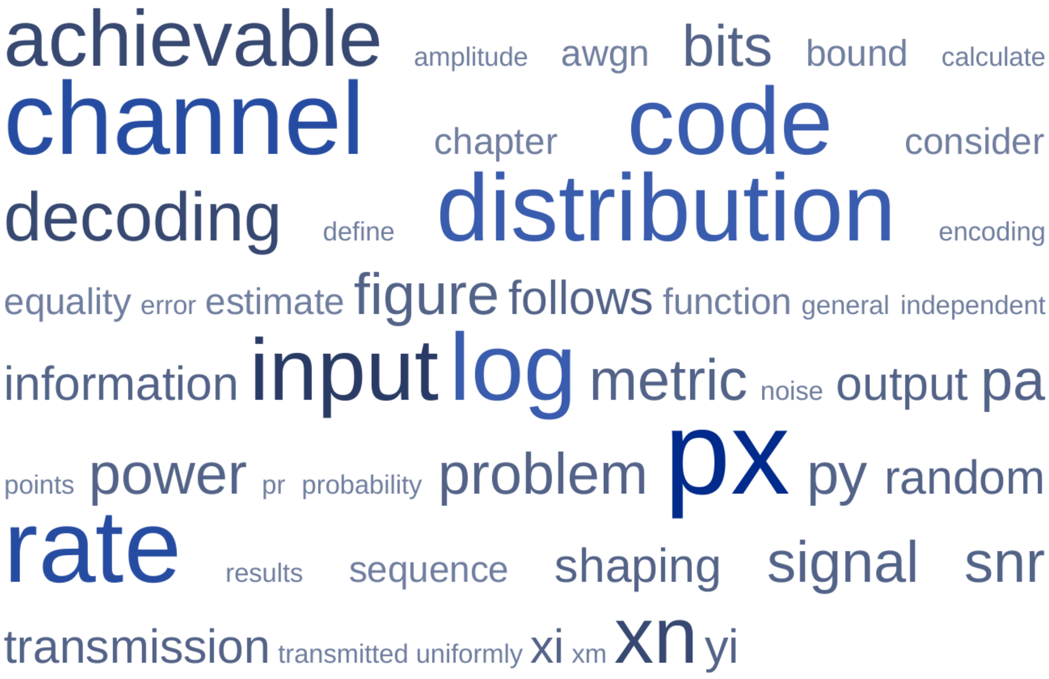
Principles of Coded Modulation

Georg Böcherer



Contents

Preface

In 1974, James L. Massey observed in his work “Coding and Modulation in Digital Communications” [1], that

> *From the coding viewpoint, the modulator, waveform channel, and demodulator together constitute a discrete channel with q input letters and q' output letters.*

He then concluded

> *...that the real goal of the modulation system is to create the “best” DMC [discrete memoryless channel] as seen by the coding system.*

These notes develop tools for designing modulation systems in Massey’s spirit.

G. Böcherer, February 2, 2018

1. Introduction

The global society currently faces a rapid growth of data traffic in the internet, which will continue in the coming decades. This puts tremendous pressure on telecommunication companies, which need innovations to provide the required digital link capacities.

In fiber-optic communication, mobile communication, and satellite communication, the data rate requirements can be met only if the digital transmitters map more than one bit to each real dimension of each time-frequency slot. This requires higher-order modulation where the digital transceivers must handle more than two signal points per real dimension.

The laws of physics dictate that by amplifying the received signal, the receiver inevitably generates noise that is unpredictable. This noise corrupts the received signal, and error correcting codes are needed to guarantee reliable communication.

Coded modulation is the combination of higher-order modulation with error correction. In these notes, basic principles of coded modulation are developed.

Chapter 2 reviews the components and figures-of-merit of digital communication systems, which are treated in detail in many text books. A good reference is [2]. The chapters 3–4 are on the discrete time additive white Gaussian noise (AWGN) channel. Using the properties of information measures developed, e.g., in [3]–[5], we analyze the effect of signal constellations and input distributions on achievable rates for reliable transmission. We observe that a discrete constellation with enough signal points and optimized distribution enables reliable transmission close to the AWGN capacity. We then use this insight in Chapter 6 to develop probabilistic amplitude shaping (PAS), a transceiver architecture to combine non-uniform input distributions with forward error correction (FEC).

In the chapters 7–10, we use information-theoretic arguments for finding achievable rates and good transceiver designs. Our derivations are guided by two main goals:

- The achievable rates are based on transceiver designs that are implementable in practice.
- The achievable rates are applicable for practical channels that we can measure, but for which we do not have a tractable analytical description.

In Chapter 11, we develop techniques to estimate achievable rates from channel measurements.

2. Digital Communication System

2.1. Transmission System

A simple model of a digital transmission system is shown in Figure 2.1. At the transmitter, data bits are mapped to a sequence of signal points, which are then transformed into a waveform. The channel distorts the transmitted waveform. At the receiver, the received waveform is transformed into a sequence of noisy signal points. This sequence is then transformed into a sequence of detected data bit. The signal points and the noisy signal points, respectively, form the interface between continuous-time and discrete-time signal processing in the communication system. Modeling and analysis of the continuous-time part and the conversion between discrete-time signals and continuous-time signals is discussed in many textbooks. For example, an introduction to the representation of continuous waveforms is given in [2, Chapter 4]. The conversion of continuous waveforms into discrete-time sequences is discussed in [6, Chapter 4].

We treat the time-continuous part of the digital transmission system as part of a channel with discrete-time input (the signal points) and discrete-time output (the noisy signal points).

2.2. Figures of Merit

To design a transmission system, we must identify the figures of merit that we are interested in. Common properties of interest are as follows.

- *Rate:* We want to transmit the data over the channel as fast as possible.
- *Reliability:* We want to recover the transmitted data correctly at the receiver.
- *Delay:* We want to be "real-time".
- *Power:* We want to use as low power as possible/we want to comply with legal power restrictions.
- *Energy:* We want to spend a small amount of energy to transmit a data packet.

To quantify the properties of interest, we consider *block-based transmission*, i.e., we jointly consider n consecutive channel uses. See Figure 2.2 for an illustration. The parameter n is called the *block length* or *number of channel uses*. Rate, reliability, delay, power, and energy can now be quantified as follows.

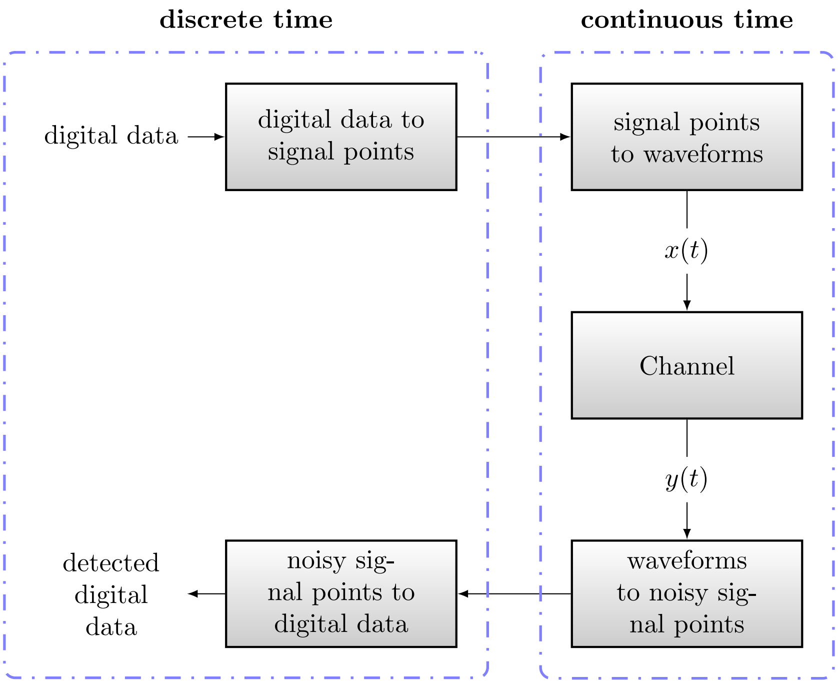

discrete time
continuous time
digital data
digital data to signal points
signal points to waveforms
$x(t)$
Channel
$y(t)$
waveforms to noisy signal points
noisy signal points to digital data
detected digital data

Figure 2.1.: A generic model of a digital communication system.

Rate The *rate* is defined as

$$R = \frac{\text{number of transmitted bits}}{\text{number of channel uses}}. \tag{2.1}$$

Since we consider n channel uses, the number of bits that we transmit per block is nR. This bit string can take 2^{nR} different values. We frequently model the data by a random variable W with alphabet $\{1, 2, \dots, 2^{nR}\}$. We call W the *message*.

Reliability At the transmitter, we encode the message W to a signal $X^n = (X_1, X_2, \dots, X_n)$. At the receiver, we choose a $\hat{W}$ from the hypotheses $1, 2, \dots, 2^{nR}$ according to some decision rule based on our observation Y^n of the signal. We loosen the requirement of "correctly recovering the data at the receiver" to "correct recovery with high probability" and we measure reliability by the *block error probability* that our decision $\hat{W}$ is equal to the transmitted message W, i.e.,

$$P_e := \Pr(W \neq \hat{W}). \tag{2.2}$$

Delay For optimal detection, before we have a decision for the first bit of the nR bits represented by W, we need to wait for the nth signal point Y_n. The delay is therefore the duration of n channel uses. (A practical system needs additional time for processing).

Power With the real signal point x, we associate the *power* x^2, see Figure 2.3 for why this makes sense. We can define the power in different ways.

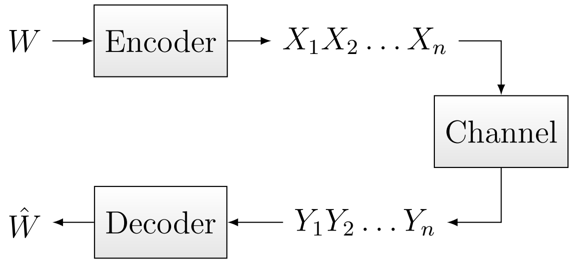


Figure 2.2.: Block-based transmission over a discrete-time channel. The message W can take the values $1, 2, \ldots, 2^{nR}$, which can be represented by nR bits. The message W is encoded to a discrete signal $X^n = x^n(W)$. The decoder detects the message from the observed channel output Y^n and outputs a decision $\hat{W}$.

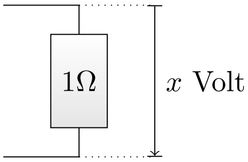


Figure 2.3.: We can think of the input x as the voltage over a unit resistance. Since power = current × voltage, we have the power $= x^2/1$ Watts by Ohm's law.

- We say a system has *peak power* at most P if

$$x^2 \leq \mathsf{P}, \quad \text{for each transmitted signal point } x. \tag{2.3}$$

- A system has *per-block power* at most P if

$$\frac{\sum_{i=1}^n x_i^2}{n} \leq \mathsf{P}, \quad \text{for each transmitted signal } x^n. \tag{2.4}$$

- A system has *average power* at most P if the *average* per-block-power is upper bounded by P, i.e., if

$$\frac{\mathbb{E}\left[\sum_{i=1}^n X_i^2\right]}{n} \leq \mathsf{P}. \tag{2.5}$$

Note that the three power constraints are ordered, namely (2.3) implies (2.4), and (2.4) implies (2.5). In particular, requiring the per-block power to be smaller than P is more restrictive than requiring an average power to be smaller than P. The most important notion of power is the average power. In the following, we will use the terms “power” and “average power” interchangeably.

Example 2.1 (BPSK)**.** Suppose the transmitted signal points take one of the two values $\{-A, A\}$. This modulation scheme is called binary phase shift keying (BPSK). For BPSK the peak power, the per-block power, and the average power are all equal to A^2, independent of the block length n.

Example 2.2 (4-ASK)**.** Suppose the transmitted signal points take on values in

$$\{-3A, -A, A, 3A\}$$

with equal probability. This modulation scheme is usually called amplitude shift keying (ASK) although both amplitude and phase are modulated. The peak power of this scheme is $9A^2$. The signal with highest per-block power is the one that has all signal points equal to $3A$ or $-3A$. The average power of the considered scheme is

$$\frac{1}{2}A^2 + \frac{1}{2}9A^2 = 5A^2 \tag{2.6}$$

which is significantly less than the peak power and the highest per-block power.

Example 2.3 (Block-Based Transmission)**.** Suppose the block length is $n = 4$ and the signal points take values in a 3-ASK constellation $\mathcal{X} = \{-1, 0, 1\}$. Suppose further that the message W is uniformly distributed on $\mathcal{W} = \{1, 2, \dots, 8\}$ and that the encoder $f\colon \mathcal{W} \to \mathcal{C} \subseteq \{-1, 0, 1\}^n$ is given by

$$\begin{aligned}
1 &\mapsto (-1, 0, 0, 1)\\
2 &\mapsto (-1, 1, 0, 0)\\
3 &\mapsto (0, -1, 0, 1)\\
4 &\mapsto (0, 0, -1, 1)\\
5 &\mapsto (0, 1, -1, 0)\\
6 &\mapsto (0, 1, 0, -1)\\
7 &\mapsto (1, -1, 0, 0)\\
8 &\mapsto (1, 0, 0, -1).
\end{aligned}$$

The rate of this scheme is

$$R = \frac{3}{4} \quad \left[\frac{\text{bits}}{\text{channel use}}\right]. \tag{2.7}$$

The peak power is 1, the per-block power is 0.5 for each code word in $\mathcal{C}$ and consequently, the average power is also equal to $\frac{1}{2}$. Let $X_1X_2X_3X_4 := f(W)$ be the transmitted signal. The distribution of X_i, $i = 1, 2, 3, 4$ is

$$P_{X_i}(-1) = P_{X_i}(1) = \frac{1}{4}, \quad P_{X_i}(0) = \frac{1}{2}. \tag{2.8}$$

Energy We quantify energy by E_b, which is the energy that we spend to transmit one bit of data. We have

$$E_b = \frac{\text{energy}}{\text{bit}} = \frac{\text{energy}}{\text{channel use}} \cdot \frac{\text{channel use}}{\text{bit}} = \frac{\text{power}}{\text{rate}} = \frac{\mathsf{P}}{R}. \tag{2.9}$$

2.3. Data Interface

The data to be transmitted is often modelled by a sequence of independent and uniformly distributed bits. For block-based transmission, this binary stream is partitioned into chunks of nR bits each, which represent the messages that we transmit per block. Each message is usually modelled as uniformly distributed. This is a good model for many kinds of digital data, especially when the data is in a compressed format such as .zip or .mp3. Separating the real world data from the digital transmission system by a *binary interface* is often referred to by *source/channel separation*. This principle allows to separately design source encoders (i.e., compression algorithms) and transmission systems. The separation of source encoding and transmission implies virtually no loss in performance, see for example [4, Section 7.13] and [2, Chapter 1].

2.4. Capacity

We now want to relate reliability, power, and rate. Consider the transmission system in Figure 2.2. It consists of an encoder that maps the nR bit message W to the signal X^n and a decoder that maps the received signal Y^n to a decision $\hat{W}$.

Definition 1 (Achievable)**.** We say the rate R is *achievable* under the power constraint P, if for each $\epsilon > 0$ and a large enough block length $n(\epsilon)$, there exists a transmission system with $2^{n(\epsilon)R}$ code words with an average power smaller or equal to P and a probability of error smaller than ϵ, i.e.,

$$P_e = \Pr(W \neq \hat{W}) < \epsilon.$$

The *capacity* of a channel is the supremum of all achievable rates. For some channels, the channel capacity can be calculated in closed form. For example, for discrete memoryless channels with input-output relation

$$P_{Y^n|X^n}(b^n|x^n) = \prod_{i=1}^{n} P_{Y|X}(b_i|a_i), \quad a^n \in \mathcal{X}^n, b^n \in \mathcal{Y}^n \tag{2.10}$$

the capacity is

$$\max_{X\colon\, \mathbb{E}(X^2)\leq \mathsf{P}} \mathbb{I}(X;Y) \tag{2.11}$$

where $\mathbb{I}(X;Y)$ is the mutual information defined in Appendix C.5. The capacity result (2.11) also holds for continuous output memoryless channels with discrete input and with continuous input. In Section 3.4, we discuss the capacity formula of the memoryless AWGN channel.

An important step in the derivation of this result is to show that for any $\delta > 0$, the value $\mathbb{I}(X;Y) - \delta$ is an achievable rate in the sense of Definition 1. This suggests to call $\mathbb{I}(X;Y)$ an approachable rate rather than an achievable rate, since the requirements of Definition 1 are only verified for $\mathbb{I}(X;Y) - \delta$, not for $\mathbb{I}(X;Y)$. In this work, we follow the common terminology ignoring this subtlety and call $\mathbb{I}(X;Y)$ an achievable rate, in consistency with literature.

For a large class of channels, including many practical channels, achievable rates can be estimated that provide a lower bound on the channel capacity. Achievability schemes operating close to capacity is the main topic of theses notes and chapters 7–10 develop such schemes in detail.

2.4.1. Channel Coding Converse for Memoryless Channels

We next state a converse result, namely that above a certain threshold, reliable communication is impossible. The derivation of converse results is in general involved; here, we only consider the special case of memoryless channels. Our goal is to attach a first operational meaning to the mutual information, which we will then study in detail for the AWGN in the chapters 3–5 and which will serve as a guidance for transmitter design in Chapter 6. The statement and proof of the converse result uses basic information measures, namely the entropy $\mathbb{H}(X)$ of a discrete random variable X, the binary entropy function $\mathbb{H}_2(P)$ of a probability P, and the differential entropy $\mathrm{h}(Y)$ of a continuous random variable Y. The definitions and basic properties of these information measures are stated in Appendix C.

Theorem 1 (Channel Coding Converse)**.** *Consider a transmission system with block length n. The message W can take the values $1, 2, \dots, 2^{nR}$. The code word $X^n = x^n(W)$ is transmitted over a memoryless channel $p_{Y|X}$. If*

$$\frac{\mathbb{H}(W)}{n} > \frac{\sum_{i=1}^{n} \mathbb{I}(X_i;Y_i)}{n} \tag{2.12}$$

then the probability of error $P_e = \Pr(W \neq \hat{W})$ is bounded away from zero.

Proof. We have

$$\begin{aligned}
\mathbb{H}_2(P_e) + P_e \log_2(2^{nR} - 1) &\overset{\text{(a)}}{\geq} \mathbb{H}(W|\hat{W}) \\
&= \mathbb{H}(W) - \mathbb{I}(W;\hat{W}) \\
&\overset{\text{(b)}}{\geq} \mathbb{H}(W) - \mathbb{I}(X^n;Y^n) \\
&\overset{\text{(c)}}{=} \mathbb{H}(W) - \left[h(Y^n) - \sum_{i=1}^{n} h(Y_i|X_i)\right] \\
&\overset{\text{(d)}}{\geq} \mathbb{H}(W) - \sum_{i=1}^{n} \left[h(Y_i) - h(Y_i|X_i)\right] \\
&= \mathbb{H}(W) - \sum_{i=1}^{n} \mathbb{I}(X_i;Y_i)
\end{aligned}$$

where (a) follows by Fano's inequality (C.18), (b) by the data-processing inequality (C.41), (c) follows because the channel is memoryless and (d) follows by the independence bound on entropy (C.12). Dividing by n, we have

$$\frac{\mathbb{H}(W)}{n} - \frac{\sum_{i=1}^{n} \mathbb{I}(X_i;Y_i)}{n} \leq \frac{\mathbb{H}_2(P_e)}{n} + P_e \frac{\log_2(2^{nR} - 1)}{n} \leq \mathbb{H}_2(P_e) + P_e R.$$

That is, if $\mathbb{H}(W)/n > \frac{\sum_{i=1}^{n} \mathbb{I}(X_i;Y_i)}{n}$ then P_e is bounded away from zero. □

2.5. Problems

Problem 2.1. For the transmission scheme of Example 2.3, calculate the direct current that results from one block transmission.

Problem 2.2. Let $X_1, X_2, \ldots, X_n$ be distributed according to the joint distribution P_{X^n} (the X_i are possibly stochastically dependent). Show that

$$\frac{\mathbb{E}\left[\sum_{i=1}^{n} X_i^2\right]}{n} = \frac{\sum_{i=1}^{n} \mathbb{E}[X_i^2]}{n}. \tag{2.13}$$

3. AWGN Channel

3.1. Summary

- The discrete time AWGN channel is
$$Y = X + Z. \tag{3.1}$$
- Y, X, Z are channel output, channel input, and noise, respectively.
- The input X and noise Z are stochastically independent.
- The noise Z is zero mean Gaussian with variance σ^2.
- A real-valued random variable X has power $\mathbb{E}(X^2)$.
- The SNR is $\mathsf{snr} = \frac{\text{input power}}{\text{noise power}}$.
- SNR in dB is $10\log_{10}\mathsf{snr}$.
- The capacity of the AWGN channel is
$$\mathsf{C}(\mathsf{snr}) = \frac{1}{2}\log_2(1+\mathsf{snr}) \quad \left[\frac{\text{bits}}{\text{channel use}}\right] \tag{3.2}$$
$$\text{inverse: } \mathsf{snr} = \mathsf{C}^{-1}(R) = 2^{2R} - 1. \tag{3.3}$$
- $\mathsf{C}(\mathsf{snr})$ is plotted in Figure 3.1.
- Phase transition at capacity: For a fixed snr^*, reliable communication at a rate R is possible, if $R < \mathsf{C}(\mathsf{snr}^*)$, and impossible, if $R > \mathsf{C}(\mathsf{snr}^*)$. For a fixed rate R^*, reliable communication is possible if $\mathsf{snr} > \mathsf{C}^{-1}(R^*)$ and impossible, if $\mathsf{snr} < \mathsf{C}^{-1}(R^*)$.

3.2. Channel Model

At each time instant i, we describe the discrete time memoryless AWGN channel by the input-output relation
$$Y_i = x_i + Z_i. \tag{3.4}$$

The noisy signal point Y_i is the sum of the signal point x_i and the noise Z_i. The noise random variables $\{Z_i, i \in \mathbf{Z}\}$ ($\mathbf{Z}$ denotes the set of integers) are independent and identically distributed (iid) according to a Gaussian density with zero mean and variance σ^2, i.e., the Z_i have the probability density function (pdf)

$$p_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{z^2}{2\sigma^2}}, \quad z \in \mathbf{R}. \tag{3.5}$$

The relation (3.4) can equivalently be represented by the conditional output pdf

$$p_{Y|X}(y|x) = p_Z(y-x). \tag{3.6}$$

We model the channel input as a random variable X that is stochastically independent of the noise Z.

Discussion Using the model (3.4) to abstract the continuous-time part of the digital transmission system can be justified in several ways. An in-depth treatment can be found in [3, Chapter 8]. In [7], it is shown with mathematical rigor that, under certain conditions, the Y_i form a sufficient statistics, i.e., all the information about X_i that is contained in the continuous-time received waveform is also contained in the discrete-time noisy signal point Y_i. This approach explicitly models the noise that gets added to the waveform as white Gaussian noise. However, we can only “see” the noise through filters. A second approach to justify (3.4) is therefore to build the digital transmission system, perform a measurement campaign, and to find a reasonable statistical model for the noise sequence $Z_i = Y_i - X_i$, where X_i is chosen at the transmitter and known and where Y_i is measured at the receiver. For instance, this approach was used in [8] to derive a channel model for ultra-wideband wireless communication. Finding the “true” channel description is often difficult. A third approach is to design systems that operate as if the channel would be an AWGN channel. This leads to a *mismatch* between the actual channel and the channel assumed by the transceiver. The mismatched approach is an effective method for designing communication systems for practical channels and we use it in the chapters 7–11. We use “the channel assumed by the transceiver” and “decoding metric” as synonyms.

3.3. Signal-to-Noise Ratio and E_b/N_0

SNR Consider the AWGN channel model

$$Y = X + Z \tag{3.7}$$

where Z is Gaussian noise with variance σ^2. We scale (3.7) by some constant κ. The resulting input-output relation is

$$\underbrace{\kappa \cdot Y}_{=:\tilde{Y}} = \underbrace{\kappa \cdot X}_{=:\tilde{X}} + \underbrace{\kappa \cdot Z}_{=:\tilde{Z}} \tag{3.8}$$

which gives us a modified channel model

$$\tilde{Y} = \tilde{X} + \tilde{Z}. \tag{3.9}$$

Suppose the power of X is $\mathbb{E}(X^2) = \mathsf{P}$. The power of $\tilde{X}$ is then $\kappa^2\mathsf{P}$, so the transmitted signal in (3.9) has a power that is different from the power of the transmitted signal in (3.7). The power of the noise has also changed, namely from σ^2 in (3.7) to $\kappa^2\sigma^2$ in (3.9). What remains constant under scaling is the SNR, which is defined as

$$\mathsf{snr} := \frac{\text{signal power}}{\text{noise power}}. \tag{3.10}$$

The SNR is unitless, so if the signal power is P Watts and the noise power is σ^2 Watts, then the SNR is P/σ^2 (unitless). The SNR is often expressed in *decibel* (dB), which is defined as

$$\text{SNR in dB} = \mathsf{snr}_{\text{dB}} = 10\log_{10}\mathsf{snr}. \tag{3.11}$$

$\boldsymbol{E_b/N_0}$ The relation of signal power and noise power can alternatively be expressed by the E_b/N_0. As defined in (2.9), E_b is the energy we spend to transmit one *message bit*. N_0 is the noise variance per two dimensions, i.e., if the noise variance per channel use is σ^2, then $N_0 = 2\sigma^2$. We can express E_b/N_0 in terms of SNR by

$$E_b/N_0 = \frac{\mathsf{P}}{R}\cdot\frac{1}{2\sigma^2} = \frac{\mathsf{P}}{\sigma^2}\cdot\frac{1}{2R} = \frac{\mathsf{snr}}{2R} \tag{3.12}$$

where R is the rate as defined in (2.1). In decibel, E_b/N_0 is

$$E_b/N_0 \text{ in dB} = 10\log_{10}(E_b/N_0). \tag{3.13}$$

3.4. Capacity

Recall the capacity result (2.11), which says that for a power constraint P, the capacity of a memoryless channel is

$$\max_{X\colon\, \mathbb{E}(X^2)\le\mathsf{P}} \mathbb{I}(X;Y). \tag{3.14}$$

We denote the maximizing input random variable by X^*. For the AWGN channel, by Problem 3.3, X^* has a Gaussian density with zero mean and variance P.

Theorem 2 (AWGN Capacity)**.** *For the AWGN channel with power constraint* P *and noise variance* σ^2*, the* capacity-power function *is*

$$\mathsf{C}(\mathsf{P}/\sigma^2) = \frac{1}{2}\log(1+\mathsf{P}/\sigma^2). \tag{3.15}$$

In other words, we have the following result.

1. *(Converse) No rate* $R > \mathsf{C}(\mathsf{P}/\sigma^2)$ *is achievable by a system with average power less or equal to* P*.*

2. *(Achievability) Any rate* $R < \mathsf{C}(\mathsf{P}/\sigma^2)$ *is achievable by a system with average power* P*.*

We provide a plot of the capacity-power function in Figure 3.1.

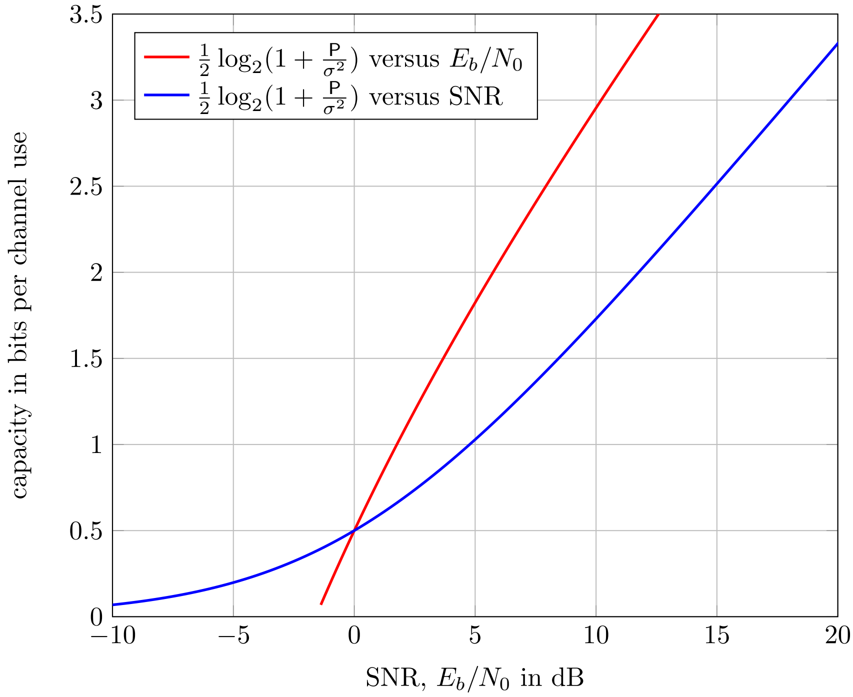


Figure 3.1.: The capacity-power function.

3.5. Problems

Problem 3.1. Let X and Z be stochastically independent Gaussian random variables with means μ_1, μ_2 and variances σ_1^2, σ_2^2. Show that $Y = X + Z$ is zero mean Gaussian with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$.
Hint: Use that for two independent real-valued random variables p_X, p_Z, the pdf of the sum is given by the convolution of p_X with p_Z.
Problem 3.2. Let X have density p_X with mean μ and variance σ^2. Let Y be Gaussian with the same mean μ and variance σ^2. Show that $\mathrm{h}(Y) \geq \mathrm{h}(X)$ with equality if and only if X has the same density as Y, i.e., if X is also Gaussian.
Hint: You can use the information inequality (C.21) in your derivation.
Problem 3.3. Consider the AWGN channel $Y = X + Z$, where Z is Gaussian with zero mean and variance σ^2, where X and Z are stochastically independent, and where $\mathbb{E}(X^2) \leq \mathsf{P}$. Show that

$$\mathbb{I}(X;Y) \leq \frac{1}{2}\log(1 + \mathsf{P}/\sigma^2). \tag{3.16}$$

For which pdf p_X is the maximum achieved?
Hint: You can use Problem 3.1 and Problem 3.2 in your derivation.
Problem 3.4. Consider two (SNR,rate) operating points $(\mathsf{snr}_1, \mathsf{C}(\mathsf{snr}_1))$ and $(\mathsf{snr}_2, \mathsf{C}(\mathsf{snr}_2))$ on the power-rate function.

1. Investigate the dependence of the SNR gap in dB on the rate gap for high SNR,

i.e., to which value does the ratio

$$\frac{10\log_{10}(\mathsf{snr}_2) - 10\log_{10}(\mathsf{snr}_1)}{\mathsf{C}(\mathsf{snr}_2) - \mathsf{C}(\mathsf{snr}_1)} \tag{3.17}$$

converge for $\mathsf{snr}_2, \mathsf{snr}_1 \to \infty$?

2. To which value does (3.17) converge for $\mathsf{P}_2, \mathsf{P}_1 \to 0$?
3. Answer the questions 1. and 2. when E_b/N_0 in dB is used in (3.17) instead of the SNR in dB.

Problem 3.5.

1. To which value does E_b/N_0 in dB converge when the SNR approaches $-\infty$?
2. What is the minimum energy we need to transmit one bit reliably over the AWGN channel?
3. How long will the transmission of one bit with minimum energy take?

4. Shaping Gaps for AWGN

For the AWGN channel, we have

$$Y = X + Z \tag{4.1}$$

with noise variance σ^2 and input power constraint P. We want to characterize the loss of mutual information of an input X and an output Y that results from not using Gaussian input. We consider the following three situations.

1. The input is not Gaussian.
2. The input is continuous and uniformly distributed.
3. The input is discrete and uniformly distributed.

See Figure 4.1 for an illustration of the considered input distributions.

4.1. Summary

- The ASK constellation with M signal points (M-ASK) is

$$\mathcal{X} = \{\pm 1, \pm 3, \dots, \pm(M-1)\}. \tag{4.2}$$

- Let X be uniformly distributed on $\mathcal{X}$.
- The channel input is ΔX where Δ is a positive real number; the input power is $\Delta^2 \,\mathbb{E}(X^2)$.
- An achievable rate is

$$\mathbb{I}(X;Y) = \mathbb{I}(X; \Delta X + Z). \tag{4.3}$$

- Achievable rates for 2^m-ASK are plotted in Figure 4.3. Observations:
  - Ungerboeck's rule of thumb [9]: The 2^m-ASK curves stay close to capacity for rates smaller than $m-1$.
  - The 2^m-ASK curves saturate at m bits for large SNR.
  - Shaping gap: with increasing m, a gap to capacity becomes apparent. The gap is caused by the uniform input distribution over the set of permitted signal points. The gap converges to 1.53 dB, asymptotically in the ASK constellation size and the SNR.

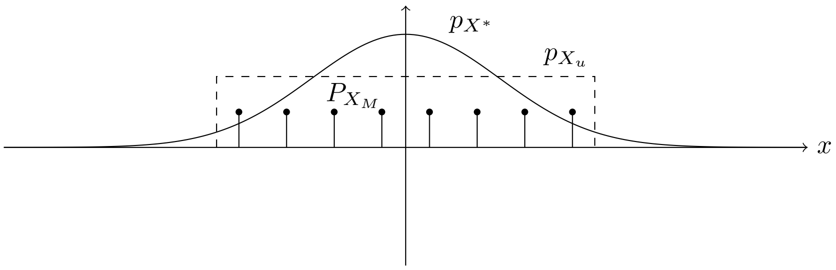


Figure 4.1.: The Gaussian density p_{X^*}, the uniform density p_{X_u} and the uniform distribution P_{X_M} on a discrete M-ASK constellation. The densities p_{X^*}, p_{X_u} and the distribution P_{X_M} have the same variance.

4.2. Non-Gaussian Input

By Problem 3.3, the capacity-achieving density p_{X^*} of the AWGN channel with power constraint P is zero mean Gaussian with variance P. When p_{X^*} is used for the channel input, the resulting output $Y^* = X^* + Z$ is zero mean Gaussian with variance $\mathsf{P} + \sigma^2$, see Problem 3.1. Let now X be a channel input, discrete or continuous, with zero mean and variance P. We expand the mutual information as

$$\mathbb{I}(X;Y) = \mathrm{h}(Y) - \mathrm{h}(Y|X). \tag{4.4}$$

For the conditional differential entropy, we have

$$\begin{aligned}
\mathrm{h}(Y|X) &\overset{\text{(a)}}{=} \mathrm{h}(Y - X|X) && (4.5)\\
&= \mathrm{h}(Z|X) && (4.6)\\
&\overset{\text{(b)}}{=} \mathrm{h}(Z) = \frac{1}{2}\log_2(2\pi e\sigma^2) && (4.7)
\end{aligned}$$

where (a) follows by (C.9) and where (b) follows because X and Z are independent.

Remark 1. We calculated the right-hand side of (4.7) using the definition of differential entropy (C.5). Since $\sigma^2 = \mathbb{E}(Z^2)$ is the noise power, it should have the unit Watts. However, the argument of the logarithm must be unitless. In Problem 4.11, we provide an alternative definition of differential entropy, which works with units. Using the alternative differential entropy definition does not alter the results presented in this chapter.

By (4.7), the term $\mathrm{h}(Y|X)$ does not depend on how the input X is distributed. However, the differential entropy $\mathrm{h}(Y)$ does depend on the distribution of X. For continuous X, the density of Y is given by

$$p_Y(y) = \int_{-\infty}^{\infty} p_X(x)p_Z(y-x)\,\mathrm{d}x = (p_X \star p_Z)(y) \tag{4.8}$$

that is, p_Y is the *convolution* of p_X and p_Z. If X is discrete, the density p_Y is given by

$$p_Y(y) = \sum_{x \in \mathcal{X}} P_X(x)p_Z(y-x). \tag{4.9}$$

The differential entropy $\mathrm{h}(Y)$ is a functional of the density p_Y and it is in general difficult to derive closed-form expressions. We next write $\mathrm{h}(Y)$ as

$$\mathrm{h}(Y) \overset{(a)}{=} \mathrm{h}(Y^*) - \mathbb{D}(p_Y \| p_{Y^*}) \tag{4.10}$$

where (a) is shown in Problem 4.4. We can now write $\mathbb{I}(X;Y)$ as

$$\begin{aligned}
\mathbb{I}(X;Y) &= \mathrm{h}(Y) - \mathrm{h}(Y|X) && (4.11)\\
&\overset{(a)}{=} \mathrm{h}(Y) - \mathrm{h}(Z) && (4.12)\\
&\overset{(b)}{=} \mathrm{h}(Y^*) - \mathbb{D}(p_Y \| p_{Y^*}) - \mathrm{h}(Z) && (4.13)\\
&\overset{(c)}{=} [\mathrm{h}(Y^*) - \mathrm{h}(Y^*|X^*)] - \mathbb{D}(p_Y \| p_{Y^*}) && (4.14)\\
&= \mathsf{C}(\mathsf{P}/\sigma^2) - \mathbb{D}(p_Y \| p_{Y^*}) && (4.15)
\end{aligned}$$

where we used (4.7) in (a) and (c) and (4.10) in (b). We make the following two observations.

- The loss of mutual information when using X instead of X^* is the informational divergence $\mathbb{D}(p_Y \| p_{Y^*})$ between the resulting output distributions.
- The loss can be small even if X differs significantly from X^* as long as the resulting output distribution p_Y is similar to p_{Y^*}. See also Problem 4.6.

We next consider two constraints on the input. First, we let X be uniformly distributed on a continuous finite interval, and second, we let X be uniformly distributed on a discrete ASK constellation.

4.3. Uniform Continuous Input

Let X_u be an input that is uniformly distributed on a finite interval $[-A, A]$ where A is such that the variance of X_u is equal to P. Denote by Y_u the corresponding output. Recall that the noise variance is σ^2 and the SNR is $\mathsf{snr} = \mathsf{P}/\sigma^2$.

Lower Bound

We bound

$$\begin{aligned}
\mathbb{I}(X_u;Y_u) &\overset{(a)}{=} \mathsf{C}(\mathsf{snr}) - \mathbb{D}(p_{Y_u} \| p_{Y^*}) && (4.16)\\
&\overset{(b)}{\geq} \mathsf{C}(\mathsf{snr}) - \mathbb{D}(p_{X_u} \| p_{X^*}) && (4.17)\\
&\overset{(c)}{=} \mathsf{C}(\mathsf{snr}) - [\mathrm{h}(X^*) - \mathrm{h}(X_u)] && (4.18)\\
&\overset{(d)}{=} \mathsf{C}(\mathsf{snr}) - \frac{1}{2}\log_2 \frac{\pi e}{6} && (4.19)
\end{aligned}$$

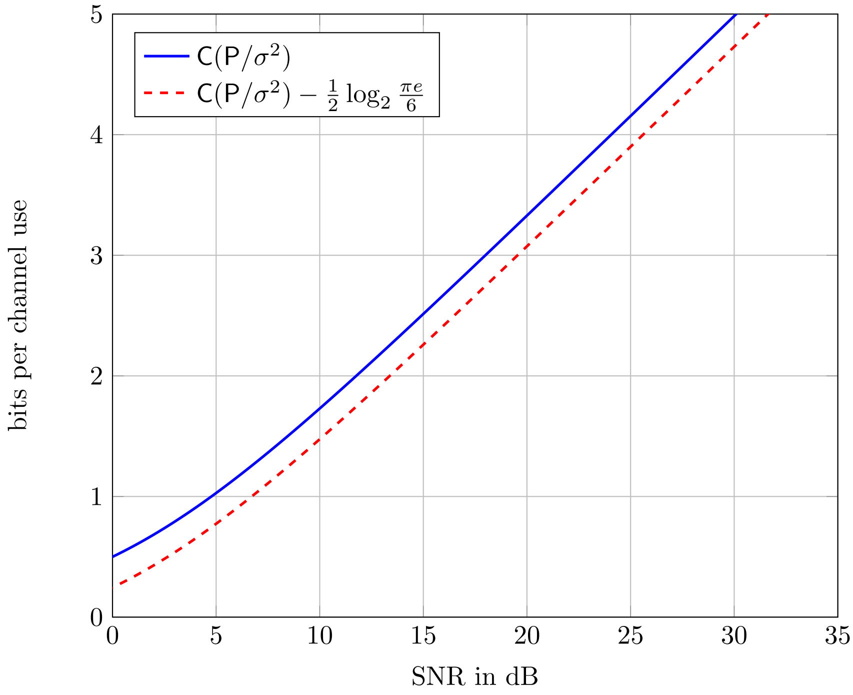


Figure 4.2.: The dashed line shows the lower bound (4.19) for continuous uniformly distributed input.

where (a) follows by (4.15), where (b) follows by Problem 4.6, where (c) follows by Problem 4.4, and where (d) follows by Problem 4.5. The inequality $\mathbb{D}(p_{X_u}\|p_{X^*}) \geq \mathbb{D}(p_{Y_u}\|p_{Y^*})$ that we used in (b) is a data processing inequality, see [5, Lemma 3.11]. We make the following observation.

- The loss of mutual information because of a uniform input density is at most $\frac{1}{2}\log_2\frac{\pi e}{6}$ independent of the SNR snr. The value $\frac{1}{2}\log_2\frac{\pi e}{6}$ is sometimes called the *shaping gap*.

We can also express (4.19) as

$$\mathbb{I}(X_u;Y_u) = \mathrm{h}(X_u) - \mathrm{h}(X_u|Y_u) \tag{4.20}$$

$$\geq \mathrm{h}(X_u) - \frac{1}{2}\log_2\left(2\pi e\frac{\mathsf{P}}{1+\mathsf{P}/\sigma^2}\right) \tag{4.21}$$

which shows that the conditional entropy of X_u is bounded from above by

$$\mathrm{h}(X_u|Y_u) \leq \frac{1}{2}\log_2\left(2\pi e\frac{\mathsf{P}}{1+\mathsf{P}/\sigma^2}\right). \tag{4.22}$$

Upper Bound

We next want to show that the shaping gap (4.19) is tight, i.e., that for large SNR, the lower bound (4.19) holds with equality. To this end, we will derive an upper bound for

$\mathbb{I}(X_u;Y_u)$ that approaches (4.19) from above when the SNR approaches infinity. Note that because X_u is uniformly distributed on $[-A,A]$, it fulfills the peak-power constraint

$$|X_u| \leq A. \tag{4.23}$$

We can thus use the following mutual information upper-bound for input with peak power A:

$$\operatorname{supp} p_X \subseteq [-A,A] \Rightarrow \mathbb{I}(X;Y) \leq \log_2\left(1+\sqrt{\frac{2A^2}{\pi e \sigma^2}}\right). \tag{4.24}$$

This bound is stated in [10] and proven in [11], see also [12]. Note that the bound holds also for non-uniformly distributed input.

Since $\mathbb{E}(X_u^2) = A^2/3 = \mathsf{P}$, we have $A^2 = 3\mathsf{P}$. We can now bound the shaping gap from below by

$$\begin{aligned}
\mathsf{C}(\mathsf{P}/\sigma^2) - \mathbb{I}(X_u;Y_u) &\geq \mathsf{C}(\mathsf{P}/\sigma^2) - \log_2\left(1+\sqrt{\frac{6\mathsf{P}}{\pi e \sigma^2}}\right) &(4.25)\\
&= \frac{1}{2}\log_2(1+\mathsf{snr}) - \frac{1}{2}\log_2\left(1+2\sqrt{\frac{6\mathsf{snr}}{\pi e}}+\frac{6\mathsf{snr}}{\pi e}\right) &(4.26)\\
&= \frac{1}{2}\log_2\frac{1+\mathsf{snr}}{1+2\sqrt{\frac{6\mathsf{snr}}{\pi e}}+\frac{6\mathsf{snr}}{\pi e}} &(4.27)\\
&= \frac{1}{2}\log_2\frac{\frac{1}{\mathsf{snr}}+1}{\frac{1}{\mathsf{snr}}+2\sqrt{\frac{6}{\pi e \mathsf{snr}}}+\frac{6}{\pi e}} &(4.28)\\
&\overset{\mathsf{snr}\to\infty}{\to} \frac{1}{2}\log_2\frac{\pi e}{6}. &(4.29)
\end{aligned}$$

Thus, asymptotically in the SNR, the shaping gap lower bound is tight, i.e., it approaches the upper bound $\frac{1}{2}\log_2\frac{\pi e}{6}$. Summarizing, we have

$$\begin{aligned}
\mathsf{C}(\mathsf{snr}) - \mathbb{I}(X_u;Y_u) &\leq \frac{1}{2}\log_2\frac{\pi e}{6} \quad \text{(for any SNR)} &(4.30)\\
\lim_{\mathsf{snr}\to\infty} \mathsf{C}(\mathsf{snr}) - \mathbb{I}(X_u;Y_u) &= \frac{1}{2}\log_2\frac{\pi e}{6}. &(4.31)
\end{aligned}$$

4.4. Finite Signal Constellations

For notational convenience, we deviate in this section from our standard notation (4.2) for M-ASK constellations and define

$$\mathcal{X} = \{\pm\Delta, \pm 3\Delta, \dots, \pm(M-1)\Delta\} \tag{4.32}$$

so that the channel input is X (instead of $X\Delta$). Let X_M be uniformly distributed on $\mathcal{X}$. The resulting power is

$$\mathsf{P} = \mathbb{E}(X_M^2) = \Delta^2 \frac{M^2-1}{3} \tag{4.33}$$

see Table 4.1.

Theorem 3 (Uniform Discrete Input Bound)**.** *The mutual information achieved by X_M is lower bounded by*

$$\begin{aligned}\mathbb{I}(X_M;Y_M) &\geq \frac{1}{2}\log_2\left(12\mathsf{P}\frac{M^2}{M^2-1}\right) - \frac{1}{2}\log_2\left[2\pi e\left(\frac{\mathsf{P}}{M^2-1}+\frac{\mathsf{P}}{1+\mathsf{P}/\sigma^2}\right)\right] && (4.34)\\ &> \mathsf{C}(\mathsf{snr}) - \frac{1}{2}\log_2\frac{\pi e}{6} - \frac{1}{2}\log_2\left[1+\left(\frac{2^{\mathsf{C}(\mathsf{snr})}}{M}\right)^2\right] && (4.35)\end{aligned}$$

where $\mathsf{snr} = \mathsf{P}/\sigma^2$.

Proof. We prove the theorem in Section 4.5. □

Our input X_M is suboptimal in two ways. First, it is restricted to the set $\mathcal{X}$ of M equidistant points, and second, X_M is distributed uniformly on $\mathcal{X}$. Because of our result from Section 4.3 for uniform inputs, the mutual information is bounded away from capacity by $\frac{1}{2}\log_2\frac{\pi e}{6} \approx 0.255$ bits independent of how large we choose M. Let's see how M needs to scale with $\mathsf{C}(\mathsf{snr})$ so that the resulting mutual information is within a constant gap of capacity. To keep the finite constellation loss within the order of the distribution loss of 0.255 bits, we calculate

$$\begin{aligned}\frac{1}{2}\log_2\left[1+\left(\frac{2^{\mathsf{C}(\mathsf{snr})}}{M}\right)^2\right] \leq \frac{\log_2 e}{2}\left(\frac{1}{M\cdot 2^{-\mathsf{C}(\mathsf{snr})}}\right)^2 &= \frac{1}{4} && (4.36)\\ \Leftrightarrow M &= 2^{\mathsf{C}(\mathsf{snr})+\frac{1}{2}+\frac{1}{2}\log_2\log_2 e} && (4.37)\end{aligned}$$

where we used $\log_2 x = \log_2(e)\log(x) \leq \log_2(e)(x-1)$. We conclude from (4.37) that the mutual information is within 0.5 bit of capacity if

$$\log_2 M \approx \mathsf{C}(\mathsf{snr}) + 0.77. \tag{4.38}$$

This condition can be confirmed in Figure 4.3, where we display rate curves and bounds for ASK constellations with uniformly distributed input.

4.5. Proof of Uniform Discrete Input Bound

We can expand mutual information in two ways, namely

$$\mathbb{I}(X_M;Y_M) = \mathrm{h}(Y_M) - \mathrm{h}(Y_M|X_M) = \mathbb{H}(X_M) - \mathbb{H}(X_M|Y_M) \tag{4.39}$$

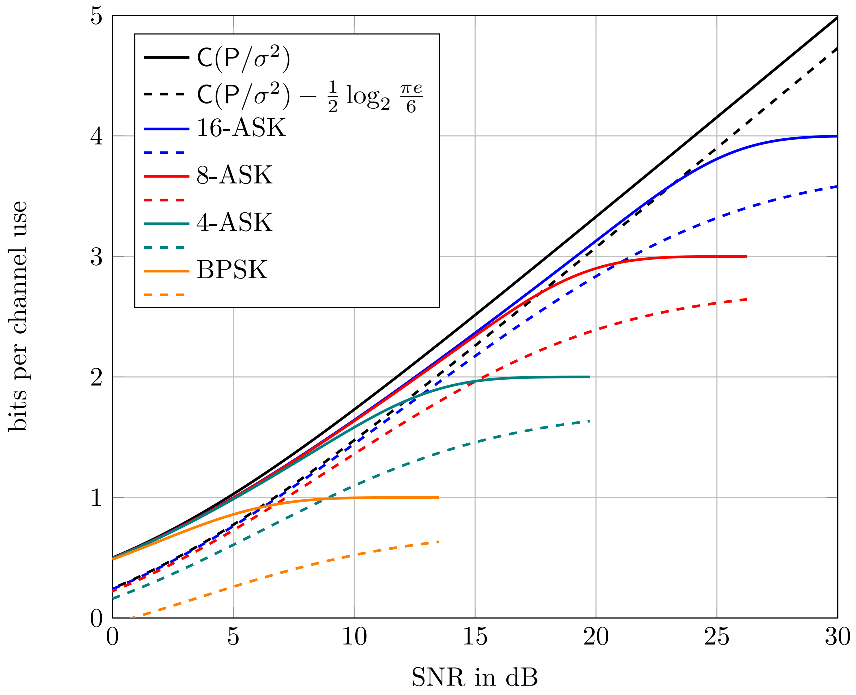


Figure 4.3.: The dotted curves show the bound (4.35) from Theorem 3. The gap between the dotted curves and the capacity-power function confirms the condition (4.38). The dashed line shows the lower bound (4.19) for continuous uniformly distributed input. Note that for low SNR, the rate curves for ASK are close to the capacity-power function, while for high SNR, the dashed curve becomes a tight lower bound, i.e., for high SNR and large ASK constellations, the shaping gap of $\frac{1}{2}\log_2\frac{\pi e}{6}$ becomes apparent.

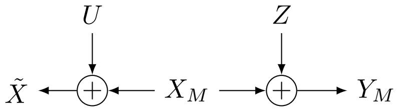


Figure 4.4.: U is uniformly distributed on $[-\Delta, \Delta]$ and $\tilde{X}$ is uniformly distributed on $[-\Delta M, \Delta M]$.

where we know

$$
\begin{aligned}
\mathrm{h}(Y_M|X_M) &= \mathrm{h}(Z) = \frac{1}{2}\log_2(2\pi e\sigma^2) &(4.40)\\
\mathbb{H}(X_M) &= \log_2 M &(4.41)
\end{aligned}
$$

but we have no insightful expressions for $\mathrm{h}(Y_M)$ and $\mathbb{H}(X_M|Y_M)$. We want to lower bound $\mathbb{I}(X_M;Y_M)$, so we can either lower bound $\mathrm{h}(Y_M)$ or we can upper bound $\mathbb{H}(X_M|Y_M)$. Following [13], we opt for upper bounding $\mathbb{H}(X_M|Y_M)$. We do this in two steps:

1. We introduce an auxiliary continuous random variable $\tilde{X}$ that is a function of X_M, so that $\mathbb{I}(X_M;Y_M) \geq \mathbb{I}(\tilde{X};Y_M)$ by the data processing inequality (C.41).
2. We upper bound $\mathrm{h}(\tilde{X}|Y_M)$ by using a conditional version of the information inequality (C.21).

Step 1: Auxiliary Continuous Input $\tilde{X}$ The random variable X_M is discrete. To make our life easier, we first introduce an auxiliary random variable. Let U be continuous and uniformly distributed on $[-\Delta, \Delta)$ and define

$$
\tilde{X} := X_M + U. \tag{4.42}
$$

We provide an illustration in Figure 4.4. From the definitions of X_M and U, it follows that $\tilde{X}$ is continuous and uniformly distributed on $[-\Delta M, \Delta M)$. Since $\tilde{X} -\!\!\circ\!\!- X_M -\!\!\circ\!\!- Y_M$ form a Markov chain, we have by the data processing inequality (C.41)

$$
\begin{aligned}
\mathbb{I}(X_M;Y_M) &\geq \mathbb{I}(\tilde{X};Y_M) &(4.43)\\
&= \mathrm{h}(\tilde{X}) - \mathrm{h}(\tilde{X}|Y_M). &(4.44)
\end{aligned}
$$

For the differential entropy $\mathrm{h}(\tilde{X})$, we have

$$
\mathrm{h}(\tilde{X}) = \log_2(2M\Delta) = \frac{1}{2}\log_2(4M^2\Delta^2) = \frac{1}{2}\log_2\frac{12M^2\mathsf{P}}{M^2-1} \tag{4.45}
$$

where we used Table 4.1. Inserting the right-hand side of (4.45) in (4.44), we have

$$
\mathbb{I}(X_M;Y_M) \geq \frac{1}{2}\log_2\frac{12M^2\mathsf{P}}{M^2-1} - \mathrm{h}(\tilde{X}|Y_M). \tag{4.46}
$$

In the next paragraph, we replace $\mathrm{h}(\tilde{X}|Y_M)$ by an insightful upper bound.

$\mathbb{E}(X_M^2) = \mathsf{P} = \frac{(M^2-1)\Delta^2}{3}$
$\mathbb{E}(U^2) = \frac{\Delta^2}{3}$
$\mathbb{E}(\tilde{X}^2) = \frac{(M\Delta)^2}{3}$

Table 4.1.: Powers of X_M, U, and $\tilde{X}$ as derived in Problem 4.3.

Step 2: Bounding $\mathrm{h}(\tilde{X}|Y_M)$. By Problem 4.7, we have

$$\mathrm{h}(\tilde{X}|Y_M) = \mathbb{E}[-\log_2 p_{\tilde{X}|Y_M}(\tilde{X}|Y_M)] \tag{4.47}$$

$$\leq \mathbb{E}[-\log_2 q(\tilde{X}|Y_M)] \tag{4.48}$$

for any $q(\cdot|\cdot)$ with the property that $q(\cdot|y)$ is a density on $\mathbf{R}$ for every $y \in \mathbf{R}$. We choose a Gaussian density with mean ky and variance s^2, i.e., we choose

$$q(x|y;k,s^2) := \frac{1}{\sqrt{2\pi}s}\exp\left[-\frac{(x-ky)^2}{2s^2}\right]. \tag{4.49}$$

This gives

$$\mathrm{h}(\tilde{X}|Y_M) \leq \log_2(e)\left(\frac{1}{2}\ln(2\pi s^2) + \frac{1}{2s^2}\mathbb{E}[(\tilde{X}-kY_M)^2]\right). \tag{4.50}$$

We calculate the expectation.

$$\mathbb{E}[(\tilde{X}-kY_M)^2] = \mathbb{E}[(X_M+U-k(X_M+Z))^2] \tag{4.51}$$

$$= \mathbb{E}[((1-k)X_M+U-kZ)^2] \tag{4.52}$$

$$\overset{\text{(a)}}{=} \mathsf{P} + \frac{\Delta^2}{3} - 2k\mathsf{P} + k^2(\mathsf{P}+\sigma^2) \tag{4.53}$$

where the reader is asked to verify (a) in Problem 4.8. We can now write the bound for the conditional entropy of $\tilde{X}$ as

$$\mathrm{h}(\tilde{X}|Y_M) \leq \log_2(e)\left(\frac{1}{2}\ln(2\pi s^2) + \frac{1}{2s^2}\left[\mathsf{P} + \frac{\Delta^2}{3} - 2k\mathsf{P} + k^2(\mathsf{P}+\sigma^2)\right]\right). \tag{4.54}$$

In Problem 4.8, we minimize this expression over the parameters k and s^2. The solution is $k = \frac{\mathsf{P}}{\mathsf{P}+\sigma^2}$ and $s^2 = \frac{\Delta^2}{3} + \frac{\mathsf{P}}{1+\mathsf{P}/\sigma^2}$ and the minimized bound is

$$\mathrm{h}(\tilde{X}|Y_M) \leq \frac{1}{2}\log_2\left[2\pi e\left(\frac{\Delta^2}{3} + \frac{\mathsf{P}}{1+\mathsf{P}/\sigma^2}\right)\right]. \tag{4.55}$$

Remark 2. We can express Δ in terms of M and P, namely

$$\Delta^2 = \frac{3\mathsf{P}}{M^2-1}. \tag{4.56}$$

Thus, we have

$$\mathrm{h}(\tilde{X}|Y_M) \leq \frac{1}{2}\log_2\left[2\pi e\left(\frac{\mathsf{P}}{M^2-1}+\frac{\mathsf{P}}{1+\mathsf{P}/\sigma^2}\right)\right] \tag{4.57}$$

$$\overset{M\to\infty}{\to} \leq \frac{1}{2}\log_2\left[2\pi e\frac{\mathsf{P}}{1+\mathsf{P}/\sigma^2}\right]. \tag{4.58}$$

Thus, as the number of signal points approaches infinity, our upper bound on $\mathrm{h}(\tilde{X}|Y_M)$ approaches our upper bound (4.22) on $\mathrm{h}(X_u|Y_u)$.

Remark 3. We can show that we have equality in (4.43), i.e., $\mathbb{I}(X_M;Y_M) = \mathbb{I}(\tilde{X};Y_M)$. The reason is that X_M is a deterministic function of $\tilde{X}$. Formally, note that $U = \tilde{X} - X_M$ and $Z = Y_M - X_M$ and by definition $\tilde{X} = X_M + U$ and $Y_M = X_M + Z$. Thus, by (C.40), we have

$$\mathbb{I}(\tilde{X};Y_M|X_M) = \mathbb{I}(\tilde{X},U;Y_M,Z|X_M) = \mathbb{I}(U;Z|X_M) \overset{\text{(a)}}{=} 0 \tag{4.59}$$

where (a) follows because X_M, Z, U are stochastically independent. Thus, we have

$$\mathbb{I}(\tilde{X};Y_M) \overset{\text{(a)}}{=} \mathbb{I}(\tilde{X},X_M;Y_M) \tag{4.60}$$

$$\overset{\text{(b)}}{=} \mathbb{I}(X_M;Y_M) + \mathbb{I}(\tilde{X};Y_M|X_M) \tag{4.61}$$

$$\overset{\text{(c)}}{=} \mathbb{I}(X_M;Y_M) \tag{4.62}$$

where (a) follows by (C.40) (X_M is a function of $\tilde{X}$), where (b) follows by (C.39), and where (c) follows by (4.59). This proves the equality.

4.6. Problems

Problem 4.1. Derive (4.8) by using the definitions of density (B.1) and expectation (B.4) and the law of total probability (B.5). *Hint:* Show that

$$\Pr(Y \leq y) = \int_{-\infty}^{y}\int_{-\infty}^{\infty} p_X(x)p_Z(\tau - x)\,\mathrm{d}x\,\mathrm{d}\tau. \tag{4.63}$$

Problem 4.2. Show that if the channel input X takes values in an M-ASK constellation, then the mutual information of channel input X and channel output Y is upper bounded by $\log_2 M$.

Problem 4.3. Let X be uniformly distributed on $[-A, A]$. Show that

$$\mathrm{Var}(X) = \frac{A^2}{3}. \tag{4.64}$$

Problem 4.4. Let X be some continuous random variable with zero mean and variance P. Let X' be zero mean Gaussian with variance P. Show that

$$\mathrm{h}(X') - \mathrm{h}(X) = \mathbb{D}(p_X \| p_{X'}). \tag{4.65}$$

Problem 4.5. Let X and X' be continuous random variables with zero mean and variance P. Suppose X is distributed uniformly on $[-d, d]$ and let X' be Gaussian.

1. Show that

$$\mathrm{h}(X) = \frac{1}{2}\log_2(12\mathsf{P}) \tag{4.66}$$

$$\mathrm{h}(X') = \frac{1}{2}\log_2(2\pi e\mathsf{P}). \tag{4.67}$$

2. Show that

$$\mathrm{h}(X') - \mathrm{h}(X) = \frac{1}{2}\log_2\frac{\pi e}{6}. \tag{4.68}$$

Problem 4.6. Let p_X and $p_{X'}$ be two densities defined on $\mathcal{X}$. Let $p_{Y|Z}$ be a conditional density where Z takes values in $\mathcal{X}$. Define

$$p_{XY}(a,b) = p_X(a)p_{Y|Z}(b|a) \tag{4.69}$$

$$p_{X'Y'}(a,b) = p_{X'}(a)p_{Y|Z}(b|a) \tag{4.70}$$

Show that

$$\mathbb{D}(p_Y\|p_{Y'}) \le \mathbb{D}(p_X\|p_{X'}). \tag{4.71}$$

Hint: Use the chain rule of informational divergence (C.25) and the information inequality (C.21).

Problem 4.7.

1. Let p_X be a density and let q be some other density with $q(a) = 0 \Rightarrow p_X(a) = 0$, i.e., $\operatorname{supp} p_X \subseteq \operatorname{supp} q$. Show that

$$\mathrm{h}(X) = \mathbb{E}[-\log_2 p_X(X)] \le \mathbb{E}[-\log_2 q(X)] \tag{4.72}$$

where all expectations are taken with respect to p_X.

2. Let now $p_{XY} = p_X p_{Y|X}$ be a joint density and for each $x \in \operatorname{supp} p_X$, let $q(\cdot|x)$ be a density on $\mathbf{R}$ with $\operatorname{supp} p_{Y|X}(\cdot|x) \subseteq \operatorname{supp} q(\cdot|x)$. Use the result from 1. to show that

$$\mathrm{h}(Y|X) = \mathbb{E}[-\log_2 p_{Y|X}(Y|X)] \le \mathbb{E}[-\log_2 q(Y|X)]. \tag{4.73}$$

Problem 4.8. Verify (4.53), (4.55), and (4.35).

Problem 4.9. Use the bounding technique from Problem (4.7) with the Ansatz (4.49) for an alternative derivation of the bound (4.19).

Problem 4.10. Show that the gap in SNR in dB between the bound (4.19) and the capacity-power function approaches approximately 1.53 dB when the SNR goes to infinity.

Problem 4.11. Let Z be zero mean Gaussian with variance σ^2 with unit Watts, i.e., Z has unit $\sqrt{\text{Watt}}$.

1. Show that if the variance σ^2 has unit Watts, then the density p_Z has unit $1/\sqrt{\text{Watt}}$.

2. Consider the probability $\Pr(a \leq Z \leq b) = \int_a^b p_Z(\tau)\,\mathrm{d}\tau$. Verify that the probability is unitless if $\mathrm{d}\tau$ has the same unit as Z.

3. Define differential entropy as alternative to (C.5) by

$$\mathrm{h_r}(Z) = \mathbb{E}\left[-\log_2[p_Z(Z)\mathrm{r}_Z]\right] \tag{4.74}$$

   where r_Z is a constant with the same unit as Z.

4. Show that

$$\mathrm{h_r}(Z) = \frac{1}{2}\log_2\frac{2\pi e\sigma^2}{\mathrm{r}_Z^2}. \tag{4.75}$$

   Note that the argument of the logarithm is unitless, fixing the issue raised in Remark 1.

5. For two continuous random variables X, Y with pdf p_{XY}, show that

$$\mathbb{I}(X;Y) = \mathrm{h_r}(X) - \mathrm{h_r}(X|Y). \tag{4.76}$$

6. (Problem 4.5 revisited) Let X and X' be continuous random variables with zero mean and variance P. Suppose X is distributed uniformly on $[-d, d]$ and let X' be Gaussian.

   a) Calculate $\mathrm{h_r}(X)$ and $\mathrm{h_r}(X')$.

   b) Show that

$$\mathrm{h_r}(X') - \mathrm{h_r}(X) = \frac{1}{2}\log_2\frac{\pi e}{6}. \tag{4.77}$$

5. Non-Uniform Discrete Input Distributions for AWGN

In this chapter, we consider non-uniform input distributions on ASK constellations.

5.1. Summary

- For ASK input, the achievable rate $\mathbb{I}(X; \Delta X + Z)$ should be maximized over the input distribution P_X *and* the constellation scaling Δ.
- The shaping gap is virtually removed if the ASK signal points are used with a sampled Gaussian distribution, which is also called the Maxwell-Boltzmann (MB) distribution.

5.2. Capacity-Achieving Input Distribution

Consider an ASK constellation with M signal points (we restrict M to even integers; in practice, M is usually a power of two) given by

$$\mathcal{X} = \{\pm 1, \pm 3, \dots, \pm(M-1)\}. \tag{5.1}$$

Let X be a random variable with distribution P_X on $\mathcal{X}$. We use X scaled by $\Delta > 0$ as the channel input of an AWGN channel. The resulting input/output relation is

$$Y = \Delta X + Z. \tag{5.2}$$

The mutual information of the channel input and channel output is

$$\begin{aligned} \mathbb{I}(\Delta X; Y) &\overset{\text{(a)}}{=} \mathbb{I}(\Delta X; \Delta X + Z) && (5.3) \\ &\overset{\text{(b)}}{=} \mathbb{I}(X; \Delta X + Z) && (5.4) \end{aligned}$$

where (a) follows by (5.2) and where (b) follows by (C.40) and because (ΔX) is a deterministic function of X and X is a deterministic function of (ΔX). If the input is subject to an average power constraint P, the scaling Δ and the distribution P_X need to be chosen such that the constraint

$$\mathbb{E}[(\Delta X)^2] \leq \mathsf{P} \tag{5.5}$$

is satisfied. The ASK capacity-power function is now given by

$$\mathsf{C}_{\text{ask}}(\mathsf{P}/\sigma^2) = \max_{\Delta, P_X\colon\, \mathbb{E}[(\Delta X)^2] \le \mathsf{P}} \mathbb{I}(X; \Delta X + Z). \tag{5.6}$$

To evaluate the ASK capacity-power function numerically, we have to maximize the mutual information $\mathbb{I}(X; \Delta X + Z)$ both over the scaling Δ of the signal points and the input distribution P_X. In the optimization we need to account for the power constraint (5.5).

1. For a fixed scaling Δ, the mutual information is concave in P_X. There is no closed form expression for the optimal input distribution, but the maximization over P_X can be done efficiently using the *Blahut-Arimoto Algorithm* [14], [15]. (In these papers, the Blahut-Arimoto algorithm is formulated for finite output alphabets and it can be easily adapted to the case of continuous output.)

2. The mutual information maximized over P_X is now a function of the scaling Δ. We optimize Δ in a second step.

3. General purpose optimization software can also be used to solve (5.6).

We denote the optimal scaling by Δ^* and the corresponding distribution by P_{X^*}.

5.3. Maxwell-Boltzmann Input Distribution

To calculate one point of the ASK capacity-power function, we need to solve (5.6) which may require too much computing power if we have to do it many times. As we will see, a suboptimal input distribution is good enough and the resulting rate-power function is very close to the ASK capacity-power function. Note that the suboptimal input distribution also provides a good starting point for accelerating the computation of the ASK capacity-power function by (5.6).

Entropy-Maximizing Input Distribution

The mutual information can be expanded as

$$\mathbb{I}(X; \Delta X + Z) = \mathbb{H}(X) - \mathbb{H}(X|\Delta X + Z). \tag{5.7}$$

For a fixed Δ, we choose the input distribution P_{X_Δ} that maximizes the input entropy subject to our power constraint, i.e., we choose

$$P_{X_\Delta} = \operatorname*{argmax}_{P_X\colon\, \mathbb{E}[(\Delta X)^2] \le \mathsf{P}} \mathbb{H}(X). \tag{5.8}$$

For each $x_i \in \mathcal{X}$, $i = 1, 2, \dots, M$, define

$$P_{X_\nu}(x_i) = A_\nu e^{-\nu x_i^2}, \qquad A_\nu = \frac{1}{\sum_{i=1}^{M} e^{-\nu x_i^2}}. \tag{5.9}$$

	target rate	SNR for $P_X^\clubsuit$	SNR for P_{X^*}	SNR Gap
16-ASK	2.9861 bits	18.0010 dB	18.0000 dB	$9.7445 \cdot 10^{-4}$ dB
32-ASK	3.9839 bits	24.0674 dB	24.0000 dB	$6.743 \cdot 10^{-2}$ dB

Table 5.1.: SNR Gap between the suboptimal input $X^\clubsuit$ and the capacity-achieving input X^*. The values for X^* were calculated by using the Blahut-Arimoto Algorithm.

	target rate	SNR uniform X	SNR $X^\clubsuit$	Gain
4-ASK	1.0000 bits	5.1180 dB	4.8180 dB	0.3000 dB
8-ASK	2.0000 bits	12.6186 dB	11.8425 dB	0.7761 dB
16-ASK	3.0000 bits	19.1681 dB	18.0911 dB	1.0770 dB
32-ASK	4.0000 bits	25.4140 dB	24.1708 dB	1.2432 dB

Table 5.2.: Shaping gains of $X^\clubsuit$ over the uniform input distribution for ASK constellations.

The distributions P_{X_ν} are called MB distributions or *sampled Gaussian distributions*. The definition of A_ν ensures that the probabilities assigned by P_{X_ν} add up to 1. In Problem 5.3, we show that P_{X_Δ} defined by (5.8) is given by

$$P_{X_\Delta}(x_i) = P_{X_\nu}(x_i) \text{ with } \nu\colon\ \mathbb{E}[(\Delta X_\nu)^2] = \mathsf{P}. \tag{5.10}$$

We show in Section 5.4 that $\mathbb{E}[(X_\nu)^2]$ is strictly monotonically decreasing in ν. Thus, the ν for which the condition (5.10) is fulfilled can be found efficiently by using the *bisection method*.

Maximizing Mutual Information

For each constellation scaling Δ, the distribution P_{X_Δ} satisfies the power constraint. We now maximize the mutual information over all input distributions from this family, i.e., we solve

$$\max_{\Delta} \mathbb{I}(X_\Delta; \Delta X_\Delta + Z). \tag{5.11}$$

We denote the best scaling by $\Delta^\clubsuit$, the resulting input distribution by $P_{X^\clubsuit}$, and the corresponding input and output by $X^\clubsuit$ and $Y^\clubsuit$, respectively. We provide numerical results in Figure 5.1 and Table 5.1. We observe that our suboptimal input $X^\clubsuit$ virtually achieves ASK capacity. In Table 5.2, we display the shaping gains of our suboptimal input $X^\clubsuit$ over uniformly distributed input. For increasing target rates and constellation sizes, the shaping gains approach the ultimate shaping gain of 1.53 dB, see Problem 4.10. In particular, for 32-ASK and 4 bits per channel use, the shaping gain is 1.24 dB.

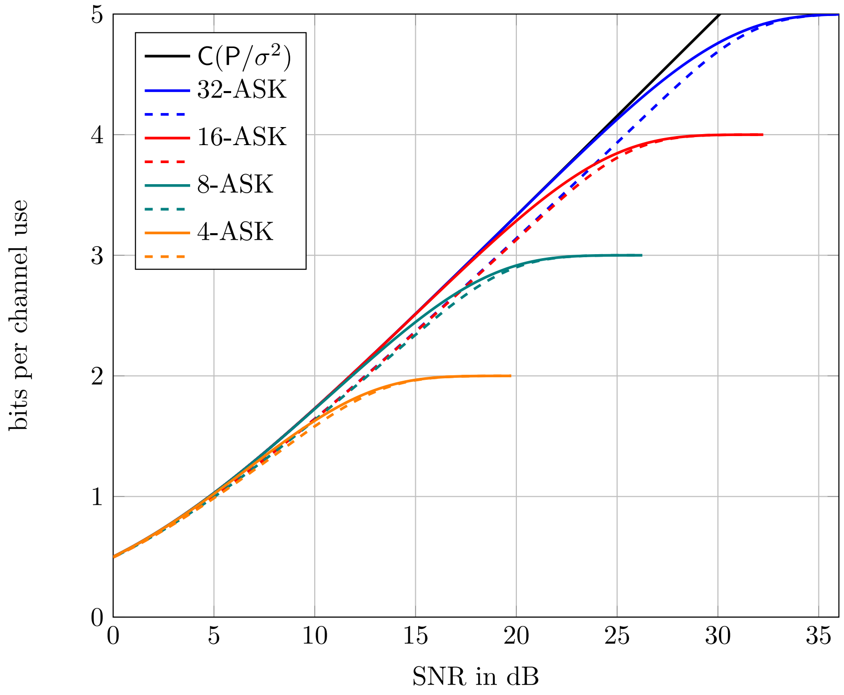


Figure 5.1.: Rate curves for the non-uniform input $X^\clubsuit$ (solid lines in color). The corresponding dashed curves display the rates achieved by uniform input. The curves for $X^\clubsuit$ are very close to the ASK capacity-power function (not displayed). See Table 5.1 for a numerical comparison of our heuristic $X^\clubsuit$ and the capacity-achieving input X^*.

5.4. Proof of Power Monotonicity

Let $f\colon \mathcal{X} \to \mathbf{R}$ be a function that assigns to each symbol $a \in \mathcal{X}$ a finite real value $f(a) \in \mathbf{R}$. For instance, in (5.9), we used the function $a \mapsto a^2$. Let X_ν be a random variable with distribution

$$P_{X_\nu}(a) = \frac{e^{-\nu f(a)}}{\sum_{a' \in \mathcal{X}} e^{-\nu f(a')}}. \tag{5.12}$$

Our aim is to show that the function $g(\nu) := \mathbb{E}[f(X_\nu)]$ is strictly monotonically decreasing in ν if

$$f(a) \neq f(a') \text{ for some } a \neq a' \in \mathcal{X}. \tag{5.13}$$

We prove this by showing that if (5.13) holds, then the first derivative of g is negative. For notational convenience, we write $\mathcal{X} = \{x_1, x_2, \dots, x_M\}$ and we define $w_i = f(x_i)$. We can then write

$$g(\nu) = \mathbb{E}[f(X_\nu)] = \frac{\sum_{i=1}^M w_i \cdot e^{-\nu w_i}}{\sum_{j=1}^M e^{-\nu w_j}} = \frac{h(\nu)}{d(\nu)}. \tag{5.14}$$

The derivative of g is

$$g'(\nu) = \frac{h'(\nu)d(\nu) - h(\nu)d'(\nu)}{[d(\nu)]^2}. \tag{5.15}$$

Since $[d(\nu)]^2 > 0$, we need to show that the numerator is negative. We have

$$\begin{aligned}&h'(\nu)d(\nu) - h(\nu)d'(\nu)\\ &= -\left[\left(\sum_{i=1}^{M} w_i^2 e^{-\nu w_i}\right)\left(\sum_{j=1}^{M} e^{-\nu w_j}\right) - \left(\sum_{i=1}^{M} w_i e^{-\nu w_i}\right)\left(\sum_{j=1}^{M} w_j e^{-\nu w_j}\right)\right].\end{aligned} \tag{5.16}$$

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

$$u_i = w_i\sqrt{e^{-\nu w_i}} \tag{5.17}$$

$$v_i = \sqrt{e^{-\nu w_i}}. \tag{5.18}$$

The numerator (5.16) now becomes

$$h'(\nu)d(\nu) - h(\nu)d'(\nu) = -[\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{v}\boldsymbol{v}^T - (\boldsymbol{u}\boldsymbol{v}^T)^2] < 0 \tag{5.19}$$

where the inequality follows by the Cauchy-Schwarz inequality (A.1) and since by assumption (5.13), $w_i \neq w_j$ for at least one pair of i, j so that $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly independent.

5.5. Problems

Problem 5.1. Consider the optimization problem (5.6).

1. What is the maximum (minimum) constellation scaling Δ that is feasible, i.e., for which maximum (minimum) value of Δ can the power constraint $\mathbb{E}[(\Delta X)^2] \leq \mathsf{P}$ be satisfied?

2. Determine the corresponding input distributions.

3. Are these distributions MB distributions (5.9)? If yes, which values does the parameter ν take?

Problem 5.2. For the ASK constellation $\mathcal{X}$ and the power constraint P, suppose P_{X^*} and Δ^* are the distribution and the constellation scaling, respectively, that achieve capacity, i.e., they are a solution of (5.6). Define

$$P_{X^\sharp}(a) = \frac{P_{X^*}(a) + P_{X^*}(-a)}{2}, \quad a \in \mathcal{X}. \tag{5.20}$$

1. Show that $P_{X^\sharp}$ is symmetric, i.e., for each $a \in \mathcal{X}$, we have $P_{X^\sharp}(a) = P_{X^\sharp}(-a)$.

2. Show that $\mathbb{E}[(\Delta^* \cdot X^\sharp)^2] \leq \mathsf{P}$.

3. Define $P_{X^-}(a) = P_{X^*}(-a)$, $a \in \mathcal{X}$. Show that $\mathbb{I}(X^-; \Delta^* \cdot X^- + Z) = \mathbb{I}(X^*; \Delta^* \cdot X^* + Z)$.
   Hint: The noise pdf p_Z is symmetric.

4. Show that $\mathbb{I}(X^\sharp; \Delta^* \cdot X^\sharp + Z) \geq \mathbb{I}(X^*; \Delta^* \cdot X^* + Z)$.
   Hint: $\mathbb{I}(X; Y)$ is concave in P_X.

5. Conclude that ASK constellations in AWGN have symmetric capacity-achieving distributions.

Problem 5.3. Consider the finite set $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$. Let f be a function that assigns to each $x_i \in \mathcal{X}$ a positive cost $f(x_i) > 0$. Define the MB distribution

$$P_{X_\nu}(x_i) = A_\nu e^{-\nu f(x_i)}, \qquad A_\nu = \frac{1}{\sum_{i=1}^n e^{-\nu f(x_i)}}. \tag{5.21}$$

1. Let P_X be some distribution on $\mathcal{X}$ with $\mathbb{E}[f(X)] = \mathsf{P}$. Choose $\nu\colon \mathbb{E}[f(X_\nu)] = \mathsf{P}$. Show that $\mathbb{H}(X) \leq \mathbb{H}(X_\nu)$ with equality if and only if $P_X = P_{X_\nu}$.

2. Let P_X be some distribution on $\mathcal{X}$ with $\mathbb{H}(X) = \mathsf{H}$. Choose $\nu\colon \mathbb{H}(X_\nu) = \mathsf{H}$. Show that $\mathbb{E}[f(X)] \geq \mathbb{E}[f(X_\nu)]$ with equality if and only if $P_X = P_{X_\nu}$.

Note that (5.9) is an instance of (5.21) for $f(x_i) = |x_i|^2$.

Problem 5.4. Let X be a discrete random variable with a distribution P_X on $\mathcal{X}$. Consider a value H with $\mathbb{H}(X) < \mathsf{H} < \log_2 |\mathcal{X}|$. Characterize the solution of the optimization problem

$$\begin{aligned} \min_{P_Y} \quad & \mathbb{D}(P_Y \| P_X) && (5.22)\\ \text{subject to} \quad & \mathbb{H}(Y) \geq \mathsf{H}. && (5.23) \end{aligned}$$

How does the solution look like when P_X is a MB distribution (5.9)?
Hint: This is a convex optimization problem.

Problem 5.5. Let $\boldsymbol{B}$ denote the Gray label of a 2^m-ASK constellation. Let the MB distribution $P_{X^\clubsuit}$ induce a distribution $P_{\boldsymbol{B}}$ via $X^\clubsuit = x_{\boldsymbol{B}}$.

1. Show that $P_{B_1}(0) = P_{B_1}(1) = \frac{1}{2}$.

2. Show that B_1 and $B_2^m = B_2 \cdots B_m$ are independent, i.e., show that

$$P_{\boldsymbol{B}}(b\boldsymbol{b}) = P_{B_1}(b) P_{B_2^m}(\boldsymbol{b}), \quad \forall b \in \{0,1\},\ \boldsymbol{b} \in \{0,1\}^{m-1}. \tag{5.24}$$

6. Probabilistic Amplitude Shaping

In this chapter, we develop the basic probabilistic amplitude shaping (PAS) architecture for combining optimized input distributions with forward error correction (FEC). We focus on the transmitter side. In Chapter 8, we analyze the decoding metric employed at the receiver and in Chapter 9, we discuss the constant composition distribution matcher (CCDM) for emulating the shaped amplitude source from uniform source bits. In Chapter 10, we derive error exponents and achievable rates for PAS with CCDM.

6.1. Summary

- The AWGN channel with 2^m-ASK constellation and symmetric input distribution P_X, i.e., $P_X(x) = P_X(-x)$, is considered.

- Probabilistic amplitude shaping (PAS): The input distribution P_X can be implemented by a transmitter that concatenates a shaped amplitude source $\boxed{P_A}$, $A = |X|$, with a systematic binary encoder with code rate $c \geq \frac{m-1}{m}$.

- The PAS transmission rate is

$$\mathsf{R}_{\text{PAS}} = \mathbb{H}(X) - (1-c)m. \tag{6.1}$$

- The PAS transmission rate is achievable if

$$\mathbb{H}(X) - (1-c)m \leq \mathbb{I}(X;Y). \tag{6.2}$$

6.2. Preliminaries

Consider the AWGN channel

$$Y = \Delta X + Z \tag{6.3}$$

where Z is zero mean Gaussian with variance one and where X is 2^m-ASK input. In Section 5.3, we have seen that optimizing Δ, P_X over the family of MB distributions results in mutual informations that are close to the AWGN capacity, when the constellation size is large enough. For the optimized parameters $\Delta^\clubsuit$, $P_{X^\clubsuit}$, we want to develop a transmitter that enables reliable transmission with a rate close to $\mathbb{I}(X^\clubsuit;Y^\clubsuit)$. Suppose our transmitter encodes message W to $X^n = x^n(W)$. Suppose further the message consists of $k = nR$ uniformly distributed bits. Then, by Theorem 1 (Channel Coding Converse), if

$$R > \frac{\sum_{i=1}^{n} \mathbb{I}(X_i; Y_i)}{n} \tag{6.4}$$

the error probability $\Pr(W \neq \hat{W})$ is bounded away from zero, for any decoding function $\hat{W} = f(Y^n)$. We therefore build a transmitter with

$$P_{X_i} = P_{X^\clubsuit}, \quad i = 1, 2, \dots, n. \tag{6.5}$$

For such a transmitter, the right-hand side of (6.4) is equal to $\mathbb{I}(X^\clubsuit; Y^\clubsuit)$, so that reliable transmission with rates close to $\mathbb{I}(X^\clubsuit; Y^\clubsuit)$ is not ruled out by (6.4). We make the following two observations:

Amplitude-Sign Factorization

We can write $X^\clubsuit$ as

$$X^\clubsuit = A \cdot S \tag{6.6}$$

where $A = |X^\clubsuit|$ is the *amplitude* of the input and where $S = \operatorname{sign}(X^\clubsuit)$ is the *sign* of the input. By (5.1), the amplitudes take values in

$$\mathcal{A} := \{1, 3, \dots, 2^m - 1\}. \tag{6.7}$$

We see from (5.9) that the distribution $P_{X^\clubsuit}$ is symmetric around zero, i.e., we have

$$P_{X^\clubsuit}(x) = P_{X^\clubsuit}(-x) \tag{6.8}$$

and therefore, A and S are stochastically independent and S is uniformly distributed, i.e., we have

$$P_{X^\clubsuit}(x) = P_A(|x|) \cdot P_S(\operatorname{sign}(x)), \quad \forall x \in \mathcal{X} \tag{6.9}$$

$$P_S(1) = P_S(-1) = \frac{1}{2}. \tag{6.10}$$

We call this property of $P_{X^\clubsuit}$ *amplitude-sign factorization.*

Uniform Check Bit Assumption

The second observation is on *systematic binary encoding*. A systematic *generator matrix* of an (n_c, k_c) binary code has the form

$$\boldsymbol{G} = [\boldsymbol{I}_{k_c} | \boldsymbol{P}] \tag{6.11}$$

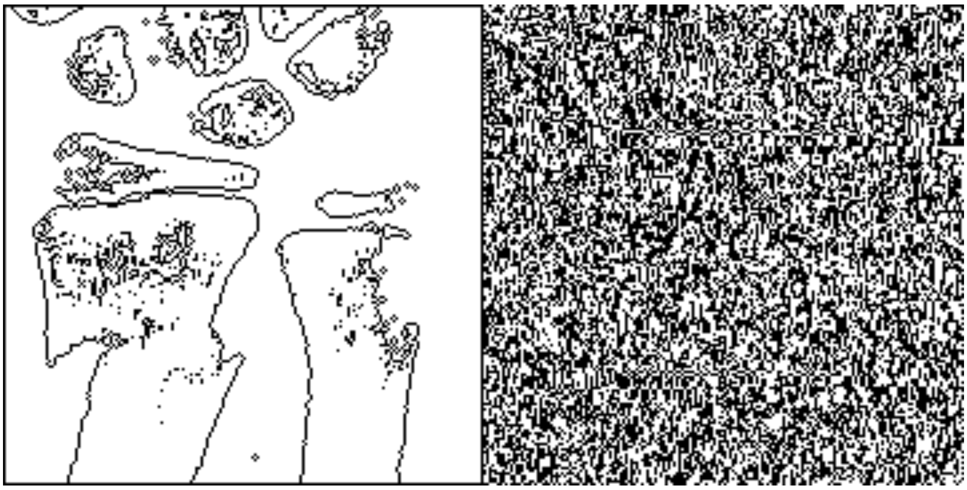


Figure 6.1.: The black and white pixels of a 180 × 180 picture are represented by 1s and 0s, respectively, and then encoded by a DVB-S2 rate 1/2 LDPC code. The resulting check bits are displayed next to the picture. The empirical distribution of the original picture is $P_{\bar{D}}(1) = 1 - P_{\bar{D}}(0) = 0.1082$ and the empirical distribution of the check bits is $P_{\bar{R}}(1) = 1 - P_{\bar{R}}(0) = 0.4970$.

where $\boldsymbol{I}_{k_c}$ is the $k_c \times k_c$ identity matrix and $\boldsymbol{P}$ is a $k_c \times (n_c - k_c)$ matrix. $\boldsymbol{P}$ is the *parity forming part* of $\boldsymbol{G}$. The generator matrix $\boldsymbol{G}$ maps k_c data bits D^{k_c} to a length n_c code word via

$$D^{k_c}\boldsymbol{G} = (D^{k_c}|R^{n_c-k_c}) \tag{6.12}$$

where $R^{n_c-k_c}$ are redundant bits that are modulo-two sums of data bits. Suppose the data bits have some distribution $P_{D^{k_c}}$. Since the encoding is systematic, this distribution is preserved at the output of the encoder. What is the distribution of the redundancy bits? To address this question, consider two independent data bits D_1 and D_2. The modulo-two sum $R = D_1 \oplus D_2$ is then more uniformly distributed than the individual summands D_1 and D_2, see Problem 6.1. This suggests that if the redundancy bits are the modulo-two sum of a large enough number of data bits, then their distribution is close to uniform. An example of this phenomenon is shown in Figure 6.1. We therefore follow [16, Section VI],[17, Chapter 5],[18, Section 7.1] and assume in the following that the redundancy bits are uniformly distributed and we call this assumption the *uniform check bit assumption*. Note that in Chapter 10, we derive with mathematical rigor error exponents and achievable rates for PAS without resorting to the uniform check bit assumption. The achievable rates in Chapter 10 coincide with the achievable rates that we state in the present chapter.

6.3. Encoding Procedure

Consider block transmission with n symbols from a 2^m-ASK constellation. Since we use binary error correcting codes, we label each of the 2^{m-1} amplitudes by a binary string

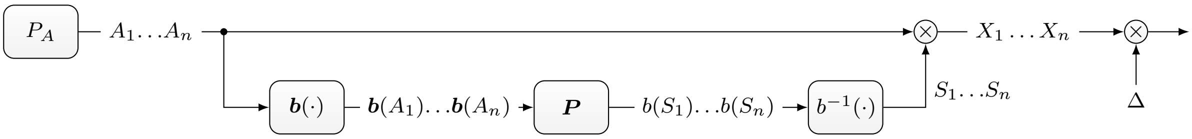


Figure 6.2.: PAS. The ASK amplitudes A_i take values in $\mathcal{A} = \{1, 3, \dots, 2^m - 1\}$. The amplitudes A_i are represented by their binary labels $\boldsymbol{b}(A_i)$. Redundancy bits $b(S_i)$ result from multiplying the binary string $\boldsymbol{b}(A_1)\boldsymbol{b}(A_2) \cdots \boldsymbol{b}(A_n)$ by the parity forming part $\boldsymbol{P}$ of a systematic generator matrix $[\boldsymbol{I}_{k_c} | \boldsymbol{P}]$. The redundancy bits $b(S_i)$ are transformed into signs S_i and multiplied with the amplitudes A_i. The resulting signal points $X_i = A_i S_i$ take values in $\mathcal{X} = \{\pm 1, \pm 3, \dots, \pm(2^m - 1)\}$. The signal points X_i are scaled by Δ and ΔX_i is transmitted over the channel.

of length $m - 1$ and we label each of the signs ± 1 by a bit, i.e., we use

$$A \mapsto \boldsymbol{b}(A) \in \{0, 1\}^{m-1} \tag{6.13}$$

$$S \mapsto b(S) \in \{0, 1\}. \tag{6.14}$$

For the sign, we use $b(-1) = 0$ and $b(1) = 1$. The choice of $\boldsymbol{b}(A)$ influences the rates that can be achieved by bit-metric decoding (BMD), i.e., the combination of a binary demapper with a binary decoder at the receiver. We discuss this in detail in Section 8.1.2. We use a rate $k_c/n_c = (m-1)/m$ binary code with systematic generator matrix $\boldsymbol{G} = [\boldsymbol{I}_{k_c} | \boldsymbol{P}]$. For block transmission with n channel uses , the block length of the code is $n_c = nm$ and the dimension of the code is $k_c = n(m-1)$. The encoding procedure is displayed in Figure 6.2. It works as follows.

1. A discrete memoryless source (DMS) $\boxed{P_A}$ outputs amplitudes $A_1, A_2, \dots, A_n$ that are iid according to P_A. We explain in Chapter 9 how the DMS $\boxed{P_A}$ can be emulated from binary data by distribution matching.

2. Each amplitude A_i is represented by its label $\boldsymbol{b}(A_i)$.

3. The resulting length $(m-1)n = k_c$ binary string is multiplied by the parity forming part $\boldsymbol{P}$ of $\boldsymbol{G}$ to generate $n_c - k_c = n$ sign labels $b(S_1)$, $b(S_2), \dots, b(S_n)$.

4. Each sign label $b(S_i)$ is transformed into the corresponding sign S_i.

5. The signal $X_i = A_i \cdot S_i$ is scaled by Δ and transmitted.

We call this procedure probabilistic amplitude shaping (PAS). Since the signs S^n are a deterministic function of the amplitudes A^n, the input symbols $X_1, X_2, \dots, X_n$ are correlated. Under the uniform check bit assumption, the marginal distributions are

$$P_{X_i}(x_i) = P_A(|x_i|) P_{S_i}(\operatorname{sign}(x_i)) \tag{6.15}$$

$$= P_A(|x_i|) \frac{1}{2} \tag{6.16}$$

$$= P_{X^\clubsuit}(x_i) \tag{6.17}$$

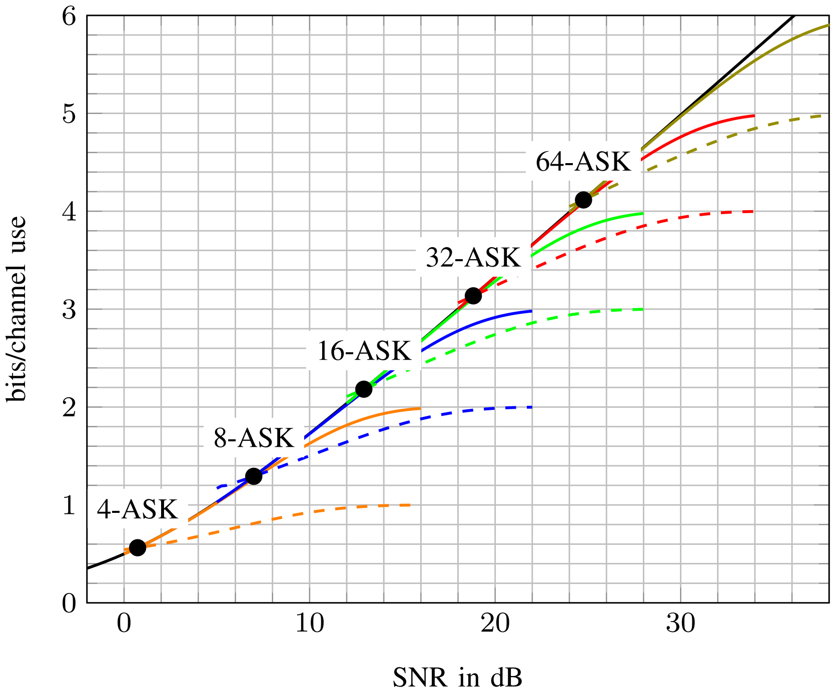


Figure 6.3.: The mutual information curves (solid) and the transmission rate curves (dashed) for ASK. The optimal operating points for rate $(m-1)/m$ codes are indicated by dots.

that is, if the uniform check bit assumption holds, then PAS has the desired property (6.5).

6.4. Optimal Operating Points

We study the rates at which reliable transmission is possible with our scheme. By (6.4), reliable communication at rate R is achievable only if

$$R < \frac{\sum_{i=1}^{n} \mathbb{I}(X_i; Y_i)}{n} = \mathbb{I}(X; Y) = \mathbb{I}(AS; Y). \tag{6.18}$$

Since A^n represents our data, our transmission rate is

$$R = \frac{\mathbb{H}(A^n)}{n} = \mathbb{H}(A) \quad \left[\frac{\text{bits}}{\text{channel use}}\right] \tag{6.19}$$

and condition (6.18) becomes

$$\mathbb{H}(A) < \mathbb{I}(AS; Y). \tag{6.20}$$

In Figure 6.3, both the mutual information $\mathbb{I}(AS; Y)$ (solid lines) and transmission rate $\mathbb{H}(A)$ (dashed lines) are displayed for $4, 8, 16, 32$, and 64-ASK. For high enough SNR, the

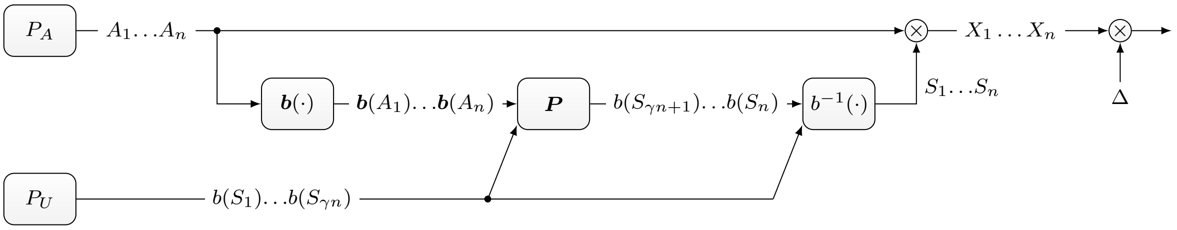


Figure 6.4.: Extension of PAS to code rates higher than $(m-1)/m$. The fraction γ of the signs is used for data, which is modelled as the output of a Bernoulli-1/2 DMS P_U.

mutual information saturates at m bits and the transmission rate saturates at $m-1$ bits. Optimal error correction for block length $n \to \infty$ would operate where the transmission rate curve crosses the mutual information curve. These crossing points are indicated by dots in Figure 6.3. Since the code rate is $c = (m-1)/m$, the transmission rate curve can be written as

$$\mathsf{R}_{\text{PAS}} = \mathbb{H}(A) = \mathbb{H}(X) - 1 = \mathbb{H}(X) - (1-c)m \quad \left[\frac{\text{bits}}{\text{channel use}}\right]. \tag{6.21}$$

6.5. PAS for Higher Code Rates

We observe in Figure 6.3 that the ASK mutual information curves stay close to the capacity $\mathsf{C}(\mathsf{P}/\sigma^2)$ over a certain range of rates above the optimal operating points. We therefore extend our PAS scheme to enable the use of code rates higher than $(m-1)/m$ on 2^m-ASK constellations. We achieve this by using some of the signs S_i for uniformly distributed data bits. We illustrate this extension of the PAS scheme in Figure 6.4. Let γ denote the fraction of signs used for data bits. We interpret γn uniformly distributed data bits as sign labels $b(S_1) \cdots b(S_{\gamma n})$. These γn bits and the $(m-1)n$ bits from the amplitude labels are encoded by the parity forming part of a systematic rate c generator matrix, which generates the remaining $(1-\gamma)n$ sign labels. The code rate can be expressed in terms of m and γ as

$$c = \frac{m-1+\gamma}{m}. \tag{6.22}$$

For a given code rate c, the fraction γ is given by

$$\gamma = 1-(1-c)m. \tag{6.23}$$

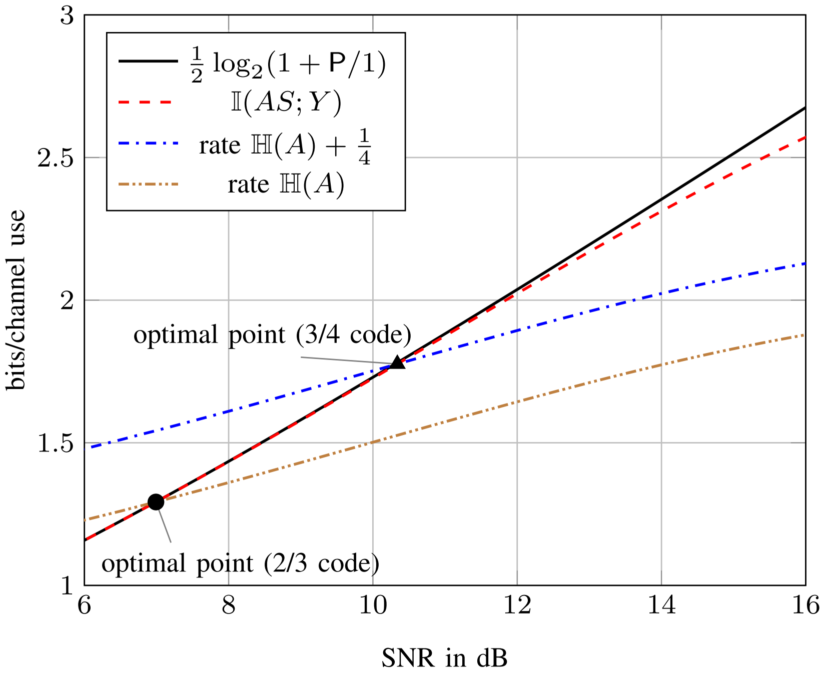


Figure 6.5.: Optimal operating points of 8-ASK for PAS ($c = 2/3$) and extended PAS ($c = 3/4$).

Since a fraction γ of the signs now carries information, the transmission rate of the extended PAS scheme is given by

$$\begin{aligned}\mathsf{R}_{\text{PAS}} &= \frac{\mathbb{H}(A^n)+\mathbb{H}(S^{\gamma n})}{n} = \mathbb{H}(A)+\gamma \\ &= \mathbb{H}(X)-1+1-(1-c)m \\ &= \mathbb{H}(X)-(1-c)m \quad \left[\frac{\text{bits}}{\text{channel use}}\right]. \end{aligned} \tag{6.24}$$

The optimal operating point is then given by the crossing of the rate curve $\mathbb{H}(A)+\gamma$ and the mutual information curve. In Figure 6.5, we display for 8-ASK the optimal operating points for $c = 2/3$ and $c = 3/4$.

6.6. PAS Example

Example 6.1 (PAS)**.** Suppose we use 4-ASK with the constellation $\mathcal{X} = \{-3,-1,1,3\}$. We consider block transmission with a block length of $n = 4$ channel uses. The desired input distribution is

$$P_X(-1) = P_X(1) = \frac{3}{8}, \quad P_X(-3) = P_X(3) = \frac{1}{8}. \tag{6.25}$$

The possible amplitudes are 1 and 3, which should occur with probabilities

$$P_A(1) = \frac{3}{4}, \quad P_A(3) = \frac{1}{4}. \tag{6.26}$$

For the amplitudes and the sign, we respectively use the labelings

$$\boldsymbol{b}(1) = 1, \quad \boldsymbol{b}(3) = 0 \tag{6.27}$$

$$b(-1) = 0, \quad b(1) = 1. \tag{6.28}$$

The labeling of the signal points in $\mathcal{X}$ is $\text{label}(x) = b[\text{sign}(x)]\boldsymbol{b}(|x|)$, e.g., $\text{label}(-3) = 00$ and $\text{label}(1) = 11$. To emulate an amplitude DMS, we now introduce the idea of a distribution matcher (DM): Our data are two independent and uniformly distributed bits $D_1 D_2$. We map the data bits to sequences of amplitudes by the mapping

$$00 \mapsto (3,1,1,1) =: \boldsymbol{a}(1) \tag{6.29}$$

$$01 \mapsto (1,3,1,1) =: \boldsymbol{a}(2) \tag{6.30}$$

$$10 \mapsto (1,1,3,1) =: \boldsymbol{a}(3) \tag{6.31}$$

$$11 \mapsto (1,1,1,3) =: \boldsymbol{a}(4). \tag{6.32}$$

This mapping is an instance of the constant composition distribution matcher (CCDM), which we discuss in detail in Chapter 9. By this mapping, each amplitude A_i, $i = 1,2,3,4$ indeed has the desired amplitude distribution P_A, i.e.,

$$P_{A_i}(1) = 1 - P_{A_i}(3) = \frac{3}{4}. \tag{6.33}$$

The binary representation of the amplitudes are by (6.27)

$$\boldsymbol{b}(1) = (0,1,1,1) \tag{6.34}$$

$$\boldsymbol{b}(2) = (1,0,1,1) \tag{6.35}$$

$$\boldsymbol{b}(3) = (1,1,0,1) \tag{6.36}$$

$$\boldsymbol{b}(4) = (1,1,1,0). \tag{6.37}$$

We use the binary linear block code with systematic generator matrix

$$\boldsymbol{G} = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{array}\right] = [\boldsymbol{I}|\boldsymbol{P}]. \tag{6.38}$$

The resulting redundancy vectors are

$$
\begin{aligned}
\boldsymbol{r}(1) &= \boldsymbol{b}(1)\boldsymbol{P} = (1,0,0,1) && (6.39)\\
\boldsymbol{r}(2) &= \boldsymbol{b}(2)\boldsymbol{P} = (0,1,1,0) && (6.40)\\
\boldsymbol{r}(3) &= \boldsymbol{b}(3)\boldsymbol{P} = (0,1,0,1) && (6.41)\\
\boldsymbol{r}(4) &= \boldsymbol{b}(4)\boldsymbol{P} = (1,0,1,0). && (6.42)
\end{aligned}
$$

We apply the inverse labeling function b^{-1} to the redundancy vectors to obtain the sign vectors

$$
\begin{aligned}
\boldsymbol{s}(1) &= (1,-1,-1,1) && (6.43)\\
\boldsymbol{s}(2) &= (-1,1,1,-1) && (6.44)\\
\boldsymbol{s}(3) &= (-1,1,-1,1) && (6.45)\\
\boldsymbol{s}(4) &= (1,-1,1,-1). && (6.46)
\end{aligned}
$$

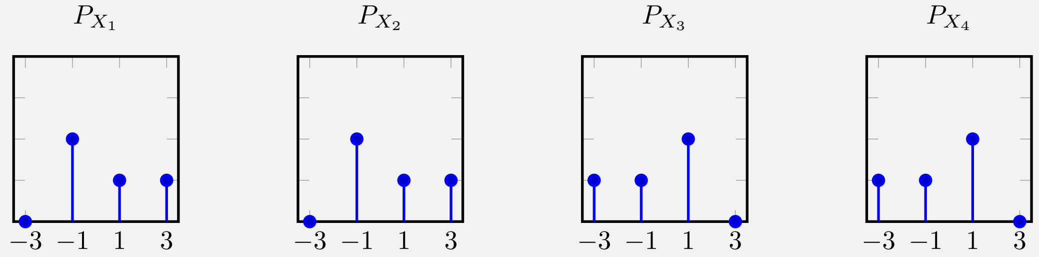
The resulting sign distribution is the desired uniform distribution

$$P_{S_i}(1) = P_{S_i}(0) = \frac{1}{2}, \quad i = 1,2,3,4. \tag{6.47}$$

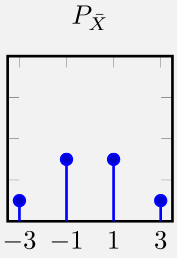
By entrywise multiplying the amplitude vectors and the sign vectors, we finally obtain the signals

$$
\begin{aligned}
\boldsymbol{x}(1) &= (3,-1,-1,1) && (6.48)\\
\boldsymbol{x}(2) &= (-1,3,1,-1) && (6.49)\\
\boldsymbol{x}(3) &= (-1,1,-3,1) && (6.50)\\
\boldsymbol{x}(4) &= (1,-1,1,-3). && (6.51)
\end{aligned}
$$

The marginal distributions of the transmitted signal points are



As we can see, the distributions deviate from the target distribution (6.25). *Note:* By averaging over the code symbols, we get



which is actually equal to the target distribution (6.25).

6.7. Problems

Problem 6.1. Let D_1 and D_2 be two independent binary random variables with distributions P and Q, respectively, and define $R = D_1 \oplus D_2$, where $\oplus$ denotes modulo two addition. Without loss of generality, assume that

$$P(0) \leq P(1), \quad Q(0) \leq Q(1). \tag{6.52}$$

Show that P_R is more uniform than P and Q, i.e., show that

$$\min\{P(1), Q(1)\} \geq P_R(0) \geq \max\{P(0), Q(0)\}. \tag{6.53}$$

Problem 6.2. Let $X^n = X_1 \dots X_n$ be independent and uniformly distributed binary random variables. Let Z be a binary random variable independent of X^n. Define Y^n by $Y_i \oplus Z$, $i = 1, 2, \dots, n$, where $\oplus$ denotes modulo two addition. Calculate $\mathbb{I}(X^n; Y^n)$.

Problem 6.3. The discrete Fourier transform (DFT) of vectors in the two-dimensional vector space $\mathbf{R} \times \mathbf{R}$ is given by

$$\boldsymbol{p} = (p_1, p_2) \multimap \tilde{\boldsymbol{p}} := (p_1 + p_2, p_1 - p_2) \tag{6.54}$$

$$= (p_1, p_2) \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{6.55}$$

The *circular convolution* of two vectors in $\mathbf{R}^2$ is

$$\boldsymbol{p} \star \boldsymbol{q} = (p_1 q_1 + p_2 q_2, p_2 q_1 + p_1 q_2). \tag{6.56}$$

1. Calculate the DFT of the uniform distribution $P_U = [P_U(0), P_U(1)] = (1/2, 1/2)$.

2. Show the correspondence

$$\boldsymbol{p} \star \boldsymbol{q} \multimap \tilde{\boldsymbol{p}} \circ \tilde{\boldsymbol{q}} := (\tilde{p}_1 \tilde{q}_1, \tilde{p}_2 \tilde{q}_2). \tag{6.57}$$

3. Let B_1 and B_2 be two independent binary random variables with $P_{B_1} = P_{B_2} = P_B$. Define $R = B_1 \oplus B_2$. Show that $P_R = P_B \star P_B$ and calculate the DFT of P_R.

4. Consider now $R = B_1 \oplus B_2 \oplus \cdots \oplus B_d$ with the B_i iid with $P_{B_i} = P_B$ and $P_B(0) \neq 0$ and $P_B(1) \neq 0$. Calculate the DFT $\tilde{P}_R$ of P_R and show that $\tilde{P}_R \overset{d\to\infty}{\to} \tilde{P}_U$. Conclude that R is approximately uniformly distributed for large enough d.

Problem 6.4.

4-ASK constellation	$\{-3, -1, 1, 3\}$
number of channel uses	$n = 20$
data sign bit fraction	$\gamma = \frac{1}{2}$

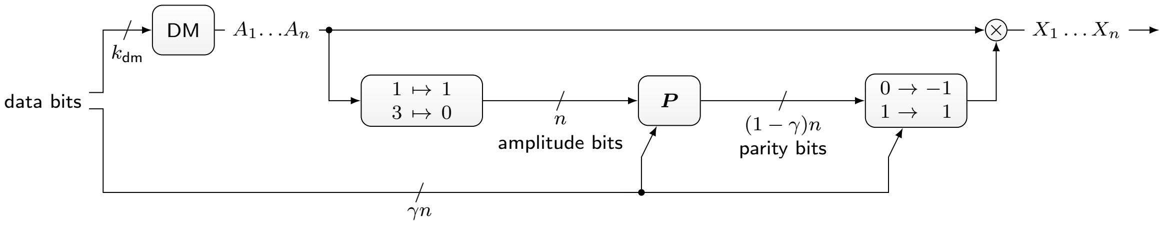


Figure 6.6.: PAS transmitter.

In the amplitude sequences at the distribution matcher (DM) output, the amplitude 1 appears 15 times and the amplitude 3 appears 5 times.

1. How many bits k_{dm} can the DM encode?

   Assume from now on $k_{\text{dm}} = 12$ bits.

2. What is the transmission power P of the system?

3. The symbols X_i, $i = 1, 2, \ldots, n$ are transmitted over an AWGN channel and the output observed at the receiver is

$$Y_i = X_i + Z_i, \qquad i = 1, 2, \ldots, 20$$

   where the Z_i are independent of the input and independent and identically distributed according to a zero mean Gaussian density with variance σ^2. For which value of σ^2 is the signal-to-noise-ratio (SNR) equal to 6 dB?

4. Calculate the transmission rate of the system and compare it to the AWGN capacity at 6 dB.

5. You have two binary forward error correction (FEC) codes, one with rate 2/3 and one with rate 3/4. Which of of the two codes is used in the system and what is the dimension of the matrix $\boldsymbol{P}$ in Figure 6.6?

6. The matcher is replaced by the mapping $0 \mapsto 3$, $1 \mapsto 1$. To operate at the same transmission rate and the same transmission power as the former system, the FEC code rate is changed to c and the symbols X_i are scaled by Δ prior to transmission. Calculate c and Δ. You can assume that the data bits are uniformly distributed.

7. Achievable Rates

In this chapter, we take an information-theoretic perspective and use random coding arguments following Gallager's error exponent approach [3, Chapter 5] to derive achievable rates for general channels and decoding metrics. The results derived in this chapter form the foundation for several other chapters.

- In Chapter 8, we instantiate the general metric of this chapter for a number of specific decoding metrics, including bit-metrics, interleaving, and hard-decision decoding.
- In Chapter 10, we derive error exponents and achievable rates for PAS, a practical transceiver architecture developed in Chapter 6.
- In Chapter 11, we discuss how to estimate achievable rates for channels that potentially have memory.

We consider a layered architecture consisting of a FEC layer, where the receiver detects the transmitted code word, and a shaping layer, whose task is to encode into code words that have a desired distribution.

7.1. FEC Layer

In Figure 7.1, we display the random coding experiment for analyzing the FEC layer. The details are as follows.

- The channel is discrete-time with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$. We derive our results assuming a continuous-valued output. Our results also apply for discrete output alphabets.
- Random coding: For indices $w = 1, 2, \ldots, |\mathcal{C}|$, we generate code words $C^n(w)$ with the $n|\mathcal{C}|$ entries independent and uniformly distributed on $\mathcal{X}$. The code is
$$\mathcal{C} = \{C^n(1), C^n(2), \ldots, C^n(|\mathcal{C}|)\}. \tag{7.1}$$
- The code rate is $R_\text{c} = \frac{\log_2(|\mathcal{C}|)}{n}$ and equivalently, we have $|\mathcal{C}| = 2^{nR_\text{c}}$ code words.
- We consider a non-negative decoding metric q on $\mathcal{X} \times \mathcal{Y}$ and we define the memoryless metric
$$q^n(x^n, y^n) := \prod_{i=1}^{n} q(x_i, y_i), \quad x^n \in \mathcal{X}^n, y^n \in \mathcal{Y}^n. \tag{7.2}$$

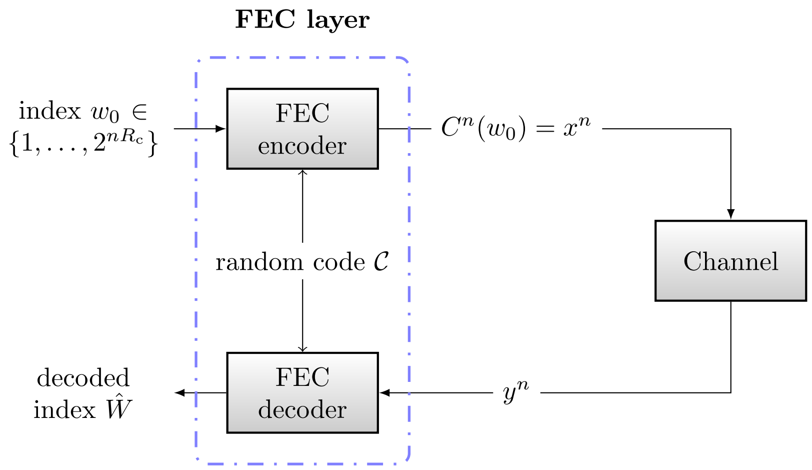


Figure 7.1.: The random coding experiment for bounding the decoding error probability of the FEC layer.

For the channel output y^n, we let the receiver decode with the rule

$$\hat{W} = \operatorname*{argmax}_{w\in\{1,\ldots,2^{nR_\mathrm{c}}\}} \prod_{i=1}^{n} q\left[C_i(w), y_i\right]. \tag{7.3}$$

- We consider the decoding error probability

$$P_e = \Pr\left[\hat{W} \neq w_0 \middle| C^n(w_0) = x^n, Y^n = y^n\right] \tag{7.4}$$

where w_0 is the index of the transmitted code word, $C^n(w_0) = x^n$ is the transmitted code word, y^n is the channel output sequence, and $\hat{W}$ is the decoded index at the receiver. Note that the code words $C^n(w)$, $w \neq w_0$ against which the decoder attempts to decode are random and the transmitted code word $C^n(w_0) = x^n$ and the channel output y^n are deterministic.

7.1.1. Achievable Code Rate

Theorem 4. *Suppose code word $C^n(w_0) = x^n$ is transmitted and let y^n be a channel output sequence. With high probability for large n, the decoder can recover the index w_0 from the sequence y^n if*

$$R_\mathrm{c} < \hat{T}_\mathrm{c}(x^n, y^n, q) = \frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{q(x_i, y_i)}{\sum_{a\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(a, y_i)} \tag{7.5}$$

$$= \log_2 |\mathcal{X}| - \underbrace{\frac{1}{n}\sum_{i=1}^{n}\left[-\log_2 \frac{q(x_i, y_i)}{\sum_{a\in\mathcal{X}} q(a, y_i)}\right]}_{\text{uncertainty}} \tag{7.6}$$

- The factor $1/|\mathcal{X}|$ in (7.5) reflects that the code word entries are generated uniformly at random in the random coding experiment.
- In (7.6), the achievable code rate is split into $\log_2|\mathcal{X}|$, which corresponds to the rate that can be achieved in the absence of noise, and the uncertainty, which quantifies how much we need to back off in terms of code rate because of the presence of noise.

Example 7.1. Consider finite input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$, respectively, and let $x^n \in \mathcal{X}^n$, $y^n \in \mathcal{Y}^n$ be two arbitrary sequences. Suppose the decoder knows the empirical input distribution P_{x^n} (see Appendix C.1) given by

$$P_{x^n}(a) = \frac{|\{i\colon x_i = a\}|}{n}. \tag{7.7}$$

It measures the output distribution P_{y^n} and uses as metric

$$q(a,b) = P_{x^n}(a)P_{y^n}(b), \quad a \in \mathcal{X}, b \in \mathcal{Y}. \tag{7.8}$$

For this metric, the achievable code rate is

$$\hat{T}_\mathrm{c}(x^n, y^n, P_{x^n}P_{y^n}) = \mathbb{D}(P_{x^n}\|P_U) = \mathbb{H}(P_U) - \mathbb{H}(P_{x^n}) \tag{7.9}$$

where P_U is the uniform distribution on $\mathcal{X}$. By the information inequality (C.21), the achievable code rate is positive for non-uniform input x^n, and is zero for uniform input. Suppose next the decoder knows the empirical channel law $P_{y^n|x^n}$ and uses the metric

$$q(a,b) = P_{x^n}(a)P_{y^n|x^n}(b|a), \quad a \in \mathcal{X}, b \in \mathcal{Y}. \tag{7.10}$$

The achievable code rate now becomes

$$\hat{T}_\mathrm{c}(x^n, y^n, P_{x^n}P_{y^n|x^n}) = \mathbb{H}(P_U) - \mathbb{H}(P_{x^n|y^n}|P_{y^n}). \tag{7.11}$$

Note that since conditioning does not increase entropy (C.15), we have

$$\hat{T}_\mathrm{c}(x^n, y^n, P_{x^n}P_{y^n}) \le \hat{T}_\mathrm{c}(x^n, y^n, P_{x^n}P_{y^n|x^n}). \tag{7.12}$$

Example 7.2. Consider an AWGN channel with BPSK, i.e., the input alphabet is $\mathcal{X} = \{-1, 1\}$. Let x^n be some arbitrary BPSK sequence and let

$$y^n = x^n + z^n \tag{7.13}$$

where z^n is a real-valued noise sample sequence. The decoder uses the decoding metric

$$q(a,b) = P_X(a)e^{ab}, \quad a \in \{-1,1\}, b \in \mathbf{R}. \tag{7.14}$$

By Problem 7.1, the achievable code rate is

$$\hat{T}_\mathrm{c}(x^n, y^n, q) = \mathbb{H}(P_U) - \frac{1}{n}\sum_{i=1}^{n}\left[-\log_2 P_{X|Y}(x_i|y_i)\right] \tag{7.15}$$

where XY are random variables distributed according to

$$P_X(a)p_{Y|X}(b|a) = P_X(a)p_Z(b-a), \quad a \in \{-1,1\}, y \in \mathbf{R} \tag{7.16}$$

and where p_Z is a zero mean Gaussian density with variance $\sigma^2 = 1$.

Example 7.3. Suppose for a channel measurement x^n, y^n, we want to estimate the achievable code rate of a specific binary low-density parity-check (LDPC) code with a soft-decision decoder. Assume the channel input alphabet is binary, i.e., $\mathcal{X} = \{0,1\}$. For each channel use $i = 1, 2, \dots, n$, we calculate the L-value

$$L_i = \log\frac{q(y_i, 0)}{q(y_i, 1)}. \tag{7.17}$$

We discuss decoding metrics in Chapter 8 and L-values in Problem 8.1. For now, simply note that if $L_i > 0$, the decoding metric prefers 0 over 1 and if $L_i < 0$, then it prefers 1 over 0. The metric of a specific binary code word b^n is

$$\sum_{i=1}^{n} L_i(-1)^{b_i} \tag{7.18}$$

and the decoder looks for the code word b^n that has the highest metric. If the transmitted sequence x^n was a code word, then we can pass L^n to the decoder and if the decoded code word is equal to x^n, then the code rate of the LDPC code is achievable for the measurement x^n, y^n.

Suppose now that the transmitted sequence x^n was not a code word. This situation occurs in particular when we want to determine achievable code rates for a given measurement by trying out different codes of different rates. We can use the following trick: for a code word b^n and the transmitted sequence x^n, we calculate the *scrambling sequence*

$$s^n = (x_1 \oplus b_1)\cdots(x_n \oplus b_n). \tag{7.19}$$

Note that this implies that adding the scrambling sequence to the transmitted sequence results in a code word, because $x_i \oplus s_i = b_i$. We now transform the soft information L^n by

$$\tilde{L}_i = L_i(-1)^{s_i}, \quad i = 1, \dots, n \tag{7.20}$$

and we pass $\tilde{L}^n$ to the decoder. We then add s^n to the decoder output $\hat{b}^n$ and we check if the result is equal to the transmitted sequence x^n. This procedure can also be used when the channel input alphabet is not binary, for instance by using bit-metric decoding (BMD), see Section 8.1.2.

Proof of Theorem 4. We consider the setup in Figure 7.1, i.e., we condition on the event that index w_0 was encoded to $C^n(w_0) = x^n$ and that sequence y^n was output by the channel. For notational convenience, we assume without loss of generality that $w_0 = 1$. We have the implications

$$\begin{aligned} \hat{W} \neq 1 &\Rightarrow \hat{W} = w' \neq 1 && (7.21)\\ &\Rightarrow L(w') := \frac{q^n(C^n(w'), y^n)}{q^n(x^n, y^n)} \geq 1 && (7.22)\\ &\Rightarrow \sum_{w=2}^{|\mathcal{C}|} L(w) \geq 1. && (7.23) \end{aligned}$$

If event $\mathcal{A}$ implies event $\mathcal{B}$, then $\Pr(\mathcal{A}) \leq \Pr(\mathcal{B})$. Therefore, we have

$$\begin{aligned} \Pr(\hat{W} \neq 1 | C^n(1) = x^n, Y^n = y^n) &\leq \Pr\left[\sum_{w=2}^{|\mathcal{C}|} L(w) \geq 1 \middle| C^n(1) = x^n, Y^n = y^n\right] && (7.24)\\ &\leq \mathbb{E}\left[\sum_{w=2}^{|\mathcal{C}|} L(w) \middle| C^n(1) = x^n, Y^n = y^n\right] && (7.25)\\ &= q^n(x^n, y^n)^{-1}\,\mathbb{E}\left[\sum_{w=2}^{|\mathcal{C}|} q^n(C^n(w), y^n)\right] && (7.26)\\ &= (|\mathcal{C}| - 1) q^n(x^n, y^n)^{-1}\,\mathbb{E}\left[q^n(C^n, y^n)\right] && (7.27)\\ &\leq |\mathcal{C}| q^n(x^n, y^n)^{-1}\,\mathbb{E}\left[q^n(C^n, y^n)\right] && (7.28)\\ &= |\mathcal{C}| \frac{1}{\prod_{i=1}^n q(x_i, y_i)} \prod_{i=1}^n \mathbb{E}\left[q(C, y_i)\right] && (7.29)\\ &= |\mathcal{C}| \frac{1}{\prod_{i=1}^n q(x_i, y_i)} \prod_{i=1}^n \sum_{a \in \mathcal{X}} |\mathcal{X}|^{-1} q(a, y_i) && (7.30)\\ &= |\mathcal{C}| \prod_{i=1}^n \frac{\sum_{a \in \mathcal{X}} q(a, y_i)}{q(x_i, y_i)|\mathcal{X}|} && (7.31) \end{aligned}$$

where

- Inequality in (7.25) follows by Markov's inequality (B.6).
- Equality in (7.26) follows because for $w \neq 1$, the code word $C^n(w)$ and the transmitted code word $C^n(1)$ were generated independently so that $C^n(w)$ and $[C^n(1), Y^n]$ are independent.
- Equality in (7.27) holds because in our random coding experiment, for each index w, we generated the code word entries $C_1(w), C_2(w), \dots, C_n(w)$ iid.
- In (7.29), we used (7.2), i.e., that q^n defines a memoryless metric.

We can now write this as

$$\Pr(\hat{W} \neq 1 | C^n(1) = x^n, Y^n = y^n) \leq 2^{-n[\hat{T}_\mathrm{c}(x^n, y^n, q) - R_\mathrm{c}]} \tag{7.32}$$

$$\text{where } \hat{T}_\mathrm{c}(x^n, y^n, q) = \frac{1}{n} \sum_{i=1}^{n} \log_2 \frac{q(x_i, y_i)}{\sum_{a \in \mathcal{X}} \frac{1}{|\mathcal{X}|} q(a, y_i)} \tag{7.33}$$

For large n, the error probability upper bound is vanishingly small if

$$R_\mathrm{c} < \hat{T}_\mathrm{c}(x^n, y^n, q). \tag{7.34}$$

Thus, $\hat{T}_\mathrm{c}(x^n, y^n, q)$ is an achievable code rate, i.e., for a random code $\mathcal{C}$, if (7.34) holds, then sequence x^n can be decoded reliably from y^n with high probability. □

7.1.2. Memoryless Processes

We take a little detour and review memoryless random processes and their basic properties, which we need in the following sections of this chapter. Consider the following statement:

> If I toss a coin, the probability that I get "heads" is 1/2.

What does this mean? If I toss a coin once, I get either "heads" or "tails" but nothing in-between. To make this statement meaningful, we need to throw the coin many times. We can then count the total number of coin tosses and the number of times we got "heads". Suppose we tossed the coin n times and let n_heads be the number of times the outcome was "heads". Our probabilistic model $\Pr(\text{"heads"}) = \frac{1}{2}$ is reasonable if the observed *relative frequency* n_heads/n is close to 1/2, i.e., if

$$\frac{n_\text{heads}}{n} \approx \Pr(\text{"heads"}) = \frac{1}{2}. \tag{7.35}$$

Weak Law of Large Numbers

The weak law of large numbers (WLLN) makes the qualitative statement (7.35) precise.

Theorem 5 (WLLN)**.** *For each integer n, let $S_n = X_1 + \cdots + X_n$ where $X_1, X_2, \ldots$ are iid random variables with distribution P_X satisfying $\mathbb{E}(|X|^2) < \infty$. Then for any $\epsilon > 0$*

$$\lim_{n\to\infty} \Pr\left\{\left|\frac{S_n}{n} - \mathbb{E}(X)\right| > \epsilon\right\} = 0. \tag{7.36}$$

Proof. See, e.g., [19, Section 1.7.4]. □

When (7.36) holds, we say that S_n/n *converges to* $\mathbb{E}(X)$ *in probability* and we write

$$\frac{S_n}{n} \xrightarrow{p} \mathbb{E}(X). \tag{7.37}$$

The statements (7.37) and (7.36) are equivalent. The phenomenon of convergence in probability is also called *measure concentration*.

Example 7.4 (Relative frequencies)**.** Let $X_1, X_2, \ldots$ be coin tosses that are iid according to $P_X(h) = P_X(t) = 1/2$. Define $Z_i = \mathbb{1}(X_i = h)$. Since the X_i are iid and Z_i is a deterministic function of X_i, the random variables $Z_1, Z_2, \ldots$ are also iid and by the WLLN, we have

$$\frac{\sum_{i=1}^n Z_i}{n} \xrightarrow{p} \mathbb{E}(Z) = P_X(h)\cdot 1 + P_X(t)\cdot 0 = P_X(h). \tag{7.38}$$

Example 7.5 (Entropy)**.** Let $X_1, X_2, \ldots$ be iid according to some distribution P_X. Define $Z_i = -\log_2 P_X(X_i)$. By the WLLN, we have

$$\frac{\sum_{i=1}^n Z_i}{n} \xrightarrow{p} \mathbb{E}(Z) = \mathbb{E}[-\log_2 P_X(X)] = \mathbb{H}(X) \tag{7.39}$$

that is, $\sum_{i=1}^n Z_i/n$ converges to the entropy $\mathbb{H}(X)$ in probability.

Example 7.6 (No Measure Concentration)**.** Consider a coin with $P_{Y_1}(h) = P_{Y_1}(t) = 1/2$ and $Y_i = Y_1$, $i = 1, 2, 3, \ldots$. The two possible realizations are the sequences $hhh\cdots$ and $ttt\cdots$. At every time instance i, the marginal distribution is

$$P_{Y_i}(h) = P_{Y_1}(h) = \frac{1}{2}. \tag{7.40}$$

As function of interest, we pick $\mathbb{1}(Y = h)$ and we define $Z_i = \mathbb{1}(Y_i = h)$. We have

$\mathbb{E}(Z) = 1/2$ and for any $\epsilon < 1/2$, we have

$$\Pr\left\{\left|\frac{\sum_{i=1}^n Z_i}{n} - \mathbb{E}(Z)\right| > \epsilon\right\} = 1 \tag{7.41}$$

independent of n. This means that no measure concentration in $\mathbb{E}(Z)$ is happening.

7.1.3. Memoryless Channels

Theorem 6. *For a memoryless channel with channel law*

$$p_{Y^n|X^n}(b^n|a^n) = \prod_{i=1}^n p_{Y|X}(b_i|a_i), \quad b^n \in \mathcal{Y}^n, a^n \in \mathcal{X}^n \tag{7.42}$$

the decoder can recover x^n from the random channel output if x^n is approximately of type P_X and if

$$R_\mathrm{c} < T_\mathrm{c} = \mathbb{E}\left[\log_2 \frac{q(X,Y)}{\sum_{a\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(a,Y)}\right] \tag{7.43}$$

$$= \log_2 |\mathcal{X}| - \underbrace{\mathbb{E}\left[-\log_2 \frac{q(X,Y)}{\sum_{a\in\mathcal{X}} q(a,Y)}\right]}_{\text{uncertainty}} \tag{7.44}$$

where the expectation is taken according to $XY \sim P_X p_{Y|X}$.

Proof. We continue to assume input sequence x^n was transmitted, but we replace the specific channel output measurement y^n by the random output Y^n, distributed according to $p^n_{Y|X}(\cdot|x^n)$. The achievable code rate (7.33) evaluated in Y^n is

$$\hat{T}_\mathrm{c}(x^n, Y^n, q) = \frac{1}{n}\sum_{i=1}^n \log_2 \frac{q(x_i, Y_i)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c, Y_i)}. \tag{7.45}$$

Since Y^n is random, $\hat{T}_\mathrm{c}(x^n, Y^n, q)$ is also random. First, we rewrite (7.45) by sorting the summands by the input symbols, i.e.,

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^n \log_2 \frac{q(x_i, Y_i)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c, Y_i)} \\ &\qquad = \sum_{a\in\mathcal{X}} \frac{N(a|x^n)}{n}\left[\frac{1}{N(a|x^n)} \sum_{i\colon x_i=a} \log_2 \frac{q(a, Y_i)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c, Y_i)}\right]\end{aligned} \tag{7.46}$$

where $N(a|x^n)$ is the number of occurrences of a in x^n, see Appendix C.1. Note that identity (7.46) holds also when the channel has memory. For memoryless channels, we make the following two observations:

- Consider the inner sums in (7.46). For memoryless channels, the outputs $\{Y_i \colon x_i = a\}$ are iid according to $p_{Y|X}(\cdot|a)$. Therefore, by the WLLN (7.36), we have

$$\frac{1}{N(a|x^n)} \sum_{i\colon x_i=a} \log_2 \frac{q(a,Y_i)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c,Y_i)} \xrightarrow{p} \mathbb{E}\left[\log_2 \frac{q(a,Y)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c,Y)} \middle| X=a\right] \tag{7.47}$$

  where $\xrightarrow{p}$ denotes convergence in probability (7.37). That is, by making n and thereby $N(a|x^n)$ large, each inner sum converges in probability to a deterministic value. Note that the expected value on the right-hand side of (7.47) is no longer a function of the output sequence Y^n and is determined by the channel law $p_{Y|X}(\cdot|a)$ according to which the expectation is calculated.

- Suppose now for some distribution P_X and $\epsilon \geq 0$, the sequence x^n is in the typical set $\mathcal{T}_\epsilon^n(P_X)$, so that by Appendix C.1, we have

$$(1-\epsilon)P_X(a) \leq \frac{N(a|x^n)}{n} \leq (1+\epsilon)P_X(a), \quad a \in \mathcal{X}. \tag{7.48}$$

  We now have

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{q(x_i,Y_i)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c,Y_i)} \\
&\xrightarrow{p} \sum_{a\in\mathcal{X}} \frac{N(a|x^n)}{n} \mathbb{E}\left[\log_2 \frac{q(a,Y)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c,Y)} \middle| X=a\right] \qquad (7.49)\\
&\geq \mathbb{E}\left[\log_2 \frac{q(X,Y)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c,Y)}\right] - \epsilon \sum_{a\in\mathcal{X}} P_X(a) \left|\mathbb{E}\left[\log_2 \frac{q(a,Y)}{\sum_{c\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(c,Y)} \middle| X=a\right]\right| \\
&\qquad (7.50)
\end{aligned}$$

  where the expectation in (7.50) is calculated according to P_X and the channel law $p_{Y|X}$. In other words, (7.50) is an achievable code rate for all code words x^n that are in $\mathcal{T}_\epsilon^n(P_X)$.

□

Remark 4. In (7.47), we use $N(a|x^n)$ samples to estimate an expectation conditioned on the input $X = a$. We do so for each $a \in \mathcal{X}$. We then sum up the conditional expectations, weighting the summands by the corresponding fraction $N(a|x^n)/n$. In principle, we could also use a different strategy, e.g., we could use the same number $n/|\mathcal{X}|$ of samples for estimating each conditional expectation, and then calculate the sum with the weighting factors $N(a|x^n)/n$. In Problem 7.2, we show that under certain conditions, using the number of samples $N(a|x^n)$ that corresponds to the weighting factor is reasonable, in the sense that it minimizes the variance of the weighted sum.

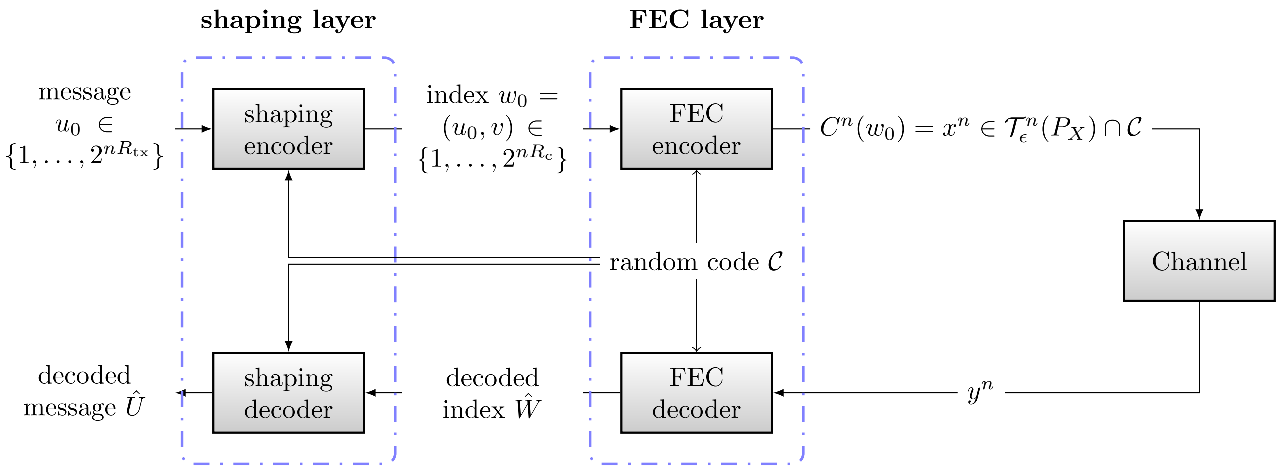

shaping layer
FEC layer
message $u_0 \in \{1,\dots,2^{nR_{\mathrm{tx}}}\}$
shaping encoder
index $w_0 = (u_0,v) \in \{1,\dots,2^{nR_{\mathrm{c}}}\}$
FEC encoder
$C^n(w_0) = x^n \in \mathcal{T}_\epsilon^n(P_X) \cap \mathcal{C}$
random code $\mathcal{C}$
Channel
decoded message $\hat{U}$
shaping decoder
decoded index $\hat{W}$
FEC decoder
y^n

Figure 7.2.: The random coding experiment with the shaping layer.

7.2. Shaping Layer

We now add a shaping layer whose task is to encode message bits into code words with a desired distribution. See Figure 7.2 for an illustration.

- Encoding: We set $R_{\mathrm{tx}} + R' = R_{\mathrm{c}}$ and double index the code words by $C^n(u,v)$, $u = 1,2,\dots,2^{nR_{\mathrm{tx}}}$, $v = 1,2,\dots,2^{nR'}$. We encode message $u \in \{1,\dots,2^{nR_{\mathrm{tx}}}\}$ by looking for a v, so that $C^n(u,v) \in \mathcal{T}_\epsilon^n(P_X)$, which is a set of typical sequences, see Appendix C.1. If we can find such v, we transmit the corresponding code word. If not, we choose some arbitrary v and transmit the corresponding code word.

- The rate is R_{tx}, since the encoder can encode $2^{nR_{\mathrm{tx}}}$ different messages.

The FEC layer works as before, in particular:

- Decoding: Recall that the receiver decodes with the rule

$$\hat{W} = \operatorname*{argmax}_{w \in \{1,\dots,2^{nR_{\mathrm{c}}}\}} \prod_{i=1}^{n} q(C_i(w), y_i). \tag{7.51}$$

  Note that the decoder evaluates the metric on all code words in $\mathcal{C}$, which includes code words that will never be transmitted because they are not in the shaping set $\mathcal{T}_\epsilon^n(P_X)$.

- Decoding error: Note that $\hat{W} = w_0$ implies $\hat{U} = u_0$, where u_0 is the encoded message and $\hat{U}$ is the decoded message. In particular, we have

$$\Pr(\hat{U} \neq u_0 | C^n(w_0) = x^n, Y^n = y^n) \leq \Pr(\hat{W} \neq w_0 | C^n(w_0) = x^n, Y^n = y^n). \tag{7.52}$$

Remark 5. The classical transceiver setup analyzed in, e.g., [3, Chapter 5 & 7],[20],[21], is as follows:

- *Random coding:* For the code $\tilde{\mathcal{C}} = \{\tilde{C}^n(1), \dots, \tilde{C}^n(2^{n\tilde{R}_\mathrm{c}})\}$, the $n \cdot 2^{n\tilde{R}_\mathrm{c}}$ code word entries are generated independently according to the distribution P_X.

- *Encoding:* Message u is mapped to code word $\tilde{C}^n(u)$.

- The decoder uses the decoding rule

$$\hat{u} = \operatorname*{argmax}_{u \in \{1,2,\dots,2^{n\tilde{R}_\mathrm{c}}\}} \prod_{i=1}^{n} q(\tilde{C}_i(u), y_i). \tag{7.53}$$

Note that here, the code word index is equal to the message, i.e., $w = u$, and consequently, the transmission rate is equal to the code rate, i.e., $R_\mathrm{tx} = \tilde{R}_\mathrm{c}$, while for the layered transceiver, we have $R_\mathrm{tx} < R_\mathrm{c}$ for non-uniform P_X.

Remark 6. In case the input distribution P_X is uniform, the layered transceiver is equivalent to the classical transceiver.

7.2.1. Achievable Encoding Rate

Lemma 1. *Encoding in the shaping layer is successful with high probability for large n if*

$$R_\mathrm{tx} < [R_\mathrm{c} - \mathbb{D}(P_X \| P_U)]^+. \tag{7.54}$$

Proof. See Section 7.3. □

If the right-hand side of (7.54) is positive, then out of the 2^{nR_c} code words, approximately $2^{n[R_\mathrm{c} - \mathbb{D}(P_X \| P_U)]}$ have approximately the distribution P_X and may be selected by the encoder for transmission. If the code rate is less than the informational divergence, then with high probability, the code does not contain any code word with approximately the distribution P_X. In this case, encoding is impossible, which corresponds to the encoding rate zero. The plus operator $[\cdot]^+ = \max\{0, \cdot\}$ ensures that this is reflected by the expression on the right-hand side of (7.54).

7.2.2. Achievable Rate

By replacing the code rate R_c in the encoding rate $[R_\mathrm{c} - \mathbb{D}(P_X \| P_U)]^+$ by the achievable code rate T_c, we arrive at an achievable rate.

Theorem 7. *An achievable rate is*

$$R = [T_\mathrm{c} - \mathbb{D}(P_X \| P_U)]^+ = \left[\mathbb{E}\left[\log_2 \frac{q(X,Y)}{\sum_{a\in\mathcal{X}} \frac{1}{|\mathcal{X}|} q(a,Y)}\right] - \mathbb{D}(P_X\|P_U)\right] \tag{7.55}$$

$$= \left[\mathbb{E}\left[\log_2 \frac{q(X,Y)\frac{1}{P_X(X)}}{\sum_{a\in\mathcal{X}} q(a,Y)}\right]\right]^+ \tag{7.56}$$

$$= \left[\mathbb{H}(X) - \underbrace{\mathbb{E}\left[-\log_2 \frac{q(X,Y)}{\sum_{a\in\mathcal{X}} q(a,Y)}\right]}_{\text{uncertainty}}\right]^+. \tag{7.57}$$

The three right-hand sides provide three different perspectives on the achievable rate.

- *Divergence perspective:* The term in (7.55) emphasizes that the random code was generated according to a uniform distribution and that of the 2^{nT_c} code words, only approximately $2^{nT_\mathrm{c}}/2^{n\mathbb{D}(P_X\|P_U)}$ code words are actually used for transmission, because the other code words very likely do not have distributions that are approximately P_X.

- *Output perspective:* In (7.56), $q(a,\cdot)/P_X(a)$ has the role of a channel likelihood given input $X = a$ assumed by the receiver, and correspondingly, $\sum_{a\in\mathcal{X}} q(a,\cdot)$ plays the role of a channel output statistics assumed by the receiver.

- *Uncertainty perspective:* In (7.57), $q(\cdot,b)/\sum_{a\in\mathcal{X}} q(a,b)$ defines for each realization b of Y a distribution on $\mathcal{X}$ and plays the role of a posterior probability distribution that the receiver assumes about the input, given its output observation. The expectation corresponds to the uncertainty that the receiver has about the input, given the output.

Example 7.7 (Example 7.1 continued)**.** For the metric

$$q(a,b) = P_{x^n}(a)P_{y^n}(b), \quad a\in\mathcal{X}, b\in\mathcal{Y}. \tag{7.58}$$

we calculated the achievable code rate

$$\hat{T}_\mathrm{c}(x^n, y^n, P_{x^n}P_{y^n}) = \mathbb{D}(P_{x^n}\|P_U). \tag{7.59}$$

By Theorem 7, this translates into the achievable rate

$$R_\mathrm{tx} = \left[\hat{T}_\mathrm{c}(x^n, y^n, P_{x^n}P_{y^n}) - \mathbb{D}(P_{x^n}\|P_U)\right]^+ = 0 \tag{7.60}$$

which makes sense, since the metric (7.58) treats channel input and output as independent. For the metric

$$q(a,b) = P_{x^n}(a)P_{y^n|x^n}(b|a), \quad a \in \mathcal{X}, b \in \mathcal{Y} \tag{7.61}$$

we calculated the achievable code rate

$$\hat{T}_{\mathrm{c}}(x^n, y^n, P_{x^n}P_{y^n|x^n}) = \mathbb{H}(P_U) - \mathbb{H}(P_{x^n|y^n}|P_{y^n}) \tag{7.62}$$

which by Theorem 7 translates into the achievable rate

$$\begin{aligned} R_{\mathrm{tx}} &= \Big[\hat{T}_{\mathrm{c}}(x^n, y^n, P_{x^n}P_{y^n|x^n}) - \mathbb{D}(P_{x^n}\|P_U)\Big]^+ && (7.63)\\ &= \mathbb{H}(P_{x^n}) - \mathbb{H}(P_{x^n|y^n}|P_{y^n}) && (7.64)\\ &= \mathbb{I}(P_{x^n}, P_{x^n|y^n}) && (7.65) \end{aligned}$$

where we dropped the $[\cdot]^+$ operator because mutual information is non-negative (C.37).

7.3. Proof of Lemma 1

To analyze the probability of successful encoding, we need a basic result on typical sequences (see also Appendix C.1). Define

$$\mu_X := \min_{a \in \operatorname{supp} P_X} P_X(a). \tag{7.66}$$

We need the following property of typical sequences.

Lemma 2 (Typicality, [22, Theorem 1.1],[23, Lemma 19])**.** *Suppose* $0 < \epsilon < \mu_X$. *We have*

$$(1-\delta_\epsilon(n, P_X))2^{n(1-\epsilon)\,\mathbb{H}(X)} \le |\mathcal{T}_\epsilon^n(P_X)| \tag{7.67}$$

where $\delta_\epsilon(P_X, n)$ *is such that* $\delta_\epsilon(P_X, n) \xrightarrow{n\to\infty} 0$ *exponentially fast in* n.

Lemma 3 (Mismatched Typicality)**.** *Suppose* $\epsilon > 0$, X^n *is emitted by the discrete memoryless source* P_X *and* $\operatorname{supp} P_{\tilde{X}} \subseteq \operatorname{supp} P_X$. *We have*

$$(1-\delta_\epsilon(P_{\tilde{X}}, n))2^{-n[\mathbb{D}(P_{\tilde{X}}\|P_X)-\epsilon\log_2(\mu_{\tilde{X}}\mu_X)]} \le \Pr[X^n \in \mathcal{T}_\epsilon^n(P_{\tilde{X}})]. \tag{7.68}$$

Proof. For $x^n \in \mathcal{T}_\epsilon^n(P_{\tilde{X}})$, we have

$$\begin{aligned} P_X^n(x^n) &= \prod_{a \in \operatorname{supp} P_{\tilde{X}}} P_X(a)^{N(a|x^n)}\\ &\ge \prod_{a \in \operatorname{supp} P_{\tilde{X}}} P_X(a)^{n(1+\epsilon)P_{\tilde{X}}(a)}\\ &= 2^{\sum_{a \in \operatorname{supp} P_{\tilde{X}}} n(1+\epsilon)P_{\tilde{X}}(a)\log_2 P_X(a)}. && (7.69) \end{aligned}$$

Now, we have

$$
\begin{aligned}
\Pr[X^n \in \mathcal{T}_\epsilon^n(P_{\tilde{X}})] &= \sum_{x^n \in \mathcal{T}_\epsilon^n(P_{\tilde{X}})} P_X^n(x^n) && (7.70)\\
&\overset{(7.69)}{\geq} \sum_{x^n \in \mathcal{T}_\epsilon^n(P_{\tilde{X}})} 2^{\sum_{a\in\operatorname{supp} P_{\tilde{X}}} n(1+\epsilon)P_{\tilde{X}}(a)\log_2 P_X(a)} && (7.71)\\
&\overset{(7.67)}{\geq} (1-\delta_\epsilon(n,P_{\tilde{X}}))2^{n(1-\epsilon)\,\mathbb{H}(\tilde{X})}2^{\sum_{a\in\operatorname{supp} P_{\tilde{X}}} n(1+\epsilon)P_{\tilde{X}}(a)\log_2 P_X(a)} && (7.72)\\
&= (1-\delta_\epsilon(n,P_{\tilde{X}}))2^{-n[\mathbb{D}(P_{\tilde{X}}\|P_X)-\epsilon\,\mathbb{H}(\tilde{X})+\epsilon\sum_{a\in\operatorname{supp} P_{\tilde{X}}} P_{\tilde{X}}(a)\log_2 P_X(a)]} && (7.73)\\
&\geq (1-\delta_\epsilon(n,P_{\tilde{X}}))2^{-n[\mathbb{D}(P_{\tilde{X}}\|P_X)+\epsilon\log_2(\mu_{\tilde{X}}\mu_X)]}. && (7.74)
\end{aligned}
$$

□

We can now analyze our encoding strategy. By (7.68), we have

$$\Pr[C^n(u,v) \in \mathcal{T}_\epsilon^n(P_X)] \geq [1-\delta_\epsilon(P_X,n)]2^{-n[\mathbb{D}(P_X\|P_U)-\epsilon\log_2(\mu_X\mu_U)]} \tag{7.75}$$

where by (7.66),

$$
\begin{aligned}
\mu_X &= \min_{a:\, P_X(a)>0} P_X(a) && (7.76)\\
\mu_U &= \min_{a:\, P_U(a)>0} P_U(a). && (7.77)
\end{aligned}
$$

Note that since P_U is uniform on $\mathcal{X}$, we have $\mu_U = 1/|\mathcal{X}|$. For large enough n, we have $\delta_\epsilon(P_X,n) \leq 1/2$. The probability to generate $2^{nR'}$ sequences $C^n(u,v)$, $v = 1,2,\dots,2^{nR'}$, that are *not* in $\mathcal{T}_\epsilon^n(P_X)$ is thus bounded from above by

$$\left(1-\frac{1}{2}2^{-n[\mathbb{D}(P_X\|P_U)-\epsilon\log_2(\mu_X\mu_U)]}\right)^{2^{nR'}} \leq \exp\left[-\frac{1}{2}2^{-n[\mathbb{D}(P_X\|P_U)-\epsilon\log_2(\mu_X\mu_U)]}2^{nR'}\right] \tag{7.78}$$

where (7.78) follows by $(1-r)^s \leq \exp(-rs)$. This probability tends to zero doubly exponentially fast for any $\epsilon > 0$ if

$$R' > \mathbb{D}(P_X\|P_U) + \epsilon\log_2\frac{1}{\mu_X\mu_U}. \tag{7.79}$$

Thus, as long as

$$R_{\text{tx}} < R_{\text{c}} - \mathbb{D}(P_X\|P_U) \tag{7.80}$$

our encoding strategy works with high probability.

7.4. Problems

Problem 7.1. Show (7.15) in Example 7.2.

Problem 7.2. Suppose for a memoryless channel $p_{Y|X}$ with input alphabet $\mathcal{X} = \{0,1\}$, you want to estimate

$$F = \mathbb{E}[f(X,Y)] \tag{7.81}$$

for the input distribution $P_X(0) = 1 - P_X(1)$ using n samples. Assume for the conditional variances $\text{Var}[f(X,Y)|X=0] = \text{Var}[(f(X,Y)|X=1] = \sigma^2$. You use an input sequence x^n with $n_0 = N(0|x^n)$ zeros and $n - n_0 = N(1|x^n)$ ones and you calculate the estimate

$$\hat{F} = P_X(0) \cdot \frac{1}{n_0} \sum_{i\colon x_i=0} f(0,y_i) + P_X(1) \cdot \frac{1}{n-n_0} \sum_{i\colon x_i=1} f(1,y_i). \tag{7.82}$$

1. Calculate the variance $\text{Var}(\hat{F})$.
2. Assuming n_0 is a real variable, for which value of n_0 is your variance expression minimized?

8. Decoding Metrics

In this chapter, we elaborate on the achievable rate for memoryless channels stated in Theorem 7. In Section 8.1, we consider the optimal choice of decoding metrics and in Section 8.2, we discuss how to assess the performance of given metrics.

8.1. Metric Design

By the information inequality (C.21), we know that

$$\mathbb{E}[-\log_2 P_Z(X)] \geq \mathbb{E}[-\log_2 P_X(X)] = \mathbb{H}(X) \tag{8.1}$$

with equality if and only if $P_Z = P_X$. We now use this observation to choose optimal metrics.

8.1.1. Mutual Information

Suppose we have no restriction on the decoding metric q. To maximize the achievable rate, we need to minimize the uncertainty term in (7.57). We have

$$\begin{aligned}\mathbb{E}\left[-\log_2 \frac{q(X,Y)}{\sum_{a\in\mathcal{X}} q(a,Y)}\right] &= \mathbb{E}\left[\mathbb{E}\left[-\log_2 \frac{q(X,Y)}{\sum_{a\in\mathcal{X}} q(a,Y)}\middle| Y\right]\right] &(8.2)\\ &\overset{(8.1)}{\geq} \mathbb{E}\left[\mathbb{E}\left[-\log_2 P_{X|Y}(X|Y)\middle| Y\right]\right] &(8.3)\\ &= \mathbb{H}(X|Y) &(8.4)\end{aligned}$$

with equality if we use the posterior probability distribution as metric, i.e.,

$$q(a,b) = P_{X|Y}(a|b), \quad a \in \mathcal{X}, b \in \mathcal{Y}. \tag{8.5}$$

Note that this choice of q is not unique, in particular, $q(a,b) = P_{X|Y}(a|b)P_Y(b)$ is also optimal, since the factor $P_Y(b)$ cancels out. For the optimal metric, the achievable rate is

$$R_{\mathsf{ps}}^{\mathsf{opt}} = [\mathbb{H}(X) - \mathbb{H}(X|Y)]^+ = \mathbb{I}(X;Y) \tag{8.6}$$

where we dropped the $(\cdot)^+$ operator because by the information inequality, mutual information is non-negative.

Discussion

In [3, Chapter 5 & 7], the achievability of mutual information is shown using the classical transceiver of Remark 5 with the likelihood decoding metric $q(a,b) = p_{Y|X}(b|a)$, $a \in \mathcal{X}, b \in \mathcal{Y}$. Comparing the classical transceiver with layered probabilistic shaping (PS) for a common rate R_{tx}, we have

$$\text{classical transceiver: } \hat{w} = \operatorname*{argmax}_{w \in \{1,2,\dots,2^{nR_{\text{tx}}}\}} \prod_{i=1}^{n} p_{Y|X}(y_i|\tilde{c}_i(w)) \tag{8.7}$$

$$\begin{aligned} \text{layered PS: } \hat{w} &= \operatorname*{argmax}_{w \in \{1,2,\dots,2^{n[R_{\text{tx}}+\mathbb{D}(P_X\|P_U)]}\}} \prod_{i=1}^{n} P_{X|Y}(c_i(w)|y_i) \\ &= \operatorname*{argmax}_{w \in \{1,2,\dots,2^{n[R_{\text{tx}}+\mathbb{D}(P_X\|P_U)]}\}} \prod_{i=1}^{n} p_{Y|X}(y_i|c_i(w))P_X(c_i(w)) \end{aligned} \tag{8.8}$$

Comparing (8.7) and (8.8) suggests the following interpretation:

- The classical transceiver uses the prior information by evaluating the *likelihood density* $p_{Y|X}$ on the code $\tilde{\mathcal{C}}$ that contains code words with distribution P_X. The code $\tilde{\mathcal{C}}$ has size $|\tilde{\mathcal{C}}| = 2^{nR_{\text{tx}}}$.

- Layered PS uses the prior information by evaluating the *posterior distribution* on all code words in the 'large' code $\mathcal{C}$ that contains mainly code words that do not have distribution P_X. The code $\mathcal{C}$ has size $|\mathcal{C}| = 2^{n[R_{\text{tx}}+\mathbb{D}(P_X\|P_U)]}$.

Remark 7. The code $\tilde{\mathcal{C}}$ of the classical transceiver is in general non-linear, since the set of vectors with distribution P_X is non-linear. It can be shown that all the presented results for layered PS also apply when $\mathcal{C}$ is a random linear code. In this case, layered PS evaluates a metric on a linear set while the classical transceiver evaluates a metric on a non-linear set.

8.1.2. Bit-Metric Decoding

Suppose the channel input is a binary vector $\boldsymbol{B} = B_1 \cdots B_m$ and the receiver uses a bit-metric, i.e.,

$$q(\boldsymbol{a}, y) = \prod_{j=1}^{m} q_j(a_j, y). \tag{8.9}$$

In this case, we have for the uncertainty term in (7.57)

$$\mathbb{E}\left[-\log_2 \frac{q(\boldsymbol{B},Y)}{\sum_{\boldsymbol{a}\in\{0,1\}^m} q(\boldsymbol{a},Y)}\right] = \mathbb{E}\left[-\log_2 \frac{\prod_{j=1}^m q_j(B_j,Y)}{\sum_{\boldsymbol{a}\in\{0,1\}^m} \prod_{j=1}^m q_j(a_j,Y)}\right] \tag{8.10}$$

$$= \mathbb{E}\left[-\log_2 \frac{\prod_{j=1}^m q_j(B_j,Y)}{\prod_{j=1}^m \sum_{a\in\{0,1\}} q_j(a,Y)}\right] \tag{8.11}$$

$$= \mathbb{E}\left[-\sum_{j=1}^m \log_2 \frac{q_j(B_j,Y)}{\sum_{a\in\{0,1\}} q_j(a,Y)}\right] \tag{8.12}$$

$$= \sum_{j=1}^m \mathbb{E}\left[-\log_2 \frac{q_j(B_j,Y)}{\sum_{a\in\{0,1\}} q_j(a,Y)}\right] \tag{8.13}$$

where equality in (8.11) follows by (A.9). For each $j = 1,\dots,m$, we now have

$$\mathbb{E}\left[-\log_2 \frac{q_j(B_j,Y)}{\sum_{a\in\{0,1\}} q_j(a,Y)}\right] = \mathbb{E}\left[\mathbb{E}\left[-\log_2 \frac{q_j(B_j,Y)}{\sum_{a\in\{0,1\}} q_j(a,Y)}\middle| Y\right]\right] \tag{8.14}$$

$$\geq \mathbb{H}(B_j|Y) \tag{8.15}$$

with equality if

$$q_j(a,b) = P_{B_j|Y}(a|b), \quad a\in\{0,1\}, b\in\mathcal{Y}. \tag{8.16}$$

The achievable rate becomes the BMD rate

$$R_{\mathsf{ps}}^{\mathsf{bmd}} = \left[\mathbb{H}(\boldsymbol{B}) - \sum_{j=1}^m \mathbb{H}(B_j|Y)\right]^+ \tag{8.17}$$

which we first stated in [24] and discuss in detail in [25, Section VI.]. In [26], we prove the achievability of (8.17) for discrete memoryless channels. For independent bit-level $B_1, B_2, \dots, B_m$, the BMD rate can be also written in the form [27]

$$R_{\mathsf{ps}}^{\mathsf{bmd,ind}} = \sum_{j=1}^m \mathbb{I}(B_j;Y). \tag{8.18}$$

8.1.3. Interleaved Coded Modulation

Suppose we have a vector channel with input $\boldsymbol{X} = X_1\cdots X_m$ with distribution $P_{\boldsymbol{X}}$ on the input alphabet $\mathcal{X}^m$ and output $\boldsymbol{Y} = Y_1\cdots Y_m$ with distributions $P_{\boldsymbol{Y}|\boldsymbol{X}}(\cdot|\boldsymbol{a})$, $\boldsymbol{a}\in\mathcal{X}^m$, on the output alphabet $\mathcal{Y}^m$. We consider the following situation:

- The Y_i are potentially correlated, in particular, we may have $Y_1 = Y_2 = \cdots = Y_m$.

- Despite the potential correlation, the receiver uses a memoryless metric q defined on $\mathcal{X} \times \mathcal{Y}$, i.e., a vector input $\boldsymbol{x}$ and a vector output $\boldsymbol{y}$ are scored by

$$q^m(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{m} q(x_i, y_i). \tag{8.19}$$

The reason for this decoding strategy may be an interleaver between encoder output and channel input that is reverted at the receiver but not known to the decoder. We therefore call this scenario *interleaved coded modulation*.

Using the same approach as for bit-metric decoding, we have

$$\begin{aligned}\frac{1}{m}\mathbb{E}\left[-\log_2 \frac{\prod_{i=1}^m q(X_i, Y_i)}{\sum_{\boldsymbol{a}\in\mathcal{X}^m}\prod_{i=1}^m q(a_i, Y_i)}\right] &= \frac{1}{m}\mathbb{E}\left[-\log_2 \frac{\prod_{i=1}^m q(X_i, Y_i)}{\prod_{i=1}^m \sum_{a\in\mathcal{X}} q(a, Y_i)}\right] &(8.20)\\ &= \frac{1}{m}\sum_{i=1}^{m}\mathbb{E}\left[-\log_2 \frac{q(X_i, Y_i)}{\sum_{a\in\mathcal{X}} q(a, Y_i)}\right] &(8.21)\end{aligned}$$

where equality in (8.20) follows by (A.9). Expression (8.21) is not very insightful. We could optimize q for, say, the ith term, which would be

$$q(a, b) = P_{X_i|Y_i}(a|b), \quad a \in \mathcal{X}, b \in \mathcal{Y} \tag{8.22}$$

but this would not be optimal for the other terms. We therefore choose a different approach. Let I be a random variable uniformly distributed on $\mathcal{I} = \{1, 2, \ldots, m\}$ and define $X = X_I$, $Y = Y_I$. Then, we have

$$\begin{aligned}\frac{1}{m}\sum_{i=1}^{m}\mathbb{E}\left[-\log_2 \frac{q(X_i, Y_i)}{\sum_{a\in\mathcal{X}} q(a, Y_i)}\right] &= \mathbb{E}\left[-\log_2 \frac{q(X_I, Y_I)}{\sum_{a\in\mathcal{X}} q(a, Y_I)}\right] &(8.23)\\ &= \mathbb{E}\left[-\log_2 \frac{q(X, Y)}{\sum_{a\in\mathcal{X}} q(a, Y)}\right]. &(8.24)\end{aligned}$$

Thus, the optimal metric for interleaving is

$$q(a, b) = P_{X|Y}(a|b) \tag{8.25}$$

which can be calculated from

$$P_X(a)p_{Y|X}(b|a) = \sum_{j=1}^{m}\frac{1}{m}P_{X_j}(a)p_{Y_j|X_j}(b|a). \tag{8.26}$$

The achievable rate becomes

$$R_{\mathsf{ps}}^{\mathsf{icm}} = \left[\mathbb{H}(\boldsymbol{X}) - m\,\mathbb{H}(X|Y)\right]^+. \tag{8.27}$$

When the input is a binary vector $\boldsymbol{B} = B_1 \cdots B_m$, we get the bit-interleaved coded modulation (BICM) rate

$$R_{\mathsf{ps}}^{\mathsf{bicm}} = \left[\mathbb{H}(\boldsymbol{B}) - m\,\mathbb{H}(B|Y)\right]^+. \tag{8.28}$$

8.2. Metric Assessment

Suppose a decoder is constrained to use a specific metric q. In this case, our task is to assess the metric performance by calculating a rate that can be achieved. If q is a non-negative metric, an achievable rate is

$$R_{\mathsf{ps}}(q) = \left[\mathbb{H}(X) - \mathbb{E}\left[-\log_2 \frac{q(X,Y)}{\sum_{a\in\mathcal{X}} q(a,Y)}\right]\right]^+. \tag{8.29}$$

However, higher rates may also be achievable by q. The reason for this is as follows: suppose we have another metric $\tilde{q}$ that scores the code words in the same order as metric q, i.e., we have

$$\tilde{q}(a_1,b) > \tilde{q}(a_2,b) \Leftrightarrow q(a_1,b) > q(a_2,b), \quad a_1,a_2 \in \mathcal{X}, b \in \mathcal{Y}. \tag{8.30}$$

Then, $R_{\mathsf{ps}}(\tilde{q})$ is also achievable by q. An example for a order preserving transformation is $\tilde{q}(a,b) = e^{q(a,b)}$. For a non-negative metric q, another order preserving transformation is $\tilde{q}(a,b) = q(a,b)^s$ for $s > 0$. We may now find a better achievable rate for metric q by calculating for instance

$$\max_{s>0} R_{\mathsf{ps}}(q^s). \tag{8.31}$$

In the following, we will say that two metrics q and $\tilde{q}$ are equivalent if and only if the order-preserving condition (8.30) is fulfilled.

Example 8.1 (AWGN Channel with BPSK)**.** Consider a BPSK constellation $\mathcal{X} = \{-1,1\}$ and uniformly distributed input, i.e., $P_X(-1) = P_X(1) = \frac{1}{2}$. The channel output is

$$Y = |h| \cdot X + Z$$

where $|h|$ is a positive real number and where Z is zero mean Gaussian with variance σ^2. At the receiver, the decoder uses the metric

$$q(b,a) = b \cdot a, \quad b \in \mathbf{R}, a \in \mathcal{X}.$$

Note that the decoder does not make use of the channel parameters $|h|, \sigma^2$. We transform the metric into the equivalent non-negative metric $e^{sq(b,a)} =: \tilde{q}(b,a)$ with $s > 0$ and calculate the uncertainty term for $\tilde{q}$. We have

$$\mathbb{E}\left[-\log_2 \frac{\tilde{q}(Y,X)}{\sum_{a\in\{-1,1\}} \tilde{q}(Y,a)}\right] = \mathbb{E}\left[\mathbb{E}\left[-\log_2 \frac{\tilde{q}(Y,X)}{\sum_{a\in\{-1,1\}} \tilde{q}(Y,a)}\middle| Y\right]\right]$$

For $Y = b$, the inner expectation is

$$\mathbb{E}\left[-\log_2 \frac{\tilde{q}(Y,X)}{\sum_{a\in\{-1,1\}} \tilde{q}(Y,a)}\middle| Y = b\right]$$

and it is calculated for $a' \in \{-1, 1\}$ according to

$$\begin{aligned} P_{X|Y}(a'|b) &= \frac{P_X(a')p_{Y|X}(b|a')}{p_Y(b)} \\ &= \frac{\frac{1}{2}p_Z(b-|h|a')}{\frac{1}{2}p_Z(b+|h|)+\frac{1}{2}p_Z(b-|h|)} \\ &= \frac{p_Z(b-|h|a')}{p_Z(b+|h|)+p_Z(b-|h|)} \\ &= \frac{e^{-\frac{(b-|h|a')^2}{2\sigma^2}}}{e^{-\frac{(b+|h|)^2}{2\sigma^2}}+e^{-\frac{(b-|h|)^2}{2\sigma^2}}} \\ &= \frac{e^{\frac{|h|a'b}{\sigma^2}}}{e^{-\frac{|h|b}{\sigma^2}}+e^{\frac{|h|b}{\sigma^2}}}. \end{aligned}$$

Note that for $s = |h|/\sigma^2$, we also have

$$\frac{\tilde{q}(b,a')}{\sum_{a\in\{-1,1\}} \tilde{q}(b,a)} = \frac{e^{\frac{|h|a'b}{\sigma^2}}}{e^{-\frac{|h|b}{\sigma^2}}+e^{\frac{|h|b}{\sigma^2}}}.$$

Thus, we have

$$\begin{aligned} \mathbb{E}\left[-\log_2 \frac{\tilde{q}(Y,X)}{\sum_{a\in\{-1,1\}} \tilde{q}(Y,a)}\middle| Y = b\right] &\overset{\text{(a)}}{\geq} \mathbb{E}\left[-\log_2 P_{X|Y}(X|b)\middle| Y = b\right] \\ &= \mathbb{H}(X|Y=b) \end{aligned}$$

with equality in (a) if $s = |h|/\sigma^2$. Thus, for this choice of s, the uncertainty term is $\mathbb{H}(X|Y)$ and the achievable rate is $\mathbb{I}(X;Y)$. Observations:

- Although the original metric $q(b,a) = a{\cdot}b$ does not take the channel parameters $|h|, \sigma^2$ into account, it is optimal because it achieves the mutual information $\mathbb{I}(X;Y)$.

- The optimal value $s^* = |h|/\sigma^2$ recovers the channel parameters.

8.2.1. Generalized Mutual Information

Suppose the input distribution is uniform, i.e., $P_X(a) = 1/|\mathcal{X}|, a \in \mathcal{X}$. In this case, we have

$$\max_{s>0} R_{\mathsf{ps}}(q^s) = \max_{s>0} \left[\mathbb{E}\left[\log_2 \frac{q(X,Y)^s \frac{1}{P_X(X)}}{\sum_{a\in\mathcal{X}} q(a,Y)^s}\right]\right]^+ \tag{8.32}$$

$$= \max_{s>0} \mathbb{E}\left[\log_2 \frac{q(X,Y)^s}{\sum_{a\in\mathcal{X}} P_X(a) q(a,Y)^s}\right] \tag{8.33}$$

where we could move $P_X(a)$ under the sum, because P_X is by assumption uniform, and where we could drop the $(\cdot)^+$ operator because for $s = 0$, the expectation is zero. The expression in (8.33) is called generalized mutual information (GMI) in [20] and was shown to be an achievable rate for the classical transceiver. This is in line with Remark 6, namely that for uniform input, layered PS is equivalent to the classical transceiver. For non-uniform input, the GMI and (8.32) may differ, i.e., we may not have equality in (8.33).

Discussion

Suppose for a non-uniform input distribution P_X and a metric q, the GMI evaluates to R, implying that a classical transceiver can achieve R. Can also layered PS achieve R, possibly by using a different metric? The answer is yes. Define

$$\tilde{q}(a,b) = q(a,b)P_X(a)^{\frac{1}{s}}, \quad a \in \mathcal{X}, b \in \mathcal{Y} \tag{8.34}$$

where s is the optimal value maximizing the GMI. We calculate a PS achievable rate for $\tilde{q}$ by analyzing the equivalent metric $\tilde{q}^s$. We have

$$R_{\mathsf{ps}} = \left[\mathbb{E}\left[\log_2 \frac{\tilde{q}^s(X,Y)\frac{1}{P_X(X)}}{\sum_{a\in\mathcal{X}} \tilde{q}^s(a,Y)}\right]\right]^+ \tag{8.35}$$

$$= \left[\mathbb{E}\left[\log_2 \frac{q^s(X,Y)}{\sum_{a\in\mathcal{X}} P_X(a) q^s(a,Y)}\right]\right]^+ \tag{8.36}$$

$$= R \tag{8.37}$$

which shows that R can also be achieved by layered PS. It is important to stress that this requires a change of the metric: for example, suppose q is the Hamming metric of a hard-decision decoder (see Section 8.2.3). In general, this does *not* imply that also $\tilde{q}$ defined by (8.34) is a Hamming metric.

8.2.2. LM-Rate

For the classical transceiver of Remark 5, the work [21] shows that the so-called LM-Rate defined as

$$R_{\mathrm{LM}}(s,r)=\left[\mathbb{E}\left[\log_2\frac{q(X,Y)^s r(X)}{\sum_{a\in\operatorname{supp}P_X}P_X(a)q(a,Y)^s r(a)}\right]\right]^+ \tag{8.38}$$

is achievable, where $s>0$ and where r is a function on $\mathcal{X}$. By choosing $s=1$ and $r(a)=1/P_X(a)$, we have

$$\begin{aligned}R_{\mathrm{LM}}(1,1/P_X)&=\left[\mathbb{E}\left[\log_2\frac{q(X,Y)\frac{1}{P_X(X)}}{\sum_{a\in\operatorname{supp}P_X}q(a,Y)}\right]\right]^+ &\quad(8.39)\\ &\geq\left[\mathbb{E}\left[\log_2\frac{q(X,Y)\frac{1}{P_X(X)}}{\sum_{a\in\mathcal{X}}q(a,Y)}\right]\right]^+ &\quad(8.40)\\ &=R_{\mathsf{ps}} &\quad(8.41)\end{aligned}$$

with equality in (8.40) if $\operatorname{supp}P_X=\mathcal{X}$. Thus, formally, our achievable rate can be recovered from the LM-Rate. We emphasize that [21] shows the achievability of the LM-Rate for the classical transceiver of Remark 5, and consequently, R_{LM} and R_{ps} have different operational meanings, corresponding to achievable rates of two different transceiver setups, with different random coding experiments, and different encoding and decoding strategies.

8.2.3. Hard-Decision Decoding

Hard-decision decoding consists of two steps. First, the channel output alphabet is partitioned into disjoint decision regions

$$\mathcal{Y}=\bigcup_{a\in\mathcal{X}}\mathcal{Y}_a,\quad \mathcal{Y}_a\cap\mathcal{Y}_b=\emptyset\text{ if }a\neq b \tag{8.42}$$

and a quantizer ω maps the channel output to the channel input alphabet according to the decision regions, i.e.,

$$\omega\colon\mathcal{Y}\to\mathcal{X},\quad \omega(b)=a\Leftrightarrow b\in\mathcal{Y}_a. \tag{8.43}$$

Second, the receiver uses the Hamming metric of $\mathcal{X}$ for decoding, i.e., we have

$$q(a,\omega(y))=\mathbb{1}(a,\omega(y))=\begin{cases}1, & \text{if } a=\omega(y)\\ 0, & \text{otherwise.}\end{cases} \tag{8.44}$$

We next derive an achievable rate by analyzing the equivalent metric $e^{s\mathbb{1}(\cdot,\cdot)}$, $s>0$. For the uncertainty term, we have

$$\mathbb{E}\left[-\log_2\frac{e^{s\mathbb{1}[X,\omega(Y)]}}{\sum_{a\in\mathcal{X}} e^{s\mathbb{1}[a,\omega(Y)]}}\right] = \mathbb{E}\left[-\log_2\frac{e^{s\mathbb{1}[X,\omega(Y)]}}{|\mathcal{X}|-1+e^s}\right] \tag{8.45}$$

$$= -\Pr[X=\omega(Y)]\log_2\frac{e^s}{|\mathcal{X}|-1+e^s} - \Pr[X\neq\omega(Y)]\log_2\frac{1}{|\mathcal{X}|-1+e^s} \tag{8.46}$$

$$= -(1-\epsilon)\log_2\frac{e^s}{|\mathcal{X}|-1+e^s} - \epsilon\log_2\frac{1}{|\mathcal{X}|-1+e^s} \tag{8.47}$$

$$= -(1-\epsilon)\log_2\frac{e^s}{|\mathcal{X}|-1+e^s} - \sum_{\ell=1}^{|\mathcal{X}|-1}\frac{\epsilon}{|\mathcal{X}|-1}\log_2\frac{1}{|\mathcal{X}|-1+e^s} \tag{8.48}$$

where we defined $\epsilon=\Pr(X\neq\omega(Y))$. By (8.1), the last line is maximized by choosing

$$s\colon 1-\epsilon=\frac{e^s}{|\mathcal{X}|-1+e^s} \text{ and } \frac{\epsilon}{|\mathcal{X}|-1}=\frac{1}{|\mathcal{X}|-1+e^s} \tag{8.49}$$

which is achieved by

$$e^s=\frac{(|X|-1)(1-\epsilon)}{\epsilon}. \tag{8.50}$$

With this choice for s, we have

$$-(1-\epsilon)\log_2(1-\epsilon)-\sum_{\ell=1}^{|\mathcal{X}|-1}\frac{\epsilon}{|\mathcal{X}|-1}\log_2\frac{\epsilon}{|\mathcal{X}|-1}$$

$$=\underbrace{-(1-\epsilon)\log_2(1-\epsilon)-\epsilon\log_2\epsilon}_{=:\mathbb{H}_2(\epsilon)}+\epsilon\log_2(|\mathcal{X}|-1) \tag{8.51}$$

$$=\mathbb{H}_2(\epsilon)+\epsilon\log_2(|\mathcal{X}|-1) \tag{8.52}$$

where $\mathbb{H}_2(\cdot)$ is the binary entropy function. The term (8.52) corresponds to the conditional entropy of a $|\mathcal{X}|$-ary symmetric channel with uniform input, see Figure 8.1 for an illustration. We conclude that by hard-decision decoding, we can achieve

$$R_{\mathsf{ps}}^{\mathsf{hd}}=[\mathbb{H}(X)-[\mathbb{H}_2(\epsilon)+\epsilon\log_2(|\mathcal{X}|-1)]]^+ \tag{8.53}$$

where

$$\epsilon=1-\Pr[X=\omega(Y)] \tag{8.54}$$

$$=1-\sum_{a\in\mathcal{X}}P_X(a)\int_{\mathcal{Y}_a}p_{Y|X}(\tau|a)\,\mathrm{d}\tau. \tag{8.55}$$

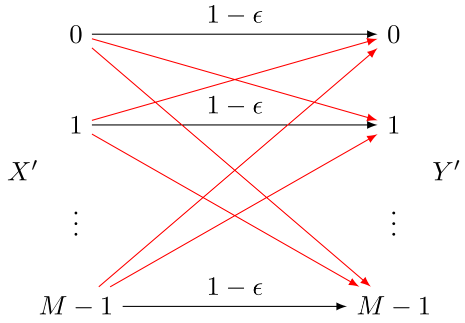


Figure 8.1.: The M-ary symmetric channel. Each red transition has probability $\frac{\epsilon}{M-1}$. Note that for $M = 2$, the channel is the binary symmetric channel. For uniformly distributed input X', we have $\mathbb{H}(X'|Y') = \mathbb{H}_2(\epsilon) + \epsilon \log_2(M-1)$.

8.2.4. Binary Hard-Decision Decoding

Suppose the channel input is the binary vector $\boldsymbol{B} = B_1 \cdots B_m$ and the decoder uses m binary quantizers, i.e., we have

$$\mathcal{Y} = \mathcal{Y}_{0j} \cup \mathcal{Y}_{1j}, \quad \mathcal{Y}_{1j} = \mathcal{Y} \setminus \mathcal{Y}_{0j} \tag{8.56}$$

$$\omega_j \colon \mathcal{Y} \to \{0,1\}, \quad \omega_j(b) = a \Leftrightarrow b \in \mathcal{Y}_{ja}. \tag{8.57}$$

The receiver uses a binary Hamming metric, i.e., we have

$$q(a,b) = \mathbb{1}(a,b), \quad a,b \in \{0,1\} \tag{8.58}$$

$$q^m(\boldsymbol{a},\boldsymbol{b}) = \sum_{j=1}^{m} \mathbb{1}(a_j,b_j) \tag{8.59}$$

and we analyze the equivalent metric

$$e^{sq^m(\boldsymbol{a},\boldsymbol{b})} = \prod_{j=1}^{m} e^{s\mathbb{1}(a_j,b_j)}, \quad s > 0. \tag{8.60}$$

Since the decoder uses the same metric for each bit-level $j = 1, 2, \dots, m$, binary hard-decision decoding is an instance of interleaved coded modulation, which we discussed in Section 8.1.3. Thus, defining the auxiliary random variable I uniformly distributed on $\{1, 2, \dots, m\}$ and

$$B = B_I, \quad \hat{B} = \omega_I(Y) \tag{8.61}$$

we can use the interleaved coded modulation result (8.24). We have for the normalized uncertainty term

$$\frac{1}{m}\mathbb{E}\left[-\log_2 \frac{\prod_{j=1}^m e^{s\mathbb{1}[B_j,\omega_j(Y)]}}{\sum_{\boldsymbol{a}\in\{0,1\}^m}\prod_{j=1}^m e^{s\mathbb{1}[a_j,\omega_j(Y)]}}\right]$$

$$\overset{(8.24),(8.61)}{=} \mathbb{E}\left[-\log_2 \frac{e^{s\mathbb{1}(B,\hat{B})}}{\sum_{a\in\{0,1\}} e^{s\mathbb{1}(a,\hat{B})}}\right] \tag{8.62}$$

$$= -\Pr(B=\hat{B})\log_2\frac{e^s}{e^s+1} - \underbrace{\Pr(B\neq\hat{B})}_{=:\epsilon}\log_2\frac{1}{e^s+1} \tag{8.63}$$

$$\overset{(8.1)}{\geq} \mathbb{H}(\epsilon) \tag{8.64}$$

with equality if

$$s\colon \frac{1}{e^s+1}=\epsilon. \tag{8.65}$$

Thus, with a hard decision decoder, we can achieve

$$R_{\mathsf{ps}}^{\mathsf{hd,bin}} = [\mathbb{H}(\boldsymbol{B}) - m\,\mathbb{H}_2(\epsilon)]^+ \tag{8.66}$$

where

$$\epsilon = \sum_{j=1}^m \frac{1}{m}\sum_{a\in\{0,1\}} P_{B_j}(a)\int_{\mathcal{Y}_{ja}} p_{Y|B_j}(\tau|a)\,\mathrm{d}\tau. \tag{8.67}$$

For uniform input, the rate becomes

$$R_{\mathsf{uni}}^{\mathsf{hd,bin}} = m - m\,\mathbb{H}_2(\epsilon) = m[1-\mathbb{H}_2(\epsilon)]. \tag{8.68}$$

8.3. Problems

Problem 8.1. For a binary input B and real-valued output Y, a decoder uses the metric

$$q(y,b) = (-1)^b L(y) = (1-2b)L(y), \quad y\in\mathbf{R}, b\in\{0,1\}$$

where

$$L(y) = \log\frac{P_B(0)p_{Y|B}(y|0)}{P_B(1)p_{Y|B}(y|1)}.$$

1. Show that

$$P_{B|Y}(0|y) = \frac{e^{\frac{L(y)}{2}}}{e^{\frac{L(y)}{2}}+e^{-\frac{L(y)}{2}}}, \quad P_{B|Y}(1|y) = \frac{e^{-\frac{L(y)}{2}}}{e^{\frac{L(y)}{2}}+e^{-\frac{L(y)}{2}}}.$$

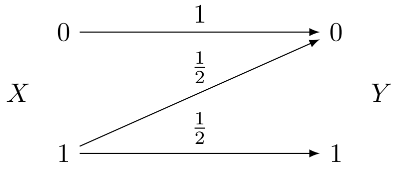


Figure 8.2.: The Z-channel.

2. To calculate an achievable rate, we evaluate the uncertainty term in (7.57) for the equivalent non-negative metric $\tilde{q}(y,b) = e^{sq(y,b)}$. For which s is the uncertainty term equal to the minimum value $\mathbb{H}(B|Y)$?

Problem 8.2. Consider the Z-channel in Figure 8.2. The input distribution is uniform, i.e., $P_X(0) = P_X(1) = \frac{1}{2}$.

1. Calculate the output distribution P_Y and the conditional input distributions $P_{X|Y}(\cdot|0)$ and $P_{X|Y}(\cdot|1)$.

2. Calculate the mutual information $\mathbb{I}(X;Y)$.

3. Specify a decoding metric $q_1(\cdot,\cdot)$ that achieves $R = \mathbb{I}(X;Y)$.

4. Transform q_1 into an equivalent metric $\tilde{q}_1$ that takes the values 0 and 1.
Hint: Two metrics are equivalent if they imply the same decoding rule.

5. The receiver now assumes the channel is a binary symmetric channel and uses metric $q_2(0,0) = q_2(1,1) = 1$ and $q_2(0,1) = q_2(1,0) = -1$. Show that this metric can achieve $\mathbb{H}(X) - \mathbb{H}_2[\Pr(X = Y)]$, where

$$\mathbb{H}_2(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)$$

is the binary entropy function.
Hint: Analyze the equivalent non-negative metric $\exp[sq_2(x,y)]$ with $s > 0$.

6. Calculate $\mathbb{H}(X) - \mathbb{H}_2[\Pr(X = Y)]$ and the achievable rate degradation that results from using q_2 instead of q_1.

7. Transform q_2 into an equivalent metric $\tilde{q}_2$ that takes the values 0 and 1.
Hint: Two metrics are equivalent if they imply the same decoding rule.

9. Distribution Matching

In Chapter 7 and 10, we repeatedly encountered situations where a functional of an input vector x^n depends only on its empirical distribution, and not on the particular ordering of its entries, i.e., any permuted version of x^n results in the same system behavior. This suggests to encode information into the permutation of x^n. Mapping source bits to sequences with a desired distribution is called distribution matching. In this chapter, we discuss the CCDM [28] and analyze its rate. The CCDM is a fundamental component of the transceiver architecture that we will define in Chapter 6 and analyze in Chapter 10.

9.1. Types

Types as defined in Appendix C.1 play a central role for the CCDM. We therefore start this chapter by reviewing types and their properties.

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.,

$$N(a|x^n) = \Bigl|\Bigl\{i \in \{1,2,\dots,n\}\colon x_i = a\Bigr\}\Bigr|, \quad a \in \mathcal{X}. \tag{9.1}$$

The empirical distribution of x^n is

$$P_{x^n}(a) = \frac{N(a|x^n)}{n}, \quad a \in \mathcal{X}. \tag{9.2}$$

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

$$P_X(a) = \frac{n_a}{n}, \quad a \in \mathcal{X}. \tag{9.3}$$

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 $\mathcal{T}^n(P_X)$.

Note that we have defined the distribution P_X as the empirical distribution of a length n sequence x^n. Often, we start with a distribution that we obtained, e.g., from maximizing a mutual information, and are then interested in length n sequences with the optimal distribution. This is only approximately possible, since in general, the optimal distribution is not n-type. In Section 9.3, we discuss the quantization of arbitrary distributions to n-types.

9.2. Constant Composition Distribution Matcher (CCDM)

A code $\mathcal{C} \subseteq \mathcal{T}^n(P_X)$ is called a constant composition code. A CCDM encodes length k bit strings to $\mathcal{T}^n(P_X)$, i.e., it implements a fixed-to-fixed length mapping into $\mathcal{T}^n(P_X)$.

Example 9.1. Consider the 4-ASK constellation $\mathcal{X} = \{\pm 1, \pm 3\}$ and the block length $n = 4$. The amplitudes are $\mathcal{A} = \{1, 3\}$. Suppose the desired amplitude distribution is

$$P_A(1) = \frac{3}{4}, \quad P_A(3) = \frac{1}{4}. \tag{9.4}$$

The probabilities are integer multiples of $1/4$, so P_A is a 4-type. The corresponding 4-type class is

$$\mathcal{T}^4(P_A) = \{(1,1,1,3),(1,1,3,1),(1,3,1,1),(3,1,1,1)\} \tag{9.5}$$

where in each sequence, the amplitudes 1 and 3 occur $n_1 = 3$ and $n_3 = 1$ times, respectively. A CCDM maps length k binary strings to sequences in $\mathcal{T}^4(P_A)$. There are 4 sequences in $\mathcal{T}^4(P_A)$, so we have

$$k = \log_2 |\mathcal{T}^4(P_A)| = 2. \tag{9.6}$$

The following look-up table (LUT) defines a CCDM.

$$\begin{aligned} 00 &\mapsto (1,1,1,3), \quad & 01 &\mapsto (1,1,3,1),\\ 10 &\mapsto (1,3,1,1), \quad & 11 &\mapsto (3,1,1,1). \end{aligned} \tag{9.7}$$

The mapping is one-to-one and therefore invertible on its image.

9.2.1. Rate

The rate of a CCDM is

$$\frac{k}{n} \quad \left[\frac{\text{input bits}}{\text{output symbols}}\right]. \tag{9.8}$$

We are interested in making the rate as large as possible, so we set the input length to the highest value for which a one-to-one mapping into $\mathcal{T}^n(P_X)$ exists. This value is given by

$$k = \lfloor \log_2 |\mathcal{T}^n(P_X)| \rfloor . \tag{9.9}$$

Suppose the output alphabet is $1, 2, \ldots, M$. Let n_i be the number of occurrences of i in each sequence in $\mathcal{T}^n(P_X)$, i.e.,

$$n_i = nP_X(i), \quad i = 1, 2, \ldots, M. \tag{9.10}$$

We can now write the size of $\mathcal{T}^n(P_X)$ as

$$|\mathcal{T}^n(P_X)| = \frac{n!}{n_1!n_2!\cdots n_M!} = \binom{n}{n_1, n_2, \dots, n_M} \tag{9.11}$$

where the term on the right-hand side is called a multinomial coefficient. For the CCDM rate, we have

$$R_{\text{ccdm}}(P_X, n) = \frac{k}{n} = \frac{\lfloor \log_2 \mathcal{T}^n(P_X) \rfloor}{n} = \frac{\left\lfloor \log_2 \binom{n}{n_1,\dots,n_M} \right\rfloor}{n}. \tag{9.12}$$

In Problem 9.1, we show how the logarithm of the multinomial coefficient and the CCDM rate can be calculated numerically even for very large n. By Problem 9.4, the CCDM rate is bounded above by the entropy of P_X, i.e.,

$$R_{\text{ccdm}}(P_X, n) \le \mathbb{H}(P_X). \tag{9.13}$$

We define the CCDM rate loss by

$$R_{\text{loss}}(P_X, n) = \mathbb{H}(P_X) - R_{\text{ccdm}}(P_X, n). \tag{9.14}$$

The next theorem characterizes how fast the rate loss approaches zero.

Theorem 8. *Let P_X be an n'-type and let n be a multiple of n.*

1. *The rate loss is bounded above and below by $\frac{\log n}{n}$, i.e.,*

$$0 < \lim_{n\to\infty} \frac{R_{\text{loss}}(P_X, n)}{\frac{\log n}{n}} < \infty \tag{9.15}$$

that is, $R_{\text{loss}}(P_X, n) \in \Theta(\frac{\log n}{n})$.

2. *We have*

$$R_{\text{ccdm}}(P_X, n) \xrightarrow{n\to\infty} \mathbb{H}(P_X). \tag{9.16}$$

Proof. We prove (9.15) in Section 9.4. The limit (9.16) is a consequence of (9.15). □

Remark 8. For binary output alphabets, (9.15) is proven in [29]. For arbitrary alphabets, (9.16) is proven in [28].

Example 9.2 (Example 9.1 continued)**.** The CCDM rate is

$$R_{\text{ccdm}}(P_A, 4) = \frac{k}{n} = \frac{\log_2 4}{4} = \frac{1}{2}. \tag{9.17}$$

The entropy of P_A is

$$\mathbb{H}(P_A) = 0.8113. \tag{9.18}$$

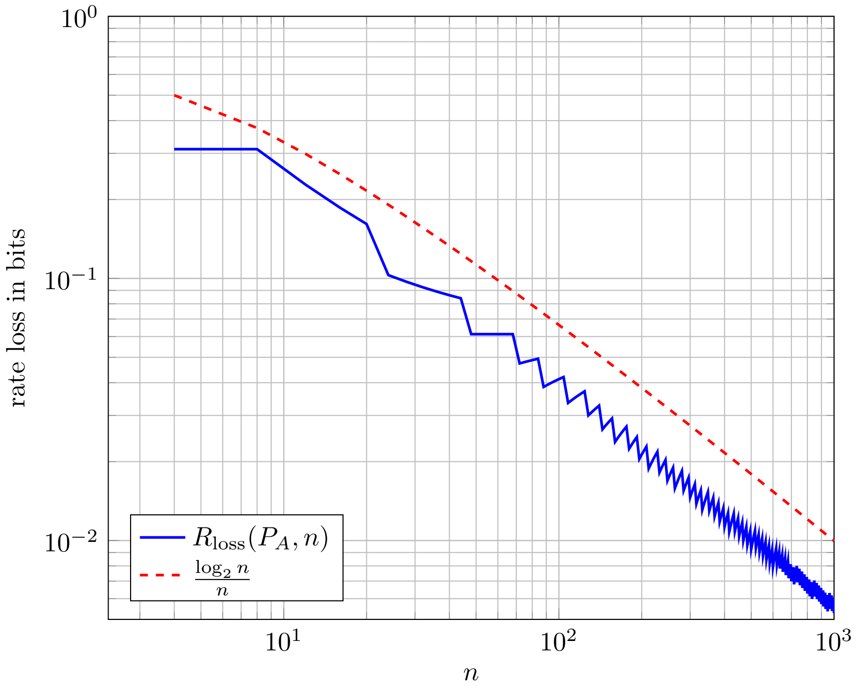


Figure 9.1.: CCDM rate loss in Example 9.2.

Consequently, the rate loss is

$$R_{\text{loss}}(P_A, 4) = 0.3113 \quad \left[\frac{\text{bit}}{\text{amplitude}}\right]. \tag{9.19}$$

For $n = 10\,000$, the CCDM rate is

$$R_{\text{ccdm}}(P_A, 10\,000) = \frac{8106}{10\,000} = 0.8106 \tag{9.20}$$

which corresponds to a rate loss of approximately 7×10^{-4} bits per amplitude. In Figure 9.1, we display the rate loss for $n = 4, 8, \ldots, 1 \times 10^3$.

9.2.2. Implementation

Theorem 8 suggests that achieving a small rate loss requires a large block length, which is confirmed in Example 9.2. Since the size of $\mathcal{T}^n(P_X)$ grows exponentially with n, the implementation of the CCDM by a LUT as in (9.7) becomes infeasible because of memory limitations. In [28], we propose an algorithm based on arithmetic coding that performs the CCDM mapping without storing $\mathcal{T}^n(P_X)$. For binary output alphabets, an algorithm similar to [28] was proposed in [30].

9.3. Distribution Quantization

Suppose we want to use the distribution P_X in our system by using a CCDM with output length n, however, P_X is not n-type. Thus, we must approximate P_X by an n-type distribution $P_{X'}$. Here, we will quantify how good $P_{X'}$ approximates P_X by the variational distance, which is given by

$$\|P_X - P_{X'}\|_1 = \sum_{a \in \mathcal{X}} |P_X(a) - P_{X'}(a)|. \tag{9.21}$$

Other measures, e.g., the informational divergence can be used; see for example [31]. We will first argue why the variational distance is a reasonable choice for our purposes and we will then state a simple algorithm to find an n-type approximation and bound the approximation error.

9.3.1. n-Type Approximation for CCDM

Two important parameters for system design are power and rate. Suppose the variational distance of P_X and $P_{X'}$ is equal to δ. Let $a_{\max}$ be the symbol in $\mathcal{X}$ of largest power. The power resulting from using $P_{X'}$ is then bounded above and below by

$$\mathbb{E}(X^2) - \delta a_{\max}^2 \leq \mathbb{E}(X'^2) \leq \mathbb{E}(X^2) + \delta a_{\max}^2. \tag{9.22}$$

In particular, as the variational distance δ approaches zero, the power $\mathbb{E}(X'^2)$ approaches the desired power $\mathbb{E}(X^2)$. Next, consider the asymptotic CCDM rate given by the entropy. By the continuity of entropy (C.19), if $\delta \leq \frac{1}{2}$, we have

$$\mathbb{H}(P_X) + \delta \log_2 \frac{\delta}{|\mathcal{X}|} \leq \mathbb{H}(P_{X'}) \leq \mathbb{H}(P_X) - \delta \log_2 \frac{\delta}{|\mathcal{X}|}. \tag{9.23}$$

Again, as δ approaches zero, the rate $\mathbb{H}(P_{X'})$ approaches the desired rate $\mathbb{H}(P_X)$.

9.3.2. n-Type Approximation Algorithm

The following algorithm calculates an n-type approximation $P_{X'}$ for an arbitrary distribution P_X.

1. For each $a \in \mathcal{X}$, calculate

   $$Q(a) = \frac{\lfloor nP_X(a) \rfloor}{n} \tag{9.24}$$

   and define

   $$L = n - \sum_{a \in \mathcal{X}} Q(a)n. \tag{9.25}$$

   Note that by definition, L is an integer.

2. For L symbols with largest approximation error $P_X(a) - Q(a)$, assign $P_{X'}(a) = Q(a) + \frac{1}{n}$. For the remaining symbols, assign $P_{X'}(a) = Q(a)$.

The algorithm immediately implies

$$|P_{X'}(a) - P_X(a)| < \frac{1}{n}, \quad a \in \mathcal{X} \tag{9.26}$$

and

$$\|P_{X'} - P_X\|_1 < \frac{|\mathcal{X}|}{n}. \tag{9.27}$$

Example 9.3. Consider the distribution $P_X(0) = 1 - P_X(1) = 1/\pi$ and $n = 1 \times 10^3$. The rounding step of the algorithm yields

$$Q(0) = \frac{318}{1000}, \quad Q(1) = \frac{681}{1000} \tag{9.28}$$

and $L = 1000 - 318 - 681 = 1$. The approximation errors are

$$P_X(0) - Q(0) \approx 3.1 \times 10^{-4}, \quad P_X(1) - Q(1) \approx 6.9 \times 10^{-4} \tag{9.29}$$

so we increase $Q(1)$ by $1/n$ and leave $Q(0)$ unchanged. The resulting n-type approximation is

$$P_{X'}(0) = \frac{318}{1000}, \quad P_{X'}(1) = \frac{682}{1000}. \tag{9.30}$$

The variational distance is

$$\|P_{X'} - P_X\|_1 \approx 6.1977 \times 10^{-4}. \tag{9.31}$$

Remark 9. In [31], it is shown that the above algorithm is optimal in terms of variational distance, and furthermore, the bound (9.27) is tightened.

9.4. Proof of Theorem 8

We now prove that the rate loss of the CCDM is $\Theta(\frac{\log n}{n})$.

In our proof, we will use Stirling's formula (A.5), which provides the upper bound

$$n! < \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} e^{\frac{1}{12n}}. \tag{9.32}$$

To get rid of the $e^{\frac{1}{12n}}$ term, which depends on n, we use instead

$$n! \leq e \cdot n^{n+\frac{1}{2}} e^{-n}. \tag{9.33}$$

For $n = 1$, (9.33) holds with equality, and for $n \geq 2$, we have $\sqrt{2\pi}e^{\frac{1}{12n}} < e$, so (9.33) follows by (9.32). Stirling's formula (A.5) provides the lower bound

$$n! > \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}e^{\frac{1}{12n+1}} > \sqrt{2\pi} \cdot n^{n+\frac{1}{2}}e^{-n}. \tag{9.34}$$

Let P_X be an n-type on an alphabet of size M. We will need in our proof the identity

$$\frac{n^n}{n_1^{n_1}n_2^{n_2}\cdots n_M^{n_M}} = 2^{n\,\mathbb{H}(P_X)} \tag{9.35}$$

which is derived in Problem 9.3.

For the size of the n-type class, we now have

$$\begin{aligned}|\mathcal{T}^n(P_X)| = \frac{n!}{n_1!\cdots n_M!} &< \frac{en^{n+\frac{1}{2}}e^{-n}}{\prod_{i=1}^M \sqrt{2\pi}n_i^{n_i+\frac{1}{2}}e^{-n_i}} &(9.36)\\ &= \frac{e}{(2\pi)^{\frac{M}{2}}} \cdot \frac{n^{n+\frac{1}{2}}}{\prod_{i=1}^M n_i^{n_i+\frac{1}{2}}} \cdot \frac{e^{-n}}{e^{-(n_1+n_2+\cdots+n_M)}} &(9.37)\\ &= \underbrace{\frac{e}{(2\pi)^{\frac{M}{2}}}}_{=:K_1} \cdot 2^{n\,\mathbb{H}(P_X)}\sqrt{\frac{n}{\prod_{i=1}^M n_i}} &(9.38)\end{aligned}$$

where

- (9.36) follows by using (9.33) and (9.34),
- (9.38) follows by (9.35) and $n_1 + \cdots + n_M = n$.

Taking the logarithm and dividing by n, we get

$$\begin{aligned}\frac{\log_2|\mathcal{T}^n(P_X)|}{n} &\leq \mathbb{H}(P_X) + \frac{\log_2 K_1}{n} + \frac{1}{2n}\log_2\frac{n}{\prod_{i=1}^M P_X(i)n} &(9.39)\\ &= \mathbb{H}(P_X) + \frac{\log_2 K_1 - \frac{1}{2}\log_2\prod_{i=1}^M P_X(i)}{n} - \frac{M-1}{2}\frac{\log_2 n}{n}. &(9.40)\end{aligned}$$

Along the same lines, we get the lower bound

$$\frac{\log_2|\mathcal{T}^n(P_X)|}{n} \geq \mathbb{H}(P_X) + \frac{\log_2 K_2 - \frac{1}{2}\log_2\prod_{i=1}^M P_X(i)}{n} - \frac{M-1}{2}\frac{\log_2 n}{n}. \tag{9.41}$$

where

$$K_2 = \frac{\sqrt{2\pi}}{e^M}. \tag{9.42}$$

Finally, the CCDM rate is bounded below and above by

$$\frac{\log_2|\mathcal{T}^n(P_X)|}{n} - \frac{1}{n} < \frac{\lfloor\log_2|\mathcal{T}^n(P_X)|\rfloor}{n} \leq \frac{\log_2|\mathcal{T}^n(P_X)|}{n} \tag{9.43}$$

The rate loss

$$R_{\text{loss}}(P_X, n) = \mathbb{H}(P_X) - \frac{\lfloor \log_2 |\mathcal{T}^n(P_X)| \rfloor}{n} \tag{9.44}$$

is now by (9.43) and (9.40) bounded above by $\frac{\log n}{n}$ asymptotically, i.e., $R_{\text{loss}}(P_X, n) \in \mathcal{O}(\frac{\log n}{n})$, and by (9.43) and (9.41), it is bounded below by $\frac{\log n}{n}$ asymptotically, i.e., $R_{\text{loss}}(P_X, n) \in \Omega(\frac{\log n}{n})$. This implies the desired result

$$R_{\text{loss}}(P_X, n) \in \Theta\left(\frac{\log n}{n}\right). \tag{9.45}$$

9.5. Problems

Problem 9.1. In this problem, we calculate a CCDM rate. We consider the following setup.

- 8-ASK constellation.
- $n = 21\,600$ channel uses.
- Amplitude distribution

$$P_A(1) = \frac{13\,258}{n}, \quad P_A(3) = \frac{6550}{n}, \quad P_A(5) = \frac{1599}{n}, \quad P_A(7) = \frac{193}{n}.$$

1. Calculate the entropy $\mathbb{H}(P_A)$.

A CCDM can encode $k = \lfloor \log_2 |\mathcal{T}^n(P_A)| \rfloor$ source bits. Our next task is to calculate this number.

2. Define $P_A(a) = c_a/n$. Argue that

$$|\mathcal{T}^n(P_A)| = \binom{n}{c_1, c_3, c_5, c_7}.$$

3. Express the multinomial coefficient as the product of binomial coefficients.
4. $n!$ is an incredibly large number. How can you calculate $\log_2 \binom{n}{k}$ avoiding large numbers?
5. Combine solutions 3. and 4. to derive a strategy for efficiently calculating $\log_2 |\mathcal{T}^n(P_A)|$.
6. Compare $R_{\text{ccdm}} = \lfloor \log_2 |\mathcal{T}^n(P_A)| \rfloor / n$ to $\mathbb{H}(P_A)$.

Problem 9.2. Let P_X be an n-type on the alphabet $\mathcal{X}$ and suppose $k = \log_2 |\mathcal{T}^n(P_X)|$ is an integer, so that the CCDM encodes onto $\mathcal{T}^n(P_X)$. Let U^n be uniformly distributed on $\{0,1\}^k$ and define $C^n = \text{ccdm}(U^k)$. Show that the marginal distributions P_{C_i} are equal to P_X, i.e.,

$$P_{C_i}(a) = P_X(a), \quad a \in \mathcal{X}, i = 1, 2, \ldots, n. \tag{9.46}$$

Problem 9.3. Show the identity (9.35).
Problem 9.4. Use (9.40) to show (9.13).

10. Error Exponents

In this chapter, we use the techniques developed in Chapter 7 to derive error exponents and achievable rates for the PAS transceiver architecture that we developed in Chapter 6. This chapter extends the results in Chapter 7 as follows.

- The non-constructive shaping layer in Section 7.2 based on random coding is replaced by a constructive shaping layer using the CCDM developed in Chapter 9.
- The results of this chapter also hold when the FEC layer uses a random linear code.
- The error exponents can be used for finite block length.

The results of this chapter rely on the assumption that following the PAS principle, the input distribution of interest decomposes into two independent distributions, namely a potentially non-uniform 'amplitude' distribution and a uniform 'sign' distribution. In this sense, the results of this chapter are less general than the results in Chapter 7.

10.1. FEC Layer

We consider the following transceiver setup:

- (As in Section 7.1) The channel is discrete-time with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$. We derive our results assuming a continuous-valued output. Our results also apply for discrete output alphabets.
- (**Different** from Section 7.1) Random coding: For indices $w = 1, 2, \dots, |\mathcal{C}|$, we generate code words $C^n(w)$ according to a general set of distributions $P_{C^n(w)}$. In particular, we permit different distributions for different indices and we also allow dependence among the code word entries of the same code word. The code is
$$\mathcal{C} = \{C^n(1), C^n(2), \dots, C^n(|\mathcal{C}|)\}. \tag{10.1}$$
- (As in Section 7.1) The code rate is $R_\mathrm{c} = \frac{\log_2(|\mathcal{C}|)}{n}$ and equivalently, we have $|\mathcal{C}| = 2^{nR_\mathrm{c}}$ code words.
- (As in Section 7.1) We consider a non-negative decoding metric q on $\mathcal{X} \times \mathcal{Y}$ and we define the memoryless metric
$$q^n(x^n, y^n) := \prod_{i=1}^{n} q(x_i, y_i), \quad x^n \in \mathcal{X}^n, y^n \in \mathcal{Y}^n. \tag{10.2}$$

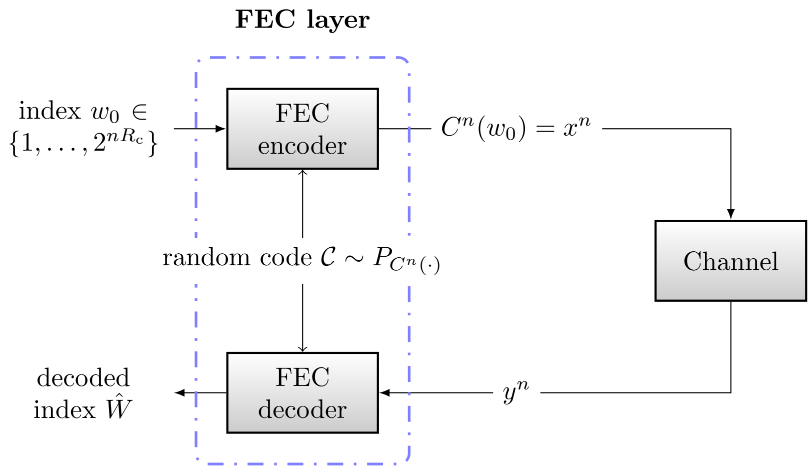


Figure 10.1.: Random coding experiment for code rate error exponent. In difference to Figure 7.1, the code words generated for each index w according to an individual distribution $P_{C^n(w)}$.

For the channel output y^n, we let the receiver decode with the rule

$$\hat{W} = \operatorname*{argmax}_{w \in \{1,\dots,2^{nR_c}\}} \prod_{i=1}^{n} q(C_i(w), y_i). \tag{10.3}$$

- (As in Section 7.1) We consider the decoding error probability

$$P_e = \Pr(\hat{W} \neq w_0 | C^n(w_0) = x^n, Y^n = y^n) \tag{10.4}$$

where w_0 is the index of the transmitted code word, $C^n(w_0) = x^n$ is the transmitted code word, y^n is the channel output sequence, and $\hat{W}$ is the decoded index at the receiver. Note that the code words $C^n(w)$, $w \neq w_0$ against which the decoder attempts to decode are random and the transmitted code word $C^n(w_0) = x^n$ and the channel output y^n are deterministic.

10.1.1. Code Rate Error Exponent

We consider the setting in Figure 10.1, i.e., we condition on that index w_0 was encoded to $C^n(w_0) = x^n$. For notational convenience, we assume without loss of generality $w_0 = 1$. We analyze the error probability $\Pr(\hat{W} \neq 1 | C^n(1)) = x^n, Y^n = y^n)$, averaged over the random code. Note that we have $C^n(1) = x^n$ and for $w = 2, 3, \dots, |\mathcal{C}|$, we have $C^n(w) \sim P_{C^n(w)}$.

We have the implications

$$\hat{W} \neq 1 \Rightarrow \hat{W} = w' \neq 1 \tag{10.5}$$

$$\Rightarrow L(w') := \frac{q^n(C^n(w'), y^n)}{q^n(x^n, y^n)} \geq 1 \tag{10.6}$$

$$\Rightarrow \sum_{w=2}^{|\mathcal{C}|} L(w) \geq 1 \tag{10.7}$$

$$\Rightarrow \Big[\sum_{w=2}^{|\mathcal{C}|} L(w)\Big]^\rho \geq 1, \qquad \rho \geq 0. \tag{10.8}$$

We will use ρ for the same purposes as in [3, Chapter 5]: We will optimize over ρ to maximize the error exponent and to identify achievable rates. If event $\mathcal{A}$ implies event $\mathcal{B}$, then $\Pr[\mathcal{A}] \leq \Pr[\mathcal{B}]$. Therefore, we have

$$\begin{aligned}
&\Pr(\hat{W} \neq 1 | C^n(1) = x^n, Y^n = y^n) \\
&\leq \Pr\left\{ \Big[\sum_{w=2}^{|\mathcal{C}|} L(w)\Big]^\rho \geq 1 \,\middle|\, C^n(1) = x^n, Y^n = y^n \right\} && (10.9)\\
&\leq \mathbb{E}\left\{ \Big[\sum_{w=2}^{|\mathcal{C}|} L(w)\Big]^\rho \,\middle|\, C^n(1) = x^n, Y^n = y^n \right\} && (10.10)\\
&= q^n(x^n, y^n)^{-\rho}\, \mathbb{E}\left[\Big[\sum_{w=2}^{|\mathcal{C}|} q^n[C^n(w), y^n]\Big]^\rho \right] && (10.11)
\end{aligned}$$

where

- the inequality in (10.10) follows by Markov's inequality (B.6),
- equality in (10.11) follows because for $w \neq 1$, the code word $C^n(w)$ and the transmitted code word $C^n(1)$ were generated independently so that $C^n(w)$ and $[C^n(1), Y^n]$ are independent.

Observe that $z \to z^\rho$ is for $0 \leq \rho \leq 1$ a concave function and for $1 \leq \rho$ a convex function, see Figure 10.2. We therefore restrict the parameter ρ to

$$0 \leq \rho \leq 1 \tag{10.12}$$

so that $(\cdot)^\rho$ is concave and by Jensen's inequality (A.7), we have $\mathbb{E}(Z^\rho) \leq \mathbb{E}(Z)^\rho$. We use this to bound (10.11) further:

$$\begin{aligned}
(10.11) &\leq q^n(x^n, y^n)^{-\rho}\, \mathbb{E}\left[\sum_{w=2}^{|\mathcal{C}|} q^n[C^n(w), y^n]\right]^\rho && (10.13)\\
&= |\mathcal{C}|^\rho q^n(x^n, y^n)^{-\rho}\, \mathbb{E}\left[\frac{1}{|\mathcal{C}|}\sum_{w=2}^{|\mathcal{C}|} q^n[C^n(w), y^n]\right]^\rho . && (10.14)
\end{aligned}$$

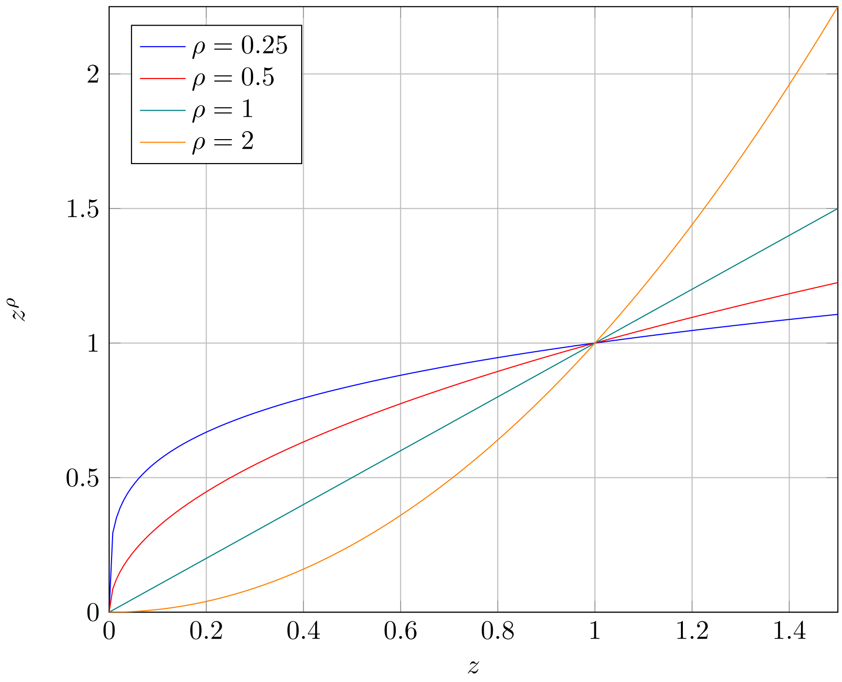


Figure 10.2.: The function $z \mapsto z^\rho$ is concave for $0 \leq \rho \leq 1$ and it is convex for $1 \leq \rho$.

10.1.2. PAS Code Rate Error Exponent

So far, we have analyzed the error probability of decoding a transmitted code word $C^n(w_0) = x^n$ against the random code $\mathcal{C} \sim P_{C^n(\cdot)}$. We next consider a specific instance of the random coding experiment $P_{C^n(\cdot)}$, following the probabilistic amplitude shaping (PAS) principle that we developed in Chapter 6. This will allow us to explicitly state an encoder and to quantify into how many distinct code words we can encode. In this way, we will be able to associate a rate with the error exponent.

We now make the assumption that the input alphabet decomposes into two parts $\mathcal{X} = \mathcal{A} \times \mathcal{S}$. We represent the code word index by $w = a^n s^{\gamma n}$, i.e., code size and code rate are respectively given by

$$|\mathcal{C}| = |\mathcal{A}|^n |\mathcal{S}|^{\gamma n} \tag{10.15}$$

$$R_\mathrm{c} = \log_2 |\mathcal{A}| + \gamma \log_2 |\mathcal{S}|. \tag{10.16}$$

Example 10.1 (Amplitude Shift Keying)**.** The 8-ASK constellation

$$\mathcal{X} = \{\pm 1, \pm 3, \pm 5, \pm 7\} \tag{10.17}$$

decomposes into an amplitude set and a sign set via

$$\mathcal{A} = \{1,3,5,7\}, \quad \mathcal{S} = \{-1,1\} \tag{10.18}$$
$$\mathcal{A} \times \mathcal{S} \to \mathcal{X}\colon (a,s) \mapsto a \cdot s \tag{10.19}$$
$$\mathcal{X} \to \mathcal{A} \times \mathcal{S}\colon x \mapsto (|x|, \operatorname{sign}(x)). \tag{10.20}$$

For a 2^m-ASK constellation, the code rate is

$$R_\text{c} = \log_2 |\mathcal{A}| + \gamma \log_2 |\mathcal{S}| = m - 1 + \gamma. \tag{10.21}$$

Example 10.2 (Quadrature Amplitude Modulation)**.** The 16-quadrature amplitude modulation (QAM) constellation

$$\mathcal{X} = \{\pm 1 \pm j, \pm 1 \pm 3j, \pm 3 \pm j, \pm 3 \pm 3j\} \tag{10.22}$$

decomposes into

$$\mathcal{A} = \{(\pm 1, \pm 3)\}, \quad \mathcal{S} = \{(\pm 1, \pm 1)\} \tag{10.23}$$
$$\mathcal{A} \times \mathcal{S} \to \mathcal{X}\colon (a,s) \mapsto a_1 s_1 + j a_2 s_2 \tag{10.24}$$
$$\mathcal{X} \to \mathcal{A} \times \mathcal{S}\colon x \mapsto [(|\operatorname{Re}(x)|, |\operatorname{Im}(x)|), (\operatorname{sign}[\operatorname{Re}(x)], \operatorname{sign}[\operatorname{Im}(x)])]. \tag{10.25}$$

Note that this is equivalent to interpreting 16-QAM as the Cartesian product of two 4-ASK constellations.

Our random coding experiment is as follows. We encode the index $w \in \mathcal{A}^n \times \mathcal{S}^{\gamma n}$ to

$$w = a^n s^{\gamma n} \mapsto C^n(w) = a^n S^n(w) \tag{10.26}$$

where the $S_i(w)$, $i = 1, 2, \ldots, n$ are independent and uniformly distributed on $\mathcal{S}$. Note that this corresponds to partially systematic encoding: the part a^n of the code word index appears in the code word. Next, we condition on an index $w_0 = a^n s^{\gamma n}$. This index selects a code word $a^n S^n$, which consists of the deterministic part a^n and a random part S^n, which is stochastically independent of w_0. In terms of the random coding distribution $P_{C^n(\cdot)}$, we have

$$P_{C^n(a^n s^{\gamma n})}(\alpha^n \sigma^n) = \begin{cases} \frac{1}{|\mathcal{S}|^n}, & \text{if } \alpha^n = a^n \\ 0, & \text{otherwise} \end{cases}, \quad a^n, \alpha^n \in \mathcal{A}^n, s^{\gamma n} \in \mathcal{S}^{\gamma n}, \sigma^n \in \mathcal{S}^n \tag{10.27}$$

where $P_{C^n(w_0)}$ indeed depends on w_0. The PAS coding experiment is displayed in Figure 10.3. We have the following differences to the coding experiment in Figure 10.1:

1. For the transmitted code word, we condition on $C^n(w_0) = a^n S^n$, which is random.
2. Since the transmitted code word is random, also the channel output Y^n is random. We therefore analyze the average error probability, where the average is over S^n and Y^n.

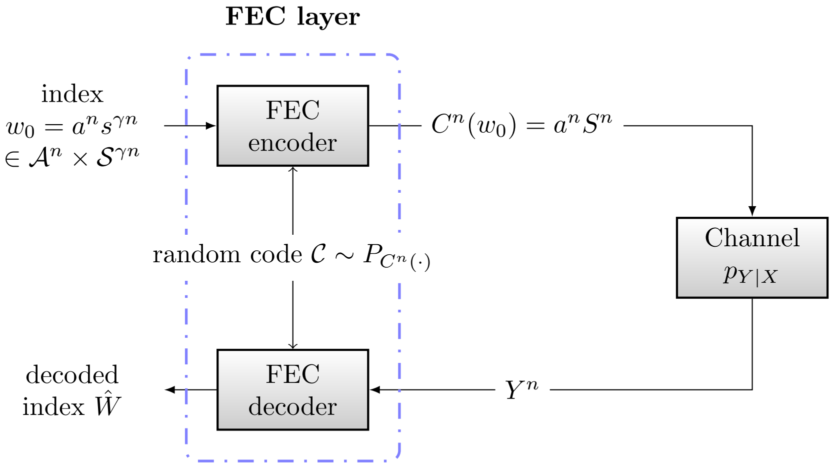


Figure 10.3.: PAS random coding experiment.

3. We assume the channel is memoryless with channel law $p_{Y|X}$.

We now have

$$\Pr(\hat{W} \neq w_0 | A^n = a^n) = \mathbb{E}\left[\Pr(\hat{W} \neq w_0 | X^n = a^n S^n, Y^n)\right] \tag{10.28}$$

$$\overset{(10.14)}{\leq} |\mathcal{C}|^\rho \, \mathbb{E}\left\{ q^n(a^n S^n, Y^n)^{-\rho} \, \mathbb{E}\left[\frac{1}{|\mathcal{C}|} \sum_{\substack{w \in \mathcal{A}^n \times \mathcal{S}^{\gamma n} \\ w \neq w_0}} q^n[C^n(w), Y^n] \middle| Y^n \right]^\rho \middle| A^n = a^n \right\}. \tag{10.29}$$

We develop the innermost expectation in (10.29). We have

$$
\begin{aligned}
&\mathbb{E}\left[\frac{1}{|\mathcal{C}|}\sum_{\substack{w\in\mathcal{A}^n\times\mathcal{S}^{\gamma n}\\ w\neq w_0}} q^n[C^n(w),Y^n]\middle| Y^n\right]\\
&\qquad= \mathbb{E}\left[\frac{1}{|\mathcal{C}|}\sum_{\substack{w\in\mathcal{A}^n\times\mathcal{S}^{\gamma n}\\ w\neq w_0}} q^n[a^n(w)S^n(w),Y^n]\middle| Y^n\right] && (10.30)\\
&\qquad= \frac{1}{|\mathcal{A}|^n|\mathcal{S}|^{\gamma n}}\sum_{\substack{w\in\mathcal{A}^n\times\mathcal{S}^{\gamma n}\\ w\neq w_0}}\sum_{t^n\in\mathcal{S}^n}\frac{1}{|\mathcal{S}|^n}q^n[a^n(w)t^n,Y^n] && (10.31)\\
&\qquad= \frac{1}{|\mathcal{S}|^{\gamma n}}\sum_{\substack{w\in\mathcal{A}^n\times\mathcal{S}^{\gamma n}\\ w\neq w_0}}\sum_{t^n\in\mathcal{S}^n}\frac{1}{|\mathcal{X}|^n}q^n[a^n(w)t^n,Y^n] && (10.32)\\
&\qquad= \frac{1}{|\mathcal{S}|^{\gamma n}}\sum_{\substack{w\in\mathcal{A}^n\times\mathcal{S}^{\gamma n}\\ w\neq w_0}}\sum_{t^n\in\mathcal{S}^n}\prod_{i=1}^n\frac{1}{|\mathcal{X}|}q[a_i(w)t_i,Y_i] && (10.33)\\
&\qquad\leq \frac{1}{|\mathcal{S}|^{\gamma n}}\sum_{a^n\in\mathcal{A}^n}\sum_{s^{\gamma n}\in\mathcal{S}^{\gamma n}}\sum_{t^n\in\mathcal{S}^n}\prod_{i=1}^n\frac{1}{|\mathcal{X}|}q(a_it_i,Y_i) && (10.34)\\
&\qquad= \sum_{a^n\in\mathcal{A}^n}\sum_{t^n\in\mathcal{S}^n}\prod_{i=1}^n\frac{1}{|\mathcal{X}|}q(a_it_i,Y_i) && (10.35)\\
&\qquad= \prod_{i=1}^n\sum_{a\in\mathcal{A}}\sum_{t\in\mathcal{S}}\frac{1}{|\mathcal{X}|}q(at,Y_i) && (10.36)\\
&\qquad= \prod_{i=1}^n\sum_{c\in\mathcal{X}}\frac{1}{|\mathcal{X}|}q(c,Y_i),\quad \text{with}\quad a^nS^n\rightarrow\boxed{\text{Channel}}\rightarrow Y^n && (10.37)
\end{aligned}
$$

where

- (10.30) follows by (10.26),
- (10.31) follows because for each $w\neq w_0$, $S^n(w)$ is independent of Y^n and because the entries of the $S^n(w)$ are independent and uniformly distributed on $\mathcal{S}$,
- (10.32) and (10.37) use $\mathcal{X}=\mathcal{A}\times\mathcal{S}$,
- (10.33) follows because the metric q^n is memoryless,
- the inequality in (10.34) follows since the sum now includes index w_0.

Remark 10. Note that we would get (10.37) also for a random code with its $n|\mathcal{C}|$ entries independent and uniformly distributed on $\mathcal{X}$. Our PAS random code can thus be interpreted as a partially systematic version of a uniformly distributed random code.

Remark 11. Note that we have not yet made use of the assumption that the channel is memoryless, in particular, (10.37) also holds for channels with memory.

By inserting (10.37) in the outer expectation in (10.29) we have

$$
\begin{aligned}
&\mathbb{E}\left\{\prod_{i=1}^{n} q(a_i S_i, Y_i)^{-\rho}\left[\sum_{c\in\mathcal{X}}\frac{1}{|\mathcal{X}|}q(c,Y_i)\right]^{\rho}\middle| A^n = a^n\right\}\\
&\qquad = \prod_{i=1}^{n}\mathbb{E}\left\{q(a_i S_i, Y_i)^{-\rho}\left[\sum_{c\in\mathcal{X}}\frac{1}{|\mathcal{X}|}q(c,Y_i)\right]^{\rho}\middle| A_i = a_i\right\} && (10.38)\\
&\qquad = \prod_{i=1}^{n}\mathbb{E}\left\{q(a_i S, Y)^{-\rho}\left[\sum_{c\in\mathcal{X}}\frac{1}{|\mathcal{X}|}q(c,Y)\right]^{\rho}\middle| A = a_i\right\} && (10.39)\\
&\qquad = \prod_{\alpha\in\mathcal{A}}\mathbb{E}\left\{q(\alpha S, Y)^{-\rho}\left[\sum_{c\in\mathcal{X}}\frac{1}{|\mathcal{X}|}q(c,Y)\right]^{\rho}\middle| A = a\right\}^{N(\alpha|a^n)} && (10.40)
\end{aligned}
$$

where

- equality in (10.38) follows because by assumption, the channel is memoryless, so conditioned on $A^n = a^n$, S_iY_i and S_jY_j are independent for $i \neq j$,
- for the ith factor in (10.39), $a_iS \to \boxed{p_{Y|X}} \to Y$,
- for the αth factor in (10.40), $\alpha S \to \boxed{p_{Y|X}} \to Y$,
- in (10.40),

$$N(\alpha|a_n) = |\{i\colon a^n = \alpha\}| = \text{number of occurrences of letter } \alpha \text{ in } a^n. \tag{10.41}$$

We denote the empirical distribution of a^n by P_A and write $N(\alpha|a^n) = nP_A(\alpha)$. Evaluating $-\frac{1}{n}\log_2(\cdot)$ in (10.40) gives

$$\tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho, P_A, q) = \sum_{a\in\mathcal{A}} P_A(a)\log_2\mathbb{E}\left\{\left[\frac{q(aS,Y)}{\sum_{c\in\mathcal{X}}\frac{1}{|\mathcal{X}|}q(c,Y)}\right]^{\rho}\middle| A = a\right\}. \tag{10.42}$$

The PAS error exponent and the bound on decoding error probability are respectively

$$
\begin{aligned}
\mathrm{E}_{\mathrm{PAS}}(R_{\mathrm{c}}, \rho, P_A, q) &= \tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho, P_A, q) - \rho R_{\mathrm{c}} && (10.43)\\
\Pr(\hat{W} \neq W | A^n = a^n) &\leq 2^{-n\mathrm{E}_{\mathrm{PAS}}(R_{\mathrm{c}},\rho,P_A,q)}. && (10.44)
\end{aligned}
$$

10.1.3. PAS with Random Linear Coding

Suppose we have $|\mathcal{X}| = 2^m$, $|\mathcal{A}| = 2^{m-1}$, and $|\mathcal{S}| = 2$, for instance, $\mathcal{X}$ may be a 2^m-ASK constellation. We represent $\mathcal{A}$ by a binary label $\boldsymbol{b} = b_1\cdots b_{m-1} \in \{0,1\}^{m-1}$ and $\mathcal{S}$ by

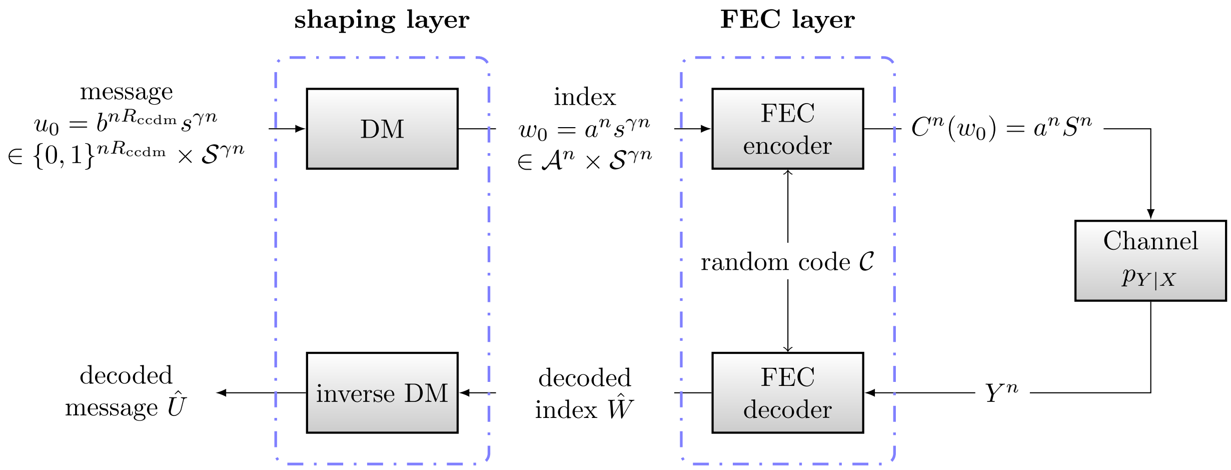

shaping layer
FEC layer
message
$u_0 = b^{nR_{\text{ccdm}}} s^{\gamma n}$
$\in \{0,1\}^{nR_{\text{ccdm}}} \times \mathcal{S}^{\gamma n}$
DM
index
$w_0 = a^n s^{\gamma n}$
$\in \mathcal{A}^n \times \mathcal{S}^{\gamma n}$
FEC encoder
$C^n(w_0) = a^n S^n$
Channel
$p_{Y|X}$
random code $\mathcal{C}$
decoded message $\hat{U}$
inverse DM
decoded index $\hat{W}$
FEC decoder
Y^n

Figure 10.4.: The PAS ensemble. The shaping layer is deterministic. The DM maps nR_{ccdm} input bits to n symbols in $\mathcal{A}$ using a CCDM and it copies γn input symbols in $\mathcal{S}$ unchanged to its output.

$s \in \{0,1\}$. For this scenario, we next show that our results for PAS also hold for random linear coding. This can be generalized to other scenarios in a straight forward manner. For the considered scenario, the PAS random coding (10.26) is

$$\boldsymbol{w} = \boldsymbol{b}^n s^{\gamma n} \mapsto \boldsymbol{b}^n S^n(\boldsymbol{w}), \quad \boldsymbol{w} \in \{0,1\}^{(m-1+\gamma)n} \tag{10.45}$$

where the entries of $S^n(\boldsymbol{w})$, $\boldsymbol{w} \in \{0,1\}^{(m-1+\gamma)n}$ are independent and uniformly distributed on $\{0,1\}$. We used two properties in our derivation:

1. In (10.11) and (10.31), we used that for $\boldsymbol{w} \neq \boldsymbol{w}'$, $S^n(\boldsymbol{w})$ and $S^n(\boldsymbol{w}')$ are independent.

2. In (10.31), we used that for each $\boldsymbol{w} \in \{0,1\}^{(m-1+\gamma)n}$, the n entries of $S^n(\boldsymbol{w})$ are independent and uniformly distributed on $\{0,1\}$.

The results we derived apply for any coding experiment that fulfills properties 1. and 2. In particular, consider the random linear coding

$$\boldsymbol{w} = \boldsymbol{b}^n s^{\gamma n} \mapsto (\boldsymbol{b}^n, \boldsymbol{w}\boldsymbol{P}) \tag{10.46}$$

where $\boldsymbol{P}$ is an $(m-1+\gamma)n \times n$ matrix with its entries independent and uniformly distributed on $\{0,1\}$. By Problem 10.2, the random linear code (10.46) fulfills properties 1. and 2. Thus, the PAS error exponent and the PAS achievable rate also hold for the random linear coding (10.46).

10.2. Shaping Layer

We display the full PAS scheme including the shaping layer in Figure 10.4. Our goal is to associate a rate with the PAS error exponent. Recall that the code rate is

$$R_{\mathrm{c}} = \log_2|\mathcal{A}| + \gamma \log_2|\mathcal{S}|. \tag{10.47}$$

The PAS error exponent $\mathrm{E}_{\mathrm{PAS}}(R_c, \rho, P_A, q)$ depends on the type of the amplitude sequence a^n. Thus, the number of code words for which it holds is

$$(\text{number of sequences } a^n \text{ of type } P_A) \times |\mathcal{S}|^{\gamma n}. \tag{10.48}$$

Using the CCDM we developed in Chapter 9, we can encode $R_{\mathrm{ccdm}}(P_A, n) \cdot n$ bits to length n sequences of type P_A. In total, we can encode

$$[R_{\mathrm{ccdm}}(P_A, n) + \gamma \log_2|\mathcal{S}|] \cdot n \text{ bits} \tag{10.49}$$

and the rate is

$$R_{\mathrm{tx}} = R_{\mathrm{ccdm}}(P_A, n) + \gamma \log_2|\mathcal{S}| \quad \left[\frac{\text{bits}}{\text{symbol}}\right]. \tag{10.50}$$

Remark 12. We can express the code rate in terms of the transmission rate by

$$R_{\mathrm{c}} = R_{\mathrm{tx}} + [\log_2|\mathcal{A}| - R_{\mathrm{ccdm}}(P_A, n)]\,. \tag{10.51}$$

Note that since $R_{\mathrm{ccdm}}(P_A, n) < \log_2|\mathcal{A}|$, the code rate is larger than the transmission rate. This makes sense, since of all code words, we only transmit those that have amplitude type P_A.

The following theorem summarizes our findings.

Theorem 9. *The PAS scheme can operate at the rate*

$$R_{\mathrm{tx}} = R_{\mathrm{ccdm}}(P_A, n) + \gamma \log_2|\mathcal{S}| \quad \left[\frac{\text{bits}}{\text{symbol}}\right] \tag{10.52}$$

with the decoding error probability bounded from above by

$$2^{-\mathrm{E}_{\mathrm{PAS}}(R_{\mathrm{c}}, \rho, P_A, q)}. \tag{10.53}$$

10.3. PAS Achievable Rates

By making n large, the error probability bound approaches zero, as long as the error exponent is positive, which is the case when $\tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho, P_A, q)$ is larger than ρR_{c}.

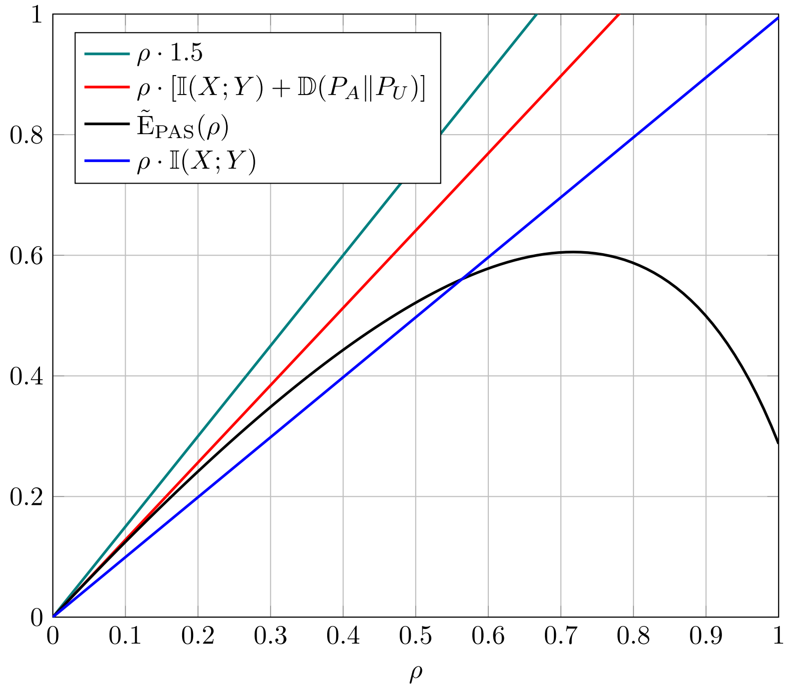


Figure 10.5.: Visualization of error exponent. By increasing the code rate R_c, the slope of $R_\mathrm{c}\rho$ becomes steeper, until it exceeds $\tilde{\mathrm{E}}_{\mathrm{PAS}}$ everywhere. The phase transition occurs for $R_\mathrm{c} = \mathbb{I}(X;Y) + \mathbb{D}(A\|U)$ in $\rho = 0$.

Example 10.3. Consider an AWGN channel

$$Y = X + Z \tag{10.54}$$

where $X = A{\cdot}S$ is 4-ASK input with alphabet $\mathcal{X} = \{\pm 1, \pm 3\}$, amplitude distribution $P_A(1) = \frac{4}{5}$, $P_A(3) = \frac{1}{5}$, uniform sign distribution $P_S(1) = P_S(-1) = \frac{1}{2}$ and where Z is zero mean Gaussian with variance $\sigma^2 = \mathbb{E}(X^2)/3$, i.e., the SNR is 3 and the AWGN channel capacity is $\frac{1}{2}\log_2(1 + \mathsf{snr}) = 1$. For the decoding metric

$$q(a,b) = P_X(a) p_{Y|X}(b|a). \tag{10.55}$$

We plot $\tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho)$ and ρR_c in Figure 10.5. We make the following two observations:

- The function $\tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho)$ is concave in ρ.
- $\tilde{\mathrm{E}}_{\mathrm{PAS}}(0) = 0$.

This implies that the largest code rate is given by the slope of $\tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho)$ in $\rho = 0$.

We now show that our observations in Example 10.3 are true in general. First,

$$\tilde{\mathrm{E}}_{\mathrm{PAS}}(0) = 0 \tag{10.56}$$

follows directly by (10.42).

Theorem 10. $\tilde{\mathrm{E}}_{\mathrm{PAS}}(0)$ *is concave for* $\rho \in \mathbf{R}$.

Proof. In 10.4, we show the non-negativity of the second derivative of a more general expression, from which the concavity of $\tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho)$ follows. □

By (10.56) and the concavity of $\tilde{\mathrm{E}}_{\mathrm{PAS}}$, it follows that for $0 \leq \rho \leq 1$, the largest slope of $\tilde{\mathrm{E}}_{\mathrm{PAS}}$ is in $\rho = 0$. To calculate the slope, define

$$Z_a = \left\{ \log \frac{q(Y, aS)}{\sum_{c \in \mathcal{X}} \frac{1}{|\mathcal{X}|} q(Y, c)} \middle| A = a \right\}. \tag{10.57}$$

We now have

$$\begin{aligned} \tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho, P_A, q) &= \sum_{a \in \mathcal{A}} P_A(a) \log_2 \mathbb{E}(e^{\rho Z_a}) && (10.58)\\ &= \sum_{a \in \mathcal{A}} P_A(a) \frac{\log[\mathsf{mgf}_{Z_a}(\rho)]}{\log 2} && (10.59) \end{aligned}$$

where mgf is the moment generating function (MGF) (B.7) and by Problem 10.1, we have

$$\begin{aligned} \frac{\partial}{\partial \rho} \tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho, P_A, q)\Big|_{\rho=0} &= \sum_{a \in \mathcal{A}} P_A(a) \frac{\mathbb{E}(Z_a)}{\log 2} && (10.60)\\ &= \mathbb{E}\left[\log_2 \frac{q(X, Y)}{\sum_{c \in \mathcal{X}} \frac{1}{|\mathcal{X}|} q(c, Y)} \right]. && (10.61) \end{aligned}$$

Example 10.4. We continue with Example 10.3. For the decoding metric $q(c, b) = P_X(c) p_{Y|X}(b|c)$, an achievable code rate is

$$\begin{aligned} T_{\mathrm{c}} &= \frac{\partial}{\partial \rho} \tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho, P_A, q)\Big|_{\rho=0} && (10.62)\\ &= \mathbb{E}\left[\log_2 \frac{P_X(X) p_{Y|X}(Y|X)}{\sum_{c \in \mathcal{X}} \frac{1}{|\mathcal{X}|} P_X(c) p_{Y|X}(b|c)} \right] && (10.63)\\ &= \mathbb{I}(X; Y) + \mathbb{D}(P_X \| P_U) && (10.64) \end{aligned}$$

where P_U is the uniform distribution on $\mathcal{X}$. By Theorem 8, for large n a CCDM generates type P_A sequences with rate $R_{\mathrm{ccdm}}(P_A, \infty) = \mathbb{H}(A)$. The rate is

$$\begin{aligned} R_{\mathrm{tx}} = \mathbb{H}(A) + \gamma &= \mathbb{H}(A) + R_{\mathrm{c}} - m + 1 && (10.65)\\ &= R_{\mathrm{c}} - \mathbb{D}(P_X \| P_U) && (10.66) \end{aligned}$$

and the achievable rate is

$$R = T_{\mathrm{c}} - \mathbb{D}(P_X \| P_U) = \mathbb{I}(X;Y) \tag{10.67}$$

i.e., PAS asymptotically achieves $\mathbb{I}(X;Y)$.

10.4. Proof of Concavity of $\tilde{\mathrm{E}}_{\mathrm{PAS}}$

Consider the function

$$g(x) = \log\left(\sum_{i=1}^{n} a_i b_i^x\right) = \log\left(\sum_{i=1}^{n} a_i e^{x\log b_i}\right) \tag{10.68}$$

where the a_i and b_i are non-negative. We verify that $g(x)$ is convex on $\mathbf{R}$ by showing that the second derivative is non-negative. We have

$$\frac{\partial}{\partial x} g(x) = \frac{\sum_{i=1}^n \log(b_i) a_i e^{x\log b_i}}{\sum_{i=1}^n a_i e^{x\log b_i}} \tag{10.69}$$

$$\frac{\partial^2}{\partial x^2} g(x) = \frac{(\sum_{i=1}^n \log(b_i)^2 a_i e^{x\log b_i})(\sum_{i=1}^n a_i e^{x\log b_i}) - (\sum_{i=1}^n \log(b_i) a_i e^{x\log b_i})^2}{(\sum_{i=1}^n a_i e^{x\log b_i})^2} \tag{10.70}$$

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

$$u_i = \log(b_i)\sqrt{a_i e^{x\log b_i}} \tag{10.71}$$

$$v_i = \sqrt{a_i e^{x\log b_i}}. \tag{10.72}$$

The numerator of the second derivative is now

$$\boldsymbol{u}\boldsymbol{u}^T \boldsymbol{v}\boldsymbol{v}^T - (\boldsymbol{u}\boldsymbol{v}^T)^2 \tag{10.73}$$

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

10.5. Problems

Problem 10.1.

1. Show that

$$\left.\frac{\partial}{\partial r}\log[\mathsf{mgf}_X(r)]\right|_{r=0} = \mathbb{E}(X) \tag{10.74}$$

$$\left.\frac{\partial^2}{\partial r^2}\log[\mathsf{mgf}_X(r)]\right|_{r=0} = \mathrm{Var}(X) \tag{10.75}$$

where mgf_X is the MGF (B.7) of X.

2. Use (10.59) and (10.75) to show that $\tilde{\mathrm{E}}_{\mathrm{PAS}}(\rho)$ is concave in $\rho = 0$.

Problem 10.2. Consider the random linear code (10.46).

1. Show that for $\boldsymbol{w} \neq w'$, $w\boldsymbol{P}$ and $\boldsymbol{w}'\boldsymbol{P}$ are independent.

2. Show that the n entries of $\boldsymbol{wP}$ are independent and uniformly distributed.

11. Estimating Achievable Rates

In this chapter, we develop tools to estimate achievable rates for channels that potentially have memory and for which we lack a complete analytical description.

11.1. Preliminaries

Let's recall the main results on achievable code rates that we derived in Chapter 7:

- An input sequence x^n can be recovered with high probability from an output sequence y^n by a rate R_c random code with decoding metric q, if n is large and

$$R_\mathrm{c} < \hat{T}_\mathrm{c}(x^n, y^n, q) = \log_2 |\mathcal{X}| - \underbrace{\frac{1}{n}\sum_{i=1}^{n}\left[-\log_2 \frac{q(x_i, y_i)}{\sum_{a\in\mathcal{X}} q(a, y_i)}\right]}_{\text{uncertainty } U_\mathrm{c}}. \tag{11.1}$$

- For a memoryless channel $p_{Y|X}$, a sequence approximately of type P_X can be recovered with high probability from the random channel output Y^n by a rate R_c random code with decoding metric q if n is large and

$$R_\mathrm{c} < T_\mathrm{c}(P_X, p_{Y|X}, q) = \log_2 |\mathcal{X}| - \underbrace{\mathbb{E}\left[-\log_2 \frac{q(X, Y)}{\sum_{a\in\mathcal{X}} q(a, Y)}\right]}_{\text{uncertainty } U_\mathrm{c}}. \tag{11.2}$$

The terms (11.1) and (11.2) are fundamentally different:

- (11.1) works for any channel, but the calculated value applies only to the specific measurement x^n, y^n. In general, (11.1) tells us nothing about the code rate that is achievable by a sequence $\tilde{x}^n$ different from x^n. Consequently, we cannot attach an achievable transmission rate to the achievable code rate $\hat{T}_\mathrm{c}(x^n, y^n, q)$.

- (11.2) applies for memoryless channels and holds for any sequence x^n of type P_X. Consequently, we can attach an achievable transmission rate to the achievable code rate $T_\mathrm{c}(P_X, p_{Y|X}, q)$.

In practice, we would like to combine the advantages of (11.1) and (11.2): for a not necessarily memoryless channel, we would like to estimate an uncertainty that applies for the whole shaping set, i.e., the set of sequences that our transmitter may output. A strategy to accomplish this is as follows. From a measurement x^n, y^n of a not necessarily memoryless channel, we calculate the uncertainty estimate

$$\hat{U}_{\mathrm{c}} = \frac{1}{n} \sum_{i=1}^{n} \left[-\log_2 \frac{q(x_i, y_i)}{\sum_{a \in \mathcal{X}} q(a, y_i)} \right]. \tag{11.3}$$

From the uncertainty estimate, we calculate various achievable rates:

- Achievable code rate for sequence x^n:

$$\hat{T}_{\mathrm{c}} = \log_2 |\mathcal{X}| - \hat{U}_{\mathrm{c}}. \tag{11.4}$$

- Achievable rate for ideal layered PS:

$$\hat{R} = \left[\mathbb{H}(P_X) - \hat{U}_{\mathrm{c}}\right]^+ \tag{11.5}$$

  where P_X is the type of the test sequence x^n.

- Achievable rate

$$\hat{R} = \left[\mathsf{R}_{\mathrm{dm}} - \hat{U}_{\mathrm{c}}\right]^+ \tag{11.6}$$

  where R_{dm} quantifies the size of the shaping set from which the test sequence x^n is chosen.

This is only meaningful, if the test sequence x^n is representative for the shaping set, in the sense that picking another test sequence from the same shaping set leads to a similar uncertainty estimate. We discuss the choice of a representative test sequence in the next section.

11.2. Estimating Uncertainty

Suppose we can transmit one input sequence x^n over a channel and measure the resulting output sequence y^n. Suppose further that we lack an analytical description of the channel, in particular, the channel may have memory. Our task is to estimate the channel uncertainty when type P_X code words are transmitted. The channel uncertainty can then be used as a benchmark for code design. The actual block length of the designed code may be longer or shorter than the length n of our measurement. This means that to a large extend, our estimate should not depend on n. Furthermore, the estimate should be representative for all sequences in the shaping set. The following three tests may be used to detect possible dependencies:

T1. How does the estimate $\hat{U}_{\mathrm{c}}(x_1^j, y_1^j, q)$, $j \leq n$ depend on the length j?

T2. How does the estimate $\hat{U}_{\mathrm{c}}(x_{j-w+1}^j, y_{j-w+1}^j, q)$, $w \leq j \leq n$ depend on window position j and window size w?

T3. How do the outcomes of tests T1 and T2 depend on the test sequence x^n?

To gain an intuition for good test sequences, we next consider T1, T2, and T3 for memoryless AWGN.

11.2.1. High and Low SNR

Consider an AWGN channel with an ASK input alphabet $\mathcal{X}$, a test sequence $x^n \in \mathcal{X}^n$ and the decoding metric

$$q(x,y) = p_{Y|X}(y|x)P_X(x). \tag{11.7}$$

Low SNR

Suppose the SNR is very low. This means that $p_{Y|X}(y|x_1)/p_{Y|X}(y|x_2) \approx 1$ for any pair $x_1, x_2 \in \mathcal{X}$ and $y \in \mathbf{R}$, so that the ith sample contributes to the uncertainty estimate by

$$\begin{aligned}
-\log_2\left[\frac{p_{Y|X}(y_i|x_i)P_X(x_i)}{\sum_{a\in\mathcal{X}} p_{Y|X}(y_i|a)P_X(a)}\right] &= -\log_2\left[\frac{P_X(x_i)}{\sum_{a\in\mathcal{X}}\frac{p_{Y|X}(y_i|a)}{p_{Y|X}(y_i|x_i)}P_X(a)}\right] && (11.8)\\
&\xrightarrow{\mathsf{snr}\to 0} -\log_2\left[\frac{P_X(x_i)}{\sum_{a\in\mathcal{X}} P_X(a)}\right] && (11.9)\\
&= -\log_2 P_X(x_i). && (11.10)
\end{aligned}$$

See Problem 11.1. The limit (11.10) implies the following:

- If P_X is uniform on $\mathcal{X}$ then $P_X(x_i) = 1/|\mathcal{X}|$ for all $i = 1, 2, \ldots, n$ and there is no dependency on the number of samples and window size and position in tests T1. and T2., respectively. The uncertainty estimate evaluates to $\log_2|\mathcal{X}|$.
- If P_X is non-uniform, we may observe strong dependencies in tests T1. and T2. See the next example.

Example 11.1 (8-ASK at −10 dB). We consider the 8-ASK constellation

$$\mathcal{X} = \{\pm 1, \pm 3, \pm 5, \pm 7\}. \tag{11.11}$$

Our test sequence x^n has length $n = 100\,000$ and type

$$\begin{aligned}
n\cdot P_X(7) &= n\cdot P_X(-7) = 2114\\
n\cdot P_X(5) &= n\cdot P_X(-5) = 7189\\
n\cdot P_X(3) &= n\cdot P_X(-3) = 16\,255\\
n\cdot P_X(1) &= n\cdot P_X(-1) = 24\,442.
\end{aligned} \tag{11.12}$$

The test sequence x^n is sorted by probabilities, i.e., the symbols appear in the order $-7, 7, -5, 5, -3, 3, -1, 1$, e.g., the first 2114 entries of x^n have the value -7 and the last 24 442 entries are equal to 1. We sample a noise sequence z^n at −10 dB SNR and calculate $y^n = x^n + z^n$.

In Figure 11.1, we display the results of tests T1 and T2 for window size $w = 10\,000$. The uncertainty estimate depends heavily on j, because it decides on how

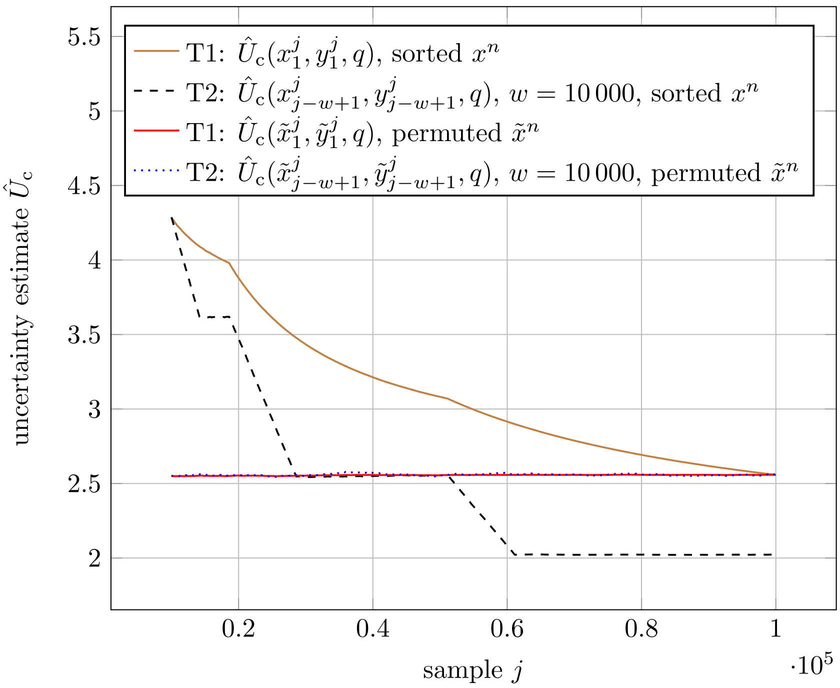


Figure 11.1.: Uncertainty estimates at -10 dB SNR. See Example 11.1 for explanations.

much the different symbols contribute to the estimate. We can use (11.10) to explain the T2 curve. At the right, the window covers only -1s and 1s. By (11.10), the uncertainty estimate is thus approximately

$$-\log_2 P_X(-1) = -\log_2 P_X(1) = -\log_2 \frac{24\,442}{100\,000} \approx 2.0326 \tag{11.13}$$

which corresponds very well to the displayed value.

Next, we generate a second test sequence $\tilde{x}^n$ by randomly permuting the entries of the sorted sequence x^n. Using the same noise as before, we calculate $\tilde{y}^n = \tilde{x}^n + z^n$ and repeat the tests. We observe in Figure 11.1 that for the permuted sequence $\tilde{x}^n$, there is almost no dependency on j. Note that for $j = 100\,000$, the T1 estimate of the sorted sequence coincides with the T1 and T2 estimates of the permuted sequence.

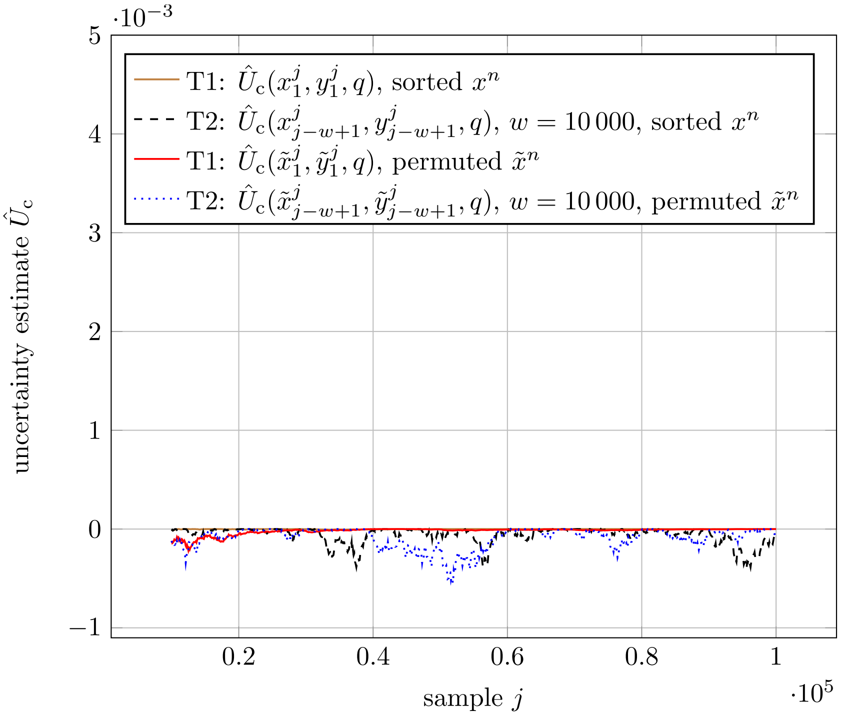


Figure 11.2.: Uncertainty estimates at 30 dB SNR. See Example 11.2 for explanations.

High SNR

In the high SNR limit, the ith sample x_i, y_i contributes to the uncertainty estimate by

$$-\log_2 \frac{p_{Y|X}(y_i|x_i)P_X(x_i)}{\sum_{a\in\mathcal{X}} p_{Y|X}(y_i|a)P_X(a)} = -\log_2 \frac{P_X(x_i)}{\sum_{a\in\mathcal{X}} \frac{p_{Y|X}(y_i|a)}{p_{Y|X}(y_i|x_i)} P_X(a)} \tag{11.14}$$

$$= -\log_2 \frac{P_X(x_i)}{P_X(x_i) + \sum_{a\in\mathcal{X}\setminus x_i} \frac{p_{Y|X}(y_i|a)}{p_{Y|X}(y_i|x_i)} P_X(a)} \tag{11.15}$$

$$\xrightarrow{\mathsf{snr}\to\infty} -\log_2 \frac{P_X(x_i)}{P_X(x_i)} \tag{11.16}$$

$$= 0. \tag{11.17}$$

See also Problem 11.1. Thus, for high SNR, the AWGN channel uncertainty is 0 independent of the test sequence. This means that in test T1., we do not see any dependency of the uncertainty estimate on the number of samples, and in test T2., we see no dependency on the window size and no dependency on the window position.

> **Example 11.2** (Example 11.1 continued)**.** We consider the same setup as in Example 11.1, now with a noise sequence sampled at 30 dB SNR. We display the resulting uncertainty estimates in Figure 11.2. As predicted by (11.17), for both

test sequences, the uncertainty estimates are close to zero, independent of j.

11.3. Estimating Uncertainty with Good Test Sequences

Our aim is to develop good test sequences that perform well under tests T1–T3, without the issues we observed in Example 11.1. We start with two examples illustrating the concept of good sequences.

Example 11.3. The sequence $x_1^8 = 00001111$ has type $P_{x_1^8}(0) = P_{x_1^8}(1) = 1/2$ while the sub-sequence x_1^4 has type $P_{x_1^4}(0) = 1 - P_{x_1^4}(1) = 1$. We therefore say that x_1^8 is a *bad sequence.*

Example 11.4. Suppose we have $P_X(0) = 1 - P_X(1) = 1/3$ and we estimate $\mathbb{H}(P_X)$ from a sequence x^n by

$$\hat{H}_\ell^j = \frac{1}{j-\ell+1}\sum_{i=\ell}^{j}[-\log_2 P_X(x_i)]. \tag{11.18}$$

For $x_1^9 = 000111111$, we get

$$\hat{H}_1^9 = \mathbb{H}(P_X) \approx 0.9183. \tag{11.19}$$

and

$$\hat{H}_1^3 = \log_2 3 \approx 1.5850 > \mathbb{H}(P_X) \tag{11.20}$$

$$\hat{H}_4^6 = \hat{H}_7^9 = \log_2 \frac{3}{2} \approx 0.5850 < \mathbb{H}(P_X) \tag{11.21}$$

and we conclude that x_1^9 is a bad sequence. Using $\tilde{x}_1^9 = 001100010$ yields

$$\hat{H}_1^3 = \hat{H}_4^6 = \hat{H}_7^9 = \mathbb{H}(P_X) \tag{11.22}$$

and we therefore call $\tilde{x}_1^9$ a *good sequence.* We have achieved this by putting a restriction on $\tilde{x}^n$: we divide $\tilde{x}^n$ into three chunks, each of three bits, and require that each chunk is of type P_X. This makes our estimates less dependent on the considered sub-sequence.

Note that we reduce the number of strings that we can choose from. For the bad sequence, we have $\binom{9}{3} = 84$ options, while for the good sequence, we have only $\binom{3}{1}^3 = 27$ options. This restriction will reduce the achievable rate that we can attach to uncertainty estimates.

We define a good test sequence as follows:

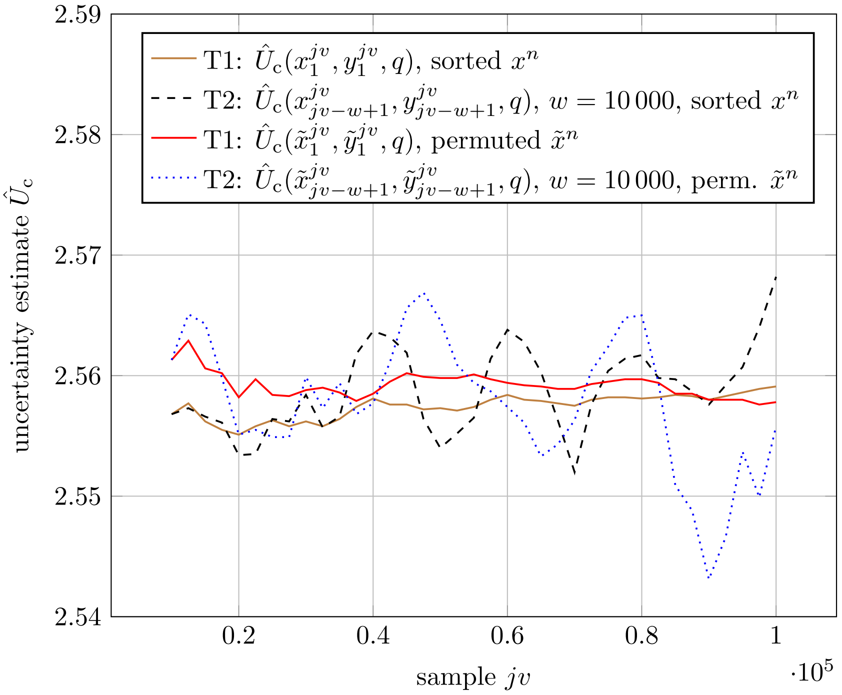


Figure 11.3.: Uncertainty estimates at $-10\,\mathrm{dB}$ SNR using good test sequences. See Example 11.5 for explanations. Figure 11.1 shows the corresponding estimates for bad sequences.

- $x^n = x_1^v x_{v+1}^{2v} \cdots x_{n-v+1}^n$ where
  - Each chunk $x_{(j-1)v+1}^{jv}$, $j = 1, 2, \ldots, n/v$ is of type P_X.
  - The chunk length v is large enough so that the rate loss $\mathbb{H}(P_X) - \frac{\log_2 |\mathcal{T}^v(P_X)|}{v}$ is small (see Section 9.2.1 for a detailed discussion of rate loss).

To check if the estimate is unlikely to apply for the whole shaping set, we perform the following tests.

T1. Calculate and plot $\hat{U}_{\mathrm{c}}(x_1^{jv}, y_1^{jv}, q)$, $j = 1, 2, \ldots, n/v$.

T2. Calculate and plot $\hat{U}_{\mathrm{c}}(x_{jv-w+1}^{jv}, y_{jv-w+1}^{jv}, q)$, $w/v \leq j \leq n/v$ and w an integer multiple of v.

T3. Do T1 and T2 for two different good test sequences.

Example 11.5 (Example 11.1 continued)**.** We choose $v = 2500$ and

$$\begin{aligned} w \cdot P_X(7) = w \cdot P_X(-7) &= 54 \\ w \cdot P_X(5) = w \cdot P_X(-5) &= 180 \\ w \cdot P_X(3) = w \cdot P_X(-3) &= 406 \\ w \cdot P_X(1) = w \cdot P_X(-1) &= 611. \end{aligned} \tag{11.23}$$

Note that (11.23) quantizes (11.12), see Section 9.3 for details. We use the same window size $w = 10\,000$ as in examples 11.1 and 11.2. For each $j = w/v, \dots, n/v$, the subsequence x_{jv-w+1}^{jv} covers $w/v = 4$ chunks, each of type P_X. For the sorted sequence, we sort the symbols in each chunk by increasing probability as in Example 11.1. For the permuted sequence, we choose for each chunk an individual random permutation. The resulting uncertainty estimates are displayed in Figure 11.3. In comparison to Figure 11.1, we notice that using good sequences has drastically reduced the dependency of the estimates on j and the considered sequence (sorted vs. permuted).

11.4. Calculating Achievable Rate Estimates

The uncertainty U_c has the following interpretation: from the observation at the decoder, we can infer a subset $\hat{\mathcal{W}} \subset \mathcal{X}^n$ of size $|\hat{\mathcal{W}}| = 2^{nU_\text{c}}$ that very likely contains the sequence that was actually transmitted. Thus, the FEC code should partition $\mathcal{X}^n$ into disjoint partitions of size $|\hat{\mathcal{W}}| = 2^{nU_\text{c}}$ that each contain one code word. This corresponds to the achievable code rate

$$T_\text{c} = \log_2 \frac{|\mathcal{X}|^n}{2^{nU_\text{c}}} = \log_2 |\mathcal{X}| - U_\text{c}. \tag{11.24}$$

The transmitter encodes now into the shaping set containing concatenations of n/v chunks, each of type P_X. That is, only the fraction $|\mathcal{T}^v(P_X)|^{\frac{n}{v}}/|\mathcal{X}|^n$ of sequences is used. An achievable rate is therefore

$$R_\text{tx} = \left[\log_2 \left(\frac{|\mathcal{T}^v(P_X)|^{\frac{n}{v}}}{|\mathcal{X}|^n} \frac{|\mathcal{X}|^n}{2^{nU_\text{c}}}\right)\right]^+ = \left[\frac{\log_2 |\mathcal{T}^v(P_X)|}{v} - U_\text{c}\right]^+. \tag{11.25}$$

A similar derivation for PAS achievable rates is discussed Problem 11.2.

11.5. Problems

Problem 11.1. Consider an 8-ASK constellation $\mathcal{X} = \{\pm1, \pm3, \pm5, \pm7\}$. Suppose the input power is fixed to P, the noise variance is σ^2 so that the SNR is $\mathsf{snr} = \mathsf{P}/\sigma^2$. The conditional output distribution of an AWGN channel is

$$p_{Y|X}(y|a) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y-a)^2}{2\sigma^2}\right], \quad a \in \mathcal{X}. \tag{11.26}$$

1. Show the limit in (11.17) explicitly by considering the limit $\sigma^2 \to 0$.

2. Show the limit in (11.10) explicitly by considering the limit $\sigma^2 \to \infty$.

Problem 11.2. Derive an achievable rate for PAS following the procedure described in Section 11.4. Assume the set of permissible sequences is the set of (amplitude,sign) pairs $\mathcal{T}^n(P_A) \times \{-1, 1\}^n$ where P_A is a distribution on ASK amplitudes.

A. Mathematical Background

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

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

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

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

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

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

B. Probability

- **Probability density function** (pdf) p_X:

$$\Pr(X \leq x) = \int_{-\infty}^{x} p_X(\tau) d\tau. \tag{B.1}$$

- **Bayes' Rule:**

$$p_{XY}(xy) = p_X(x) p_{Y|X}(y|x). \tag{B.2}$$

- **Independence:**

$$X \text{ and } Y \text{ are independent } \Leftrightarrow p_{XY}(xy) = p_X(x) p_Y(y), \quad \forall x, y. \tag{B.3}$$

- **Expectation for a real-valued function f:**

$$\mathbb{E}[f(X)] = \int_{-\infty}^{\infty} p_X(x) f(x) dx. \tag{B.4}$$

- **Law of total probability:** Let $\mathcal{A}$ be some event and let X be a random variable.

$$\Pr(\mathcal{A}) = \mathbb{E}[\Pr(\mathcal{A}|X)] \tag{B.5}$$

  where $\Pr(\mathcal{A}|X)$ is a random variable with realizations $\Pr(\mathcal{A}|X = x)$.

- **Markov's inequality, [19, Section 1.6.1]:** Let X be a non-negative random variable, i.e., $\Pr(X < 0) = 0$. Then for $a > 0$

$$\Pr(X \geq a) \leq \frac{\mathbb{E}(X)}{a}. \tag{B.6}$$

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

$$\mathsf{mgf}_X(r) = \mathbb{E}(e^{rX}) \tag{B.7}$$

$$\frac{\partial^k}{\partial r^k} \mathsf{mgf}_X(r) \bigg|_{r=0} = \mathbb{E}(X^k). \tag{B.8}$$

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

C. Information Theory

C.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.,

$$N(a|x^n) = \left|\left\{i \in \{1, 2, \ldots, n\} \colon x_i = a\right\}\right|, \quad a \in \mathcal{X}. \tag{C.1}$$

The empirical distribution of x^n is

$$P_{x^n}(a) = \frac{N(a|x^n)}{n}, \quad a \in \mathcal{X}. \tag{C.2}$$

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

$$P_X(a) = \frac{n_a}{n}, \quad a \in \mathcal{X}. \tag{C.3}$$

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 $\mathcal{T}^n(P_X)$.

Typical Sequences We use *letter-typical* sequences as defined in [22, Sec. 1.3]. Consider a distribution P_X on a finite alphabet $\mathcal{X}$. For $x^n \in \mathcal{X}^n$. We say x^n is ϵ-letter-typical with respect to P_X if for each letter $a \in \mathcal{X}$,

$$(1 - \epsilon)P_X(a) \leq \frac{N(a|x^n)}{n} \leq (1 + \epsilon)P_X(a), \quad \forall a \in \mathcal{X}. \tag{C.4}$$

Let $\mathcal{T}_\epsilon^n(P_X)$ be the set of all sequences x^n that fulfill (C.4). The sequences (C.4) are called *typical* in [33, Sec. 3.3],[34, Sec. 2.4] and *robust typical* in [23, Appendix].

C.2. Differential Entropy

- **Differential entropy:**

$$\mathrm{h}(X) := \mathbb{E}[-\log_2 p_X(X)]. \tag{C.5}$$

- **Scaling for real-valued X:**
$$\mathrm{h}(aX) = \mathrm{h}(X) + \log_2 |a|. \tag{C.6}$$

- **Translation:**
$$\mathrm{h}(X + b) = \mathrm{h}(X). \tag{C.7}$$

- **Conditional differential entropy:**
$$\mathrm{h}(Y|X) := \mathbb{E}[-\log_2 p_{Y|X}(Y|X)]. \tag{C.8}$$

- **Function of variables:**
$$\mathrm{h}\left[Y + f(X)\middle|X\right] = \mathrm{h}(Y|X). \tag{C.9}$$

- **Chain rule:**
$$\mathrm{h}(X, Y) = \mathrm{h}(X) + \mathrm{h}(Y|X) = \mathrm{h}(Y) + \mathrm{h}(X|Y). \tag{C.10}$$

- **Conditioning does not increase entropy:**
$$\mathrm{h}(X) \geq \mathrm{h}(X|Y). \tag{C.11}$$

- **Independence bound:**
$$\mathrm{h}(X, Y) \leq \mathrm{h}(X) + \mathrm{h}(Y). \tag{C.12}$$

C.3. Entropy

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

- **Entropy:**
$$\mathbb{H}(P_X) = \mathbb{H}(X) := \mathbb{E}[-\log_2 P_X(X)]. \tag{C.13}$$

- **Conditional Entropy:**
$$\mathbb{H}(P_{X|Y}|P_Y) = \mathbb{H}(X|Y) := \mathbb{E}[-\log_2 P_{X|Y}(X|Y)]. \tag{C.14}$$

- **Conditioning does not increase entropy:** We have
$$\mathbb{H}(X) \geq \mathbb{H}(X|Y) \tag{C.15}$$
with equality if and only if X and Y are stochastically independent.

- **Function of variables:** Let f, g be functions. Then
$$\mathbb{H}(X|Y) = \mathbb{H}(X, f(X,Y)|Y, g(Y)). \tag{C.16}$$

- **Relation to differential entropy:** Properties (C.10), (C.11), and (C.12) also hold for entropy.

- **Binary entropy function:** $0 \leq p \leq 1$.
$$\mathbb{H}_2(p) := -p \log_2 p - (1-p)\log_2(1-p). \tag{C.17}$$

- **Fano's Inequality:** Random variable X with distribution P_X on $\mathcal{X}$. Decision $\hat{X}$, joint distribution $P_{X\hat{X}}$, error probability $P_e = \Pr(X \neq \hat{X})$.
$$\mathbb{H}_2(P_e) + P_e \log_2(|\mathcal{X}| - 1) \geq \mathbb{H}(X|\hat{X}). \tag{C.18}$$

- **Continuity:** Distributions P_X, $P_{X'}$ on finite set $\mathcal{X}$. Suppose $\|P_X - P_{X'}\|_1 = \delta \leq \frac{1}{2}$. Then
$$|\,\mathbb{H}(P_X) - \mathbb{H}(P_{X'})| \leq -\delta \log_2 \frac{\delta}{|\mathcal{X}|}. \tag{C.19}$$

C.4. Informational Divergence

- **Informational divergence:**
$$\mathbb{D}(p_X \| p_Y) := \mathbb{E}\left[\log_2 \frac{p_X(X)}{p_Y(X)}\right] \tag{C.20}$$

- **Information inequality:**
$$\mathbb{D}(p_X \| p_Y) \geq 0 \tag{C.21}$$
with equality if and only if $p_X = p_Y$.

- **Asymmetric:** in general, we have
$$\mathbb{D}(p_X \| p_Y) \neq \mathbb{D}(p_Y \| p_X). \tag{C.22}$$

- **Conditional informational divergence:**
$$\mathbb{D}(p_{Y_1|X_1} \| p_{Y_2|X_2} | p_{X_1}) := \mathbb{E}\left[\log_2 \frac{p_{Y_1|X_1}(Y_1|X_1)}{p_{Y_2|X_2}(Y_1|X_1)}\right]. \tag{C.23}$$

- **Chain rule:**
$$\mathbb{D}(p_{X_1Y_1} \| p_{X_2Y_2}) = \mathbb{D}(p_{X_1} \| p_{X_2}) + \mathbb{D}(p_{Y_1|X_1} \| p_{Y_2|X_2} | p_{X_1}) \tag{C.24}$$
$$= \mathbb{D}(p_{Y_1} \| p_{Y_2}) + \mathbb{D}(p_{X_1|Y_1} \| p_{X_2|Y_2} | p_{Y_1}). \tag{C.25}$$

- **Discrete random variables:** All properties of the informational divergence stated above also hold for discrete random variables; replace densities, e.g., p_X, by distributions, e.g., P_X. The properties also hold for mixed random variables, e.g., if $X = X_c X_d \sim p_{X_c} P_{X_d}$, replace p_X by $p_{X_c} P_{X_d}$.

C.5. Mutual Information

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

$$\begin{aligned} \mathbb{I}(X;Y) &:= \mathbb{D}(p_{XY} \| p_X p_Y) \qquad \text{(C.26)}\\ &= \mathbb{D}(p_{Y|X} \| p_Y | p_X) \qquad \text{(C.27)}\\ &= \mathbb{D}(p_{X|Y} \| p_X | p_Y) \qquad \text{(C.28)}\\ &= \mathrm{h}(Y) - \mathrm{h}(Y|X) \qquad \text{(C.29)}\\ &= \mathrm{h}(X) - \mathrm{h}(X|Y). \qquad \text{(C.30)} \end{aligned}$$

  - X discrete, Y continuous:

$$\begin{aligned} \mathbb{I}(X;Y) &:= \mathbb{D}(P_X p_{Y|X} \| P_X p_Y) \qquad \text{(C.31)}\\ &= \mathbb{D}(P_{X|Y} \| P_X | p_Y) \qquad \text{(C.32)}\\ &= \mathbb{D}(p_{Y|X} \| p_Y | P_X) \qquad \text{(C.33)}\\ &= \mathrm{h}(Y) - \mathrm{h}(Y|X) \qquad \text{(C.34)}\\ &= \mathbb{H}(X) - \mathbb{H}(X|Y). \qquad \text{(C.35)} \end{aligned}$$

  - Other combinations of discrete/continuous accordingly.

- **Independence Test:**

$$X \text{ and } Y \text{ are independent } \Leftrightarrow \mathbb{I}(X;Y) = 0. \tag{C.36}$$

- **Non-Negative:**

$$\mathbb{I}(X;Y) \geq 0. \tag{C.37}$$

- X, Y, Z form a **Markov chain** if X and Z are independent conditioned on Y, i.e.,

$$X \multimap\!\!- Y \multimap\!\!- Z \Leftrightarrow \mathbb{I}(X;Z|Y) = 0. \tag{C.38}$$

- If $X = f(Y)$, then $X \multimap\!\!- Y \multimap\!\!- Z$

- **Chain Rule:**

$$\mathbb{I}(XZ;Y) = \mathbb{I}(Z;Y) + \mathbb{I}(X;Y|Z). \tag{C.39}$$

- **Function of variables:** Let f, g, r be functions. Then

$$\mathbb{I}(X, f(X,Z); Y, g(Y,Z) | Z, r(Z)) = \mathbb{I}(X;Y|Z). \tag{C.40}$$

- **Data Processing Inequality:** Suppose $X \multimap\!\!- Y \multimap\!\!- Z$. Then

$$\begin{aligned} \mathbb{I}(X;Y) &\geq \mathbb{I}(X;Z)\\ \mathbb{I}(Y;Z) &\geq \mathbb{I}(X;Z). \end{aligned} \tag{C.41}$$

Bibliography

[1] J. L. Massey, "Coding and modulation in digital communications," in *Proc. Int. Zurich Seminar Commun.*, 1974.

[2] R. G. Gallager, *Principles of Digital Communication.* Cambridge University Press, 2008.

[3] ——, *Information Theory and Reliable Communication.* John Wiley & Sons, Inc., 1968.

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

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

[6] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed. Prentice Hall, 2010.

[7] A. Lapidoth, *A Foundation in Digital Communication*, 1st ed. Cambridge University Press, 2009.

[8] U. G. Schuster and H. Bölcskei, "Ultrawideband channel modeling on the basis of information-theoretic criteria," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2464–2475, Jul. 2007.

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

[10] A. L. McKellips, "Simple tight bounds on capacity for the peak-limited discrete-time channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Chicago, IL, USA, Jun. 2004, pp. 348–348.

[11] A. Thangaraj, G. Kramer, and G. Böcherer, "Capacity bounds for discrete-time, amplitude-constrained, additive white Gaussian noise channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4172–4182, Jul. 2017.

[12] B. Rassouli and B. Clerckx, "An upper bound for the capacity of amplitude-constrained scalar AWGN channel," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 1924–1926, Oct. 2016.

[13] L. Ozarow and A. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Nov. 1990.

[14] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 14–20, Jan. 1972.

[15] R. Blahut, “Computation of channel capacity and rate-distortion functions,” *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460–473, Jul. 1972.

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

[17] E. A. Ratzer, “Error-correction on non-standard communication channels,” PhD thesis, University of Cambridge, 2003.

[18] G. Böcherer, “Capacity-achieving probabilistic shaping for noisy and noiseless channels,” PhD thesis, RWTH Aachen University, 2012. [Online]. Available: http://www.georg-boecherer.de/capacityAchievingShaping.pdf.

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

[20] 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, pp. 228–239, 1993.

[21] 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, pp. 2315–2328, Nov. 2000.

[22] G. Kramer, “Topics in multi-user information theory,” *Foundations and Trends in Comm. and Inf. Theory*, vol. 4, no. 4–5, pp. 265–444, 2007.

[23] A. Orlitsky and J. R. Roche, “Coding for computing,” *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

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

[25] 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, pp. 4651–4665, Dec. 2015.

[26] G. Böcherer, “Achievable rates for shaped bit-metric decoding,” *arXiv preprint*, 2016. [Online]. Available: http://arxiv.org/abs/1410.8075.

[27] 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, pp. 2756–2765, Jun. 2009.

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

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

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

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

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

[33] J. L. Massey, "Applied digital information theory I," lecture notes, ETH Zurich, [Online]. Available: `http://www.isiweb.ee.ethz.ch/archive/massey_scr/adit1.pdf`.

[34] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

Acronyms

ASK amplitude shift keying

AWGN additive white Gaussian noise

BICM bit-interleaved coded modulation

BMD bit-metric decoding

BPSK binary phase shift keying

CCDM constant composition distribution matcher

DFT discrete Fourier transform

DM distribution matcher

DMS discrete memoryless source

FEC forward error correction

GMI generalized mutual information

iid independent and identically distributed

LDPC low-density parity-check

LUT look-up table

MB Maxwell-Boltzmann

MGF moment generating function

PAS probabilistic amplitude shaping

pdf probability density function

PS probabilistic shaping

QAM quadrature amplitude modulation

SNR signal-to-noise ratio

WLLN weak law of large numbers

Index