Probabilistic Amplitude Shaping

Other titles in Foundations and Trends® in Communications and Information Theory

Reed-Muller Codes
Emmanuel Abbe, Ori Sberlo, Amir Shpilka and Min Ye
ISBN: 978-1-63828-144-3

Topics and Techniques in Distribution Testing: A Biased but Representative Sample
Clément L. Canonne
ISBN: 978-1-63828-100-9

Codes in the Sum-Rank Metric: Fundamentals and Applications
Umberto Martínez-Peñas, Mohannad Shehadeh and Frank R. Kschischang
ISBN: 978-1-63828-030-9

Codes for Distributed Storage
Vinayak Ramkumar, S. B. Balaji, Birenjith Sasidharan, Myna Vajha, M. Nikhil Krishnan and P. Vijay Kumar
ISBN: 978-1-63828-024-8

Rank-Metric Codes and Their Applications
Hannes Bartz, Lukas Holzbaur, Hedongliang Liu, Sven Puchinger, Julian Renner and Antonia Wachter-Zeh
ISBN: 978-1-63828-000-2

Common Information, Noise Stability, and Their Extensions
Lei Yu and Vincent Y. F. Tan
ISBN: 978-1-63828-014-9

Probabilistic Amplitude Shaping

Georg Böcherer
Huawei Technologies
georg.boecherer@ieee.org

the essence of knowledge
Boston — Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

G. Böcherer. *Probabilistic Amplitude Shaping.* Foundations and Trends® in Communications and Information Theory, vol. 20, no. 4, pp. 390–511, 2023.

ISBN: 978-1-63828-179-5

Foundations and Trends® in Communications and Information Theory

Volume 20, Issue 4, 2023

Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Communications and Information Theory publishes survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2023, Volume 20, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328 . Also available as a combined paper and online subscription.

Contents

Probabilistic Amplitude Shaping

Georg Böcherer

Huawei Technologies, Germany; georg.boecherer@ieee.org

ABSTRACT

Probabilistic amplitude shaping (PAS) proposed in Böcherer, Steiner, Schulte [24] is a practical architecture for combining non-uniform distributions on higher-order constellations with off-the-shelf forward error correction (FEC) codes. PAS consists of a distribution matcher (DM) that imposes a desired distribution on the signal point amplitudes, followed by systematic FEC encoding, preserving the amplitude distribution. FEC encoding generates additional parity bits, which select the signs of the signal points. At the receiver, FEC decoding is followed by an inverse DM. PAS quickly had a large industrial impact, in particular in fiber-optic communications. This monograph details the practical considerations that led to the invention of PAS and provides an information-theoretic assessment of the PAS architecture. Because of the separation into a shaping layer and an FEC layer, the theoretic analysis of PAS requires new tools. On the shaping layer, the cost penalty and rate loss of finite length DMs is analyzed. On the FEC layer, achievable FEC rates are derived. Using mismatched decoding, achievable rates are studied for decoding metrics of practical importance. Combining the findings, it is shown that PAS with linear codes is capacity-achieving on a class of discrete input channels. Open questions for future study are discussed.

Georg Böcherer (2023), "Probabilistic Amplitude Shaping", Foundations and Trends® in Communications and Information Theory: Vol. 20, No. 4, pp 390–511. DOI: 10.1561/0100000111.

Preface

Almost 10 years ago, we simulated for the first time a communication system architecture that we later called Probabilistic Amplitude Shaping[1], the title of this monograph.

How to use this monograph All readers should read Section 1: it discusses the line of thoughts that led to the invention of PAS, outlines this monograph, and provides pointers to the literature.

The **theorist** may then read the discussion sections provided at the ends of Sections 2–5. The discussion sections summarize the sections, provide pointers to the literature, and outline open problems for future study. Also of interest to the theorist may be some of the proof techniques. For instance, the study of cost and rate scaling of distribution matchers[2] in Section 2.6, the layered probabilistic shaping (PS) random code ensemble in Section 3.1, the "any channel" achievable forward error correction (FEC) rate in Theorem 3.2, and the derivation of the PAS error exponent in Section 5.

The **practitioner** may implement the formulas provided throughout the monograph for numerical evaluation as guidance for designing PAS systems for industrial application. He/she may consult the PAS webpage (see below) to check for available implementations and may also consider contributing his/her implementations. For instance, one may use the

[1]We introduced the name Probabilistic Amplitude Shaping (PAS) in [24].

[2]We introduced the term Distribution Matching (DM) in [22].

formulas from Section 2 for choosing a DM class and dimensioning DM input and output lengths for trading rate and cost against latency and complexity. Also, the practitioner may implement the cross-equivocation formulas from Section 4 to compare the performance limits of binary and nonbinary codes, to choose between hard-decision and soft-decision, or to select the resolution for quantized soft-decision decoding. Similarily, one may implement the PAS achievable rate formulas from Section 5 for assessing the performance penalty caused by a constrained FEC rate, or for plotting PAS rate limits for finite length at a required reliability. Also of practical interest are the PAS system parameters FEC overhead, shaping set rate, and PS overhead as discussed in Sections 3.3 and 4.1.

For the **lecturer**, the cross-equivocation formalism from Section 4 may be of interest. Besides the basic decoding metrics discussed in this monograph, one can easily come up with many more variations, which according to my own teaching experience provide a rich source for homework and exam questions.

The **machine learning engineer** may find interest in the cross-equivocation formalism from Section 4. The underlying empirical cross-equivocation defined in Section 3 is identical to the cross-entropy loss frequently used in machine learning. Thus, the discussion in this monograph may provide the machine learning engineer with an interesting communication system perspective on the cross-entropy loss.

Webpage One shortcoming of this monograph is an insufficient number of plots with numerical evaluations for illustrating the developed concepts. I just did not have the time to add all the illustrations I would like to have. I have therefore set up a webpage[3] to accompany this monograph, for the following purposes:

- To host implementations of formulas and algorithms provided by the community.
- To share numerical plots of performance evaluations provided by the community.
- To publish the errata of this monograph.

[3]https://github.com/gbsha/PAS

I hope this provides an effective alternative to providing numerical evaluations in the monograph.

Acknowledgments Prof. Valdemar da Rocha and Prof. Cecilio Pimentel suggested to me as a master thesis topic the study of the discrete noiseless channel at their chair at the Federal University of Pernambuco. This triggered my interest in constrained coding and led to my study of variable length DM algorithms during my PhD at Prof. Rudolf Mathar's chair at the RWTH Aachen University. The work of Prof. David MacKay and his students (in particular the MacKay-Neal codes[4] and the sparse-dense codes[5]) inspired me to combine DM and FEC. Prof. Alex Alvarado brought my interest to the study of bit-interleaved coded modulation. Coded modulation in general was brought to my attention by Gottfried Ungerboeck when I served as his teaching assistant during the first months of my postdoc at Prof. Gerhard Kramer's chair at the Technical University of Munich.

The invention of PAS resulted in an exciting time with great people. Some memories are: Studying variable length DMs with Rana Ali Amjad and Sebastian Baur; Prof. Stephan ten Brink looking at an early PAS diagram and understanding it faster than anyone else before or after; a discussion with Irina Bocharova and Boris Kudryashov that led to the development of constant composition distribution matching (CCDM) by Patrick Schulte; the first implementation of PAS for a simulated optical transmission with Tobias Fehenberger; Gianluigi Liva asking whether one could change the DM distribution to adjust the PAS rate; the first PAS optical transmission experiment with Fred Buchali and Prof. Laurent Schmalen; the Bell Labs Prize 2015 together with Fabian Steiner and Patrick Schulte; Prof. Richard Wesel suggesting to change "rate-compatible" for "rate-matched" in the title of the PAS paper; working with Bernhard Geiger on quantization for distribution synthesis; Tobias Prinz developing polar coded PAS; the suggestion of Prof. Frans Willems to use sequences up to a maximum cost for DM, which led to the development of minimum cost distribution matching

[4]MacKay [54, Section VI].
[5]Ratzer [61, Chapter 5].

(MCDM) by several groups; the Johann-Philipp-Reis-Preis 2017; the collaboration with Prof. Neri Merhav on error exponents for layered PS; Huijian Zhang and Zhuhong Zhang appreciating the invention of PAS.

Prof. Frank Kschischang proposed this monograph to Prof. Alexander Barg, the editor in chief of this journal. Prof. Alexander Barg and publisher Mike Casey showed great patience during the making of this monograph. Two anonymous reviewers provided very valuable comments on a first version.

I thank you all.

GeorgBöcherer
Munich, Germany
May 2023

1

Probabilistic Amplitude Shaping

In this section, we discuss the line of thoughts that led to the invention of probabilistic amplitude shaping (PAS). The key ingredients are three tools that have been available to the communications engineer already for some time. These three tools are: first, the additive white Gaussian noise (AWGN) capacity formula [71], second, powerful capacity-approaching binary low-density parity-check (LDPC) codes [37] and the possibility to simulate them on a personal computer [54], and third, the bit-interleaved coded modulation (BICM) architecture [27]. We briefly discuss the capacity formula in Section 1.1.1, binary forward error correction (FEC) in Section 1.1.2, word error rates (WERs) and bit error rates (BERs) in Section 1.1.3 and BICM in Section 1.2. With these tools at hand, the thought process that leads to PAS is rather of practical than theoretic nature. The steps consist in successive modifications of a practical system for simulating WERs of a binary FEC in AWGN. We discuss these modifications in Section 1.3. The PAS architecture raises several design questions, which we list in Section 1.4 and address in greater detail in the following sections of this monograph.

1.1 Preliminaries

1.1.1 AWGN Capacity

The real-valued discrete time AWGN channel is

$$Y_i = X_i + Z_i, \quad i = 1, 2, \dots, n \tag{1.1}$$

where the Y_i, X_i, and Z_i are outputs, inputs, and noise, respectively. Inputs and noise are independent and the Z_i are independent zero mean Gaussian with variance σ^2, i.e.,

$$p_{Z_i}(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{z^2}{2\sigma^2}}. \tag{1.2}$$

The input is subject to an average power constraint

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i^2) \le \mathsf{P}. \tag{1.3}$$

The capacity of the AWGN channel is

$$\max_{P_X:\ \mathbb{E}(X^2)\le \mathsf{P}} \mathbb{I}(X;Y) = \frac{1}{2}\log_2\left(1 + \frac{\mathsf{P}}{\sigma^2}\right) \tag{1.4}$$

where $\mathbb{I}(X;Y)$ denotes the mutual information of X and Y, see (A.3.5). The ratio P/σ^2 is called the signal-to-noise ratio (SNR). The capacity-achieving density of the AWGN is zero mean Gaussian with variance P.

1.1.2 Binary Linear FEC

Parity Check Matrix To protect a block $\boldsymbol{c} = c_1 \dots c_n$ of n bits against errors, a linear FEC code imposes m_{fec} linear constraints on $\boldsymbol{c}$. Each constraint requires that a certain subset of the n bits in $\boldsymbol{c}$ add to an even number, i.e., zeros in the binary field. The constraints are therefore called *parity checks*. The ith parity check is compactly written as a length n row vector $\boldsymbol{h}_i = h_{i1} \dots h_{in}$ and the vector $\boldsymbol{c}$ must fulfill $\boldsymbol{c}\boldsymbol{h}_i^{\mathrm{T}} = 0$. Arranging m_{fec} parity checks in a matrix results in the *parity check matrix* $\boldsymbol{H}$ with transpose

$$\boldsymbol{H}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{h}_1^{\mathrm{T}} & \boldsymbol{h}_2^{\mathrm{T}} & \cdots & \boldsymbol{h}_{m_{\text{fec}}}^{\mathrm{T}} \end{bmatrix} \tag{1.5}$$

and $\boldsymbol{c}$ is a codeword if and only if it fulfills all m_{fec} parity checks, i.e.,

$$\boldsymbol{c}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{0}. \tag{1.6}$$

This defines the linear FEC code

$$\mathcal{C} := \left\{\boldsymbol{c} \in \{0,1\}^n : \boldsymbol{c}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{0}\right\}. \tag{1.7}$$

Systematic Encoding It is convenient for the last m_{fec} columns of $\boldsymbol{H}$ to be linearly independent, which can always be achieved, when $\boldsymbol{H}$ is full rank, by suitable rearrangement of columns. Then, the matrix is of the form

$$\boldsymbol{H} = [\boldsymbol{Q}|\boldsymbol{R}] \tag{1.8}$$

where the $m_{\text{fec}} \times m_{\text{fec}}$ matrix $\boldsymbol{R}$ is full rank and invertible. Systematic encoding of k bits $\boldsymbol{u}$ can now be done in two steps.

1. Calculate $\boldsymbol{s} = \boldsymbol{u}\boldsymbol{Q}^{\mathrm{T}}$.
2. Solve $\boldsymbol{p}\boldsymbol{R}^{\mathrm{T}} = \boldsymbol{s} \Rightarrow \boldsymbol{p} = \boldsymbol{s}(\boldsymbol{R}^{\mathrm{T}})^{-1}$.

The vector $\boldsymbol{c} = [\boldsymbol{u}|\boldsymbol{p}]$ is then a codeword, i.e., it fulfills $\boldsymbol{c}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{0}$. A convenient way to represent systematic encoding is via a systematic generator matrix

$$\boldsymbol{G} = [\boldsymbol{I}|\boldsymbol{P}] \tag{1.9}$$

where $\boldsymbol{I}$ is a $k \times k$ identity matrix and $\boldsymbol{P} = \boldsymbol{Q}^{\mathrm{T}}(\boldsymbol{R}^{\mathrm{T}})^{-1}$. We can now compactly write systematic encoding by the multiplication of $\boldsymbol{u}$ with $\boldsymbol{G}$, i.e.,

$$\boldsymbol{u}\boldsymbol{G} = [\boldsymbol{u}|\boldsymbol{p}] = \boldsymbol{c}. \tag{1.10}$$

Since $\boldsymbol{p} = \boldsymbol{u}\boldsymbol{P}$, we call $\boldsymbol{P}$ the *parity forming part* of $\boldsymbol{G}$.

1.1.3 Word- and Bit Error Rate

The performance of FEC codes are usually characterized either by their WER or by their BER. While information theorists mainly use WER, e.g., for channel capacity, communications engineers mainly use BER.

In the remainder of this section, we consider WER, for the sake of simplicity. The obtained insights hold similarly for BER. We next define WER and BER formally and relate them to each other.

Consider a binary code with codeword length n. Suppose $\#\{\mathrm{W}\}$ codewords were transmitted and after decoding, $\#\{\mathrm{WE}\}$ word errors occured. The WER is then

$$\mathrm{WER} = \frac{\#\{\mathrm{WE}\}}{\#\{\mathrm{W}\}}. \tag{1.11}$$

The number of transmitted bits is $\#\{\mathrm{B}\} = n \cdot \#\{\mathrm{W}\}$. In each erroneous codeword, the number of erroneous bits is at least one and at most n. Thus, the number of bit errors is bounded as

$$\#\{\mathrm{WE}\} \leq \#\{\mathrm{BE}\} \leq n \cdot \#\{\mathrm{WE}\}. \tag{1.12}$$

The BER is

$$\mathrm{BER} = \frac{\#\{\mathrm{BE}\}}{\#\{B\}} \tag{1.13}$$

and bounded by

$$\frac{1}{n}\mathrm{WER} \leq \mathrm{BER} \leq \mathrm{WER}. \tag{1.14}$$

In particular, the BER is upper bounded by the WER, so if we design a communication link with low WER, we can guarantee that it has a low BER, too.

Another way to relate the two error rates is to consider error exponents. Suppose we have a family of FEC codes where we can choose the codeword length n as large as we want. Denote the corresponding error rates by $\mathrm{WER}(n)$ and $\mathrm{BER}(n)$. Then, the word and bit error exponents are respectively

$$E_W = \lim_{n\to\infty} -\frac{\log \mathrm{WER}(n)}{n} \tag{1.15}$$

$$E_B = \lim_{n\to\infty} -\frac{\log \mathrm{BER}(n)}{n}. \tag{1.16}$$

By (1.14), E_B is lower bounded by E_W and to bound E_B from above, consider

$$-\frac{\log \mathrm{BER}(n)}{n} \leq -\frac{\log\left(\frac{1}{n}\mathrm{WER}(n)\right)}{n} \tag{1.17}$$
$$= -\frac{\log \mathrm{WER}(n)}{n} + \frac{\log n}{n} \tag{1.18}$$
$$\xrightarrow{n\to\infty} -\frac{\log \mathrm{WER}(n)}{n} \tag{1.19}$$

which implies that E_B is also upper bounded by E_W. Consequently, the bit error exponent is equal to the word error exponent.

1.2 Bit-Interleaved Coded Modulation

1.2.1 BPSK in AWGN

Full System

This diagram lays out a coded transmission over the AWGN channel using a binary FEC code with a soft decision (SD) decoder. Let's go through the components from left to right.

Information bits b^k are encoded by a systematic encoder, which appends parity bits p^{n-k}. Together, information and parity bits form the codeword c^n. The coded bits are then mapped to binary phase shift keying (BPSK) symbols by the binary mapping

$$0 \mapsto x(0) = -1 \tag{1.20}$$
$$1 \mapsto x(1) = 1. \tag{1.21}$$

The BPSK symbols are transmitted over the channel and the channel output is

$$y_i = x_i + z_i, \quad i = 1, \ldots, n \tag{1.22}$$

where the z_i are independent zero mean Gaussians with variance σ^2. The demapper calculates the soft-decisions

$$\ell_i = \log \frac{p_{Y|B}(y_i|0)}{p_{Y|B}(y_i|1)} = \log \frac{p_{Y|X}(y_i|-1)}{p_{Y|X}(y_i|+1)}, \quad i = 1, \ldots, n \tag{1.23}$$

and the decoder outputs its decision $\hat{c}^n = \hat{b}^k \hat{p}^{n-k}$. For our discussion in this section, three ways to calculate the decision $\hat{c}^n$ from the soft-decision ℓ^n (or equivalently, from the likelihoods $p_{Y|B}(y_i|0)$ and $p_{Y|B}(y_i|1)$) are relevant.

1. To assess performance limits, we consider the mutual information $\mathbb{I}(B;Y)$ for uniformly distributed input bits. The achievability of mutual information is proven, e.g., in [39, Chapter 5], by considering a random code ensemble and the maximum-likelihood (ML) decision rule
$$\hat{c}^n = \underset{c^n \in \mathcal{C}}{\arg\max} \sum_{i=1}^{n} \ell_i (1 - 2c_i) \tag{1.24}$$
which minimizes the WER. We discuss decision rules and achievable rates for non-uniformly distributed input in detail in Sections 3, 4, and 5.

2. A bitwise maximum a posteriori probability (MAP) decoder, see, e.g., [62, Section 2.5.1], uses the decision rule
$$\begin{aligned} \hat{c}_i &= \underset{b \in \{0,1\}}{\arg\max}\, P_{B_i|Y^n}(b|y^n) & (1.25)\\ &= \underset{b \in \{0,1\}}{\arg\max} \sum_{\substack{c^n \in \mathcal{C}\\ c_i = b}} P_{B^n|Y^n}(c^n|y^n) & (1.26)\\ &= \underset{b \in \{0,1\}}{\arg\max} \sum_{\substack{c^n \in \mathcal{C}\\ c_i = b}} \prod_{j=1}^{n} p_{Y|B}(y_j|c_j). & (1.27) \end{aligned}$$

3. A practical LDPC decoder approximates the bitwise MAP rule by message passing on a graph with cycles. All simulation results presented in this section were obtained by using the DVB-S2 rate 4/5 LDPC code with parameters specifed in Table 1.1.

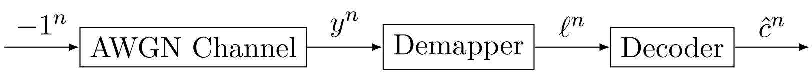
We evaluate the peformance by estimating the WER
$$\mathrm{WER} = \Pr(\hat{C}^n \neq C^n) \tag{1.28}$$
by Monte Carlo simulation. We display the WER curve in Figure 1.1, and we show the operating point at $\mathrm{WER} = 1 \times 10^{-3}$ in Figure 1.2.

Table 1.1: Parameters of the DVB-S2 LDPC code.

R_{fec}	4/5
n	64 800
k	51 840
m_{fec}	12 960
decoding algorithm	belief propagation
number of iterations	50

We note that the operating point is ≈ 0.6 dB away from the BPSK limit $\mathbb{I}(B;Y)$ and ≈ 1.6 dB away from capacity. Later in this section, we will use the 0.6 dB gap to the BPSK limit as a rough estimate of the FEC penalty of the considered code.

All Zero Codeword



If we are only interested in evaluating the WER and don't need a fully functioning system, we can simplify our setup. For BPSK in AWGN, the two input symbols -1 and $+1$ are affected equally by noise. Therefore, since the FEC code is linear, the WER does not depend on the transmitted codeword. All linear codes have the all-zero vector as codeword, so that we can remove the encoder and the mapper and transmit the -1^n vector. The WER is then

$$\Pr(\hat{C}^n \neq 0^n) \tag{1.29}$$

which we can estimate by Monte Carlo simulation. As expected, in Figures 1.1 and 1.2, the all-zero codeword WER is on top of the full system WER. The all-zero codeword system has several advantages.

1. We need to write less code for implementing it.
2. The simulation runs faster since unnecessary calculations are skipped.
3. We can evaluate FEC codes for which we have a decoder but no encoder.

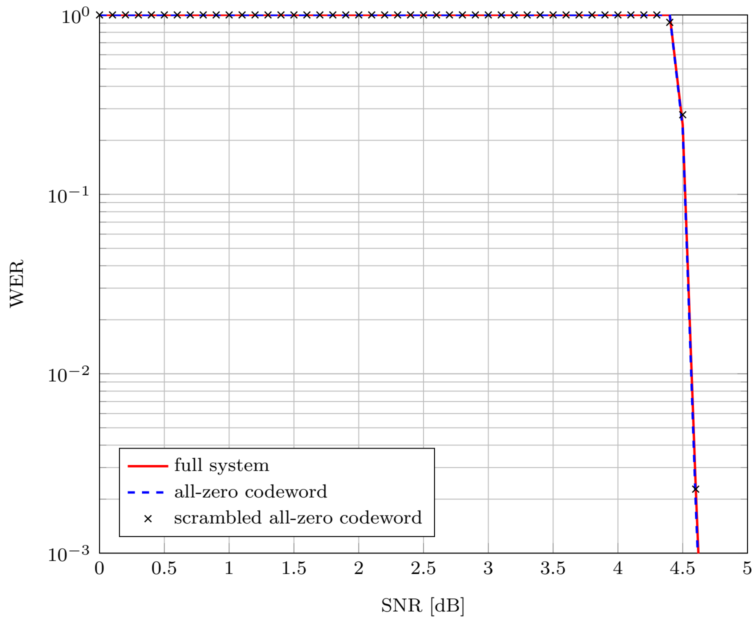


Figure 1.1: WER of BPSK.

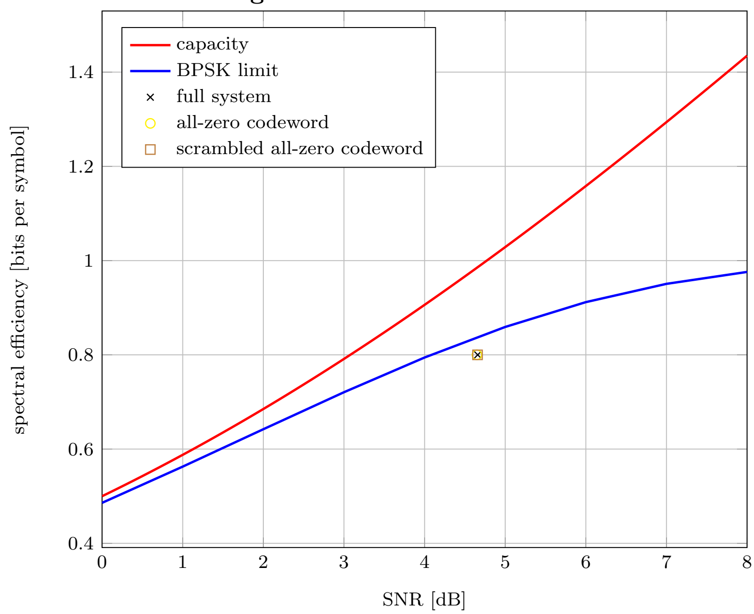
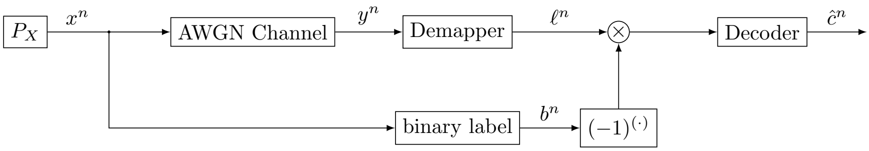


Figure 1.2: $\mathrm{WER} = 1 \times 10^{-3}$ operating point of BPSK.

Scrambled All-Zero Codeword



Instead of transmitting always -1, we can also sample the BPSK symbols independently with distribution $P_X(-1) = P_X(1) = \frac{1}{2}$. The binary label b^n of the random sequence x^n is unlikely to be a codeword. Therefore, we interpret b^n as a scrambling sequence that was applied to the all-zero codeword. Accordingly, we must descramble the demapper output ℓ^n before we pass it to the decoder. The WER is now again

$$\Pr(\hat{C}^n \neq 0^n) \tag{1.30}$$

and we estimate it by Monte Carlo simulation. As expected, in Figures 1.1 and 1.2, the scrambled all-zero codeword WER is on top of the full system WER.

1.2.2 Bit-Interleaved Coded Modulation

As we can see in Figure 1.2, for sufficiently high SNR, the BPSK limit flattens out and the gap to capacity becomes arbitrarily large. We therefore need to use constellations larger than BPSK, which is called higher-order modulation. BICM [27] provides the appropriate framework for combining higher-order modulation with binary FEC. For specifying a BICM system, we first need some definitions.

Amplitude Shift Keying We use amplitude shift keying (ASK) constellations with M symbols

$$\mathcal{X} = \{\pm 1, \pm 3, \dots, \pm(M-1)\} \tag{1.31}$$

where $M = 2^m$ for some integer m. Note that $M = 2$ recovers BPSK.

Bitwise Demapping We associate with each symbol $x \in \mathcal{X}$ a binary label $b^m = \phi(x) \in \{0,1\}^m$. The jth bit level is $b_j = \phi_j(x)$. Define the symbol sets

$$\mathcal{X}_b^j = \{x \in \mathcal{X} \colon \phi_j(x) = b\}, \quad j = 1, \dots, m,\ b \in \{0,1\}. \tag{1.32}$$

Table 1.2: The BRGC for 8-ASK.

symbol x	label $\phi(x)$
-7	000
-5	001
-3	011
-1	010
1	110
3	111
5	101
7	100

For each bit level j, the constellation $\mathcal{X}$ is partitioned into $\mathcal{X}_0^j$ with symbols where bit level j is 0, and $\mathcal{X}_1^j$ where bit level j is 1. The demapper calculates

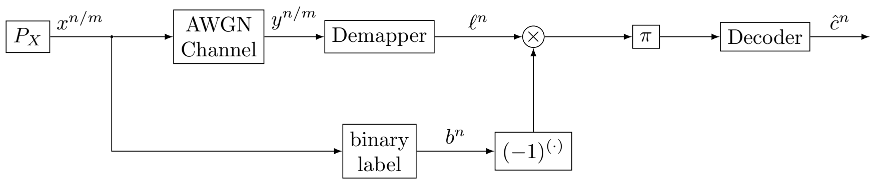
$$\ell_{ji} = \log \frac{P_{Y|B_j}(y_i|0)}{P_{Y|B_j}(y_i|1)} = \log \frac{\sum_{x\in\mathcal{X}_0^j} p_{Y|X}(y_i|x)}{\sum_{x\in\mathcal{X}_1^j} p_{Y|X}(y_i|x)}$$
$$j = 1, \ldots, m,\ i = 1, \ldots, n/m. \tag{1.33}$$

The ℓ_{ji} are reindexed to a length n vector ℓ^n and passed to the LDPC decoder, which outputs the decision $\hat{c}^n$. The internal processing of the LDPC decoder only depends on ℓ^n and not on whether ℓ^n was calculated for BPSK with one bit level or BICM with more than one bit level.

Gray Code BICM works best when the binary label ϕ is a Gray code, i.e., when any pair of neighboring symbols in $\mathcal{X}$ have labels that differ in only 1 bit level. For $M = 8$, a Gray code is listed in Table 1.2, specifically, a binary reflected Gray code (BRGC). We note that for bit level 1, we have one decision boundary, as all negative symbols have $b_1 = 0$ and all positive symbol have $b_1 = 1$. On the other hand, bit level 3 has three decision boundaries. This indicates that bit level 3 is affected more by noise than bit level 1.

Interleaver As the different bit levels have different reliability, their distribution over the codeword may affect performance. In BICM, an interleaver takes care of how bit levels map to coded bits. Here, we simply use a bit interleaver π that we sample randomly once and then leave it fixed. We discuss interleaver design in more detail in Section 1.4.3.

BICM System with Scrambled All-Zero Codeword



This diagram shows a system for simulating the WER of BICM. We note that compared to the scrambled all-zero codeword BPSK system, not much has changed. The only differences are

- The demapper function (1.33) for calculating ℓ^n, which is more complex than before.
- The interleaver π, which distributes the different bit levels uniformly over the codeword.
- The source P_X, which now samples the x_i uniformly from a 2^m-ASK constellation.
- The length of the channel input sequence, which is reduced from n to n/m, as each symbol is labelled by m bits.

Note that for $m = 1$ bit levels, we recover the BPSK system we considered before. Note that the WER is always given by (1.28), independent of the number of bit levels. We show the operating point for $\text{WER} = 1 \times 10^{-3}$ in Figure 1.3 and we also plot the BICM limit

$$\sum_{i=1}^{m} \mathbb{I}(B_i; Y). \tag{1.34}$$

We provide a derivation of (1.34) in Section 4.4.2. We observe that the gap to the BICM achievable rate is $\approx 0.6\,\text{dB}$, similar to the FEC penalty we observed for BPSK. However, the BICM achievable rate

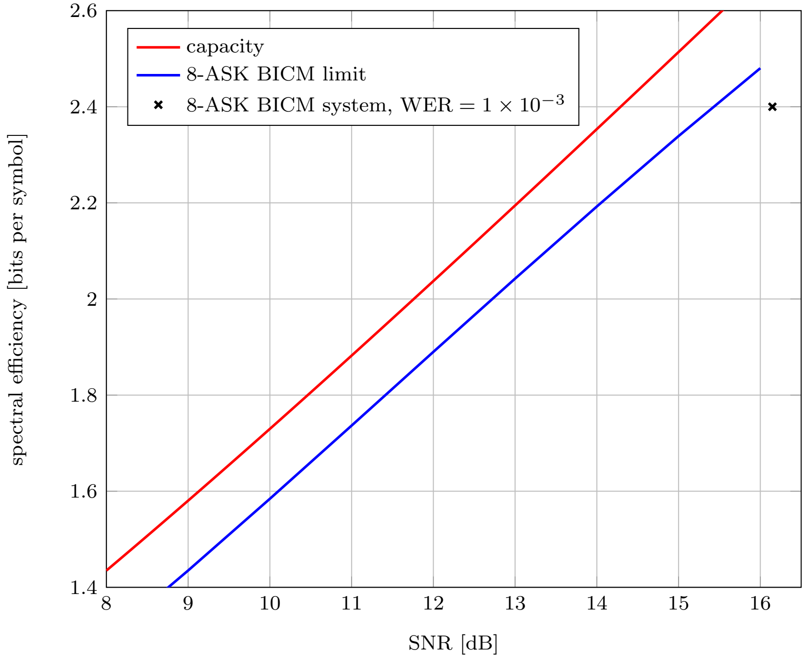


Figure 1.3: WER $= 1 \times 10^{-3}$ operating point of 8-ASK BICM.

itself has a gap of 1.2 dB to capacity. The achievable rate gap is pretty constant over the range of considered SNR values. Thus, it is unlikely to overcome this gap by using a larger constellation. Two alternative options for reducing the gap of the operating point to capacity are as follows.

1. Reduce the FEC penalty.
2. Use a non-uniform symbol distribution.

Because it is much simpler, let's focus on the second option.

1.3 Probabilistic Amplitude Shaping

Taking again a look at the BICM diagram, we note that we can change the probability distribution P_X and evaluate the WER, without affecting

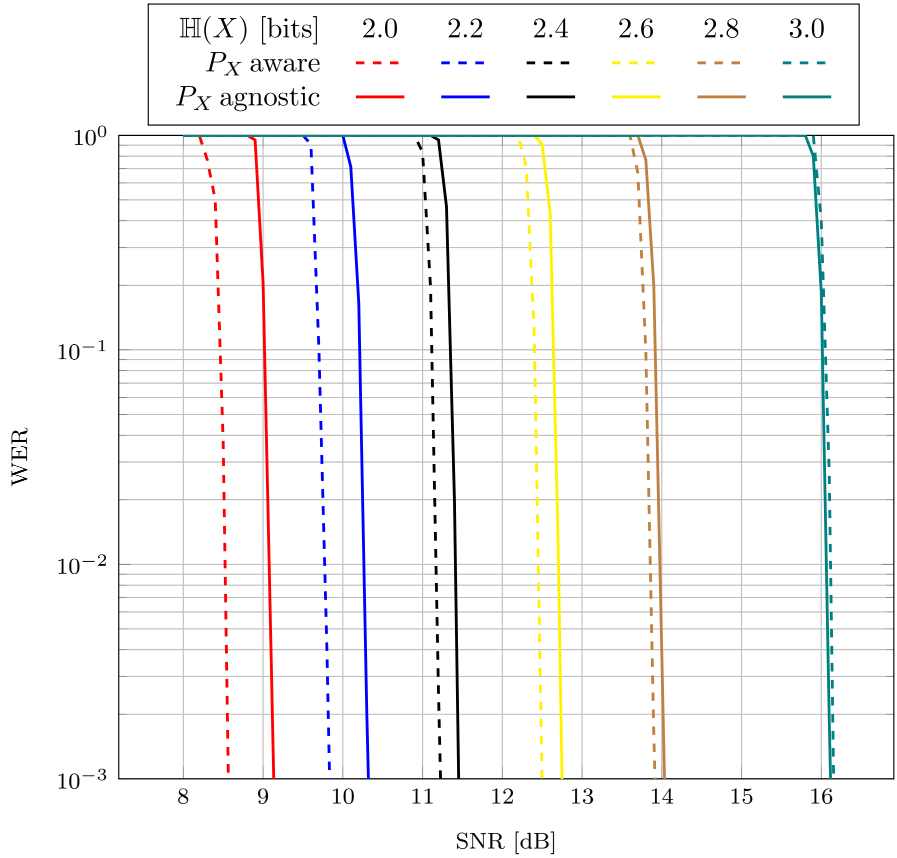


Figure 1.4: WER using demapper (1.33), agnostic of P_X, and demapper (1.36), aware of P_X.

any other part of the system. As the Gaussian density is capacity-achieving for AWGN, we choose a sampled Gaussian density, i.e.,

$$P_X(x) = \frac{e^{-\nu x^2}}{\sum_{a\in\mathcal{X}} e^{-\nu a^2}}, \quad x \in \mathcal{X}. \tag{1.35}$$

Following [51, Section IV.], we call (1.35) a Maxwell-Boltzmann (MB) distribution. The parameter $\nu \geq 0$ controls the shaping degree. For $\nu = 0$, P_X is uniform, and for $\nu \to \infty$, the probability mass concentrates on the two innermost points -1 and $+1$. We quantify the shaping degree by the entropy $\mathbb{H}(X)$ in bits. We now evaluate the WER curves for $\mathbb{H}(X) = 2.0, 2.1, \dots, 3.0$ bits.

1.3.1 WER

We observe in Figure 1.4 (solid lines) that by lowering $\mathbb{H}(X)$, the SNR required for achieving WER $= 1 \times 10^{-3}$ is also lowered, using the same FEC code and decoder. The reason is that if we fix the noise variance and we decrease the entropy $\mathbb{H}(X)$, we also decrease the transmit power and thereby the SNR, while the distance between neighboring signal points remains unchanged. Equivalently, at the same SNR, lower entropy translates into larger distance.

We note that the demapper (1.33) is not aware of the input distribution P_X. To make the prior P_X available to the decoder, we modify the demapper to

$$\ell_{ji} = \log \frac{\sum_{x\in\mathcal{X}_0^j} P_X(x) p_{Y|X}(y_i|x)}{\sum_{x\in\mathcal{X}_1^j} P_X(x) p_{Y|X}(y_i|x)}$$
$$j = 1, \dots, m, \ i = 1, \dots, n/m. \tag{1.36}$$

We display the resulting WER curves in Figure 1.4 (dashed lines). We note that the WER curves are shifted to the left and the SNR required for WER $= 1 \times 10^{-3}$ is lowered further by up to 0.6 dB.

1.3.2 Spectral Efficiency

We now would like to display the WER $= 1 \times 10^{-3}$ operating point in the SNR versus spectral efficiency (SE) plot to evaluate the gap to capacity. However, For $\mathbb{H}(X) < m$, what is the SE?

Note that as the FEC code is unchanged, the decoder still decodes against a code of rate R_{fec}, which corresponds to $R_{\text{fec}}m$ bits per symbol. For the unshaped case, the SE is mR_{fec}, which we can rewrite as

$$\text{SE} = m - m(1 - R_{\text{fec}}). \tag{1.37}$$

Here, m is the SE of an uncoded system, and $m(1 - R_{\text{fec}})$ is the FEC redundancy. For the shaped case, the uncoded SE is $\mathbb{H}(X)$ and in analogy to (1.37), we may guess the coded SE is

$$\text{SE} \stackrel{?}{=} \mathbb{H}(X) - m(1 - R_{\text{fec}}). \tag{1.38}$$

If entropy is very small, the right-hand side may become negative, which is not a meaningful value, so we modify our guess to

$$\text{SE} = [\mathbb{H}(X) - m(1 - R_{\text{fec}})]^+ . \tag{1.39}$$

In Figure 1.5 we plot required SNR versus SE assuming the correctness of (1.39). We note that below 14 dB of SNR, the curve is almost within the FEC penalty of capacity! This is an exciting observation!

1.3.3 Probabilistic Amplitude Shaping

We have to address two urgent questions:

1. How can we verify the spectral efficiency claim (1.39)?
2. How can we encode?

First, we observe that in our system, the decoder effectively decodes the binary labels of shaped symbols. Thus, we need to place the encoder between the shaped source and the channel. Using a systematic encoder, at least the information part is left unchanged by the encoder, so we may draw the following preliminary diagram.

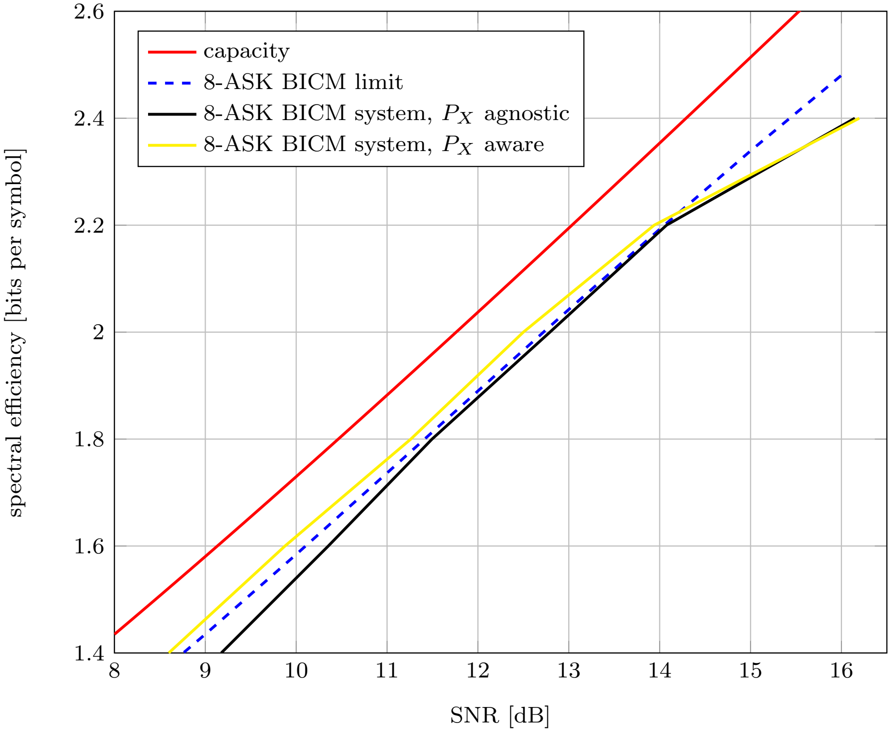
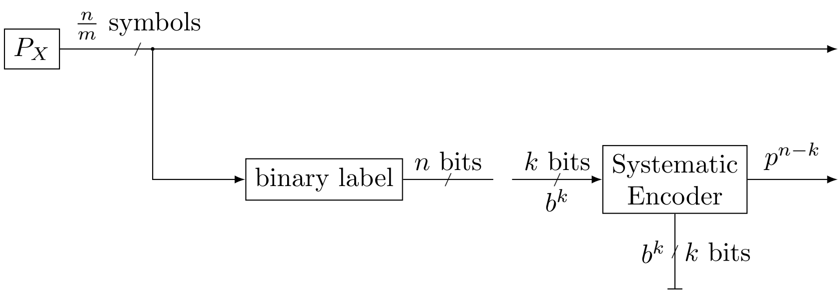
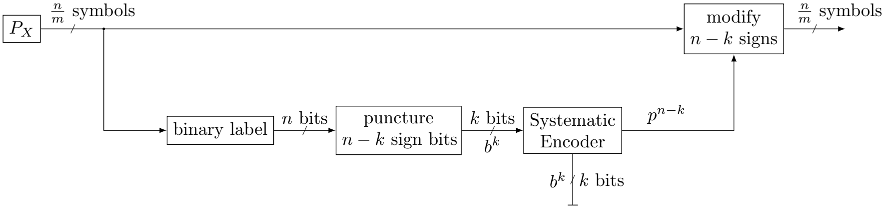


Figure 1.5: WER $= 1 \times 10^{-3}$ operating points. The P_X agnostic demapper uses (1.33) to calculate bitwise soft-decisions, while the P_X aware demapper uses (1.36).



The information bits b^k are left unchanged by systematic encoding and no further processing is required. We indicate this by the terminated encoder output in the diagram. In contrast, the parity bits p^{n-k} are newly generated by the encoder and do require further processing. This diagram has 2 issues. First, the encoder gets n bits at its input, while it only accepts k bits. Second, we must modulate the $n - k$ parity bits onto the transmitted signal somehow. The key observation is now that *we cannot impose any specific distribution onto the parity bits.* Looking at the Gray label in Table 1.2, we note that bit level 1 decides on the sign, and consequently, the transmitted power and thereby the received SNR does not depend on the distribution of bit level 1. A quick fix for the two issues is therefore as follows:

1. Mark $n - k$ sign bit positions.
2. Puncture these marked positions before the encoder, reducing the number of bits from n to k, as required.
3. Modify the $n - k$ signs corresponding to the marked positions according to the parity bits p^{n-k} output by the systematic encoder.



We are now in a position to calculate the SE. Note that for ASK constellations (1.31) the MB distribution P_X (1.35) can be factorized into amplitude A and sign S via

$$P_X(x) = P_A(|x|)P_S(\text{sign}(x)) \tag{1.40}$$

$$= P_A(|x|)\frac{1}{2}. \tag{1.41}$$

In terms of entropy, this corresponds to

$$\mathbb{H}(X) = \mathbb{H}(A) + \mathbb{H}(S) = \mathbb{H}(A) + 1. \tag{1.42}$$

The total amount of information per codeword is thus

$$\frac{n}{m}\,\mathbb{H}(A) + \left(\frac{n}{m} - (n-k)\right)\mathbb{H}(S) = \frac{n}{m}\left[\mathbb{H}(X) - \frac{(n-k)m}{n}\right] \tag{1.43}$$

$$= \frac{n}{m}\left[\mathbb{H}(X) - m(1 - R_{\text{fec}})\right] \tag{1.44}$$

which confirms the SE we postulated in (1.39). Note that on the right-hand side of (1.43), only the information bit carrying signs are counted, not the signs carrying parity bits. Thus, this SE calculation does not assume any specific distribution of the parity bits.

Having confirmed the SE, the complete PAS architecture is only a few steps away. We need to:

1. Separate the source into n/m amplitudes and $n/m-(n-k)$ signs.
2. Remove the sign puncturer.
3. Replace the sign polluter by a sign multiplexer.
4. Add an interleaver.

The diagram in Figure 1.6 shows the complete PAS architecture as we proposed it in [24].

1.4 PAS Components

At several points during the development of PAS in this section, we made design choices based on intuition, which require further study. In the following, we discuss some of them and if possible, we provide pointers to the parts of this monograph, where they are discussed in more detail.

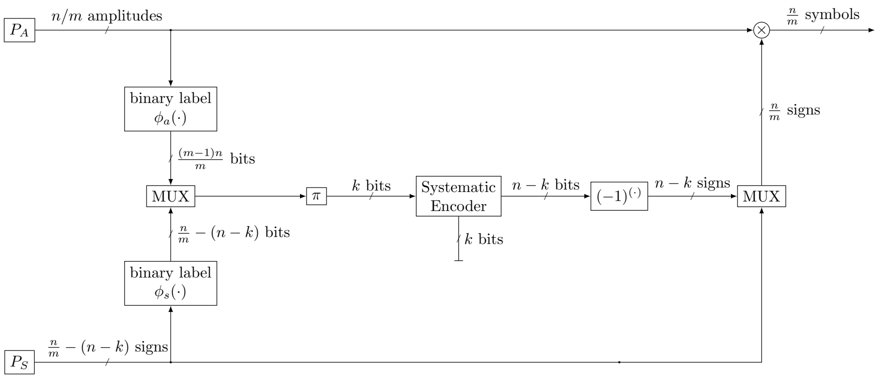


Figure 1.6: The PAS architecture as proposed in [24].

1.4.1 Distribution Matcher

A key ingredients of PAS is the amplitude source P_A, which generates amplitudes according to a desired distribution. To quantify the SE of PAS, we postulated the information content of this source to be $\mathbb{H}(A)$ bits per amplitude. In a practical system, we need to replace the amplitude source P_A by a distribution matcher (DM), which maps k uniformly distributed input bits to n amplitudes with distribution P_A. In Section 2, we study DMs in detail. The key result of Section 2 is that optimal DMs have an inherent rate loss $\mathbb{H}(A) - k/n$ that scales as $\frac{\log n}{n}$ and a cost penalty (e.g., increased average power) that also scales as $\frac{\log n}{n}$. On the downside, this requires the use of DMs that process sufficiently many amplitudes jointly. On the positive side, the rate $\mathbb{H}(A)$ can indeed be achieved, by a sufficiently long DM.

Remark 1.1. Because of the inherent rate loss, DMs operate at a rate that is *below* the entropy of the generated amplitude distribution P_A. This implies that a source decoder for a discrete memoryless source (DMS) P_A cannot be used as a DM, as it would operate at a rate *above* the entropy of P_A. We revisit this observation in Sections 2.4.2 and 2.5.6.

1.4.2 Achievable Spectral Efficiency

We postulated for PAS the SE

$$\mathrm{SE} = [\mathbb{H}(X) - m(1 - R_{\mathrm{fec}})]^{+} \tag{1.45}$$

where $m = \log_2 |\mathcal{X}|$ is the logarithmic size of the channel input alphabet $\mathcal{X}$ and where $[\cdot]^{+} = \max\{\cdot, 0\}$. In Section 3, we study what SEs are achievable by a PAS-like architecture that consists of two layers, namely the shaping layer and the FEC layer. The two layers are reflected in the achievable SEs, namely, it decomposes into two parts. The first part is the shaping set rate R_{ss}, which is bounded as

$$mR_{\mathrm{ss}} \leq \mathbb{H}(X) \tag{1.46}$$

and which can achieve this bound for sufficiently large n. The second part is the achievable FEC rate, which is given by

$$m(1 - R^{*}_{\mathrm{fec}}) = \mathbb{H}(X|Y) \tag{1.47}$$

that is, for $R_{\mathrm{fec}} < R^{*}_{\mathrm{fec}}$ and sufficiently large n, reliable communication is possible. The two parts together provide an achievable SE.

The use of a linear code is a key aspect of PAS. In Section 5, we derive an achievable SE for PAS using a random linear code. Again, this achievable SE consists of two parts. The shaping layer part is basically the rate of the employed DM (which, by Section 2, is asymptotically optimal). The FEC part recovers (1.47).

Both for the PAS-like architecture considered in Section 3 and the PAS architecture considered in Section 5, we find that $\mathbb{I}(X;Y)$ is an achievable SE, which shows that PAS is capacity-achieving for a certain class of discrete input channels.

1.4.3 Interleaver Design for Practical FEC

In our derivation of PAS we used an intra-codeword "random interleaver" (because we did not know better). In Section 3.4.3, we show that under ML-like decoding, the achievable FEC rate is invariant under intra-codeword interleaving and conclude that intra-codeword interleaver design should be considered part of practical FEC code design, accounting for suboptimal decoding. In Section 4.4.3, we revisit the interleaver question and derive the optimal decoding metric for the case when the interleaver is not known to the decoder. The design of interleavers for PAS has been considered for different families of FEC codes.

LDPC Codes When using an already designed binary LDPC code with higher order modulation, one may optimize the interleaver separately as done, e.g., in [47]. This approach was used in [16, Section V.D], [6, Section V.B], and [24, Section VIII] for optimizing the interleaver for PAS with DVB-S2 LDPC codes. In [76], [77], the joint design of LDPC codes and interleavers for PAS is considered.

Product Codes The PAS interleaver design for product codes based on algebraic component codes is considered for hard-decision decoding in [72] and for soft-decision decoding in [20].

Spatially Coupled Codes (This family of codes is known under many different names, see, e.g., [79, Section I]). PAS is combined with spatially coupled LDPC codes in [14], [15]. The PAS interleaver design for staircase codes [75] under hard-decision decoding is considered in [73]. A similar design can be used for continuously interleaved algebraic component codes under hard-decision decoding [64] and under soft-decision decoding, e.g., the oFEC code [74]. Usually, the PAS interleaver design is simpler for spatially coupled codes than for product codes.

Polar Codes The work [59] designs a PAS interleaver for polar codes, a family of FEC codes proposed in [3], [78]. The work [48] points out that polar codes inherently allow for probabilistic shaping. Various strategies for polar coding with probabilistic shaping are evaluated in [63]. We note that [63] evaluates polar coded PAS for constant composition distribution matching (CCDM). As we detail in Section 2, minimum cost distribution matcher (MCDM) performs significantly better than CCDM for short output lengths, so the comparison of [48] and polar coded PAS with MCDM is an important topic for future work.

1.4.4 Decoding Metrics

We observed that switching the demapper from calculating $\log \frac{p_{Y|B}(y|0)}{p_{Y|B}(y|1)}$ to calculating $\log \frac{P_{B|Y}(0|y)}{P_{B|Y}(1|y)}$ improved the WER of PAS. In Section 4, we derive optimal decoding metrics for several practically relevant scenarios, including bitwise demapping and hard-decision decoding.

1.4.5 Optimal Input Distribution

By our findings in Section 3 and Section 5, PAS can achieve

$$\text{SE} = \mathbb{I}(X;Y). \tag{1.48}$$

We may therefore assume that the optimal input distribution for PAS is

$$P_{X^*} = \arg\max_{P_X} \mathbb{I}(X;Y). \tag{1.49}$$

This, however, is only true if we can also choose the FEC rate freely. In practical applications, however, the FEC rate is often determined by the available FEC engine. In this case, the layered nature of the PAS architecture as reflected by the achievable SE expression needs to be taken into account. The optimization problem is then

$$\begin{aligned} \underset{P_X}{\text{maximize}} \quad & \mathbb{H}(X) && (1.50)\\ \text{subject to} \quad & \mathbb{H}(X|Y) \leq m(1-R_{\text{fec}}). && (1.51) \end{aligned}$$

This is a concave objective with a convex constraint. The Lagrangian to be maximized is

$$\mathbb{H}(X) - \lambda\,\mathbb{H}(X|Y) \tag{1.52}$$

which is the sum of a concave and a convex function. For $\lambda = 1$, fortunately, we have

$$\mathbb{H}(X) - \mathbb{H}(X|Y) = \mathbb{I}(X;Y) \tag{1.53}$$

which is known to be concave in P_X. However, for $\lambda > 1$, this may not be the case. Finding optimal distributions for FEC rate constrained PAS is interesting and of practical relevance, and we leave it for future research.

For $i = 1, \dots, n$, define

$$u_i = \log(b_i)\sqrt{a_i e^{x \log b_i}} \tag{5.88}$$

$$v_i = \sqrt{a_i e^{x \log b_i}}. \tag{5.89}$$

The numerator of the second derivative is now

$$\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{v}\boldsymbol{v}^T - (\boldsymbol{u}\boldsymbol{v}^T)^2 \tag{5.90}$$

which is non-negative, by the Cauchy-Schwarz inequality (A.1). The derivation above also holds if the sum over i is replaced by an integral over some variable τ.

5.7 Discussion

In [1], achievable rates are derived for PAS assuming a random source, in place of a DM. The work [43] analyzes PAS using typicality.

The PAS error exponent that we derived in this section has several appealing properties, for instance, it holds for linear codes, it explicitly uses a DM, and it provides an error bound for finite length. Somewhat unsatisfactory is that we had to use a CCDM so that all amplitudes have equal composition. As we have seen in Section 2, CCDM loses significantly compared to MCDM for finite length. Thus, a finite length analysis that allows for the use of an MCDM is interesting to study.

Appendices

A

Preliminaries

A.1 Mathematics

Cauchy-Schwarz Inequality

For two row vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbf{R}^M$, the Cauchy-Schwarz inequality is

$$\boldsymbol{u}\boldsymbol{u}^T \boldsymbol{v}\boldsymbol{v}^T - (\boldsymbol{u}\boldsymbol{v}^T) \geq 0 \tag{A.1}$$

with equality if and only if $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly dependent.

Big O Notation

- f is bounded below by g asymptotically:

$$f \in \Omega(g) \Leftrightarrow \liminf_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| > 0. \tag{A.2}$$

- f is bounded above by g asymptotically:

$$f \in \mathcal{O}(g) \Leftrightarrow \liminf_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| < \infty. \tag{A.3}$$

- f is bounded above and below by g asymptotically:

$$f \in \Theta(g) \Leftrightarrow f \in \Omega(g) \text{ and } f \in \mathcal{O}(g). \tag{A.4}$$

Stirling's Formula

By [36, Section II.9],

$$\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} e^{\frac{1}{12n}}. \tag{A.5}$$

Convexity

- A real-valued function f is *convex* on the interval $[A, B] \subseteq \mathbf{R}$ if for each $x_1, x_2 \in [A, B]$ and $0 \leq \lambda \leq 1$, we have

$$f[\lambda x_1 + (1-\lambda)x_2] \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

- The function f is *concave* on $[A, B]$ if $-f$ is convex on $[A, B]$.

- Let X be a random variable with support $[A, B]$. Jensen's inequality states that for f convex on $[A, B]$, we have

$$f[\mathbb{E}(X)] \leq \mathbb{E}[f(X)]. \tag{A.6}$$

For f concave on $[A, B]$, Jensen's inequality states that

$$f[\mathbb{E}(X)] \geq \mathbb{E}[f(X)]. \tag{A.7}$$

Sum-of-Products and Product-of-Sums

Consider m sets $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_m$. The Cartesian product of the m sets is the set of ordered m tuples

$$\mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_m = \{\boldsymbol{a} = (a_1, a_2, \ldots, a_m) | a_i \in \mathcal{X}_i, i = 1, 2, \ldots, m\}. \tag{A.8}$$

We now have the following sum-of-products as product-of-sums identity:

$$\sum_{\boldsymbol{a} \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_m} \prod_{j=1}^{m} a_j = \prod_{j=1}^{m} \sum_{a \in \mathcal{X}_j} a. \tag{A.9}$$

Example A.1. Consider

$$m = 2, \quad \mathcal{X}_1 = \{b, c\}, \quad \mathcal{X}_2 = \{d, e, f\}.$$

We have

$$\sum_{\boldsymbol{a}\in\mathcal{X}_1\times\mathcal{X}_2}\prod_{j=1}^{2} a_j = bd+be+bf+cd+ce+cf$$

$$\prod_{j=1}^{2}\sum_{a\in\mathcal{X}_j} = (b+c)(d+e+f) = bd+be+bf+cd+ce+cf.$$

Example A.2. We often encounter the case when $\mathcal{X}_j$ is the set of probabilities defined by a distribution P_{X_j} on an alphabet $\mathcal{X}$, i.e.,

$$\mathcal{X}_j = \{P_{X_j}(a)|a\in\mathcal{X}\}.$$

In particular, the sets $\mathcal{X}_j$ are all of the same size, i.e., $|\mathcal{X}_1| = |\mathcal{X}_2| = \cdots = |\mathcal{X}_m| = |\mathcal{X}|$. The Cartesian product of m copies of $\mathcal{X}$ is

$$\mathcal{X}^m = \underbrace{\mathcal{X}\times\mathcal{X}\times\cdots\times\mathcal{X}}_{m \text{ times}}$$

The sum-of-products as product-of-sums identity can now be written as

$$\begin{aligned}\sum_{\boldsymbol{p}\in\mathcal{X}_1\times\cdots\mathcal{X}_m}\prod_{j=1}^{m} p_j &= \sum_{\boldsymbol{a}\in\mathcal{X}^m}\prod_{j=1}^{m} P_{X_j}(a_j)\\ &= \prod_{j=1}^{m}\sum_{a\in\mathcal{X}} P_{X_j}(a).\end{aligned}$$

A.2 Probability

- **Probability distribution** P_X on discrete set $\mathcal{X}$:

$$\forall x\in\mathcal{X}\colon\ \Pr(X=x) = P_X(x). \tag{A.10}$$

- **Probability density function** (pdf) p_X on real numbers $\mathbf{R}$:

$$\forall x\in\mathbf{R}\colon\ \Pr(X\le x) = \int_{-\infty}^{x} p_X(\tau)d\tau. \tag{A.11}$$

- **Markov's inequality, [38, Section 1.6.1]:** Let X be a non-negative random variable, i.e., $\Pr(X<0)=0$. Then for $a>0$

$$\Pr(X\ge a)\le\frac{\mathbb{E}(X)}{a}. \tag{A.12}$$

- **Moments:** Real-valued random variable X, positive integer k.

$$\mathsf{mgf}_X(r) = \mathbb{E}(e^{rX}) \tag{A.13}$$

$$\frac{\partial^k}{\partial r^k}\mathsf{mgf}_X(r)\bigg|_{r=0} = \mathbb{E}(X^k). \tag{A.14}$$

$\mathsf{mgf}_X(r)$ is the moment generating function (MGF) of X and $\mathbb{E}(X^k)$ is the kth *moment* of X.

A.3 Information Theory

A.3.1 Types and Typical Sequences

Types Consider a sequence $x^n = x_1 x_2 \cdots x_n$ with entries in a finite alphabet $\mathcal{X}$. Let $N(a|x^n)$ be the number of times letter $a \in \mathcal{X}$ occurs in x^n, i.e.,

$$N(a|x^n) = \left|\left\{i \in \{1, 2, \dots, n\} : x_i = a\right\}\right|, \quad a \in \mathcal{X}. \tag{A.15}$$

The empirical distribution of x^n is

$$P_{x^n}(a) = \frac{N(a|x^n)}{n}, \quad a \in \mathcal{X}. \tag{A.16}$$

Since every permutation of x^n has the same empirical distribution, we define $n_a = N(a|x^n)$ and write

$$P_X(a) = \frac{n_a}{n}, \quad a \in \mathcal{X}. \tag{A.17}$$

Note that every probability $P_X(a)$, $a \in \mathcal{X}$, is an integer multiple of $1/n$. The distribution P_X is therefore called an n-type. The set of all length n sequences with empirical distribution P_X is called the type class of the n-type P_X and denoted by $\mathcal{T}^n(P_X)$.

A.3.2 Differential Entropy

- **Differential entropy:**

$$\mathrm{h}(X) := \mathbb{E}[-\log_2 p_X(X)]. \tag{A.18}$$

- **Independence bound:**

$$\mathrm{h}(X, Y) \leq \mathrm{h}(X) + \mathrm{h}(Y). \tag{A.19}$$

A.3.3 Entropy

Random variable X with distribution P_X on finite set $\mathcal{X}$.

- **Entropy:**

$$\mathbb{H}(P_X) = \mathbb{H}(X) := \mathbb{E}[-\log_2 P_X(X)]. \tag{A.20}$$

- **Conditional Entropy, Equivocation:**

$$\mathbb{H}(P_{X|Y}|P_Y) = \mathbb{H}(X|Y) := \mathbb{E}[-\log_2 P_{X|Y}(X|Y)]. \tag{A.21}$$

- **Relation to differential entropy:** Properties (A.19) also hold for entropy.

- **Continuity:** Distributions P_X, $P_{X'}$ on finite set $\mathcal{X}$. Suppose $\|P_X - P_{X'}\|_1 = \delta \leq \frac{1}{2}$. Then

$$|\,\mathbb{H}(P_X) - \mathbb{H}(P_{X'})| \leq -\delta \log_2 \frac{\delta}{|\mathcal{X}|}. \tag{A.22}$$

- **Cross-Entropy:** P_X, Q_X distributions on $\mathcal{X}$.

$$\mathbb{X}(P_X \| Q_X) = \mathbb{E}[-\log_2 Q_X(X)]. \tag{A.23}$$

- **Information inequality:**

$$\mathbb{X}(P_X \| Q_X) \geq \mathbb{H}(P_X) \tag{A.24}$$

  with equality if and only if $Q_X = P_X$.

- **Cross-Equivocation:** $P_{X|Y}(\cdot|b)$ distribution on $\mathcal{X}$ for each $b \in \mathcal{Y}$. $Y \sim p_Y$.

  - $Q_{X|Y}(\cdot|b)$ distribution on $\mathcal{X}$ for each $b \in \mathcal{Y}$.

$$\mathbb{X}(P_{X|Y} \| Q_{X|Y} | p_Y) = \mathbb{E}[-\log_2 Q_{X|Y}(X|Y)]. \tag{A.25}$$

  - $q(\cdot,\cdot)$ non-negative function on $\mathcal{X} \times \mathcal{Y}$.

$$\mathsf{X}(q, X, Y) = \mathbb{E}\left[-\log_2 \frac{q(X,Y)}{\sum_{a \in \mathcal{X}} q(a,Y)}\right]. \tag{A.26}$$

A.3.4 Informational Divergence

- **Informational divergence:**

$$\mathbb{D}(p_X \| p_Y) := \mathbb{E}\left[\log_2 \frac{p_X(X)}{p_Y(X)}\right] \tag{A.27}$$

- **Information inequality:**

$$\mathbb{D}(p_X \| p_Y) \geq 0 \tag{A.28}$$

with equality if and only if $p_X = p_Y$.

A.3.5 Mutual Information

- **Mutual Information:**
  - X, Y continuous:

$$\begin{aligned} \mathbb{I}(X;Y) &:= \mathbb{D}(p_{XY} \| p_X p_Y) && \text{(A.29)} \\ &= \mathbb{D}(p_{Y|X} \| p_Y | p_X) && \text{(A.30)} \\ &= \mathbb{D}(p_{X|Y} \| p_X | p_Y) && \text{(A.31)} \\ &= \mathrm{h}(Y) - \mathrm{h}(Y|X) && \text{(A.32)} \\ &= \mathrm{h}(X) - \mathrm{h}(X|Y). && \text{(A.33)} \end{aligned}$$

  - X discrete, Y continuous:

$$\begin{aligned} \mathbb{I}(X;Y) &:= \mathbb{D}(P_X p_{Y|X} \| P_X p_Y) && \text{(A.34)} \\ &= \mathbb{D}(P_{X|Y} \| P_X | p_Y) && \text{(A.35)} \\ &= \mathbb{D}(p_{Y|X} \| p_Y | P_X) && \text{(A.36)} \\ &= \mathrm{h}(Y) - \mathrm{h}(Y|X) && \text{(A.37)} \\ &= \mathbb{H}(X) - \mathbb{H}(X|Y). && \text{(A.38)} \end{aligned}$$

  - Other combinations of discrete/continuous accordingly.

B

Acronyms

ASK amplitude shift keying

AWGN additive white Gaussian noise

BER bit error rate

BIACM bit-interleaver-agnostic coded modulation

BICM bit-interleaved coded modulation

BMD bit-metric decoding

BPSK binary phase shift keying

BRGC binary reflected Gray code

BSC binary symmetric channel

CCDM constant composition distribution matching

DM distribution matcher

DMS discrete memoryless source

FEC forward error correction

GHC geometric Huffman coding

GMI generalized mutual information

IACM interleaver-agnostic coded modulation

ID informational divergence

iid independent and identically-distributed

ILD invertible low-divergence

LDPC low-density parity-check

LUT lookup table

MAP maximum a posteriori probability

MB Maxwell-Boltzmann

MCDM minimum cost distribution matcher

MGF moment generating function

ML maximum-likelihood

PAS probabilistic amplitude shaping

PS probabilistic shaping

QAM quadrature amplitude modulation

SD soft decision

SE spectral efficiency

SNR signal-to-noise ratio

VD variational distance

WER word error rate

References

[1] R. A. Amjad, "Information rates and error exponents for probabilistic amplitude shaping," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Guangzhou, China, 2018.

[2] R. A. Amjad and G. Böcherer, "Fixed-to-variable length distribution matching," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1511–1515, Istanbul, Turkey, 2013.

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

[4] S. Baur and G. Böcherer, "Arithmetic distribution matching," in *Proc. Int. ITG Conf. Syst. Commun. Coding (SCC)*, pp. 1–6, Hamburg, Germany, 2015.

[5] G. Böcherer, "Capacity-achieving probabilistic shaping for noisy and noiseless channels," Ph.D. dissertation, RWTH Aachen University, 2012. URL: http://www.georg-boecherer.de/capacityAchievingShaping.pdf.

[6] G. Böcherer, "Labeling non-square QAM constellations for one-dimensional bit-metric decoding," *IEEE Commun. Lett.*, vol. 18, no. 9, 2014, pp. 1515–1518.

[7] G. Böcherer, "Achievable rates for shaped bit-metric decoding," *arXiv preprint*, 2016. URL: http://arxiv.org/abs/1410.8075.

[8] G. Böcherer, "Principles of coded modulation," Habilitation thesis, Technical University of Munich, 2018. URL: http://www.georg-boecherer.de/bocherer2018principles.pdf.

[9] G. Böcherer and R. A. Amjad, "Fixed-to-variable length resolution coding for target distributions," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Seville, Spain, 2013.

[10] G. Böcherer and R. A. Amjad, "Informational divergence and entropy rate on rooted trees with probabilities," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 176–180, Honolulu, HI, USA, 2014.

[11] G. Böcherer and B. C. Geiger, "Optimal quantization for distribution synthesis," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, 2016, pp. 6162–6172.

[12] G. Böcherer, P. Schulte, and F. Steiner, "High throughput probabilistic shaping with product distribution matching," *arXiv preprint*, 2017. URL: http://arxiv.org/abs/1702.07510.

[13] G. Böcherer, P. Schulte, and F. Steiner, "Probabilistic shaping and forward error correction for fiber-optic communication systems," *J. Lightw. Technol.*, vol. 37, no. 2, 2019, pp. 230–244.

[14] G. Böcherer, F. Steiner, and P. Schulte, "Fast probabilistic shaping implementation for long-haul fiber-optic communication systems," in *Proc. Eur. Conf. Optical Commun. (ECOC)*, Gothenburg, Sweden, 2017.

[15] G. Böcherer, F. Steiner, and P. Schulte, "Fast probabilistic shaping implementation for long-haul fiber-optic communication systems," in *Proc. Eur. Conf. Optical Commun. (ECOC)*, Gothenburg, Sweden, 2017. URL: http://www.georg-boecherer.de/bocherer2017fast_slides.pdf.

[16] G. Böcherer, "Probabilistic signal shaping for bit-metric decoding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 431–435, Honolulu, HI, USA, 2014.

[17] G. Böcherer, "Lecture notes on variable length coding," 2016. URL: http://www.georg-boecherer.de/bocherer2016variable.pdf.

[18] G. Böcherer, "Integration of probabilistic shaping and forward error correction: Spectral efficiency, rate, overhead," in *CNRS/GdR ISIS Workshop on Coding, Modulation, and Signal Processing for Optical Communications, Telecom Paris Tech*, Paris, France, 2019. URL: http://www.georg-boecherer.de/bocherer2019integration_slides.pdf.

[19] G. Böcherer, "Achievable rates for probabilistic shaping," *arXiv preprint*, URL: https://arxiv.org/abs/1707.01134v5.

[20] G. Böcherer, F. Diedolo, and F. Pittala, "Label extension for 32QAM: The extra bit for a better FEC performance-complexity tradeoff," in *Proc. Eur. Conf. Optical Commun. (ECOC)*, Brussels, Belgium, 2020.

[21] G. Böcherer, D. Lentner, A. Cirino, and F. Steiner, "Probabilistic parity shaping for linear codes," *arXiv preprint*, 2019. URL: https://arxiv.org/abs/1902.10648.

[22] G. Böcherer and R. Mathar, "Matching dyadic distributions to channels," in *Proc. Data Compression Conf. (DCC)*, pp. 23–32, Snowbird, UT, USA, 2011.

[23] G. Böcherer, P. Schulte, and F. Steiner, "Probabilistic shaping: A random coding experiment," in *Proc. Int. Zurich Seminar Commun.*, ETH Zurich, pp. 12–14, 2020.

[24] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth efficient and rate-matched low-density parity-check coded modulation," *IEEE Trans. Commun.*, vol. 63, no. 12, 2015, pp. 4651–4665.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, 2004.

[26] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, "Rate adaptation and reach increase by probabilistically shaped 64-QAM: An experimental demonstration," *J. Lightw. Technol.*, vol. 34, no. 8, 2016.

[27] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, 1998, pp. 927–946.

[28] J. Cho, L. Schmalen, and P. J. Winzer, "Normalized generalized mutual information as a forward error correction threshold for probabilistically shaped QAM," in *Proc. Eur. Conf. Optical Commun. (ECOC)*, Gothenburg, Sweden, 2017.

[29] J. Cho, "Prefix-free code distribution matching for probabilistic constellation shaping," *IEEE Trans. Commun.*, vol. 68, no. 2, 2020, pp. 670–682.

[30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., 2006.

[31] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge University Press, 2011.

[32] I. Csiszár and P. C. Shields, "Information theory and statistics: A tutorial," *Found. Trends Comm. Inf. Theory*, vol. 1, no. 4, 2004, pp. 417–528.

[33] M. Dia, V. Aref, and L. Schmalen, "A compressed sensing approach for distribution matching," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1266–1270, Vail, Colorado, USA, 2018.

[34] T. Fehenberger, D. S. Millar, T. Koike-Akino, K. Kojima, and K. Parsons, "Multiset-partition distribution matching," *IEEE Trans. Commun.*, vol. 67, no. 3, 2019, pp. 1885–1893.

[35] T. Fehenberger, D. S. Millar, T. Koike-Akino, K. Kojima, and K. Parsons, "Parallel-amplitude architecture and subset ranking for fast distribution matching," *IEEE Trans. Commun.*, vol. 68, no. 4, 2020, pp. 1981–1990.

[36] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume I.* John Wiley & Sons, Inc, 1968.

[37] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, 1962, pp. 21–28.

[38] R. G. Gallager, *Stochastic processes: theory for applications.* Cambridge University Press, 2013.

[39] R. G. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, Inc., 1968.

[40] A. Ganti, A. Lapidoth, and E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, 2000, pp. 2315–2328.

[41] B. C. Geiger and G. Böcherer, "Greedy algorithms for optimal distribution approximation," *Entropy*, vol. 18, no. 7, 2016, pp. 1–10. DOI: 10.3390/e18070262.

[42] Y. C. Gültekin, "Enumerative sphere shaping techniques for short blocklength wireless communications," Ph.D. dissertation, Technische Universiteit Eindhoven, 2020.

[43] Y. C. Gültekin, A. Alvarado, and F. M. Willems, "Achievable information rates for probabilistic amplitude shaping: An alternative approach via random sign-coding arguments," *Entropy*, vol. 22, no. 7, 2020.

[44] Y. C. Gültekin, T. Fehenberger, A. Alvarado, and F. M. Willems, "Probabilistic shaping for finite blocklengths: Distribution matching and sphere shaping," *Entropy*, vol. 22, no. 5, 2020, p. 581.

[45] Y. C. Gültekin, W. J. van Houtum, S. Şerbetli, and F. M. Willems, "Constellation shaping for IEEE 802.11," in *Proc. IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Montreal, Quebec, Canada, 2017.

[46] Y. C. Gültekin, F. M. Willems, W. van Houtum, and S. Serbetli, "Approximate enumerative sphere shaping," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 676–680, Vail, Colorado, USA, 2018.

[47] C. Häger, A. Graell i Amat, F. Brännström, A. Alvarado, and E. Agrell, "Improving soft FEC performance for higher-order modulations via optimized bit channel mappings," *Optics Express*, vol. 22, no. 12, 2014, pp. 14 544–14 558.

[48] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric models," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, 2013, pp. 7829–7838.

[49] G. Kaplan and S. Shamai (Shitz), "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment," *AEÜ*, vol. 47, no. 4, 1993, pp. 228–239.

[50] G. Kramer, "Divergence scaling for distribution matching," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1159–1163, Melbourne, Australia, 2021.

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

[52] A. Lempel, S. Even, and M. Cohn, "An algorithm for optimal prefix parsing of a noiseless and memoryless channel," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, 1973, pp. 208–214.

[53] Y. Lomnitz and M. Feder, "A simpler derivation of the coding theorem," *arXiv preprint arXiv:1205.1389*, 2012.

[54] D. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, 1999, pp. 399–431.

[55] A. Martinez, A. Guillén i Fàbregas, G. Caire, and F. Willems, "Bit-interleaved coded modulation revisited: A mismatched decoding perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, 2009, pp. 2756–2765.

[56] N. Merhav and G. Böcherer, "Codebook mismatch can be fully compensated by mismatched decoding," *IEEE Trans. Inf. Theory*, vol. 69, no. 4, 2023, pp. 2152–2164.

[57] M. Pikus, "Finite-precision and multi-stream distribution matching," Ph.D. dissertation, Technical University of Munich, 2019.

[58] M. Pikus and W. Xu, "Bit-level probabilistically shaped coded modulation," *IEEE Commun. Lett.*, vol. 21, no. 9, 2017, pp. 1929–1932.

[59] T. Prinz, P. Yuan, G. Böcherer, F. Steiner, O. Iscan, R. Böhnke, and W. Xu, "Polar coded probabilistic amplitude shaping for short packets," in *IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, Sapporo, Japan, 2017.

[60] T. V. Ramabadran, "A coding scheme for m-out-of-n codes," *IEEE Trans. Commun.*, vol. 38, no. 8, 1990, pp. 1156–1163.

[61] E. A. Ratzer, "Error-correction on non-standard communication channels," Ph.D. dissertation, University of Cambridge, 2003.

[62] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, 2001, pp. 619–637.

[63] C. Runge, T. Wiegart, and D. Lentner, "Improved list decoding for polar-coded probabilistic shaping," *arXiv preprint*, 2023. URL: https://arxiv.org/abs/2305.07962v1.

[64] M. Scholten, T. Coe, and J. Dillard, "Continuously-interleaved BCH (CI-BCH) FEC delivers best in class NECG for 40G and 100G metro applications," in *National Fiber Optic Engineers Conference*, NTuB3, 2010.

[65] P. Schulte and G. Böcherer, "Constant composition distribution matching," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, 2016, pp. 430–434.

[66] P. Schulte, "Algorithms for distribution matching," Ph.D. dissertation, Technische Universität München, 2020.

[67] P. Schulte, R. A. Amjad, T. Wiegart, and G. Kramer, "Invertible low-divergence coding," *IEEE Trans. Inf. Theory*, vol. 68, no. 1, 2021, pp. 178–192.

[68] P. Schulte and B. C. Geiger, "Divergence scaling of fixed-length, binary-output, one-to-one distribution matching," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 3075–3079, Aachen, Germany, 2017.

[69] P. Schulte and F. Steiner, "Divergence-optimal fixed-to-fixed length distribution matching with shell mapping," *IEEE Wireless Commun. Letters*, vol. 8, no. 2, 2019, pp. 620–623.

[70] P. Schulte, F. Steiner, and G. Böcherer, "Four dimensional probabilistic shaping for fiber-optic communication," in *Proc. Signal Process. Photonic Commun. (SPPCOM)*, New Orleans, Louisiana, USA, 2017.

[71] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, 1948, 379–423 and 623–656.

[72] A. Sheikh, A. Graell i Amat, and A. Alvarado, "On product codes with probabilistic amplitude shaping for high-throughput fiber-optic systems," *IEEE Commun. Lett.*, vol. 24, no. 11, 2020, pp. 2406–2410.

[73] A. Sheikh, A. Graell i Amat, G. Liva, and F. Steiner, "Probabilistic amplitude shaping with hard decision decoding and staircase codes," *J. Lightw. Technol.*, vol. 36, no. 9, 2018, pp. 1689–1697.

[74] M. A. Sluyski, *Open ROADM MSA 3.01 W-port digital specification (200G-400G)*, 2019. URL: https://tinyurl.com/openroadm.

[75] B. P. Smith, A. Farhood, A. Hunt, F. R. Kschischang, and J. Lodge, "Staircase codes: FEC for 100 Gb/s OTN," *J. Lightw. Technol.*, vol. 30, no. 1, 2012, pp. 110–117.

[76] F. Steiner, G. Böcherer, and G. Liva, "Protograph-based LDPC code design for shaped bit-metric decoding," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, 2016, pp. 397–407.

[77] F. Steiner, "Coding for higher-order modulation and probabilistic shaping," Ph.D. dissertation, Technical University of Munich, 2020.

[78] N. Stolte, "Recursive codes with the Plotkin construction and their decoding," Ph.D. dissertation, Technical University of Darmstadt, 2002.

[79] A. Y. Sukmadji, U. Martínez-Peñas, and F. R. Kschischang, "Zipper codes," *J. Lightw. Technol.*, vol. 40, no. 19, 2022, pp. 6397–6407.

[80] T. Yoshida, M. Karlsson, and E. Agrell, "Performance Metrics for Systems With Soft-Decision FEC and Probabilistic Shaping," *IEEE Photon. Technol. Lett.*, vol. 29, no. 23, 2017, pp. 2111–2114.

[81] T. Yoshida, M. Karlsson, and E. Agrell, "Hierarchical distribution matching for probabilistically shaped coded modulation," *J. Lightw. Technol.*, vol. 37, no. 6, 2019, pp. 1579–1589.