{\alpha_1}, n_{\alpha_2}, \dots, n_{\alpha_i}\}$ represents a typical composition of CCDM output, and obeys the probability distribution $P_{A'}(\alpha_i)$. The pairwise optimization means that the single output amplitude sequence of the CCDM does not follow the probability distribution $P_{A'}(\alpha_i)$ —that is, $C \neq C_{typ}$. However, the average composition of many output amplitude sequences is equal to C_{typ} , with all output amplitude sequences obeying the target probability $P_{A'}(\alpha_i)$ as a whole. The purpose of pairwise optimization is to find components that satisfy (5) [20], where l represents the N_{comp} possible composition of the output amplitude sequence and c_l represents the number of occurrences of the sequence whose composition is C_l in the output sequence; its value does not exceed $M(C_l)$ at most. Here, we only consider the case of pairwise complementarity, not the case of three or more compositions of complementarity.

$$\frac{\sum_l^{N_{comp}} c_l \cdot C_l}{\sum_l^{N_{comp}} c_l} = C_{typ} \tag{5}$$

For example, the length of the output amplitude sequence of CCDM is $n = 10$; the amplitude is taken from (a_1, a_2, a_3, a_4) , assuming the probability distribution after quantization is $P_{A'} = (0.4, 0.3, 0.2, 0.1)$; $C_{typ} = (4, 3, 2, 1)$; and the information entropy $H(A) = 1.85$ bit/symbol. If only C_{typ} is considered in the output composition of CCDM, $M(C_{typ}) = 12,600$, it can map $k = \lceil \log_2(C_{typ}) \rceil = 13$ input bits, the rate is 1.3 bit/symbol, and the rate loss is 0.55 bit. When considering the composition of $C_1 = (4, 2, 3, 1)$ and $C_2 = (4, 4, 1, 1)$, although a single output amplitude sequence is not equal to $P_{A'}$, but $M(C_1) = M(C_2) = 6300$, the output amplitude sequences of C_1 and C_2 can be complementary, and the overall obeys the probability distribution $P_{A'}$; thus, the total output amplitude sequence is equal to $12,600 + 6300 + 6300 = 25,200$, which can map 14-bit input binary sequence. Compared with only considering C_{typ} , the rate is increased by 0.1 bit/symbol [20].

3.2. Implementation of Parallel Distribution Matcher Base on CCDM

Based on the above principle, the benefit of pairwise optimization can be exploited to improve the rate penalty of CCDM. For simplicity, only two complementary components are considered here satisfying (6).

$$C_l + \overline{C_l} = 2C_{typ} \tag{6}$$

All complementary pairs can be found by exhaustive methods, but when n is relatively large, there are many complementary pairs and it is easy to miss. A summary formula is given below, which can be used to locate all complementary pairs $\{C_l, C_2\}$ that meet (6)

on a regular basis. When $n = 10$, $C_{typ} = (4, 3, 2, 1)$, $2C_{typ} = (8, 6, 4, 2)$, consider the following polynomial:

$$\begin{aligned}
 &x_1 + x_2 + x_3 + x_4 = 10 \quad 0 \leq x_1 \leq 8, \quad 0 \leq x_2 \leq 6, \quad 0 \leq x_3 \leq 4, \quad 0 \leq x_4 \leq 2 \\
 &\underbrace{(x^0 + x^1 + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8)}_{x_1} * \underbrace{(x^0 + x^1 + x^2 + x^3 + x^4 + x^5 + x^6)}_{x_2} \\
 &* \underbrace{(x^0 + x^1 + x^2 + x^3 + x^4)}_{x_3} * \underbrace{(x^0 + x^1 + x^2)}_{x_4} \tag{7}
 \end{aligned}$$

For example, x^8 in x_1 , x^1 in x_2 , x^1 in x_3 , and x^0 in x_4 is equal to $x^{8+1+1+0} = x^{10}$ when multiplied; so, the composition $C_1 = (8, 1, 1, 0)$ is one of the cases, and the corresponding complementary pair is $\bar{C}_1 = (0, 5, 3, 2)$. This polynomial multiplication enables us to find regularly all the possible compositions and the corresponding complementary pairs.

Figure 2 shows the block diagram of the parallel distribution matcher. The length of the input binary sequence is $k = p_l + k_a + k_b$, which the first p_l bits utilized to select complementary pairs, and the remaining $k_a + k_b$ bits are serial-to-parallel conversion as needed. In the upper $CCDM_1$, the bit sequence of length k_a is mapped to the 2^{k_a} amplitude sequences in the set $M(C_l)$ by arithmetic coding, and in the lower $CCDM_2$, the bit sequence of length k_b is also mapped to the 2^{k_b} amplitude sequences in the set $M(\bar{C}_l)$. After the parallel-to-serial conversion of the generated two-paths amplitude sequence, the component composition is $2C_{typ}$, which obeys the target probability distribution $P_{A'}$.

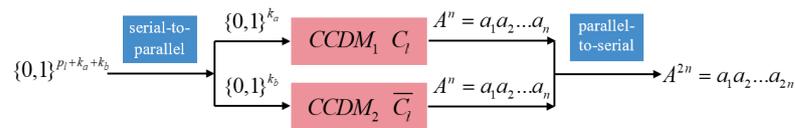


Figure 2. The block diagram of the parallel distribution matcher.

The specific implementation steps are as follows:

(1) Set the length of the output blocks of the two CCDMs to be n , quantify the target probability distribution $P_A \rightarrow P_{A'}$ to ensure that the number of occurrences of each amplitude is an integer, and obtain C_{typ} and $k_{typ} = \lfloor \log_2 M(C_{typ}) \rfloor$.

(2) According to (7), find all qualified complementary pairs $C_l, M(\bar{C}_l)$, denoted as N_{pair} ; calculate $M(C_l)$ and $M(\bar{C}_l)$, and obtain the mappable input bit sequence length $k_a + k_b = \lfloor \log_2 M(C_l) \rfloor + \lfloor \log_2 M(\bar{C}_l) \rfloor$; sort N_{pair} complementary pairs from large to small according to the size of $k_a + k_b$ —the complementary pairs of $k_a + k_b < k_{typ}$ are preferentially discarded because such complementary pairs will reduce the overall rate loss, and there are N'_{pair} types of complementary pairs left after discarding; to obtain $p_l = \log_2 N'_{pair}$, a binary sequence of length p_l and 2^{p_l} complementary pairs maintain a one-to-one mapping relationship.

(3) At the input end, for a string of input binary bit sequences, first determine which complementary pair is used by the two CCDMs according to the size of the first p_l bits. Then, make a serial-parallel conversion after obtaining $k_a + k_b$; the corresponding amplitude sequences are obtained through arithmetic coding, respectively. Finally, the two amplitude sequences are converted in parallel to serial, which can then be employed in the subsequent probability amplitude shaping system to achieve the probability shaping of QAM signals.

(4) The appropriate complementary pairs are determined by counting the frequency of occurrence of each amplitude in n amplitude sequences while the receiver performs inverse mapping, so that p_l can be obtained. k_a and k_b can be obtained by CCDM inverse mapping, thereby restoring the original data information.

4. Simulation Results

4.1. Rate Loss

This section numerically studies the rate loss of parallel distribution matcher based on CCDM. Take the probability distribution after quantization as $P_A = \{0.4, 0.3, 0.2, 0.1\}$ as an example to demonstrate the calculation process of the rate loss. When $n = 10$, $C_{typ} = (4, 3, 2, 1)$, there are 49 complementary pairs (including C_{typ}) that satisfy the (5), the specific parameters of each complementary pair are shown in Table 1. Only the first 32 kinds of complementary pairs sorted by $k_a + k_b$ are listed in Table 1, because there are 49 kinds of complementary pairs, which are equivalent to 97 cases (C_{typ} is equivalent to only one case) and can index 6-bit binary data at most.

Table 1. The specific parameters of each complementary pair when $n = 10$, $C_{typ} = (4,3,2,1)$.

Pair Number	C_l	\bar{C}_l	$k_a + k_b$	Pair Number	C_l	\bar{C}_l	$k_a + k_b$
1	(4, 3, 2, 1)	(4, 3, 2, 1)	13 + 13	17	(3, 2, 4, 1)	(5, 4, 0, 1)	13 + 10
2	(4, 3, 3, 0)	(4, 3, 1, 2)	12 + 13	18	(2, 4, 2, 1)	(6, 1, 2, 1)	12 + 11
3	(4, 4, 1, 1)	(4, 2, 3, 1)	12 + 13	19	(2, 3, 3, 2)	(6, 3, 1, 0)	14 + 9
4	(3, 4, 2, 1)	(5, 2, 2, 1)	13 + 12	20	(4, 4, 0, 2)	(4, 2, 4, 0)	11 + 11
5	(3, 3, 2, 2)	(5, 3, 2, 0)	14 + 11	21	(4, 5, 0, 1)	(4, 1, 4, 1)	10 + 12
6	(3, 3, 3, 1)	(5, 3, 1, 1)	13 + 12	22	(2, 5, 3, 0)	(6, 1, 1, 2)	11 + 11
7	(4, 4, 2, 0)	(4, 2, 2, 2)	11 + 13	23	(2, 3, 4, 1)	(6, 3, 0, 1)	19 + 9
8	(3, 5, 1, 1)	(5, 1, 3, 1)	12 + 12	24	(3, 5, 0, 2)	(5, 1, 4, 0)	11 + 10
9	(3, 4, 1, 2)	(5, 2, 3, 0)	13 + 11	25	(2, 5, 1, 2)	(6, 1, 3, 0)	12 + 9
10	(3, 4, 3, 0)	(5, 2, 1, 2)	12 + 12	25	(2, 4, 4, 0)	(6, 2, 0, 2)	11 + 10
11	(3, 2, 3, 2)	(5, 4, 1, 0)	14 + 10	27	(2, 2, 4, 2)	(6, 4, 0, 0)	14 + 7
12	(2, 4, 2, 2)	(6, 2, 2, 0)	14 + 10	28	(1, 5, 3, 1)	(7, 1, 1, 1)	12 + 9
13	(2, 4, 3, 1)	(6, 2, 1, 1)	13 + 11	29	(1, 4, 3, 2)	(7, 2, 1, 0)	13 + 8
14	(4, 5, 1, 0)	(4, 1, 3, 2)	10 + 13	30	(3, 6, 1, 0)	(5, 0, 3, 2)	9 + 11
15	(3, 5, 2, 0)	(5, 1, 2, 2)	11 + 12	31	(3, 1, 4, 2)	(5, 5, 0, 0)	13 + 7
16	(3, 3, 4, 0)	(5, 3, 0, 2)	12 + 11	32	(2, 6, 1, 1)	(6, 0, 3, 1)	11 + 9

When the first 32 complementary pairs are chosen, $p_l = 6$, the average total number of indexable bits is $6 \times 64 + 26 \times 10 + 24 \times 14 + 23 \times 12 + 22 \times 20 + 20 \times 7 = 1852$, the rate $R = 1852/64/20 = 1.447$ bit/symbol, and the rate of CCDM is 1.3 bit/symbol, increasing by 0.147 bit/symbol. When the first 17 complementary pairs are selected, $p_l = 5$ and the average total number of indexable bits is $5 \times 32 + 26 + 25 \times 10 + 24 \times 14 + 23 \times 7 = 933$, resulting in a rate of $R = 933/32/20 = 1.459$ bit/symbol, an increase of 0.159 bit/symbol. When only the first three complementary pairs are selected, $p_l = 2$, the average total number of indexable bits is $2 \times 4 + 26 + 25 \times 3 = 109$, and the rate $R = 109/4 \times 20 = 1.36$ bit/symbol, increasing by 0.06 bit/symbol. In the worst case, the complementary pair may not be chosen, and the parallel CCDM is equivalent to a single CCDM.

Figure 3 shows rate loss over block length for the probability distribution as $P_A = \{0.4, 0.3, 0.2, 0.1\}$. The rates under different a are shown in Table 2. We observe that the parallel distribution matcher achieves a lower rate loss than CCDM for all block lengths n. When the block length $n = 100$, the rate of CCDM is 1.75 bit/symbol and the rate loss is 0.1, while when the parallel distribution matcher is $n = 70$, the maximum rate can reach 1.76 bit/symbol ($p_l = 8$) and the minimum is 1.74 bit/symbol ($p_l = 4$), saving 30% on block length n. When the block length of the parallel distribution matcher is $n = 100$, the maximum rate loss is 0.0847 bit/symbol and the minimum rate loss is 0.0693 bit/symbol, both better than CCDM’s rate loss of 0.1 bit/symbol.

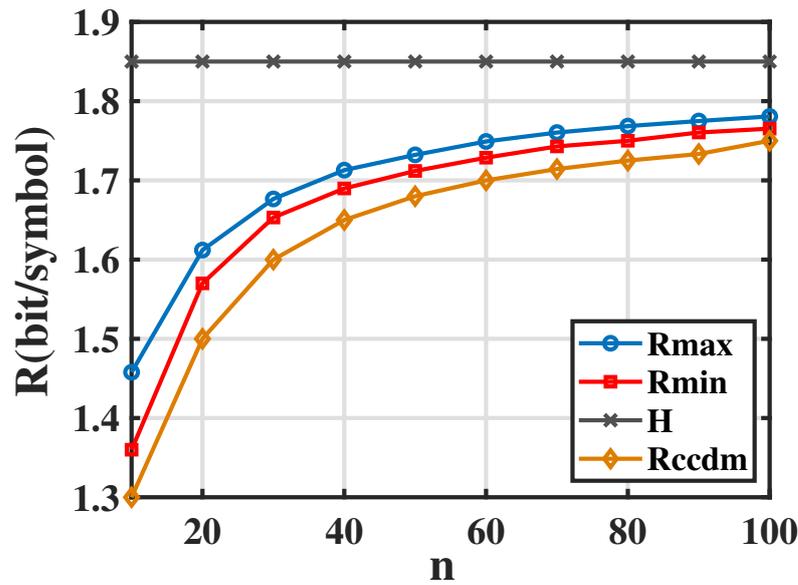


Figure 3. Rate over block length n for CCDM and parallel distribution matcher (Rmax and Rmin); the probability distribution is $P_A = \{0.4, 0.3, 0.2, 0.1\}$.

Table 2. Rate of CCDM and parallel DM under different block lengths n.

n	Parallel DM p_l = 8	Parallel DM p_l = 7	Parallel DM p_l = 6	Parallel DM p_l = 5	Parallel DM p_l = 4	CCDM
10	–	–	1.447	1.458	1.4375	1.3
20	–	–	1.612	1.606	1.589	1.5
30	–	1.677	1.673	1.665	1.653	1.6
40	1.713	1.711	1.706	1.699	1.69	1.65
50	1.732	1.73	1.725	1.72	1.712	1.68
60	1.7491	1.7462	1.7425	1.7352	1.7287	1.7
70	1.7604	1.7569	1.7521	1.747	1.743	1.7143
80	1.7685	1.7649	1.7611	1.7563	1.75	1.725
90	1.7749	1.7721	1.7679	1.7636	1.7604	1.733
100	1.7807	1.7767	1.7734	1.7702	1.7653	1.75

4.2. Generalized Mutual Information and BER Results

The Generalized mutual information (GMI) can be used to express the achievable information rate for bit-metric decoding. In this section, the generalized mutual information and bit error rate performance of the probability shaping signal using parallel distributed matcher are analyzed using the PAS scheme. The advantages of the proposed structure are demonstrated when compared with the PS-QAM signal using CCDM.

The GMI for bit-interleaved coded modulation is represented by (8) (Equation (13) of [3] and Equation (7) of [28]). In Monte Carlo simulations of N samples, GMI can be expressed as (9) (Equation (8) of [28]). Considering the rate loss of DM, the achievable information rate is shown in (10) (Equation (15) of [24]).

$$GMI(X; Y) = H(X) - \sum_{j=1}^m H(B_j|Y) \tag{8}$$

$$GMI \approx \frac{1}{N} \sum_{k=1}^N [-\log_2 P_X(x_k)] - \frac{1}{N} \sum_{k=1}^N \sum_{i=1}^m [\log_2(1 + e^{(-1)^{b_{k,i}} \Lambda_{k,i}})] \tag{9}$$

$$\Lambda_{k,i} = \log \frac{\sum_{x \in \mathcal{X}_1^i} p_{Y|X}(y_k|x) P_X(x)}{\sum_{x \in \mathcal{X}_0^i} p_{Y|X}(y_k|x) P_X(x)} = \log \frac{p_{Y|B_i}(y_k|1)}{p_{Y|B_i}(y_k|0)} + \log \frac{P_{B_i}(1)}{P_{B_i}(0)}$$

$$GMI_{DM} = GMI - R_{loss} \tag{10}$$

We consider a dual-polarization PS-64QAM modulation system with $P(1, 3, 5, 7) = (0.4, 0.3, 0.2, 0.1)$, the baud rate per polarization is 28 GBaud. The simulation block diagram and parameters in VPI are shown in Figure 4. In Figure 5, GMI in bits per symbol is shown over OSNR in dB for PS-64QAM. Under the same OSNR, the greater n ($n = 30, 50, 100$) is, the smaller the rate loss is, and the larger the generalized mutual information is for the PS-64QAM signal utilizing the parallel distribution matcher. The OSNR required by the parallel distribution matcher is less than that using CCDM when n is 100 and the GMI is the same. When GMI is 4 bit/symbol, compared with the uniformly distributed 64QAM signal, the OSNR of the PS-64QAM signal using CCDM is reduced by 0.35 dB, and the OSNR of the PS-64QAM signal using the parallel distribution matcher is reduced by 0.47 dB, which is a 0.12 dB improvement compared with CCDM. When the block length of CCDM increases from $n = 100$ to $n = 50,000$, the OSNR only reduces by 0.33 dB; so, the parallel structure has a relatively large improvement of 0.12 dB.

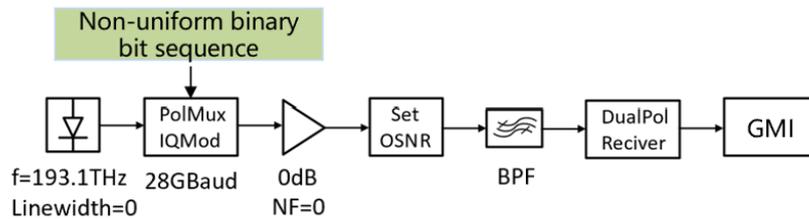


Figure 4. The GMI simulation block diagram and parameters in VPI.

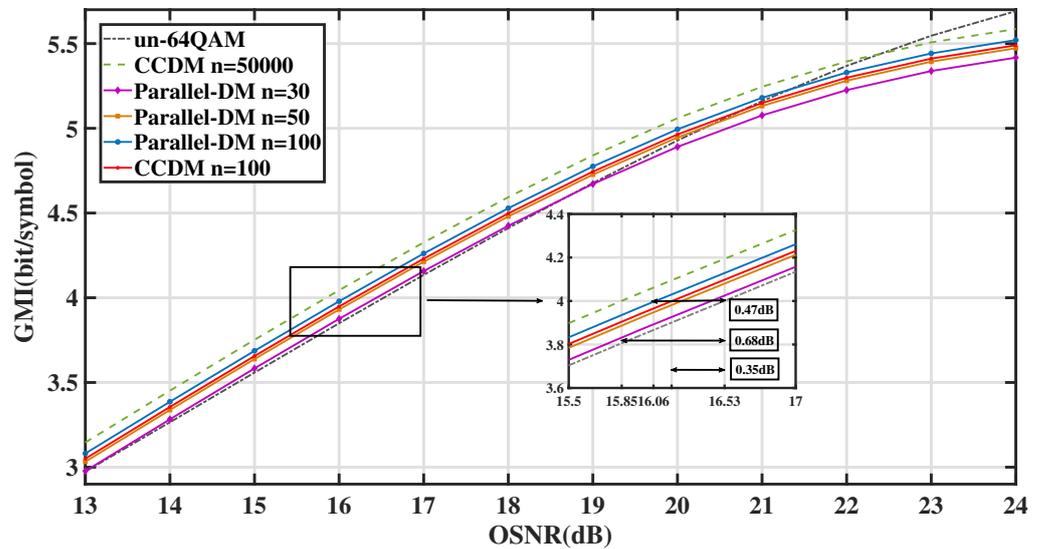


Figure 5. GMI in bit/symbol over OSNR in dB for bit-metric decoding and PS-64QAM. The inset zooms into the region around GMI = 4 bit/symbol, where the parallel distribution matcher of length $n = 100$ is 0.47 dB more power-efficient than uniform 64QAM, and is 0.12 dB more power-efficient than PS-64QAM using CCDM.

Figure 6 demonstrates the maximum and minimum GMI values for PS-64QAM signals using this parallel distribution matcher at different block lengths when the OSNR is 15, 16, and 17 dB. When OSNR = 15 dB, GMI = 3.655 bit/symbol; OSNR = 16 dB, GMI = 3.948 bit/symbol; and OSNR = 17 dB, GMI = 4.23 bit/symbol. The needed block length n of the parallel distribution matcher is only 60, although the required block length of the CCDM is 100; so, the block length can be saved by 40%. The matcher can attain the same performance as CCDM while using fewer blocks.

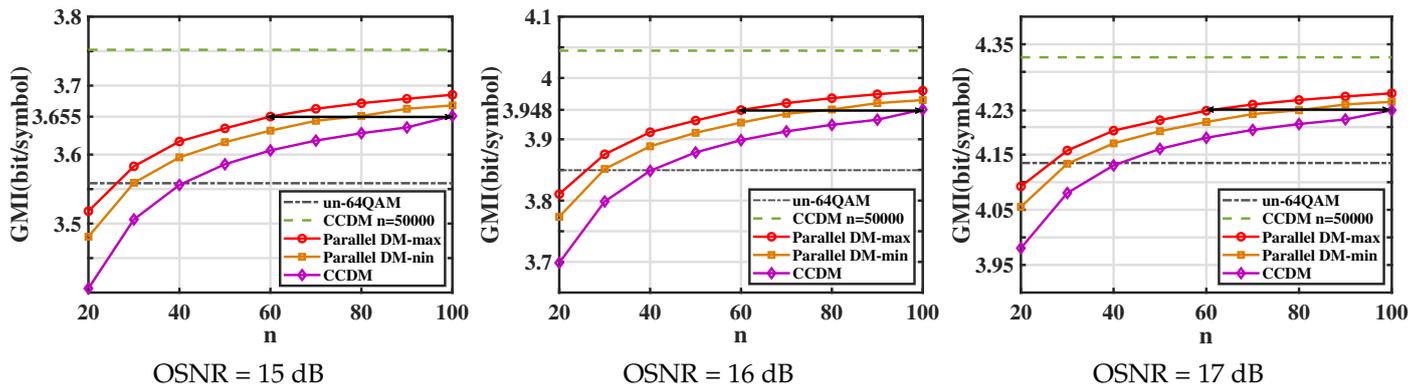


Figure 6. GMI in bit/symbol over block length n when OSNR = 15, 16, and 17 dB. When the GMI is the same, the parallel distribution matcher required block length n is only 60, while the required block length of CCDM is 100; so, the block length can be saved by 40%.

In terms of system bit error rate, the probability distribution of the final PS-QAM signal is the same whether utilizing the parallel distribution matcher or CCDM; the only difference is the number of PS-64QAM symbols generated. If $n = 50$ and the input bit sequence length is 58,128, the parallel structure can generate 33,600 PS-64QAM symbols, while the CCDM can generate 34,600 PS-64QAM symbols. When $n = 100$ and the input bit sequence length is 62,300, the parallel structure can generate 35,000 PS-64QAM symbols and the CCDM can generate 35,600 PS-64QAM symbols. Consequently, the bit error rate performance of the two is the same in the back-to-back situation.

When the transmission distance is long in the optical fiber transmission link, the amplifier spontaneous emission accumulates continuously and the number of symbols can have a slight impact on the bit error rate. Simulation block diagram and parameters of standard single-mode optical fiber transmission are shown in Figure 7. Figure 8 shows the BER performance under different transmission distances; the fiber is a standard single-mode fiber (SSMF) with $\alpha = 0.2$ dB/km, $\gamma = 1.3$ (W · km)⁻¹, and $D = 17$ ps/nm/km; each span of length 100 km is followed by an Erbium-doped fiber amplifier with a noise figure of 3.8 dB. Laser phase noise and polarization mode dispersion are not included in the simulation as perfect compensation is assumed. The transmission distance of PS-64QAM signals using the paired optimized parallel distribution matcher is marginally greater than that of PS-64QAM signals using CCDM at the forward error correction threshold of -2.42 for $n = 50$ and $n = 100$.

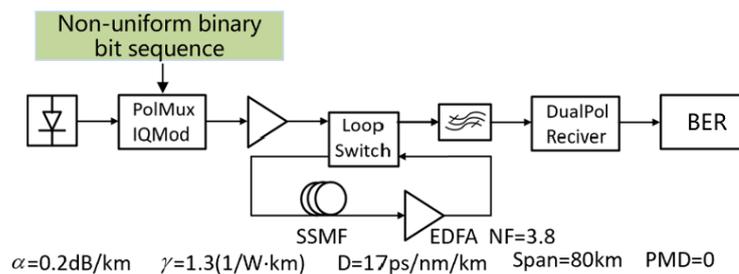


Figure 7. Simulation block diagram and parameters of standard single-mode optical fiber transmission.

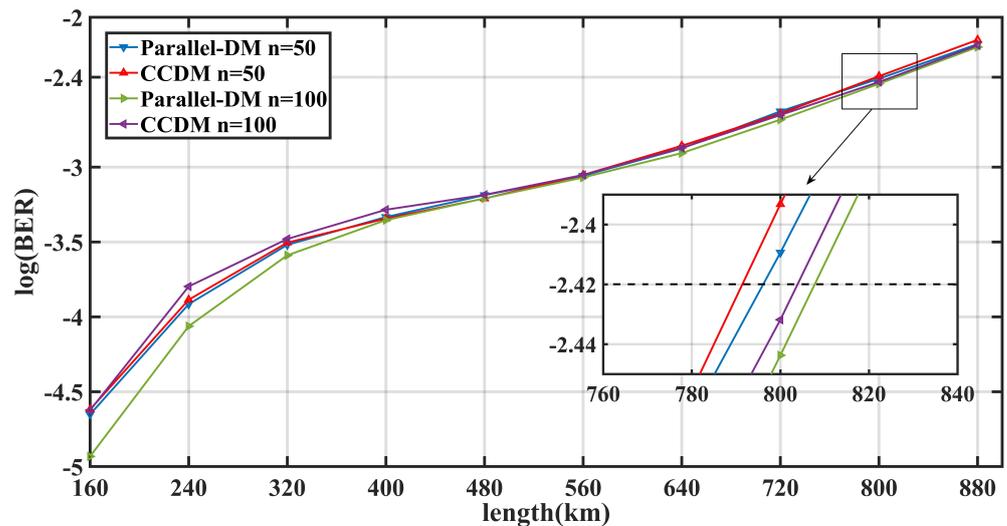


Figure 8. The BER performance over transmission distance in the standard single-mode fiber transmission link. The transmission distance of PS-64QAM signals using the parallel distribution matcher is slightly larger than that of PS-64QAM signals using CCDM.

5. Conclusions and Discussion

As a key part of probability shaping technology, the distribution matcher has significant importance and application value for research. At present, the widely used constant composition distribution matcher is a progressive optimization scheme. It is only when the length of the output symbols is infinity long that the rate loss of CCDM is zero. In addition, it adopts arithmetic coding, which is a highly serial coding method. It represents the input and output sequences, respectively, by dividing intervals between intervals. When the number of input and output sequences is large, the intervals to be divided become greater and the boundary between intervals becomes blurred, which is difficult to distinguish and easy to make mistakes when mapping. Therefore, it is necessary to propose a distribution matcher structure with good performance under short block length.

In this paper, we have proposed a novel parallel architecture distribution matcher, which is an improvement over conventional CCDM at short output block lengths. The output sequence has various compositions compared with the constant composition of the CCDM because of the parallel structure and pairwise optimization of the two CCDMs. As a result, the proposed structure has a lower rate loss when the output block lengths (n is less than 100) are the same and the output block lengths can be lowered by up to 30% with the same rate loss. In the simulation of an optical communication system of PS-64QAM signal, compared with a CCDM signal with the same generalized mutual information, the PS-64QAM signal employing this structure requires a lower OSNR, the block length can be reduced by 40%, and the transmission distance is increased. In detail, When GMI is 4 bit/symbol, compared with the uniformly distributed 64QAM signal, the OSNR of the PS-64QAM signal using CCDM is reduced by 0.35 dB, and the OSNR of the PS-64QAM signal using the parallel distribution matcher is reduced by 0.47 dB, which is a 0.12 dB improvement. The improvement in transmission distance is small, and it is expected to further improve the transmission distance in combination with digital signal processing technology, which will be our future research direction.

Other parallel distribution matcher schemes, such as parallel-amplitude distribution matcher [24], realize the function of symbol-level DM through multiple-bit-level DM. Therefore, $M-1$ bit-level DM is required to map sequences with M output amplitude values. For high-order QAM format, the number of bit-level DM required increases with the increase in modulation order, and the overall structure of the system is complex. In addition, the multiset-partition distribution [20], which also adopts the idea of pairwise optimization, needs to build a Huffman tree structure to index different complementary pairs and uses serial processing for the input data. The coding process is complex and

error-prone. In contrast, the scheme proposed in this paper is a two-stage parallel structure. No matter how high the order of the QAM signal is, only two CCDMs are required. At the same time, the input data are processed in the two CCDMs after serial–parallel conversion. The receiver performs a corresponding parallel–serial conversion. There is no complex coding process; so, the structure is simple.

Author Contributions: Conceptualization, Y.Z. (Yao Zhang); methodology, Y.Z. (Yao Zhang) and H.W.; software and validation, Y.Z. (Yao Zhang); writing—original draft preparation, Y.Z. (Yao Zhang); supervision and project administration, H.W., Y.J. and Y.Z. (Yu Zhang); funding acquisition, H.W., Y.J. and Y.Z. (Yu Zhang). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key Research and Development Program of China (Grant No. 2019YFB1803601) and the National Natural Science Foundation of China (Grant No. 62021005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the editors and anonymous reviewers for giving valuable suggestions that have helped to improve the quality of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cisco, V.N.I. Global IP Traffic Forecast, 2017–2022. In *Cisco System*; Cisco: San Jose, CA, USA, 2018.
2. Cisco, U. *Cisco Annual Internet Report (2018–2023) White Paper*; Acessado Em; Cisco: San Jose, CA, USA, 2021.
3. Cho, J.; Winzer, P.J. Probabilistic constellation shaping for optical fiber communications. *J. Lightwave Technol.* **2019**, *37*, 1590–1607. [[CrossRef](#)]
4. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
5. Kschischang, F.R.; Pasupathy, S. Optimal nonuniform signaling for Gaussian channels. *IEEE Trans. Inf. Theory.* **1993**, *39*, 913–929. [[CrossRef](#)]
6. Gallager, R.G. *Information Theory and Reliable Communication*; Wiley: New York, NY, USA, 1968; Volume 588.
7. Raphaeli, D.; Gurevitz, A. Constellation shaping for pragmatic turbo-coded modulation with high spectral efficiency. *IEEE Trans. Commun.* **2004**, *52*, 341–345. [[CrossRef](#)]
8. Yankov, M.; Forchhammer, S.; Larsen, K.J.; Christensen, L.P. Rate-adaptive constellation shaping for near-capacity achieving turbo coded BICM. In Proceedings of the 2014 IEEE International Conference on Communications (ICC), Sydney, Australia, 10–14 June 2014; pp. 2112–2117.
9. Kaimalettu, S.; Thangaraj, A.; Bloch, M.; McLaughlin, S.W. Constellation shaping using LDPC codes. In Proceedings of the 2007 IEEE International Symposium on Information Theory, Nice, France, 24–29 June 2007; pp. 2366–2370.
10. Le Goff, S.Y.; Sharif, B.S.; Jimaa, S.A. Bit-interleaved turbo-coded modulation using shaping coding. *IEEE Commun. Lett.* **2005**, *9*, 246–248. [[CrossRef](#)]
11. Böcherer, G.; Mathar, R.; Jimaa, S.A. Operating LDPC codes with zero shaping gap. In Proceedings of the 2011 IEEE Information Theory Workshop (ITW), Paraty, Brazil, 16–20 October 2011; pp. 330–334.
12. Böcherer, G.; Steiner, F.; Schulte, P. Bandwidth efficient and rate-matched low-density parity-check coded modulation. *IEEE Trans. Commun.* **2015**, *63*, 4651–4665. [[CrossRef](#)]
13. Böcherer, G.; Schulte, P.; Steiner, F. Probabilistic shaping and forward error correction for fiber-optic communication systems. *J. Lightwave Technol.* **2019**, *37*, 230–244. [[CrossRef](#)]
14. Ghazisaeidi, A.; de Jauregui Ruiz, I.F.; Rios-Müller, R.; Schmalen, L.; Tran, P.; Brindel, P.; Renaudier, J. Advanced C+ L-band transoceanic transmission systems based on probabilistically shaped PDM-64QAM. *J. Lightwave Technol.* **2017**, *35*, 1291–1299. [[CrossRef](#)]
15. Idler, W.; Buchali, F.; Schmalen, L.; Lach, E.; Braun, R.P.; Böcherer, G.; Steiner, F. Field trial of a 1 Tb/s super-channel network using probabilistically shaped constellations. *J. Lightwave Technol.* **2017**, *35*, 1399–1406. [[CrossRef](#)]
16. Cho, J.; Chen, X.; Chandrasekhar, S.; Raybon, G.; Dar, R.; Schmalen, L.; Grubb, S. Trans-atlantic field trial using high spectral efficiency probabilistically shaped 64-QAM and single-carrier real-time 250-Gb/s 16-QAM. *J. Lightwave Technol.* **2017**, *36*, 103–113. [[CrossRef](#)]
17. Kong, M.; Yu, J.; Chien, H.C.; Wang, K.; Li, X.; Shi, J.; Chen, Y. WDM Transmission of 600G Carriers over 5600 km with Probabilistically Shaped 16QAM at 106 Gbaud. In Proceedings of the Optical Fiber Communication Conference (OFC) 2019, San Diego, CA, USA, 3–7 March 2019; pp. 1–3.

18. Kong, M.; Wang, K.; Ding, J.; Zhang, J.; Li, W.; Shi, J.; Yu, J. 640-Gbps/carrier WDM transmission over 6,400 km based on PS-16QAM at 106 Gbaud employing advanced DSP. *J. Lightwave Technol.* **2020**, *39*, 55–63. [[CrossRef](#)]
19. Schulte, P.; Böcherer, G. Constant composition distribution matching. *IEEE Trans. Inf. Theory.* **2015**, *62*, 430–434. [[CrossRef](#)]
20. Fehenberger, T.; Millar, D.S.; Koike-Akino, T.; Kojima, K.; Parsons, K. Multiset-partition distribution matching. *IEEE Trans. Commun.* **2018**, *67*, 1885–1893. [[CrossRef](#)]
21. Fehenberger, T.; Millar, D.S.; Koike-Akino, T.; Kojima, K.; Parsons, K.; Griesser, H. Huffman-coded sphere shaping and distribution matching algorithms via lookup tables. *J. Lightwave Technol.* **2020**, *38*, 2826–2834. [[CrossRef](#)]
22. Pikus, M.; Xu, W. Bit-level probabilistically shaped coded modulation. *IEEE Commun. Lett.* **2017**, *21*, 1929–1932. [[CrossRef](#)]
23. Koganei, Y.; Sugitani, K.; Nakashima, H.; Hoshida, T. Optimum bit-level distribution matching with at most $O(N^3)$ implementation complexity. In Proceedings of the Optical Fiber Communication Conference (OFC) 2019, San Diego, CA, USA, 3–7 March 2019; p. M4B-4.
24. Fehenberger, T.; Millar, D.S.; Koike-Akino, T.; Kojima, K.; Parsons, K. Parallel-amplitude architecture and subset ranking for fast distribution matching. *IEEE Trans. Commun.* **2020**, *68*, 1981–1990. [[CrossRef](#)]
25. Yoshida, T.; Karlsson, M.; Agrell, E. Hierarchical distribution matching for probabilistically shaped coded modulation. *J. Lightwave Technol.* **2019**, *37*, 1579–1589. [[CrossRef](#)]
26. Goossens, S.; Van der Heide, S.; Hout, M.V.D.; Amari, A.; Gultekin, Y.C.; Vassilieva, O.; Kim, I.; Ikeuchi, T.; Willems, F.M.J.; Alvarado, A.; et al. First experimental demonstration of probabilistic enumerative sphere shaping in optical fiber communications. In Proceedings of the 24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC), Fukuoka, Japan, 7–11 July 2019; pp. 1–3.
27. Gültekin, Y.C.; Fehenberger, T.; Alvarado, A.; Willems, F.M.J. Probabilistic shaping for finite blocklengths: Distribution matching and sphere shaping. *J. Entropy.* **2020**, *22*, 581. [[CrossRef](#)] [[PubMed](#)]
28. Fehenberger, T.; Alvarado, A.; Böcherer, G.; Hanik, N. On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel. *J. Lightwave Technol.* **2016**, *34*, 5063–5073. [[CrossRef](#)]