University of Alberta

Extremal Problems of Error Exponents and Capacity of Duplication Channels

by

Mahdi Ramezani

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Communications

Department of Electrical and Computer Engineering

Spring 2013
Edmonton, Alberta

To Niloufar and my family ...

Abstract

One of the most stunning results of information theory is the channel coding theorem addressing the maximum rate of reliable communication over a noisy channel, known as channel capacity. In this thesis, we consider two problems emerging from the classic channel coding theorem. First, we study the extremal problems of the channel reliability function, which is the exponent with which the probability of making a wrong decision vanishes. To this end, we introduce a set of fundamental channels which exhibit significant monotonicity properties and invoke the theory of Chebychev systems to utilize such properties. We show that the binary symmetric channel (BSC) and binary erasure channel (BEC), which happen to be among the fundamental channels, are the two extremes of the channel reliability function. Also, we show that given a rate and a probability of error as a performance measure, BSC (BEC) needs the longest (shortest) code length to achieve such performance.

While the first problem is pure theoretical, the second problem addresses a challenging practical scenario. The most fundamental assumption in the classic channel coding theorem is that we receive as many symbols as we send. In reality, however, this is not always true, e.g., a miss-sampling at a conventional receiver might duplicate a symbol. The extra symbol confuses a receiver as it has no clue about the position of duplication. Such scenarios are collectively known as channels with synchronization errors. Unlike their classic counterparts, there is only little known about either the capacity or coding techniques for channels with synchronization errors, even in their simplest forms. In this part, we study the duplication channel by introducing a series expansion for its capacity.

Acknowledgements

"It is good to have an end to a journey; but it is the journey that matters in the end." Here is why I enjoyed my journey: It was my great pleasure to work and collaborate with Masoud during my Ph.D. and Masters studies. I feel very fortunate to be among those whose supervisor is not just an advisor, but also a friend and a caring teacher. His open approach towards research and level of professionalism have created a comforting atmosphere for everyone working with him. I would like to sincerely thank him for being very supportive before and after I moved to Toronto through our weekly cyber meetings, for encouraging me to take courses from other disciplines, and for funding my research during the last steps of my Ph.D. studies.

I would like to thank my dissertation committee, Chintha Tellambura, Hai Jiang, Mike Kouritzin and Fady Alajaji for carefully reading my thesis and providing invaluable feedback.

I give special thanks to my teachers from the Department of Mathematical and Statistical Sciences at the University of Alberta: Byron Schmuland, Alexander Litvak, Tahir Choulli, Mike Kouritzin and Alexander Melnikov. Also, I greatly appreciate the help of Richard Zalik from Auburn University with the theory of Chebychev systems.

My doctoral studies have been supported financially by generous funding from the Alberta Ingenuity (now Alberta Innovates) and iCORE. I truly appreciate their support.

I have immensely benefitted from the discussions that I had with my brilliant friends/colleagues: Raman, Payam, Arash, Moslem, Sahar and Kaveh. I also enjoyed the company of my fellow officemates: Gayan, Zohreh, Damith, Hossein, Madushanka, Wei and Chao.

Despite all the ups and downs, I thoroughly enjoyed my journey because of having amazing friends: Amirhossein, Arash, Saina, Raman, Payam, Moslem, Marjan, Hamid, Kaveh, Sahar, Mahdi, Zohreh and Mahmoud.

My stay in Toronto has been incredible mainly due to my in-law family. I am greatly indebted to my parents-in-law who literally treated me like their own son. I enjoyed our nightly lifesaving card games that helped me to free my mind from research problems. Also, I would like to thank the rest of my in-law family for creating warm and intimate family moments that I have been missing since I left my family in Iran.

My family is the source of my strength. Although I live far away, I always have them in my heart. They were very supportive in spite of the fact that I had to leave them.

And of course, the biggest joy of my journey was meeting my future wife. I am truly grateful for having my caring and understanding wife, Niloufar, who believed in me and stood by my side through my rough times. I wish one day I could return her dedication and support. This thesis is especially dedicated to her.

Table of Contents

List of Tables

List of Figures

List of Symbols

Symbol	Definition	First Use
$\mathcal{A}$	Non-empty set	1
$\mathbf{c}$	Vector	2
C	Channel capacity	3
R	Code rate	3
P_e	Probability of decoding error	3
$E(\cdot)$	Channel reliability function	3
R_{cr}	Critical rate of channel	3
X^n	Vector-valued random variable	7
$\mathbb{P}$	Generic probability measure	7
p	Duplication probability	7
$C(p)$	Duplication capacity at duplication probability p	7
$P_{Y\|X}$	Channel's conditional probability measure	9
$\bar{\mathbb{R}}$	Set of extended real numbers	9
$\mathcal{A}^n$	n-time Cartesian product of set $\mathcal{A}$	9
P_X	A probability assignment on outcomes of X	9
BSC(ϵ)	A BSC with crossover probability ϵ	10
BEC(ϵ)	A BEC with erasure probability ϵ	10
$\mathbf{G}$, $\mathbf{W}$	Channel transition matrix	10
$\bar{\epsilon}$	$1-\epsilon$ for $\epsilon \in [0,1]$	11
a	L-density	11
A	L-distribution	11
$\mathsf{H}_t(\cdot)$	Heavyside distribution with a jump at t	11
$\mathbf{1}_{\mathcal{A}}$	The indicator function of set $\mathcal{A}$	11
$\Delta_t(\cdot)$	Mass point at t	11
$\mathsf{a}_{\mathrm{BSC}(\epsilon)}$	L-density of a BSC(ϵ)	12
$\mathsf{a}_{\mathrm{BEC}(\epsilon)}$	L-density of a BEC(ϵ)	12
$\mathbb{R}$	Set of real numbers	12
$\|\mathsf{a}\|$	$\|L\|$-density	13
g	P-density	14
$\mathbb{E}$	Expected value	15

List of Abbreviations

Abbreviation	Definition	First Use
BSC	Binary symmetric channel	2
BEC	Binary erasure channel	2
LLR	Log-likelihood ratio	9
RHS	Right hand side	9
LHS	Left hand side	9
BIAWGN	Binary-input additive white Gaussian noise	10
DMC	Discrete memoryless channel	10
MBIOS	Memoryless binary-input output-symmetric	11
T–System	Chebychev system	16
CT–System	Complete Chebychev system	16
ET–System	Extended Chebychev system	18
ECT–System	Extended complete Chebychev system	18
DCT	Dominated convergence theorem	22
MMSE	Minimum mean-squared error	28
SFB	Shulman-Feder bound	40
i.i.d.	Independent and identically distributed	60
TV	Total variation	63
LDPC	Low-density parity-check	92

Chapter 1

Introduction

The information era has begun by the seminal work of Claude Shannon [1] through which he introduced a set of mathematical tools, known as *information theory*, to model data transmission. Communication is subject to channel's imperfection. The effect of channel can be in the form of distorting signals, corrupting a stream of bits, etc. Since then, the attempt of channel coding theory has been finding practical schemes to battle the channel's noise. In today's applications, there is hardly one without a channel coding technique embedded in it. Talking on a mobile phone, streaming video content over the internet or even watching a Blu-ray disk are among everyday activities that enjoy the power of channel coding. Information theory, by finding the fundamental limits and extremes of possibilities in channel coding, sheds light to design of more efficient solutions. This thesis studies some of these fundamental limits and extreme cases.

1.1 The Classic Model of Communication

The Shannon's model of communication is:

1. A priori-unknown *message* is selected randomly from the set $\{1, \ldots, M\}$.
2. A *noisy channel* takes an input symbol from the input alphabet $\mathcal{X}$, applies a random transformation and puts out a symbol from the output alphabet $\mathcal{Y}$.

3. An *encoder* maps a message from the message set to a word of channel's input alphabets, i.e., $e : \{1, \ldots, M\} \mapsto \mathcal{X}^n$. We call $e(i) = \mathbf{c}_i \in \mathcal{X}^n$ the codeword of length n associated with the message i. The set of codewords $\{\mathbf{c}_1, \ldots, \mathbf{c}_M\}$ is called a code or codebook of length n. Since messages are chosen uniformly random, the required number of bits per channel use to describe the code denotes the *code rate* and equals $R = \frac{1}{n} \log_2 M$ bits.

4. A *decoder* tries to recover the original sent message from the channel outputs, i.e., $d : \mathcal{Y}^n \mapsto \{1, \ldots, M\}$. Note that the decoder is provided with the codebook that the encoder uses to encode messages at the transmitter side. The ultimate goal is to minimize probability of making decoding errors while maximizing the code rate (equivalently, maximizing M).

In this setup, it is assumed that we receive exactly as many symbols as we send over. This means that if we feed the channel with 100 symbols, we will receive 100 symbols at the receiver. Shannon showed that there exists a maximum code rate, called *channel capacity*, below which reliable communication is possible [2]. By reliable communication we mean that the probability of making a wrong decision vanishes as n grows.

Consider the binary symmetric channel (BSC) where $\mathcal{X} = \mathcal{Y} = \{0, 1\}$. A BSC flips every input bit with probability p known as the crossover probability. BSC is used to model hard-decision decoding scenarios. According to Shannon, the maximum rate of reliable communication over BSC is $1 + p \log_2 p + (1 - p) \log_2(1 - p)$ bits per channel use. Fig. 1.1 presents a graphical sketch of BSC. The binary erasure channel (BEC) is another channel of our interest where $\mathcal{X} = \{0, 1\}$ and $\mathcal{Y} = \{0, e, 1\}$. It is used to model scenarios where the decoder has the option of refusing to make a decision by putting out an erasure symbol. BEC is also shown in Fig. 1.1. As it can be seen, the decoder declares an erasure (e) with probability ϵ. The capacity of a BEC is $1 - \epsilon$. In the rest of this chapter, we will see several properties of BSC and BEC. We observe that although these channels look very simple, they exhibit fundamental properties.

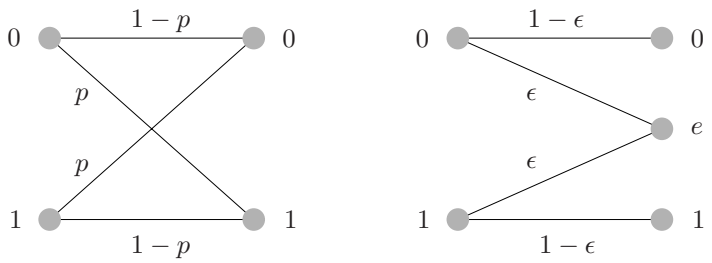

Figure 1.1: Left: A BSC with crossover probability p. Right: A BEC with erasure probability ϵ.

1.2 Thesis Overview

The contents of the current thesis revolves around studying the simplest, yet most *fundamental*, noisy channels. In this section, we briefly overview two scenarios where these fundamental channels appear. We keep this overview simple and leave the details to future chapters.

1.2.1 Extremal Problems of Error Exponents

Shannon showed that reliable communication is possible at any rate below the capacity, denoted by C. This means that using random coding, probability of making a wrong decision vanishes as $n \to \infty$. However, it is not clear *how fast the probability of error goes to zero*. To elaborate, let us denote a code with three major parameters: (n, R, P_e) representing the code length, code rate and maximum probability of decoding error. Shannon result shows that for any rate less than the capacity, there exists a random code (in fact a sequence of codes indexed by n) such that $P_e \to 0$ as $n \to \infty$. Define $P_e^*(n, R)$ as the least probability of error among all random codes of length n and rate R. It is desired to find a channel–dependent exponent $E(R)$ such that for $R \leq C$, $P_e^*(n, R) \approx 2^{-nE(R)}$. Clearly, this representation suggests that $P_e^*(n, R) \to 0$ when $R < C$. $E(R)$, called channel's *reliability function*, indicates how fast the probability of error vanishes at rate R [3]. Since reliable communication is impossible for rates greater than the capacity, we have $E(R) = 0,\ R \geq C$.

Unfortunately, the reliability function is only known for $R_{cr} \leq R \leq C$ where R_{cr} is the critical rate which is a function of the channel. There are

several lower and upper bounds known for the reliability function, mostly for the range of $0 \leq R < R_{cr}$. Given a set of channels, we would like to find out which channel has the maximum (minimum) of reliability function. In other words, we are seeking a channel with the fastest (slowest) converging probability of error at a given rate. To make a fair comparison, we assume that all given channels have the same capacity.

In this thesis, we address the following questions:

1. Among all binary-input symmetric channels of the same capacity, which channel has the maximum (minimum) of the error exponent?

2. Among all binary-input symmetric channels of the same capacity, which channel does exhibit the maximum (minimum) of the critical rate?

3. How does one modify these extremes when an additional constraint is imposed, e.g., a constraint on the uncoded probability of error of the channel?

In a similar way, one can define $R^*(n, P_e)$ as the maximum rate of transmission that is possible using a random code of length n such that the original codeword can be recovered with probability at least $1-P_e$. According to Shannon, $R^*(n, P_e) \to C$ when $n \to \infty$ and $P_e \to 0$. A very interesting question is remained to be answered: *Which channel needs the highest (lowest) code length to achieve a desired rate/probability of error?*

To answer all these questions, we define a set of simple channels with equal capacity, which includes BSC and BEC, that spans the space of symmetric channels. For obvious reasons, we call it the set of *basis channels*. Then, we invoke the theory of Chebychev systems from approximation theory to study monotonicity properties among the basis channels. We show that extreme channels for several error exponents lie in the set of basis channels.

1.2.2 Duplication Channels

As it was mentioned before, in the classic model of communication, we suppose that the number of sent and received symbols exactly match. In other words,

we assume that there is a perfect *synchronization* between the transmitter and receiver. In modern communication systems, timing recovery methods are essential and used to overcome synchronization problems. In these systems, achieving perfect synchronization is not possible even with such recovery methods. In reality, there are several applications suffering from imperfect synchronization. For example, in a mobile communication system where the clock of the transmitter and receiver are not synchronized, a miss-sampling at the receiver may result in a deletion of a symbol which can hinder the overall performance. As the modern communication systems become faster and operate at lower signal-to-noise ratios the synchronization errors become more relevant. The situation gets particularly worse when a coding method is involved. In this case, a single uncorrected duplication/deletion error can result in a burst of errors. Synchronization errors happen more frequently in the magnetic recording systems (hard disk drive) where due to improper head movements, one unit of data might be read/written either twice or none during a read/write process [4–7]. When synchronization is not perfect, the decoder is not aware of the position of duplications/deletions. Channels with synchronizations can be found in other sciences, e.g., symbols from DNA and RNA sequences are deleted and duplicated in genetic processes. Studying channels with synchronization errors may give us important insight into these genetic processes [8, 9]. These applications demand a whole new theory regarding capacity analysis and coding methods. Unfortunately, unlike their "synchronized" counterparts, there is only little known about either the capacity or coding for channels with synchronization errors [10–27].

It is important to note that in case of duplications, the extra bits should not be mistaken with the redundancy that we add in coding methods. While the former is random and is caused by the channel, the latter is intentional and is used to add data protection.

Dobrushin [28] defines a channel with synchronization errors as a channel transforming every input symbol to a word (of possibly zero length). This means that symbols might be deleted from or inserted into the input sequence. The classic model does not capture this characteristic of channels with syn-

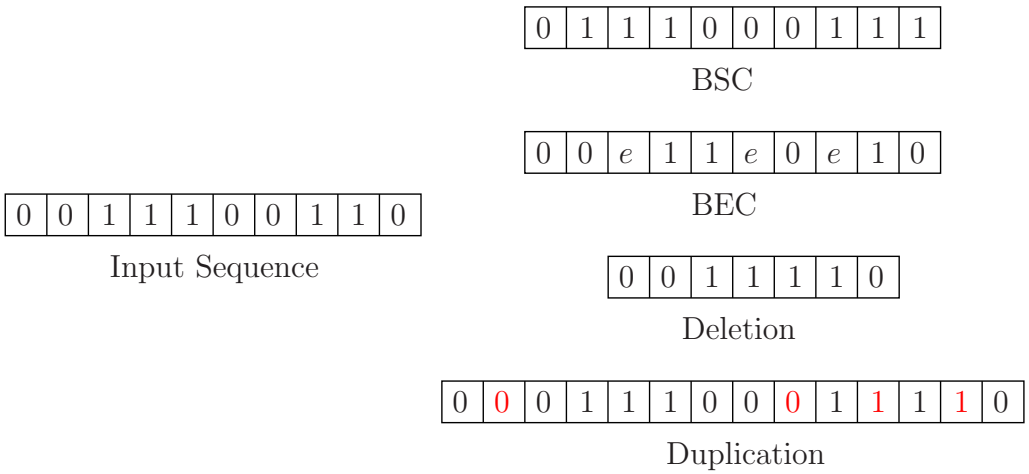

Figure 1.2: The input sequence on the left is fed to different channels. The corresponding output sequences are shown on the right. In the case of a duplication channel, the duplicated bits are shown in red.

chronization errors. The simplest form of such channels are the binary *deletion* and binary *duplication* channels, where each input bit is either transmitted intact or deleted (duplicated) with some probability.

Fig. 1.2 shows how binary symmetric, binary erasure, binary deletion and binary duplication channels act on an input sequence. It is important to note that in the case of BEC, when an erasure happens, the decoder declares an erasure and an erasure symbol is received. However, in a deletion channel, the actual bit is deleted and the decoder has no idea about the position of deleted bits. In the case of duplication channel, the decoder does not know which bit is duplicated unless for example, there is a single zero bordered by ones.

Two fundamental questions arise here:

1. How well can one communicate reliably over such channels? In other words, if it exists, what is the capacity?

2. Is there any efficient coding method?

In this thesis, we focus on the first problem. The interested reader may look at [8, 29] for literature context of coding methods for synchronization channels.

To see where the difficulty in finding the capacity of such channels comes

from, let us examine the ultimate goal in finding the capacity, i.e., analyzing the maximum likelihood decoding. Let $X^n = (X_1, \ldots, X_n)$ be the channel's input. We denote the received sequence corresponding to X^n by $Y(X^n)$. The reason behind such notation is that the length of received sequence is a random variable which can be greater than, equal to, or less than n. We have $\mathbb{P}(Y(X^n)|X^n) \propto \aleph(X^n, Y(X^n))$ where $\aleph(X^n, Y(X^n))$ denotes the number of insertion/deletion patterns that transform X^n into $Y(X^n)$. There we have a decoding algorithm: take the received sequence, count how many times it appears as a super/subsequence of each codeword, and output the codeword with the largest count [8]. This means that the decoding performance is limited by the efficiency of the counting process as n grows. It turns out that the analysis of $\aleph(X^n, Y(X^n))$ is very complicated and there is no efficient method to do so. Hence, most of the research done in this venue has been around studying subclasses of synchronization channels. Even for the simplest models, i.e., the binary deletion and duplication channels, the single letter characterization of capacity is unknown.

In this thesis, we study the single letter characterization of capacity of duplication channels. Let $C(p)$ denote the capacity of a binary duplication channel that duplicates each input bit with probability p. $C(p)$ is unknown except for $p = 0$ and $p = 1$ where $C(0) = C(1) = 1$ bit per channel use. We find a series expansion of $C(p)$ around $p = 0$ and show that some upper and lower bounds of $C(p)$ match up to a term of order p. Moreover, we will see that surprisingly $C(p)$ is not a symmetric function of p. This observation leeds to some interesting system design perspectives.

1.3 Thesis Organization

Chapter 2 covers the preliminary background material to understand the rest of this thesis. The set of equal-capacity basis channels and their properties are studied thoroughly in Chapter 3. The extremal problems of classic error exponents are solved in details in Chapter 4, where we use the theory of

Chebychev systems to solve such extremal problems[1]. Finally, in chapter 5, we study the capacity of duplication channels in terms of a series expansion for small duplication probabilities[2].

1.4 Notations Guide

$\mathbb{P}$ is reserved for a generic probability measure. The expected value and variance are designated by $\mathbb{E}$ and $\mathbb{V}$. A nonempty set is denoted by calligraphic letters $\mathcal{A}$, $\mathcal{B}$, ... A set difference, cardinality and an n-time Cartesian product of a set are shown by $\mathcal{A}\backslash\mathcal{B}$, $|\mathcal{A}|$, and $\mathcal{A}^n$ respectively. Small bold letters indicate vectors, while capital bold letters show matrices. log and $\log_2$ denote the natural logarithm and the logarithm in base 2. The binary entropy function in bits is shown by $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$ for $x \in [0,1]$. Also, $h^{-1} : [0, \frac{1}{2}] \mapsto [0, \frac{1}{2}]$ indicates the inverse binary entropy function. A mass point at $x = p$ is $\Delta_p(x)$. Also, for $\epsilon \in [0,1]$, $\bar{\epsilon} = 1-\epsilon$. The set of real numbers, non-negative real numbers, extended real numbers and natural numbers are designated by $\mathbb{R}$, $\mathbb{R}_+$, $\bar{\mathbb{R}}$ and $\mathbb{N}$. $\mathbf{1}_{\mathcal{A}}$ is the indicator function of set $\mathcal{A}$. The first, second and nth, $n \geq 3$ derivatives of a function f are presented by f', f'' and $f^{(n)}$, respectively. A BSC with crossover probability ϵ, $\epsilon \in [0,1]$, is shown by BSC(ϵ). Also, a BEC with erasure rate of ϵ, $\epsilon \in [0,1]$, is denoted by BEC(ϵ). $x_n \uparrow x^*$ ($x_n \downarrow x^*$) indicates a monotonically increasing (decreasing) sequence converging to x^*.

[1]The results of Chapter 3 and 4 have been submitted to IEEE Transactions on Information Theory.

[2]The results of Chapter 5 have been accepted for publication in IEEE Transactions on Communications.

Chapter 2

Preliminaries

In this chapter, we present preliminary background material used in Chapter 3 and 4. We leave the preliminaries for channels with synchronization errors to Chapter 5. If the reader decides to follow the results on duplication channels, they may fast forward to Chapter 5.

2.1 Symmetric Channels

Let X and Y denote the input and output of a binary-input channel. We denote the channel's conditional probability measure by $P_{Y|X} : \{-1, +1\} \mapsto \mathcal{Y}$, where $\mathcal{Y}$ is a subset of $\bar{\mathbb{R}}$. Let $X^n = (X_1, \ldots, X_n)$. We say that a channel is memoryless if each channel use is independent of other instances, i.e., for all $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$

$$P_{Y^n|X^n}(y^n|x^n) = \prod_{k=1}^{n} P_{Y_k|X_k}(y_k|x_k).$$

The log-likelihood ratio (LLR) of the channel output y is given by the function $l : \mathcal{Y} \mapsto \bar{\mathbb{R}}$ defined as

$$l(y) = \log \frac{P_{X|Y}(+1|y)}{P_{X|Y}(-1|y)} = \log \frac{P_{Y|X}(y|+1)}{P_{Y|X}(y|-1)} + \log \frac{P_X(+1)}{P_X(-1)}.$$

Strictly speaking, the first term on the right hand side (RHS) denotes the LLR. While the second term on the RHS is called a priori LLR, the left hand side (LHS) is called a posteriori LLR. In this thesis, we assume an equiprobable input, meaning that a posteriori LLR is equal to the LLR [30].

LLR values are extensively used when one is to estimate the reliability of a decision. For example, consider a binary-input additive white Gaussian (BIAWGN) channel: $Y = X + Z$ where X is uniformly distributed and Z is a Gaussian random variable with mean zero and variance σ^2. A hard decoder declares $x = +1$ when the received value is positive and declares $x = -1$ otherwise. However, it does not matter to a hard decoder whether $y > 0$ is very large or very small; it will decode the received value to $x = +1$ anyways. However, a soft decoder computes the LLR value as

$$l(y) = \log \frac{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-1)^2}{2\sigma^2}\right)}{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y+1)^2}{2\sigma^2}\right)} = \frac{2y}{\sigma^2}. \tag{2.1}$$

Instead of making a hard decision, LLR values are usually combined to enhance the reliability. Now, a larger y produces a more confident decision. Moreover, $l(y)$ contains channel's characteristic σ^2: the noisier the channel, the less reliable the estimation. We will see that for a broad class of channels, the probabilistic behaviour of $l(y)$ fully characterizes the channel.

It is clear that $L = l(Y)$ is a random variable. In fact, L is a *sufficient statistic* for estimating X given Y. This means that an optimal decoder can be based on $l(y)$ instead of y itself [31]. Furthermore, according to the data processing inequality, the mutual information between X and Y is equal to the mutual information between X and L.

For a discrete memoryless channel (DMC) where $\mathcal{Y}$ is countable, the transition matrix is a matrix whose rows indicate inputs and columns indicate outputs. The element at the row corresponding to x and the column corresponding to y is $P_{Y|X}(y|x)$. For example, the transition matrices of BSC(ϵ) and BEC(ϵ) are

$$\mathbf{G}_{\mathrm{BSC}(\epsilon)} = \begin{array}{c} \\ +1 \\ -1 \end{array} \begin{array}{c} \begin{array}{cc} +1 & -1 \end{array} \\ \begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix} \end{array}$$

are

$$\mathbf{G}_{\mathrm{BEC}(\epsilon)} = \begin{array}{c} \\ +1 \\ -1 \end{array} \begin{array}{c} \begin{array}{ccc} +1 & e & -1 \end{array} \\ \begin{bmatrix} 1-\epsilon & \epsilon & 0 \\ 0 & \epsilon & 1-\epsilon \end{bmatrix} \end{array}.$$

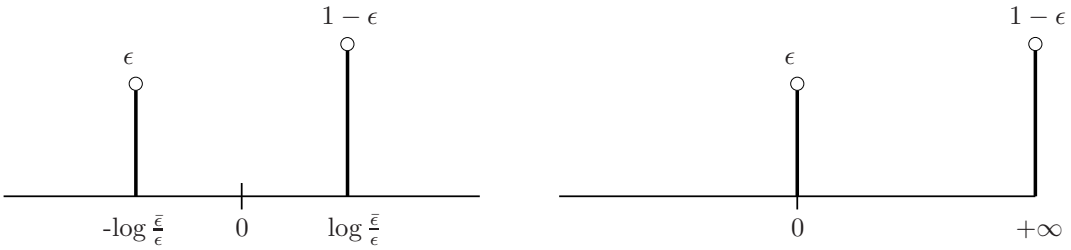

Figure 2.1: The L-densities of a BSC(ϵ) (left) and a BEC(ϵ) (right).

Definition 2.1 [Symmetric Channels [2]]**:** A channel is said to be symmetric if the rows of the channel transition matrix are permutations of each other and the columns are permutations of each other.

Although we cannot define a transition matrix for a general binary-input channel $P_{Y|X} : \{-1, +1\} \mapsto \mathcal{Y}$, $\mathcal{Y} \subset \bar{\mathbb{R}}$, we could partition the transition probability matrix for a countable output alphabet into sub matrices of the form

$$\begin{array}{c} \\ +1 \\ -1 \end{array} \begin{array}{c} \begin{array}{cc} \quad y & \quad -y \end{array} \\ \left[\begin{array}{cc} P_{Y|X}(y|+1) & P_{Y|X}(-y|+1) \\ P_{Y|X}(y|-1) & P_{Y|X}(-y|-1) \end{array}\right] \end{array}.$$

Now, if $P_{Y|X}(y|-1) = P_{Y|X}(-y|+1)$ for all $y \in \mathcal{Y}$, then each row (column) will be a permutation of each other row (column), hence a symmetric channel.

Definition 2.2 [MBIOS Channel]**:** A memoryless binary-input channel $P_{Y|X} : \mathcal{X} \mapsto \mathcal{Y}$ is said to be output-symmetric if $P_{Y|X}(y|-1) = P_{Y|X}(-y|+1)$ for all $y \in \mathcal{Y}$. Such channel is called a memoryless binary-input output-symmetric (MBIOS) channel.

2.1.1 L-densities

In this section, we borrow most of the notations and definitions from [31]. Let Y denote the observation and $L = l(Y)$ the LLR associated with Y. Also, let a be the density and A be the distribution of L given $X = +1$. We say that a is an *L-density* and A is an *L-distribution*. Let $\mathsf{H}_t(x) = \mathbf{1}_{\{x \geq t\}}$ be the Heavyside distribution and $\Delta_t(x)$ be its associated density. The L-density and L-distribution of BSC(ϵ) and BEC(ϵ) are given by

$$\mathsf{a}_{\mathrm{BSC}(\epsilon)}(x) = \epsilon\Delta_{-\log\frac{1-\epsilon}{\epsilon}}(x) + (1-\epsilon)\Delta_{\log\frac{1-\epsilon}{\epsilon}}(x),$$
$$\mathsf{a}_{\mathrm{BEC}(\epsilon)}(x) = \epsilon\Delta_0(x) + (1-\epsilon)\Delta_{+\infty}(x)$$

and

$$\mathsf{A}_{\mathrm{BSC}(\epsilon)}(x) = \epsilon\mathsf{H}_{-\log\frac{1-\epsilon}{\epsilon}}(x) + (1-\epsilon)\mathsf{H}_{\log\frac{1-\epsilon}{\epsilon}}(x),$$
$$\mathsf{A}_{\mathrm{BEC}(\epsilon)}(x) = \epsilon\mathsf{H}_0(x).$$

Also, according to (2.1), the L-density of a BIAWGN channel is Gaussian with mean $\frac{2}{\sigma^2}$ and variance $\frac{4}{\sigma^2}$. The L-densities of a BSC(ϵ) and a BEC(ϵ) are shown in Fig. 2.1.

Definition 2.3 [Symmetry of L-density [31]]**:** We call an L-density symmetric if for all $x \in \bar{\mathbb{R}}$, $\mathsf{a}(-x) = e^{-x}\mathsf{a}(x)$. Equivalently, we call an L-distribution symmetric if[1]

$$\int f(x)d\mathsf{A}(x) = \int e^{-x}f(-x)d\mathsf{A}(x)$$

for all bounded continuous function f such that $f(-x)e^{-x}$ is also bounded.

It is easy to check that $\mathsf{a}_{\mathrm{BSC}(\epsilon)}$ and $\mathsf{a}_{\mathrm{BEC}(\epsilon)}$ are symmetric densities. It is important to note that since LLR values can be infinite with a positive probability, we have to tweak the definition of a distribution to capture such characteristic. To do so, let $\mathcal{A}_L$ be the space of right-continuous, non-decreasing functions A defined over $\mathbb{R}$ satisfying

$$\lim_{x\to-\infty}\mathsf{A}(x) = 0, \quad \lim_{x\to+\infty}\mathsf{A}(x) \le 1.$$

According to Kolmogorov [32], we can associate a random variable L to each distribution $\mathsf{A} \in \mathcal{A}_L$ such that $L \in (-\infty, +\infty]$. This is quite similar to the conventional definition of a distribution, except that we allow L to have some probability mass at $+\infty$ by letting $\lim_{x\to+\infty}\mathsf{A}(x) \le 1$. The integral $\int g(x)d\mathsf{A}(x)$ for an L-distribution $\mathsf{A} \in \mathcal{A}_L$ and any non-negative continuous function g is interpreted as $\int g(x)\mathsf{a}(x)dx$, where a is A's corresponding L-density. However, if $\lim_{x\uparrow+\infty}\mathsf{A}(x) < 1$ and the limit $\lim_{x\to+\infty}g(x)$ exists, then we have to include

[1]Integrals are understood in the Lebesgue-Stieltjes sense.

the term

$$\left(1 - \lim_{x\uparrow+\infty} \mathsf{A}(x)\right) \lim_{x\to+\infty} g(x)$$

in the definition of $\int g(x)d\mathsf{A}(x)$. For example,

$$\int \log_2(1+e^{-x})d\mathsf{A}_{\mathrm{BEC}(\epsilon)}(x) = \int \log_2(1+e^{-x})\mathsf{a}_{\mathrm{BEC}(\epsilon)}(x)dx + (1-\epsilon)\times 0 = \epsilon.$$

Theorem 2.1 [Symmetry of L-distributions for MBIOS Channels [31, Theorem 4.26]]**:** Let A denote the L-distribution of an MBIOS channel. Then A is symmetric.

The L-density completely characterizes its associated MBIOS channel. In this thesis, we refer to an MBIOS channel by its associated L-density.

2.1.2 P-densities

Let a be the L-density of a symmetric channel. There are two interesting facts:

1. Due to symmetry, the negative tail of an L-density can be made from the positive tail.

2. An insightful observation about a is that it can be seen as a probabilistic weighting of the family $\{\mathsf{a}_{\mathrm{BSC}(\epsilon)}\}_{\epsilon\in[0,1]}$. As an example, a BEC can be seen as a weighted combination of BSC(0) ($\mathsf{a}_{\mathrm{BSC}(0)}(x) = \Delta_{+\infty}(x)$) and BSC($\frac{1}{2}$) ($\mathsf{a}_{\mathrm{BSC}(\frac{1}{2})}(x) = \Delta_0(x)$). This fact has been used originally in [33] and later in [34–36].

The first fact suggests that an L-density (assuming $X = +1$) carries redundant information and can be reproduced from the density of $|L|$ which is independent of X. To elaborate, let $|\mathsf{a}|$ denote the density of $|L|$. We say $|\mathsf{a}|$ is an $|L|$-density. We have

$$|\mathsf{a}|(x) = \mathsf{a}(x) + \mathsf{a}(-x) = \mathsf{a}(x)(1+e^{-x}),\ x \geq 0. \tag{2.2}$$

Therefore, a can be recovered from $|\mathsf{a}|$ by

$$\mathsf{a}(x) = \mathbf{1}_{\{x\geq 0\}}\frac{|\mathsf{a}|(x)}{1+e^{-x}} + \mathbf{1}_{\{x\leq 0\}}\frac{|\mathsf{a}|(-x)}{1+e^{-x}},\ x \in (-\infty, +\infty].$$

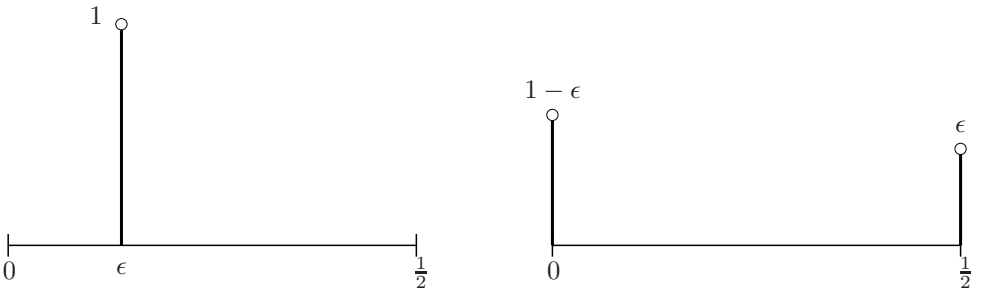

Figure 2.2: The P-densities of a BSC(ϵ) (left) and a BEC(ϵ) (right).

Having used the first fact, we exploit the second fact in this way: $|\mathsf{a}|$ can be seen as a probabilistic weighting of $\{|\mathsf{a}_{\mathrm{BSC}(\epsilon)}|\}_\epsilon$. Note that

$$|\mathsf{a}_{\mathrm{BSC}(\epsilon)}|(x) = \Delta_{\log\frac{1-\epsilon}{\epsilon}}(x),\ x \geq 0. \tag{2.3}$$

implying that ϵ must fall in the interval $[0, \frac{1}{2}]$ instead of $[0, 1]$. This ensures that we only sweep the positive tail of a. Therefore, the family

$$\{|\mathsf{a}_{\mathrm{BSC}(\epsilon)}|\}_{\epsilon\in[0,\frac{1}{2}]}$$

is rich enough to reproduce any $|L|$-density and consequently any L-density. According to (2.3), there is a random variable $P \in [0, \frac{1}{2}]$ such that $|L| = \log\frac{1-P}{P}$ or, equivalently

$$P = \frac{1}{1+e^{|L|}}. \tag{2.4}$$

In order to quantify the "probabilistic weighting of $|\mathsf{a}_{\mathrm{BSC}}|$", we have to find the density of P from the $|L|$-density by the change of variables given in (2.4) and using (2.2) as

$$\begin{aligned}\mathsf{g}(p) &= \frac{1}{p(1-p)}|\mathsf{a}|\left(\log\tfrac{1-p}{p}\right)\\ &= \frac{1}{p(1-p)^2}\mathsf{a}\left(\log\tfrac{1-p}{p}\right),\ p \in [0, \tfrac{1}{2}].\end{aligned} \tag{2.5}$$

In this dissertation, we call g a *P-density*. Similar to an L-density, a P-density completely characterizes its associated MBIOS channel. Also, one can define the associated P-distribution in the conventional way. We will use P-densities extensively throughout Chapter 3 and 4.

P-densities imply that an MBIOS channel can be seen as a collection of BSCs with crossover probabilities $p \in [0, \frac{1}{2}]$, each happening with probability $\mathsf{g}(p)$. For example as shown in Fig. 2.2, BSC(ϵ) is expressed by $\mathsf{g}(p) = \Delta_\epsilon(p)$ and BEC(ϵ) has a P-density of the form

$$\mathsf{g}(p) = (1-\epsilon)\Delta_0(p) + \epsilon\Delta_{\frac{1}{2}}(p).$$

Remark 2.1: Although we used the random variable P in a technical context, we have to add that $P = P(Y)$ has an important meaning. According to [36], P represents the uncoded probability of error upon receiving Y, i.e.,

$$P = \min\{P_{X|Y}(+1|Y), P_{X|Y}(-1|Y)\}.$$

Remark 2.2: In terms of optimization, dealing with P-densities is much easier than L-densities because not only the support of P is finite, but also its density does not impose any symmetry constraints.

2.1.3 Functionals over Symmetric Densities

In this section, we present three functional over the L densities which will be used in Chapter 3 and 4.

The capacity of a symmetric channel is achieved by a uniform input distribution [2]. Therefore, capacity of an MBIOS channel in bits per channel use is

$$\begin{aligned} C &= \mathbb{E}\left[\log_2 \frac{P_{Y|X}(Y|X)}{P_Y(Y)}\right] \\ &= \mathbb{E}\left[\log_2 \frac{2P_{Y|X}(Y|X)}{P_{Y|X}(Y|X) + P_{Y|X}(Y|-X)}\right] \\ &= \mathbb{E}\left[\log_2 \frac{2}{1+e^{-L}}\right]. \end{aligned} \tag{2.6}$$

Given $X = +1$, a decision error occurs at an optimal receiver whenever $L(y) < 0$ with probability one or $L(y) = 0$ with probability one half. Taking into account all realizations of y, we can infer that the *uncoded probability of error* of an MBIOS channel, given $X = +1$, is the area under the negative tail of its

L-density plus a half of the mass at the origin, i.e.,

$$\begin{aligned} P_{u,e} &= \int \mathbf{1}_{\{x<0\}} d\mathsf{A}(x) + \lim_{\delta \downarrow 0} \frac{1}{2} \int \mathbf{1}_{\{-\delta<x<\delta\}} d\mathsf{A}(x) \\ &= \lim_{\delta \downarrow 0} \mathbb{E}\left[\mathbf{1}_{\{L<0\}} + \frac{1}{2}\mathbf{1}_{\{-\delta<L<\delta\}}\right] \\ &= \frac{1}{2}\,\mathbb{E}\left[\exp\left(-\frac{1}{2}(L+|L|)\right)\right]. \end{aligned} \tag{2.7}$$

The Bhattacharyya parameter of an MBIOS channel is

$$B = \mathbb{E}[e^{-\frac{L}{2}}].$$

It can be seen that in fact $\frac{1}{2}B$ is the tightest Chernoff bound on the probability of error [37].

2.2 Chebychev Systems

In this section, we briefly mention definitions and key results of the theory of Chebychev systems. The reader is referred to [38, 39] for more information.

In the classic moment problem, one tries to find a measure with the prescribed moments, e.g., mean, variance, etc, where the underlying system of interest is the set of polynomials $\{1, t, t^2, \ldots\}$. The theory of Chebychev systems generalizes this problem to other sets of functions exhibiting a *specific* structure.

Definition 2.4 [T–System]**:** A set of real-valued continuous functions $\mathcal{U} = \{u_0, \ldots, u_n\}$ defined on the interval $[a, b]$ is called a *Chebychev system* or T–*system* if the determinant

$$\mathsf{D}(u_0, \ldots, u_n; t_0, \ldots, t_n) = \begin{vmatrix} u_0(t_0) & u_0(t_1) & \cdots & u_0(t_n) \\ u_1(t_0) & u_1(t_1) & \cdots & u_1(t_n) \\ \vdots & \vdots & & \vdots \\ u_n(t_0) & u_n(t_1) & \cdots & u_n(t_n) \end{vmatrix} \tag{2.8}$$

does not vanish for any $a \le t_0 < t_1 < \ldots < t_n \le b$. A T–system is called a *Complete Chebychev system* or *CT-system* if $\{u_0, u_1, \ldots, u_k\}$ is a T–system for $k = 0, 1, \ldots, n$.

The letter T in "T-system" stands for Tchebycheff, the Russian spelling of Chebychev . It can be seen that for $0 \leq t_0 < t_1 < \ldots < t_n \leq 1$, the determinant $\mathsf{D}(1, t, \ldots, t^n; t_0, \ldots, t_n)$, which was mentioned before, is a Vandermonde determinant, hence strictly positive.

Remark 2.3: From the continuity of $\mathsf{D}(u_0, \ldots, u_n; t_0, \ldots, t_n)$, it is deduced that $\mathcal{U}$ is a T–system if and only if the determinant of (2.8) maintains a single sign for any choice of t_k's. Without loss of generality, we may multiply each function u_k by a $+1$ or -1 to have a positive determinant. Such T–systems are called T_+–systems. From now on, we assume that such multiplications have already been made and the system $\mathcal{U}$ is a T_+–system.

Remark 2.4: The definition of T–systems can be extended to any proper interval. Moreover, a T–system on a proper interval $\mathcal{I} \subset \mathbb{R}$ is a T–system on any proper interval embedded in $\mathcal{I}$.

Remark 2.5: An equivalent characterization of T–systems is expressed in terms of the number of zeros of an arbitrary linear combination of $\{u_0, \ldots, u_n\}$. A system $\mathcal{U}$ is a T–system on a proper interval $\mathcal{I}$ if and only if any function in the linear space spanned by $\mathcal{U}$ has at most n zeros in $\mathcal{I}$. The term *polynomial* is used to refer to a function in the span of $\mathcal{U}$. We will be using the polynomial characterization of T–systems whenever it suits our problem.

If the system is sufficiently smooth, we may allow equalities among t_i's [38]. Let $\mathcal{C}^i(\mathcal{I})$, $i \geq 0$, be the class of functions defined on the proper interval $\mathcal{I}$ that posses i continuous derivatives. If the functions $u_i \in \mathcal{C}^n(\mathcal{I})$, for any set $t_0 \leq t_1 \leq \cdots \leq t_n$ of points in $\mathcal{I}$, the determinant $\mathsf{D}^*(u_0, \ldots, u_n; t_0, \ldots, t_n)$ is defined to be the determinant of the RHS of (2.8), where for each set of consecutive t_k, the corresponding columns are replaced by the successive derivatives evaluated at the point. For example,

$$\mathsf{D}^*(u_0, u_1, u_2; t_0, t_1, t_1) = \begin{vmatrix} u_0(t_0) & u_0(t_1) & u_0'(t_1) \\ u_1(t_0) & u_1(t_1) & u_1'(t_1) \\ u_2(t_0) & u_2(t_1) & u_2'(t_1) \end{vmatrix}$$

and $\mathsf{D}^*(u_0, \ldots, u_n; t, \ldots, t) = \mathsf{W}(u_0, \ldots, u_n)(t)$ is the Wronskian of $\mathcal{U}$:

$$\mathsf{W}(\mathcal{U})(t) = \mathsf{W}(u_0, \ldots, u_n)(t) = \begin{vmatrix} u_0(t) & u_0'(t) & u_0''(t) & \cdots & u_0^{(n)}(t) \\ u_1(t) & u_1'(t) & u_1''(t) & \cdots & u_1^{(n)}(t) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_n(t) & u_n'(t) & u_n''(t) & \cdots & u_n^{(n)}(t) \end{vmatrix}, \ t \in \mathcal{I},$$

where $u_i^{(n)}$ is the nth derivative of u_i [40].

Definition 2.5 [Extended T–System]**:** The system $\mathcal{U}$ is called an *Extended Chebychev system* or *ET*–system on $\mathcal{I}$ provided that for any set $t_0 \leq \cdots \leq t_n$ of the points in $\mathcal{I}$, $\mathsf{D}^*(u_0, \ldots, u_n; t_0, \ldots, t_n) > 0$, and it is called an *Extended Complete Chebychev system* or *ECT–system* if $\{u_0, u_1, \ldots, u_k\}$ is an *ET*–system on $\mathcal{I}$ for all $k = 0, \ldots, n$.

Theorem 2.2 [[38, Theorem 1.1, Sec. XI.1]]**:** Let $\mathcal{U}$ be of class $\mathcal{C}^n(\mathcal{I})$. Then, $\mathcal{U}$ is an *ECT*–system on $\mathcal{I}$ if and only if for $k = 0, \ldots, n$ we have $\mathsf{W}(u_0, \ldots, u_k) > 0$ on $\mathcal{I}$.

In the sequel, all integrals are understood in the Lebesgue-Stieltjes sense. Let $\mathbf{c} = (c_0, c_1, \ldots, c_n) \in \mathbb{R}^{n+1}$.

Definition 2.6 [Moment Space]**:** The moment space $\mathcal{M}_{n+1}$ associated with the T–system $\mathcal{U}$ is the convex cone

$$\mathcal{M}_{n+1} = \left\{ \mathbf{c} : c_i = \int_a^b u_i(t) d\sigma(t), \sigma \in \Sigma, 0 \leq i \leq n \right\},$$

where Σ is the set of all non-decreasing right continuos functions (distributions) of bounded variation.

The definition of σ falls into the general class of distributions. To deal with probability measures, we introduce an extra moment by letting $u_0(t) = 1$, i.e., $\int_a^b d\sigma(t) = 1$. Define the set

$$\Sigma(\mathbf{c}) = \left\{ \sigma \in \Sigma : \int_a^b u_i(t) d\sigma(t) = c_i, 0 \leq i \leq n \right\}.$$

If $\mathbf{c} \in \text{Int } \mathcal{M}_{n+1}$, the set $\Sigma(\mathbf{c})$ contains infinite distributions. $\Sigma(\mathbf{c})$ contains only one distribution when $\mathbf{c} \in \partial\mathcal{M}_{n+1}$, i.e., when $\mathbf{c}$ is a boundary point

of $\mathcal{M}_{n+1}$ [38]. Let $\sigma \in \Sigma(\mathbf{c})$ be a distribution with finitely many points of increase, i.e.,

$$d\sigma(t) = \sum_{j=1}^{m} \kappa_j \Delta_{t_j}(t),$$

where κ_j is the mass at point t_j. Then a representation of $\mathbf{c}$ associated with σ is

$$c_i = \int_a^b u_i(t) d\sigma(t) = \sum_{j=1}^{m} \kappa_j u_i(t_k), 0 \leq i \leq n,$$

where the points $\{t_j\}_{j=1}^m$ are called the *roots of representation*. The *index of representation* is defined as $\sum_{j=1}^m \epsilon(t_k)$ where the index function is $\epsilon(t) = 2$ for $t \in (a, b)$ and $\epsilon(t) = 1$ for $t = a, b$.

Lemma 2.1 [[39, Theorem 4.1, Sec. III.4]]: $\mathbf{c} \in \partial\mathcal{M}_{n+1}$ if and only if it admits a representation of index not greater than n.

Definition 2.7 [Principle and Canonical Representation]: Let $\mathbf{c} \in \text{Int } \mathcal{M}_{n+1}$. A representation of $\mathbf{c}$ of index $n+1$ is called *principle*. A representation of index $n+2$ is called *canonical*. If b is a root of representation, it is called *upper* principle/canonical, if not it is called *lower* principle/canonical.

Theorem 2.3 [Roots of Principle Representation [39, p. 77]]: There are only two possible types of principle representation:

1) $n = 2\nu - 1$:

   - lower principle: $\{t_0, \ldots, t_{\nu-1}\}$ where $a < t_0 < \cdots < t_{\nu-1} < b$
   - upper principle: $\{t_0, \ldots, t_\nu\}$ where $a = t_0 < t_1 < \cdots < t_{\nu-1} < t_\nu = b$

2) $n = 2\nu$:

   - lower principle: $\{t_0, \ldots, t_\nu\}$ where $a = t_0 < t_1 < \cdots < t_\nu < b$
   - upper principle: $\{t_0, \ldots, t_\nu\}$ where $a < t_0 < t_1 < \cdots < t_{\nu-1} < t_\nu = b$

Let $\mathcal{U}$ be a T–system and Ω be a continuous function. The following theorem gives the extremes of $\int_a^b \Omega(t) d\sigma(t)$ when $\sigma \in \Sigma(\mathbf{c})$ and $\mathbf{c} \in \text{Int } \mathcal{M}_{n+1}$:

Theorem 2.4 [Extremal Distributions [38, Theorem 1.1, Sec. III.I]]**:** Let $\mathbf{c} \in$ Int $\mathcal{M}_{n+1}$. If both $\mathcal{U}$ and the augmented system $\mathcal{U} \cup \{\Omega\}$ are T–systems, then

$$\sup_{\sigma \in \Sigma(\mathbf{c})} \int_a^b \Omega(t) d\sigma(t)$$

is attained uniquely for σ^*, the measure corresponding to the upper principle representation of $\mathbf{c}$ and

$$\inf_{\sigma \in \Sigma(\mathbf{c})} \int_a^b \Omega(t) d\sigma(t)$$

is attained uniquely for σ_*, the measure corresponding to the lower principle representation of $\mathbf{c}$.

Remark 2.6: Note that as long as the augmented system is a T–system, the optimizing distributions σ_* and σ^* are independent of the objective function Ω.

Chapter 3

Fundamental Basis Channels

3.1 Introduction

In this chapter, we introduce a set of equal-capacity symmetric channels over which any MBIOS channel can be decomposed. We call this the set of *basis channels* and denote it by $\mathcal{G}(C)$ to emphasize its dependency on a given capacity. We will show that there are several monotonicity properties among the set of basis channels. These properties will be exploited in Chapter 4. In Appendix E, we will present an application of the set of basis channels in designing universal codes over MBIOS channels of the same capacity.

3.1.1 From L-densities Towards P-densities

In Section 2.1.2, it was shown that for a given symmetric channel, the P-density can be obtained from the channel's L-density. In this section, we elaborate the migration from P-densities to L-densities. To this end, we show how to take an expected value with respect to an L-density using the corresponding P-density.

Lemma 3.1 [Change of Measures]**:** Let φ be a continuous function. We have

$$\mathbb{E}[\varphi(L)] = \mathbb{E}\left[(1-P)\varphi\left(\log\tfrac{1-P}{P}\right) + P\varphi\left(-\log\tfrac{1-P}{P}\right)\right].$$

Proof: By symmetry of a, we have

$$\begin{aligned}\mathbb{E}[\varphi(L)] &= \int_{\mathbb{R}} \varphi(l)\mathsf{a}(l)dl \\ &= \int_0^\infty \varphi(l)\mathsf{a}(l)dl + \int_0^\infty e^{-l}\varphi(-l)\mathsf{a}(l)dl \\ &= \int_0^\infty [\varphi(l) + e^{-l}\varphi(-l)]\mathsf{a}(l)dl.\end{aligned}$$

Now, taking $l = \log\frac{1-p}{p}$ for $p \in [0, \frac{1}{2}]$ and using (2.5), we obtain the desired form by

$$\begin{aligned}\mathbb{E}[\varphi(L)] &= \int_0^{\frac{1}{2}} \left[\varphi\left(\log\tfrac{1-p}{p}\right) + \frac{p}{1-p}\varphi\left(-\log\tfrac{1-p}{p}\right)\right] \mathsf{a}\left(\log\tfrac{1-p}{p}\right) \frac{dp}{p(1-p)} \\ &= \int_0^{\frac{1}{2}} \left[(1-p)\varphi\left(\log\tfrac{1-p}{p}\right) + p\varphi\left(-\log\tfrac{1-p}{p}\right)\right] \mathsf{g}(p)dp.\end{aligned}$$

□

An immediate application of Lemma 3.1 is the representation of the functionals over L-densities introduced in Section 2.1.3 using P-densities. The capacity of an MBIOS channel in bits per channel use is

$$C = 1 - \mathbb{E}[h(P)],$$

where we used Lemma 3.1. From now on, we denote the equivocation of the channel by $H = 1 - C$ [41]. Also, by (2.7), the probability of error of an MBIOS channel given $X = +1$ is

$$P_{u,e} = \lim_{\delta\downarrow 0} \mathbb{E}\left[\mathbf{1}_{\{L<0\}} + \tfrac{1}{2}\mathbf{1}_{\{-\delta<L<\delta\}}\right] = \mathbb{E}[P],$$

where we used Lemma 3.1 and the Lebesgue's dominated convergence theorem (DCT) [42].

Note that the derivation of above equations is rather intuitive considering the probabilistic interpretation of an L-density as a family of BSCs[1]. By the concavity of the binary entropy function and the fact that it lays above $2p$, $p \in [0, \frac{1}{2}]$, it is straightforward to see that by the Jensen's inequality [36]

$$2\,\mathbb{E}[P] \le \mathbb{E}[h(P)] \le h(\mathbb{E}[P]),$$

[1]Recall that the uncoded probability of error and capacity of BSC(ϵ) are $P_{u,e} = \epsilon$ and $C = 1 - h(\epsilon)$.

where the left (right) bound is tight for a BEC (BSC). Thus, for a fixed capacity C, the error probability is bounded by

$$h^{-1}(H) \leq P_{u,e} \leq \frac{H}{2}. \tag{3.1}$$

Finally, the Bhattacharyya parameter of an MBIOS channel can be shown as

$$B = \mathbb{E}\left[2\sqrt{P(1-P)}\right].$$

We have [31, Lemma 4.64]

$$2P_{u,e} \leq B \leq 2\sqrt{P_{u,e}(1-P_{u,e})},$$

where the left (right) inequality is tight for a BEC (BSC).

3.1.2 New T–Systems

We mentioned in Remark 2.2 that since the support of a P-density is finite, it is a perfect candidate for optimization problems including symmetric channels. We will see that the theory of Chebychev systems is a great tool for solving such optimization problems analytically. In this thesis, we will exploit several T–systems in our development. The following theorem summarizes those systems. The proof is postponed to Appendix A.

Theorem 3.1 [New T–Systems]**:** Let $\mu : [0, \frac{1}{2}] \times \mathbb{R}_+ \mapsto [0, 1]$ be a function defined as

$$\mu(x, \rho) = 2^{-\rho}\left(x^{\frac{1}{1+\rho}} + (1-x)^{\frac{1}{1+\rho}}\right)^{1+\rho}.$$

The following systems are T–systems over $[0, \frac{1}{2}]$:

$$\begin{aligned}
&\mathcal{U}_0 : \{1, h(p)\} \\
&\mathcal{U}_1 : \{1, h(p), p\} \\
&\mathcal{U}_2 : \{1, h(p), -\left(p^{a_1} + (1-p)^{a_1}\right)^{b_1}\left(p^{a_2} + (1-p)^{a_2}\right)^{b_2}\}, \\
&\qquad 0 < a_1, a_2 < 1,\ 0 < b_1, b_2,\ a_1 b_1 + a_2 b_2 = 1 \\
&\mathcal{U}_3 : \{1, h(p), p, \mu(p, \rho)\},\ \rho > 0 \\
&\mathcal{U}_4 : \{1, h(p), 1 - (1-2p)^2\}
\end{aligned}$$

$$\mathcal{U}_5 : \Big\{1, h(p), -\lim_{\rho\to 1}\frac{\partial}{\partial\rho}\mu(p,\rho)\Big\}$$
$$\mathcal{U}_6 : \Big\{1, h(p), -p^s(1-p)^{1-s} - p^{1-s}(1-p)^s\Big\},\ s\in(0,1)$$
$$\mathcal{U}_7 : \Big\{1, h(p), -p\log_2^2 p - (1-p)\log_2^2(1-p)\Big\}$$
$$\mathcal{U}_8 : \Big\{1, h(p), p, p\log_2^2 p + (1-p)\log_2^2(1-p)\Big\}$$

Proof: See Appendix A. □

For simplicity, we drop the arguments wherever possible, e.g., we write $\{1, h\}$ instead of $\mathcal{U}_0$.

3.2 Equal-Capacity Basis Channels

Let $\mathcal{S}(C)$ be the set of MBIOS channels of capacity C for some $C \in (0,1)$ (we consider non-trivial cases). We showed in [43] that there exists a set of equal-capacity basis channels such that every channel from $\mathcal{S}(C)$ is included in their convex hull. We will study the monotonicity and extremal properties of these channels. Let $\eta = h^{-1}(H)$ be the crossover probability of the BSC with capacity C. As shown in [43], a basis channel is constructed by mixing two BSCs with crossover probabilities x and y. The P-density of such channel is

$$\mathrm{g}_{x,y}(p) = \gamma(x,y)\Delta_x(p) + \bar{\gamma}(x,y)\Delta_y(p), \tag{3.2}$$

where $x \le \eta \le y$, $(x,y) \in \mathcal{D}(C)$, and

$$\mathcal{D}(C) = \big\{(x,y) \in [0,\eta]\times[\eta,1/2] : \eta = h^{-1}(H)\big\}.$$

It is important to note that $x \le \eta \le y$ must hold, since the capacity of BSC(x) (BSC(y)) is greater (less) than or equal to C. Therefore, there exits a convex combination of them with capacity C. The coefficient $\gamma(x,y)$ represents the probability of BSC(x) ($1-\gamma(x,y)$ corresponds to BSC(y)) and is uniquely determined such that the capacity of the basis channel $\mathrm{g}_{x,y}$ is equal to C, i.e.,

$$\gamma(x,y) = \frac{h(y)-H}{h(y)-h(x)}, \quad \forall (x,y)\in\mathcal{D}(C)\backslash\{x=y=\eta\}. \tag{3.3}$$

By the construction given in (3.2), a basis channel is an MBIOS channel in essence. A typical basis channel is shown in Fig. 3.1 which demonstrates that a basis channel is essentially a convex combination of two BSCs with *different capacities*, each happening with its associated probability γ and $\bar{\gamma}$, while the mixture has a fixed capacity C. The transition probability matrix of a $\mathsf{g}_{x,y}$ is

$$\mathbf{G}_{x,y} = \begin{bmatrix} \alpha\bar{x} & \alpha x & \bar{\alpha}\bar{y} & \bar{\alpha}y \\ \alpha x & \alpha\bar{x} & \bar{\alpha}y & \bar{\alpha}\bar{y} \end{bmatrix}, \tag{3.4}$$

where the rows and columns are the channel inputs and outputs, respectively. Note that if x (or y) is equal to η, then $\gamma(x,y) = 1$ ($\gamma(x,y) = 0$) in (3.2). We denote the set of equal-capacity basis channels by

$$\mathcal{G}(C) = \left\{ \mathsf{g}_{x,y}(p) : (x,y) \in \mathcal{D}(C) \right\}.$$

Clearly, BEC and BSC belong to this set and $\mathcal{G}(C) \subset \mathcal{S}(C)$. The following theorem is by [43].

Theorem 3.2 [Equal-Capacity Decomposition [43]]**:** Any MBIOS channel with capacity C and P-density g can be decomposed as

$$\mathsf{g}(p) = \iint_{\mathcal{D}(C)} \mathsf{g}_{x,y}(p) P_{\mathsf{g}}(x,y) dxdy,$$

for some probability assignment P_{g} over $\mathcal{D}(C)$.

3.2.1 Monotonicity and Extremal Distributions

We shall see that studying monotonicity properties of $\mathcal{G}(C)$ will prove extremely helpful in Chapter 4. We denote the expected value operator taken under $\mathsf{g}_{x,y}$ by $\mathbb{E}_{x,y}[\cdot]$. To elaborate why we are interested in monotonicity properties of the basis channels, assume that we are supposed to maximize $\mathbb{E}[\varphi(P)]$ for some continuous function $\varphi : [0, \frac{1}{2}] \to \mathbb{R}$. By the Fubini-Tonelli's theorem [42], we have

$$\begin{aligned} \mathbb{E}[\varphi(P)] &= \int_0^{\frac{1}{2}} \varphi(p) \mathsf{g}(p) dp \\ &= \int_0^{\frac{1}{2}} \varphi(p) \iint_{\mathcal{D}} P_{\mathsf{g}}(x,y) \mathsf{g}_{x,y}(p) dxdydp \\ &= \iint_{\mathcal{D}} P_{\mathsf{g}}(x,y)\, \mathbb{E}_{x,y}[\varphi(P)] dxdy. \end{aligned}$$

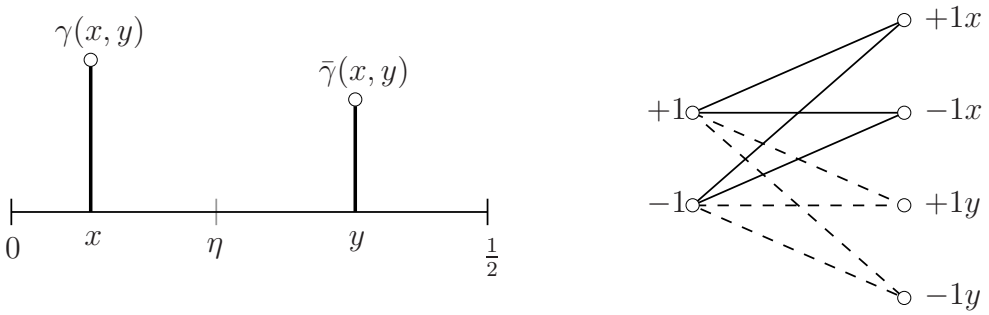

Figure 3.1: Left: The P-density representation of a basis channel. Right: The same basis channel is represented graphically, where each output is labeled according to the corresponding subchannel. Solid lines show BSC(x) happening with probability $\gamma(x,y)$ while dashed lines indicate the connections of BSC(y) happening with probability $\bar{\gamma}(x,y)$. Probabilities are according to $\mathbf{G}_{x,y}$ given in (3.4).

Assume that all of the basis channels obey $\mathbb{E}_{x,y}[\varphi(P)] \leq \mathbb{E}_{x^*,y^*}[\varphi(P)]$ for some fixed x^* and y^*. Then

$$\mathbb{E}[\varphi(P)] \leq \iint_{\mathcal{D}} P_{\mathsf{g}}(x,y)\, \mathbb{E}_{x^*,y^*}[\varphi(P)] dxdy = \mathbb{E}_{x^*,y^*}[\varphi(P)].$$

Therefore, under certain conditions on φ, we might be able to find extremes of $\mathbb{E}[\varphi(P)]$ by studying the extremes of $\mathbb{E}_{x,y}[\varphi(P)]$. The following lemma establishes monotonicity of the basis channels:

Lemma 3.2 [Extremes and Monotonicity of Basis Channels]**:** Let $\varphi : [0, \frac{1}{2}] \to \mathbb{R}$ be a continuous function such that $\{1, h, \varphi\}$ be a T–system. Then, $\sup \mathbb{E}[\varphi(P)]$ (resp. $\inf \mathbb{E}[\varphi(P)]$) over all MBIOS channels in $\mathcal{S}(C)$ is achieved by a BEC (resp. BSC). Moreover, the following monotonicity properties hold:

$$\frac{\partial}{\partial x}\, \mathbb{E}_{x,y}[\varphi(P)] \leq 0, \quad \frac{\partial}{\partial y}\, \mathbb{E}_{x,y}[\varphi(P)] \geq 0.$$

Proof: Optimizing $\mathbb{E}[\varphi(p)]$ over the set of MBIOS channels with the same capacity imposes two constraints on the P-density of the channel: the distribution must exhibit capacity C (or equivalently, it must exhibit equivocation equal to $H = 1 - C$) and it must be a probability distribution. In other words, we are seeking optimizing distributions for

$$\begin{array}{ll} \text{opt} & \mathbb{E}[\varphi(P)] \\ \text{s.t.} & \mathbb{E}[1] = 1 \\ & \mathbb{E}[h(P)] = H. \end{array}$$

According to Theorem 3.1, $\{1, h\}$ is a T–system. We claim that the point $\mathbf{c} = (1, H)$ is not a boundary point of $\mathcal{M}_2$. In fact, if it was, by Lemma 2.1 it would admit a representation of index one. This is impossible unless either $C = 0$ or $C = 1$. Therefore, $\mathbf{c} \in \text{Int } \mathcal{M}_2$. By Theorem 2.4, if the augmented system $\{1, h, \varphi\}$ is a T–system, then $\sup_{\Sigma(\mathbf{c})} \mathbb{E}[\varphi(P)]$ over all MBIOS channels of capacity C is attained uniquely for a distribution associated with the upper principle representation of $\mathbf{c}$ for $n = 1$ which is the BEC(H) (see Theorem 2.3). Similarly, $\inf_{\Sigma(\mathbf{c})} \mathbb{E}[\varphi(P)]$ is attained for a distribution associated with the lower principle representation of $\mathbf{c}$ for $n = 1$ which is the BSC(η).

To prove the monotonicity properties, let us restrict φ to the interval $[x, y]$, where $0 < x \leq \eta \leq y < \frac{1}{2}$. Clearly, $\{1, h, \varphi\}$ is a T–system on $[x, y]$ too. Therefore, BSC(η) is the minimizer and $\mathsf{g}_{x,y}$ is the maximizer of $\mathbb{E}[\varphi(P)]$. Let

$$\mathcal{S}_{x,y}(C) = \left\{ \mathsf{g} : P_{\mathsf{g}}(x', y') = 0, \ \forall (x', y') \in [0, x) \times (y, \tfrac{1}{2}] \right\}.$$

Now, we restrict φ to $[x - \delta_1, y + \delta_2]$ for some small positive δ_1, δ_2. Clearly, $\mathcal{S}_{x,y} \subset \mathcal{S}_{x-\delta_1, y+\delta_2}$ and

$$\sup_{\mathcal{S}_{x-\delta_1, y+\delta_2}(C)} \mathbb{E}[\varphi(P)] \geq \sup_{\mathcal{S}_{x,y}(C)} \mathbb{E}[\varphi(P)].$$

However, these supremum values are attained by $\mathsf{g}_{x,y}$ and $\mathsf{g}_{x-\delta_1, y+\delta_2}$, i.e.,

$$\mathbb{E}_{x-\delta_1, y+\delta_2}[\varphi(P)] \geq \mathbb{E}_{x,y}[\varphi(P)].$$

Since δ_1 and δ_2 are arbitrary, we get the result by

$$\lim_{\delta_1 \to 0} \frac{\mathbb{E}_{x,y}[\varphi(P)] - \mathbb{E}_{x-\delta_1, y}[\varphi(P)]}{\delta_1} \leq 0$$

and

$$\lim_{\delta_2 \to 0} \frac{\mathbb{E}_{x,y+\delta_2}[\varphi(P)] - \mathbb{E}_{x,y}[\varphi(P)]}{\delta_2} \geq 0.$$

□

3.2.2 Ordering on $\mathcal{G}(C)$

An important consequence of Theorem 3.1 and Lemma 3.2 is that one can order the set of basis channels $\mathcal{G}(C)$ and set of equal-capacity MBIOS channels $\mathcal{S}(C)$

with respect to those functionals defined over L and P-densities that satisfy the conditions of Lemma 3.2. Here, we exemplify the ordering among the set of basis channels. The probability of error of an MBIOS channel is given by $\varphi(p) = p$ which forms a T–system with $\{1, h\}$ (see $\mathcal{U}_1$ of Theorem 3.1). Therefore, the minimum (maximum) of probability of error over $\mathcal{S}(C)$ and $\mathcal{G}(C)$ is given by the BSC (BEC). We have obtained this result using the Jensen's inequality in (3.1). Moreover, among the basis channels, as we move inward (by increasing x and decreasing y, see Fig. 3.1), the probability of error will decrease monotonically. This means that for $x_1 \leq x_2 \leq \eta \leq y_2 \leq y_1$, $\mathbb{E}_{x_1,y_1}[p] \geq \mathbb{E}_{x_2,y_2}[p]$.

As we have seen in Section 2.1.3, the Bhattacharyya parameter of an MBIOS channel is given by $\varphi(p) = 2\sqrt{p(1-p)}$. By Theorem 3.1, $\{1, h, -\varphi\}$ is a T–system by taking $s = \frac{1}{2}$ in $\mathcal{U}_6$. Thus, the minimum (maximum) of probability of error over $\mathcal{S}(C)$ and $\mathcal{G}(C)$ is given by the BEC (BSC), i.e.,

$$\frac{1}{2}(1 - C) \leq B \leq 2\sqrt{\eta(1-\eta)}.$$

As final example, the mean-squared error of $\hat{X}(Y)$, an estimator of X given Y, is a random variable given by

$$\begin{aligned}\mathbb{E}[(X - \hat{X}(Y))^2|Y] &= 1 - 2\,\mathbb{E}[X\hat{X}(Y)|Y] + \hat{X}^2(Y) \\ &= 1 - 2\hat{X}(Y)\,\mathbb{E}[X|Y] + \hat{X}^2(Y),\end{aligned}$$

where we used the fact that $\hat{X}(Y)$ is measurable with respect to the σ-algebra generated by Y [44]. The minimum of this error is achieved by the minimum mean-squared error (MMSE) estimator

$$\hat{X}_{\text{MMSE}}(Y) = \mathbb{E}[X|Y] = P_{X|Y}(+1|Y) - P_{X|Y}(-1|Y).$$

Using the Bayes' theorem and assuming a uniform priori distribution, we have

$$\begin{aligned}\hat{X}_{\text{MMSE}}(Y) &= \frac{P_{Y|X}(Y|+1) - P_{Y|X}(Y|-1)}{P_{Y|X}(Y|+1) + P_{Y|X}(Y|-1)} \\ &= \frac{e^{L(Y)} - 1}{e^{L(Y)} + 1} \\ &= \tanh\tfrac{L(Y)}{2}.\end{aligned}$$

Table 3.1: Summary of the extremes of various functionals over $\mathcal{G}(C)$ and $\mathcal{S}(C)$

Functional	$\varphi(p)$	$\inf \mathbb{E}[\varphi(P)]$	$\sup \mathbb{E}[\varphi(P)]$	Monotonicity of $\mathbb{E}_{x,y}[\varphi(P)]$
$P_{u,e}$	p	BSC	BEC	$\frac{\partial}{\partial x} \leq 0,\ \frac{\partial}{\partial y} \geq 0$
B	$2\sqrt{p(1-p)}$	BEC	BSC	$\frac{\partial}{\partial x} \geq 0,\ \frac{\partial}{\partial y} \leq 0$
MMSE	$1-(1-2p)^2$	BSC	BEC	$\frac{\partial}{\partial x} \leq 0,\ \frac{\partial}{\partial y} \geq 0$

By Lemma 3.1, the MMSE can be expressed using the P-density of the channel as

$$\text{MMSE} = 1 - \mathbb{E}[\hat{X}^2(Y)] = 1 - \mathbb{E}\left[\tanh^2 \tfrac{L}{2}\right] = 1 - \mathbb{E}[(1-2P)^2].$$

The extremes of MMSE can be obtained in a similar fashion. Table 3.1 summarizes the extremes and monotonicity properties of some functionals over symmetric densities.

Chapter 4

Extremal Problems of Error Exponents

4.1 Introduction

Consider a DMC with a conditional probability measure $P_{Y|X} : \mathcal{X} \mapsto \mathcal{Y}$ where $\mathcal{X}$ and $\mathcal{Y}$ are the input and output sets. A codebook is a set of M codewords $(\mathbf{c}_1, \ldots, \mathbf{c}_M)$ where each $\mathbf{c}_i \in \mathcal{X}^n$ and n is the code length. The code rate, measured in bits, is $R = \frac{1}{n} \log_2 M$. According to Shannon's model of communication, a message from the set $\{1, \ldots, M\}$ is mapped to one of the codewords and is passed through the channel. At the receiving side, the decoder produces an estimate of the original message by observing the channel outputs. Assuming equiprobable messages, the average probability of error is

$$P_e = \frac{1}{M} \sum_{i=1}^{M} P_{e,i},$$

where $P_{e,i}$ denotes the probability of decoding error, averaged over the channel imperfections, when message i has been sent. We denote a code with three major parameters: (n, R, P_e) representing the code length, code rate and the probability of error. Let $R^*(n, P_e)$ be the maximum rate of transmission that is possible using a codebook of length n such that the original codeword can be recovered with probability at least $1 - P_e$ [45]. Shannon showed that

$$\lim_{P_e \to 0} \liminf_{n \to \infty} R^*(n, P_e) = C,$$

where C is the channel capacity in bits per channel use. For a DMC, $C = \sup_{P_X} I(X;Y)$ where $I(X;Y)$ is the mutual information between X and Y,

and P_X is a probability assignment on $\mathcal{X}$. Let

$$P_e^*(n, R) = \inf\{P_e : \exists(n, R, P_e)\text{–code}\}$$

denote the least probability of error among all codebooks of length n and rate R. An important characteristic of a DMC is the *channel reliability function*, defined as

$$E(R) = \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^*(n, R).$$

and is the exponent with which the probability of error for a given rate may be made to vanish with increasing n [3]. The reliability function is a function of channel and is known exactly for $R_{cr} \leq R \leq C$, where R_{cr} is called the *critical rate*. Since reliable communication is impossible for rates greater than the capacity, we have $E(R) = 0$, $R \geq C$. There are several lower and upper bounds known for the channel reliability function. In this chapter, we study the extremes of these bounds over the set of MBIOS channels. More specifically, we consider the random coding error exponent, expurgated error exponent, sphere-packing error exponent, and exponent for the erasure/list decoding. The following questions will be addressed in this chapter:

- Among all MBIOS channels of the same capacity, which channel does have the maximum (minimum) of the error exponent? In other words, given a sequence of codes $\{(n, R, P_e)\}_n$, $R < C$, which channel does have the fastest (slowest) vanishing probability of error?

- Among all MBIOS channels of the same capacity, which channel does exhibit the maximum (minimum) of the critical rate?

- How does one modify these extremes when an additional constraint is imposed, e.g., a constraint on the uncoded probability of error of the channel?

The analysis of the reliability function ($P_e^*(n, R)$) and channel capacity ($R^*(n, P_e)$) are asymptotic in code length. A very interesting question is remained to be answered: How does the code length affect $P_e^*(n, R)$ and $R^*(n, P_e)$ for a finite code length? In other words, how does one choose the

code length to achieve a desired rate/probability of error? Polyanskiy *et al.* [46] answer this question by re-introducing an important characteristic of channel called *channel dispersion*, V, which measures the stochastic variability of the channel relative to a deterministic channel with the same capacity [46][1]. For a large enough code length, it is proved that[2]

$$R^*(n, P_e) \approx C - \sqrt{\frac{V}{n}} Q^{-1}(P_e),$$

from which, one can approximate the required code length for the given rate/probability of error. Similar to the error exponents, one can find the channel, among all symmetric channels of the same capacity, which maximizes (minimizes) the channel dispersion, hence the maximum (minimum) of required code length.

To answer all these questions, we introduce a mathematical framework translating these problems into well-defined optimization problems. We define a set of equal-capacity basis channels over which every MBIOS channel can be decomposed. Then, we invoke the theory of Chebychev systems from approximation theory to solve these optimization problems analytically. We find several monotonicity properties among the basis channels and show that the set of basis channels is rich enough to answer these questions. It turns out that the BEC and BSC, which are two instances of basis channels, are these extremes when the only constraint is the capacity. We will show how one can modify these extremes when other constraints are imposed. In those cases, the extreme channels are still among the basis channels. This shows that the theory of Chebychev systems and the set of basis channels are powerful tools and can be used to solve even more complicated extremal problems of symmetric channels.

There are two related works. Fàbregas *et al.* [49] studied the Gallager's random coding error exponent. Using the properties of convex functions, it is shown that BSC and BEC are the extremes. The work of Alsan [50] shows the same extremality of BSC and BEC. The author then applies this result in the

[1] Channel dispersion was previously used in [47, 48] in a similar context.

[2] $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$

context of channel polarization [51]. In this thesis, we present a framework which not only provides these results, but also allows for solving extremal problems of other error exponents as well as solving the extremal problems under extra constraints.

The application of Chebychev systems to solve extremal problems related to symmetric channels has been considered by [52] and partly by [36] to solve extremal problems of information combining. The problems considered in this thesis, however, are different in nature and require totally different Chebychev systems.

4.2 Extremal Problems of Error Exponents

In this section, we shall solve several extremal problems related to the reliability function. Lower bounds on the reliability function including the Gallager's random coding error exponent and expurgated error exponent will be considered. The sphere-packing error exponent is studied as an upper bound on the reliability function. In addition to the conventional decoder, we also study the error exponent for decoders with erasure/list option. The extremal problems of channel dispersion, which can be obtained via the reliability function, will be solved at the end of this section.

4.2.1 Gallager's Random Coding Error Exponent

Let $P_{Y|X} : \mathcal{X} \mapsto \mathcal{Y}$ be a DMC. The average probability of decoding error with an $(n, R \text{ [nats]}, P_e)$ code under maximum-likelihood decoding is bounded by [3, Theorem 5.6.2]

$$P_e \leq \exp\left(-nE_r(R)\right),$$

where

$$E_r(R) = \sup_{\rho \in [0,1]} \left(E_0(\rho) - \rho R\right)$$

is called the *random coding error exponent*, $E_0(\rho) = \sup_{P_X} E_0(\rho, P_X)$ and

$$E_0(\rho, P_X) = -\log \sum_{y \in \mathcal{Y}} \left[\sum_{x \in \mathcal{X}} P_X(x) P_{Y|X}(y|x)^{\frac{1}{1+\rho}} \right]^{1+\rho}.$$

As we mentioned in Section 2.1, any MBIOS channel is completely characterized by its L-density or P-density. In this section, we derive an alternative representation for $E_0(\rho)$ based on the channel's P-density. We will utilize this representation to solve the extremal problems of random coding error exponent.

For symmetric channels, the uniform input distribution maximizes both $I(X;Y)$ and $E_0(\rho, P_X)$ [3, Theorem 4.5.2, 5.6.5]. We write $E_0(\rho) = E_0(\rho, P_X \equiv \frac{1}{2})$. The next lemma shows the representation of random coding error exponent using the P-density of the channel.

Lemma 4.1 [Representation of E_0 by a P-density]**:** For an MBIOS channel, we have

$$E_0(\rho) = -\log \mathbb{E}[\mu(P, \rho)],\ \rho \geq 0,$$

where

$$\mu(p, \rho) = 2^{-\rho}\left(p^{\frac{1}{1+\rho}} + (1-p)^{\frac{1}{1+\rho}}\right)^{1+\rho},\ p \in [0, \tfrac{1}{2}].$$

Proof: Assuming that the channel has a continuous output alphabet, we can change the summation to integral. We have

$$\begin{aligned}
\exp\left(-E_0(\rho)\right) &= \int \left(\sum_x P_X(x) P_{Y|X}(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho} dy \\
&= \int \left(\frac{1}{2} + \frac{1}{2} e^{-\frac{L(y)}{1+\rho}}\right)^{1+\rho} P_{Y|X}(y|+1) dy \\
&\overset{(a)}{=} \mathbb{E}\left[2^{-(1+\rho)}\left(1 + e^{-\frac{L}{1+\rho}}\right)^{1+\rho}\right] \\
&\overset{(b)}{=} \mathbb{E}\left[2^{-\rho}\left(P^{\frac{1}{1+\rho}} + (1-P)^{\frac{1}{1+\rho}}\right)^{1+\rho}\right],
\end{aligned}$$

where (a) holds by the change of measures[3] $P_{Y|X}(y|+1)dy = a(l)dl$ and (b) follows from Lemma 3.1.

□

Now that we have the representation of Gallager's error exponent using the P-density of channel, we are ready to find its extreme distributions, which are

[3]Note that the LLR is a sufficient statistic for X given Y.

the optimizing distributions of

$$\begin{array}{ll} \text{opt} & E_r(R) \\ \text{s.t.} & \mathbb{E}[1] = 1 \\ & \mathbb{E}[h(P)] = H. \end{array}$$

It is noteworthy that since

$$E_r(R) = \sup_{\rho \in [0,1]} (E_0(\rho) - \rho R),$$

the optimizer of the above problem is also the optimizer of the following problem provided that the optimizing distribution is independent of ρ:

$$\begin{array}{ll} \text{opt} & E_0(\rho) \\ \text{s.t.} & \mathbb{E}[1] = 1 \\ & \mathbb{E}[h(P)] = H. \end{array}$$

As we will see later in Theorems 4.1, 4.2 and 4.3, this is the case and those distributions do not depend on ρ. Note that maximizing $E_r(R)$ is equivalent to maximizing $E_0(\rho) = -\log \mathbb{E}[\mu(P, \rho)]$ and it leads to minimizing $\mathbb{E}[\mu(P, \rho)]$ since the logarithm is a monotonically increasing function. The following theorem shows the extremes of the Gallager's error exponent.

Theorem 4.1 [Extremes of E_r]**:** Among all equal-capacity MBIOS channels with capacity C, the minimum (maximum) of the error exponent $E_r(R)$ for every rate below the capacity is achieved by BSC(H) (resp. BSC(η)).

Proof: First, note that according to Theorem 3.1, the system

$$\begin{aligned} u_0(p) &= 1 \\ u_1(p) &= h(p) \\ u_2(p) &= -\mu(p, \rho), \ \rho \geq 0 \end{aligned}$$

is a T–system. To see this, let $a_1 = a_2 = \frac{1}{1+\rho}$, $b_1 = 1$ and $b_2 = \rho$ in $\mathcal{U}_2$. Now, by Theorem 2.4, the maximum (minimum) of $\mathbb{E}[\mu(P, \rho)]$ is obtained by the BSC (BEC). Since the extremes are independent of ρ, this means that the maximum (minimum) of E_0 and E_r are achieved by the BEC (BSC) of capacity C.

□

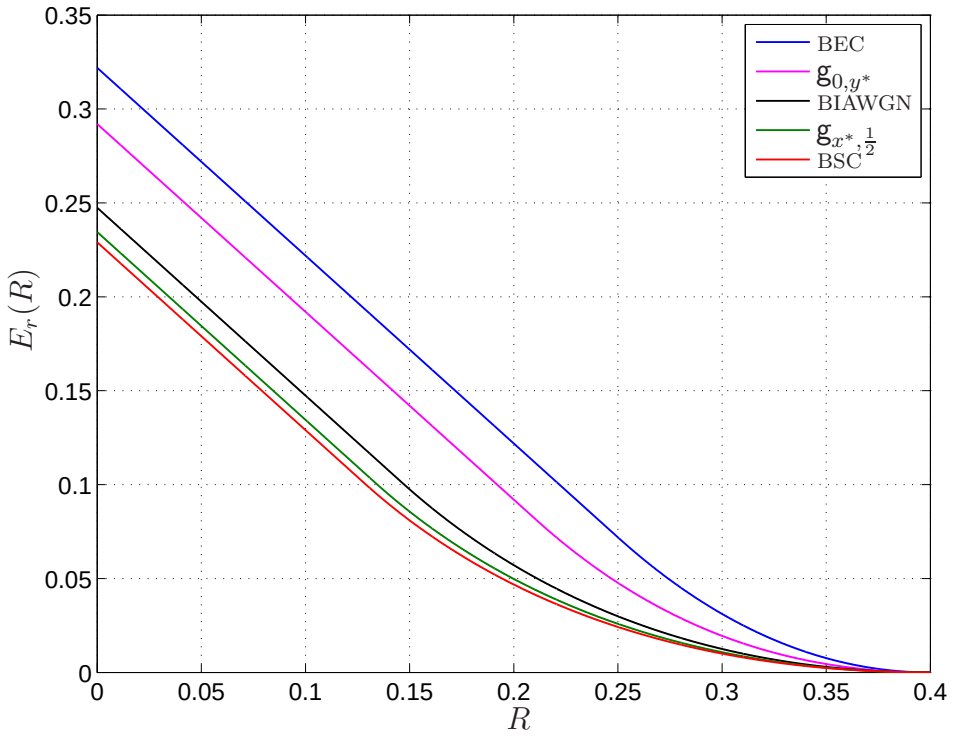

Figure 4.1: Comparison of the extremes of E_r for MBIOS channels of capacity $C = 0.4$ and probability of error $P_0 = 0.1921$ ($= P_{u,e}$ of the corresponding BIAWGN channel of capacity 0.4).

Theorem 4.1 shows extremal properties of the BSC and BEC. When there is a constraint on the channel's probability of error, the BEC and BSC are no longer the extremes of random coding error exponent. For example, the extremes of E_r among all MBIOS channels in $\mathcal{S}(C)$ with the probability of error equal to error probability of the Gaussian channel in $\mathcal{S}(C)$ contains neither the BEC nor the BSC. Mathematically speaking, we are interested in the solution of optimization problems of the following kind:

$$\begin{array}{ll} \text{opt} & E_r(R) \\ \text{s.t.} & \mathbb{E}[1] = 1 \\ & \mathbb{E}[h(P)] = H \\ & \mathbb{E}[P] \geq P_0 \text{ (or } \mathbb{E}[P] \leq P_0) \end{array}$$

In other words, we are seeking a single MBIOS channel from $\mathcal{S}(C)$ which satisfies a constraint on the probability of error and maximizes (minimizes) E_r for every $R < C$. It is important to point out that according to (3.1), P_0 must satisfy $h^{-1}(H) < P_0 < \frac{H}{2}$, otherwise the condition on $\mathbb{E}[P]$ will be relaxed

automatically. Furthermore, if P_0 is either $\frac{H}{2}$ or η, then the only feasible channel will be BEC and BSC, respectively.

We will see how the set of basis channels facilitates solving such complicated optimization problems. The extremes are stated in the following theorems:

Theorem 4.2 [Constrained Minimum of E_r]**:** Among all the equal-capacity MBIOS channels with probability of error greater (resp. less) than P_0, the minimum of E_r for every rate below the capacity is achieved by $\mathsf{g}_{x^*,\frac{1}{2}}$ (resp. $\mathsf{g}_{\eta,\eta} = \mathrm{BSC}(\eta)$), where x^* is the unique solution of

$$h(x) = 1 - C\frac{1-2x}{1-2P_0},\ x \in (0,\eta). \tag{4.1}$$

Theorem 4.3 [Constrained Maximum of E_r]**:** Among all the equal-capacity MBIOS channels with probability of error less (resp. greater) than P_0, the maximum of E_r for every rate below the capacity is achieved by g_{0,y^*} (resp. $\mathsf{g}_{0,\frac{1}{2}} = \mathrm{BEC}(H)$), where y^* is the unique solution of

$$h(y) = y\frac{H}{P_0},\ y \in (P_0,\tfrac{1}{2}). \tag{4.2}$$

Proof: [Theorem 4.2 and 4.3] See Section 4.4 □

In Fig. 4.1, the error exponents for various channels from $\mathcal{S}(0.4)$ are depicted. According to Theorem 4.1, the BEC and BSC have the maximum and minimum of E_r, respectively. By Theorem 4.2 and 4.3, if the set of channels is limited to those whose error probability is equal to 0.1921, i.e., the error probability of the BIAWGN$\in \mathcal{S}(0.4)$, then the maximum and minimum of E_r are attained by the basis channels g_{0,y^*} and $\mathsf{g}_{x^*,\frac{1}{2}}$, respectively. As it can be seen, there is a visual ordering among basis channel: as we move inward from the BEC (by increasing x and decreasing y), E_r gets smaller until we reach the BSC.

Critical rate

The stationary $\rho \in [0,1]$ for maximizing $E_0(\rho) - \rho R$ is obtained by

$$R = \frac{\partial E_0(\rho)}{\partial \rho}.$$

By [3, Theorem 5.6.3], $E_0(\rho, P_X)$ is a concave function for all $\rho \geq 0$ and P_X, particularly, for the optimal input distribution, i.e., $E_0(\rho) = E_0(\rho, P_X \equiv \frac{1}{2})$ is a concave function of ρ. This means that

$$\left.\frac{\partial E_0(\rho)}{\partial \rho}\right|_{\rho=1} \leq R \leq \left.\frac{\partial E_0(\rho)}{\partial \rho}\right|_{\rho=0} \text{[nats]}.$$

From Lemma 4.1 and Lemma C.1, we have

$$\begin{aligned}\left.\frac{\partial E_0(\rho)}{\partial \rho}\right|_{\rho=0} &= \frac{-1}{\mathbb{E}[\mu(P,0)]}\,\mathbb{E}\left[\lim_{\rho\to 0}\frac{\partial \mu(P,\rho)}{\partial \rho}\right] \\ &= \mathbb{E}[1-h(P)]\log 2 \\ &= C\log 2 \text{ [nats]},\end{aligned}$$

where C is the capacity in bits per channel use. The point $\left.\frac{\partial E_0(\rho)}{\partial \rho}\right|_{\rho=1}$ is called *critical rate* of the channel and is denoted by R_{cr}. Note that for $0 \leq R \leq R_{cr}$, $\rho = 1$ is the optimal choice and

$$E_r(R) = E_0(1) - R.$$

In other words, the slope of the exponent is -1 for $0 \leq R \leq R_{cr}$, increases monotonically for $R_{cr} \leq R \leq C$ [nats] and approaches zero as R approaches C [53].

In the following Theorem, we find the extremes of critical rate over $\mathcal{S}(C)$.

Theorem 4.4 [Extremes and Monotonicity of $R_{cr,x,y}$]**:** The maximum (minimum) of critical rate over the set of MBIOS channels with capacity C is attained by the BEC (BSC). Furthermore, let $R_{cr,x,y}$ be the critical rate of $\mathsf{g}_{x,y}$. The following monotonicity holds:

$$\frac{\partial R_{cr,x,y}}{\partial x} \leq 0, \;\; \frac{\partial R_{cr,x,y}}{\partial y} \geq 0.$$

Proof: By Lemma C.1, we obtain

$$R_{cr} = \frac{-1}{\mathbb{E}[\mu(P,1)]}\,\mathbb{E}\left[\lim_{\rho\to 1}\frac{\partial \mu(P,\rho)}{\partial \rho}\right] = \frac{\mathbb{E}[\mu^{\circ}(P)]}{\mathbb{E}[\mu_{\circ}(P)]},$$

where $\mu_{\circ}(p) = 1 + 2\sqrt{p(1-p)}$, and

$$\begin{aligned}\mu^{\circ}(p) &= -2 \times \lim_{\rho\to 1}\frac{\partial \mu(P,\rho)}{\partial \rho} \\ &= \left(\sqrt{p}+\sqrt{\bar{p}}\right)\left(\sqrt{p}\log\sqrt{p}+\sqrt{\bar{p}}\log\sqrt{\bar{p}}\right) + \left(\sqrt{p}+\sqrt{\bar{p}}\right)^2 \log\frac{2}{\sqrt{p}+\sqrt{\bar{p}}}.\end{aligned}$$

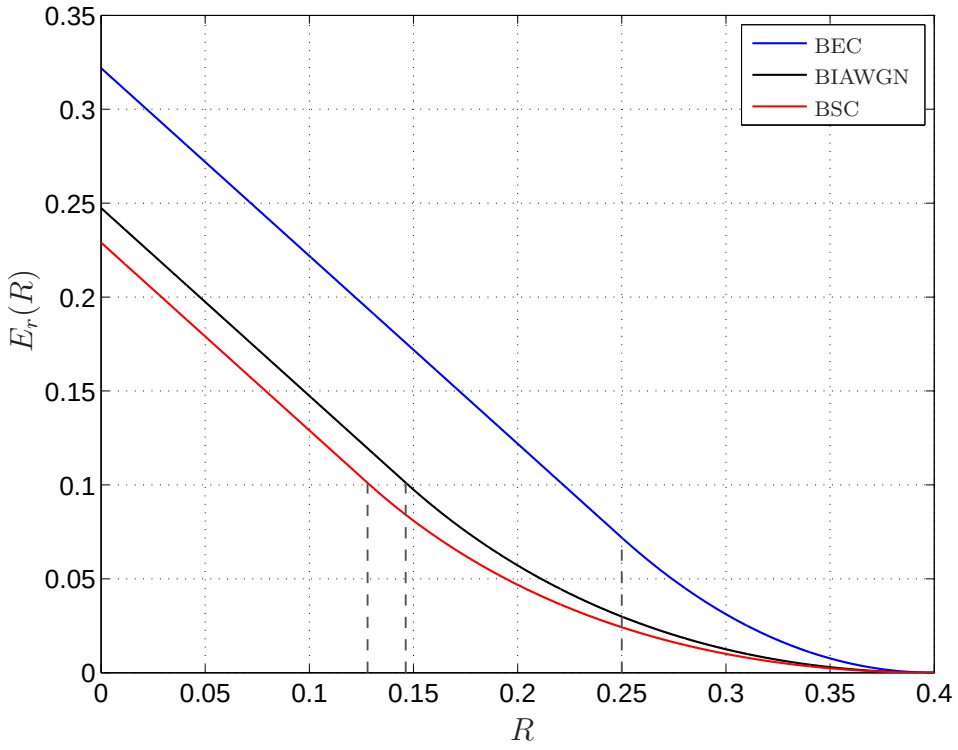

Figure 4.2: Comparison of critical rates for three channels from $\mathcal{S}(0.4)$. Dashed lines show the critcal rates.

Unlike the previous cases where we optimized a single expectation, in this case we are dealing with a ratio of expectations. According to Theorem 3.1, the system $\{1, h, \mu^\circ\}$ is a T–system implying that[4]

$$\mathbb{E}_{\text{BSC}}[\mu^\circ(P)] \leq \mathbb{E}[\mu^\circ(P)] \leq \mathbb{E}_{\text{BEC}}[\mu^\circ(P)].$$

Moreover, by taking $s = \frac{1}{2}$ in $\mathcal{U}_6$ of Theorem 3.1 and subtracting one, the system $\{1, h, -\mu_\circ\}$ is a T–system , i.e.,

$$\mathbb{E}_{\text{BEC}}[\mu_\circ(P)] \leq \mathbb{E}[\mu_\circ(P)] \leq \mathbb{E}_{\text{BSC}}[\mu_\circ(P)].$$

Since $\mu_\circ$ is strictly positive, we conclude that the BEC and BSC are the extremes. To show the monotonicity, we observe that

$$R_{cr,x,y} = \frac{\mathbb{E}_{x,y}[\mu^\circ(P)]}{\mathbb{E}_{x,y}[\mu_\circ(P)]}$$

[4] $\mathbb{E}_{\text{BSC}} \equiv \mathbb{E}_{\eta,\eta}$ and $\mathbb{E}_{\text{BEC}} \equiv \mathbb{E}_{0,\frac{1}{2}}$.

and

$$\frac{\partial R_{cr,x,y}}{\partial x} = \frac{1}{\mathbb{E}^2_{x,y}[\mu_\circ(P)]}\Big(\mathbb{E}_{x,y}[\mu_\circ(P)] \underbrace{\frac{\partial}{\partial x}\mathbb{E}_{x,y}[\mu^\circ(P)]}_{\leq 0} - \mathbb{E}_{x,y}[\mu^\circ(P)] \underbrace{\frac{\partial}{\partial x}\mathbb{E}_{x,y}[\mu_\circ(P)]}_{\geq 0}\Big) \leq 0,$$

where we used the monotonicity properties of the basis channels and the fact that μ° and $\mu_\circ$ are positive[5]. A similar approach applies to the derivative with respect to y.

□

Shulman-Feder bound (SFB)

So far, we have considered the *random* coding error exponent. What happens if the code has some specific structure, e.g., a linear block code? This leads to a natural question: *How does one modify the random error exponent for a linear block code to incorporate the specific structure?*

Let $\mathcal{C}$ be a specific binary linear code of length n and rate R. Also, let $\{A_\ell\}_{i=0}^n$ be its weight distribution, i.e.,

$$A_\ell = |\{\text{codewords with Hamming distance} = \ell\}|.$$

According to the SFB [54], the average block error probability of $\mathcal{C}$ over an MBIOS channel is bounded by

$$P_e \leq \exp\left(-nE_r\left(R + \log\frac{\alpha(\mathcal{C})}{n}\right)\right),$$

where E_r is the random coding error exponent and $\alpha(\mathcal{C})$ is a function of the distance spectrum of the code. It is the maximal ratio of the distance spectrum of a code and the average distance spectrum. In other words,

$$\alpha(\mathcal{C}) = \max_{1\leq\ell\leq n} \frac{A_\ell}{2^{nR}-1}\frac{2^n}{\binom{n}{\ell}},$$

[5]To see this, note that

$$\mu^\circ(p) = \left(\sqrt{p}+\sqrt{\bar{p}}\right)\left[\left(\sqrt{p}\log\sqrt{p}+\sqrt{\bar{p}}\log\sqrt{\bar{p}}\right) + \left(\sqrt{p}+\sqrt{\bar{p}}\right)\log\frac{2}{\sqrt{p}+\sqrt{\bar{p}}}\right].$$

A simple application of Jensen's inequality with function $x\log x$, $x\in[0,\frac{1}{2}]$ to the expression in the brackets shows positivity.

where $\frac{\binom{n}{\ell}}{2^n}$ is the probability of choosing a codeword of length n with the weight ℓ using a uniform distribution, and $\frac{A_\ell}{2^{nR}-1}$ is the same probability under the assumption that we choose randomly one of the (nonzero) codewords in the code [54]. Therefore, if the code's spectrum is close to the spectrum of a random code, then we have $\alpha(\mathcal{C}) \approx 1$ and we get the same random coding error exponent.

Corollary 4.1 [Extremes of SFB]**:** The extremes of the SFB of a linear code $\mathcal{C}$ over the MBIOS channels in $\mathcal{S}(C)$ is the same as the extremes of the random coding error exponent for any rate R such that $R + \log \frac{\alpha(\mathcal{C})}{n} < C$.

The SFB can be generalized to an ensemble of (rather than a specific) binary linear codes. Let $\bar{\mathcal{C}}$ denote an ensemble of linear codes. Let $\mathcal{U} \in \{1, 2, \ldots, n\}$ and its complement $\mathcal{U}^c = \{1, 2, \ldots, n\} \backslash \mathcal{U}$. According to Miller and Burshtein [55, Theorem 1], the block error probability under maximum likelihood decoding is bounded by

$$P_e \leq \sum_{\ell \in \mathcal{U}} \bar{A}_\ell B^\ell + 2^{-nE_r\left(R + \log \frac{\alpha(\bar{\mathcal{C}})}{n}\right)},$$

where B is the Bhattacharyya parameter of the channel, $\bar{A}_\ell$ denotes the average number of codewords of Hamming weight ℓ in the ensemble and

$$\alpha(\bar{\mathcal{C}}) = \max_{\ell \in \mathcal{U}^c} \frac{\bar{A}_\ell}{2^{nR} - 1} \frac{2^n}{\binom{n}{\ell}}.$$

Corollary 4.2 [Extremes of Miller-Burshtein Bound]**:** The extremes of the Miller-Burshtein bound of an ensemble of linear codes $\bar{\mathcal{C}}$ over the MBIOS channels in $\mathcal{S}(C)$ is the same as the extremes of the random coding error exponent for any rate R such that $R + \log \frac{\alpha(\bar{\mathcal{C}})}{n} < C$.

4.2.2 Expurgated Exponent

The construction of the error exponent $E_r(R)$ is based on the argument of random codeword generation which chooses codewords independently according to some fixed input distribution. This means that there exist poor codewords with high probability of error in an ensemble of random codes. It turns out

that such poor codewords have an adverse effect on the random coding error exponent at low rates. By expurgating those codewords, we can reduce their contribution in the overall probability of error [3]. The following theorem establishes existence of the expurgated ensemble.

Theorem 4.5 [Expurgated Exponent [3, Theorem 5.7.1]]**:** For any DMC channel, let n be a positive integer and R be any positive number. There exist codes of length n and rate R nats such that for $m = 1, \ldots, e^{nR}$,

$$P_{e,m} \leq \exp\left(-nE_{ex}\left(R + \frac{\log 4}{n}\right)\right),$$

where the function E_{ex} is called the *expurgated* error exponent and is given by

$$E_{ex}(R_0) = \sup_{\rho \geq 1} \left(E_x(\rho) - \rho R_0\right), \tag{4.3}$$

$E_x(\rho) = \sup_{P_X} E_x(\rho, P_X)$, and

$$E_x(\rho, P_X) = -\rho \log \sum_{x,x'} P_X(x) P_X(x') \left(\int \sqrt{P_{Y|X}(y|x) P_{Y|X}(y|x')} dy\right)^{\frac{1}{\rho}}.$$

It has been shown that the expurgated error exponent is greater than the random error exponent at low rates which means that it provides a tighter lower bound on the channel reliability function. In the case of MBIOS channels, we can find the extremes of the expurgated error exponent by identifying the proper T–systems.

Theorem 4.6 [Extremes of E_{ex}]**:** For an MBIOS channel with a Bhattacharyya parameter B,

$$E_x(\rho) = -\log\left(\tfrac{1}{2} + \tfrac{1}{2} B^{\frac{1}{\rho}}\right)^{\rho}, \ \rho \geq 1.$$

Moreover, among all MBIOS channels of capacity C, the maximum (resp. minimum) of E_{ex} is achieved by the BEC (resp. BSC).

Proof: Consider an MBIOS channel with capacity C. Let $w = P_X(+1)$. For

$x \neq x'$, we have

$$\begin{aligned}\int \sqrt{P_{Y|X}(y|x)P_{Y|X}(y|x')}dy &= \int \sqrt{\frac{P_{Y|X}(y|-1)}{P_{Y|X}(y|+1)}}P_{Y|X}(y|+1)dy \\ &= \mathbb{E}[e^{-\frac{L}{2}}] \\ &= \mathbb{E}[2\sqrt{p(1-p)}] = B.\end{aligned}$$

Therefore,

$$\begin{aligned}\exp\left(-E_x(\rho, P_X)/\rho\right) &= w^2 + (1-w)^2 + 2w(1-w)B^{\frac{1}{\rho}} \\ &= 1 + 2w(1-w)\left(B^{\frac{1}{\rho}} - 1\right),\end{aligned}$$

which is symmetric with respect to $w \to 1-w$. Since $B \leq 1$, $w = \frac{1}{2}$ maximizes $E_x(\rho, P_X)$, thus

$$E_x(\rho) = -\log\left(\tfrac{1}{2} + \tfrac{1}{2}B^{\frac{1}{\rho}}\right)^{\rho}.$$

According to Theorem 3.1 (take $s = \frac{1}{2}$ in $\mathcal{U}_6$), the maximum and minimum of the Bhattacharyya parameter over all MBIOS channels with the same capacity is achieved by BSC and BEC, respectively. Thus, the maximum and minimum of E_x (for all $\rho \geq 1$) and E_{ex} is achieved by the BEC(H) and BSC(η), respectively.

□

The analysis of critical rate for the expurgated error exponent is similar to the random coding error exponent. Solving (4.3) for R_0 gives the parametric representation of rate in terms of E_x as

$$R_0 = \frac{\partial E_x(\rho)}{\partial \rho} = \log\frac{2}{1 + B^{\frac{1}{\rho}}} + \frac{\log B}{\rho\left(1 + B^{\frac{-1}{\rho}}\right)}.$$

Simple calculus shows that $E_x(\rho)$ is a concave function of ρ:

$$\frac{\partial^2 E_x(\rho)}{\partial \rho^2} = -\frac{B^{\frac{1}{\rho}}\log^2 B}{\rho^3\left(1 + B^{\frac{1}{\rho}}\right)^2} \leq 0,$$

which means that

$$0 = \lim_{\rho\to\infty}\frac{\partial E_x(\rho)}{\partial \rho} \leq R_0 \leq \left.\frac{\partial E_x(\rho)}{\partial \rho}\right|_{\rho=1},$$

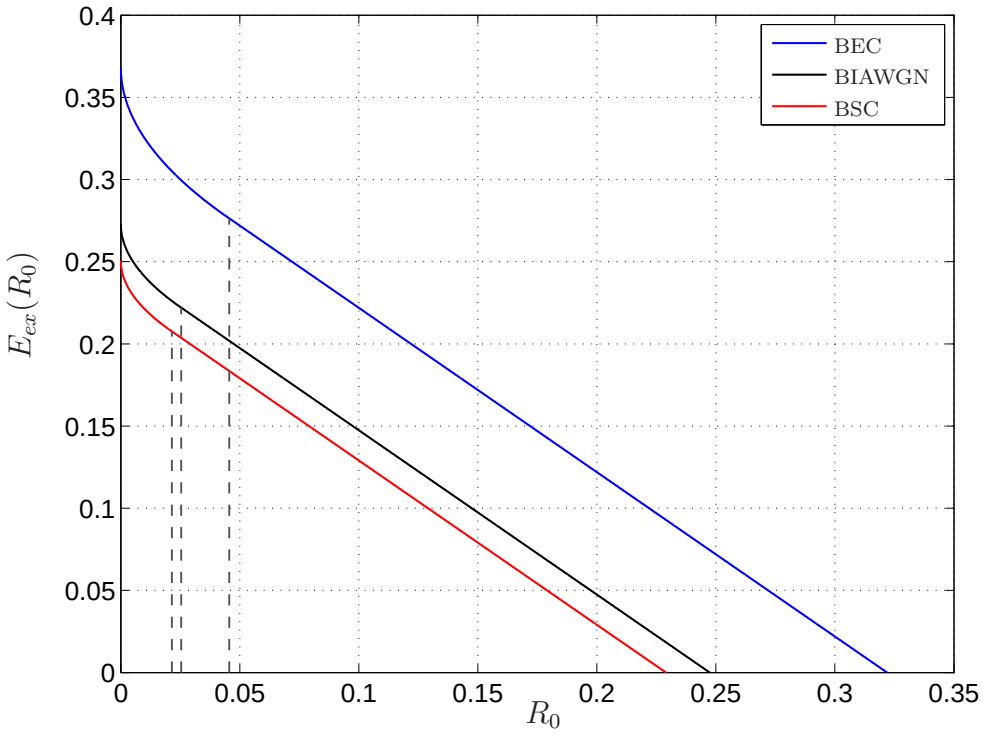

Figure 4.3: Comparison of expurgated error exponents for MBIOS channels of capacity $C = 0.4$. Dashed lines indicate the expurgated critical rate. R_0 is in bits.

where $R_{cr,ex} = \frac{\partial E_x(\rho)}{\partial \rho}\big|_{\rho=1}$ is the expurgated critical rate. Therefore, the expurgated error exponent monotonically decreases for $0 \leq R_0 \leq R_{cr,ex}$ and decreases linearly as $E_x(1) - R_0$ for $R_{cr,ex} \leq R_0 \leq E_x(1)$, where $E_x(1)$ is the R_0-intercept and is equal to $\log \frac{2}{1+B}$. The expurgated critical rate can be expressed as

$$R_{cr,ex} = \log \frac{2}{1+B} + \frac{B \log B}{1+B}.$$

Theorem 4.7 [Extremes of Expurgated Critical Rate]**:** Among all MBIOS channels of capacity C, the BEC (BSC) maximizes (minimizes) the expurgated critical rate.

Proof: The proof is straightforward by considering the facts that the maximum (minimum) of the Bhattacharyya parameter over all MBIOS channels

with the same capacity is achieved by the BSC (BEC), $0 < B < 1$ and

$$\frac{\partial R_{cr,ex}}{\partial B} = \frac{\log B}{(1+B)^2} < 0.$$

□

Fig. 4.3 compares E_{ex} for three channels from $\mathcal{S}(0.4)$. As it can be seen from the figure, BEC (BSC) maximizes (minimizes) both the expurgated error exponent and the expurgated critical rate.

4.2.3 Sphere-Packing Exponent

So far, we have only considered the lower bounds on the reliability function. Sphere-packing bound provides an upper bound for the reliability function. The following theorem summarizes the sphere-packing bound.

Theorem 4.8 [Sphere-Packing Bound [3]]**:** Let $P_{Y|X} : \mathcal{X} \mapsto \mathcal{Y}$ be a DMC. Any $(n, R \text{ [nats]}, P_e)$ code satisfies

$$P_e \geq \exp\left(-n(E_{sp}(R - o_1(n)) + o_2(n))\right),$$

where

$$E_{sp}(R) = \sup_{\rho>0}(E_0(\rho) - \rho R),$$

and $E_0(\rho)$ is given in Section 4.2.1. Moreover,

$$o_1(n) = \frac{\log 8}{n} + \frac{|\mathcal{X}| \log n}{n},$$
$$o_2(n) = \frac{\log 8}{n} + \sqrt{\frac{2}{n}} \log \frac{e^2}{P_{\min}},$$

where $P_{\min} = \min\{P_{Y|X}(y|x) : P_{Y|X}(y|x) > 0\}$.

Similar to random-coding error exponent, E_{sp} is a non-negative convex function of R. In fact, using the same analysis as in Section 4.2.1, one can see that for symmetric channels

$$0 = \lim_{\rho\to\infty} \frac{\partial E_0(\rho)}{\partial \rho} \leq R \leq \left.\frac{\partial E_0(\rho)}{\partial \rho}\right|_{\rho=0} = C \text{ [nats]}.$$

The extremes of sphere-packing error exponent is given by the following theorem. The proof is straightforward using the results of Section 4.2.1.

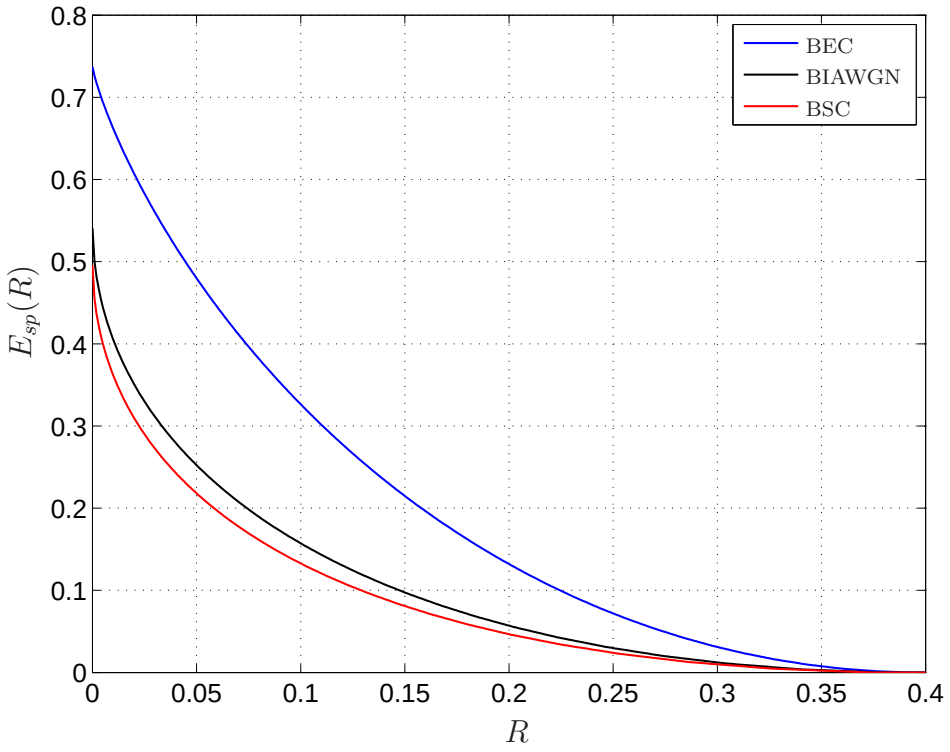

Figure 4.4: Comparison of sphere-packing exponents for MBIOS channels of capacity $C = 0.4$. R is in bits.

Theorem 4.9 [Extremes of Sphere-Packing Bound]**:** Among all equal-capacity MBIOS channels with capacity C, the minimum (maximum) of the sphere-packing error exponent $E_{sp}(R)$ for every rate below the capacity is achieved by BEC(H) (resp. BSC(η)).

Sphere-packing error exponents for MBIOS channels of capacity 0.4 bits per channel use are compared in Fig. 4.4

4.2.4 Erasure/List Decoding Exponent

In this section, we will consider two decoders, in addition to the ordinary decoder of the classic setup in Section 4.1. These decoders prove beneficial when the transmitted data contains some redundancy, when a feedback channel is available, or when further stages of coding (concatenation) are contemplated [56].

1. The decoder has the option of not deciding at all, of rejecting all esti-

mates. The resulting output is called an erasure. We have an undetected error only if the decoder makes an estimate, and it is wrong.

2. The decoder has the option of putting out more than one estimate. The resulting output is called a list. Only if the correct codeword is not on the list do we have a list error [56].

Let $\mathcal{R}_m$, defined over the space of channel outputs, be the region such that if the channel output falls into it, the decoder declares the message $m \in \{1, \ldots, M\}$ as an estimate of what has been sent. In an ordinary decoder, the decision regions are disjoint and occupies the whole space (the channel output falls into only one of the regions). With the erasure option, the decision regions are still disjoint but there are channel outputs that belong to none of the regions. With the list option, the decision regions overlap and a channel output may belong to a list of regions.

Let E_2 be the event of undetected errors. The probability of undetected errors of an ordinary decoder is

$$\mathbb{P}(E_2) = \sum_m \sum_{\mathbf{y} \in \mathcal{R}_m} \sum_{m' \neq m} P_{X^n, Y^n}(\mathbf{c}_{m'}, \mathbf{y}),$$

which is minimized by the maximum a posteriori decoding rule. The performance of a decoder with erasure is characterized in terms of probability of undetected errors and probability of erasure. Define E_1 as an event which the received $\mathbf{y}$ does not fall in the decision region $\mathcal{R}_m$, corresponding to the transmitted code word $\mathbf{c}_m$ [56]. When E_1 occurs, either an undetected error or an erasure occurs. Therefore, the probability of E_1 is

$$\mathbb{P}(E_1) = \sum_m \sum_{y \notin \mathcal{R}_m} P_{X^n, Y^n}(\mathbf{c}_m, \mathbf{y}) = \mathbb{P}(E_2) + \mathbb{P}(E),$$

where E is the erasure event. There is a fundamental tradeoff between these two probabilities; letting more erasure happen on unreliable receptions will decrease the probability of undetected errors; equivalently, there is a tradeoff between $\mathbb{P}(E_1)$ and $\mathbb{P}(E_2)$. The optimal decision criterion can be obtained using the Neyman-Pearson lemma [2].

Theorem 4.10 [Forney [56, Theorem 1]]**:** Let T be an arbitrary number. There is no set of decision regions other than

$$\mathcal{R}_m = \left\{ \mathbf{y} : \frac{P_{X^n,Y^n}(\mathbf{c}_m, \mathbf{y})}{\sum_{m' \neq m} P_{X^n,Y^n}(\mathbf{c}_{m'}, \mathbf{y})} \geq e^{nT} \right\}, \tag{4.4}$$

which gives both a lower $\mathbb{P}(E_1)$ and $\mathbb{P}(E_2)$ than this region does.

The arbitrary parameter T governs the relative magnitudes of $\mathbb{P}(E_1)$ and $\mathbb{P}(E_2)$; clearly as T increases, $\mathbb{P}(E_1)$ increases while $\mathbb{P}(E_2)$ decreases, since the decision regions $\mathcal{R}_m$ shrink. One can observe that T must be positive, in order for the decision regions to be necessarily disjoint [56].

The performance of the decoder with list option is characterized via two parameters: the probability of list error where the transmitted codeword is not in the list and the average number of incorrect messages on the list. It is shown in [56] that with a slight modification (allowing the summation in $\mathbb{P}(E_2)$ to be performed over the overlapped regions), these parameters are respectively given by $\mathbb{P}(E_1)$ and $\mathbb{P}(E_2)$. Note that in this case the optimum tradeoff is still given by Theorem 4.10 with the exception of T being negative to allow the regions overlap.

Similar to an ordinary decoder, one can find the error exponents for $\mathbb{P}(E_1)$ and $\mathbb{P}(E_2)$ using the technique of Gallager [3]. The following theorem by Forney states so.

Theorem 4.11 [Forney [56, Theorem 2]]**:** There is a block code of length n and rate R nats such that when the likelihood ratio criterion of (4.4) is used with a threshold T, one can simultaneously obtain

$$\mathbb{P}(E_1) < \exp\left(-nE_1(R,T)\right),$$
$$\mathbb{P}(E_2) < \exp\left(-nE_2(R,T)\right),$$

where $E_1(R,T)$ is given at high rates by

$$E_1(R,T) = \sup_{0 \leq s \leq \rho \leq 1} (E_0(s,\rho) - \rho R - sT),$$

where $E_0(s,\rho) = \max_{P_X} E_0(s,\rho,P_X)$, and

$$E_0(s,\rho,P_X) = -\log \int \left(\sum_x P_X(x) P_{Y|X}(y|x)^{1-s} \right) \left(\sum_{x'} P_X(x') P_{Y|X}(y|x')^{\frac{s}{\rho}} \right)^{\rho} dy,$$

and at low rates by

$$E_1(R,T) = \sup_{s\in[0,1],\rho\geq 1} (E_x(s,\rho) - \rho R - sT),$$

where $E_x(s,\rho) = \max_{P_X} E_x(s,\rho,P_X)$, and

$$E_x(s,\rho,P_X) = -\rho \log \sum_x \sum_{x'} P_X(x) P_X(x') \left(\int P_{Y|X}(y|x)^{1-s} P_{Y|X}(y|x')^s dy \right)^{\frac{1}{\rho}},$$

and $E_2(R,T) = E_1(R,T) + T$.

The following theorem gives the extremes of $E_1(R,T)$ and $E_2(R,T)$ over the set of symmetric channels of capacity C.

Theorem 4.12 [Extremes of E_1 and E_2]**:** For any MBIOS channel, we have

$$E_0(s,\rho) = -\log \mathbb{E}[\mu(P,s,\rho)],$$

where

$$\mu(p,s,\rho) = 2^{-\rho} \left(p^{1-s} + (1-p)^{1-s} \right) \left(p^{\frac{s}{\rho}} + (1-p)^{\frac{s}{\rho}} \right)^{\rho}, \ 0 \leq s \leq \rho \leq 1.$$

and

$$E_x(s,\rho) = -\log \left(\tfrac{1}{2} + \tfrac{1}{2} \mathbb{E}^{\frac{1}{\rho}} \left[P^s (1-P)^{1-s} + P^{1-s}(1-P)^s \right] \right)^{\rho}, \ \rho \geq 1, \ s \in [0,1].$$

Moreover, among all the MBIOS channels with capacity C, the maximum (minimum) of $E_1(R,T)$ and $E_2(R,T)$ is attained by the BEC (BSC) for all rates $0 \leq R < C$ [nats] and arbitrary T.

Proof: Using the techniques used in Lemma 4.1 and by letting $w = P_X(+1)$, one can see that at high rates

$$\begin{aligned}
\exp(-E_0(s,\rho,P_X)) &= \mathbb{E}\left[\left(w + (1-w) e^{-(1-s)L} \right) \left(w + (1-w) e^{-\frac{s}{\rho} L} \right)^{\rho} \right] \\
&= \mathbb{E}\Big[\left(wP^{1-s} + (1-w)(1-P)^{1-s} \right) \left(wP^{\frac{s}{\rho}} + (1-w)(1-P)^{\frac{s}{\rho}} \right)^{\rho} \\
&\quad + \left(w(1-P)^{1-s} + (1-w)P^{1-s} \right) \left(w(1-P)^{\frac{s}{\rho}} + (1-w)P^{\frac{s}{\rho}} \right)^{\rho} \Big]
\end{aligned}$$

will not change if $w \to 1-w$. In Appendix B, it is proved that the function inside the brackets will be minimized by a uniform P_X, i.e., $w = \frac{1}{2}$. Therefore, we obtain

$$E_0(s,\rho) = -\log \mathbb{E}\left[2^{-\rho}\left(P^{1-s}+(1-P)^{1-s}\right)\left(P^{\frac{s}{\rho}}+(1-P)^{\frac{s}{\rho}}\right)^{\rho}\right].$$

Similarly, for low rates we have

$$\exp(-E_x(s,\rho,P_X)/\rho) = w^2+(1-w)^2+w(1-w)\left(\mathbb{E}^{\frac{1}{\rho}}[e^{-sL}]+\mathbb{E}^{\frac{1}{\rho}}[e^{-(1-s)L}]\right),$$

which is symmetric with respect to $w \to 1-w$, hence minimized by a uniform input distribution. Thus,

$$E_x(s,\rho) = -\log\left(\tfrac{1}{2}+\tfrac{1}{2}\,\mathbb{E}^{\frac{1}{\rho}}\left[P^s(1-P)^{1-s}+P^{1-s}(1-P)^s\right]\right)^{\rho}.$$

The extreme part is proved through Lemma 4.1 and Theorem 3.1 ($\mathcal{U}_6$ and $\mathcal{U}_2$ with $a_1 = 1-s$, $a_2 = \frac{s}{\rho}$, $b_1 = 1$ and $b_2 = \rho$).

□

4.2.5 Channel Dispersion

As we mentioned in Section 4.1, channel dispersion plays an important role in the non-asymptotic analysis of $R^*(n,P_e)$ and $P_e^*(n,R)$. Polyanskiy *et al.* defined channel dispersion (measured in squared information units per channel use) as

$$\begin{aligned} V &= \lim_{P_e\to 0}\limsup_{n\to\infty} n\left(\frac{C-R^*(n,P_e)}{Q^{-1}(P_e)}\right)^2 \\ &= \lim_{P_e\to 0}\limsup_{n\to\infty} n\frac{\left(C-R^*(n,P_e)\right)^2}{2\log\frac{1}{P_e}}. \end{aligned} \tag{4.5}$$

The rationale for this definition is the following expansion, valid for a number of different channels (the $P_e > 0$ is fixed and $n \to \infty$) [46]:

$$R^*(n,P_e) = C - \sqrt{\frac{V}{n}}Q^{-1}(P_e) + o(n).$$

In this section, we explore the extremes of channel dispersion among channels with a given capacity. The main key to our analysis is studying the second derivative of the reliability function. An alternative way to obtain the results

of this section is the representation of the channel dispersion via the variance of information density. This definition is from [47,48]. This approach is explored in Appendix D.

Recall the definition of channel dispersion given in (4.5). At rates close to capacity, when n is large and P_e is small enough, one can obtain the following approximation:

$$V \approx \frac{(C-R)^2}{2E(R)}.$$

The fact that the reliability function behaves parabolically for rates near the capacity was known to Shannon (see Fig. 18 of [46]). Therefore, channel dispersion is obtained by the second derivative of the reliability function at the capacity [46] as

$$V = \lim_{R \to C} \left(\frac{\partial^2}{\partial R^2} E(R) \right)^{-1}.$$

Lemma 4.2 [Channel Dispersion]**:** For an MBIOS channel, channel dispersion in squared bits per channel use is

$$V = \mathbb{E}[\nu(P)] - H^2,$$

where

$$\nu(p) = p \log_2^2 p + (1-p) \log_2^2 (1-p), \ p \in [0, \tfrac{1}{2}].$$

Proof: For rates above the critical rate, the reliability function is known and is equal to the random coding error exponent. We have[6]

$$V = \lim_{R \to C} \left(\frac{\partial^2}{\partial R^2} E_r(R) \right)^{-1} = - \lim_{\rho \to 0} \frac{\partial^2}{\partial \rho^2} E_0(\rho).$$

[6]We have $R = \frac{\partial E_0}{\partial \rho}$ and $\frac{\partial R}{\partial \rho} = \frac{\partial^2 E_0}{\partial \rho^2}$. Moreover, since $E_r(R) = E_0(\rho) - \rho \frac{\partial E_0}{\partial \rho}$, we have

$$\frac{\partial E_r}{\partial \rho} = -\rho \frac{\partial^2 E_0}{\partial \rho^2}, \ \frac{\partial E_r}{\partial R} = \frac{\frac{\partial E_r}{\partial \rho}}{\frac{\partial R}{\partial \rho}} = -\rho,$$

which gives [3]

$$\frac{\partial^2 E_r}{\partial R^2} = -\frac{\partial \rho}{\partial R} = -\left(\frac{\partial^2 E_0}{\partial \rho^2} \right)^{-1}.$$

According to Appendix C.1, one can now obtain

$$\begin{aligned} V &= \lim_{\rho\to 0} \frac{1}{\mathbb{E}^2[\mu(P,\rho)]} \left(\mathbb{E}[\mu(P,\rho)] \frac{\partial^2}{\partial\rho^2} \mathbb{E}[\mu(P,\rho)] - \left(\frac{\partial}{\partial\rho} \mathbb{E}[\mu(P,\rho)] \right)^2 \right) \\ &= \frac{1}{\mathbb{E}^2[\mu(P,0)]} \left(\mathbb{E}[\mu(P,0)] \, \mathbb{E}\left[\lim_{\rho\to 0} \frac{\partial^2}{\partial\rho^2} \mu(P,\rho) \right] - \mathbb{E}^2 \left[\lim_{\rho\to 0} \frac{\partial}{\partial\rho} \mu(P,\rho) \right] \right) \\ &= \mathbb{E}\left[\lim_{\rho\to 0} \frac{\partial^2}{\partial\rho^2} \mu(P,\rho) \right] - \mathbb{E}^2 \left[\lim_{\rho\to 0} \frac{\partial}{\partial\rho} \mu(P,\rho) \right], \end{aligned}$$

resembling a variance expression. Also, we have

$$\begin{aligned} \lim_{\rho\to 0} \frac{\partial^2}{\partial\rho^2} \mu(p,\rho) &= \log^2 2 + p \log^2 p + (1-p) \log^2 (1-p) - 2h(p) \log^2 2 \\ &= \log^2 2 \, (1 + \nu(p) - 2h(p)) \, , \end{aligned}$$

where

$$\nu(p) = p \log_2^2 p + (1-p) \log_2^2 (1-p), \ p \in [0, \tfrac{1}{2}],$$

and

$$\lim_{\rho\to 0} \frac{\partial}{\partial\rho} \mu(p,\rho) = -\log 2 (1 - h(p)).$$

Therefore, channel dispersion in bits squared per channel use is

$$V = 1 + \mathbb{E}[\nu(P)] - 2\,\mathbb{E}[h(P)] - C^2 = \mathbb{E}[\nu(P)] - H^2.$$

□

For a BSC with the cross probability $p \in [0,1]$, it can be shown that the dispersion is

$$V(p) = p(1-p) \log_2^2 \frac{p}{1-p}.$$

One might speculate that since every MBIOS channel can be expressed as a stochastic BSC, the channel dispersion for an arbitrary MBIOS channel can be calculated by the expected value of $V(P)$. This is in fact not true and $V \neq \mathbb{E}[V(P)]$ because the variance operator is not linear on the information density (see Appendix D). Clearly, BSC is the only MBIOS channel with such a property. In fact, Lemma 4.2 shows that $V = \mathbb{E}[\nu(P) - H^2]$ for all MBIOS channels.

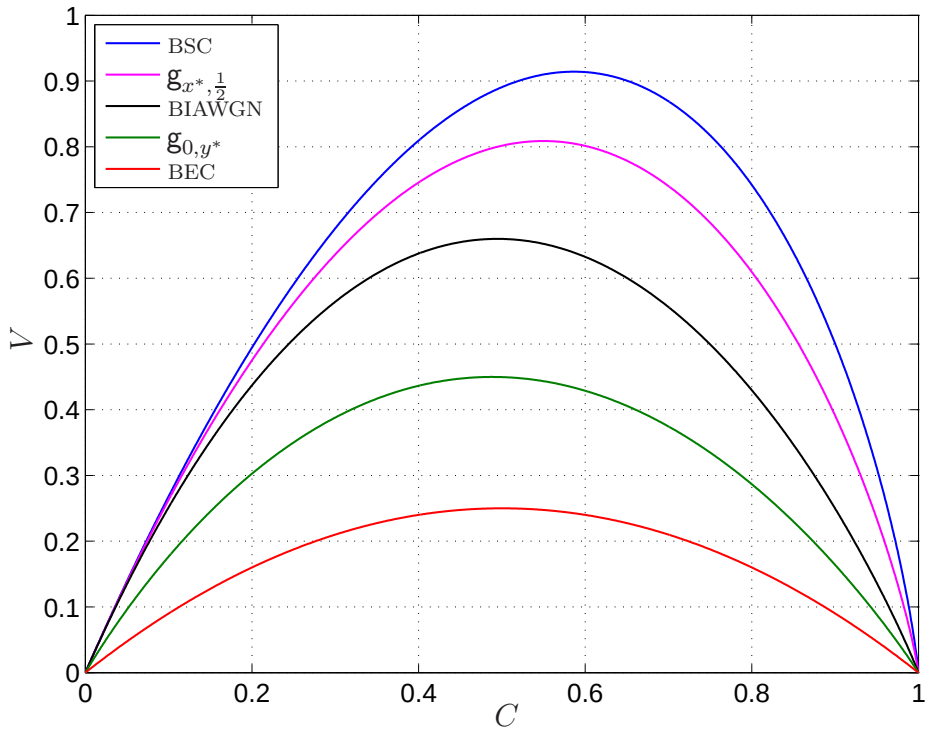

Figure 4.5: Sketch of channel dispersion versus the capacity. BSC and BEC are the extremes.

The following theorem summarizes the extremes of channel dispersion. It is important to note that by Remark 2.6 and Theorem 3.1, the extremes of channel dispersion (with or without the constraint of probability of error) are the same as the extremes of $\mathbb{E}[\mu(P,\rho)]$.

Theorem 4.13 [Extremes of Channel Dispersion]**:** Among all MBIOS channels of capacity C, the maximum (minimum) of the channel dispersion is achieved by the BSC (BEC). Let x^* and y^* be defined as (4.1) and (4.2). When there is a probability of error constraint, the extremes will be modified as:

1. $\mathbb{E}[P] \geq P_0$: The maximum and minimum of channel dispersion are given by $\mathsf{g}_{x^*,\frac{1}{2}}$ and BEC(H), respectively.

2. $\mathbb{E}[P] \leq P_0$: The maximum and minimum of channel dispersion are given by BSC(η) and g_{0,y^*}, respectively

Channel dispersion for several channels are depicted in Fig. 4.5 versus the

capacity. For each value of capacity, BSC and BEC are the ultimate extremes. We fix $P_0 = Q\left(\frac{1}{\sigma}\right)$ where σ is the standard deviation of the BIAWGN with the given capacity. Among all MBIOS channels with the given capacity and given probability of error, $\mathsf{g}_{x^*,\frac{1}{2}}$ and g_{0,y^*} are the extremes. We have seen this behaviour when we studied the extremes of $\mathbb{E}[\mu(P,\rho)]$ with capacity and probability of error constraints. As we mentioned before, this similarity comes from Remark 2.6 and shows the power of Chebychev systems in analyzing such optimization problems.

Remark 4.1: As we mentioned in Section 4.1, it is proved that for a large code length, $R^*(n,P_e) \approx C - \sqrt{V/n}Q^{-1}(P_e)$. This means that to achieve a fraction δ of capacity, i.e., $R = \delta C$, at probability of decoding error of P_e, the required code length is approximately

$$n \gtrapprox \left(\frac{Q^{-1}(P_e)}{1-\delta}\right)^2 \frac{V}{C^2}.$$

For a fixed probability of error and a given capacity, it can be deduced from Theorem 4.13 that BSC needs the largest code length and BEC requires the shortest codewords. We are going to compare the require code length to achieve $\delta = 80\%$ of the capacity of the BSC, BIAWGN and BEC from $\mathcal{S}(0.5)$ at 10^{-3} probability of error. Approximately, the required code length is 240 for BEC, 630 for BIAWGN and 850 for BSC to satisfy the desired performance. For a detailed study, one can see [46] (see Figures 1, 3 and 6).

4.3 Discussions and Summary

Discussion 4.1 [*Handling other constraints*]**:** We observed in this chapter that one can solve extremal problems of error exponents when a constraint on the channel's probability of error is imposed. It can be proved that other constraints can be handled in a similar way. For example, the maximum and minimum of random coding error exponents among all channels from $\mathcal{S}(C)$ whose Bhattacharyya parameter (or MMSE) satisfies some constraint can be obtained by identifying the proper T–system. Results, similar to Theorem 4.2 and 4.3, are possible to obtain.

Discussion 4.2 [*Constraint on probability of error*]**:** Although we only solved the extremal problems with probability of error constraint for random coding error exponent, it is possible to obtain similar results for other exponents we considered. The same ordering among basis channels holds for other exponents as well. For brevity, we did not mention such results.

In this chapter, we solved the extremal problems of various error exponents over the set of MBIOS channels. We showed that BSC and BEC are the two extreme channels of several error exponents, which bound the reliability function. Based on this evidence, we conjecture the following:

Conjecture 4.1 [Extremes of Reliability Function]**:** Among all MBIOS channels of capacity C, the maximum and minimum of the channel reliability function

$$E(R) = \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^*(n,R),$$

for any rate below the capacity are achieved by the BEC and BSC, respectively.

Note that the channel reliability function is known for $R_{cr} \leq R \leq C$ and the exponents we considered bound it for $0 \leq R \leq R_{cr}$. It was shown that the set of basis channels with their monotonicity properties is a powerful tool to analyze problems of error exponents. We believe that this method can be used with other extremal problems of symmetric densities provided that the proper T–system is identified correctly.

4.4 Proof of Theorem 4.2 and 4.3

First, we solve the following related optimization problem:

$$O_= : \begin{array}{ll} \text{opt} & E_r(R) \\ \text{s.t.} & \mathbb{E}[1] = 1 \\ & \mathbb{E}[h(P)] = H \\ & \mathbb{E}[P] = \tilde{P} \end{array}$$

First, note that if $\tilde{P} = \eta$ ($\tilde{P} = H/2$), then only the BSC (BEC) will satisfy the constraints. Thus, we exclude the BSC and BEC from the search space

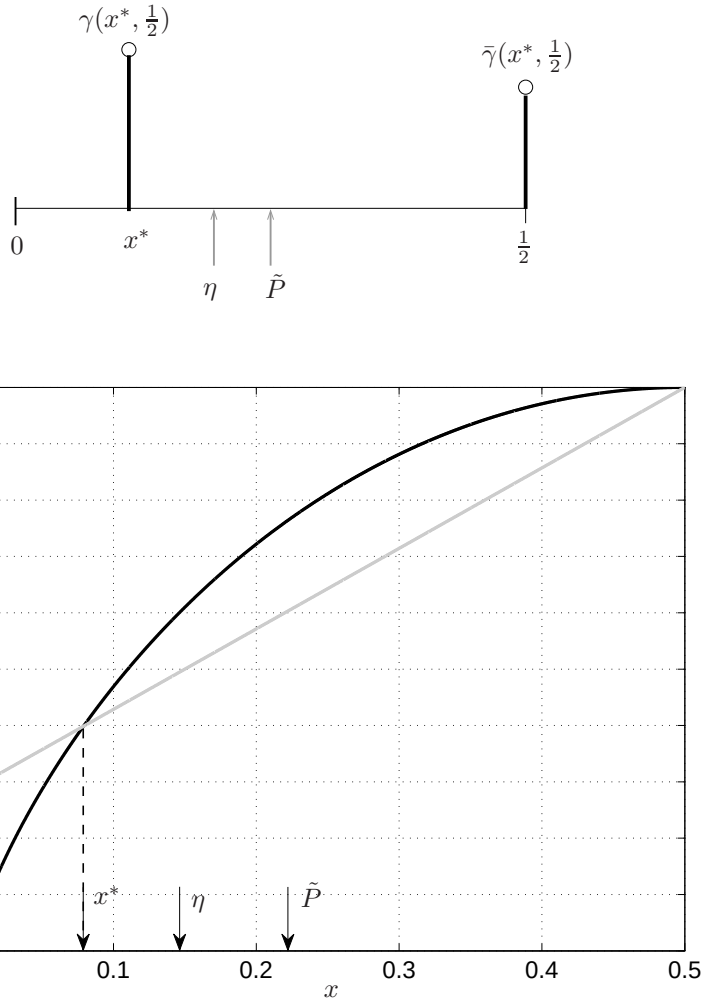

Figure 4.6: A sample plot of $h(x) = 1 - C\frac{1-2x}{1-2\tilde{P}}$ for $C = 0.4$ and $\tilde{P} = 0.22$. The intersection point is $x^* = 0.0790 < \eta = 0.1461$.

by assuming $\eta < \tilde{P} < H/2$ or equivalently, $2\tilde{P} < H < h(\tilde{P})$[7]. Next, according to Theorem 3.1, $u_0(p) = 1$, $u_1(p) = h(p)$, $u_2(p) = p$ form a T–system. The point $\mathbf{c} = (1, H, \tilde{P})$ cannot be a boundary point of $\mathcal{M}_3$ because if it was, by Lemma 2.1, it would admit a representation of index not greater than two. This implies that one of the following cases will happen:

1. A representation of index one corresponds to either a useless channel, i.e., $\mathsf{g}(p) = \Delta_{\frac{1}{2}}(p)$, or a perfect channel, i.e., $\mathsf{g}(p) = \Delta_0(p)$, either of which is not possible as $H \in (0, 1)$.

[7]If $\tilde{P} > H/2$ or $\tilde{P} < \eta$, the constraint on the probability of error is automatically relaxed. In this case, Theorem 4.1 is applied.

2. A representation of index two corresponds to either a BSC ($H = h(\tilde{P})$) or a BEC ($H = 2\tilde{P}$).

Thus, $\mathbf{c}$ is an inner point of $\mathcal{M}_3$. By Theorem 3.1, the augmented system with $u_3(p) = \mu(p, \rho)$ is a T–system. Therefore, according to Theorem 2.4, $\sup_{\Sigma(\mathbf{c})} \mathbb{E}[\mu(P, \rho)]$ is achieved by the distribution associated with the upper principle representation of $\mathbf{c}$, i.e., a mass point at $p = \frac{1}{2}$ and a mass point at an interior of $[0, \frac{1}{2}]$. Since this representation must exhibit capacity C, we conclude that the optimizing channel is the basis channel

$$\mathsf{g}_{x^*,\frac{1}{2}}(p) = \gamma(x^*, \tfrac{1}{2})\Delta_{x^*}(p) + \bar{\gamma}(x^*, \tfrac{1}{2})\Delta_{\frac{1}{2}}(p),$$

where $x^* \in (0, \eta)$ is obtained by replacing $\gamma(x^*, \frac{1}{2})$ from (3.3) in the constraint on the probability of error, i.e.,

$$\gamma(x^*, \tfrac{1}{2})x^* + \frac{1}{2}\bar{\gamma}(x^*, \tfrac{1}{2}) = \tilde{P},$$

and is given in (4.1). To show that (4.1) has a unique solution in the interval $(0, \eta)$, let $f(x) = h(x) - 1 + C\frac{1-2x}{1-2\tilde{P}}$. f is a continuous function in $[0, \frac{1}{2}]$ and by (3.1), $f(0) = \frac{2\tilde{P}-H}{1-2\tilde{P}} < 0$, and $f(\eta) = 2C\frac{\tilde{P}-\eta}{1-2\tilde{P}} > 0$. According to the intermediate value theorem, there exists $x^* \in (0, \eta)$ such that $f(x^*) = 0$. Therefore, (4.1) has a unique solution in $(0, \eta)$. The line given in (4.1) passes the point $(\frac{1}{2}, 1)$ and its slope is $\frac{2C}{1-2\tilde{P}} > 2$ as shown in Fig. 4.6.

On the other hand, $\inf_{\Sigma(\mathbf{c})} \mathbb{E}[\mu(P, \rho)]$ is attained by the distribution associated with the lower principle representation of $\mathbf{c}$, i.e., a mass point at $p = 0$ and a mass point at an interior point of $[0, \frac{1}{2}]$. Similarly, the optimizing channel is the following basis channel

$$\mathsf{g}_{0,y^*}(p) = \gamma(0, y^*)\Delta_0(p) + \bar{\gamma}(0, y^*)\Delta_{y^*}(p),$$

where $y^* \in (\eta, \frac{1}{2})$ is obtained by plugging $\gamma(0, y^*)$ from (3.3) into the constraint on the probability of error. The solution of (4.2) is unique in $(\tilde{P}, \frac{1}{2})$. To see this, let $f(y) = y\frac{H}{\tilde{P}} - h(y)$. f is a continuous function in $[\tilde{P}, \frac{1}{2}]$ with $f(\tilde{P}) = H - h(\tilde{P}) < 0$ and $f(\frac{1}{2}) = \frac{H}{2\tilde{P}} - 1 > 0$. Thus, by the intermediate value theorem, there exists $y^* \in (\tilde{P}, \frac{1}{2})$ such that $f(y^*) = 0$. Therefore, (4.2) has a unique solution in $(\tilde{P}, \frac{1}{2})$. A sample plot is shown in Fig. 4.7.

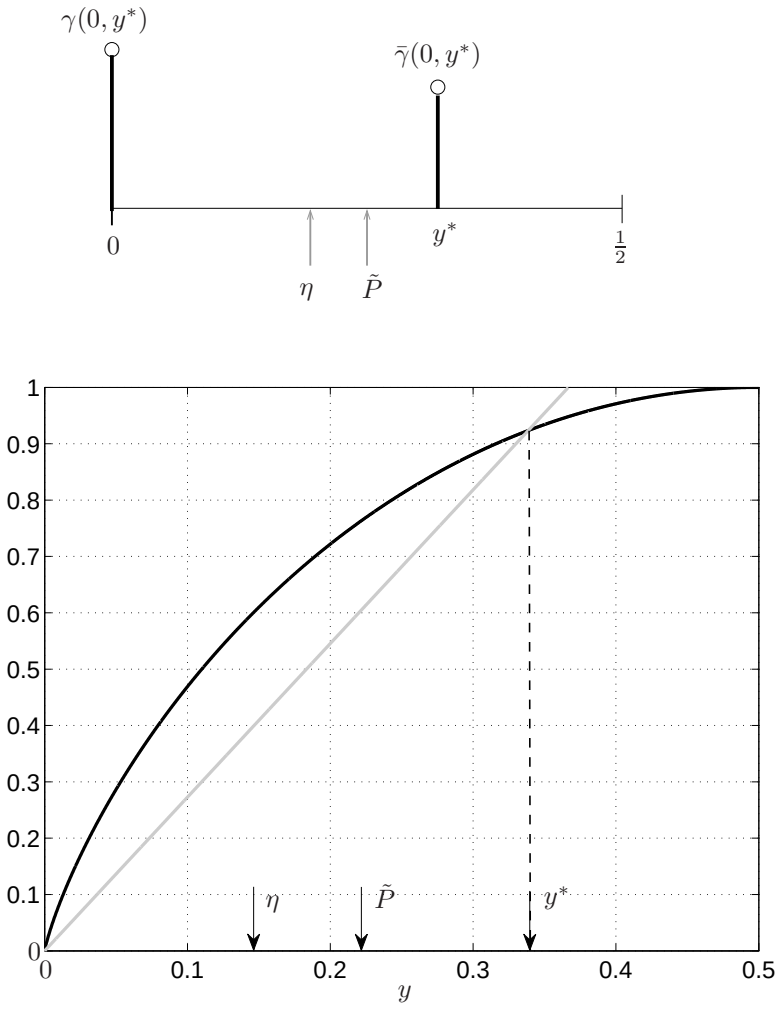

Figure 4.7: A sample plot of $h(y) = y\frac{H}{\tilde{P}}$ for $C = 0.4$ and $\tilde{P} = 0.22$. The intersection point is $y^* = 0.3386 > \tilde{P} = 0.22$.

It is important to note that the optimizing channels are independent of ρ which means that they give the extremes of $O_=$ too. Now, we prove Theorem 4.2 and 4.3. Since there is a same logic behind both proofs, we only prove Theorem 4.2, i.e.,

$$O_{\geq}: \begin{array}{ll} \sup & \mathbb{E}[\mu(P,\rho)] \\ \text{s.t.} & \mathbb{E}[1] = 1 \\ & \mathbb{E}[h(P)] = H \\ & \mathbb{E}[P] \geq P_0. \end{array}$$

Consider a collection of optimization problems $O_=$ indexed by $\tilde{P}$, where $\tilde{P} \in [P_0, \frac{H}{2}]$ (see (3.1)). Let $\mathsf{g}_{x^*(\tilde{P}),\frac{1}{2}}$ be the maximizer of $\mathbb{E}[\mu(P,\rho)]$ where

$x^*(\tilde{P})$ is the unique solution of

$$h(x) = 1 - C\frac{1-2x}{1-2\tilde{P}}$$

in $(0, \eta)$. This means that solving $O_{\geq}$ is accomplished by maximizing $\mathbb{E}[\mu(P,\rho)]$ over the set of basis channels

$$\left\{ \mathtt{g}_{x^*(\tilde{P}),\frac{1}{2}} : \tilde{P} \in [P_0, \tfrac{H}{2}] \right\},$$

where $\tilde{P} = \mathbb{E}_{x^*(\tilde{P}),\frac{1}{2}}[P]$. It can be seen that $\mathbb{E}_{x^*(\tilde{P}),\frac{1}{2}}[\mu(P,\rho)]$ is a decreasing function of $\tilde{P}$ because according to Theorem 3.1 and Lemma 3.2

$$\begin{aligned}
\frac{d\,\mathbb{E}_{x^*(\tilde{P}),\frac{1}{2}}[\mu(P,\rho)]}{d\tilde{P}} &= \frac{\partial\,\mathbb{E}_{x^*(\tilde{P}),\frac{1}{2}}[\mu(P,\rho)]}{\partial x^*(\tilde{P})} \left(\frac{d\tilde{P}}{dx^*(\tilde{P})} \right)^{-1} \\
&= \underbrace{\frac{\partial\,\mathbb{E}_{x^*(\tilde{P}),\frac{1}{2}}[\mu(P,\rho)]}{\partial x^*(\tilde{P})}}_{>0} \left(\underbrace{\frac{\partial\,\mathbb{E}_{x^*(\tilde{P}),\frac{1}{2}}[P]}{\partial x^*(\tilde{P})}}_{<0} \right)^{-1} < 0, \qquad (4.6)
\end{aligned}$$

Therefore the maximizer of $O_{\geq}$ is obtained when $\tilde{P} = P_0$, i.e., $\mathtt{g}_{x^*(P_0),\frac{1}{2}} = \mathtt{g}_{x^*,\frac{1}{2}}$.

For the other case, we wish to solve

$$O_{\leq} : \begin{array}{ll} \sup & \mathbb{E}[\mu(P,\rho)] \\ \text{s.t.} & \mathbb{E}[1] = 1 \\ & \mathbb{E}[h(P)] = H \\ & \mathbb{E}[P] \leq P_0. \end{array}$$

The same argument holds again: solving $O_{\leq}$ is accomplished by maximizing $\mathbb{E}[\mu(P,\rho)]$ over the set of basis channels

$$\left\{ \mathtt{g}_{x^*(\tilde{P}),\frac{1}{2}} : \tilde{P} \in [\eta, P_0] \right\},$$

where $\tilde{P} = \mathbb{E}_{x^*(\tilde{P}),\frac{1}{2}}[P]$. Thus, according to (4.6), $O_{\leq}$ is maximized when $P_0 = \eta$, i.e., $\mathtt{g}_{\eta,\frac{1}{2}} \equiv$BSC$(\eta)$[8].

[8]It is easy to see that $\gamma(\eta, \frac{1}{2}) = 1$ and there will be no mass point at $p = \frac{1}{2}$.

Chapter 5

Capacity of Duplication Channels

5.1 Introduction

5.1.1 Channels with Synchronization Errors

A channel with synchronization errors is defined by Dobrushin [28] as a channel that independently transforms every input symbol to a word of random (possibly zero) length. As it was mentioned in Chapter 1, since insertion/deletion channels are generally difficult to deal with, there have been various attempts to understand them by studying subclasses of such channels that ease the analysis. Most of the research on channels with synchronization errors have been on the deletion channel which deletes input bits independently with some probability. For a general survey on deletion channels, the reader is referred to [8].

In this chapter, we study the *independent and identically distributed (i.i.d.) duplication channel*, the simplest insertion channel, which duplicates each input symbol independently with a certain probability. We focus on the binary case where the duplication probability is denoted by p and capacity is shown by $C = C(p)$. Generalization of the results to the non-binary cases is straightforward.

Sticky channels is a subclass of insertion channels where each symbol is duplicated multiple times according to a probability distribution on positive integers. These channels are studied in [57] where numerical upper and lower

bounds on the capacity of sticky channels are presented. For a single duplication channel (the channel of our interest), it is trivial that an i.i.d. Bernoulli($\frac{1}{2}$) process achieves the capacity of one bit per channel use when the duplication probability is either $p = 0$ or $p = 1$. It has been pointed out in [57] that numerical optimization shows that for values of p close to either 0 or 1, the optimal distribution is very close to the i.i.d. Bernoulli($\frac{1}{2}$) process. We shall prove this observation by providing a series expansion of the capacity around $p = 0$ of the form

$$C = 1 + p \log_2 p + \alpha_1 p + \alpha_2 p^2 + \cdots$$

We show that an i.i.d. Bernoulli($\frac{1}{2}$) process achieves the capacity up to term p. Furthermore, we find tight lower bounds on α_2 and conjecture the optimal process to improve the series expansion up to term p^2. Capacity behaviour when $p \to 1$ will be discussed and some achievable rates will be presented. We will show that when $p \to 1$, duplication capacity takes a series expansion of the form

$$C = 1 + 2(1-p)^2 \log_2(1-p) + \beta_2 (1-p)^2 + \cdots$$

5.1.2 A Naive Approach Towards the Capacity

Dobrushin [28] showed that channels with synchronization errors are information stable and their capacity is expressed as:

$$C = \lim_{n \to \infty} C_n, \quad C_n = \sup_{P_{X^n}} \frac{1}{n} I(X^n; Y(X^n)), \tag{5.1}$$

where $X^n = (X_1, \ldots, X_n)$ is passed through the channel and $Y(X^n)$ is the received sequence of symbols upon sending X^n. Since the length of the received sequence is a random variable, we denote the output sequence by $Y(X^n)$. Also, $I(X^n; Y(X^n))$ denotes the mutual information between X^n and $Y(X^n)$ [2]. It is easy to check that if we eliminate all synchronization errors, (5.1) reduces to $C = \sup_{P_X} I(X;Y)$.

Let $\mathcal{X} = \{0, 1\}$ be the channel input set and W_n be a channel that accepts n input bits and duplicates each bit independently with probability p, i.e.,

$$W_n : \{0,1\}^n \mapsto \bigcup_{k=n}^{2n} \{0,1\}^k,$$

where the number of output bits varies between n and $2n$. According to (5.1), finding the duplication capacity is equivalent to finding the limit of $\frac{1}{n}C(W_n)$ as $n \to \infty$, where $C(W_n) = nC_n$ indicates the capacity of W_n.

Define $q = 1 - p$. Let us examine a few of these channels. The capacity of W_1, whose transition matrix is

$$\mathbf{W}_1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 00 & 11 \end{array} \\ \begin{bmatrix} q & 0 & p & 0 \\ 0 & q & 0 & p \end{bmatrix} \end{array},$$

is one bit per channel use, which is achieved by a uniform distribution on $\mathcal{X}$. Similarly, the transition matrix of W_2 is as follows:

$$\mathbf{W}_2 = \begin{array}{c} \\ 00 \\ 01 \\ 10 \\ 11 \end{array} \begin{array}{c} \begin{array}{cccccccccccccc} 00 & 01 & 10 & 11 & 000 & 001 & 011 & 100 & 110 & 111 & 0000 & 0101 & 1010 & 1111 \end{array} \\ \begin{bmatrix} q^2 & 0 & 0 & 0 & 2pq & 0 & 0 & 0 & 0 & 0 & p^2 & 0 & 0 & 0 \\ 0 & q^2 & 0 & 0 & 0 & pq & pq & 0 & 0 & 0 & 0 & p^2 & 0 & 0 \\ 0 & 0 & q^2 & 0 & 0 & 0 & 0 & pq & pq & 0 & 0 & 0 & p^2 & 0 \\ 0 & 0 & 0 & q^2 & 0 & 0 & 0 & 0 & 0 & 2pq & 0 & 0 & 0 & p^2 \end{bmatrix} \end{array}$$

Note that there are elements of $\bigcup_{k=n}^{2n}\{0,1\}^k$ which will not be produced during the duplication process of $\mathcal{X}^n$, hence not shown in the transition matrix. Clearly, each possible output of W_2 is only connected to one of the inputs which shows that $C(W_2) = 2$ bits per channel use. However, as we proceed with W_3, the capacity is no longer trivial, because, for example, 01111 can be produced from either 011 (one duplication pattern) or 0111 (three possible duplication patterns). In this case, $C(W_3)$ depends on the value of p. To obtain the capacity numerically, one can use the well-known Blahut-Arimoto algorithm [58, 59]. However, the size of $\mathbf{W}_n$ grows exponentially with n which makes the calculation of $C(W_n)$ infeasible for $n > 10$.

Although duplication capacity is expressed as (5.1), it is difficult to obtain its *single-letter* characterization $C(p)$ from (5.1).

It is shown in [60] that in fact $C = \inf_{n\geq 1} C_n$, thus $\frac{1}{n}C(W_n) = C_n$ for $n \geq 1$ upper bounds the duplication capacity.

Remark 5.1: In the case of deletion channels, one can define a similar set of channels $\{V_n\}_{n\geq 1}$, where $V_n : \{0,1\}^n \mapsto \bigcup_{k=0}^{n}\{0,1\}^k$ accepts n input bits and deletes each bit with some probability. The maximum value of n for which computing $C(V_n)$ is feasible is significantly larger than the one for duplication channels. This is because of the fact that the size of $\mathbf{V}_n$ is of order $O(2^n)$ compared to $O(2^{2n})$ for $\mathbf{W}_n$. Studying the capacity of V_n for $n = 1, 2, \ldots$ provides upper bounds on deletion capacity. This approach is thoroughly studied by Fertonani and Duman [15] by introducing various genie-aided scenarios.

5.1.3 Notations Guide

Binary processes are shown by letters $\mathbb{X}, \mathbb{Y}, \ldots$[1] For a binary sequence, a *run* of zeros (ones) is defined as the maximal chunk of zeros (ones) bordered by ones (zeros). $H(\mathbb{X})$ stands for the entropy rate of the process $\mathbb{X}$ [2]. $\mathcal{S}$ designates the set of stationary and ergodic processes. For a sequence of random variables, $X^n = (X_1, \ldots, X_n)$ and $X_n^m = (X_n, \ldots, X_m)$. Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be two discrete probability measures defined on the same probability space. We denote the total variation distance between $\mathbb{P}_1$ and $\mathbb{P}_2$ by

$$\| \mathbb{P}_1 - \mathbb{P}_2 \|_{\text{TV}} = \frac{1}{2} \sum_x |\mathbb{P}_1(x) - \mathbb{P}_2(x)| .$$

Also, the Kullback–Leibler divergence between $\mathbb{P}_1$ and $\mathbb{P}_2$ is [2]

$$D(\mathbb{P}_1 \,\|\, \mathbb{P}_2) = \sum_x \mathbb{P}_1(x) \log_2 \frac{\mathbb{P}_1(x)}{\mathbb{P}_2(x)}.$$

A binomial distribution of size n and success probability p is shown by $B(n, p)$. The big-O and little-o notations are defined as follows:

- Let $f : [0,1] \to \mathbb{R}$ and $g : [0,1] \to \mathbb{R}_+$ be two functions. We write $f(p) = O(g(p))$ if there exists $M > 0$ such that $|f(p)| \leq Mg(p)$ for all $p \in [0,1]$.

[1] We follow the notations of [60].

- Let $f : \mathbb{N} \to \mathbb{R}$ and $g : \mathbb{N} \to \mathbb{R}_+$. We write $f(n) = o(g(n))$ as $n \to \infty$ if and only if
$$\limsup_{n\to\infty} \left| \frac{f(n)}{g(n)} \right| = 0.$$

5.2 Setup

Let $\mathbb{X}$ be a binary process $\{X_n\}_{-\infty}^{\infty}$ as the input to the duplication channel. We are going to study such binary sequences in terms of their runs. We characterize the duplication channel by an i.i.d. binary process $\mathbb{Q}$ indicating whether the corresponding bit of $\mathbb{X}$ is duplicated ($Q_n = 1$) or not ($Q_n = 0$). The probability of duplication is denoted by $p \in [0, 1]$ i.e., for all n, $\mathbb{P}(Q_n = 1) = p$.

The binary process $\mathbb{Y} = \mathbb{Y}(\mathbb{X}, \mathbb{Q})$ is the result of passing $\mathbb{X}$ through the duplication channel. Table 5.1 shows an example of such processes. Although not present in reality, a character "|" is used to distinguish consecutive runs. As it can be seen, an input bit is duplicated wherever the corresponding bit in $\mathbb{Q}$ is one. The bits introduced by the channel are shown in gray. Because of the nature of the duplication process, we can map each run of $\mathbb{Y}$ uniquely to a run in $\mathbb{X}$. Assume that the vector X^n is passed through the duplication channel. The length of the received sequence $Y(X^n)$ will be

$$M_n = n + \sum_{k=1}^{n} \mathbf{1}_{\{Q_k=1\}}. \tag{5.2}$$

Thus, $M_n - n \sim B(n, p)$ and $\mathbb{E}[M_n] = n(1 + p)$.

The mutual information between $\mathbb{X}$ and $\mathbb{Y}$ is denoted by

$$I(\mathbb{X}) = \lim_{n\to\infty} \frac{I_n}{n}, \quad I_n = I(X^n; Y(X^n)).$$

For any $\mathbb{X} \in \mathcal{S}$, the process $\mathbb{Y}$ is also stationary and ergodic as $\mathbb{Q}$ is an i.i.d process. We denote runs in $\mathbb{X}$ and $\mathbb{Y}$ by $\mathcal{L}$ and $\mathcal{T}$, respectively, and their lengths by L and T where according to (5.2), $T = L + G$ and G is a binomial $B(L, p)$ variable.

For a stationary and ergodic process $\mathbb{X}$, we indicate the limit of the empirical distribution of run lengths by $P_L(\ell)$, $\ell = 1, 2, \ldots$ Following the notation

Table 5.1: An example of processes $\mathbb{X}$, $\mathbb{Q}$, and $\mathbb{Y}$.

$\mathbb{X}$	$\ldots\|000\|11\|0\|111\|00\|\ldots$
$\mathbb{Q}$	$\ldots 101\ 11\ 0\ 001\ 01\ldots$
$\mathbb{Y}$	$\ldots\|0000\|1111\|0\|1111\|000\|\ldots$

of [60], we call P_L the block perspective run length distribution. Assuming a finite average run-length $\mathbb{E}[L] < \infty$, the run length distribution of the run at position 0, denoted by L_0, is simply obtained by

$$P_{L_0}(\ell) = \frac{\ell P_L(\ell)}{\mathbb{E}[L]},\ \ell \in \mathbb{N} \tag{5.3}$$

and is called the bit perspective run length distribution [60]. Clearly, for the process $\mathbb{X}^*$, i.e., an i.i.d. Bernoulli($\frac{1}{2}$) process, the run length distribution is geometrically distributed. We denote the corresponding block and bit perspective run length distributions by $P_L^* = 2^{-\ell}$ and $P_{L_0}^* = \ell 2^{-(\ell+1)}$ for $\ell \geq 1$. Correspondingly, we denote the expected value with respect to the process $\mathbb{X}^*$ by $\mathbb{E}^*$. Moreover, it is not difficult to see that the run length of the received symbols is distributed as

$$P_T(\ell) = \sum_{k=\lceil \ell/2 \rceil}^{\ell} P_L(k) \binom{k}{\ell-k} p^{\ell-k}(1-p)^{2k-\ell},\ \ell \in \mathbb{N},$$

where we use the fact that $T = L + G$ and $G \sim B(L, p)$. The letters L, T, and G will be used for the random variables drawn according to such distributions.

Before presenting the main results, we generalize the statement in [60, Lemma II.2] to a general class of insertion/deletion channels. Let D be a random variable taking non-negative integer values with probability

$$\mathbb{P}(D = k) = d_k, \quad k = 0, 1, 2, \ldots$$

Consider a general class of insertion/deletion channels where each input bit is repeated multiple times based on an outcome of D, i.e., if $D = k$, the corresponding bit will be repeated k times. Note that $D = 0$ implies a deletion and $D = 2$ means a duplication. Also, for a channel without any synchronization errors, $D = 1$ all the times. The capacity of such channels can be calculated by (5.1). In the following lemma, we show that the set of stationary and ergodic

processes suffices to achieve the capacity of such channels. Moreover, the role of limit and supremum in (5.1) can be exchanged.

Lemma 5.1 [Sufficiency of Stationary Ergodic Processes]**:** Consider a memoryless channel that repeats each input bit multiple times according to an outcome of D, a random variable on non-negative integers such that $\mathbb{E}[D], H(D) < \infty$. There exists a binary stationary and ergodic process $\mathbb{X}$ that achieves the capacity of this channel. More precisely,

$$C = \sup_{\mathbb{X} \in \mathcal{S}} I(\mathbb{X}),$$

where

$$I(\mathbb{X}) = \lim_{n \to \infty} \frac{I_n}{n}, \quad I_n = I(X^n; Y(X^n)).$$

Proof: See Section 5.5.1. □

5.3 Main Results

In this section, we provide the main results regarding the capacity of duplication channels. We only state the results here and leave their technical proofs to Section 5.5.

First, we start by providing a lower bound to the capacity as a consequence of Lemma 5.1:

Theorem 5.1 [Information Rate of $\mathbb{X} \in \mathcal{S}$]**:** The capacity of a binary duplication channel is lower bounded by

$$I(\mathbb{X}) = \frac{H(T)}{\mathbb{E}[L]} - h(p) + \frac{\mathbb{E}\left[\log_2 \binom{L}{G}\right]}{\mathbb{E}[L]}, \tag{5.4}$$

where $\mathbb{X}$ can be any stationary and ergodic process with i.i.d. runs distributed according to some P_L defined on $\mathbb{N}$.

Proof: See Section 5.5.2. □

Next, the following theorem shows that the process $\mathbb{X}^*$, an i.i.d. Bernoulli($\frac{1}{2}$) process, achieves the capacity of a binary duplication channel for small values of duplication probability.

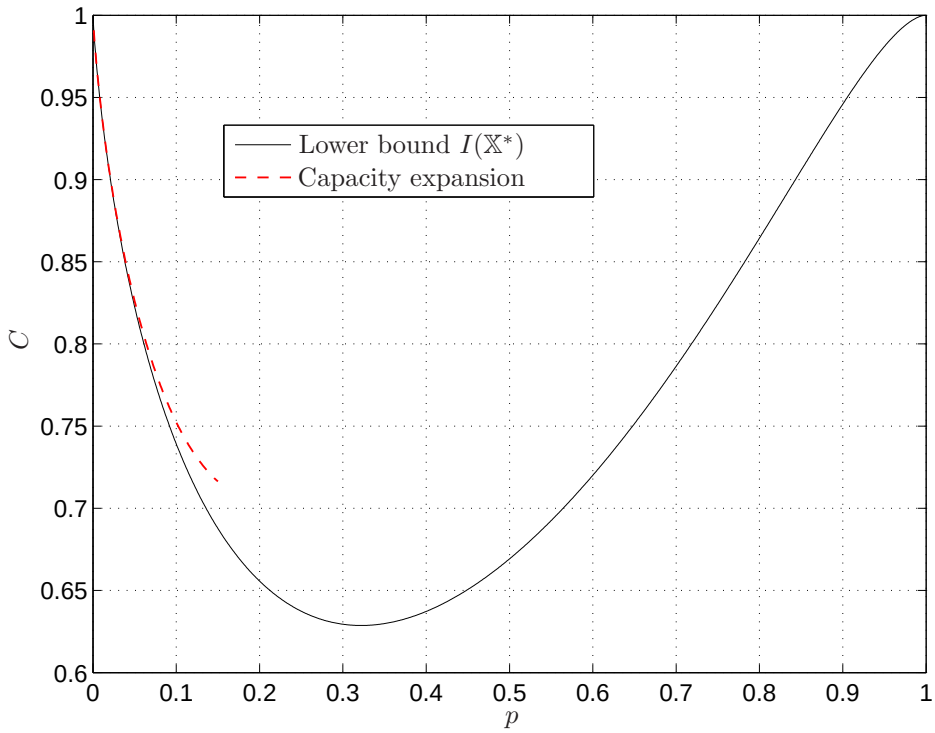

Figure 5.1: The comparison of the lower bound of Theorem 5.1 evaluated for the process $\mathbb{X}^*$ and the series expansion given in Theorem 5.2 without the $O(p^{3/2-\epsilon})$ term.

Theorem 5.2 [Capacity Expansion around $p = 0$]**:** The capacity of a binary duplication channel is

$$C = 1 + p \log_2 p + \alpha_1 p + O(p^{3/2-\epsilon})$$

for some small $\epsilon > 0$, where

$$\alpha_1 = \log_2(2/e) + \sum_{\ell=1}^{\infty} 2^{-(\ell+1)} \ell \log_2 \ell \approx 0.8458362348.$$

Furthermore, the capacity is achieved by the process $\mathbb{X}^*$.

Proof: See Section 5.5.3. □

Theorem 5.2 shows that the Bernoulli($\frac{1}{2}$) process is optimal for small duplication probabilities as conjectured in [57]. Fig. 5.1 compares the lower bound on the duplication capacity based on Theorem 5.1 for $\mathbb{X} = \mathbb{X}^*$ and the series expansion given in Theorem 5.2.

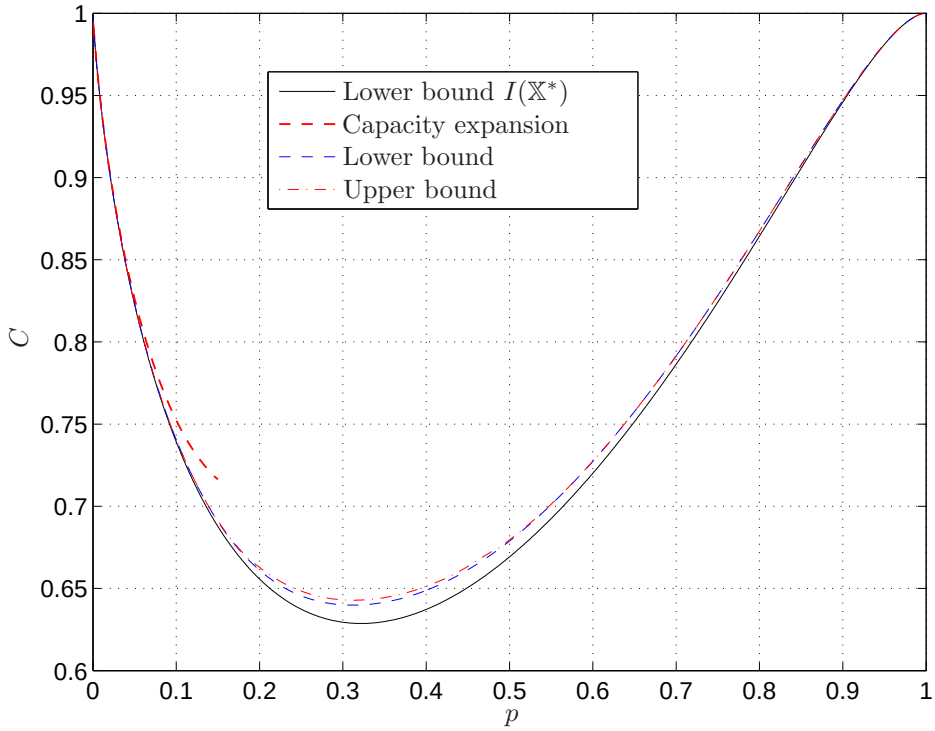

Figure 5.2: The comparison of $I(\mathbb{X}^*)$ and the bounds proposed by [57].

Discussion 5.1 [*Comparison with the deletion capacity*]**:** A comparison of the result of Theorem 5.2 with the result of [60] regarding the deletion channel shows that both capacity expansions are quite similar, except for the coefficient of the linear part. In fact, the capacity of a deletion channel is $C = 1+d\log_2 d-A_1 d+O(d^{3/2-\epsilon})$, where d is the deletion probability and $A_1 = \log_2(2e) - \sum_{\ell=1}^{\infty} 2^{-(\ell+1)}\ell\log_2\ell$ [60]. Moreover, $\alpha_1 + A_1 = 2$.

Discussion 5.2 [*Comparison with other bounds*]**:** To compare Theorem 5.2 with the numerical results reported by [57] for sticky channels, the upper and lower bounds proposed by [57] have to be implemented using the Jimbo-Kunisawa iterative algorithm [61]. These bounds are based on a mapping that transforms a sticky channel into a DMC with integer alphabets. Mitzenmacher in [57, Theorem 2.1] showed that the capacity of a sticky channel is equal to the capacity per unit cost of the corresponding integer-alphabet DMC, where the cost of sending ℓ bits is ℓ [62,63]. The comparison given in Fig. 5.2, shows that the numerical bounds are very tight and in agreement with Theorem 5.2.

5.3.1 Lower Bounds on α_2

By Theorem 5.2, we have obtained the capacity of duplication channels up to the first term in the series expansion. In this section, we provide lower bounds on α_2 by limiting $\mathbb{X}$ to specific *p-dependent* processes from $\mathcal{S}$. Then, we optimize these bounds to obtain tight lower bounds.

In this section, we use the fact that for an arbitrary run-length distribution P_L defined on $\mathbb{N}$ [60], we have

$$D(P_L \| P_L^*) = \sum_{\ell=1}^{\infty} P_L(\ell)(\ell + \log_2 P_L(\ell)) = \mathbb{E}[L] - H(L).$$

Therefore, we may write

$$\frac{H(T)}{\mathbb{E}[L]} = (1+p)\left(1 - \frac{D(P_T \| P_L^*)}{\mathbb{E}[T]}\right). \tag{5.5}$$

In what follows, to lower bound α_2, we restrict $\mathbb{X} \in \mathcal{S}$ to a class of stationary and ergodic perturbations in P_L^* and the class of symmetric first-order Markov processes. A similar approach is used for deletion channels in [64]. Although our methodology is totally different, we use the same symbols as in [64] to distinguish such processes.

Perturbation in P_L^*

In this section, by limiting the process $\mathbb{X} \in \mathcal{S}$ to the class of stationary and ergodic perturbations in P_L^*, we find a tight lower bound on α_2.

From the proof of Theorem 5.1 (see (5.14) and (5.15)), one can deduce that for any process $\mathbb{X} \in \mathcal{S}$ with non-i.i.d. runs we have

$$I(\mathbb{X}) < \frac{H(T)}{\mathbb{E}[L]} - h(p) + \frac{\mathbb{E}\left[\log_2 \binom{L}{G}\right]}{\mathbb{E}[L]} \leq C.$$

This shows that the capacity of duplication channels is achieved with a stationary and ergodic process with i.i.d. runs. For small duplication probabilities, the optimal run-length should not be very different than P_L^*. Let $P_L^\dagger(\ell) = 2^{-\ell}(1 + \epsilon_0(\ell, p))$ be a perturbation in P_L^* such that the following conditions hold:

$$\lim_{p \to 0} \epsilon_0(\ell, p) = 0, \ \ell \in \mathbb{N} \tag{5.6a}$$

$$\sum_{\ell=1}^{\infty} 2^{-\ell}\epsilon_0(\ell,p) = 0 \tag{5.6b}$$

Condition (5.6a) ensures that there is no perturbation at $p = 0$ and condition (5.6b) is required to make $P_L^\dagger$ a valid probability measure.

A Taylor series expansion of ϵ_0 around $p = 0$ gives $\epsilon_0(\ell, p) = p\epsilon(\ell) + O(p^2)$ where $\epsilon(\ell) = \lim_{p\to 0} \frac{\partial \epsilon_0}{\partial p}$. We take

$$P_L^\dagger(\ell) = 2^{-\ell}(1 + p\epsilon(\ell)),\ \ell \in \mathbb{N},$$

and let

$$\mathbb{E}^*[\epsilon(L)] = 0 \tag{5.7}$$

in order to satisfy (5.6b). We denote a process with i.i.d. run-lengths according to $P_L^\dagger$ with $\mathbb{X}^\dagger$. Similarly, $\mathbb{E}^\dagger$ is reserved for the expected value operator under $P_L^\dagger$. In order to find the tightest bound on α_2, we wish to find $\epsilon(\ell)$ such that it maximizes $I(\mathbb{X}^\dagger)$. The following theorem does so:

Theorem 5.3 [Optimal Perturbation]**:** Let $\mathbb{X}^\dagger$ be a stationary and ergodic process with i.i.d. runs distributed according to $P_L^\dagger(\ell) = 2^{-\ell}(1+p\epsilon(\ell))$, $\ell \in \mathbb{N}$, a perturbation in P_L^* such that $\mathbb{E}^*[\epsilon(L)] = 0$. The maximum of $I(\mathbb{X}^\dagger)$ over all eligible $\epsilon(\ell)$'s is achieved by

$$\epsilon_*(\ell) = \ell \log \ell - \ell(\tfrac{1}{2}\tau + 1) + 2,\ \ell \in \mathbb{N},$$

where $\tau = \mathbb{E}^*[L \log L] \approx 1.78628364$. Moreover, the maximum of $I(\mathbb{X}^\dagger)$ is equal to

$$I(\mathbb{X}^\dagger) = 1 + p\log_2 p + \alpha_1 p + \alpha_2^\dagger p^2 + O(p^3),$$

where α_1 is the same as Theorem 5.2,

$$\begin{aligned}
\alpha_2^\dagger &= \alpha_2^* + \frac{\mathbb{E}^*[\epsilon_*^2(L)]}{4\log 2} \\
&= \frac{1}{4\log 2}\Bigg(\mathbb{E}^*[L^2 \log^2 L] - (4+\tau)\,\mathbb{E}^*[L^2 \log L] \\
&\qquad + 2\,\mathbb{E}^*\left[\tbinom{L}{2} \log \tbinom{L}{2}\right] + \frac{3}{2}\tau^2 + 8\tau - 2\Bigg) \approx -2.74700879,
\end{aligned}$$

and α_2^* is defined as

$$\alpha_2^* = -\frac{1}{\log 2} - \frac{1}{2}\,\mathbb{E}^*\left[L(L-1)\log_2 L - \tbinom{L}{2}\log_2\tbinom{L}{2}\right] \approx -2.87845825. \quad (5.8)$$

Proof: See Section 5.5.4.

□

First-order Markov processes

In this section, we limit $\mathbb{X} \in \mathcal{S}$ to the class of symmetric first-order Markov processes to find a lower bound on α_2. Consider a symmetric first-order symmetric Markov source depicted in Fig. 5.3 where $\mathbb{P}(X_i = b|X_{i-1} = b) = \delta$, $b = 0, 1$. Let $\mathbb{X}^\circ$ denote such a process. Clearly, if either of states is selected uniformly at random (the stationary distribution), the process $\mathbb{X}^\circ$ will be stationary and ergodic. The runs of $\mathbb{X}^\circ$ are i.i.d and distributed as

$$P_L^\circ(\ell) = (1-\delta)\delta^{\ell-1},\ \ell \in \mathbb{N}.$$

It is important to note that $\mathbb{X}^\circ$ coincides with $\mathbb{X}^*$ when $\delta = \frac{1}{2}$. Therefore, we expect that $\delta - \frac{1}{2} = O(p)$ when the duplication probability is small enough. We wish to find the optimal $O(p)$ term to maximize $I(\mathbb{X}^\circ)$. We denote the expected value operator under P_L° by $\mathbb{E}^\circ$.

Theorem 5.4 [Optimal Markov Source]**:** Consider the class of symmetric first-order Markov sources where δ is the probability of transition from 0 to 1 and vice versa. The optimal p-dependent value of δ to maximize $I(\mathbb{X}^\circ)$ is

$$\delta^* = \frac{1}{2} + \Omega p,$$
$$\Omega = \frac{1}{4}\,\mathbb{E}^*[L(L-3)\log L] - \frac{1}{2} \approx 0.10409610$$

Moreover

$$I(\mathbb{X}^\circ) = 1 + p\log_2 p + \alpha_1 p + \alpha_2^\circ p^2 + O(p^3),$$

where

$$\alpha_2^\circ = \alpha_2^* + \frac{2\Omega^2}{\log 2} \approx -2.81592784,$$

and α_2^* is given by (5.8).

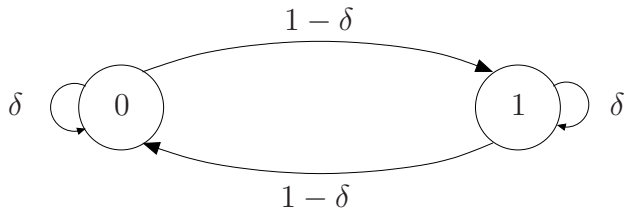

Figure 5.3: A symmetric first-order Markov process.

Proof: See Section 5.5.5. □

By letting $\epsilon(\ell) \equiv 0$ in the proof of Theorem 5.3, it can be deduced that $\alpha_2^* \approx -2.87845825$ is in fact the multiplier of p^2 in the series expansion of $I(\mathbb{X}^*)$, i.e.,

$$I(\mathbb{X}^*) = 1 + p\log_2 p + \alpha_1 p + \alpha_2^* p^2 + O(p^3).$$

Hence,

$$\alpha_2^* < \alpha_2^\circ < \alpha_2^\dagger \leq \alpha_2.$$

In other words, by limiting the process $\mathbb{X}$ to the class of perturbations satisfying conditions (5.6a) and (5.6b), and the class of first-order Markov processes, we have improved the achievable rates over the process $\mathbb{X}^*$. Although we do not prove the converse, we conjecture that the process $\mathbb{X}^\dagger$ achieves the capacity of duplication channels up to a term $O(p^3)$ in the series expansion.

Conjecture 5.1 [Capacity of Duplication Channels at $p \to 0$]**:** The capacity of a duplication channel near $p = 0$ is

$$\begin{aligned} C =& 1 + p\log_2 p + \alpha_1 p + \alpha_2^\dagger p^2 + O(p^3) \\ & \alpha_1 \approx 0.8458,\ \alpha_2^\dagger \approx -2.7470, \end{aligned}$$

which is achieved by a stationary and ergodic process with independent runs distributed according to $P_L^\dagger$ given in Theorem 5.3.

It is interesting to point out that numerical results show that

$$\sup |I(\mathbb{X}^\dagger) - \max\{I_l, I_u\}| = 5 \times 10^{-7}$$

for all $p < 0.15$, where I_l and I_u are the numerical lower and upper bounds given in Discussion 5.2.

5.3.2 Capacity at $p \to 1$

Similar to the analysis for small duplication probabilities, one can expand $I(\mathbb{X}^*)$ given in Theorem 5.1 when $p \to 1$.

Lemma 5.2 [$I(\mathbb{X}^*)$ at $p \to 1$]**:** The rate of the process $\mathbb{X}^*$ when $p \to 1$ is

$$I(\mathbb{X}^*) = 1 - O\left((1-p)^{2-\epsilon}\right)$$

for some small $\epsilon > 0$.

Proof: See Section 5.5.6. □

Since $I(\mathbb{X}^*)$ is a lower bound to the duplication capacity, this result shows that

$$1 - O((1-p)^{2-\epsilon}) \leq C \leq 1,$$

and

$$\lim_{p \to 1} \frac{dC}{dp} = 0,$$

i.e., the duplication capacity has a slope of zero at $p = 1$ (see Fig. 5.2). Although we do not know the exact capacity expression, this approximation helps us to understand the behaviour of capacity around $p = 1$. According to Theorem 5.2, the slope of capacity when $p \to 0$ is $-\infty$ which means that the capacity of duplication channels is not symmetric with respect to the probability of duplication.

Further approximation shows that

$$I(\mathbb{X}^*) = 1 + 2q^2 \log_2 q + \beta_2^* q^2 + O(q^{3-\epsilon}),$$

where $q = 1 - p$ and

$$\beta_2^* = \frac{1}{2}\,\mathbb{E}^*\left[\tbinom{L}{2} \log_2 \tbinom{L}{2}\right] - \log_2(2e) \approx 0.4417.$$

One might wonder whether it is possible to perform a perturbation analysis similar to the case of $p \to 0$. By the same logic, let $P_L^\dagger(\ell) = 2^{-\ell}\,(1 + q\epsilon(\ell))$,

$\ell \in \mathbb{N}$ such that $\mathbb{E}^*[\epsilon(L)] = 0$. Without stating the long technical proof, we have $I(\mathbb{X}^\dagger) = 1 + 2q^2 \log_2 q + \beta_2^\dagger q^2 + O(q^{3-\epsilon})$, where

$$\beta_2^\dagger = \beta_2^* - \frac{\mathbb{E}^*[\epsilon^2(L)]}{4 \log 2} \leq \beta_2^*$$

shows that the optimal choice is $\epsilon(\ell) \equiv 0$. Similar to the case of $p \to 0$, this strong evidence suggests the following conjecture:

Conjecture 5.2 [Capacity of Duplication Channels at $p \to 1$]**:** For small $\epsilon > 0$, the capacity of a duplication channel near $p = 1$ is

$$C = 1 + 2(1-p)^2 \log_2(1-p) + \beta_2^*(1-p)^2 + O\left((1-p)^{3-\epsilon}\right), \ \beta_2^* \approx 0.4417,$$

which is achieved by the process $\mathbb{X}^*$.

Discussion 5.3 [*System design perspective*]**:** Based on above discussion, capacity drops much faster for values of $p \approx 0$ compared to $p \approx 1$ (see Fig. 5.2). This means that from a practical point of view, it is more desirable to communicate over a *near perfect duplicate* channel rather than a channel which barely duplicates. To elaborate, we compare two duplication channels: One with duplication rate of ε and another one with duplication rate of $1 - \varepsilon$, where ε is very small. We decode the second channel using a simple (yet sub-optimal) decoder: Starting from left, remove the bits at even positions of every run of length two or more. For example, assuming p very close to 1, 0|11|0000|1|000 will be decoded to 0|1|00|1|00. The decoder makes an error only when two equal consecutive bits are neither duplicated which happens with probability $\frac{1}{2}\varepsilon^2$. The second channel with the sub-optimal decoder resembles a deletion channel with the deletion rate of $d = \frac{1}{2}\varepsilon^2$. As we mentioned in Section 5.3, the capacity of a deletion channel with the deletion rate of d, is $C \approx 1 + d \log_2 d - A_1 d$, where $A_1 = 2 - \alpha_1$. Therefore, the ratio of the gap to capacity of the second case and the first case is

$$\lim_{\varepsilon \to 0} \frac{d \log_2 d - A_1 d}{\varepsilon \log_2 \varepsilon + \alpha_1 \varepsilon} = \frac{1}{2} \times \lim_{\varepsilon \to 0} \frac{\varepsilon^2 \log_2\left(\frac{1}{2}\varepsilon^2\right) - A_1 \varepsilon^2}{\varepsilon \log_2 \varepsilon + \alpha_1 \varepsilon} = 0.$$

Therefore, for duplication probabilities near one, the sub-optimal decoder achieves higher rates than a duplication channel with small duplication probabilities. It is important to point out that although the sub-optimal decoder

does not capture the "zero-slope behaviour", it is a proper tool for explaining the asymmetric shape of the duplication capacity.

5.4 Summary

In this chapter, we studied the capacity of duplication channels. It was shown that the set of stationary ergodic processes suffices to achieve the capacity. Then, for small duplication probabilities, we showed that a Bernoulli($\frac{1}{2}$) process achieves the capacity up to term p in a series expansion (with an error of order $p^{3/2}$). We also gave two tight lower bounds for the next term in the series expansion by limiting the input process into two subclasses of stationary ergodic processes. We showed that the best lower bound is given by a p-dependent perturbation in a Bernoulli($\frac{1}{2}$) process. Although we did not prove its converse, we believe that the perturbed process achieves the capacity up to a term p^2.

Moreover, we provided achievable rates for the capacity when $p \to 1$. A similar perturbation analysis showed that a Bernoulli($\frac{1}{2}$) process can achieve the rate up to a term $(1-p)^2$ when $p \to 1$. We observed that unlike the case of $p \to 0$, the slope of the capacity is zero when $p \to 1$. This study showed that the duplication capacity is not a symmetric function and behaves differently for small and large duplication probabilities which can be used in practical setups.

5.5 Proofs

In this section, we provide the proofs of main results, where we exploit the renewal theory [65] in some parts. Let x^n be a realization of X^n from a stationary and ergodic process $\mathbb{X}$ with i.i.d. runs drawn according to a common distribution P_L. Assume that there are K_n runs in x^n. The process K_n is in fact a counting process, where an increment occurs by the birth of a new run. According to the renewal theory [65],

$$\lim_{n\to\infty} \frac{\mathbb{E}[K_n]}{n} = \frac{1}{\mathbb{E}[L]}.$$

Moreover, $K_n \in \{1, \ldots, n\}$ and

$$H(K_n) \leq \log_2(n) = o(n),$$

due to the fact that the entropy of any distribution on a finite alphabet is not greater than the entropy of the uniform distribution [2].

Furthermore, the law of total expectation (also known as the tower rule) will be used in this section [44, Chapter 9]. Let X be an integrable random variable, i.e., $\mathbb{E}[|X|] < \infty$ and Y be any random variable on the same probability space. Then, $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$, where the outer expected value on the RHS is taken with respect to Y. Note that $\mathbb{E}[X|Y]$ is a random variable itself. To illustrate, we calculate $\mathbb{E}[G]$ as an example. The same logic is applied to other similar cases. From Section 5.2, G is a $B(L, p)$ random variable, i.e., the size of the binomial distribution is a random variable itself. Therefore, by the law of total expectation, we get

$$\mathbb{E}[G] = \mathbb{E}[\mathbb{E}[G|L]] = \sum_{\ell=1}^{\infty} P_L(\ell) \underbrace{\mathbb{E}[G|L = \ell]}_{=p\ell} = p\,\mathbb{E}[L].$$

Instead of expanding the expected values, we write

$$\mathbb{E}[G] = \mathbb{E}[\mathbb{E}[G|L]] = \mathbb{E}[pL] = p\,\mathbb{E}[L]$$

as a proxy.

5.5.1 Proof of Lemma 5.1

The proof is similar to that of [60] with some extra considerations. For the sake of completeness, we state the complete proof. For any $\epsilon > 0$, we construct a process $\mathbb{X}_\epsilon \in \mathcal{S}$ such that $\forall N > N_0(\epsilon)$,

$$\frac{I_N}{N} \geq C - \epsilon.$$

According to (5.1), there exists n such that $C_n \geq C - \epsilon/2$. We construct $\mathbb{X}_\epsilon$ with i.i.d. blocks of length n with a common distribution $P^*_{X^n}$ that achieves C_n (see (5.1))[2]. To make $\mathbb{X}_\epsilon$ stationary, we introduce a random offset $s \in$

[2]Clearly, W_n is a DMC. Therefore, there exists an optimal input distribution achieving its capacity [3]. The optimal input distribution is denoted by $P^*_{X^n}$.

$\{1, 2, \ldots, n\}$ from which we move the first complete block to the right. Since s is uniformly random, every window of length n will have the same distribution which confirms stationarity. Also, ergodicity inherits from the construction.

Let $N = kn + r$ for some $k \in \mathbb{N}$ and $r \in \{0, 1, \ldots, n-1\}$. We have

$$X^N = (X_1^{s-1}, \underbrace{X(1), \ldots, X(k-1)}_{\text{Blocks of length } n}, X_{s+(k-1)n}^N),$$
$$Y(X^N) = (Y(X_1^{s-1}), Y(1), \ldots, Y(k-1), Y(X_{s+(k-1)n}^N)),$$

where $X(i)$'s are i.i.d. blocks of length n generated according to $P_{X^n}^*$, and $Y(i)$'s are the corresponding received blocks according to the realization of D.

Now, consider a genie-aided channel where a character "|" is placed by a genie after each segment of X^N gone through the insertion/deletion process. The output of this channel is

$$\tilde{Y}(X^N) = (Y(X_1^{s-1})|Y(1)|\cdots|Y(k-1)|Y(X_{s+(k-1)n}^N)).$$

It is important to note that since we allow deletions, a whole block might disappear, hence two adjacent "|". Since $(X(i), Y(i))$'s are i.i.d., we have

$$\begin{aligned} H(Y(X^N)|X^N) &\leq H(\tilde{Y}(X^N)|X^N) \\ &\overset{(a)}{\leq} (k-1)H(Y(1)|X(1)) + (n+r)H(D) \\ &\overset{(b)}{\leq} (k-1)H(Y(1)|X(1)) + 2nH(D), \end{aligned} \tag{5.9}$$

where (a) holds by the fact that there are $k-1$ blocks of length n and $(X(i), Y(i))$'s are i.i.d. The ambiguity of the remaining $(s-1)+(N-s-(k-1)n+1) = n+r$ bits is the uncertainty about the outcomes of D. Also, (b) follows from $r < n$.

Let M be the length of $Y(X^N)$, i.e., $M = \sum_{i=1}^N D_i$, hence $\mathbb{E}[M] = N\,\mathbb{E}[D]$. Given $Y(X^N)$, the ambiguity that remains in $\tilde{Y}(X^N)$ is the logarithm of the number of possible placements of k synchronization character "|". This is the number of non-negative integer solutions of the Diophantine equation $u_1 + u_2 + \cdots + u_k = M$ where $u_i \geq 0$ represents the number of bits between between the $(i-1)$th and ith character "|". Thus, we get

$$
\begin{aligned}
H(\tilde{Y}(X^N)|Y(X^N)) &= \mathbb{E}\left[\log_2 \binom{M+k}{k}\right] \\
&\overset{(a)}{\leq} k\,\mathbb{E}[\log_2 (e\,(1+M/k))] \\
&\overset{(b)}{\leq} k \log_2 (e(1+N\,\mathbb{E}[D]/k)) \\
&\overset{(c)}{<} k \log_2 (e(1+2n\,\mathbb{E}[D]))\,, \qquad (5.10)
\end{aligned}
$$

where (a) is due to the upper bound on the binomial coefficient $\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$, (b) follows from the concavity of logarithm and $\mathbb{E}[M] = N\,\mathbb{E}[D]$, and (c) is by $r < n$. Moreover,

$$
H(\tilde{Y}(X^N)) \geq (k-1)H(Y(1)). \qquad (5.11)
$$

Now, since $\mathbb{E}[D], H(D) < \infty$ and $C_n \leq 1$, we can bound $I(X^N; Y(X^N))$ by

$$
\begin{aligned}
I_N &= H(Y(X^N)) - H(Y(X^N)|X^N) \\
&= H(\tilde{Y}(X^N)) - H(\tilde{Y}(X^N)|Y(X^N)) - H(Y(X^N)|X^N) \\
&\geq (k-1)H(Y(1)) - k\log_2 (e\,(1+2n\,\mathbb{E}[D])) \\
&\qquad - (k-1)H(Y(1)|X(1)) - 2nH(D),
\end{aligned}
$$

where the last inequality results from combining (5.9), (5.10), and (5.11). Since $X(1)$ is a block of n bits generated by the optimal input distribution $P^*_{X^n}$ to achieve C_n given in (5.1), $H(Y(1)) - H(Y(1)|X(1)) = nC_n$. Thus, we continue as

$$
\begin{aligned}
I_N &\geq (k-1)nC_n - k\log_2 (e\,(1+2n\,\mathbb{E}[D])) - 2nH(D) \\
&= (k+1)nC_n - k\log_2 (e\,(1+2n\,\mathbb{E}[D])) - 2n(H(D)+C_n) \\
&\geq NC_n - k\log_2 (e\,(1+2n\,\mathbb{E}[D])) - 2n(H(D)+1),
\end{aligned}
$$

where we used the fact that $(k+1)n > N$ and $C_n \leq 1$. This implies that $I_N/N \geq C_n - \epsilon/2$ provided that

$$
\frac{1}{n}\log_2 (e(1+2n\,\mathbb{E}[D])) < \frac{\epsilon}{4}, \quad \frac{2n}{N}(H(D)+1) < \frac{\epsilon}{4},
$$

i.e., $N > N_0(\epsilon) = \frac{8n}{\epsilon}(1 + H(D))$. By construction, $C_n \geq C - \epsilon/2$. Thus

$$\frac{I_N}{N} \geq C - \epsilon.$$

According to Fekete's lemma, the limit $\lim_{n\to\infty} I_n/n$ exists [60]. Since ϵ is arbitrary and $I_N/N \leq C_N$ (see (5.1)), letting $\epsilon \to 0$ results in $C = \sup_{\mathbb{X}\in\mathcal{S}} I(\mathbb{X})$.

5.5.2 Proof of Theorem 5.1

The process $\mathbb{X}$ consists of i.i.d. runs, so does $\mathbb{Y}$. By Lemma 5.1, any process $\mathbb{X} \in \mathcal{S}$ gives a lower bound on the duplication capacity. We have

$$I(\mathbb{X}) = \lim_{n\to\infty} \frac{1}{n}(H(Y(X^n)) - H(Y(X^n)|X^n)).$$

A sample realization of $\mathbb{X}$ will look like

$$x^n = \ldots 111 \overbrace{0000}^{\mathcal{L}_1} \overbrace{111111}^{\mathcal{L}_2} \overbrace{000}^{\mathcal{L}_3} 11 \ldots \tag{5.12}$$

and causes a realization of $\mathbb{Y}$ as

$$y(x^n) = \ldots 111 \overbrace{000000}^{\mathcal{T}_1} \overbrace{111111111}^{\mathcal{T}_2} \overbrace{000}^{\mathcal{T}_3} 11 \ldots \tag{5.13}$$

Note that each run in $y(x^n)$ ($\mathcal{L}_i$) is uniquely mapped to a run in x^n ($\mathcal{T}_i$). Assume that there are K_n runs in x^n. Given each run $\mathcal{L}_i$, the uncertainty that remains in the corresponding run $\mathcal{T}_i$ is the ambiguity on the number of duplications, i.e., $H(G_i)$. Therefore, using the fact that (L_i, T_i)'s and G_i's are i.i.d., we obtain

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} H(Y(X^n)|X^n) &= \lim_{n\to\infty} \frac{1}{n} \sum_{x^n} p(x^n) H(Y(X^n)|X^n = x^n) \\
&= \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left[H(T_1^{K_n}|L_1^{K_n})\right] \\
&= \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^{K_n} H(G_i)\right] \\
&= \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{K_n} H(G_i)\middle| K_n\right]\right] \\
&= \lim_{n\to\infty} \frac{\mathbb{E}[K_n]}{n} H(G) \\
&= \frac{H(G)}{\mathbb{E}[L]}.
\end{aligned} \tag{5.14}$$

Also, since $G \sim B(L,p)$ has the binomial distribution, the average run length at the output of the channel is

$$\mathbb{E}[T] = \mathbb{E}[L+G] = \mathbb{E}[L] + \mathbb{E}[\mathbb{E}[G|L]] = (1+p)\,\mathbb{E}[L].$$

We continue evaluating $H(G)/\,\mathbb{E}[L]$ as:

$$\begin{aligned}
\frac{H(G)}{\mathbb{E}[L]} &= -\frac{1}{\mathbb{E}[L]}\,\mathbb{E}\left[\log_2\left(\binom{L}{G}p^G(1-p)^{L-G}\right)\right] \\
&= -\frac{1}{\mathbb{E}[L]}\left(\mathbb{E}[G]\log_2 p + \mathbb{E}[L-G]\log_2(1-p) + \mathbb{E}\left[\log_2\binom{L}{G}\right]\right) \\
&= \frac{1}{\mathbb{E}[L]}\left(\mathbb{E}[L]h(p) - \mathbb{E}\left[\log_2\binom{L}{G}\right]\right) \\
&= h(p) - \frac{\mathbb{E}\left[\log_2\binom{L}{G}\right]}{\mathbb{E}[L]}.
\end{aligned}$$

Moreover, one may characterize $Y(X^N)$ by its run lengths $(T_1,\ldots,T_{K_n},K_n)$[3]; except that it might start with either 0 or 1. Since $(T_1,\ldots,T_{K_n},K_n)$ characterizes $Y(X^n)$ within one bit of ambiguity and T_i's are i.i.d., one may continue as follows:

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}H(Y(X^n)) &= \lim_{n\to\infty}\frac{1}{n}\Bigg(H(Y(X^n),T_1^{K_n},K_n) - \underbrace{H(T_1^{K_n},K_n|Y(X^N))}_{=0}\Bigg) \\
&= \lim_{n\to\infty}\frac{1}{n}\Bigg(\underbrace{H(Y(X^N)|T_1^{K_n},K_n)}_{=1} + H(T_1^{K_n}|K_n) + \underbrace{H(K_n)}_{=o(n)}\Bigg) \\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\mathbb{P}(K_n=j)\,\underbrace{H(T_1^{K_n}|K_n=j)}_{=jH(T)} \\
&= \lim_{n\to\infty}\frac{\mathbb{E}[K_n]}{n}H[T] \\
&= \frac{H[T]}{\mathbb{E}[L]}. \qquad (5.15)
\end{aligned}$$

Therefore, the capacity of duplication channels is lower bounded by

$$I(\mathbb{X}) = \frac{H[T]}{\mathbb{E}[L]} - h(p) + \frac{\mathbb{E}\left[\log_2\binom{L}{G}\right]}{\mathbb{E}[L]}.$$

[3]Here, the length of the sequence of run lengths is random itself.

5.5.3 Proof of Theorem 5.2

In order to prove this theorem, first, we calculate $I(\mathbb{X}^*)$. Then, we prove that for any other process $\mathbb{X} \in \mathcal{S}$, $I(\mathbb{X})$ cannot be greater than $I(\mathbb{X}^*)$. Note that for simplicity of presentation, we use the symbol ϵ for all the polynomial approximations with the big-O notation.

Achievability

For small values of p, we can approximate the binary entropy function as

$$\begin{aligned} h(p) &= -p\log_2 p + (1-p)\sum_{k=1}^{\infty}\frac{p^k}{k}\log_2 e \\ &= -p\log_2 p + p\log_2 e + O(p^2). \end{aligned} \tag{5.16}$$

Now, we are going to approximate the remaining terms of the lower bound given in Theorem 5.2 for $\mathbb{X}^*$. Clearly, the run lengths are distributed geometrically according to $P_L^*(\ell) = 2^{-\ell}$, $\ell = 1, 2, \ldots$ The following moments, calculated by simple calculus, will prove handy as we go on:

$$\mathbb{E}^*[L] = 2,\ \mathbb{E}^*[L^2] = 6,\ \mathbb{E}^*[L^3] = 26.$$

Since G is a binomial $B(L, p)$, we get

$$\begin{aligned} \frac{\mathbb{E}\left[\log_2 \binom{L}{G}\right]}{\mathbb{E}[L]} &= \frac{1}{\mathbb{E}[L]}\,\mathbb{E}\left[\sum_{g=0}^{L}\binom{L}{g}\log_2\binom{L}{g}p^g(1-p)^{L-g}\right] \\ &= \frac{1}{\mathbb{E}[L]}\,\mathbb{E}\left[\sum_{g=0}^{L}\binom{L}{g}\log_2\binom{L}{g}p^g\left(1+O(p)\right)\right] \\ &= p\frac{\mathbb{E}[L\log_2 L]}{\mathbb{E}[L]} + O(p^2) \\ &= p\,\mathbb{E}[\log_2 L_0] + O(p^2), \end{aligned} \tag{5.17}$$

where the last equality holds by (5.3). Note that (5.17) is valid for any run length distribution. Moreover, the probability of receiving a run of length $\ell = 1, 2, \ldots$ can be written as

$$\begin{aligned} P_T^*(\ell) &= \sum_{k=\lceil \ell/2\rceil}^{\ell} 2^{-k}\binom{k}{\ell-k}p^{\ell-k}(1-p(2k-\ell)+O(p^2)) \\ &= 2^{-\ell}(1-\ell p) + 2^{-\ell+1}(\ell-1)p + O(p^2) \\ &= 2^{-\ell}\left(1+p(\ell-2)\right) + O(p^2), \end{aligned}$$

which implies that the ambiguity in the received run lengths will be

$$\begin{aligned} H(T) &= -\sum_{\ell=1}^{\infty} 2^{-\ell}(1+(\ell-2)p+O(p^2))\log_2\left(2^{-\ell}(1+(\ell-2)p+O(p^2))\right) \\ &= \sum_{\ell=1}^{\infty} 2^{-\ell}(1+(\ell-2)p+O(p^2))\left(\ell-(\ell-2)p\log_2 e+O(p^2)\right) \\ &= 2(1+p)+O(p^2). \end{aligned} \tag{5.18}$$

Proof of achievability is complete due to the fact that Theorem 5.1 provides a lower bound on the capacity and by combining (5.16), (5.17) (using $\mathbb{E}^*$), and (5.18), i.e.,

$$I(\mathbb{X}^*) = 1 + p\log_2 p + \alpha_1 p + O(p^2),$$

where

$$\begin{aligned} \alpha_1 &= 1 - \log_2 e + \mathbb{E}^*[\log_2 L_0] \\ &= \log_2(2/e) + \sum_{\ell=1}^{\infty} 2^{-(\ell+1)}\ell\log_2\ell. \end{aligned}$$

Converse

Before proving the converse, it is worth mentioning that by the achievability part, we can focus on those stationary and ergodic processes for which $I(\mathbb{X}) \geq 1 - p\log_2 1/p$ (otherwise, $I(\mathbb{X}) \leq I(\mathbb{X}^*)$ and we are automatically done). Furthermore, because $H(\mathbb{X}) \geq I(\mathbb{X})$, we can confine our attention to those processes whose entropy rate satisfies

$$H(\mathbb{X}) \geq 1 - p\log_2 1/p.$$

This restriction allows us to use the results of [60, Lemmas IV.1, IV.2, IV.3].

Lemma 5.3 [Distance between P_{L_0} and $P^*_{L_0}$ [60, Lemmas IV.3]]**:** There exists $p_0 > 0$ such that for any $\mathbb{X} \in \mathcal{S}$ with $H(\mathbb{X}) \geq 1 - p\log_2 1/p$ and any $p < p_0$,

$$\|P_{L_0} - P^*_{L_0}\|_{\text{TV}} = O(\sqrt{p(\log_2 1/p)^3}). \tag{5.19}$$

Lemma 5.3 demonstrates that when the probability of duplication is small enough, the distance between the run length distributions of $\mathbb{X}$ and $\mathbb{X}^*$ cannot

be large. In the sequel, we use the fact that for any $\epsilon > 0$, $\log_2 1/p = O(p^{-\epsilon})$, i.e., $\log_2 1/p$ can be absorbed into any power of p.

Take any stationary and ergodic process $\mathbb{X}$, not necessarily composed of i.i.d. runs. We have to upper bound $I(\mathbb{X})$ in order to prove the converse. First of all, $Y(X^n)$ contains $M_n \in \{n, n+1, \ldots, 2n\}$ bits (see (5.2)) and does not necessarily exhibit i.i.d. runs. We have

$$\begin{aligned}
H(Y(X^n)) &= H(Y_1^{M_n}|M_n) + H(M_n) \\
&= \sum_{j=n}^{2n} \mathbb{P}(M_n = j) H(Y_1^{M_n}|M_n = j) + H(M_n) \\
&\leq \sum_{j=n}^{2n} \mathbb{P}(M_n = j) \sum_{i=1}^{j} H(Y_i) + H(M_n) \\
&\overset{(a)}{\leq} \mathbb{E}[M_n] + \log_2(n+1) \\
&\overset{(b)}{=} n(1+p) + o(n),
\end{aligned}$$

where (a) is due to $H(Y_i) \leq 1$ and (b) follows from (5.2). On the other hand, we have

$$\begin{aligned}
H(Y(X^n)|X^n) &= H(Q^n|X^n) - H(Q^n|X^n, Y(X^n)) \\
&= nh(p) - \sum_{x^n, y(x^n)} \mathbb{P}(x^n, y(x^n)) H(Q^n|x^n, y(x^n)).
\end{aligned}$$

Given a pair of sequences $(x^n, y(x^n))$ (see (5.12), (5.13)), the ambiguity remaining in the duplication sequence Q^n is the total uncertainty about the duplication patterns given all runs in x^n and $y(x^n)$. Assume that there are K_n runs in X^n and let $Q(i)$ be the part of the duplication process corresponding to the ith run, $i = 1, \ldots, K_n$, of X^n. One can continue as

$$\begin{aligned}
H(Q^n|X^n, Y(X^n)) &= \mathbb{E}\left[H(Q^n|L_1^{K_n}, T_1^{K_n})\right] \\
&\leq \mathbb{E}\left[\sum_{i=1}^{K_n} H(Q(i)|L_i, T_i)\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{K_n} \log \binom{L_i}{G_i}\right] \\
&= \mathbb{E}[K_n]\, \mathbb{E}\left[\log \binom{L}{G}\right],
\end{aligned}$$

where equality holds if X^n has i.i.d. runs (in other words, if (L_i, T_i)'s are independent). Combining this with (5.17) and using a renewal process argument, we obtain that[4]

$$I(\mathbb{X}) \leq 1 + p\log_2 p + p(\log_2(2/e) + \mathbb{E}[\log_2 L_0]) + O(p^2). \tag{5.20}$$

Now, we are left to bound $\mathbb{E}[\log_2 L_0]$ from above. To do so, let $\mathcal{S}_{L^*}$ denote the set of stationary ergodic processes in which with probability one, no run has length more than L^*, i.e., $L \leq L^*$ almost surely. To obtain the corresponding process in $\mathcal{S}_{L^*}$, i.e., $\mathbb{X}_{L^*}$, one has to flip the L^*+1 consecutive bit in $\mathbb{X}$. This limitation will help us to bound $\mathbb{E}[\log_2 L_0]$ using (5.19), i.e., for any $\epsilon > 0$, we obtain that

$$\begin{aligned}\big|\,\mathbb{E}[\log_2 L_0] - \mathbb{E}^*[\log_2 L_0]\big| &\leq \sum_{\ell=1}^{L^*} \log_2 \ell \,\big|P_{L_0}(\ell) - P^*_{L_0}(\ell)\big| \\ &\leq \|P_{L_0} - p^*_{L_0}\|_{\text{TV}} \log_2 L^* \\ &\leq K p^{1/2-\epsilon} \log_2 L^* \end{aligned} \tag{5.21}$$

for some large enough $K < \infty$. The following lemma is directly inferred from [60, Lemma III.2]:[5]

Lemma 5.4 [Distance between $I(\mathbb{X})$ and $I(\mathbb{X}_{L^*})$ [60, Lemma III.2]]**:** For any $\epsilon > 0$, there exists $p_0(\epsilon)$ such that the following happens for all $p < p_0(\epsilon)$. For any $\mathbb{X} \in \mathcal{S}$ such that $H(\mathbb{X}) > 1 - p\log_2 1/p$ and for any $L^* > \log_2 1/p$, there exists $\mathbb{X}_{L^*} \in \mathcal{S}_{L^*}$ such that

$$|I(\mathbb{X}) - I(\mathbb{X}_{L^*})| \leq 2p^{1/2-\epsilon} L^{*-1} \log_2 L^*. \tag{5.22}$$

This Lemma shows that by restricting the run length, we will not lose too much information rate, provided that L^* is large enough. Note that (5.20) is valid for any stationary and ergodic process. Therefore, $I(\mathbb{X}_{L^*})$ can be bounded using (5.20) and (5.21) as

$$I(\mathbb{X}_{L^*}) \leq 1 + p\log_2 p + \alpha_1 p + O(p^{3/2-\epsilon}\log_2 L^*) + O(p^2). \tag{5.23}$$

[4]As noted before, (5.17) is valid for any run length distribution.

[5]Our result differs from [60, Lemma III.2] in $H(F) \leq 2n(1+p)h(\alpha)+o(n) \leq 4nh(\alpha)+o(n)$.

Taking $L^* = \lfloor 1/p \rfloor$ and by adding (5.22) to (5.23), since $O(p^2) = O(p^{3/2-\epsilon})$, we get

$$\begin{aligned} I(\mathbb{X}) &\leq 1 + p\log_2 p + \alpha_1 p + \underbrace{2p^{3/2-\epsilon}\log_2 1/p}_{O(p^{3/2-\epsilon})} + \underbrace{O(p^{3/2-\epsilon'}\log_2 1/p)}_{O(p^{3/2-\epsilon})} + O(p^2) \\ &= 1 + p\log_2 p + \alpha_1 p + O(p^{3/2-\epsilon}), \end{aligned}$$

which shows that for any process $\mathbb{X} \in \mathcal{S}$, $I(\mathbb{X}) \leq I(\mathbb{X}^*)$. This completes the proof.

5.5.4 Proof of Theorem 5.3

The average run-length of $\mathbb{X}^\dagger$ is

$$\mathbb{E}^\dagger[L] = 2 + p\,\mathbb{E}^*[L\epsilon(L)].$$

Also, the received run-length distribution can be written as

$$\begin{aligned} P_T^\dagger(\ell) &= \sum_{k=\lceil \ell/2 \rceil}^{\ell} 2^{-k}(1 + p\epsilon(k))\binom{k}{\ell-k} p^{\ell-k}(1-p)^{2k-\ell} \\ &= 2^{-\ell}(1 + p\delta_1(\ell) + p^2\delta_2(\ell)) + O(p^3), \end{aligned}$$

where for $\ell \in \mathbb{N}$,

$$\begin{aligned} \delta_1(\ell) &= \epsilon(\ell) + \ell - 2 \\ \delta_2(\ell) &= 2(\ell-1)\epsilon(\ell-1) - \ell\epsilon(\ell) + \frac{1}{2}(\ell^2 - 9\ell + 16). \end{aligned}$$

The following moments are straightforward to obtain using condition (5.7) and the fact that for every $\varphi : \mathbb{N} \cup \{0\} \mapsto \mathbb{R}$ such that $\varphi(0) = 0$, $2\,\mathbb{E}^*[\varphi(L-1)] = \mathbb{E}^*[\varphi(L)]$:

$$\begin{aligned} \mathbb{E}^*[\delta_1(L)] &= 0 \\ \mathbb{E}^*[\delta_1^2(L)] &= \mathbb{E}^*[\epsilon^2(L)] + 2\,\mathbb{E}^*[L\epsilon(L)] + 2 \\ \mathbb{E}^*[\delta_2(L)] &= 2 \end{aligned}$$

A Taylor expansion of $D(P_T^\dagger \| P_L^*)$ around P_L^* gives

$$\begin{aligned} D(P_T^\dagger \| P_L^*) &= \sum_{\ell=1}^{\infty} P_T^\dagger(\ell) \left(\ell + \log_2 P_T^\dagger(\ell) \right) \\ &= \frac{1}{\log 2} \sum_{\ell=1}^{\infty} \left(P_T^\dagger(\ell) - 2^{-\ell} \right) + \frac{1}{2 \log 2} \sum_{\ell=1}^{\infty} 2^\ell \left(P_T^\dagger(\ell) - 2^{-\ell} \right)^2 + O(p^3) \\ &= \frac{1}{\log 2} \left(p\, \mathbb{E}^*[\delta_1(L)] + p^2\, \mathbb{E}^*[\delta_2(L)] + \frac{1}{2} p^2\, \mathbb{E}^*[\delta_1^2(L)] \right) + O(p^3) \\ &= \frac{p^2}{2 \log 2} \left(\mathbb{E}^*[\epsilon^2(L)] + 2\, \mathbb{E}^*[L\epsilon(L)] + 6 \right) + O(p^3), \end{aligned}$$

where we used the fact that $\left| P_T^\dagger(\ell) - 2^{-\ell} \right| = O(p)$. Also, we have

$$\mathbb{E}^\dagger[T] = (1+p)(2 + p\, \mathbb{E}^*[L\epsilon(L)]) = 2 + O(p).$$

Thus, by (5.5) we yield

$$\frac{H(T)}{\mathbb{E}^\dagger[L]} = 1 + p - \frac{p^2}{4 \log 2} \left(\mathbb{E}^*[\epsilon^2(L)] + 2\, \mathbb{E}^*[L\epsilon(L)] + 6 \right) + O(p^3).$$

Similarly, one can obtain

$$\begin{aligned} \mathbb{E}^\dagger \left[\log_2 \binom{L}{G} \right] =& p\, \mathbb{E}^*[L \log_2 L] + p^2 \Big(\mathbb{E}^*[L\epsilon(L) \log_2 L] \\ & - \mathbb{E}^* \left[L(L-1) \log_2 L - \tbinom{L}{2} \log_2 \tbinom{L}{2} \right] \Big). \end{aligned}$$

The reciprocal of $\mathbb{E}^\dagger[L]$ can be approximated as

$$\frac{1}{\mathbb{E}^\dagger[L]} = \frac{1}{2 \left(1 + \frac{1}{2} p\, \mathbb{E}^*[L\epsilon(L)] \right)} = \frac{1}{2} - \frac{p}{4}\, \mathbb{E}^*[L\epsilon(L)] + O(p^2),$$

provided that $p |\, \mathbb{E}^*[L\epsilon(L)]| < 2$. Combining these results, we get

$$\begin{aligned} \frac{\mathbb{E}^\dagger \left[\log_2 \binom{L}{G} \right]}{\mathbb{E}^\dagger[L]} &= \frac{p}{2}\, \mathbb{E}^*[L \log_2(L)] + \frac{p^2}{2} \Bigg(\mathbb{E}^*[L\epsilon(L) \log_2 L] \\ &- \frac{1}{2}\, \mathbb{E}^*[L\epsilon(L)]\, \mathbb{E}^*[L \log_2 L] - \mathbb{E}^* \left[L(L-1) \log_2 L - \tbinom{L}{2} \log_2 \tbinom{L}{2} \right] \Bigg) + O(p^3). \end{aligned}$$

The information rate of process $\mathbb{X}^\dagger$, using Theorem 5.1, will be

$$I(\mathbb{X}^\dagger) = 1 + p \log_2 p + \alpha_1 p + \alpha_2^\dagger p^2 + O(p^3),$$

where α_1 is the same as Theorem 5.2 and $\alpha_2^\dagger = \alpha_2^* + \frac{\alpha_{2,1}}{4\log 2}$ and

$$\alpha_2^* = -\frac{1}{\log 2} - \frac{1}{2}\,\mathbb{E}^*\left[L(L-1)\log_2 L - \tbinom{L}{2}\log_2 \tbinom{L}{2}\right] \approx -2.87845825,$$
$$\alpha_{2,1} = -\,\mathbb{E}^*[\epsilon^2(L)] + 2\,\mathbb{E}^*[L\epsilon(L)\log L] - \mathbb{E}^*[L\epsilon(L)](2 + \mathbb{E}^*[L\log L]).$$

It is clear that $\alpha_2^\dagger$ is a lower bound of α_2. We wish to maximize $\alpha_2^\dagger$ to obtain the tightest lower bound. We are left to find the optimal $\epsilon(\ell)$, denoted by $\epsilon_*(\ell)$, that maximizes $\alpha_{2,1}$ (and $\alpha_2^\dagger$) provided that $\mathbb{E}^*[\epsilon(L)] = 0$. To do so, first note that

$$\tau = \mathbb{E}^*[L\log L] \approx 1.78628364$$

is independent of $\epsilon(\ell)$. Then, we rewrite $\alpha_{2,1}$ as

$$\alpha_{2,1} = \sum_{\ell=1}^{\infty} 2^{-\ell}\left(-\epsilon^2(\ell) + 2\ell\epsilon(\ell)\log\ell - \ell\epsilon(\ell)(2+\tau)\right)$$

and use the Lagrange method. The optimal point can be obtained via

$$-2\epsilon_*(\ell) + 2\ell\log\ell - \ell(2+\tau) + \theta = 0,\ \ell = 1, 2, \ldots,$$

where $\theta > 0$ is the Lagrange multiplier. Now, by constraint (5.7), we obtain $\theta = 4$ independent of ℓ. Therefore, the optimal *p-dependent* run-length distribution is

$$P^\dagger(\ell) = 2^{-\ell}(1 + p\epsilon_*(\ell)),\ \ell \in \mathbb{N}$$
$$\epsilon_*(\ell) = \ell\log\ell - \ell(\tfrac{1}{2}\tau + 1) + 2.$$

It is easy to see that

$$\alpha_{2,1} = -\,\mathbb{E}^*\left[(\epsilon(L) - \epsilon_*(L))^2\right] + \mathbb{E}[\epsilon_*^2(L)],$$

which gives the maximum value of $\alpha_2^\dagger$ as

$$\begin{aligned}\alpha_2^\dagger &= \alpha_2^* + \frac{\mathbb{E}^*[\epsilon_*^2(L)]}{4\log 2}\\ &= \frac{1}{4\log 2}\Bigg(\mathbb{E}^*[L^2\log^2 L] - (4+\tau)\,\mathbb{E}^*[L^2\log L]\\ &\qquad + 2\,\mathbb{E}^*\left[\tbinom{L}{2}\log\tbinom{L}{2}\right] + \frac{3}{2}\tau^2 + 8\tau - 2\Bigg)\\ &\approx -2.74700879.\end{aligned}$$

One can verify that $\mathbb{E}^*[L\epsilon(L)] \approx 0.42$, satisfying $p|\,\mathbb{E}^*[L\epsilon(L)]| < 2$.

5.5.5 Proof of Theorem 5.4

Let $x = \delta - \frac{1}{2} = O(p)$. A Taylor expansion of $P_L^\circ(\ell) = (1-\delta)\delta^{\ell-1}$ around $\delta = \frac{1}{2}$ gives

$$P_L^\circ(\ell) = 2^{-\ell}\left(1 + x\delta_1(\ell) + x^2\delta_2(\ell)\right) + O(p^3),$$

where

$$\delta_1(\ell) = 2(\ell-2),\ \delta_2(\ell) = 2(\ell-1)(\ell-4),\ \ell \in \mathbb{N}.$$

The received run-length distribution can be written as

$$P_T^\circ(\ell) = 2^{-\ell}\Bigg(1 + x\delta_1(\ell) + x^2\delta_2(\ell) + p\left(\ell - 2 + x(\ell^2 - 6\ell + 6)\right) + \frac{p^2}{2}\left(\ell^2 - 9\ell + 16\right)\Bigg) + O(p^3).$$

A Taylor expansion of $D(P_T^\circ \| P_L^*)$ around P_L^* gives

$$\begin{aligned} D(P_T^\circ \| P_L^*) &= \sum_{\ell=1}^{\infty} P_T^\circ(\ell)\left(\ell + \log_2 P_T^\circ(\ell)\right) \\ &= \frac{1}{\log 2}\sum_{\ell=1}^{\infty}\left(P_T^\circ(\ell) - 2^{-\ell}\right) + \frac{1}{2\log 2}\sum_{\ell=1}^{\infty} 2^\ell\left(P_T^\circ(\ell) - 2^{-\ell}\right)^2 + O(p^3) \\ &= \frac{1}{\log 2}\left(4x^2 + 4xp + 3p^2\right) + O(p^3), \end{aligned}$$

where we used $x = O(p)$ and $\left|P_T^\circ(\ell) - 2^{-\ell}\right| = O(p)$, and the following moments:

$$\mathbb{E}^*[\delta_1(L)] = 0$$
$$\mathbb{E}^*[\delta_2(L)] = 0$$
$$\mathbb{E}^*[\delta_1^2(L)] = 8$$

Also, we have

$$\mathbb{E}^\circ[T] = (1+p)(2 + O(p)) = 2 + O(p),$$

which gives the reciprocal of $\mathbb{E}^\circ[T]$ by

$$\frac{1}{\mathbb{E}^\circ[T]} = \frac{1}{2} + O(p).$$

Thus, we get

$$\begin{aligned}\frac{H(T)}{\mathbb{E}^\circ[L]} &= (1+p)\left(1-\frac{D(P_L^\circ\|P_L^*)}{\mathbb{E}^\circ[T]}\right)\\ &= 1+p-\frac{1}{2\log 2}\left(4x^2+4xp+3p^2\right)+O(p^3)\\ &= 1-\frac{2x^2}{\log 2}+p\left(1-\frac{2x}{\log 2}\right)-\frac{3p^2}{2\log 2}+O(p^3).\end{aligned}$$

Next, one can obtain

$$\begin{aligned}\mathbb{E}^\circ\left[\log_2\binom{L}{G}\right] &= p\,\mathbb{E}^*[L\log_2 L]+2xp\,\mathbb{E}^*[L(L-2)\log_2 L]\\ &\quad -p^2\,\mathbb{E}^*\left[L(L-1)\log_2 L-\tbinom{L}{2}\log_2\tbinom{L}{2}\right].\end{aligned}$$

The reciprocal of $\mathbb{E}^\circ[L]$ can be approximated in a similar way as

$$\frac{1}{\mathbb{E}^\circ[L]}=\frac{1}{2\,(1+2x+O(p^2))}=\frac{1}{2}-x+O(p^2).$$

Combining these results, we get

$$\begin{aligned}\frac{\mathbb{E}^\circ\left[\log_2\binom{L}{G}\right]}{\mathbb{E}^\circ[L]} &= p\left(\mathbb{E}^*[L\log_2(L)]+x\,\mathbb{E}^*[L(L-3)\log_2 L]\right)\\ &\quad -\frac{p^2}{2}\left(\mathbb{E}^*\left[L(L-1)\log_2 L-\tbinom{L}{2}\log_2\tbinom{L}{2}\right]\right)+O(p^3).\end{aligned}$$

Using Theorem 5.1, the information rate of process $\mathbb{X}^\circ$ will be

$$\begin{aligned}I(\mathbb{X}^\circ) &= 1-\frac{2x^2}{\log 2}+p\log_2 p+p\Big(\alpha_1+\frac{x}{\log 2}\left(\mathbb{E}^*[L(L-3)\log L]-2\right)\Big)\\ &\quad +\alpha_2^* p^2+O(p^3),\end{aligned}$$

where α_1 is the same as Theorem 5.2 and α_2^* is given by (5.8).

To find the optimal x, we take the derivative and obtain

$$x=\frac{p}{4}\left(\mathbb{E}^*[L(L-3)\log L]-2\right).$$

Therefore, the optimal p-dependent transition probability for a first-order Markov process is

$$\begin{aligned}\delta^* &= \frac{1}{2}+\Omega p,\\ \Omega &= \frac{1}{4}\,\mathbb{E}^*[L(L-3)\log L]-\frac{1}{2}\approx 0.10409610.\end{aligned}$$

If we plug the optimal value back into $I(\mathbb{X}^\circ)$, we obtain

$$I(\mathbb{X}^\circ) = 1 + p\log_2 p + \alpha_1 p + \alpha_2^\circ p^2 + O(p^3),$$

where

$$\alpha_2^\circ = \alpha_2^* + \frac{2\Omega^2}{\log 2} \approx -2.81592784.$$

5.5.6 Proof of Lemma 5.2

To simplify the analysis, it is important to note that the second term on the right-hand side of (5.4), i.e., $h(p)$, is symmetric about $p = \frac{1}{2}$. The third term is also symmetric around $p = \frac{1}{2}$ because $\binom{L}{G} = \binom{L}{L-G}$ and a binomial distribution is symmetric with respect to its success probability (see (5.17)). This means that the approximations in (5.16) and (5.17) are valid for $p \to 1$ if we change p to $q = 1 - p$. Therefore,

$$h(p) = -q\log_2 q + q\log_2 e + O(q^2), \tag{5.24}$$

and

$$\frac{\mathbb{E}^*\left[\log_2 \binom{L}{G}\right]}{\mathbb{E}^*[L]} = q\,\mathbb{E}^*[L\log_2 L] + O(q^2). \tag{5.25}$$

We are left with the first term of $I(\mathbb{X}^*)$ in Theorem 5.1. We derive an approximation of P_T^* for the even and odd values of t, separately. Let $t = 2k$, $k \geq 1$ and $q = 1 - p$. We have

$$\begin{aligned} P_T^*(2k) &= \sum_{\ell=k}^{2k} 2^{-\ell} \binom{\ell}{2k-\ell} (1-q)^{2k-\ell} q^{2\ell-2k} \\ &= \sum_{\ell=k}^{2k} 2^{-\ell} \binom{\ell}{2k-\ell} \left(1 - (2k-\ell)q + O(q^2)\right) q^{2\ell-2k} \\ &= 2^{-k}\,(1 - kq) + O(q^2). \end{aligned}$$

Now, for $t = 2k-1$, $k \geq 1$ we have

$$\begin{aligned} P_T^*(2k-1) &= \sum_{\ell=k}^{2k-1} 2^{-\ell} \binom{\ell}{2k-\ell-1} (1-q)^{2k-\ell-1} q^{2\ell-2k+1} \\ &= \sum_{\ell=k}^{2k-1} 2^{-\ell} \binom{\ell}{2k-\ell-1} \left(1 - (2k-\ell-1)q + O(q^2)\right) q^{2\ell-2k+1} \\ &= 2^{-k} kq + O(q^2). \end{aligned}$$

To obtain the entropy of the received run-length, we divide the summation into the odd and even parts as

$$H(T) = H_e(T) + H_o(T),$$

where

$$\begin{aligned}
H_e(T) &= -\sum_{k=1}^{\infty} P_T^*(2k) \log_2 P_T^*(2k) \\
&= -\sum_{k=1}^{\infty} \left(2^{-k}(1-kq) + O(q^2)\right) \log_2 \left(2^{-k}(1-kq) + O(q^2)\right) \\
&= -\sum_{k=1}^{\infty} \left(2^{-k}(1-kq) + O(q^2)\right) \left(\log_2 2^{-k} - kq \log_2 e + O(q^2)\right) \\
&= \sum_{k=1}^{\infty} \left(2^{-k}k + 2^{-k}kq \log_2 e - 2^{-k}k^2 q\right) + O(q^2) \\
&= 2\left(1 + q \log_2 e - 3q\right) + O(q^2),
\end{aligned}$$

and

$$\begin{aligned}
H_o(T) &= -\sum_{k=1}^{\infty} P_T^*(2k-1) \log_2 P_T^*(2k-1) \\
&= -\sum_{k=1}^{\infty} \left(2^{-k}kq + O(q^2)\right) \log_2 \left(2^{-k}kq + O(q^2)\right) \\
&= -\sum_{k=1}^{\infty} \left(2^{-k}kq + O(q^2)\right) \left(\log_2 q + \log_2(2^{-k}k) + O(q)\right) \\
&= -\sum_{k=1}^{\infty} \left(2^{-k}kq \log_2 q + 2^{-k}k \log_2(2^{-k}k)q\right) + O(q^2 \log_2 1/q) + O(q^2) \\
&= 2\left(-q \log_2 q - q \sum_{k=1}^{\infty} 2^{-k-1} k \log_2 k + 3q\right) + O(q^{2-\epsilon}),
\end{aligned}$$

where we used the fact that for small $\epsilon > 0$ and $q \to 0$, $\log_2 1/q = O(q^{-\epsilon})$. Thus,

$$\frac{H(T)}{\mathbb{E}^*[L]} = 1 - q \log_2 q + q(\log_2 e - \mathbb{E}^*[L \log_2 L]) + O(q^{2-\epsilon}). \tag{5.26}$$

By adding (5.24), (5.25) and (5.26), we get $I(\mathbb{X}^*) = 1 - O\left(q^{2-\epsilon}\right)$[6].

[6]The negative sign is used to emphasize on $C \leq 1$.

Chapter 6

Conclusion and Future Work

In this chapter, we summarize the contributions of the current thesis and present some possible future research directions.

6.1 Extremal Problems of Error Exponents

We solved the extremal problems of several classical error exponents over the set of MBIOS channels of the same capacity. The key ingredients of our analysis are the set of equal-capacity basis channels and the theory of Chebychev systems. It was shown that properly identified Chebychev systems show a detailed ordering among the set of basis channels. Using these orderings, we are able to solve extremal problems of error exponents with additional constraints. It was shown that the BEC and BSC are the two extremes of the error exponents considered in Chapter 4. The framework introduced in this thesis can be used to solve extremal problems of other error exponents such as the joint source-channel coding error exponent [66]. Moreover, it would be interesting to solve the extremal problems of error exponents using the calculus of variations.

The set of basis channels can be used for studying other extremal problems over the set of MBIOS channels of the same capacity. For example, the transmission of low-density parity-check (LDPC) codes takes place on an MBIOS channel [37]. It would be of a great practical and theoretical interest to find the extremal densities of one iteration of belief propagation decoding of LDPC codes [67]. This could lead to a universal design over all MBIOS channels of

the same capacity [43, 68]. In fact, this is a generalization of the information combining problem which addresses the extremal densities in a half iteration of belief propagation [34, 36, 52, 69, 70].

6.2 Capacity of Duplication Channels

The capacity of the duplication channel for small duplication probabilities was studied. We started by presenting an analytical lower bound on the duplication capacity. We showed that a Bernoulli($\frac{1}{2}$) process achieves the capacity up to term p in a series expansion (with an error of order $p^{3/2}$). To find a more refined expansion, two lower bounds for the next term in the series expansion were introduced. We believe that the one obtained from a perturbation in a Bernoulli($\frac{1}{2}$) process is the capacity, although we have not proved its converse statement. A similar analysis was carried out for $p \to 1$.

The i.i.d. duplication channel is a great model to analyze and gain insight into the behaviour of channels with synchronization errors. In practice, it is unlikely to have independent successive duplications. Therefore, it is more realistic to consider a channel model with memory for duplications. For example, the duplication process $\mathbb{Q}$ can be generated by a first-order Markov chain where with a high probability, the duplication state remains 0 and if it is in state 1, with a high probability it goes back to 0.

In this work, we studied the duplication capacity by means of a series expansion up to a term of order p^2. Finding a systematic way to update the optimal input distribution to achieve higher order terms is a great path to follow.

It is of great interest to study a duplication channel that also flips the bits with some probability. This is equivalent to a duplication channel concatenated by a BSC. Similar scenarios are studied in [12, 71–73].

Genie-aided bounds on the capacity, similar to the work of Fertonani and Duman [15], can be obtained for duplication channels. However, it can be shown that these bound are not as tight as the bounds introduced in [57]. Improving such genie-aided bounds for duplication channels when the dupli-

cation probability is neither close to zero nor close to one is another possible research direction.

Bibliography

[1] M. Mitzenmacher, "Capacity bounds for sticky channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 72–77, Jan. 2008.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. Journal*, vol. 27, pp. 379–423 and 623–656, Oct. 1948.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 2006.

[4] R. G. Gallager, *Information Theory and Reliable Communication.* New York, NY: John Wiley and Sons, 1968.

[5] A. Iyengar, P. Siegel, and J. Wolf, "Write channel model for bit-patterned media recording," *IEEE Trans. Magn.*, vol. 47, no. 1, pp. 35–45, Jan. 2011.

[6] M. Inoue and H. Kaneko, "Deletion/insertion/reversal error correcting codes for bit-patterned media recording," in *IEEE Int. Symp. Defect and Fault Tolerance in VLSI and Nanotech. Sys. (DFT)*, Oct. 2011, pp. 286–293.

[7] A. R. Krishnan and B. Vasic, "Coding for correcting insertions and deletions in bit-patterned media recording," in *IEEE Global Telecommn. Conf. (GLOBECOM)*, Dec. 2011, pp. 1–5.

[8] S. Zhang, K. Cai, M. Lin-Yu, J. Zhang, Z. Qin, K. K. Teo, W. E. Wong, and E. T. Ong, "Timing and written-in errors characterization for bit patterned media," *IEEE Trans. Magn.*, vol. 47, no. 10, pp. 2555–2558, Oct. 2011.

[9] M. Mitzenmacher, "A survey of results for deletion channels and related synchronization channels," *Probability Surveys*, vol. 6, pp. 1–33, 2009.

[10] A. Motahari, G. Bresler, and D. Tse, "Information theory of DNA sequencing," *submitted to IEEE Trans. Inf. Theory*, Jul. 2012.

[11] M. C. Davey and D. J. C. Mackay, "Reliable communication over channels with insertions, deletions, and substitutions," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 687–698, Feb. 2001.

[12] V. I. Levenshtein, "Binary codes capable of correcting spurious insertions and deletions of ones," *Problems Inf. Transmission,*, vol. 1, no. 1, pp. 12–25, 1965.

[13] S. Diggavi and M. Grossglauser, "On information transmission over a finite buffer channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1226–1237, Mar. 2006.

[14] S. Diggavi, M. Mitzenmacher, and H. D. Pfister, "Capacity upper bounds for the deletion channel," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Nice, Jun. 2007, pp. 1716–1720.

[15] J. Ullman, "On the capabilities of codes to correct synchronization errors," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 95–105, Jan. 1967.

[16] D. Fertonani and T. M. Duman, "Novel bounds on the capacity of the binary deletion channel," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2753–2765, Jun. 2010.

[17] M. Mitzenmacher and E. Drinea, "A simple lower bound for the capacity of the deletion channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4657–4660, Oct. 2006.

[18] N. J. A. Sloane, "On single-deletion-correcting codes," in *Ohio State University*, 2001, pp. 273–291.

[19] Z. Liu and M. Mitzenmacher, "Codes for deletion and insertion channels with segmented errors," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 224–232, Jan. 2010.

[20] E. Drinea and M. Mitzenmacher, "On lower bounds for the capacity of deletion channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4648–4657, Oct. 2006.

[21] E. A. Ratzer and D. J. MacKay, "Codes for channels with insertions, deletions and substitutions," in *2nd Int. Symp. on Turbo Codes and Related Topics*, 2000, pp. 149–156.

[22] S. Konstantinidis, "Relationships between different error-correcting capabilities of a code," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 2065–2069, Jul. 2001.

[23] V. I. Levenshtein, "Efficient reconstruction of sequences," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 2–22, Jan. 2001.

[24] K. S. Zigangirov, "Sequential decoding for a binary channel with dropouts and insertions," *Problems Inf. Transmission,*, vol. 5, no. 2, pp. 17–22, 1969.

[25] N. Vvedenskaya and R. Dobrushin, "The computation on a computer of the channel capacity of a line with symbol drop-out," *Problems Inf. Transmission,*, vol. 4, no. 3, pp. 76–79, 1968.

[26] A. Kavcic and R. Motwani, "Insertion/deletion channels: reduced-state lower bounds on channel capacities," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jun./Jul. 2004.

[27] R. Venkataramanan, S. Tatikonda, and K. Ramchandran, "Achievable rates for channels with deletions and insertions," *submitted to IEEE Trans. Inf. Theory*, Jul. 2011.

[28] H. Mirghasemi and A. Tchamkerten, "On the capacity of the one-bit deletion and duplication channel," in *Submitted to Allerton 2012*, Oct. 2012.

[29] R. L. Dobrushin, "Shannon's theorems for channels with synchronization errors," *Problems Inf. Transmission,*, vol. 3, no. 4, pp. 11–26, 1967.

[30] R. Yazdani, "Reliable communications under limited knowledge of the channel," Ph.D. dissertation, University of Alberta, Edmonton, AB, Canada, 2012.

[31] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, 3rd ed. Springer, 2004.

[32] T. Richardson and R. Urbanke, *Modern Coding Theory*. New York: Cambridge University Press, 2008.

[33] A. N. Kolmogorov, *Foundations of the Theory of Probability*, 2nd ed. Chelsea Pub. Co., Jun. 1960.

[34] N. Chayat and S. Shamai (Shitz), "Extension of an entropy property for binary input memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 1077–1079, Sep. 1989.

[35] I. Land, P. Hoeher, S. Huettinger, and J. Huber, "Bounds on information combining," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 612–619, Feb. 2005.

[36] C.-C. Wang, S. Kulkarni, and H. V. Poor, "Finite-dimensional bounds on and binary LDPC codes with belief propagation decoders," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 56–81, Jan. 2007.

[37] I. Sutskover, S. Shamai, and J. Ziv, "Constrained information combining: theory and applications for LDPC coded systems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1617–1643, May 2007.

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

[39] S. Karlin and W. J. Studden, *Tchebycheff systems: with applications in analysis and statistics*. New York: Interscience, 1966.

[40] M. G. Krěin and A. A. Nudeĺman, *The Markov Moment Problem and Extremal Problems*. Providence, Rhode Island: American Math. Society, 1977.

[41] R. A. Zalik, "Some properties of Chebyshev systems," *J. Comput. Analysis and Appl.*, vol. 13, pp. 20–26, 2011.

[42] M. Hellman and J. Raviv, "Probability of error, equivocation, and the Chernoff bound," *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 368–372, Jul. 1970.

[43] R. Gariepy and W. Ziemer, *Modern Real Analysis*. Brooks/Cole Pub. Co., 1994.

[44] A. Sanaei, M. Ramezani, and M. Ardakani, "Identical-capacity channel decomposition for design of universal LDPC codes," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 1972–1981, Jul. 2009.

[45] D. Williams, *Probability with Martingales*. Cambridge University Press, 2008.

[46] Y. Polyanskiy, "Channel coding: non-asymptotic fundamental limits," Ph.D. dissertation, Princeton University, Princeton, NJ, USA, 2010.

[47] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[48] R. L. Dobrushin, "Mathematical problems in the shannon theory of optimal coding of information," *4th Berkeley Symp. Math., Stat., and Prob.*, vol. 1, pp. 211–252, 1961.

[49] V. Strassen, "Asymptotische abschätzungen in shannon's informationstheorie," *Trans. 3rd Prague Conf. Inf. Theory*, pp. 689–723, 1962.

[50] A. Guillén i Fàbregas, I. Land, and A. Martinez, "Extremes of random coding error exponents," in *IEEE Int. Symp. Inf. Theory (ISIT)*, St. Petersburg, Russia, Jul. 2011.

[51] M. Alsan, "Extremality properties for Gallager's random coding exponent," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, USA, Jul. 2012, pp. 2954–2958.

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

[53] Y. Jiang, A. A., R. Koetter, and S. A. C., "Extremal problems of information combining," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 51–71, Jan. 2008.

[54] E. Arıkan, "Channel combining and splitting for cutoff rate improvement," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 628–639, Feb. 2006.

[55] N. Shulman and M. Feder, "Random coding techniques for nonrandom codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2101–2104, Sep. 1999.

[56] G. Miller and D. Burshtein, "Bounds on the maximum-likelihood decoding error probability of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2696–2710, Nov. 2001.

[57] J. Forney, G., "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, pp. 206–220, Mar. 1968.

[58] 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.

[59] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460–473, Jul. 1972.

[60] Y. Kanoria and A. Montanari, "On the deletion channel with small deletion probability," in *IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2010, pp. 1002–1006.

[61] M. Jimbo and K. Kunisawa, "An iteration method for calculating the relative capacity," *Inf. Contr.*, vol. 43, no. 2, pp. 216–223, 1979.

[62] K. A. S. Abdel-Ghaffar, "Capacity per unit cost of a discrete memoryless channel," *IEE Electronics Letters*, vol. 29, no. 2, pp. 142–144, Jan. 1993.

[63] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, Sep. 1990.

[64] Y. Kanoria and A. Montanari, "Optimal coding for the deletion channel with small deletion probability," *submitted to IEEE Trans. Inf. Theory*, May 2011.

[65] M. Mitzenmacher and E. Upfal, *Probability and Computing : Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[66] Y. Zhong, F. Alajaji, and L. L. Campbell, "On the joint source-channel coding error exponent for discrete memoryless systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1450–1468, Apr. 2006.

[67] I. Sutskover, "Iterative decoding of low-density parity-check codes over compound channels," Ph.D. dissertation, Technion-Israel Institute of Technology, Haifa, Israel, 2005.

[68] I. Sason and B. Shuval, "On universal LDPC code ensembles over memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5182–5202, Aug. 2011.

[69] I. Sutskover, S. Shamai, and J. Ziv, "Extremes of information combining," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1313–1325, 2005.

[70] I. Land and J. Huber, "Information combining," *Foundations and Trends in Commun. and Inf. Theory*, vol. 3, no. 3, pp. 227–330, Nov. 2006.

[71] M. Rahmati and T. M. Duman, "Achievable rates for noisy channels with synchronization errors," *submitted to IEEE Trans. Inf. Theory*, Mar. 2012.

[72] ——, "Analytical lower bounds on the capacity of insertion and deletion channels," *submitted to IEEE Trans. Inf. Theory*, Jun. 2012.

[73] H. Mercier, V. Tarokh, and F. Labeau, "Bounds on the capacity of discrete memoryless channels corrupted by synchronization and substitution errors," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4306–4330, Jul. 2012.

[74] I. Sason, "On universal properties of capacity-approaching LDPC code ensembles," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 2956–2990, Jul. 2009.

[75] R. Felowe and G. A. Harris, "A note on generalized Vandermonde determinants," *SIAM J. Matrix Anal. Appl.*, vol. 14, no. 4, pp. 1146–1151, Oct. 1993.

[76] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge, MA: MIT Press, 1963.

[77] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[78] S.-Y. Chung, "On the construction of some capacity-approaching coding schemes," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2000.

[79] S.-Y. Chung, J. Forney, G. D., T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 db of the Shannon limit," *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58–60, Feb. 2001.

[80] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Improved low-density parity-check codes using irregular graphs," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 585–598, Feb. 2001.

[81] T. Richardson and R. Urbanke, "Fixed points and stability of density evolution," *Commun. in inf. and sys.*, vol. 4, no. 1, pp. 103–116, 2004.

Appendix A

Proof of Theorem 3.1

This appendix is devoted to proving that the systems introduced in Theorem 3.1 are T–systems. We will take different approaches suited for each set. Before starting the proof of Theorem 3.1, we prove the following lemma:

Lemma A.1 [CT–Systems and ECT–Systems]**:** The system $\{1, u_1, \ldots, u_n\}$, $u_k \in \mathcal{C}^n([a,b])$, $k = 1, \ldots, n$, is a CT–system on $[a,b]$ if $\{u_1', \ldots, u_n'\}$ is an ECT–system on (a,b).

Proof: Consider the polynomial

$$f(t) = \mathsf{D}(1, u_1, \ldots, u_n; t_0, \ldots, t_{n-1}, t),\ a \le t_0 < \cdots < t_{n-1} < t \le b.$$

Clearly, $f(t_{n-1}) = 0$. An application of the mean-value theorem to f gives

$$\mathsf{D}(1, u_1, \ldots, u_n; t_0, \ldots, t_n) = (t_n - t_{n-1}) \begin{vmatrix} 1 & \cdots & 1 & 0 \\ u_1(t_0) & \cdots & u_1(t_{n-1}) & u_1'(\zeta_n) \\ \vdots & & \vdots & \vdots \\ u_n(t_0) & \cdots & u_n(t_{n-1}) & u_n'(\zeta_n) \end{vmatrix},$$

where $t_{n-1} < \zeta_n < t_n \le b$. Another application of mean-value theorem gives

$$\mathsf{D}(1, u_1, \ldots, u_n; t_0, \ldots, t_n) = (t_n - t_{n-1})(t_{n-1} - t_{n-2}) \times \begin{vmatrix} 1 & \cdots & 1 & 0 & 0 \\ u_1(t_0) & \cdots & u_1(t_{n-2}) & u_1'(\zeta_{n-1}) & u_1'(\zeta_n) \\ \vdots & & \vdots & \vdots & \vdots \\ u_n(t_0) & \cdots & u_n(t_{n-2}) & u_1'(\zeta_{n-1}) & u_n'(\zeta_n) \end{vmatrix},$$

where $t_{n-2} < \zeta_{n-1} < t_{n-1}$. Repeating the same procedure leads to

$$\mathsf{D}(1, u_1, \ldots, u_n; t_0, \ldots, t_n) = \mathsf{D}(u_1', \ldots, u_n'; \zeta_1, \ldots, \zeta_n) \prod_{k=1}^{n} (t_k - t_{k-1}),$$

where $a < \zeta_1 < \cdots < \zeta_n < b$. If $\{u'_1, \ldots, u'_n\}$ is an ECT–system on (a, b), it would also be a CT–system on (a, b) and $\mathsf{D}(u'_1, \ldots, u'_k; \zeta_1, \ldots, \zeta_k) > 0$ for $a < \zeta_1 < \cdots < \zeta_k < b$, $k = 1, \ldots, n$. Since the choice of $a \leq t_0 < \cdots < t_n \leq b$ is arbitrary, $\mathsf{D}(u'_1, \ldots, u'_k; \zeta_1, \ldots, \zeta_k) > 0$ for any $a < \zeta_1 < \cdots < \zeta_k < b$, $k = 1, \ldots, n$, and $\{u_1, \ldots, u_n\}$ is a CT–system on $[a, b]$. □

Lemma A.1 is the key step in proving Theorem 3.1. Note that

$$\mathsf{W}(1, u_1, \ldots, u_k)(p) = \mathsf{W}(u'_1, \ldots, u'_k)(p)$$

for $k = 1, \ldots, n$. Therefore, a system is a T–system on $[0, \frac{1}{2}]$ if we show that

$$\mathsf{W}(1, u_1, \ldots, u_k)(p) > 0, \ p \in (0, \tfrac{1}{2}), \ k = 1, \ldots, n. \tag{A.1}$$

To prove Theorem 3.1, we introduce a new variable defined as $q = \frac{1-p}{p}$, for $p \in (0, \frac{1}{2})$. Clearly, $q > 1$. After the change of variables, we will have functions of the form $u_k(q) = \frac{\lambda_k(q)}{1+q}$ for some continuous function λ_k. Using the Jacobian $\frac{dp}{dq} = -(1+q)^2$, we get

$$u'_k(q) = -(1+q)^2 \times \frac{d}{dq}\left(\frac{\lambda_k(q)}{1+q}\right).$$

In the same manner, the rest of the derivatives are

$$\begin{aligned} u'_k(q) &= \lambda_k(q) - (1+q)\lambda'_k(q) \\ u''_k(q) &= (1+q)^3\lambda''_k(q) \\ u^{(3)}_k(q) &= -(1+q)^4(3\lambda''_k(q) + (1+q)\lambda^{(3)}_k(q)). \end{aligned}$$

For example, for the binary entropy function

$$h(q) = \frac{(1+q)\log(1+q) - q\log q}{(1+q)\log 2}, \ q > 1.$$

Therefore,

$$\begin{aligned} h'(q) &= \frac{\log q}{\log 2}, \\ h''(q) &= -\frac{(1+q)^2}{q\log 2}, \\ h^{(3)}(q) &= \frac{(q-1)(1+q)^3}{q^2\log 2}. \end{aligned}$$

Let f be a continuous function on $q > 1$. A sufficient condition for f to be strictly positive is[1]

$$\lim_{q \to 1} f(q) = 0 \text{ and } f'(q) > 0. \tag{A.2}$$

We will encounter several instances of such functions for which we need to prove strict positiveness. In those cases, condition (A.2) will be used frequently to get the result. It is important to note that by change of variables, we do not change the interval over which the T–systems are defined. It is just the matter of calculation that we rather use q instead of p for some of the systems. We now prove Theorem 3.1.

$\mathcal{U}_0 : \{u_0, u_1\}$ where

$$u_0(p) = 1$$
$$u_1(p) = h(p).$$

Since u_1 is a strictly increasing function on $[0, \frac{1}{2}]$, for $0 \leq p_0 < p_1 \leq \frac{1}{2}$, we obtain

$$\mathsf{D}(1, u_1; p_0, p_1) = h(p_1) - h(p_0) > 0,$$

hence a T–system.

$\mathcal{U}_1 : \{u_0, u_2, u_2\}$ where

$$u_0(p) = 1$$
$$u_1(p) = h(p)$$
$$u_2(p) = p.$$

We use the polynomial characterization of T–systems (Remark 2.5). Let f_1 be a polynomial in the linear space spanned by $\mathcal{U}_2$, i.e., $f_1(p) = a_0 u_0(p) + a_1 u_1(p) + a_2 u_2(p)$ for $a_i \in \mathbb{R}$, $i = 0, 1, 2$. If $a_1 = 0$, then f_1 will be a linear function having at most one zero in $[0, \frac{1}{2}]$. If $a_1 \neq 0$, $f_1''(p) = a_1 h''(p)$ shows that f_1 is either convex or concave. Therefore, it cannot have more than two

[1]It is noteworthy that q is strictly greater than 1. Therefore, one may replace the limit condition with $\lim_{q \to 1^+} f(q) > 0$.

zeros in $[0, \frac{1}{2}]$, hence a T–system. An alternative proof is by the fact that $\mathsf{W}(1, u_1)(p) = h'(p) > 0$ on $[0, \frac{1}{2})$ and $\mathsf{W}(1, u_1, u_2)(p) = -h''(p) > 0$ on $[0, \frac{1}{2}]$, hence by condition (A.1), $\mathcal{U}_1$ is a T–system.

$\mathcal{U}_2 :\ \{u_0, u_1, u_2\}$ where

$$\begin{aligned} u_0(p) &= 1 \\ u_1(p) &= h(p) \\ u_2(p) &= -\left(p^{a_1} + (1-p)^{a_1}\right)^{b_1} \left(p^{a_2} + (1-p)^{a_2}\right)^{b_2}, \\ &\qquad 0 < a_i < 1,\ 0 < b_i,\ i = 1, 2,\ a_1 b_1 + a_2 b_2 = 1. \end{aligned}$$

We have

$$u_2(q) = -\frac{\left(1 + q^{a_1}\right)^{b_1} \left(1 + q^{a_2}\right)^{b_2}}{1+q},\ q > 1$$

and

$$\begin{aligned} \lambda_2'(q) &= -q^{-1} \left(1 + q^{a_1}\right)^{b_1 - 1} \left(1 + q^{a_2}\right)^{b_2 - 1} \\ &\qquad \times \left[a_1 b_1 q^{a_1} (1 + q^{a_2}) + a_2 b_2 q^{a_2} (1 + q^{a_1})\right] \\ \lambda_2''(q) &= q^{-2} \left(1 + q^{a_1}\right)^{b_1 - 2} \left(1 + q^{a_2}\right)^{b_2 - 2} \left[a_1 b_1 \bar{\alpha}_1 q^{a_1} (1 + q^{a_2})^2 \right. \\ &\qquad \left. + a_2 b_2 \bar{\alpha}_2 q^{a_2} (1 + q^{a_1})^2 + a_1 a_2 b_1 b_2 \left(q^{a_1} - q^{a_2}\right)^2 \right]. \end{aligned}$$

Therefore, the Wronskian can be written as

$$\begin{aligned} \mathsf{W}(\mathcal{U}_2)(q) &= u_1'(q) u_2''(q) - u_1''(q) u_2'(q) \\ &= \frac{(1+q)^2}{q \log 2} \left(q(1+q) \lambda_2''(q) \log q + \lambda_2(q) - (1+q) \lambda_2'(q)\right) \\ &= \frac{(1+q)^2}{q^2 \log 2} \left(1 + q^{a_1}\right)^{b_1 - 2} \left(1 + q^{a_2}\right)^{b_2 - 2} f_2(q), \end{aligned}$$

where

$$\begin{aligned} f_2(q) &= a_1 b_1 (1 + q^{a_2})^2 t^{a_1} f_{2, a_1}(q) + a_2 b_2 (1 + q^{a_1})^2 q^{a_2} f_{2, a_2}(q) \\ &\qquad + a_1 a_2 b_1 b_2 (1+q) \left(q^{a_1} - q^{a_2}\right)^2 \log q, \end{aligned}$$

and

$$f_{2,\alpha}(q) = (1-\alpha)(1+q) \log q - (1 + q^{\alpha})(q^{1-\alpha} - 1),\ \alpha \in [0, 1).$$

Note that we excluded $\alpha = 1$ because $f_{2,\alpha=1} \equiv 0$. It is clear that if $f_{2,\alpha}$ is strictly positive, then $\mathsf{W}(\mathcal{U}_2)(q) > 0$. In order to show that, we use condition (A.2): $\lim_{q\to 1} f_{2,\alpha} = 0$ and

$$f'_{2,\alpha}(q) = (1-\alpha)(\log q + 1 + q^{-1}) - ((1-\alpha)q^{-q-1} + 1 - \alpha q^{\alpha-1}).$$

An application of condition (A.2) to $f'_{2,\alpha}$, shows that $\lim_{q\to 1} f'_{2,\alpha}(q) = 0$ and

$$f''_{2,\alpha}(q) = \frac{1-\alpha}{q^2}\left(q - 1 - \alpha(q^\alpha - q^{1-\alpha})\right).$$

Now, one can see that for a fixed q, $\sup_\alpha \alpha(q^\alpha - q^{1-\alpha})$ is by $\alpha = 1$. Thus

$$\alpha(q^\alpha - q^{1-\alpha}) < q - 1,$$

which shows that $f''_{2,\alpha}$, $f'_{2,\alpha}$ and $f_{2,\alpha}$ are strictly positive, hence $\mathcal{U}_2$ is a T–system.

$\mathcal{U}_3$: $\{u_0, u_1, u_2, u_3\}$ where

$$\begin{aligned}
u_0(p) &= 1\\
u_1(p) &= h(p)\\
u_2(p) &= p\\
u_3(p) &= 2^{-\rho}\left(p^{\frac{1}{1+\rho}} + (1-p)^{\frac{1}{1+\rho}}\right)^{1+\rho},\ \rho > 0.
\end{aligned}$$

We know that $\mathsf{W}(1, u_1)$ and $\mathsf{W}(1, u_1, u_2)$ are strictly positive on $(0, \frac{1}{2})$. Also, it is straightforward to see that

$$\begin{aligned}
\lambda''_3(q) &= -\frac{\rho 2^{-\rho} q^{\frac{1}{1+\rho}-2}(1+q)^3}{1+\rho}\left(1 + q^{\frac{1}{1+\rho}}\right)^{\rho-1},\\
\lambda^{(3)}_3(q) &= \frac{\rho 2^{-\rho} q^{\frac{1}{1+\rho}-3}(1+q)^4}{(1+\rho)^2}\left(1 + q^{\frac{1}{1+\rho}}\right)^{\rho-2}\\
&\qquad \times \left[(\rho+2)\left(q - q^{\frac{1}{1+\rho}}\right) + (1+2\rho)\left(q^{\frac{1}{1+\rho}+1} - 1\right)\right].
\end{aligned}$$

We have

$$\begin{aligned}
\mathsf{W}(\mathcal{U}_3)(q) &= u^{(3)}_1(q)u''_3(q) - u^{(3)}_3(q)u''_1(q)\\
&= (1+q)^3\lambda''_3(q)\left(h^{(3)}(q) + 3(1+q)h''(q)\right) + (1+q)^5 h''(q)\lambda^{(3)}_3(q)\\
&= \frac{\rho 2^{-\rho} q^{\frac{1}{1+\rho}-4}(1+q)^6}{(1+\rho)^2}\left(1 + q^{\frac{1}{1+\rho}}\right)^{\rho-2}\left[q - q^{\frac{1}{1+\rho}} + r(q^{\frac{1}{1+\rho}+1} - 1)\right] > 0.
\end{aligned}$$

Therefore, $\mathcal{U}_3$ is a T–system.

$\mathcal{U}_4$: $\{u_0, u_1, u_2\}$ where

$$\begin{aligned} u_0(p) &= 1 \\ u_1(p) &= h(p) \\ u_2(p) &= 1 - (1-2p)^2. \end{aligned}$$

The binary entropy function can be expressed using the following Taylor series expansion [74]:

$$h(p) = 1 - \sum_{k=1}^{\infty} h_k (1-2p)^{2k}, \ h_k = \frac{1}{2k(2k-1)\log 2}, \ p \in [0,1].$$

Let $0 \leq p_0 < p_1 < p_2 \leq \frac{1}{2}$. By multiplying the second and third rows by -1 and adding the first row, we have

$$\begin{aligned} \mathsf{D}\left(u_0, u_1, u_2; p_0, p_1, p_2\right) &= \mathsf{D}\left(1, 1-u_1, 1-u_2; p_0, p_1, p_2\right) \\ &= \sum_{k=1}^{\infty} h_k \, \mathsf{D}\left(1, (1-2p)^{2k}, (1-2p)^2; p_0, p_1, p_2\right) \\ &\overset{(a)}{=} \sum_{k=2}^{\infty} h_k \, \mathsf{D}\left(1, (1-2p)^2, (1-2p)^{2k}; p_2, p_1, p_0\right) \\ &\overset{(b)}{=} \sum_{k=2}^{\infty} h_k \prod_{0 \leq i < j \leq 2} (a_j - a_i) \sum_{\substack{i_0+i_1+i_2=k-2 \\ i_0,i_1,i_2 \geq 0}} a_0^{i_0} a_1^{i_1} a_2^{i_2} \\ &> 0, \end{aligned}$$

where (a) follows from switching rows 2 and 3 and columns 1 and 3, and (b) follows from letting $a_i = (1-2p_{2-i})^2, \ i = 0, 1, 2$, $a_0 < a_1 < a_2$, and using the generalized Vandermonde determinant [75]. Thus, $\mathcal{U}_4$ is a T–system.

$\mathcal{U}_5 : \{u_0, u_1, u_2\}$ where

$$\begin{aligned} u_0(p) &= 1 \\ u_1(p) &= h(p) \\ u_2(p) &= \frac{1}{2}\left(\sqrt{p} + \sqrt{\bar{p}}\right)\left(\sqrt{p}\log\sqrt{p} + \sqrt{\bar{p}}\log\sqrt{\bar{p}}\right) \\ &\quad + \frac{1}{2}\left(\sqrt{p} + \sqrt{\bar{p}}\right)^2 \log\frac{2}{\sqrt{p} + \sqrt{\bar{p}}}. \end{aligned}$$

By change of variables, we have

$$\lambda_2(q) = \frac{1}{2}\left(1+\sqrt{q}\right)\left(\log 4 - 2\left(1+\sqrt{q}\right)\log\left(1+\sqrt{q}\right) + \sqrt{q}\log(4q)\right)$$
$$\lambda_2'(q) = \frac{1}{4}q^{-\frac{1}{2}}\left(-4(q-1)\log\left(1+\sqrt{q}\right) - 2\sqrt{q}\log q + (q-1)\log(16q)\right)$$
$$\lambda_2''(q) = \frac{1}{8}q^{-\frac{3}{2}}(1+q)^3\left(2 + 4\log\left(1+\sqrt{q}\right) - \log(16q)\right).$$

The Wronskian can be written as

$$\mathsf{W}(\mathcal{U}_5)(q) = \frac{q^{-\frac{3}{2}}}{4\log 2}(1+q)^2 f_5(q),$$

where

$$f_5(q) = (\sqrt{q}-1)^2\log q + 2\log\left(\frac{1+\sqrt{q}}{2q^{\frac{1}{4}}}\right) f_5^\circ(q),$$

and $f_5^\circ(q) = 2 - 2q + (1+q)\log q$. It is easy to check that by condition A.1, $f_5^\circ > 0$. Since $1+\sqrt{q} > 2q^{\frac{1}{4}}$, the Wronskian is positive and $\mathcal{U}_5$ is a T–system.

$\mathcal{U}_6:\ \{u_0, u_1, u_2\}$ where

$$u_0(p) = 1$$
$$u_1(p) = h(p)$$
$$u_2(p) = -p^s(1-p)^{1-s} - p^{1-s}(1-p)^s,\ s \in (0,1).$$

We have

$$u_2(q) = -\frac{q^s + q^{1-s}}{1+q},$$

$\lambda_2'(q) = -q^{-1}(sq^s + (1-s)q^{1-s})$, and $\lambda_2''(q) = s(1-s)q^{-2}(q^s + q^{1-s})$. The Wronskian can be expressed as

$$\mathsf{W}(\mathcal{U}_6)(q) = \frac{(1+q)^2}{q^2} f_6(q),$$

where

$$f_6(q) = s(1-s)(1+q)(q^s+q^{1-s})\log q + (1+q)(sq^s + (1-s)q^{1-s}) - q(q^s+q^{1-s}).$$

Note that applying condition (A.2), $\lim_{q\to 1} f_6(q) = 0$ and

$$f_6'(q) = (1-s)q^{-s} f_{6,s}(x)\big|_{x=q^{2s}} + sq^{-(1-s)} f_{6,1-s}(x)\big|_{x=q^{2(1-s)}},$$

where

$$f_{6,s}(x) = ((1+s)x + 1 - s)\log\sqrt{x} - x + 1,\ x > 1.$$

Again, applying condition (A.2), we have $\lim_{x\to 1} f_{6,s}(x) = 0$ and

$$\begin{aligned} f'_{6,s}(x) &= \frac{1}{2}\left[(1+s)\log x + 1 + s + (1-s)x^{-1}\right] - 1 \\ &\geq \frac{1}{2}\left[(1+s)(1-x^{-1}) + 1 + s + (1-s)x^{-1}\right] - 1 \\ &= s(1-x^{-1}) > 0. \end{aligned}$$

Therefore, $f_{6,s}$, f'_6, and f_6 are strictly positive meaning that $\mathcal{U}_6$ is a T–system.

$\mathcal{U}_7:\ \{u_0, u_1, u_2\}$ where

$$\begin{aligned} u_0(p) &= 1 \\ u_1(p) &= h(p) \\ u_2(p) &= -p\log_2^2 p - (1-p)\log_2^2(1-p). \end{aligned}$$

By change of variables, we have

$$u_2(q) = -\frac{\log^2(1+q) + q\log^2(1+q^{-1})}{(1+q)\log^2 2},$$

and

$$\begin{aligned} \lambda'_2(q) &= -\frac{(1+q)\log^2(1+q^{-1}) + 2\log q}{(1+q)\log^2 2} \\ \lambda''_2(q) &= \frac{2(1+q)(\log(q+1) - 1) - 2\log q}{q(1+q)^2\log^2 2}. \end{aligned}$$

The Wronskian is obtained by

$$\begin{aligned} \mathsf{W}(\mathcal{U}_7)(q) &= \frac{(1+q)^2}{q}\left[q(1+q)\lambda''_2(q)\log q + \lambda_2(q) - (1+q)\lambda'_2(q)\right] \\ &= \frac{q^2-1}{q\log^3 2}\log^2 q > 0. \end{aligned}$$

Therefore, $\mathcal{U}_7$ is a T–system.

$\mathcal{U}_8:\ \{u_0, u_1, u_2, u_3\}$ where

$$\begin{aligned} u_0(p) &= 1 \\ u_1(p) &= h(p) \\ u_2(p) &= p \\ u_3(p) &= p\log_2^2 p + (1-p)\log_2^2(1-p). \end{aligned}$$

Similar to $\mathcal{U}_7$, we obtain

$$\begin{aligned}
\lambda_3'(q) &= \frac{(1+q)\log^2(1+q^{-1}) + 2\log q}{(1+q)\log^2 2} \\
\lambda_3''(q) &= -\frac{2(1+q)(\log(q+1)-1) - 2\log q}{q(1+q)^2\log^2 2} \\
\lambda_3^{(3)}(q) &= \frac{2(1+q)(1+2q)\log(1+q) - 2(1+3q)\log q - 3q(1+q)}{q^2(1+q)^3\log^2 2}.
\end{aligned}$$

The Wronskian is now

$$\begin{aligned}
\mathsf{W}(\mathcal{U}_8)(q) &= u_1^{(3)}(q)u_3''(q) - u_1''(q)u_3^{(3)}(q) \\
&= (1+q)^3\lambda_3''(q)\left(h^{(3)}(q) + 3(1+q)h''(q)\right) + (1+q)^5 h''(q)\lambda_3^{(3)}(q) \\
&= \frac{2(1+q)^4}{q^3\log^3 2}(q^2 - 1 + q\log q) > 0.
\end{aligned}$$

Hence, $\mathcal{U}_8$ is a T–system.

Appendix B

A Note Regarding Theorem 4.12

For $w \in [0,1],\ 0 \le s \le \rho \le 1,\ p \in [0, \frac{1}{2}]$, define

$$\theta(w) = \left(wp^{1-s} + (1-w)(1-p)^{1-s}\right)\left(wp^{s/\rho} + (1-w)(1-p)^{s/\rho}\right)^{\rho}.$$

We would like to find the minimum value of $\zeta(w) = \theta(w) + \theta(1-w)$. Clearly, $\zeta(0) = \zeta(1) = 1$. We have

$$\begin{aligned}
\theta'(w) &= z^{\rho-1}(w)\big[(1-p)^{s/\rho}(p^{1-s} - (1-p)^{1-s}) + \rho(1-p)^{1-s}(p^{s/\rho} - (1-p)^{s/\rho}) \\
&\qquad + w(1+\rho)(p^{1-s} - (1-p)^{1-s})(p^{s/\rho} - (1-p)^{s/\rho})\big] \\
&= z^{\rho-1}(w)\big[p^{s/\rho}(p^{1-s} - (1-p)^{1-s}) + \rho p^{1-s}(p^{s/\rho} - (1-p)^{s/\rho}) \\
&\qquad - \bar{w}(1+\rho)(p^{1-s} - (1-p)^{1-s})(p^{s/\rho} - (1-p)^{s/\rho})\big],
\end{aligned}$$

where $z(w) = wp^{s/\rho} + (1-w)(1-p)^{s/\rho}$. For $w < \frac{1}{2}$, $z(w) > z(1-w)$. We obtain

$$\begin{aligned}
\zeta'(w) &= \theta'(w) - \theta'(1-w) \\
&< z^{\rho-1}(w)(2w-1)(1+\rho)(p^{1-s} - (1-p)^{1-s})(p^{s/\rho} - (1-p)^{s/\rho}) \\
&\le 0.
\end{aligned}$$

It is important to note that for $w < \frac{1}{2}$ and $p \neq \frac{1}{2}$, $\zeta' < 0$. We would not worry about the case of $p = \frac{1}{2}$ since it corresponds to a useless channel. Similarly for

$w > \frac{1}{2}$, $z(w) < z(1-w)$ and we have

$$\begin{aligned}\zeta'(w) &= \theta'(w) - \theta'(1-w) \\ &> z^{\rho-1}(1-w)(2w-1)(1+\rho)(p^{1-s} - (1-p)^{1-s})(p^{s/\rho} - (1-p)^{s/\rho}) \\ &\geq 0,\end{aligned}$$

which means that ζ decreases for $0 \leq w < \frac{1}{2}$ and increases for $\frac{1}{2} < w \leq 1$. Since $\zeta'(\frac{1}{2}) = 0$, we conclude that ζ is minimized at $w = \frac{1}{2}$.

Appendix C

Limit of $\mathbb{E}[\mu(P,\rho)]$ and Its Derivatives

In this appendix, we prove the following lemma:

Lemma C.1 [Limits of $\mathbb{E}[\mu(P,\rho)]$ and Its Derivatives]**:** For the function μ defined in Lemma 4.1, the following hold for $\rho^* \in [0,1]$:

(a) $\lim_{\rho\to\rho^*} \mathbb{E}[\mu(P,\rho)] = \mathbb{E}[\mu(P,\rho^*)]$

(b) $\lim_{\rho\to\rho^*} \frac{\partial}{\partial\rho}\mathbb{E}[\mu(P,\rho)]$ exists and equals $\mathbb{E}\left[\lim_{\rho\to\rho^*}\frac{\partial}{\partial\rho}\mu(P,\rho)\right]$.

(c) $\lim_{\rho\to\rho^*} \frac{\partial^2}{\partial\rho^2}\mathbb{E}[\mu(P,\rho)]$ exists and equals $\mathbb{E}\left[\lim_{\rho\to\rho^*}\frac{\partial^2}{\partial\rho^2}\mu(P,\rho)\right]$.

Before proceeding with the proof of Lemma C.1, we state the following lemma which will prove handy later.

Lemma C.2 [Integrability]**:** The random variables $\mu(P,\rho)$, $\frac{\partial}{\partial\rho}\mu(P,\rho)$ and $\frac{\partial^2}{\partial\rho^2}\mu(P,\rho)$ are bounded and integrable for any $\rho \in [0,1]$.

Proof: Clearly, $\mu(P,\rho)$ is bounded by one, hence integrable. We have

$$\frac{\partial}{\partial\rho}\mu(P,\rho) = -2^{-\rho}\left(P^{\frac{1}{1+\rho}} + \bar{P}^{\frac{1}{1+\rho}}\right)^{\rho}\zeta_1(P,\rho),$$

where

$$\zeta_1(p,\rho) = p^{\frac{1}{1+\rho}}\log p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}\log\bar{p}^{\frac{1}{1+\rho}} + \left(p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}\right)\log\frac{2}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}}.$$

Using the Jensen's inequality for the convex function $x \log x$, $x \in [0, \frac{1}{2}]$, it is straightforward to see that the function ζ_1 is non-negative. Moreover,

$$\zeta_1(p, \rho) \leq \left(p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}} \right) \log 2.$$

Therefore, we have

$$\left| \frac{\partial}{\partial \rho} \mu(P, \rho) \right| \leq \log 2 \times \mu(P, \rho), \ \rho \in [0, 1]$$

almost surely. For the second derivative, we obtain

$$\frac{\partial^2}{\partial \rho^2} \mu(P, \rho) = \mu(P, \rho) \frac{\zeta_2(P, \rho)}{(1+\rho)^3},$$

where

$$\begin{aligned} \zeta_2(p, \rho) = & \left(\frac{\bar{p}^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}} \right)^2 \times \rho \log^2 \frac{\bar{p}}{p} + (1+\rho)^3 \log^2 \frac{2p^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}} \\ & + \frac{\bar{p}^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}} \times \log \frac{\bar{p}}{p} \left(\log \frac{\bar{p}}{p} + 2(1+\rho)^2 \log \frac{2p^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}} \right) \end{aligned}$$

is apparently non-negative. Furthermore, we obtain

$$\begin{aligned} \frac{\zeta_2(p, \rho)}{(1+\rho)^3} \leq & \frac{\bar{p}^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}} \log^2 \frac{\bar{p}^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}}} + \log^2 \frac{2p^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}} \\ & + \frac{2\bar{p}^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}} \log \frac{\bar{p}^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}}} \log \frac{2p^{\frac{1}{1+\rho}}}{p^{\frac{1}{1+\rho}} + \bar{p}^{\frac{1}{1+\rho}}}. \end{aligned}$$

Let $x = \left(\frac{1-p}{p} \right)^{\frac{1}{1+\rho}}$, $x \geq 1$. We have

$$\begin{aligned} \frac{\zeta_2(p, \rho)}{(1+\rho)^3} \leq \zeta_2(x) &= \frac{x}{1+x} \log^2 x + \log^2 \frac{2}{1+x} + \frac{2x}{1+x} \log x \log \frac{2}{1+x} \\ &= \left(\frac{x}{x+1} \log x + \log \frac{2}{1+x} \right)^2 + \frac{x}{(1+x)^2} \log^2 x, \ x \geq 1, \end{aligned}$$

where $\zeta_2(x)$ is continuous and non-negative for $x \geq 1$, $\zeta_2(1) = 0$ and

$$\lim_{x \to \infty} \zeta_2(x) = \log^2 2.$$

Therefore, it is bounded. In fact, simple calculus shows that $x^* = 2e(\sqrt{e^2 - 1} + e) - 1 \approx 27.51$ maximizes ζ_2 and $\zeta_2(x^*) \approx 0.66 < \log 2$. Thus, the second derivative is bounded by an integrable random variable, i.e.,

$$\left| \frac{\partial^2}{\partial \rho^2} \mu(P, \rho) \right| \leq \log 2 \times \mu(P, \rho), \ \rho \in [0, 1]$$

almost surely.

□

Proof: [Lemma C.1] (a) holds by DCT and Lemma C.2. Note that treating $\rho \to \rho^*$ using sequences in n is accomplished by any monotone sequence $\{\rho_n\} \uparrow \rho^*$ or $\{\rho_n\} \downarrow \rho^*$.

For (b) and (c), we use [44, Section A16.1] to interchange (multiple) differentiation and the expected value operator. Once we took the derivatives inside the expectation, the proof will be complete by using DCT (via Lemma C.2). To apply [44, Section A16.1], we have to show that two classes which are $\mathcal{P}_b = \{\mu(P, \rho), \rho \in [0, 1]\}$ and $\mathcal{P}_c = \{\int_0^\rho \mu(P, s)ds, \rho \in [0, 1]\}$ are uniformly integrable. In both cases, uniform integrability is clear by the fact that both $\mathcal{P}_b$ and $\mathcal{P}_c$ are bounded[1].

□

[1]Note that in order to apply [44, Section A16.1] in part (c), one has to reapply [44, Section A16.1] to $\mathcal{P}_c$.

Appendix D

Channel Dispersion Revisited

In this section, with a slight abuse of notation, we write $P_{Y|X=x}$ for the measure $P_{Y|X}(\cdot|x)$. The information density of a pair of random variables X and Y measured in bits and defined on $\mathcal{X}$ and $\mathcal{Y}$ is a random variable defined as

$$i(x,y) = \log_2 \frac{dP_{XY}}{d(P_X \times P_Y)}(x,y) = \log_2 \frac{dP_{Y|X=x}}{dP_Y}(y),$$

with the understanding that if $P_{Y|X=x}$ is not absolutely continuous with respect to P_Y, we define $i(x,y) = +\infty$ for all y in the singular set, and we define $i(x,y) = -\infty$ for any y such that $\frac{dP_{Y|X=x}}{dP_Y} = 0$ [46].

For an MBIOS channel for which the input distribution is uniform, we have $P_Y(y) = \frac{1}{2}P_{Y|X=+1}(y) + \frac{1}{2}P_{Y|X=-1}(y)$. Therefore, $P_{Y|X=x} \ll P_Y$, i.e., the measure $P_{Y|X=x}$, $x = \pm 1$, is absolutely continuous with respect to P_Y.

Theorem D.1 [Sufficient Statistic and Information Density]**:** Let $P_{Y|X} : \mathcal{X} \mapsto \mathcal{Y}$ be an MBIOS channel with uniform input distribution and L be the LLR at its output. The information density between X and Y is equal to the information density between X and L, i.e., $i(X,Y) = i(X, L(Y))$. Moreover,

$$\mathbb{E}\left[\varphi(i(X,Y))\right] = \mathbb{E}\left[\varphi\left(\log_2 \tfrac{2}{1+e^{-L}}\right)\right],$$

holds for every continuos function φ.

Proof: Let x be either $+1$ or -1. Define

$$\mathcal{Y}_x = \{y \in \mathcal{Y} : P_{Y|X=x}(y) \neq 0\}.$$

We assume that all output alphabets can be reached with a positive probability, i.e., $\mathcal{Y}_{+1} \cup \mathcal{Y}_{-1} = \mathcal{Y}$. If $y \in \mathcal{Y}_x$, we have

$$\begin{aligned} i(x,y) &= \log_2 \frac{dP_{Y|X=x}}{dP_Y}(y) \\ &= \log_2 \left(\frac{1}{2} \frac{d(P_{Y|X=x} + P_{Y|X=-x})}{dP_{Y|X=x}}(y) \right)^{-1} \\ &= \log_2 \frac{2}{1+e^{-xL(y)}}. \end{aligned}$$

Otherwise, $i(x,y) = -\infty$. Let λ be the Lebesgue measure on $\mathbb{R}$. On the other hand, we have

$$\begin{aligned} i(+1,l) &= \log \frac{dP_{L|X=+1}}{dP_L}(l) \\ &= \log \frac{dP_{L|X=+1}}{d\lambda}(l) - \log \frac{dP_L}{d\lambda}(l) \\ &= \log \mathsf{a}(l) - \log \left(\tfrac{1}{2}\mathsf{a}(-l) + \tfrac{1}{2}\mathsf{a}(l) \right) \\ &= \log \frac{2}{1+e^{-l}}. \end{aligned}$$

Similarly, $i(-1,l) = \log \frac{2}{1+e^{l}}$. For $y \notin \mathcal{Y}_x$, $i(x,y) = -\infty$ corresponds to $i(x,l(y)) = -\infty$. This shows that $i(X,Y) = i(X,L(Y))$ almost surely.

Finally, We have

$$\begin{aligned} \mathbb{E}\left[\varphi(i(X,Y))\right] &= \mathbb{E}\left[\varphi(i(X,L(Y)))\right] \\ &= \frac{1}{2}\int \varphi(i(+1,l(y)))\mathsf{a}(l)dl + \frac{1}{2}\int \varphi(i(-1,l(y)))e^{-l}\mathsf{a}(l)dl \\ &= \frac{1}{2}\int \varphi(i(+1,l(y)))\mathsf{a}(l)dl + \frac{1}{2}\int \varphi(i(+1,-l(y)))e^{-l}\mathsf{a}(l)dl \\ &= \int \varphi(i(+1,l(y)))\mathsf{a}(l)dl \\ &= \mathbb{E}\left[\varphi\left(\log_2 \tfrac{2}{1+e^{-L}}\right)\right]. \end{aligned}$$

□

Note that using Lemma 3.1, one can get the P-density representation of Theorem D.1. For a memoryless channel, the channel capacity is $C = \sup_{P_X} \mathbb{E}[i(X;Y)]$. Also, according to Strassen [48, Theorem 1.2], channel dispersion is expressed as the minimum variance of the information density, i.e.,

$$V = \inf_{X:\ C=\mathbb{E}[i(X;Y)]} \mathbb{V}[i(X;Y)].$$

According to Theorem D.1, for an MBIOS channel, $C = \mathbb{E}[\log_2 \frac{2}{1+e^{-L}}]$ which is the same as (2.6). To calculate the variance of the information density with the capacity achieving input distribution, we take the same steps and arrive at the following expression for the channel dispersion in squared bits per channel use:

$$\begin{aligned} V &= \mathbb{E}[i^2(X;Y)] - \mathbb{E}^2[i(X;Y)] \\ &= \mathbb{E}\left[\log_2^2 \tfrac{2}{1+e^{-L}}\right] - C^2 \\ &= \mathbb{E}\left[\log_2^2(1+e^{-L})\right] - H^2. \end{aligned}$$

Using Lemma 3.1, we obtain exactly the same expression as the one we obtained in Section 4.2.5. Although having proved it for $n = 1, 2$, we conjecture that all cumulants of information density can be obtained via the derivatives of $E_0(\rho)$:

Conjecture D.1 [Cumulants of Information Density]**:** The following holds:

$$\kappa_n(i(X,Y)) = -\lim_{\rho \to 0} \frac{\partial^n}{\partial \rho^n} E_0(\rho),$$

where $\kappa_n(Z)$ is the nth cumulant of random variable Z.

Appendix E

Basis Channels in Other Applications

We introduced the set of basis channels in [43] for design of universal low-density parity-check (LDPC) codes [37, 76, 77] whose transmission takes place over MBIOS channels. Universal codes are of great practical and theoretical interest. On the other hand, LDPC codes are extremely powerful and their performance, if properly designed, can be very close to the channel capacity [78, 79]. Here, we briefly mention the importance of the set of basis channels in the design of universal LDPC codes.

The primary characteristic of an ensemble of LDPC codes is a pair of degree distributions (λ, ρ) which shows how the code is constructed [80]. By convergence of an LDPC code over an MBIOS channel, we mean that the probability of error under iterative belief propagation vanishes as the number of iterations goes to infinity. Based on strong supporting evidences, we conjectured in [43] that if an LDPC code converges on two MBIOS channels, it does so on any channel from the convex hull of those channels. Note that the L-density of a channel in the convex hull of two symmetric channels is a convex combination of the L-densities of those channels. Since the basis channels span the space of equal-capacity MBIOS channels, we come to the conclusion that if an LDPC code converges on the set of basis channels, it does converge on any MBIOS channel of the same capacity. A very high percentage of the capacity is achieved using this method [43]. In this appendix, assuming that the density evolution process is monotone [31], we prove the conjecture we

made in [43]. For information on channel degradation and density evolution, refer to [31].

Theorem E.1 [Convergence over the Convex Hull]**:** Let (λ, ρ) be a pair of degree distributions converging on two equal-capacity MBIOS channels with L-densities a and b. Under the assumption of monotonicity of density evolution, (λ, ρ) converges on any channel from the convex hull of a and b.

Proof: Fix s and let c be a channel from the convex hull of a and b such that $\mathsf{c} = s\mathsf{a} + (1-s)\mathsf{b}$. We prove this theorem by contradiction: assume that $\mathsf{f} \neq \Delta_\infty$ is the fixed point of the density evolution of c (see [31, Theorem 4.119]). According to [81, Theorem 2], $\mathsf{f} \hookrightarrow \mathsf{c}$, i.e., c is physically degraded with respect to the fixed point of its density evolution. We show that both a and b are physically degraded with respect to f. The transition matrix of the channel c can be written as

$$\mathbf{C} = [s\mathbf{A} \mid (1-s)\mathbf{B}],$$

where the output of each channel can be relabelled even if they represent the same symbol (a simple example in the case of basis channels is shown in Fig. 3.1). Let $\mathbf{F}$ be the transition matrix of f. From $\mathsf{f} \hookrightarrow \mathsf{c}$, there exist a channel w with transition matrix $\mathbf{W}$, such that $\mathbf{C} = \mathbf{F}\mathbf{W}$. It is straightforward to see that $\mathbf{W}$ can be decomposed as $\mathbf{W} = [s\mathbf{W}_1 \mid (1-s)\mathbf{W}_2]$ which means that both a and b are degraded with respect to f.

For some starting density d, let $\mathsf{a}_\ell(\mathsf{d})$ and $\mathsf{b}_\ell(\mathsf{d})$ denote the L-density of the ℓth iteration of density evolution of a and b, respectively. According to [31, Lemma 4.105], since a and b are degraded with respect to f, for any $\ell \geq 0$ we have

$$\mathsf{a}_\ell(\mathsf{f}) \hookrightarrow \mathsf{a}_\ell(\mathsf{a}) \text{ and } \mathsf{b}_\ell(\mathsf{f}) \hookrightarrow \mathsf{b}_\ell(\mathsf{b}).$$

Now, by [31, Lemma 4.106] and letting the number of iterations grow, we get

$$\lim_{\ell\to\infty} \mathbb{P}(\mathsf{a}_\ell(\mathsf{f})) = \lim_{\ell\to\infty} \mathbb{P}(\mathsf{b}_\ell(\mathsf{f})) = 0,$$

where we used the convergence of the code over both a and b. This implies that if we start the density evolution of either a or b from the fixed point of

c, we will have convergence too. However, after one iteration of the density evolution of c starting from its fixed point f, we obtain [31]

$$\begin{aligned}
\mathbb{P}(\mathsf{f}) &= \mathbb{P}(\mathsf{c}_1(\mathsf{f})) \\
&= \mathbb{P}(\mathsf{c} \circledast \lambda(\rho(\mathsf{f}))) \\
&= \mathbb{P}((s\mathsf{a} + (1-s)\mathsf{b}) \circledast \lambda(\rho(\mathsf{f}))) \\
&= s\,\mathbb{P}(\mathsf{a} \circledast \lambda(\rho(\mathsf{f}))) + (1-s)\,\mathbb{P}(\mathsf{b} \circledast \lambda(\rho(\mathsf{f}))) \\
&= s\,\mathbb{P}(\mathsf{a}_1(\mathsf{f})) + (1-s)\,\mathbb{P}(\mathsf{b}_1(\mathsf{f})) \\
&\leq \max\{\mathbb{P}(\mathsf{a}_1(\mathsf{f})), \mathbb{P}(\mathsf{b}_1(\mathsf{f}))\} \\
&< \mathbb{P}(\mathsf{f}),
\end{aligned}$$

which contradicts the initial assumption. Thus, $\mathsf{f} = \Delta_\infty$ and by [31, Theorem 4.119], the pair (λ, ρ) is convergent over the channel c. Since the choice of c is arbitrary, it follows that the pair (λ, ρ) is convergent over the convex hull of a and b.

□