Channel coding: non-asymptotic fundamental limits

Yury Polyanskiy

A Dissertation

Presented to the Faculty

of Princeton University

in Candidacy for the Degree

of Doctor of Philosophy

Recommended for Acceptance

by the Department of

Electrical Engineering

Advisers: H. Vincent Poor and Sergio Verdú

November, 2010

Abstract

Noise is an inalienable property of all communication systems appearing in nature. Such noise acts against the very purpose of communication, that is delivery of a data to the destination with minimal possible distortion. This creates a problem that has been addressed by various disciplines over the past century. In particular, information theory studies the question of the maximum possible rate achievable by an ideal system under certain assumptions regarding the noise generation and structural design constraints. The study of such questions, initiated by Claude Shannon in 1948, has typically been carried out in the asymptotic limit of an infinite number of signaling degrees of freedom (blocklength). Such a regime corresponds to the regime of laws of large numbers, or more generally ergodic limits, in probability theory. However, with the ever increasing demand for ubiquitous access to real time data, such as audio and video streaming for mobile devices, as well as the advent of modern sparse graph codes, one is interested in describing fundamental limits non-asymptotically, i.e. for blocklengths of the order of 1000. Study of these practically motivated questions requires new tools and techniques, which are systematically developed in this work. Knowledge of the behavior of the fundamental limits in the non-asymptotic regime enables the analysis of many related questions, such as the energy efficiency, effects of dynamically varying channel state, assessment of the suboptimality of modern codes, benefits of feedback, etc. As a result it is discovered that in several instances classical (asymptotics-based) conclusions do not hold under this more refined approach.

To Olga

Acknowledgements

I owe my deepest gratitude to my advisers: Prof. Vincent Poor for his wisdom and omnipresent support, and Prof. Sergio Verdú for teaching me to see and love the elegance of information theory behind the wall of definitions and theorems. I would like to thank the entire faculty of the Department of Electrical Engineering for creating a great learning atmosphere. I am especially indebted to Prof. Robert Calderbank for opening to me a beautiful world of algebra and for always finding time to discuss my endless questions, and to Prof. Erhan Çinlar for his warm guidance through the rugged terrain of all things stochastic. I appreciate the valuable remarks of Prof. Sanjeev Kulkarni and Prof. Paul Cuff. I am also grateful to the faculty of the Department of Mathematics for inspiration and intellectual stimuli.

I am equally indebted to my colleagues who were always ready to be either technical or caring depending on the circumstances: Eugene Brevdo, Yihong Wu, Aman Jain, Lorne Applebaum, Vaneet Aggarwal, Ankit Gupta, Maria Fresia, Sharon Betz and Arvid Wang. My great friends, Konstantin Mukhanov, Vladimir Tropin, Alex Kovalenko, Sergey Morozov, Dmitry Lakontsev, Sergey Paryshev, Konstantin Kravtsov, Grigory Ovanesyan, Andrei Malashevich, Evgeny Andriyash, and Alexey Soluyanov have been the true foundation for my life on which I could rest upon in the hard times. I also cannot emphasize enough the importantance of all the fantastic people that I was lucky to meet during these years: Dmitry Dylov, Nikolay Yampolsky, Michael Dorf, Alexander Pechen, Andrei and Lena Zhmoginov, Eugenia and Ilya Dodin, Tania Castro, Pablo Acerenza, Giacomo Bacci and Luca Scardovi among them.

Finally, it is my greatest honor to thank my family. My father, mother, sister and little brother have been that source of constant endorsement and motivation, which kept me focused, and without which I could not have achieved any of the serious goals. Most importantly, I would like to thank my wife Olga, who has been with me, side by side, throughout all the ups and downs of the last five years; her unique creativity and bright character are undoubtedly reflected in all sides of this work.

Contents

List of Figures

List of Tables

List of Abbreviations

BSC	binary symmetric channel	**i.i.d.**	independent identically distributed
BEC	binary erasure channel	**CLT**	central-limit theorem
GEC	Gilbert-Elliott channel	**MDP**	moderate deviation property
AWGN	additive white Gaussian noise	**ML**	maximum likelihood
BIAWGN	binary input AWGN	**ARQ**	automatic repeat request
SNR	signal-to-noise ratio	**LDPC**	low-density parity-check (code)
RCU	random-coding union (bound)	**VLF**	variable-length feedback (code)
DT	dependence testing (bound)	**VLFT**	VLF (code) with termination
PDF	probability density function	**FV**	fixed-to-variable (code)
CDF	cumulative distribution function		

Chapter 1

Introduction

1.1 The capacity

One of the brilliant achievements of Shannon's ground-breaking work [2] is creation of the abstract model of communication, converting many practical engineering questions into well-posed mathematical problems. Many of the methods developed for studying such problems have become known collectively as *information theory* (of channel coding). Shannon's model, as simple as it is, has withstood the test of time and critique. We now briefly describe it.

A communication problem consists of the following ingredients:

1. An apriori unknown message, which is modeled as a random variable equiprobably taking values in the set $\{1, \ldots, M\}$.

2. A channel, representing the abstraction of the noisy communication medium. The channel takes an input symbol in some alphabet $\mathcal{A}$, applies a random transformation ("adds intrinsic noise") and outputs a symbol in the alphabet $\mathcal{B}$. The channel can be used multiple times in which case the random transformation applied to each symbol in the sequence is the same[1] .

3. An encoder that maps messages into length n sequences of channel input symbols ("codewords"). The length n is known as the blocklength and the encoder is then a function $f : \{1, \ldots, M\} \to \mathcal{A}^n$.

4. A decoder that produces an estimate of the original message by observing an n-sequence of channel outputs. The decoder is a function $g : \mathcal{B}^n \to \{1, \ldots, M\}$. The pre-images $g^{-1}(j), j = 1, \ldots, M$ are known as the decoding sets.

An error happens if the decoder estimates the message incorrectly. Once the encoder and decoder are fixed, we can compute the probability of error by averaging with respect to the choice of the message and channel noise. The goal of the communication engineer is to find good encoder-decoder pairs ("codes") capable of communicating the message with

[1] Of course, different channel models are also considered, but here we restrict the presentation to *stationary memoryless* channels.

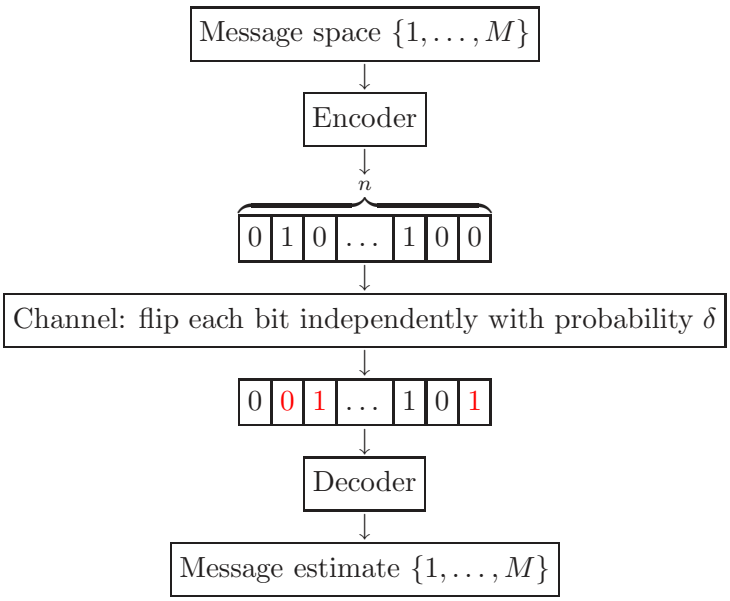


Figure 1.1: Block coding for the BSC, which acts by adding (mod 2) a binary noise with i.i.d. *Bernoulli*(δ) entries.

some required probability of error ϵ and the smallest possible blocklength n. We see that the most important parameters of the code are given by a tuple (n, M, ϵ) representing the blocklength, number of messages and the probability of error.

For the sake of illustration, we consider a particular example of the channel, the binary symmetric channel (BSC), which serves as a good model for many simple systems employing binary phase-shift keying with coherent hard-decision demodulators (wireless line of sight, or over the wire). The BSC has a binary $\{0, 1\}$ input, which is perturbed by flipping the bit with probability δ, known as the crossover probability, to produce a binary output. The schematic representation of Shannon's model of communication over the BSC is depicted on Fig. 1.1.

In this case, the goal is to select M binary n-strings and disjoint decoding sets ("balls") in the space $\{0, 1\}^n$ around them such that when the original string is transmitted the corresponding ball captures the perturbed output with probability of at least $1 - \epsilon$. Notice that to achieve a small probability of error ϵ, the encoder adds redundancy to the data: the original $\log_2 M$ data bits are mapped into a larger number of bits n. The ratio between $\frac{\log_2 M}{n}$ is known as the rate

$$R = \frac{\log_2 M}{n} \tag{1.1}$$

measured in bits per channel use. The term "rate" signifies that different channel uses typically correspond to different time instants, and therefore the blocklength n is proportional to the duration of communication.

A striking observation made by Shannon in [2] is that there exist sequences of (n, M_n, ϵ_n) codes with increasing blocklength n achieving a positive asymptotic rate

$$R = \lim_{n\to\infty} \frac{1}{n} \log_2 M_n > 0 \tag{1.2}$$

and vanishing probability of error

$$\epsilon_n \to 0\,. \tag{1.3}$$

However, not all rates R are achievable with vanishing probability of error: there is a maximal such rate, called the capacity C of the channel. For example, for the BSC with crossover probability δ the capacity is given (in bits per channel use) as

$$C(\delta) = 1 + \delta \log_2 \delta + (1-\delta)\log_2(1-\delta)\,. \tag{1.4}$$

The intuitive explanation of this phenomenon hinges on the fact that for large n perturbation of the codeword incurred by the channel is of a very restricted kind: each symbol in the codeword is perturbed independently and therefore different perturbations are very unlikely to "conspire" and produce a significant disturbance.

In order to state Shannon's result rigorously, let us fix the blocklength n and some probability of error $0 < \epsilon < 1$ and define the function

$$M^*(n,\epsilon) = \max\{M : \exists(n, M, \epsilon)\text{-code}\}\,, \tag{1.5}$$

which is the maximum number of messages that it is possible to transmit using blocklength n and such that the original message can be recovered with probability at least $1-\epsilon$. The function $M^*(n,\epsilon)$ is the non-asymptotic fundamental limit for a given communication channel. Going back to the BSC, $M^*(n,\epsilon)$ denotes the maximum number of "balls" that it is possible to pack into a space of binary n-strings $\{0,1\}^n$, where each "ball" is required to capture the probability $1-\epsilon$ when its "center" is being transmitted.

Shannon's result then states that

$$\lim_{\epsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} \log_2 M^*(n,\epsilon) = C\,, \tag{1.6}$$

where C is given by (1.4) for the BSC. In fact, Wolfowitz [3] showed that for any $0 < \epsilon < 1$ we have

$$\limsup_{n\to\infty} \frac{1}{n} \log_2 M^*(n,\epsilon) \le C\,, \tag{1.7}$$

a result known as a strong converse. Together (1.6) and (1.7) imply that for any fixed probability of error $0 < \epsilon < 1$ and $n \to \infty$ the fundamental limit satisfies

$$\log_2 M^*(n,\epsilon) = nC + o(n)\,. \tag{1.8}$$

The practical interpretation of (1.8): it is possible to send "reliably" nC data bits using n channel uses. This interpretation may then serve as a basis for system design and optimization.

1.2 Reliability function

The result (1.8) has one serious drawback: it does not suggest in any way how a fixed ϵ affects the value of the fundamental limit $\log_2 M^*(n,\epsilon)$. One frequently taken approach is to assume that the blocklength n is "large enough" to attain a situation where ϵ-dependent term $o(n)$ becomes much smaller (say, below 10%) than the leading term nC. Quite surprisingly,

however, to the best of our knowledge no systematic analysis has ever been made to estimate how large this "large enough" should be. Rigorously, we are interested in the smallest value of n such that $\frac{1}{n}\log_2 M^*(n,\epsilon) \geq 0.9C$, where ϵ is a required reliability level.

Another problem, not addressed by (1.8) is the following. Many modern applications require communicating the real-time data, such as voice, video streams or stock prices. The nature of such data puts a hard delay requirement on its delivery. For example, the physiology of human hearing (and bit rates of popular voice compressors) requires that the digitized speech be delivered in chunks no larger than 500-1000 bits in order to be perceived without noticeable (and annoying) delay. The goal of information theory is to answer what is the smallest n for which $\log_2 M^*(n,\epsilon) \geq 500$ (for some prescribed reliability level ϵ, of course). The only recipe suggested by (1.8) is to estimate $n \sim \frac{500}{C}$, which does not take into account ϵ (and is very inaccurate, as we will see).

Nevertheless, the question of the effect of probability of error ϵ on the fundamental limit $\log_2 M^*(n,\epsilon)$ has been classically addressed but in a different manner. Instead of studying the function $M^*(n,\epsilon)$ the idea is to study a related function:

$$\epsilon^*(n,R) = \inf\{\epsilon : \exists (n, 2^{nR}, \epsilon)\text{-code}\}\,, \tag{1.9}$$

which represents the smallest achievable probability of error among all codes mapping M messages to n channel inputs. Its asymptotic behavior for a fixed rate is determined by the function $E(R)$, a reliability function, defined as[2]

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 \epsilon^*(n,R) = E(R)\,. \tag{1.11}$$

Obviously by (1.7), for any $R > C$ we have $E(R) = 0$. So far, the value of $E(R)$ was established for most channels only for rates $R_c < R < C$, where R_c is called a critical rate of the channel. Notably, the question is open even for the BSC. Besides some special channels [4, 5], the landscape in the problem of reliability function has been set by Gallager [6] and Shannon, Gallager and Berlekamp [7] (see also [8] for some recent progress in the case of the BSC). For the BSC with crossover probability δ, the reliability function $E(R)$ is given by

$$E(R) = s\log_2\frac{s}{\delta} + (1-s)\log_2\frac{1-s}{1-\delta}\,, \tag{1.12}$$

where s is found as a solution to $C(s) = R$ and $C(\cdot)$ is given by (1.4), provided that $C\left(\frac{\sqrt{\delta}}{\sqrt{\delta}+\sqrt{1-\delta}}\right) < R < C(\delta)$; see [9, Section 5.6].

The meaning of (1.11) is that by restricting the rate to be strictly below capacity, $R < C$, it is possible to attain an exponentially decaying probability of error, with the optimal exponent given by $E(R)$. Although apriori fixing the rate (instead of ϵ) might seem artificial, it was quite natural in the early years of communication. For example, Bell Labs DS0/DS1 digital lines were operating at a fixed rate of 64 kbps, corresponding to

[2]This is equivalent to studying the limit

$$\lim_{n\to\infty}\frac{1}{n}\log_2 M^*(n, 2^{-nE}) \tag{1.10}$$

for a fixed $E > 0$.

a single phone line sampled at 8 kHz and 8-bit pulse-code modulated. With the advent of packet-switching networks, variable-rate compressors and higher demand for raw data throughput, however, the practice of allocating fixed rates is becoming increasingly rare. Another concern regarding the error-exponent approach is that practically it makes very little sense to believe in exponentially small estimates on probability of error in view of the crudeness of the original memoryless channel models.

Despite this, the reliability function gives the first guideline regarding the tradeoff between the probability of error, communication rate and blocklength. For example, to get an estimate on n required to achieve 90% of the capacity we can take

$$n \approx \frac{-\log_2 \epsilon}{E(0.9C)}\,, \tag{1.13}$$

which corresponds to approximating $\epsilon^*(n,R) \approx 2^{-nE(R)}$ and solving for n. Specifically, let us take the BSC with crossover probability $\delta = 0.11$ and capacity $C \approx 0.5$ bit. If we want to achieve $\epsilon = 10^{-3}$ and 90% of capacity, then the error-exponent approximation (1.13) yields

$$n \approx 4730\,. \tag{1.14}$$

1.3 Bounds

How do we know whether the approximation (1.14) is an accurate one?

All approaches discussed so far were asymptotic, either giving the limit of $\frac{1}{n}\log_2 M^*(n,\epsilon)$ at fixed ϵ as in (1.8), or of $\frac{1}{n}\log_2 \epsilon^*(n,R)$ at fixed rate (1.11). As exciting as these results are, in practice, however, we are interested in values of $M^*(n,\epsilon)$ or $\epsilon^*(n,R)$ for a finite n. Can this be computed exactly?

In principle, computation of the $M^*(n,\epsilon)$ can be performed according to the definition, since in the case of the BSC there are only finitely many codes for each blocklength and M. The caveat is that for rate R there are $\binom{2^n}{2^{nR}}$ different codes, and the direct computation becomes prohibitive already for very small values of n. In general, computation of $M^*(n,\epsilon)$ is an NP-hard problem [10]. So testing the accuracy of (1.14) directly is not possible.

If we cannot compute $M^*(n,\epsilon)$ exactly, maybe we can provide upper and lower bounds? After all, proving asymptotic results like (1.8) or (1.11) involves finding upper and lower bounds that match up asymptotically. Can we compute such bounds non-asymptotically?

This is indeed possible, and the bounds behind both (1.8) and (1.11) are computable [11–14]. The problem is that in proving asymptotic results one seeks the bounds that are general and easy to analyze asymptotically, such as Feinstein [15] or Gallager [6] lower (achievability) bounds, or Wolfowitz [3], and Shannon, Gallager and Berlekamp's sphere packing [7] upper (converse) bounds. In these cases, however, generality comes at the expense of poor non-asymptotic performance. In fact, more recent bounds, such as those developed by Csiszár and Körner [16, 17], are not as tight non-asymptotically as the cited classical bounds; see Section 2.2.1. Several authors have tried to modify the classical bounds in order to improve the non-asymptotic behavior [18, 19]. A notable exception from this picture is a case of the additive white Gaussian noise (AWGN) channel, for which Shannon has derived individual bounds [4] which are useful for both asymptotic analysis and numerical computation [20–23].

The BSC and the binary erasure channel (BEC) have also enjoyed a similar special attention in the literature [24, 25].

In this thesis we take a different approach. Instead of tweaking the classical bounds or proving specialized bounds for each and every channel, we start anew and derive novel bounds from first principles with non-asymptotic tightness in mind. This is the content of Chapter 2. The resulting general bounds provide a basis for finite blocklength analysis. Interestingly some of the novel results turn out to be both analytically tractable, e.g. prove the most general capacity formula [26], and at the same time specialize to the tightest known bounds non-asymptotically (e.g., the dependence testing (DT) bound for the binary erasure channel (BEC); see Section 3.3.1). In some cases our general bounds specialize to the best known non-asymptotic bounds which were previously derived using channel-specific methods (such as the sphere packing bound for the BSC which we derive as an application of the meta-converse; see Section 3.2.1).

Specializing the bounds to the BSC we can tightly sandwich the value of $\log_2 M^*(n,\epsilon)$ for the entire range of n. For example, for our running example of the BSC with $\delta = 0.11$ we get

$$190 \le \log_2 M^*(500, 10^{-3}) \le 193.3\,. \tag{1.15}$$

A better picture is obtained by considering Fig. 3.3 in Chapter 3, where the upper and lower bounds on $\frac{1}{n}\log_2 M^*(n,\epsilon)$ clearly illustrate the effect of convergence to capacity predicted by (1.8).

Returning to the question of the minimal blocklength needed to achieve 90% of the capacity, non-asymptotic bounds give us the following firm estimates:

$$2985 \le n \le 3106\,. \tag{1.16}$$

And we conclude therefore that the error-exponent approximation (1.14) is not accurate.

1.4 Normal approximation and beyond

How can we better predict the true value of $\log_2 M^*(n,\epsilon)$ without computing the bounds? In the case of the BSC, once the bounds are derived, a simple analysis requiring only Stirling's formula reveals that the upper and lower bounds match up to the first three terms and we obtain the following asymptotic expansion

$$\log_2 M^*(n,\epsilon) = nC(\delta) - \sqrt{nV(\delta)}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,, \tag{1.17}$$

where as usual,

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy\,, \tag{1.18}$$

and the coefficient V is referred to as *the channel dispersion* and for the BSC is given by

$$V(\delta) = \delta(1-\delta)\log_2^2 \frac{1-\delta}{\delta}\,. \tag{1.19}$$

Clearly (1.17) is a refinement of (1.8). Without the $\log n$ term, this expansion has been obtained by Weiss [27] and rediscovered recently in [28].

By dropping the $O(1)$ term in (1.17) we obtain the following *normal approximation* for the BSC:

$$\log_2 M^*(n,\epsilon) \approx nC(\delta) - \sqrt{nV(\delta)}Q^{-1}(\epsilon) + \frac{1}{2}\log n\,, \tag{1.20}$$

The quality of this approximation can be observed in Fig. 3.3. The surprising tightness of this approximation suggests that the asymptotic expansions at fixed ϵ, such as (1.17), might result in good approximations non-asymptotically. To the best of our knowledge, this approach is pioneered in this work. In particular, for the minimal blocklength needed to achieve 90% of the capacity of the BSC with $\delta = 0.11$ we obtain

$$n \approx 3150 \tag{1.21}$$

which compares much better to the true value sandwiched by (1.16) than an error-exponent approximation (1.14).

A natural question to ask now is: Does an expansion of the kind (1.17) hold true for other memoryless channels (with a different V and perhaps a different $\log n$ term)?

A positive answer was conjectured by Dobrushin [29] for a class of discrete symmetric channels , and later generalized by Strassen [1] to arbitrary discrete memoryless channels (DMCs) who showed that for $\epsilon < 1/2$ there exists a V such that for $n \to \infty$ we have

$$\log_2 M^*(n,\epsilon) = nC - \sqrt{nV}Q^{-1}(\epsilon) + O(\log n)\,. \tag{1.22}$$

Now assuming that the approximation obtained by dropping the $O(\log n)$ term is comparable in quality to a similar one obtained for the BSC, we can give a general answer to the question we started with: In order to achieve a fraction η of the capacity at probability of error ϵ one needs blocklength

$$n \gtrsim \left(\frac{Q^{-1}(\epsilon)}{1-\eta}\right)^2 \frac{V}{C^2}\,, \tag{1.23}$$

which requires knowing only two fundamental quantities associated with the channel: the capacity C and the channel dispersion V.

Motivated by the BSC example, obtaining the refined asymptotic expansions such as (1.17) occupies a bulk of Chapters 3 and 4. In particular, we elaborate on the $O(\log n)$ term for the BSC and BEC, amend Strassen's result in the case of $\epsilon > 1/2$ (his treatment of this regime contained an error) and provide some refined estimates on the $O(\log n)$ term; we analyze a certain ergodic channel with memory as well as a channel which is a non-ergodic mixture of the BSCs. We extend (1.22) to the AWGN channel and the parallel AWGN channel. In most cases we compare the normal approximation to the non-asymptotic bounds, each time obtaining an excellent match. Regarding Gaussian channels we also consider a special question of energy efficiency, asymptotically solved by Shannon [30] but virtually untouched non-asymptotically.

Having access to a tight approximation of the behavior of $\log_2 M^*(n,\epsilon)$ for finite n, we address applications to several engineering questions in Chapter 5 such as assessing the efficiency of known codes and effects of channel dynamics on the communication rate, where the analysis of the capacity term in (1.17) leads to drastically incorrect design decisions, compared to the analysis taking into account both the capacity and the dispersion terms.

Finally, in Chapter 6 we analyze the effect of feedback on memoryless channels. Shannon showed that the availability of feedback cannot increase the capacity of such channels [31]. However, we demonstrate that the non-asymptotic behavior changes dramatically in the presence of feedback. For example, instead of $n \approx 3100$ needed to achieve 90% of the capacity, as shown by (1.16), this value becomes 200 or even 20 depending on the model taken for the packet termination signaling. At the same time, we show that such savings are only possible when one considers average length: putting constraints on the excess delay nullifies the advantages of feedback even non-asymptotically (except, perhaps, for very short lengts).

Chapter 2

Bounds for general channels

The main tools required for non-asymptotic analysis of channel coding problems are introduced in this chapter. After setting the notation (Section 2.1) previous results are reviewed in Section 2.2. The problem of binary hypothesis testing, central to many of the methods in this work, is discussed in Section 2.3. Next, three main achievability bounds are derived in Sections 2.4, 2.5 and 2.6 for the average probability of error, maximal probability of error, and cost-constrained settings, respectively. Finally, Section 2.7 develops a highly general approach to proving impossibility results, a *meta-converse*, whose efficiency is demonstrated by showing that all of the relevant classical converse bounds are simple specializations of the meta-converse, and by obtaining some new results also. The material in this chapter has been presented in part in [32] and [33].

2.1 Definitions and notation

In this thesis a measurable space A, or *an alphabet*, is a set A equipped with a σ-algebra $\sigma\mathsf{A}$ of its subsets. For all spaces of finite cardinality we always assume that σ-algebra consists of all subsets. A measure is a non-negative σ-additive function $\sigma\mathsf{A} \to \mathbb{R}_+$. A transition probability kernel acting between two alphabets $T : \mathsf{A} \to \mathsf{B}$ assigns to each $x \in \mathsf{A}$ a measure $T(\cdot|x)$ on B, such that for any $E \in \sigma\mathsf{B}$ the function $T(E|x)$ is measurable with respect to $\sigma\mathsf{A}$. Every measurable function $f : \mathsf{A} \to \mathsf{B}$ can be identified with the transition probability kernel T_f as follows:

$$T_f(E|x) = 1\{f(x) \in E\}\,, \tag{2.1}$$

where $1\{\cdot\}$ is an indicator of the event. For this reason transition probability kernels can be understood as randomized functions (or maps). Similar to maps, transition probability kernels $T : \mathsf{A} \to \mathsf{B}$ and $S : \mathsf{B} \to \mathsf{W}$ can be composed to give a kernel $S \circ T : \mathsf{A} \to \mathsf{W}$ by

$$S \circ T(E|x) \stackrel{\triangle}{=} \int_{\mathsf{B}} S(E|w)T(dw|x)\,, \tag{2.2}$$

where integration is over the conditional measure $T(\cdot|x)$ on W.

A probability measure P is absolutely continuous with respect to Q, $P \ll Q$ in short, if $Q(E) = 0$ implies $P(E) = 0$. For a pair of such measures we denote by $\frac{dP}{dQ}$ a Radon-Nikodym derivative of P with respect to Q. The (Kullback-Leibler) divergence $D(P||Q)$ is

defined as

$$D(P||Q) \stackrel{\triangle}{=} \int_{\mathsf{A}} \frac{dP}{dQ} \log \frac{dP}{dQ} \cdot dQ \tag{2.3}$$

$$= \mathbb{E}_P \left[\log \frac{dP}{dQ} \right], \tag{2.4}$$

provided that $P \ll Q$, and we take $D(P||Q) = +\infty$ otherwise. Similarly we define *the divergence variance* as

$$V(P||Q) \stackrel{\triangle}{=} \int_{\mathsf{A}} \frac{dP}{dQ} \log^2 \frac{dP}{dQ} dQ - D^2(P||Q) \tag{2.5}$$

$$= \operatorname*{Var}_P \left[\log \frac{dP}{dQ} \right]. \tag{2.6}$$

The units of divergence (and other information measures) are specified by fixing a base of the logarithm in (2.3) and (2.5), which throughout this work can be chosen arbitrarily, as long as the exponent function, exp, is taken to the same base.

Definition 1 *A random transformation is given by a triplet* $(\mathsf{A}, \mathsf{B}, P_{Y|X})$ *of input and output alphabets* A *and* B*, and a transition probability kernel* $P_{Y|X} : \mathsf{A} \to \mathsf{B}$*. A channel is a sequence of random transformations* $(\mathsf{A}_n, \mathsf{B}_n, P_{Y^n|X^n})$*,* $n = 1, \ldots, \infty$*, where parameter* n *is the blocklength.*

This definition follows the approach taken in [26], so that in the applications we take A and B to be n-fold Cartesian products of some alphabets $\mathcal{A}$ and $\mathcal{B}$, and the transition kernels of the channel to be a sequence of conditional probabilities $\{P_{Y^n|X^n} : \mathcal{A}^n \to \mathcal{B}^n\}$. Thus, to focus ideas the elements of A and B (and the values of random variables X and Y) throughout subsequent sections can be viewed as vectors of fixed dimension equal to the blocklength.

For a transition probability kernel $T : \{1, \ldots, M\} \to \{1, \ldots, M\}$ we define its minimal diagonal element as

$$P_{min}(T) \stackrel{\triangle}{=} \min_{j=1,\ldots,M} T(j|j), \tag{2.7}$$

and its diagonal average as

$$P_{avg}(T) \stackrel{\triangle}{=} \frac{1}{M} \sum_{j=1}^{M} T(j|j). \tag{2.8}$$

Definition 2 *An* M*-code for the random transformation* $(\mathsf{A}, \mathsf{B}, P_{Y|X})$ *is defined by an (encoder) map* $f : \{1, \ldots, M\} \to \mathsf{A}$ *and a transition probability kernel (decoder)* $g : \mathsf{B} \to \{1, \ldots, M\}$*. The elements of the image of* f *are called codewords. For a code* (f, g) *we define its maximal probability of error*

$$\epsilon_{max}(f, g) \stackrel{\triangle}{=} 1 - P_{min}(g \circ P_{Y|X} \circ T_f) \tag{2.9}$$

$$= \max_{j=1,\ldots,M} \left(1 - P_{Y|X}(g^{-1}(j)|f(j))\right), \tag{2.10}$$

where T_f was defined in (2.1). *An M-code with $\epsilon_{max} \le \epsilon$ is said to be an* (M, ϵ)-code (maximal probability of error). *Similarly, for a code (f, g) we define its average probability of error*

$$\epsilon_{avg}(f,g) \stackrel{\triangle}{=} 1 - P_{avg}(g \circ P_{Y|X} \circ T_f) \tag{2.11}$$

$$= \frac{1}{M}\sum_{j=1}^{M}\left(1 - P_{Y|X}(g^{-1}(j)|f(j))\right) . \tag{2.12}$$

An M-code with $\epsilon_{avg} \le \epsilon$ is said to be an (M, ϵ)-code (average probability of error).

Although not a main focus of our attention, we also define a randomized M-code[1] to be a pair (f, g) of transition probability kernels $f : \{1, \dots, M\} \to \mathsf{A}$ and $g : \mathsf{B} \to \{1, \dots, M\}$. The rest of the quantities are defined analogously to Definition 2. Note that for a randomized code, the concept of the codeword is meaningless.

Definition 3 *Given a pair of random transformations* $(\mathsf{A}_1, \mathsf{B}_1, P_{Y_1|X_1})$ *and* $(\mathsf{A}_2, \mathsf{B}_2, P_{Y_2|X_2})$ *we define their product as a random transformation* $(\mathsf{A}_1 \times \mathsf{A}_2, \mathsf{B}_1 \times \mathsf{B}_2, P_{Y^2|X^2})$ *with*

$$P_{Y^2|X^2}(\cdot|x_1, x_2) = P_{Y_1|X_1}(\cdot|x_1) \times P_{Y_2|X_2}(\cdot|x_2) , \tag{2.13}$$

where the right-hand side is a product of probability measures.

As example of using Definition 3 we define the binary symmetric channel (BSC) with crossover probability $0 \le \delta \le 1$ as follows. For $n = 1$ we take input and output alphabets $\mathcal{A} = \mathcal{B} = \{0, 1\}$ and the transition probability kernel:

$$P_{Y|X}(b|a) = \begin{cases} 1-\delta, & a = b , \\ \delta, & a \ne b . \end{cases} \tag{2.14}$$

For $n > 1$ we iterate n times the product construction of Definition 3 applied to random transformation (2.14). The sequence of random transformations obtained in this way is known as the BSC. A random transformation for blocklength n is denoted $BSC(n, \delta)$ for convenience. Explicitly, $BSC(n, \delta)$ has input and output alphabets $\mathsf{A} = \mathsf{B} = \mathcal{A}^n = \mathcal{B}^n = \{0, 1\}^n$ – a space of binary strings of length n – and the kernel $P_{Y^n|X^n}$ acts by adding a binary noise Z^n independent of the input X^n:

$$Y^n = X^n + Z^n , \tag{2.15}$$

where Z^n has independent, identically distributed (i.i.d.) components with Bernoulli distribution: $\mathbb{P}[Z_i = 1] = 1 - \mathbb{P}[Z_i = 0] = \delta$.

Channels whose constituent random transformations are obtained as n-fold products of a single base random transformation are called memoryless channels. If the base random transformation acts between finite input and output alphabet then the resulting sequence is a discrete memoryless channel (DMC). Such sequences of channels parametrized by the blocklength n arise frequently in practical models of communication.

[1] More precisely, an M-code with a randomized encoder.

Definition 4 *An (M, ϵ) code for the n-th random transformation of the channel is called an (n, M, ϵ) code. We define the fundamental non-asymptotic limit for a channels as*

$$M^*(n,\epsilon) = \max\{M : \exists (n,M,\epsilon)\text{-code (maximal probability of error)}\,, \quad (2.16)$$

$$M^*_{avg}(n,\epsilon) = \max\{M : \exists (n,M,\epsilon)\text{-code (average probability of error)}\,. \quad (2.17)$$

The non-asymptotic fundamental limit $M^*(n,\epsilon)$ gives rise to a number of asymptotic quantities associated to a given channel.

Definition 5 *The ϵ-capacity C_ϵ (measured in information units per channel use) of a channel is defined as*

$$C_\epsilon \stackrel{\triangle}{=} \liminf_{n\to\infty} \frac{1}{n} \log M^*(n,\epsilon)\,. \quad (2.18)$$

Definition 6 *The capacity C (measured in information units per channel use) of a channel is defined as*

$$C \stackrel{\triangle}{=} \lim_{\epsilon\to 0} C_\epsilon\,. \quad (2.19)$$

According to this definition the capacity is the maximal rate of communication which still admits an asymptotically vanishing probability of error.

Definition 7 *The channel dispersion V (measured in squared information units per channel use) of a channel with capacity C is equal to*[2]

$$V = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \left(\frac{nC - \log M^*(n,\epsilon)}{Q^{-1}(\epsilon)} \right)^2 \quad (2.20)$$

$$= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \frac{(nC - \log M^*(n,\epsilon))^2}{2\ln\frac{1}{\epsilon}}\,. \quad (2.21)$$

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

$$\log M^*(n,\epsilon) = nC - \sqrt{nV} Q^{-1}(\epsilon) + O(\log n) \quad (2.22)$$

So Definition 7 extracts the coefficient V in this approximation, similar to how Definition 6 extracts the coefficient C. Expansions of the type (2.22) have been studied in [1, 3, 27–29] with main contributions by Dobrushin [29] and Strassen [1]; see Section 3.1.2 for more details.

The utility of defining the asymptotic quantities C and V for non-asymptotic analysis is established in this thesis by showing that for many different channels the following approximation:

$$\log M^*(n,\epsilon) \approx nC - \sqrt{nV} Q^{-1}(\epsilon) \quad (2.23)$$

gives an excellent estimate for the true value of $\log M^*(n,\epsilon)$ in the regime of practically interesting n and ϵ. In particular, the minimal blocklength required to achieve a given fraction η of capacity with a given error probability ϵ can be estimated as

$$n \gtrsim \left(\frac{Q^{-1}(\epsilon)}{1-\eta} \right)^2 \frac{V}{C^2}\,. \quad (2.24)$$

[2] This form of the definition has been proposed by S. Verdú.

Similar to how ϵ-capacity is a refinement of the capacity, the following definition is a refinement of the definition of dispersion:

Definition 8 *For a sequence of channels with ϵ-capacity C_ϵ, the ϵ-dispersion is defined for $\epsilon \in (0,1) - \{\frac{1}{2}\}$ as*

$$V_\epsilon = \limsup_{n\to\infty} \frac{1}{n} \left(\frac{nC_\epsilon - \log M^*(n,\epsilon)}{Q^{-1}(\epsilon)} \right)^2 . \tag{2.25}$$

Note that for $\epsilon < \frac{1}{2}$, approximating $\frac{1}{n}\log M^*(n,\epsilon)$ by C_ϵ is optimistic and smaller dispersion is preferable, while for $\epsilon > \frac{1}{2}$, it is pessimistic and larger dispersion is more favorable. Since $Q^{-1}(\frac{1}{2}) = 0$, it is immaterial how to define $V_{\frac{1}{2}}$ as far as the normal approximation (2.23) is concerned.

For a joint distribution P_{XY} on $\mathsf{A} \times \mathsf{B}$ we are interested in the information density[3]

$$\begin{aligned} i(x,y) &= \log \frac{dP_{XY}}{d(P_X \times P_Y)}(x,y) && (2.26)\\ &= \log \frac{dP_{Y|X=x}}{dP_Y}(y)\,. && (2.27) \end{aligned}$$

More formally, we assume that for some measure μ on B we have $P_{Y|X=x} \ll \mu$ for all $x \in \mathsf{A}$ and $P_Y \ll \mu$; then define

$$f(x,y) \stackrel{\triangle}{=} \frac{dP_{Y|X=x}}{d\mu}(y)\,, \quad g(y) \stackrel{\triangle}{=} \frac{P_Y}{d\mu}(y)\,. \tag{2.28}$$

The information density can then be defined as[4]

$$i(x,y) = \begin{cases} -\infty, & f(x,y) = 0\,, \\ +\infty, & g(y) = 0\,, \\ \log \frac{f(x,y)}{g(y)}, & f(x,y) \neq 0, g(y) \neq 0\,. \end{cases} \tag{2.29}$$

In this thesis we denote by P_X, Q, $P_{Y|X=x}$, etc. distributions of a single variable, whereas $\mathbb{P}$ is reserved for probability measure on the underlying probability spaces.

2.2 Previous work

2.2.1 Achievability results

Two main classical non-asymptotic achievability bounds are due to Feinstein [15] and Shannon [34]. We present the generalizations to the settings with input constraints due to Thomasian [35] (see also [36] and [37, (2.34)]).

[3]We take (2.27) as the definition of $i(x,y)$ in this thesis. The reason for this is that the quantity (2.26) is only defined $(P_X \times P_Y)$-almost surely. Consequently, whenever $P_X(x) = 0$, it is meaningless to talk about the distribution of $i(x,Y)$, which is inconvenient for channels with continuous alphabets.

[4]Notice that it is irrelevant how to define $i(x,y)$ for the case $f(x,y) = g(y) = 0$, since we are interested in defining $i(x,\cdot)$ only on the union of supports of $P_{Y|X=x}$ and P_Y.

Suppose that all codewords are required to belong to some set $\mathsf{F} \subset \mathsf{A}$. For example, there might be a cost $\mathsf{c}(x)$ associated with using a particular input vector x, in which case the set F might be chosen as

$$\mathsf{F} = \{x : \mathsf{c}(x) \leq P\}\,. \tag{2.30}$$

The achievability bounds are then given as follows:

Theorem 1 (Feinstein) *For any distribution P_X, and any $\gamma > 0$, there exists an (M, ϵ) code (maximal probability of error) with codewords in the set $\mathsf{F} \subset \mathcal{X}$ satisfying*

$$M \geq \gamma \left(\epsilon - \mathbb{P}\left[i(X;Y) \leq \log \gamma\right] - P_X[\mathsf{F}^c]\right)\,, \tag{2.31}$$

or equivalently,

$$\epsilon \leq \mathbb{P}\left[i(X;Y) \leq \log \gamma\right] + \frac{\gamma}{M} + P_X[\mathsf{F}^c]\,. \tag{2.32}$$

Theorem 2 (Shannon) *For any distribution P_X, and any $\gamma > 0$, there exists an (M, ϵ) code (average probability of error) with codewords in the set F such that*

$$\epsilon \leq \mathbb{P}\left[i(X;Y) \leq \log \gamma\right] + \frac{\gamma}{M-1} + P_X[\mathsf{F}^c]\,. \tag{2.33}$$

Apart from the difference between M and $M-1$, Feinstein's bound implies Shannon's bound. Note that unconstrained versions are obtained by taking $\mathsf{F} = \mathsf{A}$ in Theorems 1 and 2. Another general coding theorem result is the one due to Gallager [6].

Theorem 3 (Gallager, no cost) *For any P_X and $\lambda \in [0, 1]$, there exists an (M, ϵ) code (average probability of error) such that*

$$\epsilon \leq M^{\lambda}\, \mathbb{E}\left[\left(\mathbb{E}\left[\exp \frac{i(X, \bar{Y})}{1+\lambda} \,\middle|\, \bar{Y}\right]\right)^{1+\lambda}\right] \tag{2.34}$$

where the pair $(X, \bar{Y})$ are distributed as

$$P_{X\bar{Y}}(a, b) = P_X(a) \sum_{a' \in \mathcal{A}} P_{Y|X}(b|a') P_X(a')\,. \tag{2.35}$$

For a memoryless channel (2.34) turns, after optimization over λ, into

$$\epsilon \leq \exp\{-nE_r(R)\}\,, \tag{2.36}$$

where $R = \frac{\log M}{n}$ is a coding rate and $E_r(R)$ is Gallager's random coding exponent.

Theorem 3 admits generalization to a case with cost-constraints $\mathsf{c}(x)$, see [9]:

Theorem 4 (Gallager, with cost) *Suppose P_X is such that*

$$\sum_{x \in \mathcal{A}} \mathsf{c}(x) P_X(x) \leq P\,, \tag{2.37}$$

and consider some $\delta \in [0, P]$ such that $\mu(\delta) > 0$ with $\mu(\delta)$ defined as

$$\mu(\delta) \stackrel{\triangle}{=} P_X[P - \delta \leq \mathsf{c}(X) \leq P]\,. \tag{2.38}$$

Then for any $\lambda \in [0,1]$ and $r \geq 0$ there exists an (M,ϵ) code (average probability of error) with codewords satisfying $\mathsf{c}(x) \leq P$ and such that

$$\epsilon \leq M^{\lambda} \left(\frac{\exp(r\delta)}{\mu(\delta)}\right)^{1+\lambda} \mathbb{E}\left[\left(\mathbb{E}\left[\exp\left\{\frac{i(X,\bar{Y})}{1+\lambda} + (\mathsf{c}(X)-P)r\right\}\middle|\bar{Y}\right]\right)^{1+\lambda}\right] \tag{2.39}$$

where the pair $(X,\bar{Y})$ are distributed as

$$P_{X\bar{Y}}(a,b) = P_X(a)\sum_{a'\in\mathcal{A}} P_{Y|X}(b|a')P_X(a')\,. \tag{2.40}$$

There are also bounds specially developed for particular discrete and Gaussian channels, which are going to be discussed in Sections 3.1 and 4.1.

We do not cite here any of the "joint typicality" or type-splitting achievability results. This is because those bounds, contrary to our goal, are derived with implicit assumption of tending $n \to \infty$ and thus do not yield tight bounds.

Let us motivate this omission quantitatively. For example, in [17] Csisz'ar and Körner give an achievability bound (Theorem 1 there), which after optimization reduces to

$$\epsilon \leq \exp\{-n(E_r(R+\delta_n) - \delta_n')\}\,, \tag{2.41}$$

where R and $E_r(R)$ are the same as in (2.36). We already can see that this bound can not be better than Gallager's. Numerically, even if we neglect δ_n' we can see that (2.41) compared to Gallager's bound incurs the loss of rate by at least

$$\delta_n = (|\mathcal{A}|^2 + |\mathcal{A}|)\frac{\log n + 1}{n} + \frac{1}{n}\,. \tag{2.42}$$

For the BSC with $n = 1000$, we find that $\delta_n \approx 0.06$. Now look at Fig. 3.1, where different bounds are compared. If one subtracts 0.06 from Gallager's bound it becomes obvious that (2.41) is very far away from the contenders. The presence of δ_n' deteriorates situation even more, e.g. for $n = 1000$ we have $\exp\{n\delta_n'\} = 10^{24}$.

For these reasons in this thesis we do not consider the aforementioned achievability bounds and also corresponding type-based converse bounds (e.g., Haroutounian's [38]).

2.2.2 Converse results

Among the relevant converses we cite Fano's inequality:

Theorem 5 *Every (M,ϵ)-code (average probability of error) for a random transformation $P_{Y|X}$ satisfies*

$$\log M \leq \frac{1}{1-\epsilon}\sup_X I(X;Y) + \frac{1}{1-\epsilon}h(\epsilon) \tag{2.43}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy function.

For the maximal probability of error, Fano's inequality is significantly improved by the bound due to Wolfowitz [3].

Theorem 6 (Wolfowitz) *Every (M,ϵ)-code (maximal probability of error) must satisfy*

$$M \le \inf_{\beta>0} \beta \left(\inf_{x\in \mathsf{A}} P_{Y|X=x}\left[i(x;Y) < \log \beta \right] - \epsilon \right)^{-1} \tag{2.44}$$

provided that the right-hand side is not less than 1.

As shown in [39, Theorem 7.8.1], this bound leads to the strong converse theorem for the discrete memoryless channel (DMC)

$$\log M^*(n,\epsilon) \le nC + o(n) \qquad \forall \epsilon \in (0,1)\,. \tag{2.45}$$

Moreover, the bound (2.45) also holds in the presence of noiseless feedback [39].

The following corollary to Theorem 6 gives another converse bound which also leads to (2.45):

Theorem 7 ([9, Theorem 5.8.5]) *For an arbitrary discrete memoryless channel of capacity C and any $(n, \exp\{nR\}, \epsilon)$ code with rate $R > C$, we have*

$$\epsilon \ge 1 - \frac{4A}{n(R-C)^2} - \exp\left\{ -\frac{n(R-C)}{2} \right\}, \tag{2.46}$$

where $A > 0$ is constant independent of n or R.

We notice that Theorem 7 is in general too coarse for analyzing the finite blocklength behavior of fundamental limits.

The dual of the Shannon-Feinstein bounds in Theorems 1 and 2 (in the unconstrained setting) is given in [26].

Theorem 8 (Verdú-Han) *Every (M,ϵ)-code (average error probability) satisfies*

$$\epsilon \ge \sup_{\beta>0} \left\{ \mathbb{P}\left[i(X;Y) \le \log \beta \right] - \frac{\beta}{M} \right\}, \tag{2.47}$$

where $P_{XY} = P_X P_{Y|X}$ and P_X is the distribution on A induced by the code.

A looser bound of [40] is obtained from (2.47) by replacing the optimization over the β with a fixed choice $\beta = \frac{M}{2}$. Although Theorem 8 leads to the most general formula for the channel capacity [26], obtaining computable bounds on fundamental limits via (2.47) is challenging due to the need of solving an optimization over the set of all n-dimensional input distributions. Similar problems appear in computing a generally tighter bound given in [41]:

Theorem 9 (Poor-Verdú) *Every (M,ϵ)-code (average error probability) satisfies*

$$\epsilon \ge \sup_{\beta>0} \left(1 - \frac{\beta}{M} \right) \mathbb{P}\left[i(X;Y) \le \log \beta \right], \tag{2.48}$$

where $P_{XY} = P_X P_{Y|X}$ and P_X is the distribution on A induced by the code.

A generalization of Theorem 6 was proposed in [42] by changing the reference probability measure in the definition of $i(X;Y)$ from P_Y to an arbitrary Q_Y; see also [43, 44]:

Theorem 10 *Every (M,ϵ)-code (average error probability) satisfies*

$$\epsilon \geq \sup_{\beta>0}\left\{\inf_{P_X}\sup_{Q_Y}\mathbb{P}\left[\log\frac{dP_{XY}}{d(P_X\times Q_Y)}(X,Y)\leq\log\beta\right]-\frac{\beta}{M}\right\}. \tag{2.49}$$

Finally, for the asymptotic analysis of error exponents, the following bound [7] is crucial (see also [45] for the same bound explored in a different notation).

Theorem 11 (Shannon-Gallager-Berlekamp) *Let $P_{Y|X}:\mathcal{A}\mapsto\mathcal{B}$ be a DMC. Then any (n,M,ϵ) code (average probability of error) satisfies*

$$\epsilon \geq \exp\{-n(E_{sp}(R-o_1)+o_2)\}, \tag{2.50}$$

where

$$\begin{aligned}
R &= \frac{\log M}{n}, && (2.51)\\
E_{sp}(R) &= \sup_{\rho\geq 0}[E_0(\rho)-\rho R], && (2.52)\\
E_0(\rho) &= \max_{P_X} E_0(\rho,P_X), && (2.53)\\
E_0(\rho,P_X) &= -\log\sum_{y\in\mathcal{B}}\left[\sum_{x\in\mathcal{A}}P_X(x)P_{Y|X}(y|x)^{1/(1+\rho)}\right]^{1+\rho} && (2.54)\\
&= -\log\left(\mathbb{E}\left[\mathbb{E}\left[\exp\frac{i(\bar{X};Y)}{1+\rho}\middle|Y\right]\right]^{1+\rho}\right), && (2.55)\\
o_1 &= \frac{\log 4}{n}+\frac{|\mathcal{A}|\log n}{n}, && (2.56)\\
o_2 &= \sqrt{\frac{8}{n}}\log\frac{e}{\sqrt{P_{min}}}+\frac{\log 8}{n}, && (2.57)\\
P_{min} &= \min\{P_{Y|X}(y|x):P_{Y|X}(y|x)>0\}, && (2.58)
\end{aligned}$$

where the maximization in (2.53) *is over all probability distributions on $\mathcal{A}$; and in* (2.55), *$\bar{X}$ and Y are independent:*

$$P_{\bar{X}Y}(a,b)=P_X(a)\left(\sum_{x\in\mathcal{A}}P_{Y|X}(b|x)P_X(x)\right). \tag{2.59}$$

Although Theorem 11 proved to be key for finding the reliability function at high rates, its utility for the finite blocklength regime is questionable, mainly due to coarse estimates o_1 and o_2. For these reasons, [18] and [19] have recently tightened those estimates and also extended the bound to continuous-output channels.

2.3 Binary hypothesis testing

Many of our results and methods depend on evaluating the optimal performance of the binary hypothesis test. Consider a random variable W on W which can take one of the two distributions P and Q. A randomized test between those two distributions is a random transformation (a transition probability kernel) $P_{Z|W} : \mathsf{W} \to \{0,1\}$, where 0 indicates that the test chooses Q. The optimal performance among all such transformations is denoted by[5]

$$\beta_\alpha(P,Q) = \inf_{\substack{P_{Z|W}: \\ \sum_{w\in\mathsf{W}} P_{Z|W}(1|w)P(w) \ge \alpha}} \left[\sum_{w\in\mathsf{W}} P_{Z|W}(1|w)Q(w)\right]. \tag{2.60}$$

Thus, $\beta_\alpha(P,Q)$ gives the minimum probability of error under hypothesis Q if the probability of success under hypothesis P is at least α.

Note that $\alpha \mapsto \beta_\alpha$ is a non-decreasing, convex function of $\alpha \in [0,1]$. Indeed, for any test $P_{Z|Y}$ we define

$$\alpha = \sum_{w\in\mathsf{W}} P_{Z|W}(1|y)P(w)\,, \tag{2.61}$$

$$\beta = \sum_{w\in\mathsf{W}} P_{Z|W}(1|y)Q(w)\,. \tag{2.62}$$

Then the totality of points (α,β) for all $P_{Z|W}$ form a convex subset of $[0,1]^2$. Since β_α is a lower boundary of this set, it must be convex.

The infimum in (2.60) is guaranteed to be achieved by an optimum randomized test according to the following lemma due to Neyman and Pearson (e.g., see [46]).

Lemma 12 (Neyman-Pearson) *Consider a space W and probability measures P and Q. Then for any $\alpha \in [0,1]$ there exist $\gamma > 0$ and $\tau \in [0,1)$ such that*

$$\beta_\alpha(P,Q) = Q[Z^*_\alpha = 1]\,, \tag{2.63}$$

and where[6] *the conditional probability $P_{Z^*|W}$ is defined via*

$$Z^*_\alpha(W) = 1\left\{\frac{dP}{dQ} > \gamma\right\} + Z_\tau 1\left\{\frac{dP}{dQ} = \gamma\right\}, \tag{2.64}$$

where $Z_\tau \in \{0,1\}$ equals 1 with probability τ independent of W. The constants γ and τ are uniquely determined by solving the equation

$$P[Z^*_\alpha = 1] = \alpha\,. \tag{2.65}$$

Moreover, any other test Z satisfying $P[Z=1] \ge \alpha$ either differs from Z^*_α only on the set $\left\{\frac{dP}{dQ} = \gamma\right\}$ or is strictly larger with respect to Q: $Q[Z=1] > \beta_\alpha(P,Q)$.

[5] Here and below we write summations over alphabets, whenever it does not cause confusion. However, all of the general results in this chapter hold for non-discrete measures and uncountable alphabets.

[6] In the case in which P is not absolutely continuous with respect to Q, we can define $\frac{dP}{dQ}$ to be equal to $+\infty$ on the singular set and hence to be automatically included in every optimal test.

In other words, β_α is a piecewise linear function, joining the points

$$\begin{cases} \beta_\alpha = Q\left[\dfrac{dP}{dQ} \geq \gamma\right], \\ \alpha = P\left[\dfrac{dP}{dQ} \geq \gamma\right] \end{cases} \tag{2.66}$$

iterated over all $\gamma > 0$.

The following bounds are easy to show ([47]):

$$\beta_\alpha(P,Q) \geq \frac{1}{\gamma}\left(\alpha - P\left[\frac{dP}{dQ} \geq \gamma\right]\right) \tag{2.67}$$

$$\beta_\alpha(P,Q) \leq \frac{1}{\gamma_0} P\left[\frac{dP}{dQ} \geq \gamma_0\right] \tag{2.68}$$

$$\leq \frac{1}{\gamma_0}, \tag{2.69}$$

where $\gamma > 0$ is arbitrary and γ_0 satisfies

$$P\left[\frac{dP}{dQ} \geq \gamma_0\right] \geq \alpha. \tag{2.70}$$

For completeness we give the proofs. For an arbitrary test $P_{Z|W}$ we have:

$$Q[Z=1] \geq Q\left[\{Z=1\} \cap \left\{\frac{dP}{dQ} < \gamma\right\}\right] \tag{2.71}$$

$$\geq \frac{1}{\gamma} P\left[\{Z=1\} \cap \left\{\frac{dP}{dQ} < \gamma\right\}\right] \tag{2.72}$$

$$\geq \frac{1}{\gamma}\left(P[Z=1] - P\left[\frac{dP}{dQ} \geq \gamma\right]\right) \tag{2.73}$$

$$\geq \frac{1}{\gamma}\left(\alpha - P\left[\frac{dP}{dQ} \geq \gamma\right]\right), \tag{2.74}$$

from which (2.67) follows. To show (2.68), notice that if we denote

$$\alpha_0 \stackrel{\triangle}{=} P\left[\frac{dP}{dQ} \geq \gamma_0\right] \geq \alpha, \tag{2.75}$$

then we have

$$\beta_\alpha(P,Q) \leq \beta_{\alpha_0}(P,Q) \tag{2.76}$$

$$= Q\left[\frac{dP}{dQ} \geq \gamma_0\right] \tag{2.77}$$

$$\leq \frac{1}{\gamma_0} P\left[\frac{dP}{dQ} \geq \gamma_0\right], \tag{2.78}$$

where (2.76) follows from monotonicity of β_α, (2.77) is a consequence of Neyman-Pearson lemma, and (2.78) follows by a standard change of measure argument, see also [48]. This completes the proof of (2.68).

In general, the function $\alpha \to \beta_\alpha(P,Q)$ provides a lot of information about the relation between measures P and Q. For example $\beta_\alpha(P,Q) = \alpha$ if and only if $P = Q$; any expectation $\mathbb{E}_Q\left[f\left(\frac{dP}{dQ}\right)\right]$ can be computed via the formula:

$$\int f\left(\frac{dP}{dQ}\right) dQ = \int_0^1 \beta'(\alpha) f\left(\frac{1}{\beta'(\alpha)}\right) d\alpha\,, \tag{2.79}$$

where $\beta'(\alpha) = \frac{d\beta_\alpha}{d\alpha}$ exists almost everywhere by the Lebesgue theorem; see also [49, Theorem 11]. In particular (2.79) shows that every f-divergence [50] between P and Q can be obtained from the knowledge of β_α.

Below, the binary hypothesis testing of interest is $\mathsf{W} = \mathsf{B}$, $P = P_{Y|X=x}$ and $Q = Q_Y$, an auxiliary unconditional distribution. [7] In that case, for brevity and with a slight abuse of notation we will denote

$$\beta_\alpha(x, Q_Y) = \beta_\alpha(P_{Y|X=x}, Q_Y)\,. \tag{2.80}$$

Bounds (2.67) and (2.69) imply that β_α behaves approximately as the exponent of (the negative of) the α-th quantile of $\log\frac{dP}{dQ}$ under P. In this thesis we will mostly deal with distributions that are n-fold products of a fixed distribution. In this case $\log\frac{dP}{dQ}$ is a sum of i.i.d. random variables and the quantile behavior is governed by the central-limit theorem (CLT), or, more precisely, by the Berry-Esseen Theorem, e.g. [51, Theorem 2, Chapter XVI.5]:

Theorem 13 (Berry-Esseen) *Let the X_k, $k = 1, \ldots, n$ be independent with*

$$\mu_k = \mathbb{E}\left[X_k\right],\ \sigma_k^2 = \mathrm{Var}[X_k],\ \textit{and}\ t_k = \mathbb{E}\left[|X_k - \mu_k|^3\right]. \tag{2.81}$$

Denote $V = \sum_1^n \sigma_k^2$ and $T = \sum_1^n t_k$. Then

$$\left|\mathbb{P}\left[\frac{\sum_1^n (X_k - \mu_k)}{\sqrt{V}} \le \lambda\right] - Q(-\lambda)\right| \le 6\frac{T}{V^{3/2}}\,, \tag{2.82}$$

where $Q(x)$ is defined in (1.18).

Note that for i.i.d. X_k it is known that the factor of 6 in the right hand side can be replaced by 1 or less; see [52]. In this thesis, the exact value of the constant does not affect the results and so we take the conservative value of 6 even in the i.i.d. case.

Regarding the asymptotic behavior of the β_α, the Berry-Esseen inequality implies the following result, proved in Appendix A:

Lemma 14 *Let $\mathcal{A}$ be a measurable space with measures $\{P_i\}$ and $\{Q_i\}$, with $P_i \ll Q_i$ defined on it for $i = 1, \ldots, n$. Define two measures on $\mathcal{A}^n$: $P = \prod_{i=1}^n P_i$ and $Q = \prod_{i=1}^n Q_i$,*

[7] As we show later, it is sometimes advantageous to allow Q_Y that cannot be generated by any input distribution.

and

$$D_n = \frac{1}{n}\sum_{i=1}^{n} D(P_i||Q_i)\,, \tag{2.83}$$

$$V_n = \frac{1}{n}\sum_{i=1}^{n} V(P_i||Q_i) = \frac{1}{n}\sum_{i=1}^{n}\int\left(\log\frac{dP_i}{dQ_i}\right)^2 dP_i - D(P_i||Q_i)^2\,, \tag{2.84}$$

$$T_n = \frac{1}{n}\sum_{i=1}^{n}\int\left|\log\frac{dP_i}{dQ_i} - D(P_i||Q_i)\right|^3 dP_i\,, \tag{2.85}$$

$$B_n = 6\frac{T_n}{V_n^{3/2}}\,. \tag{2.86}$$

Assume that all quantities are finite and $V_n > 0$. Then, for any $\Delta > 0$

$$\log\beta_\alpha(P,Q) \geq -nD_n - \sqrt{nV_n}Q^{-1}\left(\alpha - \frac{B_n+\Delta}{\sqrt{n}}\right) - \frac{1}{2}\log n + \log\Delta\,, \tag{2.87}$$

$$\log\beta_\alpha(P,Q) \leq -nD_n - \sqrt{nV_n}Q^{-1}\left(\alpha + \frac{B_n}{\sqrt{n}}\right) - \frac{1}{2}\log n + \log\left(\frac{2\log 2}{\sqrt{2\pi V_n}} + 4B_n\right)\,. \tag{2.88}$$

Each bound holds provided that the argument of Q^{-1} lies in $(0,1)$.

In particular, when $P_i = P$ and $Q_i = Q$, $i = 1,\ldots,n$, $V(P||Q) > 0$ and the third moment of $\log\frac{dP}{dQ}$ is finite, we have

$$\log\beta_\alpha\left(P^n,Q^n\right) = -nD(P||Q) - \sqrt{nV(P||Q)}Q^{-1}(\alpha) - \frac{1}{2}\log n + O(1)\,. \tag{2.89}$$

If $V(P||Q) = 0$ then we trivially have

$$\log\beta_\alpha\left(P^n,Q^n\right) = -nD(P||Q) + \log\alpha\,. \tag{2.90}$$

The lower bound (2.87) holds only for n sufficiently large, while sometimes we want a firm bound, valid for all n, such as provided by the following result:

Lemma 15 *In the notation of Lemma 14, we have*

$$\log\beta_\alpha(P,Q) \geq -nD_n - \sqrt{\frac{2nV_n}{\alpha}} + \log\frac{\alpha}{2}\,. \tag{2.91}$$

A proof of this result is also found in Appendix A.

Each per-codeword cost constraint can be defined by specifying a subset $\mathsf{F} \subset \mathsf{A}$ of permissible inputs. For an arbitrary $\mathsf{F} \subset \mathsf{A}$, we define a related measure of performance for the composite hypothesis test between Q_Y and the collection $\{P_{Y|X=x}\}_{x\in\mathsf{F}}$:

$$\kappa_\tau(\mathsf{F},Q_Y) = \inf_{\substack{P_{Z|Y}:\\ \inf_{x\in\mathsf{F}} P_{Z|X}(1|x)\geq\tau}} \sum_{y\in\mathsf{B}} Q_Y(y)P_{Z|Y}(1|y)\,. \tag{2.92}$$

As long as Q_Y is the output distribution induced by an input distribution Q_X, the quantity (2.92) satisfies the bound

$$\tau Q_X[\mathsf{F}] \leq \kappa_\tau(\mathsf{F}, Q_Y) \leq \tau\,. \tag{2.93}$$

The right-hand side bound is achieved by choosing the test Z that is equal to 1 with probability τ regardless of Y. To see the left-hand bound, note that for any $P_{Z|Y}$ that satisfies the condition in (2.92), we have

$$\begin{aligned}
& \sum_{y\in\mathsf{B}} Q_Y(y) P_{Z|Y}(1|y) \\
= & \sum_{x\in\mathsf{A}}\sum_{y\in\mathsf{B}} Q_X(x) P_{Y|X}(y|x) P_{Z|Y}(1|y) && (2.94)\\
\geq & \sum_{x\in\mathsf{F}} Q_X(x) \sum_{y\in\mathsf{B}} P_{Y|X}(y|x) P_{Z|Y}(1|y) && (2.95)\\
\geq & \sum_{x\in\mathsf{F}} Q_X(x) \left\{ \inf_{x\in\mathsf{F}} \sum_{y\in\mathsf{B}} P_{Y|X}(y|x) P_{Z|Y}(1|y) \right\} && (2.96)\\
\geq & \tau Q_X[\mathsf{F}]\,. && (2.97)
\end{aligned}$$

A remark on notation: Typically we will take A and B as n-fold Cartesian products of alphabets $\mathcal{A}$ and $\mathcal{B}$. To emphasize dependence on n we will write $\beta_\alpha^n(x, Q_Y)$ and $\kappa_\tau^n(\mathsf{F}, Q_Y)$. Since Q_Y and F will usually be fixed we will simply write κ_τ^n. Also, in many cases $\beta_\alpha^n(x, Q_Y)$ will be the same for all $x \in \mathsf{F}$. In these cases we will write β_α^n.

2.4 Achievability: average probability of error

All of the upper-bounds on the average probability of error considered in this thesis invoke the original idea of Shannon [2], namely, generating the codebook randomly. Specifically, the exact value of the probability of error is given by the following expression[8] :

Theorem 16 *Denote by $\epsilon(c_1, \ldots, c_M)$ the error probability achieved by the maximum likelihood decoder with codebook $(c_1, \ldots, c_M)$. Let $X_1, \ldots, X_M$ be independent with marginal distribution P_X. Then,*

$$\mathbb{E}\left[\epsilon(X_1, \ldots, X_M)\right] = 1 - \sum_{\ell=0}^{M-1} \binom{M-1}{\ell} \frac{1}{\ell+1} \mathbb{E}\left[W^\ell Z^{M-1-\ell}\right] \tag{2.98}$$

where

$$\begin{aligned}
W &= \mathbb{P}\left[i(\bar{X};Y) = i(X;Y) \,\middle|\, X, Y\right] && (2.99)\\
Z &= \mathbb{P}\left[i(\bar{X};Y) < i(X;Y) \,\middle|\, X, Y\right] && (2.100)
\end{aligned}$$

with

$$P_{XY\bar{X}}(a,b,c) = P_X(a) P_{Y|X}(b|a) P_X(c)\,. \tag{2.101}$$

[8] This result was obtained by S. Verdú.

Proof: Since the M messages are equiprobable, upon receipt of the channel output y, the maximum likelihood decoder chooses with equal probability among the members of the set

$$\arg \max_{i=1,\ldots,M} i(c_i;y)\,.$$

Therefore, if the codebook is $(c_1,\ldots,c_M)$, and $m=1$ is transmitted, the maximum likelihood decoder will choose $\hat{m}=1$ with probability $\frac{1}{1+\ell}$ if

$$\sum_{j=2}^{M} 1\{i(c_j;y)=i(c_1;y)\} \;=\; \ell \tag{2.102}$$

$$\sum_{j=2}^{M} 1\{i(c_j;y)>i(c_1;y)\} \;=\; 0\,, \tag{2.103}$$

for $\ell=0,\ldots M-1$. If (2.103) is not satisfied an error will surely occur. Since the codewords are chosen independently with identical distributions, given that the codeword assigned to message 1 is c_1 and given that the channel output is $y\in\mathsf{B}$, the joint distribution of the remaining codewords is $P_X\times\cdots\times P_X$. Consequently, the conditional probability of correct decision is

$$\sum_{\ell=0}^{M-1}\binom{M-1}{\ell}\frac{1}{\ell+1}\left(\mathbb{P}\left[i(\bar{X};y)=i(c_1;y)\right]\right)^{\ell}\left(\mathbb{P}\left[i(\bar{X};y)<i(c_1;y)\right]\right)^{M-1-\ell} \tag{2.104}$$

where $\bar{X}$ has the same distribution as X, but is independent of any other random variable arising in this analysis. Averaging (2.104) with respect to (c_1,y) jointly distributed as $P_XP_{Y|X}$ we obtain the summation in (2.98). Had we conditioned on a message other than $m=1$ we would have obtained the same result. Therefore, the error probability averaged over messages and codebook is given by (2.98). ■

Naturally, Theorem 16 leads to an achievability upper bound since there must exist an $(M,\mathbb{E}\left[\epsilon(X_1,\ldots,X_M)\right])$ (average error probability) code. Although in some special cases it is possible to compute the value appearing in the right-hand side of (2.98), in general the required computational complexity is too high and we need to consider simpler upper bounds. Such upper bounds are the focus of the subsequent sections.

2.4.1 Random coding union (RCU) bound

Our first bound is the following:

Theorem 17 (RCU) *For an arbitrary P_X there exists an (M,ϵ) code (average probability of error) such that*

$$\epsilon\le\mathbb{E}\left[\min\left\{1,\;(M-1)\mathbb{P}\left[i(\bar{X},Y)\ge i(X,Y)\,\middle|\,X,Y\right]\right\}\right]\,, \tag{2.105}$$

where $P_{XY\bar{X}}(a,b,c)=P_X(a)P_{Y|X}(b|a)P_X(c)$.

Proof: Let us generate our codewords $X_1, \dots X_M$ as independent random variables with common distribution P_X. Let us denote by λ_j the (random) probability of error conditioned on transmitting the j-th codeword:

$$\lambda_j \stackrel{\triangle}{=} P[\text{error} \mid X = X_j]\,. \tag{2.106}$$

Then, the average probability of error is given by

$$\epsilon = \frac{1}{M}\sum_{j=1}^{M} \lambda_j\,, \tag{2.107}$$

and by symmetry we find that

$$\mathbb{E}\left[\epsilon\right] = \mathbb{E}\left[\lambda_1\right]\,. \tag{2.108}$$

We need now to average λ_1 over the random choice of codebook $X_1, \dots X_M$. The maximum likelihood decoder will decode to the codeword with maximal $i(X_j, Y)$ given the received value Y. Thus, we can upper-bound the probability of error as

$$\mathbb{E}\left[\lambda_1\right] \le \mathbb{P}\left[\bigcup_{j=2}^{M}\{i(X_j,Y) \ge i(X_1,Y)\}\right]\,. \tag{2.109}$$

(this is an inequality because some of the cases $i(X_j, Y) = i(X_1, Y)$ might have been resolved to X_1, whereas we have assumed the worst).

In (2.109) we have $P_{X_1 Y X_{2\dots}} = P_{X_1} P_{Y|X_1} P_{X_2} P_{X_3} \cdots$. Notice that we can first condition on X_1 and Y, and then take expectation over them:

$$\epsilon \le \mathbb{E}\left[\mathbb{P}\left[\bigcup_{j=2}^{M}\{i(X_j,Y) \ge i(X_1,Y)\}\,\middle|\, X_1, Y\right]\right]\,. \tag{2.110}$$

In this way, the internal conditional probability is actually a probability of $M-1$ independent events. It is natural to apply the union bound then:

$$\epsilon \le \mathbb{E}\left[\min\left\{1,\ \sum_{j=2}^{M}\mathbb{P}\left[i(X_j,Y) \ge i(X_1,Y)\,|\,X_1,Y\right]\right\}\right]\,. \tag{2.111}$$

Here we used $\min(x, 1)$ to exclude the values larger than 1. Note that all the probabilities in the $\sum_{j=2}^{M}$ are equal, and thus we can simply write

$$\epsilon \le \mathbb{E}\left[\min\left\{1,\ (M-1)\mathbb{P}\left[i(\bar{X},Y) \ge i(X,Y)\,\middle|\, X,Y\right]\right\}\right]\,. \tag{2.112}$$

■

Essentially, the only ingredients of the proof are the random-coding and the union bound (hence the name: RCU). The bound can be viewed as a generalization of Shannon's AWGN bound [4]. For symmetric channels the computational complexity of the RCU bound is rather low, e.g. $O(n^2)$ for the BSC, and hence the bound is computable even for rather

large blocklengths; see Sections 3.2 and 3.3 below. However, in general the bound (2.105) has complexity $O\left(n^{2(|\mathcal{A}|-1)|\mathcal{B}|}\right)$ and thus frequently the direct application of Theorem 17 is not possible. Below we give alternative bounds that are easier to compute and still tight enough for many purposes.

There is also a way to simplify the bound (2.105) different from the path that we adopt below. Namely, we can first use a Chernoff-type upper-bound:

$$\mathbb{P}[i(\bar{X},Y) \geq i(X,Y)\,|\,X=x,Y=y] \leq \sum_{\bar{x}} P_X(\bar{x}) \left(\frac{P_{Y|X}(y|\bar{x})}{P_{Y|X}(y|x)}\right)^{\lambda} \tag{2.113}$$

(for memoryless channels it is known that this upper-bound is exponentially tight for the optimal choice of λ). This reduces the bound to

$$\epsilon \leq \mathbb{E}\left[\min\left\{1, (M-1)\sum_{\bar{x}} P_X(\bar{x}) \left(\frac{P_{Y|X}(Y|\bar{x})}{P_{Y|X}(Y|X)}\right)^{\lambda}\right\}\right]. \tag{2.114}$$

The complexity here is just $O(n^{(|\mathcal{A}|-1)|\mathcal{B}|})$.

Additionally we can use the simple inequality $\min\{x,1\} \leq x^{\rho}$ valid for $\rho \in [0,1]$. After plugging this into (2.114) we get

$$\epsilon \leq (M-1)^{\rho}\mathbb{E}\left[\left\{\sum_{\bar{x}} P_X(\bar{x}) \left(\frac{P_{Y|X}(Y|\bar{x})}{P_{Y|X}(Y|X)}\right)^{\lambda}\right\}^{\rho}\right]. \tag{2.115}$$

As shown by Gallager, the optimum choice of λ is $1/(1+\rho)$ in which case the bound simply becomes Theorem 3:

$$\epsilon \leq (M-1)^{\rho}\sum_{y}\left\{P_X(x)P_{Y|X}(y|x)^{1/(1+\rho)}\right\}^{1+\rho}. \tag{2.116}$$

2.4.2 Dependence testing (DT) bound

Theorem 18 (DT) *For any distribution P_X on A, there exists a code with M codewords and average probability of error not exceeding*

$$\begin{aligned}\epsilon &\leq \mathbb{P}\left[i(X,Y) \leq \log\frac{M-1}{2}\right] + \frac{M-1}{2}\mathbb{P}\left[i(X,\bar{Y}) > \log\frac{M-1}{2}\right] &(2.117)\\ &= \mathbb{E}\left[\exp\left\{-\left|i(X,Y) - \log\frac{M-1}{2}\right|^{+}\right\}\right] &(2.118)\end{aligned}$$

where $P_{XY\bar{Y}}(a,b,c) = P_X(a)P_{Y|X}(b|a)P_Y(c)$.

The name "dependence testing (DT) bound" will be explained shortly, see Section 2.4.3. Before proving the theorem, we formulate and prove a useful lemma.

Lemma 19 *Consider a distribution P_X on A, a distribution $P_Y(y) = \sum_{x\in\mathsf{A}} P_{Y|X}(y|x)P_X(x)$ on B and a measurable function $\gamma : \mathsf{A} \mapsto [0,\infty]$. Then there exists an (M,ϵ) code (average probability of error) satisfying*

$$\epsilon \leq P[i(X,Y) \leq \log\gamma(X)] + \frac{M-1}{2}P[i(X,\bar{Y}) > \log\gamma(X)], \tag{2.119}$$

where $P_{XY\bar{Y}}(a,b,c) = P_X(a)P_{Y|X}(b|a)P_Y(c)$.

Remark: We demonstrate how a very slightly weaker bound can be obtained from Theorem 17. Note that in (2.105) when $i(X,Y)$ is small the term $(M-1)\mathbb{P}[\cdots]$ is probably larger than 1 and thus it is natural to upper-bound $\min\{x,1\}$ by 1. On the other hand, when $i(X,Y)$ is very large it is probably smaller than 1 and thus it is reasonable to upper-bound $\min\{x,1\}$ by x:

$$\begin{aligned}\epsilon &\leq \mathbb{E}\left[1\{i(X,Y)\leq\gamma\}+1\{i(X,Y)>\gamma\}(M-1)\mathbb{P}[i(\bar{X},Y)\geq i(X,Y)\,|\,XY]\right] && (2.120)\\ &\leq \mathbb{P}[i(X,Y)\leq\gamma]+(M-1)\mathbb{P}[i(\bar{X},Y)>\gamma]\,. && (2.121)\end{aligned}$$

This bound is only 1-bit weaker than (2.119).

Proof of Lemma 19: The idea of the proof is to average the probability of error over random codebooks generated using the distribution P_X; the decoder runs M likelihood ratio binary hypothesis tests in parallel, the j^{th} of which is between the true distribution $P_{Y|X=c_j}$ and "average noise" P_Y.

Let $\{Z_x\}_{x\in\mathsf{A}}$ be a collection of deterministic functions over B defined as

$$Z_x(y)=1\{i(x,y)>\log\gamma(x)\}\,. \tag{2.122}$$

First we describe the operation of the decoder given the codebook $\{c_i\}_{i=1}^M$; the decoder computes the values $Z_{c_j}(y)$ for the received channel output y and returns the first index j for which $Z=1$ (or 0 if all of them returned 0). In this way, the average probability of error is given as

$$\epsilon(c_1,\ldots c_M)=\frac{1}{M}\sum_{i=1}^M\lambda_i\,, \tag{2.123}$$

where

$$\lambda_j=\mathbb{P}\left[\{Z_{c_j}(Y)=0\}\bigcup_{i<j}\{Z_{c_i}(Y)=1\}\,\middle|\,X=c_j\right]\,, \tag{2.124}$$

or, using the union bound and the definition of $Z_x(y)$,

$$\lambda_j\leq\mathbb{P}[i(c_j,Y)\leq\log\gamma(c_j)\,|\,X=c_j]+\sum_{i<j}\mathbb{P}[i(c_i,Y)>\log\gamma(c_i)\,|\,X=c_j]\,. \tag{2.125}$$

We will now average each expression (2.125) over codebooks $\{c_i\}$ that are generated as (pairwise) independent random variables with distribution P_X. Then, we obtain

$$\mathbb{E}\,[\lambda_j]\leq\mathbb{P}[i(X,Y)\leq\log\gamma(X)]+(j-1)\mathbb{P}[i(X,\bar{Y})>\log\gamma(X)]\,. \tag{2.126}$$

Then from (2.123) we find that the ensemble average of ϵ satisfies

$$\mathbb{E}\,[\epsilon(c_1,\ldots c_M)]\leq\mathbb{P}[i(X,Y)\leq\log\gamma(X)]+\frac{M-1}{2}\mathbb{P}[i(X,\bar{Y})>\log\gamma(X)]\,. \tag{2.127}$$

■

Proof of Theorem 18: Notice that by taking an expectation conditioned on X in (2.119) we obtain

$$P_{Y|X=x}[i(x,Y)\leq\log\gamma(x)]+\frac{M-1}{2}P_Y[i(x,Y)>\gamma(x)]\,, \tag{2.128}$$

which is a weighted sum of two types of errors. This thus corresponds to the average error probability in a Bayesian hypothesis testing problem for which the optimal solution is the likelihood ratio test with threshold $\gamma(x) = (M-1)/2$.

We need only to show that (2.117) and (2.118) are equal[9] . To this end consider an expression

$$P\left[\frac{dP}{dQ} \leq \gamma\right] + \gamma Q\left[\frac{dP}{dQ} > \gamma\right]. \tag{2.129}$$

Note that when evaluating the second term we can drop any region N such that $Q(N) = 0$ or $P(N) = 0$ (since $\gamma \geq 0$). Thus, we can replace Q and P with different measures $\tilde{Q}$ and $\tilde{P}$ such that $\tilde{Q} \sim \tilde{P}$ and formula $\frac{d\tilde{Q}}{d\tilde{P}} = \left(\frac{d\tilde{P}}{d\tilde{Q}}\right)^{-1}$ holds. Thus, the second term can be rewritten as

$$\gamma Q\left[\frac{dP}{dQ} > \gamma\right] = \int \gamma \left(\frac{dP}{dQ}\right)^{-1} 1_{\{\frac{dP}{dQ} > \gamma\}}\, dP\,. \tag{2.130}$$

Summing with the first term in (2.129) we obtain

$$\begin{aligned} P\left[\frac{dP}{dQ} \leq \gamma\right] + \gamma Q\left[\frac{dP}{dQ} > \gamma\right] &= \int \left[1_{\{\frac{dP}{dQ} \leq \gamma\}} + \gamma \left(\frac{dP}{dQ}\right)^{-1} 1_{\{\frac{dP}{dQ} > \gamma\}}\right] dP \quad (2.131) \\ &= \int \min\left\{\gamma \left(\frac{dP}{dQ}\right)^{-1}, 1\right\} dP\,. \quad (2.132) \end{aligned}$$

Now substituting $e^{-i(X,Y)}$ for $\left(\frac{dP}{dQ}\right)^{-1}$ and $\frac{M-1}{2}$ for γ we have (2.118). ■

2.4.3 Some properties of the DT bound

The right side of the DT bound (2.117) is equal to $\frac{M+1}{2}$ times the Bayesian minimal error probability of a binary hypothesis test of dependence:

$$\begin{aligned} \mathrm{H}_1: &\quad P_{XY} &&\text{with probability } \tfrac{2}{M+1} \\ \mathrm{H}_0: &\quad P_X P_Y &&\text{with probability } \tfrac{M-1}{M+1} \end{aligned}$$

Therefore, Theorem 18 demonstrates that the *dependence testing* (DT) problem is related to the problem of constructing channel codes.

One of the properties of the DT bound that makes it particularly useful in applications is that unlike the existing bounds (2.31), (2.33), and (2.34), the bound in Theorem 18 requires no optimization of auxiliary constants. Moreover, for the case of no input constraints ($\mathsf{F} = \mathsf{A}$), Theorem 2 follows from Lemma 19 by taking $\gamma(x) = \beta$ and upper-bounding $\frac{M-1}{2}$ by M. Similarly, a recent bound in [53] can also be seen as a weakened version of Lemma 19, and is therefore provably weaker than Theorem 18 (originally published in [33]). Therefore, the DT bound is strictly stronger than Shannon's bound (2.33) and [53].

Regarding the asymptotic analysis, it can be easily seen from (2.118) that Theorem 18 can be used to prove the achievability part of the most general known channel capacity formula [26].

[9]The compact form of the DT bound given by (2.118) was proposed by S. Verdú.

For the analysis of the second term in the representation (2.117) frequently the following result comes in handy.

Lemma 20 *Let $Z_1, Z_2, \ldots, Z_n$ be independent random variables, $\sigma^2 = \sum_{j=1}^n \operatorname{Var} Z_j$ be non-zero and $T = \sum_{j=1}^n \mathbb{E}\left[|Z_j - \mathbb{E}\, Z_j|^3\right] < \infty$; then for any A*

$$\mathbb{E}\left[\exp\left\{-\sum_{j=1}^n Z_j\right\} 1_{\{\sum_{j=1}^n Z_j > A\}}\right] \leq 2\left(\frac{\log 2}{\sqrt{2\pi}} + \frac{12T}{\sigma^2}\right)\frac{1}{\sigma}\exp\{-A\}\,. \tag{2.133}$$

Proof: By Theorem 13 we have for any x and δ

$$\mathbb{P}\left[x \leq \sum_{j=1}^n (Z_j - \mathbb{E}\, Z_j) < x + \delta\right] \tag{2.134}$$

$$\leq \int_{x/\sigma}^{(x+\delta)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt + \frac{12T}{\sigma^3} \tag{2.135}$$

$$\leq \left(\frac{\delta}{\sqrt{2\pi}} + \frac{12T}{\sigma^2}\right)\frac{1}{\sigma}\,. \tag{2.136}$$

On the other hand,

$$\mathbb{E}\left[\exp\left\{-\sum_{j=1}^n Z_j\right\} 1_{\{\sum_{j=1}^n Z_j > A\}}\right] \tag{2.137}$$

$$\leq \sum_{l=0}^{\infty} \exp\{-A - l\delta\}\, \mathbb{P}\left[A + l\delta \leq \sum_{j=1}^n Z_j < A + (l+1)\delta\right]\,. \tag{2.138}$$

Using (2.136) and $\delta = \log 2$ we get (2.133) since

$$\sum_{l=0}^{\infty} 2^{-l} = 2\,. \tag{2.139}$$

■

Notice that

$$\lim_{M\to\infty} \left| i(X,Y) - \log\frac{M-1}{2}\right|^+ = 0 \tag{2.140}$$

and the convergence is monotone in M. Therefore, from (2.118) we see that as M ranges from 1 to ∞ the right-hand side of the DT bound in (2.118) grows monotonically from 0 to 1 (and is, thus, the cumulative distribution function, CDF, of some random variable). This suggests to take a different look on the expression (2.118) by defining a particular f-divergence [50] as follows:

$$\mathcal{D}_\gamma(P||Q) = \int \left[\frac{dP}{dQ} - \gamma\right]^+ dQ\,. \tag{2.141}$$

Then the DT bound (2.118) can be restated as:

$$1-\epsilon \geq \mathcal{D}_{\frac{M-1}{2}}\left(P_{XY}||P_XP_Y\right)\,. \tag{2.142}$$

Since processing does not increase f-divergence, the left-hand side of the bound (2.142) can be further lower-bounded by applying a suitable mapping of the space $\mathsf{A}\times\mathsf{B}$ into a simpler space. For example, in the case of the BSC ($\mathsf{A}=\mathsf{B}=\{0,1\}^n$) a convenient map is $\mathsf{A}\times\mathsf{B}\to\mathbb{Z}_+$ given by $(x,y)\to|x-y|$, where $|\cdot|$ is the Hamming weight.[10]

Some of the interesting properties of f-divergence $\mathcal{D}(\cdot||\cdot)$ are listed in the following:

Theorem 21 *Assuming $P\sim Q$ we have*

$$\mathcal{D}_\gamma(P||Q) \;=\; \int_0^1 P\left[\log\frac{dP}{dQ}>\log\frac{\gamma}{u}\right]du\,. \tag{2.143}$$

The function $\gamma\to\mathcal{D}_\gamma$ is non-increasing from 1 to 0 on $\mathbb{R}_+$, and

$$\lim_{\delta\to 0+}\frac{\mathcal{D}_{\gamma+\delta}-\mathcal{D}_\gamma}{\delta} \;=\; -Q\left[\frac{dP}{dQ}>\gamma\right]\,, \tag{2.144}$$

$$\lim_{\delta\to 0-}\frac{\mathcal{D}_{\gamma+\delta}-\mathcal{D}_\gamma}{\delta} \;=\; -Q\left[\frac{dP}{dQ}\geq\gamma\right]\,, \tag{2.145}$$

and hence, Lebesgue-almost everywhere

$$\frac{d\mathcal{D}_\gamma}{d\gamma}=-Q\left[\frac{dP}{dQ}\geq\gamma\right]\,. \tag{2.146}$$

The function $\gamma\to\mathcal{D}_\gamma$ contains the same information about P and Q as the function $\alpha\to\beta_\alpha(P,Q)$, according to:

$$\mathcal{D}_\gamma(P||Q)=\alpha(\gamma)-\gamma\beta_{\alpha(\gamma)}(P,Q)\,, \tag{2.147}$$

where $\alpha(\gamma)=P\left[\frac{dP}{dQ}\geq\gamma\right]$. Consequently, any other f-divergence can be expressed in terms of $\mathcal{D}_\gamma$. For example, for the divergence $D(\cdot||\cdot)$ we have: If $D(P||Q)<\infty$ then[11]

$$D(P||Q)=\log e-\int_0^\infty\log\gamma\,\frac{dD_\gamma}{d\gamma}d\gamma\,. \tag{2.148}$$

Finally, the following holds:

$$\mathcal{D}_\gamma(P||Q) \;\leq\; 1-\exp\left\{-\mathbb{E}_P\left[\left[\log\frac{dP}{dQ}-\log\gamma\right]^+\right]\right\} \tag{2.149}$$

$$\;\leq\; \frac{1}{\log e}\mathbb{E}_P\left[\left[\log\frac{dP}{dQ}-\log\gamma\right]^+\right]\,. \tag{2.150}$$

[10] This is an instance of a general idea of channel simplification: $\mathsf{A}\times\mathsf{B}$ should be mapped to the orbit space under the action of the automorphism group of the channel on $\mathsf{A}\times\mathsf{B}$, see Section 6.5.

[11] Notice that $-\frac{d\mathcal{D}_\gamma}{d\gamma}$ is a density of a probability measure on $\mathbb{R}_+$.

Remark: Upper-bound (2.149) is especially useful when $\frac{1}{n}\log\frac{dP_n}{dQ_n}$ is geometrically concentrating around $\frac{1}{n}D(P_n||Q_n)$ as $n\to\infty$.

Proof: From the definition (2.141), we have the following equivalent expressions:

$$\begin{aligned}\mathcal{D}_\gamma(P||Q) &= P\left[\frac{dP}{dQ}\geq\gamma\right]-\gamma Q\left[\frac{dP}{dQ}\geq\gamma\right] &(2.151)\\ &= 1-\int\left[\min\left\{\frac{dP}{dQ},\gamma\right\}\right]dQ &(2.152)\\ &= 1-\int\left[\min\left\{\gamma\left(\frac{dP}{dQ}\right)^{-1},1\right\}\right]dP &(2.153)\end{aligned}$$

whereas (2.143) follows from (2.153) by applying

$$\mathbb{E}\left[\min\{X,1\}\right]=\int_0^1\mathbb{P}[X\geq u]du\,, \tag{2.154}$$

which is valid for any non-negative X. The non-increasing nature of $\mathcal{D}_\gamma$ follows trivially from (2.141).

To get one-sided derivatives (2.144) and (2.145), we need to simply use representation (2.152) and notice that

$$\begin{aligned}\frac{1}{\delta}\left[\min\{x,\gamma+\delta\}-\min\{x,\gamma\}\right] &\to 1\{x>\gamma\}\,,\quad \text{as } \delta\searrow 0\,, &(2.155)\\ \frac{1}{\delta}\left[\min\{x,\gamma+\delta\}-\min\{x,\gamma\}\right] &\to 1\{x\geq\gamma\}\,,\quad \text{as } \delta\nearrow 0\,, &(2.156)\end{aligned}$$

and both converge uniformly in $x\in\mathbb{R}$.

Representation (2.147) follows from (2.151) and the Neyman-Pearson lemma (2.66).

To prove (2.148) notice that by definition of divergence we have

$$D(P||Q)=\int\frac{dP}{dQ}\log\frac{dP}{dQ}dQ=\int_{\mathbb{R}_+}x\log x\,d\tilde{Q}\,, \tag{2.157}$$

where Q' is the distribution of $\frac{dP}{dQ}$ under Q:

$$Q'\stackrel{\triangle}{=}Q\circ\left(\frac{dP}{dQ}\right)^{-1}. \tag{2.158}$$

Continuing from (2.157) we have:

$$\begin{aligned}-D(P||Q) &= \int_{\mathbb{R}_+}-x\log x\,dQ' &(2.159)\\ &= x\log xQ\left[\frac{dP}{dQ}>x\right]\Big|_0^\infty-\int_0^\infty(\log x+\log e)Q\left[\frac{dP}{dQ}>x\right]dx &(2.160)\\ &= x\log xQ\left[\frac{dP}{dQ}>x\right]\Big|_0^\infty+\int_0^\infty(\log x+\log e)\frac{d\mathcal{D}_x}{dx}dx &(2.161)\\ &= x\log xQ\left[\frac{dP}{dQ}>x\right]\Big|_0^\infty-\log e+\int_0^\infty\log x\cdot\frac{d\mathcal{D}_x}{dx}dx &(2.162)\\ &= -\log e+\int_0^\infty\log x\,d\mathcal{D}_x &(2.163)\end{aligned}$$

where (2.160) is integration by parts, (2.161) follows by applying (2.146), (2.162) holds since $\int_0^\infty d\mathcal{D}_x = -1$, and (2.163) follows because

$$\lim_{x\to\infty} Q\left[\frac{dP}{dQ} > x\right] x\log x = 0\,. \tag{2.164}$$

To prove (2.164) notice that

$$\begin{aligned} Q\left[\frac{dP}{dQ} > x\right]\cdot x\log x &= \mathbb{E}_Q\left[x\log x\cdot 1_{\{\frac{dP}{dQ}>x\}}\right] &(2.165)\\ &\le \mathbb{E}_Q\left[\frac{dP}{dQ}\log\frac{dP}{dQ}\cdot 1_{\{\frac{dP}{dQ}>x\}}\right], &(2.166)\end{aligned}$$

and since $D(P||Q) < \infty$, (2.166) tends to zero as $x\to\infty$ by the dominated convergence.

To show (2.149), recall the following inequality due to Donsker and Varadhan [54]:

$$D(P||Q) \ge \mathbb{E}_P\left[f\left(\frac{dP}{dQ}\right)\right] - \log\mathbb{E}_Q\left[\exp\left\{f\left(\frac{dP}{dQ}\right)\right\}\right], \tag{2.167}$$

where f is an arbitrary function. Applying this with $f(y) = \min\{\log y, \log\gamma\}$ we get by (2.152)

$$D(P||Q) \ge \mathbb{E}_P\left[\min\left\{\log\frac{dP}{dQ}, \log\gamma\right\}\right] - \log(1-\mathcal{D}_\gamma(P||Q))\,, \tag{2.168}$$

which after a simple algebra leads to (2.149) ■

2.5 Achievability: maximal probability of error

The details of the proof of Theorem 18 reveal that we could have generated the random codebook with only pairwise independent codewords. Thus, for some channels (e.g., discrete channels with additive noise) we can generate the codebook by imposing a distribution on the generating matrix of a linear code. Then Theorem 18 implies the existence of a linear code with average probability of error upper-bounded by (2.118). But the maximal and average probability of error coincide for a linear code decoded with a maximum likelihood decoder, and hence for additive-noise discrete channels the bound in Theorem 18 also holds in the sense of maximal probability of error (see Appendix C for further details on this approach). The following bound on maximal error probability holds in general.

Theorem 22 *For any input distribution P_X and measurable $\gamma : \mathsf{A}\to[0,\infty]$, there exists a code with M codewords such that the j-th codeword's probability of error satisfies*

$$\epsilon_j \le \mathbb{P}[i(X,Y)\le\log\gamma(X)] + (j-1)\sup_x \mathbb{P}[i(x,Y)>\log\gamma(x)]\,, \tag{2.169}$$

where the first probability is with respect to P_{XY} and the second is with respect to P_Y. In particular, the maximal probability of error satisfies

$$\epsilon \le \mathbb{P}[i(X,Y)\le\log\gamma(X)] + (M-1)\sup_x \mathbb{P}[i(x,Y)>\log\gamma(x)]\,. \tag{2.170}$$

Proof: First, we specify the operation of the decoder given the codebook $\{c_1, \ldots c_M\}$. The decoder simply computes $i(c_j, y)$ for the received channel output y and selects the first codeword c_j for which $i(c_j, y) > \log \gamma(c_j)$.

Now, let us show that we can indeed choose M codewords so that their respective probabilities of decoding error ϵ_j's satisfy (2.169). Suppose that the first codeword is equal to some $x \in \mathsf{A}$. Then the conditional probability of error under the specified decoding rule is equal to

$$\epsilon_1(x) = \mathbb{P}[i(x, Y) \leq \log \gamma(x) | X = x] \,. \tag{2.171}$$

Let us choose codeword x at random with distribution P_X. Then, the average of $\epsilon_1(x)$ is

$$\mathbb{E}\left[\epsilon_1(X)\right] = \mathbb{P}[i(X, Y) \leq \log \gamma(X)] \,. \tag{2.172}$$

Thus, there must exist at least one choice of x such that $\epsilon(x)$ is less than the right-hand side of (2.172). Call this choice c_1.

Now assume that n codewords $\{c_j\}_{j=1}^n$ have been chosen and we are to show that the $n+1$-st one can also be chosen so that (2.169) is satisfied. Denote

$$D = \bigcup_{j=1}^{n} \{y : \, i(c_j, y) > \log \gamma(c_j)\} \;\subseteq \mathsf{B} \,. \tag{2.173}$$

Suppose that the $n+1$-st codeword is equal to x. Then the conditional probability of error is

$$\epsilon_{n+1}(c_1, \ldots, c_n, x) = 1 - \mathbb{P}[\{i(x, Y) > \log \gamma(x)\} \setminus D | X = x] \,. \tag{2.174}$$

If we generate the $n+1$-st codeword randomly with probability distribution P_X then the average of ϵ_{n+1} is

$$\mathbb{E}\left[\epsilon_{n+1}(c_1, \ldots, c_n, X)\right] = \mathbb{P}[\{i(X, Y) \leq \log \gamma(X)\} \cup D] \leq \mathbb{P}[i(X, Y) \leq \log \gamma(X)] + P_Y(D) \,. \tag{2.175}$$

From (2.173) and the union bound we obtain

$$P_Y(D) \leq n \sup_{x \in \mathsf{A}} P_Y[i(x, Y) > \gamma(x)] \,. \tag{2.176}$$

Finally, we have

$$\mathbb{E}\left[\epsilon_{n+1}(c_1, \ldots, c_n, X)\right] \leq \mathbb{P}[i(X, Y) \leq \log \gamma(X)] + n \sup_{x \in \mathsf{A}} P_Y[i(x, Y) > \gamma(x)] \,. \tag{2.177}$$

Thus there must exist at least one value of X such that ϵ_{n+1} satisfies (2.169). The theorem is thus proved. ■

Remark: The proof technique of this theorem might be called *sequential random coding* because we have applied the random coding idea sequentially, codeword by codeword, instead of generating the whole codebook independently.

Some symmetric channels and choices of P_X (most notably the BEC and the BSC under equiprobable P_X) satisfy the sufficient condition in the next result.

Theorem 23 *Fix an arbitrary input distribution P_X. If the CDF $\mathbb{P}[i(x,Y) \le \alpha]$ does not depend on x for any α when Y is distributed according to P_Y, then there exists an (M,ϵ) code with maximal probability of error satisfying (for any $x \in \mathsf{A}$)*

$$\epsilon \quad \le \quad \mathbb{E}\left[\exp\left\{-[i(X,Y)-\log(M-1)]^+\right\}\right] . \tag{2.178}$$

Proof: Under the stated conditions, bound (2.170) yields

$$\epsilon \le \mathbb{P}[i(X,Y) \le \log\gamma(X)] + (M-1)\mathbb{P}[i(x,Y) > \log\gamma(x)] . \tag{2.179}$$

Thus $\gamma(x)$ can be optimized similarly to the proof of Theorem 18. ■

2.6 Achievability: input constraints

Of course, by restricting the input space A all of the achievability bounds proved so far yield bounds for the case with input constraints. However, such bounds are typically very inconvenient to use, because the auxiliary input distribution then has to be selected so that its support be on the constraint set $\mathsf{F} \subset \mathsf{A}$. For example, when $\mathsf{A} = \mathcal{A}^n$ it is convenient (analytically) to work with input distributions P_X that are obtained as n-fold products of single-letter distributions on $\mathcal{A}$. For this reason, it is advisable to find bounds which take input distributions on A but produce input-constrained codes with codewords inside F.

2.6.1 Generalization of the DT bound

Using Lemma 19 we can extend Theorem 18 to the case of input constraints as follows.

Theorem 24 *For any distribution P_X on A there exists a code with M codewords in the set F with average probability of error satisfying*

$$\epsilon \le \mathbb{P}\left[i(X,Y) \le \log\frac{M-1}{2}\right] + \frac{M-1}{2}\mathbb{P}\left[i(X,\bar{Y}) > \log\frac{M-1}{2}\right] + \mathbb{P}_X[\mathsf{F}^c] . \tag{2.180}$$

Proof: Set $\gamma(x)$ to be $\frac{M-1}{2}$ for $x \in \mathsf{F}$ and $+\infty$ for $x \in \mathsf{F}^c$. Then by Lemma 19 we have

$$\epsilon \le \mathbb{P}\left[\left\{i(X,Y) \le \log\frac{M-1}{2}\right\} \cup \left\{X \in \mathsf{F}^c\right\}\right] + \frac{M-1}{2}\mathbb{P}\left[i(X,\bar{Y}) > \log\frac{M-1}{2}, X \in \mathsf{F}\right] \tag{2.181}$$

Trivial upper-bounding yields (2.180). So by Lemma 19 we established the existence of the codebook and the decoder so that the average probability of error satisfies the required (2.180). However, in this codebook some of the codewords might fall outside the set F. On the other hand, our decoding rule is based on comparing $i(x,y)$ with the codeword-dependent threshold $\gamma(x)$. If the codeword does not belong to F then this threshold is $+\infty$. We conclude that all codewords in F^c have empty decoding sets. Thus, the average probability of error (under this suboptimal decoding) will not change if we remap all of these codewords to arbitrary $c_0 \in \mathsf{F}$. This proves the theorem. ■

Theorem 22 can be extended to the case of input constraints in the following way.

Theorem 25 *For any input distribution P_X and measurable $\gamma : \mathsf{A} \to [0,\infty]$, there exists a code with M codewords in the set F such that the maximal probability of error ϵ satisfies*

$$\epsilon P_X[\mathsf{F}] \le \mathbb{P}[i(X,Y) \le \log\gamma(X)] + (M-1)\sup_{x\in\mathsf{F}}\mathbb{P}[i(x,Y) > \log\gamma(x)] . \tag{2.182}$$

Proof: The case of $P_X[\mathsf{F}] = 0$ is trivial. Assume otherwise. The proof is similar to the proof of Theorem 22 with the only change being in how we upper-bound

$$\mathbb{E}\left[\epsilon_{n+1}(X)\right] = \mathbb{P}[\{i(X,Y) \le \log\gamma(X)\} \cup D]\,. \tag{2.183}$$

We proceed as follows

$$\mathbb{E}\left[P_X[\mathsf{F}]\epsilon_{n+1}(X)\right] = P_X[\mathsf{F}]\mathbb{P}[\{i(X,Y) \le \log\gamma(X)\} \cup D] \le \tag{2.184}$$

$$P_X[\mathsf{F}]\left(\mathbb{P}[i(X,Y) \le \log\gamma(X)] + P_Y[D]\right) = \tag{2.185}$$

$$\mathbb{E}\left[1_{\mathsf{F}}(X)\left(\mathbb{P}[i(X,Y) \le \log\gamma(X)] + P_Y[D]\right)\right]\,. \tag{2.186}$$

Now it is an elementary fact that if f and g are two non-negative functions, μ is a measure satisfying $\mu(F) > 0$ and

$$\int 1_F f\, d\mu \ge \int g\, d\mu \tag{2.187}$$

then at least at some $x^* \in F$ we have $g(x^*) \le f(x^*)$. In our case this means that there must exist at least one value of X (call it c_{n+1}) such that

$$P_X[\mathsf{F}]\epsilon_{n+1}(c_{n+1}) \le \mathbb{P}[i(X,Y) \le \log\gamma(X)] + P_Y[D]\,. \tag{2.188}$$

The rest of the proof follows that of Theorem 22 without change. ■

Comparing this result with Theorem 24 we note that (2.182) is stronger than the bound

$$\epsilon \le \mathbb{P}[i(X,Y) \le \log\gamma(X)] + (M-1)\sup_{x\in\mathsf{F}}\mathbb{P}[i(x,Y) > \log\gamma(x)] + P_X[\mathsf{F}^c]\,. \tag{2.189}$$

An immediate corollary of Theorem 25 is the following:

Corollary 26 *For any distribution P_X and any $\gamma > 0$, there exists an (M,ϵ) code (maximal probability of error) with codewords in the set $\mathsf{F} \subset \mathsf{A}$ such that*

$$M \ge 1 + \gamma\left(\epsilon P_X[\mathsf{F}] - \mathbb{P}\left[i(X;Y) < \log\gamma\right]\right)\,. \tag{2.190}$$

Proof: Apply Theorem 25 and use

$$P_Y[i(x,Y) \ge \log\gamma] \le \frac{1}{\gamma}\,. \tag{2.191}$$

■

Note that (2.190) is always stronger than a classical version of the input-constrained Feinstein's lemma (Theorem 1, see Section 2.2.1 for the history and references regarding the input-constrained version of Feinstein's lemma).

2.6.2 $\kappa\beta$ bound

Theorem 27 ($\kappa\beta$ bound) *For any $0 < \epsilon < 1$, any $0 < \tau < \epsilon$ and any distribution Q_Y on B, there exists an (M,ϵ) code with codewords chosen from $\mathsf{F} \subset \mathsf{A}$, satisfying*

$$M \ge \frac{\kappa_\tau(\mathsf{F}, Q_Y)}{\sup_{x\in\mathsf{F}}\beta_{1-\epsilon+\tau}(x, Q_Y)}\,. \tag{2.192}$$

Note: It is possible[12] that (2.192) will be of the form $M \geq \alpha/0$ with $\alpha > 0$. In this case the statement of the theorem should be understood as "(M, ϵ) codes with arbitrarily high M exist".

Proof: We first describe the operation of the decoder given a codebook $\{c_i\}_{i=1}^M$. Upon reception of $y \in \mathsf{B}$ the decoder sequentially tests whether codeword c_i was sent, where i runs from 1 to M. The test for c_i is performed as a binary hypothesis test discriminating $P_{Y|X=c_i}$ (hypothesis $\mathcal{H}_0$) against "average noise" Q_Y (hypothesis $\mathcal{H}_1$). We would like to select each such test as an optimal one within the constraint $P(\text{decide}\,\mathcal{H}_0|\mathcal{H}_0) \geq 1 - \epsilon + \tau$. To do this we define a collection of random variables $Z(x), x \in \mathsf{F}$ conditionally independent given Y and with $P_{Z(x)|Y}$ chosen so that it achieves $\beta_{1-\epsilon+\tau}(x, Q_Y)$ in (2.60). In other words,

$$P[Z(x) = 1|X = x] \geq 1 - \epsilon + \tau\,, \tag{2.193}$$

$$Q[Z(x) = 1] = \beta_{1-\epsilon+\tau}(x, Q_Y)\,, \tag{2.194}$$

where we denoted

$$Q[Z(x) = 1] \stackrel{\triangle}{=} \int_{\mathsf{B}} P_{Z(x)|Y}(1|y)dQ(y)\,. \tag{2.195}$$

The decoder applies the M independent random transformations $P_{Z(c_1)|Y}, \ldots, P_{Z(c_M)|Y}$ to the channel output Y and outputs the first index j such that $Z(c_j) = 1$, or 1 if all Z are zero.

Having specified the decoder operation we proceed to generate the codebook $\{c_i\}_{i=1}^M$. This will be done in a manner similar to the maximal coding idea of Feinstein.

At first step, choose any $c_1 \in \mathsf{F}$. Then, by (2.193) we know that the described decoder will decode c_1 with probability of at least $1 - \epsilon + \tau$ which is better than $1 - \epsilon$. So c_1 does not violate the maximum probability of error criterion. Next, suppose that j codewords have already been selected, then choose the $(j+1)$-st codeword. We can select some $x \in \mathsf{F}$ as the new codeword c_{j+1} only provided that

$$P\left[Z(x) = 1, Z(c_1) = \cdots = Z(c_j)|X = x\right] \geq 1 - \epsilon\,. \tag{2.196}$$

If we cannot find any such x then STOP; otherwise choose any x satisfying (2.196).

There are two cases. Either the process continues indefinitely, in which case there is nothing to prove, or it stops after a finite number of steps M. In the latter case, we have found an (M, ϵ) code and we need only to show that M satisfies the bound in (2.192). Note that there is a large amount of freedom in the process: each random variable $Z(c_i)$ is perhaps not uniquely defined by c_i, the choice of c_{j+1} is not unique, etc. However, the lower bound on M will be independent of all those choices[13] .

Denote

$$V_M = \max\{Z(c_j), j = 1, \ldots, M\}\,. \tag{2.197}$$

[12]For an example of such a case, take $\mathsf{A} = \mathsf{B} = [0, 1]$ with the Borel σ-algebra. Define $P_{Y|X=x}(y) = \delta_x(y)$, i.e. a point measure at $y = x$, and take Q_Y to be Lebesgue measure. Then, $\beta_\alpha(x, Q_Y) = 0$ for any x and α, and $\kappa_\tau(Q_Y) = 1$ for any $\tau > 0$.

[13]Note that we could make the procedure completely deterministic using the axiom of choice and well-ordering theorem to well-order all sets. Then, for example, at each step we can choose the first x (under the established order on F) that satisfies (2.196). This allows as to talk about *the* code constructed by Theorem 27.

Then the process stopping after M steps implies that for every $x \in \mathsf{F}$ we have

$$P\left[Z(x) = 1, V_M = 0 | X = x\right] < 1 - \epsilon\,. \tag{2.198}$$

But, by definition of $Z(x)$ and (2.193) it follows that

$$\begin{aligned} 1 - \epsilon + \tau &\leq P[Z(x) = 1 | X = x] && (2.199)\\ &\leq P[Z(x) = 1, V_M = 0 | X = x] + P[V_M = 1 | X = x] && (2.200)\\ &< 1 - \epsilon + P[V_M = 1 | X = x]\,. && (2.201) \end{aligned}$$

Thus, V_M is a random variable taking values in $\{0, 1\}$ and such that, for every $x \in \mathsf{F}$

$$P\left[V_M = 1 | X = x\right] \geq \tau\,. \tag{2.202}$$

But then, V_M defines a composite hypothesis test and, by the definition (2.92) of κ_τ we have

$$Q\left[V_M = 1\right] \geq \kappa_\tau(\mathsf{F}, Q_Y)\,. \tag{2.203}$$

Now, on the other hand

$$\begin{aligned} Q\left[V_M = 1\right] &= Q\left[\bigcup_1^M \{Z(c_j) = 1\}\right] && (2.204)\\ &\leq \sum_1^M Q\left[Z(c_j) = 1\right] && (2.205)\\ &= \sum_1^M \beta_{1-\epsilon+\tau}(c_j, Q_Y) && (2.206)\\ &\leq M \sup_{x \in \mathsf{F}} \beta_{1-\epsilon+\tau}(x, Q_Y)\,, && (2.207) \end{aligned}$$

where (2.206) follows by (2.194). Finally, (2.207) and (2.203) imply (2.192). ■

Using (2.93) in Theorem 27 we obtain a weakened but useful bound:

$$M \geq \sup_{0<\tau<\epsilon} \sup_{Q_X} \frac{\tau Q_X[\mathsf{F}]}{\sup_{x \in \mathsf{F}} \beta_{1-\epsilon+\tau}(x, Q_Y)} \tag{2.208}$$

where the supremum is over all input distributions, and Q_Y denotes the distribution induced by Q_X on the output. An interesting connection of the (weakened form of the) $\kappa\beta$ bound and the DT bound comes from the following observation. By a judicious choice of $\gamma(x)$ in Lemma 19 we could have obtained the bound (2.208) for average probability error with supremum in the denominator replaced by the average over Q_X.

In (2.60) and (2.92) we have defined β_α and κ_τ using randomized tests. Then, in Theorem 27 we have constructed the coding scheme with a randomized decoder. Correspondingly, if we define β_α and κ_τ using non-randomized tests, then the analog of Theorem 27 for a non-randomized decoder can be proved. For further details see Appendix B.

2.7 Converse bounds

In this section we develop a method for proving converse ("impossibility") results. The central idea can be summarized as follows: we take an arbitrary code for the channel $P_{Y|X}$; we prove that if used on a different channel $Q_{Y|X}$ this code must have large probability of error with a guaranteed lower bound; we then show that there is a link between the probability of error on the Q-channel and the probability of error on the P-channel; since the former is lower-bounded, so is the latter. Quite interestingly, we show that all the converse bounds mentioned in Section 2.2 as well as many new ones can be recovered as applications of the meta-converse.

2.7.1 Meta-converse: average probability of error

Theorem 28 *Consider two random transformations* $(\mathsf{A}, \mathsf{B}, P_{Y|X})$ *and* $(\mathsf{A}, \mathsf{B}, Q_{Y|X})$*. Fix a code* (f, g) *(possibly randomized encoder and decoder pair) and let* ϵ *and* ϵ' *be its average probability of error under the* P*-transformation and the* Q*-transformation, respectively. Also denote by* $P_X = Q_X$ *the probability distribution induced by the encoder* f *on the input alphabet* A*. Then we have*

$$\beta_{1-\epsilon}(P_{XY}, Q_{XY}) \leq 1 - \epsilon' \,. \tag{2.209}$$

Proof: Denote by W and $\hat{W}$ the random variable representing the input to the encoder (i.e. the message) and the output of the decoder (i.e. the message estimate), respectively. Then we have two joint distributions $P_{WXY\hat{W}}$ and $Q_{WXY\hat{W}}$ defined as follows:

$$P_{WXY\hat{W}}(w, x, y, \hat{w}) = \frac{1}{M} f(x|w) P_{Y|X}(y|x) g(\hat{w}|w) \,, \tag{2.210}$$

$$Q_{WXY\hat{W}}(w, x, y, \hat{w}) = \frac{1}{M} f(x|w) P_{Y|X}(y|x) g(\hat{w}|w) \,, \tag{2.211}$$

where $\frac{1}{M}$ represents the fact that W is equiprobable on $\{1, \ldots, M\}$. We define the following random variable

$$Z = 1\{W = \hat{W}\} \,. \tag{2.212}$$

The crucial observation is that the conditional distribution of Z given (X, Y) is the same for both P and Q; namely, we have

$$P_{Z|XY} = Q_{Z|XY} \,. \tag{2.213}$$

Indeed, we have

$$\begin{aligned}
\mathbb{P}[Z=1|X,Y] &= \sum_{j=1}^{M} \mathbb{P}[W=j,\hat{W}=j|X,Y] && (2.214)\\
&= \sum_{j=1}^{M} \mathbb{P}[W=j|X,Y]\mathbb{P}[\hat{W}=j|X,Y] && (2.215)\\
&= \sum_{j=1}^{M} \mathbb{P}[W=j|X]\mathbb{P}[\hat{W}=j|Y] && (2.216)\\
&= \sum_{j=1}^{M} \mathbb{P}[W=j|X]g(j|Y)\,, && (2.217)
\end{aligned}$$

where (2.214) is by definition of Z, (2.215) follows since under both P and Q we have a Markov chain: $W - X - Y - \hat{W}$ and therefore, conditioned on (X,Y) W and $\hat{W}$ are independent; and (2.216) is also a consequence of the Markov chain condition. Finally, (2.216) implies (2.213) since $\mathbb{P}[W=j|X]$ in each term of the sum depends only on the joint distribution of X and W, while by construction we have that $P_{W,X} = Q_{W,X}$.

Overall, $P_{Z|XY}$ defines a transition probability kernel $\mathsf{A}\times\mathsf{B}\to\{0,1\}$ and therefore constitutes a binary hypothesis test between P_{XY} and Q_{XY} satisfying

$$\begin{aligned}
\sum_{x\in\mathsf{A}}\sum_{y\in\mathsf{B}} P_{Z|XY}(1|xy)P_{XY}(x,y) &= 1-\epsilon && (2.218)\\
\sum_{x\in\mathsf{A}}\sum_{y\in\mathsf{B}} P_{Z|XY}(1|xy)Q_{XY}(x,y) &= 1-\epsilon'\,. && (2.219)
\end{aligned}$$

Therefore, by the definition of β_α in (2.60) we have

$$\beta_{1-\epsilon}(P_{XY},Q_{XY}) \le 1-\epsilon'\,. \tag{2.220}$$

■

Theorem 28 allows one to use any converse for channel $Q_{Y|X}$ to prove a converse for channel $P_{Y|X}$. It has many interesting generalizations (for example, to list-decoding and channels with feedback) and applications, whose study is outside the scope of the thesis.

Here we want to briefly explain an alternative (and a more illuminating) way to prove the crucial step (2.213). Note that for any probability space, we can define a directed acyclic graph (DAG) connecting its variables via transition probability kernels. Then following this DAG we can reconstruct the joint distribution for all the random variables. For example, in the proof of Theorem 28 the DAG was the following:

$$W \xrightarrow{f} X \overset{P_{Y|X}}{\underset{Q_{Y|X}}{\rightrightarrows}} Y \xrightarrow{g} \hat{W} \tag{2.221}$$

where the choice between $P_{Y|X}$ or $Q_{Y|X}$ is made depending on the channel used. Conversely, every DAG (with arrows marked by transition probability kernels[14]) generates a unique

[14]If a random variable does not have inbound arrows, it should be marked with a distribution, according to which it is generated; if a random variable has several inbound arrows such as

$$A \longrightarrow C \longleftarrow B \tag{2.222}$$

probability space (joint distribution on the variables, corresponding to its vertices). We say that two DAGs are equivalent if they generate the same probability space. An argument based on Markov chain condition, proves that the following two DAGs are equivalent:

$$W \xrightarrow{f} X \xrightarrow{g} Y \quad \sim \quad W \xleftarrow{f'} X \xrightarrow{g} Y\,, \tag{2.223}$$

where in the right-hand DAG, the distribution P_X is taken as $f \circ P_W$, and $f'(w|x)$ is just the Bayes rule inversion of the kernel $f(x|w)$:

$$f'(w|x) = \frac{f(x|w)P_W(w)}{\sum_{w' \in \mathsf{W}} f(x|w')P_W(w')}\,. \tag{2.224}$$

The principal observation is that f' does not depend on g in (2.223). Therefore, applying this to the DAG (2.221) and reintroducing Z as a function of $(W, \hat{W})$ we get the following equivalent DAG:

$$\begin{array}{ccccc} X & \xrightarrow{f'} & W & & \\ \big\downarrow & & & \searrow & \\ Y & \xrightarrow{g} & \hat{W} & \longrightarrow & Z \end{array} \tag{2.225}$$

and the composite arrow $(X, Y) \to Z$ does not depend on the particular arrow chosen between X and Y. This is precisely the meaning of (2.213).

A simple application of Theorem 28 yields the following result.

Theorem 29 (Converse) *Every (M, ϵ) code (average probability of error) with codewords belonging to F satisfies*

$$M \le \sup_{P_X} \inf_{Q_Y} \frac{1}{\beta_{1-\epsilon}(P_{XY}, P_X \times Q_Y)}\,, \tag{2.226}$$

where P_X ranges over all distributions on F, and Q_Y ranges over all distributions on B.

Proof: Denote the distribution of the encoder output by $\bar{P}_X$ and particularize Theorem 28 by choosing $Q_{Y|X} = Q_Y$ for an arbitrary Q_Y, in which case we obtain $\epsilon' = 1 - \frac{1}{M}$. Therefore, from (2.209) we obtain

$$\begin{aligned} \frac{1}{M} &\ge \sup_{Q_Y} \beta_{1-\epsilon}(\bar{P}_X P_{Y|X}, \bar{P}_X \times Q_Y) && (2.227)\\ &\ge \inf_{P_X} \sup_{Q_Y} \beta_{1-\epsilon}(P_{XY}, P_X \times Q_Y). && (2.228) \end{aligned}$$

■

As we will see shortly in important special cases $\beta_\alpha(x, Q_Y)$ is constant on F. In those cases the following converse is particularly useful.

Theorem 30 *Fix a probability measure Q_Y on B. Suppose that $\beta_\alpha(x, Q_Y) = \beta_\alpha(Q_Y)$ for $x \in \mathsf{F}$. Then every (M, ϵ)-code (average probability of error) satisfies*

$$M \le \frac{1}{\beta_{1-\epsilon}(Q_Y)}\,. \tag{2.229}$$

---

then the kernel should be $P_{C|AB}$, which acts from $A \times B$ to C.

Proof: The result follows from Theorem 29 and the following auxiliary result. ■

Lemma 31 *Suppose that* $\beta_\alpha(P_{Y|X=x}, Q_{Y|X=x}) = \beta_\alpha$ *is independent of* $x \in \mathsf{F}$. *Then, for any* P_X *supported on* F *we have*

$$\beta_\alpha(P_X P_{Y|X}, P_X Q_{Y|X}) = \beta_\alpha(P_{Y|X=x}, Q_{Y|X=x})\,. \tag{2.230}$$

Proof: Take a collection of optimal tests Z_x for each pair $P_{Y|X=x}$ *vs.* $Q_{Y|X=x}$, i.e.

$$\begin{aligned} P_{Y|X=x}[Z_x = 1] &\geq 1-\alpha\,, &(2.231)\\ Q_{Y|X=x}[Z_x = 1] &= \beta_{1-\alpha}\,. &(2.232)\end{aligned}$$

Then take Z_X as a test for P_{XY} *vs.* Q_{XY}. In this way we get

$$\beta_\alpha(P_X P_{Y|X}, P_X Q_{Y|X}) \leq \beta_\alpha(P_{Y|X=x}, Q_{Y|X=x})\,. \tag{2.233}$$

Since β_α is non-decreasing and convex, the reverse inequality follows from the next lemma. ■

Lemma 32 *Suppose that there is an non-decreasing convex function* $f : [0,1] \to [0,1]$ *such that for all* $x \in \mathsf{F}$ *we have*

$$\beta_\alpha(P_{Y|X=x}, Q_{Y|X=x}) \geq f(\alpha) \tag{2.234}$$

Then, for any P_X *supported on* F *we have*

$$\beta_\alpha(P_X P_{Y|X}, P_X Q_{Y|X}) \geq f(\alpha)\,. \tag{2.235}$$

Proof: Consider an arbitrary test Z such that

$$P_{XY}[Z=1] = \sum_{x\in\mathsf{A}} P_X(x) P_{Y|X=x}[Z=1] \geq \alpha\,. \tag{2.236}$$

Then observe that

$$\begin{aligned} \sum_{x\in\mathsf{A}} P_X(x) Q_{Y|X=x}[Z=1] &\geq \sum_{x\in\mathsf{A}} P_X(x)\beta_{P_{Y|X=x}[Z=1]}(P_{Y|X=x}, Q_{Y|X=x}) &(2.237)\\ &\geq \sum_{x\in\mathsf{A}} P_X(x) f(P_{Y|X=x}[Z=1]) &(2.238)\\ &\geq f(P[Z=1]) &(2.239)\\ &\geq f(\alpha)\,, &(2.240)\end{aligned}$$

where (2.238) follows from (2.235), (2.239) is by Jensen's inequality, and (2.240) follows because f is non-decreasing function of α. Therefore, taking infimum over all tests $P_{Z|XY}$, from (2.240) we obtain that

$$\beta_\alpha(P_X P_{Y|X}, P_X Q_{Y|X}) \geq f(\alpha)\,. \tag{2.241}$$

■

2.7.2 Meta-converse: maximal probability of error

To apply Theorem 28, one needs to prove a lower bound on $\beta_\alpha(P_{XY}, Q_{XY})$. However, since the distribution P_X depends on the code, obtaining a lower bound valid for all P_X is generally a hard problem. A much easier problem is computing $\beta_\alpha(P_{Y|X=x}, Q_{Y|X=x})$. Then Lemma 32 gives a lower-bound on $\beta_\alpha(P_{XY}, Q_{XY})$ (or even a precise value of β_α, if conditions of Lemma 31 are satisfied). However, for the maximal probability of error formalism, there is no such problem: namely, the function $\beta_\alpha(P_{Y|X=x}, Q_{Y|X=x})$ takes up the role of $\beta_\alpha(P_{XY}, Q_{XY})$ in Theorem 28, as follows:

Theorem 33 *Consider two random transformations* $(\mathsf{A}, \mathsf{B}, P_{Y|X})$ *and* $(\mathsf{A}, \mathsf{B}, Q_{Y|X})$. *Fix a code* (f, g) *(possibly with a randomized decoder) with codewords belonging to a constraint set* $\mathsf{F} \subseteq \mathsf{A}$. *Let* ϵ *and* ϵ' *be its maximal probability of error under the* P*-transformations and the* Q*-transformations, respectively. Then,*

$$\inf_{x \in \mathsf{F}} \beta_{1-\epsilon}(P_{Y|X=x}, Q_{Y|X=x}) \le 1 - \epsilon' . \tag{2.242}$$

Proof: Consider an (M, ϵ)-code with codewords $\{c_j \in \mathsf{F}\}_{j=1}^M$ and a randomized decoding rule $P_{Z|Y} : \mathsf{B} \mapsto \{0, \dots, M\}$. We have for some j^*

$$\sum_{b \in \mathsf{B}} P_{Z|Y}(j^*|b) Q_{Y|X}(b|j^*) = 1 - \epsilon' , \tag{2.243}$$

and at the same time

$$\sum_{b \in \mathsf{B}} P_{Z|Y}(j^*|b) P_{Y|X}(b|j^*) \ge 1 - \epsilon . \tag{2.244}$$

Consider the hypothesis test between $P_{Y|X=j^*}$ and $Q_{Y|X=j^*}$ that decides in favor of $P_{Y|X=j^*}$ only when the decoder output is j^*. By (2.244) the probability of correct decision under $P_{Y|X=j^*}$ is at least $1 - \epsilon$, and therefore

$$\begin{aligned} 1 - \epsilon' &\ge \beta_{1-\epsilon}(P_{Y|X=j^*}, Q_{Y|X=j^*}) && (2.245) \\ &\ge \inf_{x \in \mathsf{F}} \beta_{1-\epsilon}(P_{Y|X=x}, Q_{Y|X=x}) . && (2.246) \end{aligned}$$

■

Theorem 34 (Converse) *Every* (M, ϵ) *code (maximal probability of error) with codewords belonging to* F *satisfies*

$$M \le \inf_{Q_Y} \sup_{x \in \mathsf{F}} \frac{1}{\beta_{1-\epsilon}(x, Q_Y)} , \tag{2.247}$$

where the infimum is over all distributions Q_Y *on* B.

Proof: Repeat the argument of the proof of Theorem 29 replacing Theorem 28 by Theorem 33. ■

2.7.3 Applications of the meta-converse

We illustrate how Theorems 28 and 33 can be used to prove classical converse results (including all of the cited in Section 2.2):

- Fano's inequality (Theorem 5): Particularize (2.227) to the case $Q_Y = P_Y$, where P_Y is the output distribution induced by the code and the channel $P_{Y|X}$. Note that any hypothesis test is a (randomized) binary-output transformation and therefore, by the data-processing inequality for divergence we have

$$d\big(1-\epsilon\big|\big|\beta_{1-\epsilon}(P_{XY}, P_X \times P_Y)\big) \le D(P_{XY}||P_X \times P_Y)\,, \tag{2.248}$$

  where the binary divergence function satisfies

$$\begin{aligned} d(a||b) &= a\log\frac{a}{b} + (1-a)\log\frac{1-a}{1-b} && (2.249)\\ &\ge -h(a) + a\log\frac{1}{b}\,. && (2.250)\end{aligned}$$

  Using (2.249) in (2.248) we obtain

$$\log\frac{1}{\beta_{1-\epsilon}(P_{XY}, P_X \times P_Y)} \le \frac{I(X;Y)+h(\epsilon)}{1-\epsilon}\,. \tag{2.251}$$

  Fano's inequality (2.43) follows from (2.251) and (2.227).

- Information spectrum converse (Theorem 8): Replace (2.251) with (2.67), which together with (2.227) yields

$$\begin{aligned} \frac{1}{M} &\ge \beta_{1-\epsilon}(P_{XY}, P_X \times P_Y) && (2.252)\\ &\ge \sup_{\gamma>0}\frac{1}{\gamma}(\mathbb{P}[i(X;Y) < \log\gamma] - \epsilon)\,, && (2.253)\end{aligned}$$

  where P_Y is a distribution on B induced a given (M,ϵ) code. The bound (2.253) is equivalent to the converse bound (2.47). Similarly, by using a stronger bound in place of (2.67) we can derive [41]. Furthermore, by keeping the freedom in choosing Q_Y in (2.227) we can prove a stronger version of the result.

- A stronger bound due to Poor and Verdú [41], Theorem 9, can also be obtained. Indeed, since the distribution P_X induced on A by a given (M,ϵ) code is discrete with atoms of weight (at least) $\frac{1}{M}$ we must have

$$\inf_{\mathsf{A}\times\mathsf{B}}\frac{d(P_X \times P_Y)}{dP_{XY}} \ge \frac{1}{M}\,. \tag{2.254}$$

  Therefore, using the following Lemma with $\theta = \frac{1}{M}$ we obtain

$$\begin{aligned} \beta_{1-\epsilon}(P_{XY}, P_X \times P_Y) &\ge \frac{1}{\gamma}\left(1-\epsilon-\left(1-\frac{\gamma}{M}\right)\mathbb{P}[i(X;Y)\ge\gamma]\right) && (2.255)\\ &= \frac{1}{M} + \frac{1}{\gamma}\left(\left(1-\frac{\gamma}{M}\right)\mathbb{P}[i(X;Y)\ge\gamma] - \epsilon\right)\,, && (2.256)\end{aligned}$$

  which together with (2.252) implies Theorem 9.

Lemma 35 *For a pair (P, Q) of probability distributions on W and any $\gamma \geq 0$ we have*

$$\alpha \leq \gamma\beta_\alpha(P,Q) + (1-\gamma\theta)P\left[\frac{dP}{dQ} \geq \gamma\right], \tag{2.257}$$

provided that

$$0 \leq \theta \leq \inf_{x \in \operatorname{supp} Q \cap \operatorname{supp} P} \frac{dQ}{dP}. \tag{2.258}$$

Proof: Follow the proof of (2.67) and replace the lower bound on

$$P\left[\{Z=1\} \cap \left\{\frac{dP}{dQ} < \gamma\right\}\right] \tag{2.259}$$

with the one obtained from the simple identity:

$$\frac{dP}{dQ} \cdot 1\left\{\frac{dP}{dQ} < \gamma\right\} \leq \gamma + \left(\frac{dP}{dQ} - \gamma\right) t, \tag{2.260}$$

where $t = \frac{-\gamma\theta}{1-\gamma\theta}$. ■

- Wolfowitz's strong converse (Theorem 6): apply Theorem 34 with some arbitrary Q_Y. Then from (2.67) we have:

$$\inf_{x \in \mathsf{A}} \beta_\alpha(x, Q_Y) \geq \sup_{\gamma > 0} \frac{1}{\gamma}\left(\alpha - \sup_{x \in \mathsf{A}} P_{Y|X=x}\left[\frac{dP_{Y|X=x}}{dQ_Y} \geq \gamma\right]\right). \tag{2.261}$$

  Now, suppose that $Q_Y = P_Y$, then by (2.27) we conclude that Theorem 34 implies Theorem 6. Retaining the freedom of choice of Q_Y in (2.261) we obtain Theorem 10.

- Shannon-Gallager-Berlekamp (Theorem 11): Applying Theorem 34, we may first split the input space A into regions F_i such that $\beta_\alpha(x, Q_Y)$ is constant within F_i. For example, for symmetric channels and Q_Y equal to the capacity achieving output distribution, there is no need to split A since $\beta_\alpha(x, Q_Y)$ is identical for all $x \in \mathsf{A}$. For a general DMC, we apply Theorem 28 with $Q_{Y|X}$ chosen as follows. The distribution $Q_{Y|X=x^n}$ depends only on the type of x^n and is chosen optimally for each type (and depending on the coding rate). Under the Q-transformation, the decoder can at most distinguish codewords belonging to different types and therefore, we can estimate $1 - \epsilon' \leq \frac{n^{|\mathcal{A}|-1}}{M}$. Using this estimate in (2.209), the proof of Theorem 11 follows along the same lines as the proof of [45, Theorem 19] by weakening (2.209) using Chernoff-type estimates.

- Refinements to Theorem 11 in [18] and [19]: As we explained above, Theorem 11 is obtained from Theorem 34 by choosing $Q_{Y|X}$ judiciously and by performing a large deviation analysis of β_α. Reference [18] improved Theorem 11 by extending the results to the case of infinite $|\mathsf{B}|$ and by tightening the Chernoff-type estimates of [7]. A further improvement was found in [19] for the special case of input-symmetric channels by directly lower-bounding the average probability of error and avoiding the step of splitting a code into constant composition subcodes. Theorem 30 is tighter

than the bound in [19] because for symmetric channels and relevant distributions Q_Y the value of $\beta_\alpha(x, Q_Y)$ does not depend on x and therefore the average probability of error is bounded directly.

- Low-rate converse bounds on the error-exponent can be also be obtained. For example, when $P_{Y|X}$ is the $BSC(n, \delta)$ then we take $Q_{Y|X}$ as the channel that produces exactly $\lceil \mu n \rceil$ errors equiprobably out of $\binom{n}{\mu n}$ possibilities, where $0 < \mu < 1$ is some parameter. Then, a simple analysis shows that

$$\beta_\alpha(P_{Y|X=x}, Q_{Y|X=x}) \quad = \quad \left[\frac{\alpha - 1 + p_n}{p_n}\right]^+ \tag{2.262}$$

where

$$p = \binom{n}{\mu n}\delta^{\mu n}\bar{\delta}^{n-\mu n} \approx \exp\{-nd(\mu||\delta)\}\,. \tag{2.263}$$

If the code has minimum distance νn then its maximal probability of error on the Q-transformation (assuming the maximum likelihood decoding for the P-transformation) satisfies (with exponential precision):

$$\epsilon' \quad \gtrsim \quad \begin{cases} 0\,, & \mu < \nu/2\,, \\ \exp(-n\left[h(\mu) - \nu - (1-\nu)h\left(\frac{\mu-\nu/2}{1-\nu}\right)\right])\,, & \nu/2 \le \mu < 1/2\,, \\ 1\,. & 1/2 \le \mu \end{cases} \tag{2.264}$$

Then, by taking

$$\mu = \delta(1-\nu) + \frac{\nu}{2} \tag{2.265}$$

and applying Theorem 33 with (2.262) and (2.264), we obtain that the maximal probability of error over the P-transformation satisfies:

$$\epsilon \gtrsim (4\delta(1-\delta))^{\nu n/2}\,. \tag{2.266}$$

Since by the Plotkin bound $\nu \le 1/2$ for any code with positive rate $R > 0$ we obtain that any $(n, 2^{nR}, \epsilon)$ code with positive rate satisfies

$$\epsilon \gtrsim \exp\left(-n\frac{1}{2}d(\tfrac{1}{2}||\delta)\right)\,, \tag{2.267}$$

which is a well-known zero-rate bound for the BSC, see also [7].

Above we have shown that Fano's inequality is a direct consequence of the meta-converse and the data-processing inequality (2.248). Replacing divergence with another f-divergence, we get other useful bounds. For example, let us take Hellinger divergence which for $\lambda > 0$, $\lambda \ne 1$ is defined as [49]

$$D_\lambda(P||Q) \quad \stackrel{\triangle}{=} \quad \frac{1}{\lambda - 1}\mathbb{E}_Q\left[\left(\frac{dP}{dQ}\right)^\lambda - 1\right] \tag{2.268}$$

$$= \quad \frac{1}{\lambda - 1}\mathbb{E}_P\left[\left(\frac{dP}{dQ}\right)^{\lambda-1} - 1\right]\,. \tag{2.269}$$

Then following the same argument as in the derivation of Fano's inequality, we obtain: *Any* (M, ϵ) *code must satisfy for every* $\lambda > 0$*:*

$$\frac{1}{\lambda - 1}\left((1-\epsilon)^{\lambda} M^{\lambda-1} + \epsilon^{\lambda}\right) \leq \frac{1}{\lambda - 1}\mathbb{E}\left[\exp((\lambda - 1) i(X;Y))\right], \tag{2.270}$$

and in particular for $\rho > 0$*:*

$$(1-\epsilon)^{1+\rho} M^{\rho} \leq \mathbb{E}\left[\exp(\rho \cdot i(X;Y))\right]. \tag{2.271}$$

Note that as $\lambda \to 1$ inequality (2.270) converges to Fano's inequality (2.43).[15] Inequality (2.271) demonstrates that the right-hand side of (2.271) for a reliable (small ϵ) code should not be too different from M^{ρ}.

Although we do not further use (2.270) and (2.271) in this work, these inequalities can be shown to be a useful tool for proving lower bounds on the error-probability for a code with a known weight distribution (in the spirit of [55]) and analyzing error-exponents of the families of codes.

[15] Subtract $\frac{1}{\lambda-1}$ from both sides of (2.270) before taking the limit.

Chapter 3

Discrete channels

In this chapter, the general methods of Chapter 2 are particularized to various discrete channels with the aim of computing the channel dispersion, obtaining tight non-asymptotic bounds and refined asymptotic expansions. Section 3.1 reviews some of the previous work specific to discrete channels. In particular, classical bounds for the binary symmetric channel (BSC) and the binary erasure channel (BEC), as well as Strassen's asymptotic expansion are discussed. Then, Sections 3.2 and 3.3 show new results regarding the BSC and the BEC, respectively. In Section 3.4 we give a proof of (an amended version of) Strassen's theorem for the general discrete memoryless channel (DMC). The material in Sections 3.1-3.4 has been presented in part in [32,33]. Compared to [32], we show the $O(1)$ lower bound in Strassen's theorem for general DMC without additional assumptions, which follows from a stronger version of the achievability bound, Theorem 47. The full proof of the achievability bound for the exotic DMC, Theorem 51, and the refined results on the $\log n$ term, Section 3.4.5, also appear here for the first time. A simple model involving dynamically changing state, the Gilbert-Elliott channel, is addressed in Section 3.5. Finally, in Section 3.6 the idea of a normal approximation for the composite channels is demonstrated on the example of a non-ergodic mixture of two BSCs. Each section contains extensive numerical evaluations, validating both the need for and the usefulness of the knowledge of channel dispersion. The material in Sections 3.5-3.6 has been presented in part in [56, 57].

3.1 Previous work

We have already discussed in Section 2.2 a number of previously known bounds. Here we list additional results that have appeared in the information theory literature for special discrete channels, considered in this chapter.

3.1.1 Bounds for special discrete channels

For a linear code over the BSC, Poltyrev [24] proved the following upper-bound on the probability of error.

Theorem 36 (Poltyrev) *The maximal probability of error P_e under maximum likelihood*

decoding of a linear code[1] *with weight distribution*[2] $\{A_w, w = 0, \dots n\}$ *over the BSC with crossover probability* δ *satisfies*

$$P_e \;\leq\; \sum_{\ell=0}^{n} \delta^{\ell}(1-\delta)^{n-\ell} \min\left\{ \binom{n}{\ell}, \sum_{w=0}^{n} A_w B(\ell, w, n) \right\}, \tag{3.1}$$

where

$$B(\ell, w, n) \;=\; \sum_{w/2 \leq t \leq \min\{\ell, w\}} \binom{w}{t} \binom{n-w}{\ell - t}. \tag{3.2}$$

A $[k, n]$ linear code is generated by a $k \times n$ binary matrix. We can average (3.1) over an equiprobable ensemble of such matrices. Applying Jensen's inequality to pass expectation inside the minimum and noticing that $\mathbb{E}[A_w] = 2^{k-n}\binom{n}{w}$ we obtain the following achievability bound.

Theorem 37 *For a BSC with crossover probability* δ *there exists a* $[k, n]$ *linear code such that a maximum likelihood decoder has a maximal probability of error* P_e *satisfying*

$$P_e \;\leq\; \sum_{\ell=0}^{n} \delta^{\ell}(1-\delta)^{n-\ell} \min\left\{ \binom{n}{\ell}, \sum_{w=0}^{n} 2^{k-n} \binom{n}{w} B(\ell, w, n) \right\}, \tag{3.3}$$

where $B(\ell, w, n)$ *is given by* (3.2)*.*

A negligible improvement to (3.3) is possible if we average (3.1) over an ensemble of all full-rank binary matrices instead. Another modification by expurgating low-weight codewords [58] leads to a tightening of (3.3) when the rate is much lower than capacity.

For the BEC the results of [59, Theorem 9] can be used to compute the exact value of the probability of error over an ensemble of all linear codes generated by full-rank $k \times n$ binary matrices [60].

Theorem 38 (Ashikhmin) *Given a BEC with erasure probability* δ*, the average probability of error over all binary* $k \times n$ *linear codes with full-rank generating matrices chosen equiprobably is equal to*

$$P_e = \sum_{i=0}^{n} \binom{n}{i} \delta^{n-i}(1-\delta)^{i} \sum_{r=\max\{0, k-n+i\}}^{\min\{k,i\}} \begin{bmatrix} i \\ r \end{bmatrix} \begin{bmatrix} n-i \\ k-r \end{bmatrix} \begin{bmatrix} n \\ k \end{bmatrix}^{-1} 2^{r(n-i-k+r)} \left(1 - 2^{r-k}\right), \tag{3.4}$$

where

$$\begin{bmatrix} a \\ r \end{bmatrix} \triangleq \prod_{j=0}^{r-1} \frac{2^a - 2^j}{2^r - 2^j} \tag{3.5}$$

is the number of r*-dimensional subspaces of* $\mathbb{F}_2^a$*.*

[1] The same bound can be shown for a non-linear code by generalizing the notion of weight distribution. In this case, however, the upper bound only holds for the average probability of error, not the maximal.

[2] We define A_0 to be the number of 0-weight codewords in the codebook minus 1. In particular, for a linear codebook with no repeated codewords $A_0 = 0$.

Numerical evaluation and some improvements for the finite blocklengths bounds and various special channels have been made in [11, 18–20, 23, 61] among others. Of particular interest for non-asymptotic analysis is the treatment of the BSC in [12], where the authors numerically compute exact error probability averaged over the random ensemble of codebooks (for small blocklengths) and compare it with the sphere packing bound as well as with the performance of some practical codes.

3.1.2 Asymptotic expansions

The importance of studying the asymptotics of the function $M^*(n,\epsilon)$ was realized already in [2]. Shannon [2, Theorem 12] and Wolfowitz [3] originally showed that for the DMC (and later for some other memoryless channels) the following expansion holds, regardless of $\epsilon \in (0,1)$:

$$\log M^*(n,\epsilon) = nC + o(n)\,, \tag{3.6}$$

where C is the channel capacity. Later, Wolfowitz [39] improved the $o(n)$ term to $O(\sqrt{n})$. In parallel, Weiss [27] showed that for the BSC with crossover probability $\delta < \frac{1}{2}$,

$$\log_2 M^*(n,\epsilon) \le n(1-h(\delta)) - Q^{-1}(\epsilon)\sqrt{n}\sqrt{\delta-\delta^2}\log_2\frac{1-\delta}{\delta} + o(\sqrt{n})\,, \tag{3.7}$$

where Q^{-1} denotes the functional inverse of the Q-function (1.18). For symmetric DMCs, whose transition matrices are such that the rows are permutation of each other and so are the columns, Dobrushin [29, (75)] claimed without proof that[3]

$$\log M^*(n,\epsilon) = nC - Q^{-1}(\epsilon)\sqrt{nV} + O(\log n)\,, \tag{3.8}$$

where C and V are the expectation and the variance of the information density $i(X;Y)$ under the equiprobable input distribution. Dobrushin credited Pinsker for raising the important question of obtaining more terms in the asymptotic expansion, compared to (3.6).

Strassen [1] showed that (3.8) holds for an arbitrary DMC, thus generalizing and strengthening all previous results. For the general DMC, C in (3.8) is the capacity and V is the variance of the information density $i(X,Y)$ under a capacity achieving distribution P_X which also minimizes the variance V (if $\epsilon < 1/2$) or maximizes it (if $\epsilon > 1/2$). According to Definition 7, in this thesis we call V *the channel dispersion*.

For the Gilbert-Elliott channel [62, 63], considered in Section 3.5, the capacity was found by Mushkin and Bar-David [64]. The ϵ-capacity of a mixture of BSCs, considered in Section 3.6, is well known (e.g., [26, 39]), except for the points of discontinuity of C_ϵ. In determining the C_ϵ at these points recent progress was made by Kieffer in [65]. Otherwise, to the best of our knowledge the finite-blocklength analysis of these channels (regarding bounds and refined expansions of the form (3.8)) is attempted here for the first time.

[3] An incorrect $\log n$ term was also claimed.

3.2 Binary symmetric channel (BSC)

This section illustrates the application of the bounds developed so far to the BSC with crossover probability $\delta < 1/2$. Recall that for the BSC the input and output alphabets are binary, $\mathsf{A} = \mathsf{B} = \{0,1\}^n$, and the channel is defined as

$$P_{Y^n|X^n}(y^n|x^n) = \delta^{|y^n - x^n|}(1-\delta)^{n-|y^n-x^n|}\,, \tag{3.9}$$

where $|z^n|$ denotes the Hamming weight of the binary vector z^n. We will compute an achievability bound under the maximal probability of error criterion and a converse under the average one.

3.2.1 Bounds

Choosing P_X in Theorems 17 and 18 to be equiprobable on the input alphabet we obtain:

Corollary 39 *For the BSC with crossover probability δ, there exists an (n, M, ϵ) code (average probability of error) such that (RCU bound)*

$$\epsilon \le \sum_{t=0}^{n} \binom{n}{t} \delta^t (1-\delta)^{n-t} \min\left\{1,\ (M-1)\sum_{s=0}^{t}\binom{n}{s}2^{-n}\right\}\,, \tag{3.10}$$

and (DT bound)

$$\epsilon \le \sum_{t=0}^{n} \binom{n}{t} \delta^t (1-\delta)^{n-t} \min\left\{1, (M-1)2^{-n-1}(1-\delta)^{t-n}\delta^{-t}\right\}\,. \tag{3.11}$$

Whenever $M = 2^k$ for integer k, the statement holds for maximal probability of error as well.

Proof: The proof follows once we notice that with the equiprobable input distribution the information density is

$$i(x^n; y^n) = n\log(2-2\delta) + t\log\frac{\delta}{1-\delta} \tag{3.12}$$

where t is the Hamming weight of the difference between x^n and y^n. Thus, for example, for the RCU bound (2.105), we get

$$\mathbb{P}\left[i(\bar{X}^n; Y^n) \ge i(X^n; Y^n)\,\middle|\, X^n = x^n, Y^n = y^n\right] = \sum_{k=0}^{t}\binom{n}{k}2^{-n}\,, \tag{3.13}$$

since $\bar{X}^n$ is equiprobable and independent of X^n. Similarly, the DT bound (2.118) implies (3.11) after substituting (3.12).

The statement about the maximal probability of error is explained in Appendix C. ■

Note that RCU bound (3.10) has computational complexity $O(n^2)$, while the DT bound (3.11) has complexity $O(n)$ and therefore both are computable for blocklengths of practical interest. As discussed in Section 2.4.2, the DT bound (3.11) can be interpreted as $\frac{M+1}{2}$ times

the probability of error of an optimal binary hypothesis test between n fair coin flips (with prior probability $\frac{M-1}{M+1}$) and n bias-δ coin flips (with prior probability $\frac{2}{M+1}$).

Before numerical computation, we need to convert upper bounds on probability of error to lower bounds on $M^*(n,\epsilon)$. Doing so is straightforward: if we have a bound

$$\epsilon \le f(M)\,, \tag{3.14}$$

where $f(M)$ is some function of M (and n, k, δ), then we need to find the largest $M = 2^k$ such that the right-hand side of (3.14), $f(M)$, is still below the prescribed ϵ:

$$M^*(n,\epsilon) \ge \max\left\{2^k : f(2^k) \le \epsilon\right\}\,. \tag{3.15}$$

For comparison, Feinstein's lemma, with equiprobable P_X, yields the following bound:

$$M^*(n,\epsilon) \ge \sup_{t>0} 2^{nt}\left(\epsilon - P\left[Z \ge n(a-t)/b\right]\right)\,, \tag{3.16}$$

where $Z \sim B(n,\delta)$.

Gallager's random coding bound (2.34) also with equiprobable P_X, ensures that[4]

$$\log_2 M^*(n,\epsilon) \ge nE_r^{-1}\left(\frac{1}{n}\log_2\frac{1}{\epsilon}\right)\,, \tag{3.17}$$

where [9, Theorem 5.6.2, Corollary 2 and Example 1 in Section 5.6.]

$$E_r(1-h(s)) = \begin{cases} d(s||\delta)\,, & s \in (\delta, s^*]\,, \\ h(s) - 2\log s_1\,, & s > s^*\,, \end{cases} \tag{3.18}$$

and $s^* = \frac{\sqrt{\delta}}{\sqrt{\delta}+\sqrt{1-\delta}}$, $s_1 = \sqrt{\delta} + \sqrt{1-\delta}$.

Regarding Poltyrev's bound, Theorem 37, it turns out that (3.3), derived using linear codes and weight spectra, is in fact equal to (3.10) with $M-1$ replaced by 2^k. Indeed, notice that

$$\sum_{w=0}^{n}\binom{n}{w}\sum_{w/2\le t\le \min\{\ell,w\}}\binom{w}{t}\binom{n-w}{\ell-t} = \binom{n}{\ell}\sum_{s=0}^{\ell}\binom{n}{s}\,. \tag{3.19}$$

This holds since on the left we have counted all the ways of choosing two binary n-vectors X and Z such that $\mathrm{wt}(Z) = \ell$ and Z overlaps at least a half of X. The last condition is equivalent to requiring $\mathrm{wt}(X-Z) \le \mathrm{wt}(Z)$. So we can choose Z in $\binom{n}{\ell}$ ways and X in $\sum_{s=0}^{\ell}\binom{n}{s}$ ways, which is the right-hand side of (3.19). Now applying (3.19) to (3.3) yields (3.10) with $M-1$ replaced by 2^k.

We now turn our attention to the computation of the converse bound. Choosing Q_{Y^n} equiprobable on $\{0,1\}^n$, Theorem 34 yields the classical sphere packing bound (cf. [9, (5.8.19)] for an alternative expression).

[4]This bound holds for average probability of error. Fig. 3.1 shows the corresponding bound on maximal error probability where we drop the half of the codewords with worse error probability. This results in an additional term of -1 appended to the right-hand side of (3.17), while $\frac{1}{\epsilon}$ becomes $\frac{2}{\epsilon}$ therein.

Theorem 40 *For the BSC with crossover probability δ, the size of an (n, M, ϵ) code (average error probability) must satisfy*

$$M \leq \frac{1}{\beta^n_{1-\epsilon}}\,, \tag{3.20}$$

where β^n_α is defined as

$$\beta^n_\alpha = (1-\lambda)\beta_L + \lambda\beta_{L+1} \tag{3.21}$$

$$\beta_\ell = \sum_{k=0}^{\ell} \binom{n}{k} 2^{-n}\,, \tag{3.22}$$

where $0 \leq \lambda < 1$ and the integer L are defined by

$$\alpha = (1-\lambda)\alpha_L + \lambda\alpha_{L+1} \tag{3.23}$$

$$\alpha_\ell = \sum_{k=0}^{\ell-1} \binom{n}{k} (1-\delta)^{n-k}\delta^k\,. \tag{3.24}$$

Proof: To streamline notation, we denote $\beta^n_\alpha = \beta_\alpha(x^n, Q_{Y^n})$ since it does not depend on x^n, and Q_{Y^n} is fixed. Then, the Hamming weight $|Y^n|$ is a sufficient statistic for discriminating between $P_{Y^n|X^n=\mathbf{0}}$ and Q_{Y^n}. Thus, the optimal randomized test is (assuming $\delta \leq 1/2$)

$$P_{Z_0|Y^n}(1|y^n) = \begin{cases} 0, & |y^n| > L^n_\alpha\,, \\ \lambda^n_\alpha, & |y^n| = L^n_\alpha\,, \\ 1, & |y^n| < L^n_\alpha\,, \end{cases} \tag{3.25}$$

where $L^n_\alpha \in \mathbb{Z}_+$ and $\lambda^n_\alpha \in [0,1)$ are uniquely determined by the condition

$$\sum_{y^n \in \mathsf{B}} P_{Y^n|X^n}(y^n|\mathbf{0}) P_{Z_0|Y^n}(1|y^n) = \alpha\,. \tag{3.26}$$

Then we find that

$$\beta^n_\alpha = \lambda^n_\alpha \binom{n}{L^n_\alpha} 2^{-n} + \sum_{k=0}^{L^n_\alpha - 1} \binom{n}{k} 2^{-n}\,. \tag{3.27}$$

Thus, by Theorem 30

$$M^*(n,\epsilon) \leq \frac{1}{\beta^n_{1-\epsilon}}\,. \tag{3.28}$$

■

3.2.2 Asymptotic expansion

Theorem 41 *For the BSC with crossover probability δ, such that $\delta \notin \{0, \frac{1}{2}, 1\}$, the capacity C and dispersion V are equal to*

$$C(\delta) = \log 2 - h(\delta)\,, \tag{3.29}$$

$$V(\delta) = \delta(1-\delta)\log^2 \frac{1-\delta}{\delta}\,. \tag{3.30}$$

Moreover, we have

$$\log M^*(n,\epsilon) = nC - \sqrt{nV}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,, \tag{3.31}$$

regardless of whether ϵ is maximal or average probability of error.

Proof: To prove the upper-bound we have

$$\begin{aligned} \log M^*(n,\epsilon) &\leq -\log \beta^n_{1-\epsilon} && (3.32)\\ &\leq nC - \sqrt{nV}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,, && (3.33)\end{aligned}$$

where (3.32) follows by the converse bound (3.20) and (3.33) holds due to Lemma 14.

In order to obtain the right $\log n$ term in the lower (achievability) bound we must use the strongest bound, i.e. the RCU as given by (3.10), because no other bound achieves the right $\log n$ term. First, denote

$$S^k_n \stackrel{\triangle}{=} 2^{-n}\sum_{l=0}^{k}\binom{n}{l}\,. \tag{3.34}$$

Then (3.10) implies the existence of (n, M, ϵ) code (maximal probability of error) with

$$\epsilon \leq \sum_{k=0}^{n}\binom{n}{k}\delta^k(1-\delta)^{n-k}\min\left\{1,\ MS^k_n\right\}\,. \tag{3.35}$$

We are going to argue that (3.35) implies a lower-bound on M^* with a matching $\log n$ term.

Without loss of generality, assume $\delta < 1/2$; choose any $r \in (\delta, 1/2)$ and set

$$\begin{aligned} q &= \frac{r}{1-r} < 1\,, \text{ and}\\ K &= n\delta + \sqrt{n\delta(1-\delta)}Q^{-1}\left(\epsilon - \frac{B+G}{\sqrt{n}}\right)\,, \end{aligned} \tag{3.36}$$

where G is going to be defined below and B denotes as usual the Berry-Esseen constant for a binomial (n,δ) distribution. Then from Berry-Esseen Theorem we obtain

$$\sum_{k>K}\binom{n}{k}\delta^k(1-\delta)^{n-k} \leq \epsilon - \frac{G}{\sqrt{n}}\,. \tag{3.37}$$

It is also clear that for all sufficiently large n we have $K < rn$.

Now, observe the following inequality, valid for $k \in [1, n-1]$ and $j \in [-(n-k), k]$:

$$\binom{n}{k-j} \leq \binom{n}{k}\left(\frac{k}{n-k}\right)^j\,. \tag{3.38}$$

Consider any M such that $MS^K_n \leq 1$, then

$$\begin{aligned} &M\sum_{k=0}^{K}S^k_n = M\sum_{t=0}^{K}(K-t+1)\binom{n}{t}2^{-n} = M\sum_{l=0}^{K}(l+1)\binom{n}{K-l}2^{-n} \leq\\ &MS^K_n\sum_{l=0}^{K}(l+1)\left(\frac{K}{n-K}\right)^l \leq MS^K_n\sum_{l=0}^{K}(l+1)q^l \leq MS^K_n\sum_{l=0}^{\infty}(l+1)q^l \leq \frac{1}{(1-q)^2}\,. \end{aligned} \tag{3.39}$$

On the other hand we are going to prove

$$\sup_{k\in[0,n]} \binom{n}{k}\delta^k(1-\delta)^{n-k} \le \frac{G_1}{\sqrt{n}}\,. \tag{3.40}$$

First, observe that $\frac{1}{\sqrt{n}}$ is a precise estimate of the order and also that this effect is not specific to a binomial distribution but is common to all distributions which are the sums of independent random variables, see [66].

For the reason that a more general proof is given in [66] we will only outline the proof of (3.40). Namely, below we neglect the possibility that δn is not an integer. Take $k_* = \delta n$ and note that for any $k = k_* - j$ (j might be negative) we have by (3.38) that

$$\begin{aligned}\binom{n}{k}\delta^k(1-\delta)^{n-k} &\le \binom{n}{k^*}\left(\frac{\delta n}{n-\delta n}\right)^j \delta^{k_*-j}(1-\delta)^{n-k^*+j} &(3.41)\\ &\le \binom{n}{k^*}\delta^{k_*}(1-\delta)^{n-k_*}\,. &(3.42)\end{aligned}$$

Thus, the maximum in (3.40) is achieved at k_*. The rest is Stirling's approximation:

$$\sqrt{2\pi n}e^{n\ln n-n} < n! < e^{1/12}\cdot\sqrt{2\pi n}e^{n\ln n-n}\,. \tag{3.43}$$

Now set

$$G = \frac{G_1}{(1-q)^2} \tag{3.44}$$

and observe that if $MS_n^K \le 1$ then by (3.39)

$$\sum_{k=0}^{K}\binom{n}{k}\delta^k(1-\delta)^{n-k}MS_n^k \le \frac{G}{\sqrt{n}}\,. \tag{3.45}$$

We can now see that (3.35) implies that

$$M^*(n,\epsilon) \ge \frac{1}{S_n^K}\,. \tag{3.46}$$

Indeed, pick $M = \frac{1}{S_n^K}$. Then from (3.37) and (3.45) it follows that

$$\begin{aligned}&\sum_{k=0}^{n}\binom{n}{k}\delta^k(1-\delta)^{n-k}\min\left\{1,\ MS_n^k\right\} &(3.47)\\ &\le \sum_{k=0}^{K}\binom{n}{k}\delta^k(1-\delta)^{n-k} + \sum_{k>K}\binom{n}{k}\delta^k(1-\delta)^{n-k} &(3.48)\\ &\le \frac{G}{\sqrt{n}} + \epsilon - \frac{G}{\sqrt{n}} &(3.49)\\ &\le \epsilon\,. &(3.50)\end{aligned}$$

Finally, we must upper-bound $\log S_n^K$ up to $O(1)$ terms. This is just an application of (3.38):

$$\begin{aligned} S_n^K &= 2^{-n}\sum_{k=0}^{K}\binom{n}{k} && (3.51)\\ &\le 2^{-n}\binom{n}{K}\sum_{l=0}^{\infty}\left(\frac{K}{n-K}\right)^l && (3.52)\\ &\le 2^{-n}\binom{n}{K}\frac{n-K}{n-2K}\,. && (3.53)\end{aligned}$$

For n sufficiently large $n-2K$ will become larger than $n(1-2r)$, thus for such n we have $\frac{n-K}{n-2K}\le\frac{1}{1-2r}$ and hence

$$\log S_n^K \le -n\log 2+\log\binom{n}{K}+O(1)\,. \tag{3.54}$$

Using (3.43) we obtain the inequality

$$\binom{n}{K}\le\frac{e^{1/12}}{\sqrt{2\pi}}\sqrt{\frac{n}{K(n-K)}}\exp(nh(K/n))\,. \tag{3.55}$$

Plugging K from (3.36) and applying Taylor's formula to $h(p)$ implies

$$\log S_n^K \le n(h(\delta)-\log 2)+\sqrt{n\delta(1-\delta)}\log\frac{1-\delta}{\delta}Q^{-1}\left(\epsilon-\frac{B+G}{\sqrt{n}}\right)-\frac{1}{2}\log n+O(1)\,. \tag{3.56}$$

Finally, applying Taylor's formula to Q^{-1}, we conclude

$$\log S_n^K \le n(h(\delta)-\log 2)+\sqrt{n\delta(1-\delta)}\log\frac{1-\delta}{\delta}Q^{-1}(\epsilon)-\frac{1}{2}\log n+O(1)\,. \tag{3.57}$$

Substituting this into (3.46) we obtain the required expansion. ■

3.2.3 Numerical evaluation

The numerical evaluation of the RCU bound (3.10), the DT bound (3.11), Feinstein's bound (3.16), Gallager's bound (3.17) and the converse bound (3.20) is shown in Figs. 3.1 and 3.2. As we anticipated analytically, the DT bound is always tighter than Feinstein's bound. For $\delta=0.11$ and $\epsilon=0.001$, we can see in Fig. 3.1 that for blocklengths greater than $n_*\approx 150$, Theorem 18 gives better results than Gallager's bound. In fact, for large n the gap to the converse upper bound of the new lower bound is less than half that of Gallager's bound. Finally, the RCU bound (3.10) is uniformly better than all other bounds. In fact for all $n\ge 20$ the difference between (3.10) and the converse is within $3-4$ bits in $\log M$. This tendency remains for other choices of δ and ϵ. Although, as Fig. 3.2 shows, for smaller ϵ and/or δ, Gallager's bound (designed to analyze the regime of exponentially small ϵ) performs better (i.e., the value of n_* is greater). A similar relationship between the two bounds holds, qualitatively, in the case of the AWGN channel, see Section 4.4.

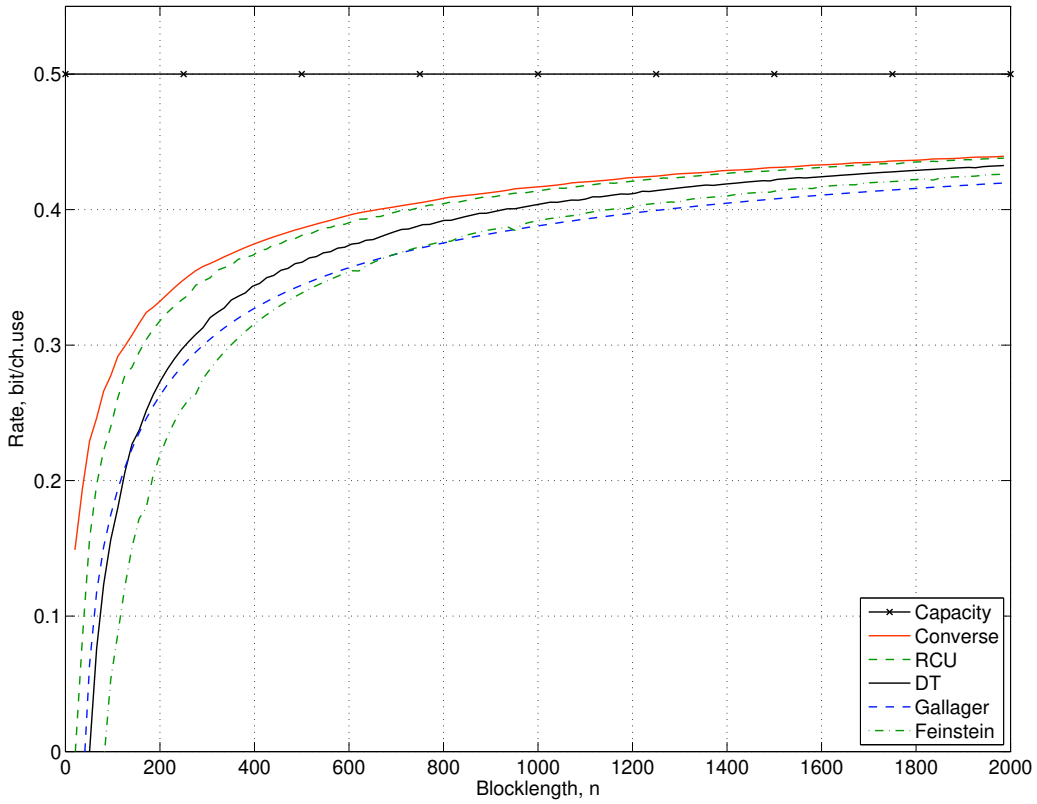


Figure 3.1: Rate-blocklength tradeoff for the BSC with crossover probability $\delta = 0.11$ and maximal block error rate $\epsilon = 10^{-3}$: comparison of the bounds.

By Theorem 41, we have

$$\log M^*(n,\epsilon) = nC - \sqrt{nV}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,, \tag{3.58}$$

where C and V are the capacity and dispersion of the BSC. Interestingly, Gallager's bound only yields the bound $nC + O(\sqrt{n})$ with a suboptimal $\sqrt{n}$ term; both Feinstein and the DT bound (Theorem 18) yield the correct $\sqrt{n}$ term, but Feinstein's bound is worse in terms of the $\log n$ term. Finally, only the RCU bound (Theorem 17) achieves the correct $\log n$ term. So we can see that asymptotic analysis of the bounds correctly predicts their relative merit observed in numerical computations on Fig. 3.1 and 3.2.

The primary use of Theorem 41 for non-asymptotic analysis is in obtaining the (refined) normal approximation:

$$\log M^*(n,\epsilon) \approx n(\log 2 - h(\delta)) - \sqrt{n\delta(1-\delta)}Q^{-1}(\epsilon)\log\frac{1-\delta}{\delta} + \frac{1}{2}\log n\,. \tag{3.59}$$

In Figs. 3.3 and 3.4 we compare the normal approximation (3.59) and the best of the upper and lower bounds, computed above. We make two conclusions from the plot: 1) knowledge of the $\log n$ term improves the precision of the general normal approximation (2.23);

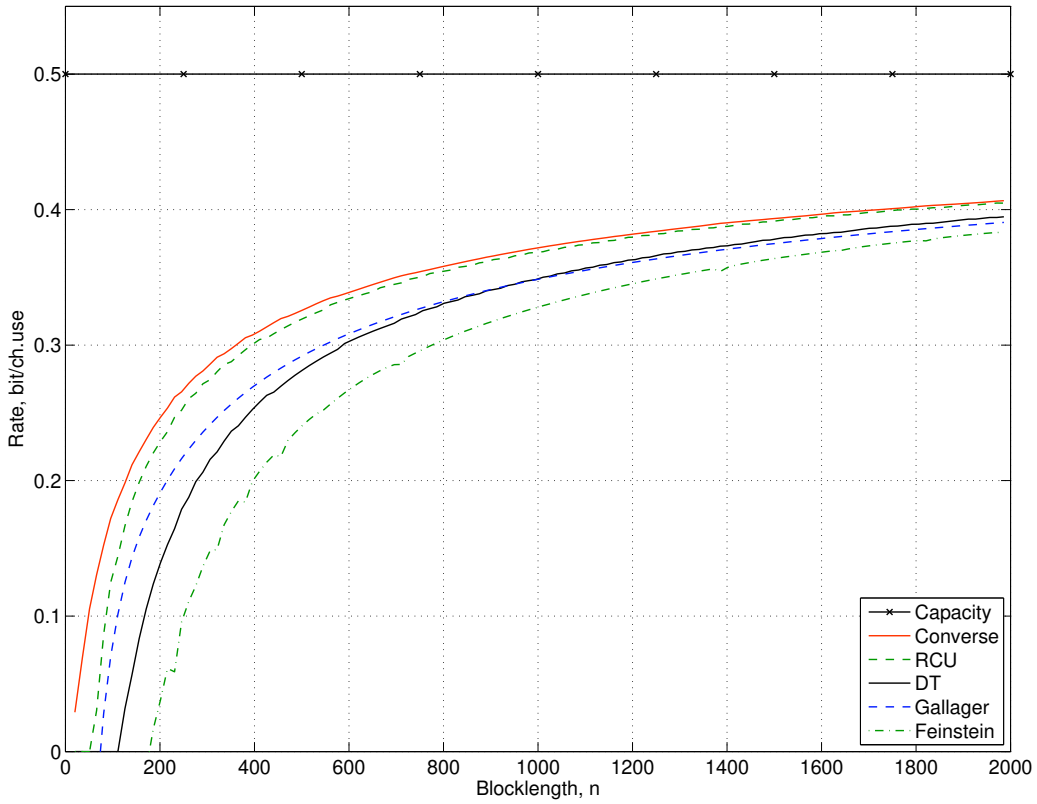


Figure 3.2: Rate-blocklength tradeoff for the BSC with crossover probability $\delta = 0.11$ and maximal block error rate $\epsilon = 10^{-6}$: comparison of the bounds.

2) although slightly pessimistic, expression (3.59) can serve as an excellent substitute for complex computations of the bounds (3.10) and (3.20).

3.3 Binary erasure channel (BEC)

This section illustrates the theory developed so far as applied to the BEC. Recall that $BEC(n, \delta)$ for blocklength n and erasure probability δ is defined as follows: the input alphabet $\mathsf{A} = \{0, 1\}^n$, the output alphabet $\mathsf{B} = \{0, e, 1\}^n$, and the channel acts as

$$\mathbb{P}_{Y^n|X^n}(y^n|x^n) = \begin{cases} \left(\frac{\delta}{1-\delta}\right)^{\#\{y_j=e\}} (1-\delta)^n, & \text{if } y^n \text{ and } x^n \text{ agree on unerased positions}\,, \\ 0, & \text{otherwise}\,. \end{cases} \tag{3.60}$$

3.3.1 Bounds

Choosing P_X in Theorems 17 and 18 to be equiprobable on the input alphabet we obtain:

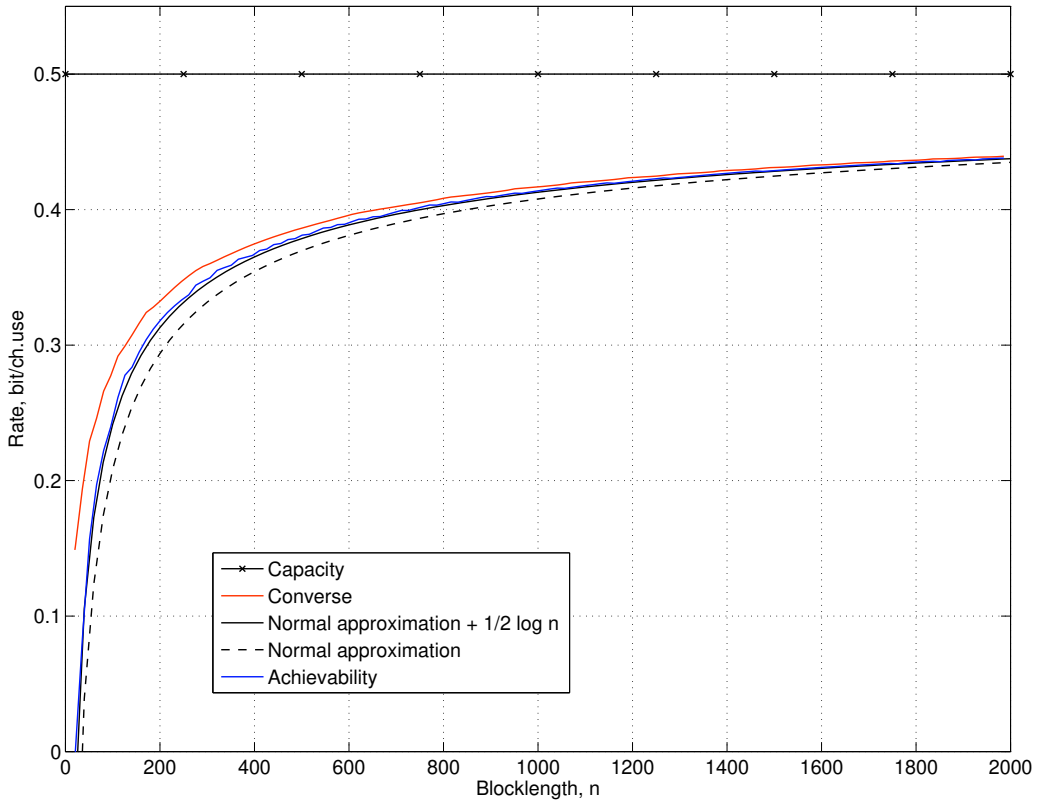


Figure 3.3: Rate-blocklength tradeoff for the BSC with crossover probability $\delta = 0.11$ and maximal block error rate $\epsilon = 10^{-3}$: normal approximation.

Corollary 42 *For the BEC with erasure probability δ, there exists an (n, M, ϵ) code (average probability of error) such that (RCU bound)*

$$\epsilon \leq \sum_{t=0}^{n} \binom{n}{t} \delta^t (1-\delta)^{n-t} 2^{-[n-t-\log_2(M-1)]^+} . \tag{3.61}$$

and (DT bound)

$$\epsilon \leq \sum_{t=0}^{n} \binom{n}{t} \delta^t (1-\delta)^{n-t} 2^{-[n-1-t-\log_2(M-1)]^+} . \tag{3.62}$$

Whenever $M = 2^k$ for integer k, the statement holds for maximal probability of error as well.

Proof: The proof follows once we notice that with equiprobable input distribution it follows that the information density equals

$$i(x^n; y^n) = \begin{cases} \#\{j : y_j \neq e\} \cdot \log 2 , & \text{if } y^n \text{ and } x^n \text{ agree on unerased positions} , \\ -\infty, & \text{otherwise} . \end{cases} \tag{3.63}$$

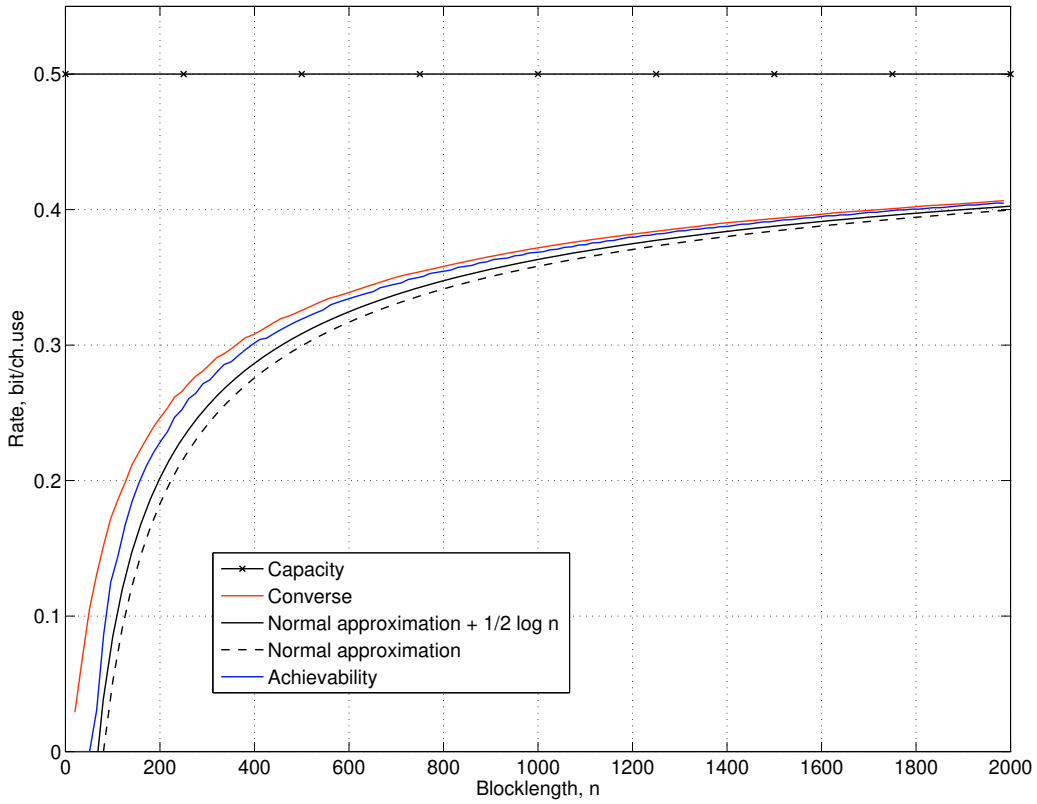


Figure 3.4: Rate-blocklength tradeoff for the BSC with crossover probability $\delta = 0.11$ and maximal block error rate $\epsilon = 10^{-6}$: normal approximation.

Since the number of erasures is distributed binomially with parameters n and δ, results follow from Theorems 17 and 18 at once.

The statement on the maximal probability of error is explained in Appendix C. ■

A number of remarks are in order:

1. For the BEC, the DT bound (3.62) is obviously strictly stronger than the RCU bound (3.61). Since the RCU bound is always stronger than Gallager's bound, we see that the DT bound in this case dominates both the RCU and Gallager's bound, and thus achieves the random-coding error-exponent.

2. With equiprobable P_{X^n} the generalization of the DT bound to the maximal probability of error, Theorem 23, yields the following upper bound on maximal error probability, see (2.178):

$$\epsilon \leq \mathbb{E}\left[2^{-(Z-\log(M-1))^+}\right], \tag{3.64}$$

where Z is binomial with parameters n and $1-\delta$, $Z \sim B(n, 1-\delta)$. Notice that this expression coincides exactly with the RCU bound (3.61), except that (3.64) upper-bounds maximal probability of error. Since Theorem 23 is stronger than Feinstein's

bound, we can immediately conclude that, for the BEC, the DT bound (3.62) applied with $M = 2^k$ dominates the RCU, Feinstein and Gallager bounds for all blocklengths and rates.

3. The average block erasure probability for a random ensemble of all $[k, n]$ linear codes is given in [25] as follows:

$$P_{RLC}(2^k) = \sum_{u=k}^{n} \binom{n}{u} \delta^{n-u}(1-\delta)^u \left[1 - \prod_{a=0}^{k-1}(1-2^{a-u})\right] + \sum_{u=0}^{k-1} \binom{n}{t} \delta^t (1-\delta)^{n-t} \,. \quad (3.65)$$

If we denote the right-hand side of (3.61) by $P_{RCU}(M)$ and the right-hand side of (3.62) by $P_{DT}(M)$ then it follows that

$$P_{DT}(2^k) \le P_{RLC}(2^k) \le P_{RCU}(2^k) \,. \quad (3.66)$$

Indeed, the left-hand inequality in (3.66) follows by applying

$$\left[1 - \prod_{a=0}^{k-1}(1-2^{a-u})\right] \ge 1 - (1-2^{k-1-u}) \quad (3.67)$$

to (3.65). The right-hand inequality in (3.66) follows from (3.65) by applying to it the following inequality:

$$\left[1 - \prod_{a=0}^{k-1}(1-2^{a-u})\right] \le 1 - (1-2^{k-u}) \,, \quad (3.68)$$

which is a consequence of $\prod_j (1-b_j x) \ge 1 - \sum_j b_j x$ (for $b_j \ge 0$ and $x \ge 0$).

In this way, (3.66) demonstrates that the DT bound (3.62) results in a tighter bound on probability of error compared to the bounds in [25]. However, the additional advantage of P_{RLC} and P_{RCU} is that they upper-bound the block *erasure* probability, and thus for those codes all errors are detected.

4. For a random ensemble of (non-linear) codebooks of size M and blocklength n, the average block erasure probability can be easily computed in closed form:

$$P_{RC}(M) = \sum_{u=0}^{n} \binom{n}{u} \delta^{n-u}(1-\delta)^u \left[1 - (1-2^{-u})^{M-1}\right] \,. \quad (3.69)$$

To compare $P_{RLC}(2^k)$ and $P_{RC}(2^k)$ we apply

$$1 - (1-2^{-t})^M < 1 - \prod_j (1 - M_j 2^{-t}) \,, \quad (3.70)$$

which holds as long as $\sum M_j = M$, to (3.69) to obtain:

$$P_{RC}(2^k) < P_{RLC}(2^k) \quad (3.71)$$

and, therefore, in this case *random coding over all codebooks is strictly better than random coding over linear codebooks only.*

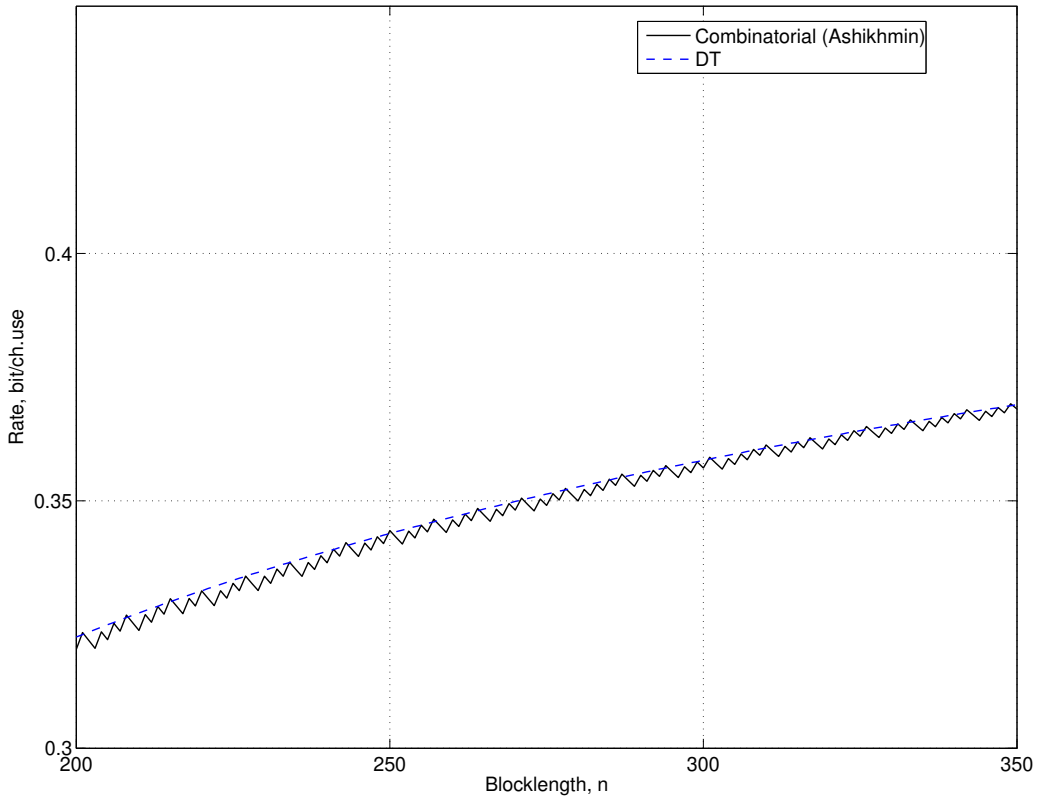


Figure 3.5: Comparison of the DT-bound (3.62) and the combinatorial bound of Ashikhmin (3.4) for the BEC with erasure probability $\delta = 0.5$ and probability of block error $\epsilon = 10^{-3}$.

5. Finally, a comparison with the BEC-specific bound of Ashikhmin (3.4) is given on Fig. 3.5. The bounds are within one bit of each other, the winner depending on a particular value of n. The zigzagging of the plot of (3.4) is a behavior common to all bounds that are restricted to integer values of $\log_2 M$. Again, computation-wise the DT bound is much more preferable: the complexity of (3.4) is $O(n^3)$, compared to $O(n)$ for the DT bound (3.62). Analytical comparison of the bound (3.4) and the DT bound is complicated, since the random ensembles are different (in particular, the random ensemble in Ashikhmin's bound does not contain codebooks with repeated codewords).

The upper bound on code size given by Theorem 34 (with capacity achieving output distribution) is improved by the following theorem, which is also stronger than a related bound, sometimes called a singleton bound, such as in [5].

Similar to Section 3.2.1, upper bounds on probability of error can be converted to lower bounds on $\log M^*(n, \epsilon)$; see (3.15).

Theorem 43 *For the BEC with erasure probability δ, the average error probability of an*

(n, M, ϵ) code satisfies

$$\epsilon \geq \sum_{\ell=\lfloor n-\log_2 M\rfloor+1}^{n} \binom{n}{\ell} \delta^\ell (1-\delta)^{n-\ell} \left(1 - \frac{2^{n-\ell}}{M}\right), \tag{3.72}$$

even if the encoder knows the location of the erasures non-causally[5] *.*

Proof: The main idea is to apply the following, self-evident result: *Suppose that X, Y and $\{error\}$ are, as usual, the input codeword, the channel output and the error event. Denote by Z a random variable (possibly dependent on X and Y). If there exists a function $\lambda(z)$ such that for any $z \in \mathsf{Z}$*

$$\mathbb{P}[error|Z = z] \geq \lambda(z), \tag{3.73}$$

then we have

$$\epsilon \geq \sum_{z\in\mathsf{Z}} \lambda(z)\mathbb{P}[Z = z]. \tag{3.74}$$

We define Z to be the number of erasures in the output Y^n. Then, Z takes values from 0 to n. The conditional channel $P_{Y|X}^{(z)}$ is then simply a noiseless channel of $n - z$ channel uses (erasures essentially only decrease blocklength). Thus, the following bound holds for the code with M codewords even if it knows the locations of the erasures non-causally:

$$P[\text{error}|Z = z] \geq \left(1 - \frac{2^{n-z}}{M}\right)^+. \tag{3.75}$$

Indeed, when $2^{n-z} > M$ the bound is obvious. Otherwise, there are $M - 2^{n-z}$ messages whose conditional probability of error is 1.

Finally, we do not need to find a lower-bound measure μ because we can calculate $P[Z = z|X = x]$ exactly for any x:

$$P[Z = z|X = x] = \binom{n}{z} \delta^z (1-\delta)^{n-z}. \tag{3.76}$$

The application of (3.74) yields the result. ■

3.3.2 Asymptotic expansion

Theorem 44 *For the BEC with erasure probability δ, we have*

$$\log M^*(n, \epsilon) = n(1-\delta)\log 2 - \sqrt{n\delta(1-\delta)}Q^{-1}(\epsilon)\log 2 + O(1), \tag{3.77}$$

regardless of whether ϵ is maximal or average probability of error.

[5]The same result holds for a q-ary erasure channel with 2^{n-l-k} replaced by q^{n-l-k}; also such a q-ary extension is achievable by q-ary maximum distance separable, MDS, codes.

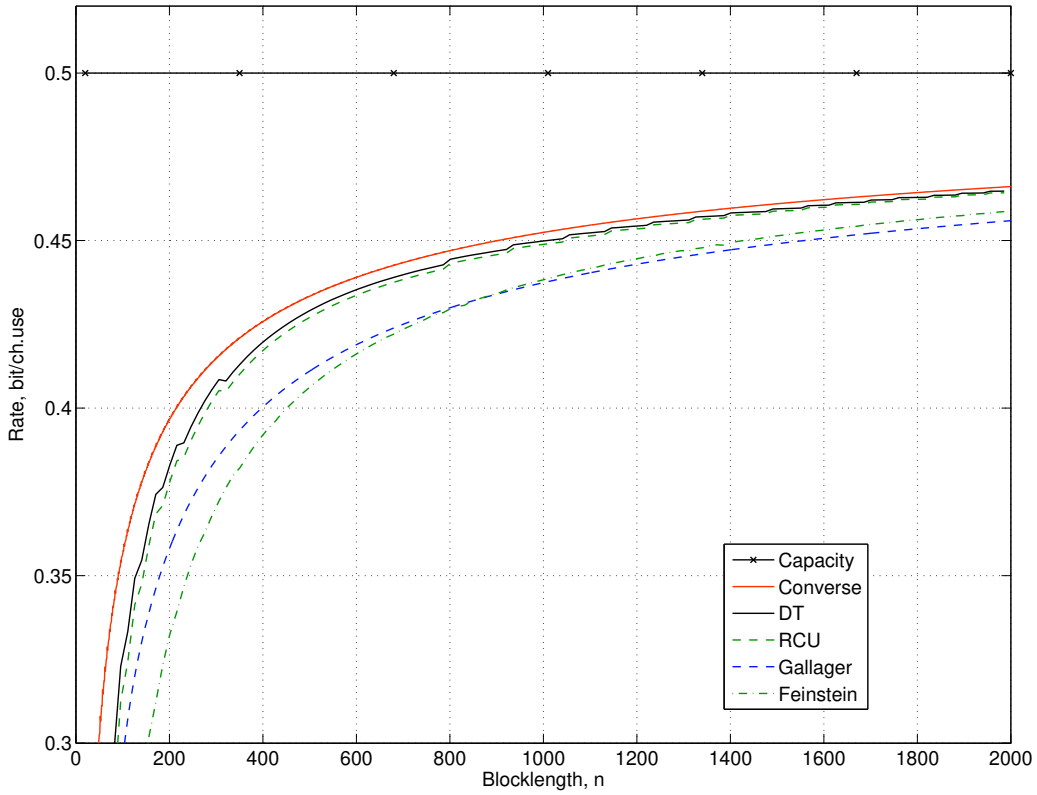


Figure 3.6: Rate-blocklength tradeoff for the BEC with erasure probability $\delta = 0.5$ and maximal block error rate $\epsilon = 10^{-3}$: comparison of the bounds.

Proof: Any (n, M, ϵ) code (average probability of error) must satisfy (3.72). Thus we must simply find M so large that the left-hand side of (3.72) is larger than a given ϵ. We can then conclude that $M^*(n, \epsilon)$ is upper-bounded by such M.

First, we observe that by (3.40)

$$\sum_{\ell=\lfloor n-\log M\rfloor+1}^{n} \binom{n}{\ell} \delta^{\ell}(1-\delta)^{n-\ell} 2^{n-\ell-\log M} \leq \frac{2G_1}{\sqrt{n}}. \tag{3.78}$$

Then, denote by B the usual Berry-Esseen constant for a binomial distribution and set

$$\log M = n(1-\delta) - \sqrt{n\delta(1-\delta)}Q^{-1}\left(\epsilon + \frac{B+2G_1}{\sqrt{n}}\right). \tag{3.79}$$

Then from the Berry-Esseen Theorem we obtain

$$\sum_{l\geq n-\log M} \binom{n}{l} \delta^{l}(1-\delta)^{n-l} \geq \epsilon + \frac{2G_1}{\sqrt{n}}. \tag{3.80}$$

Finally from (3.78) we conclude that

$$\sum_{\ell=\lfloor n-\log M\rfloor+1}^{n} \binom{n}{\ell} \delta^{\ell}(1-\delta)^{n-\ell}\left(1 - 2^{n-\ell-\log M}\right) \geq \epsilon \tag{3.81}$$

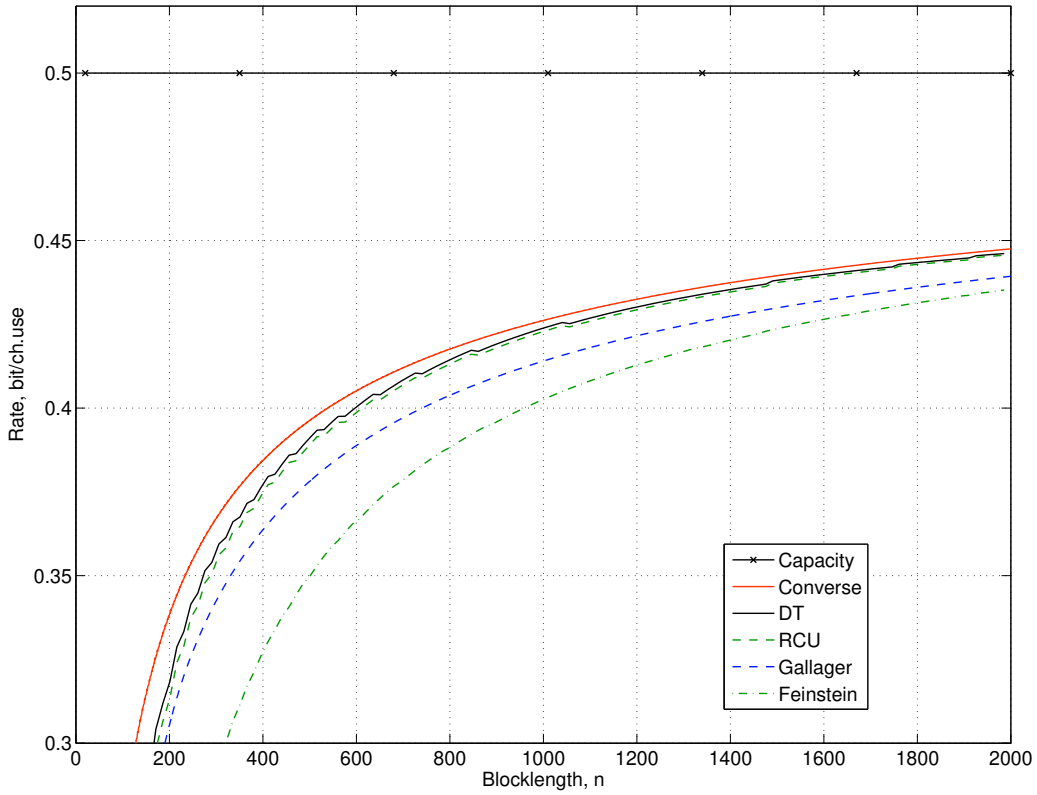


Figure 3.7: Rate-blocklength tradeoff for the BEC with erasure probability $\delta = 0.5$ and maximal block error rate $\epsilon = 10^{-6}$: comparison of the bounds.

and hence

$$\log M^*(n,\epsilon) \le n(1-\delta) - \sqrt{n\delta(1-\delta)}Q^{-1}\left(\epsilon + \frac{B+2G_1}{\sqrt{n}}\right) = n(1-\delta) - \sqrt{n\delta(1-\delta)}Q^{-1}(\epsilon) + O(1)\,, \quad (3.82)$$

where the last step is by Taylor's formula.

For the achievability part we use (3.64), which gives the bound on the maximal probability of error. We rewrite the left-hand side (3.64) as follows:

$$\sum_{k>\log M}\binom{n}{k}(1-\delta)^k\delta^{n-k}M2^{-k} + \sum_{k\le\log M}\binom{n}{k}(1-\delta)^k\delta^{n-k}\,. \quad (3.83)$$

Again, by using (3.40) we note that the first term is less than $\frac{2G_1}{\sqrt{n}}$; then by setting

$$\log M = n(1-\delta) - \sqrt{n\delta(1-\delta)}Q^{-1}\left(\epsilon - \frac{B+2G_1}{\sqrt{n}}\right) \quad (3.84)$$

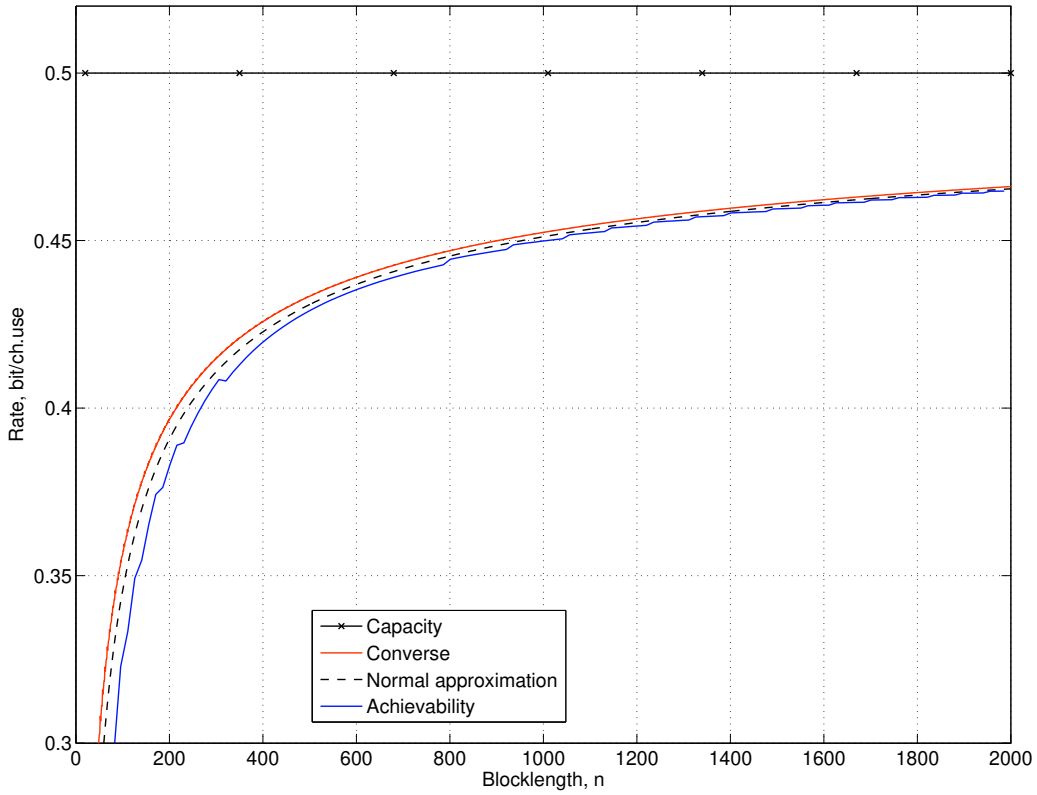


Figure 3.8: Rate-blocklength tradeoff for the BEC with erasure probability $\delta = 0.5$ and maximal block error rate $\epsilon = 10^{-3}$: normal approximation.

we make the entire expression (3.83), by Berry-Esseen, smaller than ϵ. In this way, we have established that

$$\log M^*(n,\epsilon) \geq n(1-\delta) - \sqrt{n\delta(1-\delta)}Q^{-1}\left(\epsilon - \frac{B+2G_1}{\sqrt{n}}\right). \tag{3.85}$$

After applying Taylor's formula to Q^{-1} we prove the theorem. ■

3.3.3 Numerical comparison

The numerical evaluation of the RCU bound (3.61), the DT bound (3.62), Feinstein's bound, Gallager's bound and the converse bound (3.72) is shown in Figs. 3.6 and 3.7. As discussed previously, the DT bound uniformly beats all other bounds.

According to Theorem 44, as $n \to \infty$ the fundamental limit $\log M^*(n,\epsilon)$ behaves as

$$\log_2 M^*(n,\epsilon) = n(1-\delta) - \sqrt{n\delta(1-\delta)}Q^{-1}(\epsilon) + O(1). \tag{3.86}$$

Similarly to the BSC, for the BEC Gallager's bound does not give a correct lower-order term at all; Feinstein's bound yields the correct $\sqrt{n}$ term but a suboptimal $\log n$ term; both the RCU (3.61) and the DT (3.62) bounds achieve an optimal $\log n$ term.

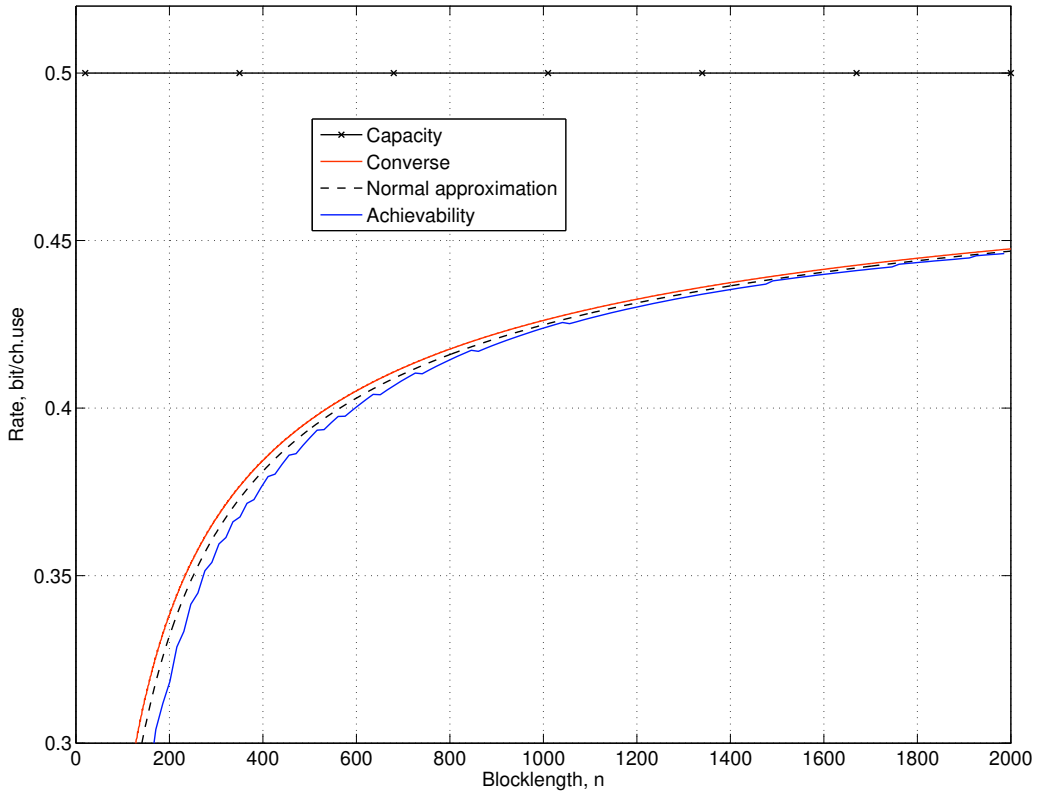


Figure 3.9: Rate-blocklength tradeoff for the BEC with erasure probability $\delta = 0.5$ and maximal block error rate $\epsilon = 10^{-6}$: normal approximation.

The value of Theorem 44 for non-asymptotic analysis is in providing an estimate of the fundamental limit at finite n (normal approximation):

$$\log_2 M^*(n,\epsilon) \approx n(1-\delta) - \sqrt{n\delta(1-\delta)}Q^{-1}(\epsilon)\,. \tag{3.87}$$

The comparison of this approximation with the sharp bounds discussed above is given in Figs 3.8 and 3.9. Again, we notice a remarkable precision of the simple formula (3.87).

3.4 General discrete memoryless channel (DMC)

The DMC has finite input alphabet $\mathcal{A}$, finite output alphabet $\mathcal{B}$, and the conditional probabilities are defined as

$$P_{Y^n|X^n}(y^n|x^n) = \prod_{i=1}^{n} W(y_i|x_i)\,, \tag{3.88}$$

where $W(y|x)$ is a conditional probability mass function from $\mathcal{A}$ to $\mathcal{B}$; for convenience we denote

$$W_x \stackrel{\triangle}{=} W(\cdot|x)\,. \tag{3.89}$$

The following functions describe the fundamental limits of the DMC:

$$M^*(n,\epsilon) = \max\{M : \exists \ (n,M,\epsilon)\text{-code (maximal probability of error)}\}, \quad (3.90)$$
$$M^*_{avg}(n,\epsilon) = \max\{M : \exists \ (n,M,\epsilon)\text{-code (average probability of error)}\}. \quad (3.91)$$

Note that the input probability distributions are elements of $\mathbb{R}^{|\mathcal{A}|}$ constrained to an $(|\mathcal{A}|-1)$-dimensional simplex. We denote this simplex by $\mathcal{P}$. In this way $\mathcal{P}$ is a compact metric space. We also emphasize its subsets ("n-types") indexed by $n = 1, \ldots$:

$$\mathcal{P}_n \stackrel{\triangle}{=} \{P \in \mathcal{P} : \ nP(x) \in \mathbb{Z}_+ \ \forall x \in \mathcal{A}\}. \quad (3.92)$$

For each fixed $P \in \mathcal{P}$ define:

- *output distribution* PW as $PW(y) = \sum_x P(x)W(y|x)$.
- for an arbitrary $Q \in \mathcal{P}$ s.t. $P \ll Q$ recall from (2.5), that the divergence variance is given by
$$V(P||Q) = \sum_x P(x)\left[\log\frac{P(x)}{Q(x)}\right]^2 - D(P||Q)^2. \quad (3.93)$$
- for an arbitrary Q_Y on $\mathcal{B}$ define *the conditional divergence variance* as
$$V(W||Q_Y|P) = \sum_x P(x)V(W_x||Q_Y). \quad (3.94)$$
Note that $V(W||Q_Y|P)$ is defined only provided that $W_x \ll Q_Y$ for P-almost all x.
- *mutual information*, $I(P,W) = \mathbb{E}[i(X;Y)]$, or
$$I(P,W) = \sum_{x,y} P(x)W(y|x)\log\frac{W(y|x)}{PW(y)}. \quad (3.95)$$
It is known that $I(P,W)$ is continuous on $\mathcal{P}$ so that
$$C = \max_{P\in\mathcal{P}} I(P,W) \quad (3.96)$$
is a well-defined, finite quantity.
- *unconditional information variance* $U(P,W) = \mathrm{Var}(i(X;Y))$, or
$$U(P,W) = \sum_{x,y} P(x)W(y|x)\left[\log\frac{W(y|x)}{PW(y)}\right]^2 - [I(P,W)]^2 \quad (3.97)$$
$$= V(P\times W||P\times PW). \quad (3.98)$$
- *conditional information variance* $V(P,W) = \mathbb{E}[\mathrm{Var}(i(X;Y)\,|\,X)]$, or
$$V(P,W) = \sum_x P(x)\left\{\sum_y W(y|x)\left[\log\frac{W(y|x)}{PW(y)}\right]^2 - [D(W_x||PW)]^2\right\} \quad (3.99)$$
$$= V(W||PW|P) \quad (3.100)$$

Since in general

$$\sum_x P(x)[D(W_x||PW)]^2 \neq [I(P,W)]^2 \tag{3.101}$$

values of $U(P,W)$ and $V(P,W)$ do not necessarily coincide. For example, take a binary-input binary-output noiseless channel and $P = \left[\frac{1}{4}, \frac{3}{4}\right]^T$ then $V(P,W) = 0$ while $U(P,W) \approx 0.47$ bit^2. We always have

$$V(P,W) \leq U(P,W)\,, \tag{3.102}$$

with the equality if and only if

$$D(W_x||PW) = I(P,W) \qquad \text{for } P\text{-almost all } x\,. \tag{3.103}$$

- *third absolute moment of the information density*

$$T(P,W) = \sum_{x,y} P(x)W(y|x) \left| \log \frac{W(y|x)}{PW(y)} - D(W_x||PW) \right|^3 . \tag{3.104}$$

- *third unconditional absolute moment of the information density*

$$T_u(P,W) = \sum_{x,y} P(x)W(y|x) \left| \log \frac{W(y|x)}{PW(y)} - I(P,W) \right|^3 . \tag{3.105}$$

- a subset of *capacity achieving distributions* Π by

$$\Pi \stackrel{\triangle}{=} \{P \in \mathcal{P} :\ I(P,W) = C\}\,. \tag{3.106}$$

  Note that $\Pi = I^{-1}(C)$ is a compact subset of $\mathcal{P}$.

- maximal and minimal conditional variances as

$$V_{max} = \max_{P\in\Pi} V(P,W) = \max_{P\in\Pi} U(P,W)\,, \tag{3.107}$$

$$V_{min} = \min_{P\in\Pi} V(P,W) = \min_{P\in\Pi} U(P,W)\,. \tag{3.108}$$

  The reason for writing max and min instead of sup and inf, as well as the right-most equalities in (3.107) and (3.108) is to be explained shortly (see Lemma 46). Note that for the purpose of defining V_{max} and V_{min} both quantities $V(P,W)$ and $U(P,W)$ are equivalent. We introduce both since one appears naturally in the achievability bound and the other in the converse.

- Define the (unique) *capacity achieving output distribution* P_Y^* by $P_Y^* = P^*W$, where P^* is any capacity achieving input distribution.

- W is an *exotic DMC* if $V_{max} = 0$ and there exists an input letter x_0 such that: a) for any capacity achieving P: $P(x_0) = 0$, b) $D(W_{x_0}||P_Y^*) = C$, and c) $V(W_{x_0}||P_Y^*) > 0$.

As usual, we take $0 \log 0 = 0$, $0 \log^2 0 = 0$ and $0 \cdot D(W_x||PW) = 0$ (note that $D(W_x||PW)$ maybe infinite).

The purpose of this Section is to give a proof of the following result:[6]

Theorem 45 *The ϵ-dispersion of the DMC W is*

$$V_\epsilon = \begin{cases} V_{min}\,, & \epsilon < 1/2\,, \\ V_{max}\,, & \epsilon > 1/2\,. \end{cases} \tag{3.109}$$

More specifically, as $n \to \infty$ we have

$$\log M^*(n,\epsilon) = nC - \sqrt{nV_\epsilon} Q^{-1}(\epsilon) + O(\log n)\,, \tag{3.110}$$

$$\log M^*_{avg}(n,\epsilon) = nC - \sqrt{nV_\epsilon} Q^{-1}(\epsilon) + O(\log n)\,, \tag{3.111}$$

unless the DMC is exotic and $\epsilon > 1/2$. In any case, we have[7]

$$\log M^*(n,\epsilon) \geq nC - \sqrt{nV_\epsilon} Q^{-1}(\epsilon) + O(1)\,. \tag{3.112}$$

For the exotic DMC and $\epsilon > 1/2$ we have

$$\log M^*(n,\epsilon) = nC + O\left(n^{\frac{1}{3}}\right)\,, \tag{3.113}$$

and the estimate of the order $n^{\frac{1}{3}}$ cannot be improved in general.

3.4.1 Comparison to Strassen [1]

Strassen [1] claims the validity of (3.110), (3.111) and (3.112) for all DMCs. For example, in [1, (1.15)] Strassen states that for any DMC and n sufficiently large we have

$$\log M^*(n,\epsilon) \leq nC - \sqrt{nV_\epsilon} + |\mathcal{B}| \log n\,. \tag{3.114}$$

This cannot be true for $\epsilon > 1/2$ as (3.113) and the counter-example in Theorem 51 show. So the converse part for $\epsilon > 1/2$ in [1] is flawed. For the case of $\epsilon < 1/2$ the result in Theorem 45 coincides with Strassen's result. In the rest of the Section we use the bounds we derived in Chapter 2 to give an alternative proof that clearly demonstrates that the expansions up to $o(\sqrt{n})$ are significantly easier to obtain than $O(\log n)$ expansions (the difference being that of using Lemma 49 instead of Lemma 48; see discussion below). In particular, with the approach of restricting to constant composition subcodes, the $O(\log n)$ converse results are impossible to obtain without precise analysis of the second-order derivatives (Hessian) of the mutual information function (see the counter-example after Lemma 48). Note that the sufficient condition given in [32, Theorem 49] is in fact unnecessary according to Strassen [1]. Also, Strassen does indeed allude to the average error probability analysis in the discussion on p. 31, contrary to what is claimed in [32, p. 2332].

[6] Recall the Definition 8 for ϵ-dispersion.

[7] This estimate of the $\log n$ term cannot be improved without additional assumptions, because the BEC has zero $\log n$ term; see Theorem 44.

3.4.2 Achievability bound

We start by showing some properties of $U(P,W)$, $V(P,W)$ and $T(P,W)$:

Lemma 46 *The functions $U(P,W)$, $V(P,W)$, $T(P,W)$ and $T_u(P,W)$ are continuous on $\mathcal{P}$. Functions $U(P,W)$ and $V(P,W)$ coincide on Π.*

Note that Lemma 46 justifies taking min and max in (3.108) and (3.107), as well as the right-most equalities therein.

Proof: First, note that $U(P,W)$, $V(P,W)$ and $T(P,W)$ are well-defined and finite. Indeed, each one is a sum of finitely many terms. We must show that every term is well-defined. This is true since, whenever $W(y|x) = 0$ or $PW(y) = 0$ or $P(x) = 0$, we have $P(x)W(y|x) = 0$ and thus

$$P(x)W(y|x)\left[\log\frac{W(y|x)}{PW(y)}\right]^2 \tag{3.115}$$

and

$$P(x)W(y|x)\left|\log\frac{W(y|x)}{PW(y)} - D(W_x||PW)\right|^3 \tag{3.116}$$

are both equal to zero by convention. On the other hand, if $P(x) > 0$ then $W_x \ll PW$ and thus $D(W_x||PW)$ is a well-defined finite quantity.

Second, take a sequence $P_n \to P$. Then we want to prove that each term in $U(P,W)$ is continuous. In other words

$$P_n(x)W(y|x)\left[\log\frac{W(y|x)}{P_nW(y)}\right]^2 \to P(x)W(y|x)\left[\log\frac{W(y|x)}{PW(y)}\right]^2 . \tag{3.117}$$

If $W(y|x) = 0$ then this is obvious. If $P_n(x) \not\to 0$ then this is also true since the argument of the logarithm is bounded away from 0 and $+\infty$. So, we assume $P_n(x) \to 0$ and we must show that then the complete quantity also tends to 0. For $P_n(x) > 0$ we notice that

$$\log\{P_n(x)W(y|x)\} \le \log P_nW(y) \le 0\,. \tag{3.118}$$

Thus,

$$|\log W(y|x) - \log P_nW(y)|^2 \le 2(\log^2 W(y|x) + \log^2\{P_n(x)W(y|x)\})\,. \tag{3.119}$$

But then,

$$0 \le P_n(x)W(y|x)\left[\log\frac{W(y|x)}{P_nW(y)}\right]^2 \le 2P_n(x)W(y|x)(\log^2 W(y|x) + \log^2\{P_n(x)W(y|x)\})\,. \tag{3.120}$$

This is also true for $P_n(x) = 0$ assuming the convention $0\log^2 0$. Now continuity follows from the fact that $x\log^2\{\alpha x\}$ is continuous for $x \in [0,1]$ when defined as 0 for $x = 0$. Thus, continuity of $U(P,W)$ is established.

To establish continuity of $V(P,W)$ we are left to prove that

$$\sum_x P(x)D(W_x||PW)^2 \tag{3.121}$$

is continuous in P. Let us expand a single term here:

$$P(x)\left[\sum_y W(y|x)\log\frac{W(y|x)}{PW(y)}\right]^2. \tag{3.122}$$

First notice that if $P_n(x) \not\to 0$ then continuity of this term follows from the fact that the argument of the logarithm is bounded away from 0 and $+\infty$ for all y with $W(y|x) > 0$. So we are left with the case $P_n(x) \to 0$. To that end let us prove the inequality for $P(x) > 0$:

$$D(W_x||PW) \le 2H(W_x) + \log\frac{1}{P(x)}\,. \tag{3.123}$$

From here continuity follows as we can see that $P_n(x)D(W_x||P_nW)^2 \to 0$ because $x\log x$ and $x\log^2 x$ are continuous at zero.

We now prove inequality (3.123). From (3.118) we see that

$$\left|\log\frac{W(y|x)}{PW(y)}\right| \le \log\frac{1}{W(y|x)} + \log\frac{1}{P(x)W(y|x)} = 2\log\frac{1}{W(y|x)} + \log\frac{1}{P(x)}\,. \tag{3.124}$$

Then,

$$D(W_x||PW) \le \sum_x W(y|x)\left|\log\frac{W(y|x)}{PW(y)}\right| \le 2H(W_x) + \log\frac{1}{P(x)}\,. \tag{3.125}$$

Thus $V(P,W)$ is continuous in P.

To establish continuity of $T(P,W)$, we again consider a single term:

$$P(x)W(y|x)\left|\log\frac{W(y|x)}{PW(y)} - D(W_x||PW)\right|^3. \tag{3.126}$$

If $W(y|x) = 0$ then this term is equal to zero regardless of P, and thus is continuous in P. Assume $W(y|x) > 0$. Take $P_n \to P$. If $P(x) \neq 0$ then $P_nW(y)$ is bounded away from 0 and thus $\log\frac{W(y|x)}{P_nW(y)}$ tends to $\log\frac{W(y|x)}{PW(y)}$. Similarly, for any y' such that $W(y'|x) > 0$ we have that $P_nW(y')$ is also bounded away from 0. Thus, $D(W_x||P_nW)$ tends to $D(W_x||PW)$.

We now assume that $P_n(x) \to 0$ and must prove that (3.126) tends to 0. Using the inequality $|a+b|^3 \le 4(|a|^3 + |b|^3)$, we obtain

$$P_n(x)W(y|x)\left|\log\frac{W(y|x)}{P_nW(y)} - D(W_x||P_nW)\right|^3 \le \tag{3.127}$$

$$4P_n(x)W(y|x)\left|\log\frac{W(y|x)}{P_nW(y)}\right|^3 + 4P_n(x)W(y|x)D^3(W_x||P_nW)\,. \tag{3.128}$$

Application of (3.123) immediately proves that the second term in the last inequality tends to zero. Continuity of the first term is established exactly like (3.117) with (3.119) replaced by

$$|\log W(y|x) - \log P_nW(y)|^3 \le 4(-\log^3 W(y|x) - \log^3\{P_n(x)W(y|x)\})\,. \tag{3.129}$$

This proves continuity of $T(P,W)$. A completely similar proof shows continuity of $T_u(P,W)$.

Finally, $V(P,W)$ and $U(P,W)$ coincide on Π for the reason that, under any capacity-achieving distribution it is known that

$$D(W_x||PW) = \mathbb{E}\,[i(X;Y)\,|\,X=x] = C \quad P\text{-a.s.}\,. \tag{3.130}$$

Indeed, then

$$U(P,W) \stackrel{\triangle}{=} \mathbb{E}\,\left[(i-\mathbb{E}\,i)^2\right] = \mathbb{E}\,\left[(i-C)^2\right] = \mathbb{E}\,\left[(i-\mathbb{E}\,[i\,|\,X])^2\right] \stackrel{\triangle}{=} V(P,W)\,. \tag{3.131}$$

The last equality is by the definition of $V(P,W)$ as the average conditional variance. ■

The fact that $U(P,W)$ and $V(P,W)$ coincide on Π is very important. Indeed, remember the classical proof of capacity and the strong converse. First, we use Feinstein's lemma to prove that $\frac{1}{n}\log M^* \gtrsim \mathbb{E}\,[i(X;Y)]$. Then, following Wolfowitz, Theorem 6, we establish that in fact the upper bound on rate depends on $\mathbb{E}\,[i(X;Y)\,|\,X]$. However, thanks to the fact (3.130) the conditional expectation is almost surely a constant.

The next result is needed (in particular) to show the achievability part of Strassen's theorem.

Theorem 47 *For any $P \in \mathcal{P}$, we have*

$$\log M^*_{avg}(n,\epsilon) \quad \geq \quad nI(P,W) - \sqrt{n\,U(P,W)}\,Q^{-1}(\epsilon) + O(1)\,, \tag{3.132}$$

if $U(P,W) > 0$ and

$$\log M^*(n,\epsilon) \geq nI(P,W) + \log\epsilon\,, \tag{3.133}$$

if $U(P,W) = 0$. If $V(P,W) > 0$ then we have

$$\log M^*(n,\epsilon) \geq nI(P,W) - \sqrt{n\,U(P,W)}\,Q^{-1}(\epsilon) + O(1)\,. \tag{3.134}$$

Remark: Note that the only case when result (3.132) is not implied by (3.134) is $V(P,W) = 0$ and $U(P,W) > 0$.

Proof: To show (3.132) select $P \in \mathcal{P}$. Let $\mathsf{A} = \mathcal{A}^n$, and choose the product measure P^n as the distribution of X^n. Passing this distribution through W^n induces a joint probability distribution on (X^n, Y^n), and the information density is the sum of independent identically distributed Z_k:

$$i(X^n;Y^n) = \sum_{k=1}^{n} \log\frac{W(Y_k|X_k)}{PW(Y_k)} = \sum_{k=1}^{n} Z_k\,. \tag{3.135}$$

The random variable Z_k has the distribution of $i(X;Y)$ when (X,Y) is distributed according to $P \times W$. Accordingly, it has mean $I(P,W)$ and variance $U(P,W)$, and its third absolute moment (being a continuous function of P, see Lemma 46) is uniformly bounded on $\mathcal{P}$:

$$\kappa = \sup_{P\in\mathcal{P}} T_u(P,W) < \infty\,. \tag{3.136}$$

Suppose that $U(P,W) = 0$, and therefore $i(X^n;Y^n) = nI(P,W)$. Then Theorem 23 asserts that there exists an (n,M,ϵ) code (maximal probability of error) for any M and

$$\epsilon \leq (M-1)\exp\{-nI(P,W)\}\,. \tag{3.137}$$

In particular, by taking $M = \lceil \exp\{-nI(P,W)\}\epsilon \rceil$ we get (3.133).

Now, assume that $U(P,W) > 0$ and denote

$$B \stackrel{\triangle}{=} \frac{6\kappa}{U(P,W)^{3/2}}\,. \tag{3.138}$$

To use the DT bound, Theorem 18, we need to prove that for some γ the following inequality holds:

$$\begin{aligned} \epsilon &\geq \mathbb{E}\left[\exp\left\{-[i(X^n;Y^n) - \log\gamma]^+\right\}\right] && (3.139)\\ &= \mathbb{P}[i(X^n;Y^n) \leq \log\gamma] && (3.140)\\ &\quad + \gamma\mathbb{E}\left[\exp\left\{-i(X^n;Y^n)\right\} 1_{\{i(X^n;Y^n)>\log\gamma\}}\right]\,. && (3.141)\end{aligned}$$

Denote for an arbitrary τ

$$\log\gamma = nI(P,W) - \tau\sqrt{nU(P,W)}\,. \tag{3.142}$$

According to Theorem 13, we have

$$\left|\mathbb{P}\left[i(X^n;Y^n) \leq \log\gamma\right] - Q(\tau)\right| \leq \frac{B}{\sqrt{n}}\,. \tag{3.143}$$

For sufficiently large n, let

$$\tau = Q^{-1}\left(\epsilon - \left(\frac{2\log 2}{\sqrt{2\pi U(P,W)}} + 5B\right)\frac{1}{\sqrt{n}}\right)\,. \tag{3.144}$$

Then, from (3.143) we obtain

$$\mathbb{P}\left[i(X^n;Y^n) \leq \log\gamma\right] \leq \epsilon - 2\left(\frac{\log 2}{\sqrt{2\pi U(P,W)}} + 2B\right)\frac{1}{\sqrt{n}}\,. \tag{3.145}$$

To bound the second term (3.141) we use Lemma 20, to obtain

$$\gamma\,\mathbb{E}\left[\exp\left\{-i(X^n;Y^n)\right\} 1_{\{i(X^n;Y^n)>\log\gamma\}}\right] \leq 2\left(\frac{\log 2}{\sqrt{2\pi U(P,W)}} + 2B\right)\frac{1}{\sqrt{n}}\,. \tag{3.146}$$

Summing (3.145) and (3.146) we prove inequality (3.139). Hence, by Theorem 18 we get

$$\begin{aligned} \log M^*_{avg}(n,\epsilon) &\geq \log\gamma && (3.147)\\ &= nI(P,W) - \tau\sqrt{nU(P,W)} && (3.148)\\ &= nI(P,W) - \sqrt{nU(P,W)}Q^{-1}(\epsilon) + O(1)\,, && (3.149)\end{aligned}$$

because according to (3.144) and the differentiability of Q^{-1} we have

$$\tau = Q^{-1}(\epsilon) + O\left(\frac{1}{\sqrt{n}}\right)\,. \tag{3.150}$$

To prove (3.134) we apply Theorem 22 with $\gamma(x^n)$ chosen as follows[8]

$$\gamma(x^n) = \begin{cases} \gamma', & \mathrm{Var}[i(X^n;Y^n)|X^n=x^n] \ge \frac{nV(P,W)}{2}, \\ +\infty, & \text{otherwise}, \end{cases} \tag{3.151}$$

where similar to (3.142) and (3.144) we choose

$$\log\gamma' = nI(P,W) - \sqrt{nU(P,W)}Q^{-1}\left(\epsilon - \left(\frac{2\log 2}{\sqrt{\pi V(P,W)}} + 7B'\right)\frac{1}{\sqrt{n}}\right), \tag{3.152}$$

where

$$B' \stackrel{\triangle}{=} \frac{2^{3/2}6\kappa}{V(P,W)^{3/2}}. \tag{3.153}$$

Theorem 22 guarantees existence of the (n, M, ϵ') code (maximal probability of error) with

$$\epsilon' \le \mathbb{P}[i(X^n,Y^n) \le \log\gamma(X^n)] + M\sup_{x^n}\mathbb{P}[i(x^n,Y^n) > \log\gamma(x^n)]. \tag{3.154}$$

The first term is upper-bounded as follows:

$$\begin{aligned} &\mathbb{P}[i(X^n,Y^n) \le \log\gamma(X^n)] && (3.155)\\ &\quad\le \mathbb{P}[i(X^n,Y^n) \le \log\gamma'] + \mathbb{P}[\gamma(X^n)=\infty] && (3.156)\\ &\quad\le \epsilon - 2\left(\frac{\log 2}{\sqrt{\pi V(P,W)}} + 3B'\right)\frac{1}{\sqrt{n}} + \mathbb{P}[\gamma(X^n)=\infty] && (3.157)\\ &\quad\le \epsilon - 2\left(\frac{\log 2}{\sqrt{\pi V(P,W)}} + 3B'\right)\frac{1}{\sqrt{n}} + \exp\{-O(n)\} && (3.158)\\ &\quad\le \epsilon - 2\left(\frac{\log 2}{\sqrt{\pi V(P,W)}} + 2B'\right), && (3.159) \end{aligned}$$

where (3.157) follows similar to (3.145) after noticing that $B' > B$ since $V(P,W) < U(P,W)$ and (3.158) is by Chernoff bound applied to a sum of bounded i.i.d. random variables:

$$\mathrm{Var}[i(X^n;Y^n)|X^n] = \sum_{j=1}^{n} V(W_{X_j}||PW), \tag{3.160}$$

and (3.159) holds for all n sufficiently large.

For the second term in (3.154) we have by Lemma 20

$$M\sup_{x^n}\mathbb{P}[i(x^n,Y^n) > \log\gamma(x^n)] \quad\le\quad \frac{M}{\gamma'}2\left(\frac{\log 2}{\sqrt{\pi V(P,W)}} + 2B'\right)\frac{1}{\sqrt{n}}. \tag{3.161}$$

[8] A similar idea of restricting the codewords to inputs with very small conditional variance appears in Strassen's [1] proof of the achievability part of his theorem.

Thus summing (3.158) and (3.161) we obtain from (3.154) an (n, M, ϵ') code with $\epsilon' \leq \epsilon$. Therefore, for all n sufficiently large

$$\log M^*(n, \epsilon) \geq \log \gamma' \tag{3.162}$$

from which (3.134) follows after invoking Taylor's expansion in (3.152). ■

Note that by using the Feinstein bound (2.31) we would not be able to derive such a strong estimate on $\log n$ term. This suboptimality in the $\log n$ term is an analytical illustration of the fact that we have already observed in Sections 3.2.3 and 3.3.3: the Feinstein bound is not tight enough for the refined analysis of $\log M^*(n, \epsilon)$ for finite n.

3.4.3 Converse bound

The following results are concerned with the behavior of the maximum of $nf(x) + \sqrt{n}g(x)$ for large n. We need them for the proof of the converse bound.

Lemma 48 *Let D be a compact metric space. Suppose $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ are continuous, then we have*

$$\max_{x \in D} \left[nf(x) + \sqrt{n}g(x) \right] = nf^* + \sqrt{n}g^* + o(\sqrt{n}) \,, \tag{3.163}$$

where

$$f^* = \max_{x \in D} f(x) \,, \tag{3.164}$$

$$g^* = \sup_{\{x:\, f(x) = f^*\}} g(x) \,. \tag{3.165}$$

Proof: Denote

$$F(x, n) = nf(x) + \sqrt{n}g(x) \tag{3.166}$$

$$F^*(n) = \max_{x \in D} F(x, n) \,. \tag{3.167}$$

Then (3.163) is equivalent to a pair of statements:

$$\lim_{n \to \infty} \frac{1}{n} F^*(n) = f^* \tag{3.168}$$

$$\lim_{n \to \infty} \frac{F^*(n) - nf^*}{\sqrt{n}} = g^* \tag{3.169}$$

which we are going to prove.

First we note that because of the compactness of D both f and g are bounded. Now

$$F(x, n) \leq nf^* + \sqrt{n}g_{max} \implies \frac{1}{n} F^*(n) \leq f^* + \frac{1}{\sqrt{n}} g_{max} \implies \limsup \frac{1}{n} F^*(n) \leq f^* \,. \tag{3.170}$$

On the other hand, if we take x^* to be any $x \in D$ maximizing $f(x)$ then

$$F^*(n) = \max_x F(x, n) \geq F(x^*, n) = nf^* + \sqrt{n}g(x^*) \,. \tag{3.171}$$

Thus

$$\liminf \frac{1}{n} F^*(n) \geq f^* \,, \tag{3.172}$$

and the first statement is proved.

Now define

$$D_1 = \{x \in D : \; f(x) = f^*\} \,, \tag{3.173}$$

which is also compact. Thus there exists a (possibly non-unique) maximum x^{**} of $g(x)$ on D_0:

$$x^{**} = \operatorname*{argmax}_{x \in D_0} g(x) \text{ and } g(x^{**}) = g^* \tag{3.174}$$

Now by definition

$$F^*(n) - nf^* \geq F(x^{**}, n) - nf^* = \sqrt{n} g^* \,. \tag{3.175}$$

Thus

$$\liminf \frac{F^*(n) - nf^*}{\sqrt{n}} \geq g^* \,. \tag{3.176}$$

On the other hand, $F(x, n)$ is continuous on D, so that

$$F^*(n) = F(x_n^*, n) \,. \tag{3.177}$$

Then notice that

$$F^*(n) - nf^* = n(f(x_n^*) - f^*) + \sqrt{n} g(x_n^*) \leq \sqrt{n} g(x_n^*) \,, \tag{3.178}$$

where the last inequality follows because $f(x_n^*) \leq f^*$. Now we see that

$$\frac{F^*(n) - nf^*}{\sqrt{n}} \leq g(x_n^*) \,. \tag{3.179}$$

On denoting

$$h(n) \stackrel{\triangle}{=} \frac{F^*(n) - nf^*}{\sqrt{n}} \,, \tag{3.180}$$

there exists a sequence $\{n_k\}$ such that

$$h(n_k) \to \limsup h(n) \,, \quad \text{as } k \to \infty \,. \tag{3.181}$$

For that sequence we have

$$h(n_k) \leq g(x_{n_k}^*) \,. \tag{3.182}$$

Since the $x_{n_k}^*$'s all lie in the compact D, there exists a convergent subsequence[9] :

$$y_l \stackrel{\triangle}{=} x_{n_{k_l}}^* \to x_0 \,. \tag{3.183}$$

We will now argue that $f(x_0) = f^*$.

[9]This is the only place where we used the metric-space nature of D. Namely we need sequential compactness to follow from compactness. Thus, in complete generality Lemma 48 holds for an arbitrary topological space D that is compact and satisfies the first axiom of countability.

As we have just shown,

$$\frac{1}{n_{k_l}} F^*(n_{k_l}) \to f^* , \tag{3.184}$$

where

$$F^*(n_{k_l}) = F(y_l, n_{k_l}) = n_{k_l} f(y_l) + \sqrt{n_{k_l}} g(y_l) . \tag{3.185}$$

Thus, since $g(x)$ is bounded

$$\lim_{l\to\infty} \frac{1}{n_{k_l}} F^*(n_{k_l}) = \lim_{l\to\infty} f(y_l) = f(x_0) , \tag{3.186}$$

where the last step follows from the continuity of f. So indeed

$$f(x_0) = f^* \quad \Longleftrightarrow \quad x_0 \in D_0 \quad \Longrightarrow \quad g(x_0) \le g^* . \tag{3.187}$$

Now we recall that

$$h(n_{k_l}) \le g(y_l) , \tag{3.188}$$

and by taking the limit as $l \to \infty$ we obtain

$$\limsup h(n) = \lim_{l\to\infty} h(n_{k_l}) \le \lim_{l\to\infty} g(y_l) = g(x_0) \le g^* . \tag{3.189}$$

So we have shown

$$\lim \frac{F^*(n) - n f^*}{\sqrt{n}} = g^* . \tag{3.190}$$

■

Remarks:

1. The message of this lemma is that, for continuous f and g,

$$\max_x \left[n f(x) + \sqrt{n} g(x) \right] \approx n f(x^{**}) + \sqrt{n} g(x^{**}) \tag{3.191}$$

   where x^{**} is found by first maximizing $f(x)$ and then maximizing $g(x)$ over the set of maximizers of $f(x)$.

2. Lemma 48 can be generalized to any finite set of "basis terms", instead of $\{n, \sqrt{n}\}$. In this case, the only requirement would be that $u_{j+1}(n) = o(u_j(n))$.

3. Lemma 48 is tight in the sense that term $o(\sqrt{n})$ can not be improved without further assumptions. Indeed, take $f(x) = -x^2$ and $g(x) = x^{1/k}$ for some $k \in \mathbb{Z}_+$ on $[-1, 1]$. Then simple calculation shows that

$$\max_{x\in[-1,1]} \left[n f(x) + \sqrt{n} g(x) \right] = \text{const} \cdot n^{\frac{k-1}{2k-1}} \tag{3.192}$$

   and the power of n can be arbitrary close to $\sqrt{n}$.

If we assume more about f and g then a stronger result can be stated. The assumptions below essentially mean that f is twice differentiable near f^* with negative-definite Hessian and g is differentiable. As the example (3.192) shows, without these additional assumptions Lemma 48 is the best possible result.

Lemma 49 *In the notation of previous lemma, denote*

$$D_0 \stackrel{\triangle}{=} \{x : f(x) = f^*\}, \text{ and } D_\delta \stackrel{\triangle}{=} \{x : d(x, D_0) \le \delta\}, \tag{3.193}$$

where $d(\cdot,\cdot)$ is a metric. Suppose that for some $\delta > 0$ and some constants $f_1 > 0$ and f_2 we have

$$f(x) - f^* \le -f_1\, d(x, D_0)^2, \text{ and} \tag{3.194}$$
$$|g(x) - g^*| \le f_2\, d(x, D_0). \tag{3.195}$$

Then,

$$\max_{x\in D} \left[nf(x) + \sqrt{n}g(x)\right] = nf^* + \sqrt{n}g^* + O(1). \tag{3.196}$$

Proof: Because of the boundedness of $g(x)$, the points x_n^* must all lie in D_δ for n sufficiently large. So, for such n we have

$$\max_{x\in D} F(x,n) = \max_{x\in D_\delta} F(x,n) \tag{3.197}$$

(we use the notation from the above proof).

Continue as follows:

$$\max_{x\in D_\delta} F(x,n) = nf^* + \sqrt{n}g^* + \left[n(f(x_n^*) - f^*) + \sqrt{n}(g(x_n^*) - g^*))\right]. \tag{3.198}$$

We can now bound the term in brackets by using conditions in the lemma:

$$0 \le \left[n(f(x_n^*) - f^*) + \sqrt{n}(g(x_n^*) - g^*))\right] \le -f_1\left(\sqrt{n}d(x_n^*, D_0)\right)^2 + f_2\left(\sqrt{n}d(x_n^*, D_0)\right). \tag{3.199}$$

Now we see that we have a quadratic polynomial in variable $y \stackrel{\triangle}{=} \sqrt{n}d(x_n^*, D_0)$. Since $f_1 > 0$ it has a maximum equal to $\frac{f_2^2}{4f_1^2}$. Then,

$$0 \le \left[n(f(x_n^*) - f^*) + \sqrt{n}(g(x_n^*) - g^*))\right] \le \frac{f_2^2}{4f_1^2} \tag{3.200}$$

and we see that residual term is $O(1)$. This establishes (3.196). ■

To prove the converse bound we introduce the following definitions:

- For any $P_0 \in \mathcal{P}_n$ denote *a type* of elements $x^n \in \mathcal{A}^n$ by

$$T_{P_0}^n \stackrel{\triangle}{=} \{x^n : \forall a \in \mathcal{A} : \sum_{i=1}^n 1_{\{x_i = a\}} = P_0(a)\}. \tag{3.201}$$

- For any n and $P_0 \in \mathcal{P}_n$ define:

$$M_{P_0}^*(n,\epsilon) = \max\{M : \exists(n, M, \epsilon)\text{-code with codewords } \in T_{P_0}^n\}, \tag{3.202}$$

  where ϵ is the maximal probability of error

Now we are ready to prove the following theorem.

Theorem 50 *Fix a channel W. If $\epsilon \in (0, 1/2]$ then there exist a number $N_0 \geq 1$ and a constant $F > 0$ such that for all $n \geq N_0$ and $P_0 \in \mathcal{P}_n$ we have*

$$\log M^*_{P_0}(n, \epsilon) \leq nC - \sqrt{nV_{min}}Q^{-1}(\epsilon) + \frac{1}{2}\log n + F\,. \tag{3.203}$$

If $\epsilon \in (1/2, 1)$ then there exist a number $N_0 \geq 1$ and a constant $F > 0$ such that for all $n \geq N_0$ and $P_0 \in \mathcal{P}_n$ we have

$$\log M^*_{P_0}(n, \epsilon) \leq nC - \sqrt{nV_{max}}Q^{-1}(\epsilon) + \frac{1}{2}\log n + F \tag{3.204}$$

unless the channel is an exotic DMC in which case we have only

$$\log M^*_{P_0}(n, \epsilon) \leq nC + Fn^{1/3}\,. \tag{3.205}$$

Proof: We must consider four cases separately:

1. $\epsilon \leq 1/2$ and $V_{min} > 0$.
2. $\epsilon \leq 1/2$ and $V_{min} = 0$.
3. $\epsilon > 1/2$ and $V_{max} > 0$.
4. $\epsilon > 1/2$ and $V_{max} = 0$.

It is instructive to begin with some general remarks. For simplicity of notation we denote elements of $\mathsf{A} = \mathcal{A}^n$ and $\mathsf{B} = \mathcal{B}^n$ by x and y without superscripts. The aim is to use Theorem 34 with $\mathsf{F}_n = T^n_{P_0}$. To do so we need to select a distribution P_{Y^n} on $\mathcal{A}^n$ and compute $\inf_{x \in T^n_{P_0}} \beta^n_\alpha(x, P^n_Y)$. Notice that the theorem is concerned only with codebooks over some fixed type. So, if P_{Y^n} is a product distribution then $\beta^n_\alpha(x, P_Y)$ does not depend on $x \in T^n_{P_0}$ and thus

$$\beta^n_\alpha(x, P_Y) = \beta^n_\alpha(P_Y)\,. \tag{3.206}$$

For this reason we will simply write $\beta^n_\alpha(P_Y)$, and even β^n_α, since the distribution P_Y will be apparent. After these remarks, treatment of the regular case (Case 1) becomes a straightforward application of Theorem 34, while for Case 2 we need an original method proposed by Strassen.

Case 1. Denote the closed δ-neighborhood of Π as

$$\Pi_\delta \stackrel{\triangle}{=} \{P \in \mathcal{P} : d(P, \Pi) \leq \delta\}\,. \tag{3.207}$$

Here $d(\cdot, \cdot)$ denotes Euclidean distance between vectors of $\mathbb{R}^{|\mathcal{A}|}$.

We fix some $\delta > 0$ to be determined. First, we find δ_1 small enough so that everywhere on Π_{δ_1} we have $V(P, W) \geq V_{min}/2$. This is possible by the continuity of $V(P, W)$.

Without loss of generality, we can assume that $\mathcal{B}$ does not have inaccessible outputs, i.e. for every $y_0 \in \mathcal{B}$ there is an $x_0 \in \mathcal{A}$ such that $W(y_0|x_0) > 0$. Then, it is well known that

for any $P_1, P_2 \in \Pi$ the output distributions coincide, i.e. $P_1W = P_2W = P_Y^*$, and also that this unique P_Y^* dominates all $W(\cdot|x)$. Since all outputs are accessible, this implies that

$$P_Y^*(y) > 0, \ \ \forall y \in \mathcal{B}\,. \tag{3.208}$$

Now for each y, the function $PW(y)$ is linear in the input distribution P, and thus there is some $\delta_2 > 0$ such that in the closed δ_2-neighborhood of Π we have $PW(y) > 0$ for all $y \in \mathcal{B}$. Set $\delta = \min(\delta_1, \delta_2)$.

Fix n and $P_0 \in \mathcal{P}_n$. Choose the distribution P_Y on $\mathcal{A}^n$ as the n-fold product of P_0W. Also set $\alpha = 1 - \epsilon$. Then by Theorem 34 and the argument above we have

$$\log M_{P_0}^*(n, \epsilon) \le -\log \beta_\alpha^n(x, P_Y) \tag{3.209}$$

where x is any element of $T_{P_0}^n$.

The idea for lower-bounding β_α^n is to apply Lemma 14 if $P_0 \in \Pi_\delta$ and Lemma 15 otherwise. In both cases, $P_i = Q_{Y|X=x_i}$ and $Q_i = P_0W$. Note that there are $nP_0(1)$ occurrences of $P_{Y|X=1}$ among the P_i's, $nP_0(2)$ occurrences of $P_{Y|X=2}$ etc. Thus, the quantities defined in Lemma 14 become

$$D_n = I(P_0, W)\,, \text{ and } V_n = V(P_0, W)\,. \tag{3.210}$$

Suppose that $P_0 \in \mathcal{P}_n \setminus \Pi_\delta$; then, applying Lemma 15 we obtain

$$\log M_{P_0}^*(n, \epsilon) \le -\log \beta_\alpha^n \le nI(P_0, W) + \sqrt{\frac{2nV(P_0, W)}{1-\epsilon}} + \log \frac{1-\epsilon}{2}\,. \tag{3.211}$$

Denote,

$$C' = \sup_{P \in \mathcal{P} \setminus \Pi_\delta} I(P, W) < C\,, \quad M_V = \max_{P \in \mathcal{P}} V(P, W) < \infty\,. \tag{3.212}$$

Then, continuing the bound, we have

$$\log M_{P_0}^*(n, \epsilon) \le nC' + \sqrt{\frac{2M_V}{1-\epsilon}}\sqrt{n} + \log \frac{1-\epsilon}{2}\,. \tag{3.213}$$

Since $C' < C$ we can see that, even with $F = 0$, there exists N_1 such that for all $n \ge N_1$ the right-hand side of (3.213) is below the right-hand side of (3.203). So this proves the theorem for $P_0 \in \mathcal{P}_n \setminus \Pi_\delta$.

Now, consider $P_0 \in \Pi_\delta$. Remember, that T_n in Lemma 14 is in fact

$$T_n = \sum_{x,y} P_0(x)W(y|x) \left| \log \frac{W(y|x)}{P_0W(y)} - D(W_x||P_0W) \right|^3 = T(P_0, W)\,. \tag{3.214}$$

As shown in Lemma 46 function $T(P_0, W)$ is continuous on $\mathcal{P}$ and thus has a finite upper-bound:

$$T_n \le M_T < \infty\,. \tag{3.215}$$

On the other hand, over Π_δ we have $V(P_0, W) \ge V_{min}/2 > 0$. In summary, we can upper bound B_n in Lemma 14 as

$$B_n \stackrel{\triangle}{=} 6\frac{T_n}{V_n^{3/2}} \le M_B \stackrel{\triangle}{=} \frac{6 \cdot 2^{3/2} M_T}{V_{min}^{3/2}}\,. \tag{3.216}$$

Thus we are ready to apply Lemma 14, namely to use (2.87) with $\Delta = M_B - B_n + 1 \geq 1$ and to conclude that, for n sufficiently large,

$$\log M^*_{P_0}(n,\epsilon) \leq nI(P_0,W) + \sqrt{nV(P_0,W)}Q^{-1}\left(\alpha - \frac{M_B+1}{\sqrt{n}}\right) + \frac{1}{2}\log n\,. \tag{3.217}$$

For n large, depending on M_B, we can expand Q^{-1} using Taylor's formula. In this way, we can conclude that there is a constant F_1 such that

$$Q^{-1}\left(\alpha - \frac{M_B+1}{\sqrt{n}}\right) \leq Q^{-1}(\alpha) + \frac{F_1}{\sqrt{n}}\,. \tag{3.218}$$

Then for such n and a constant F_2 (remember $V(P_0,W) \leq M_V$), we have

$$\log M^*_{P_0}(n,\epsilon) \leq nI(P_0,W) + \sqrt{nV(P_1,W)}Q^{-1}(\alpha) + \frac{1}{2}\log n + F_2\,. \tag{3.219}$$

To conclude the proof we must maximize the right-hand side over $P_0 \in \Pi_\delta$. Note that this is exactly the case treated in Lemmas 48 and 49. We want to use the latter one and need to check its conditions. From the definitions of $I(P,W)$ and $V(P,W)$ we can see that on Π_δ they are infinitely differentiable functions. This is because all terms $\log\frac{W(y|x)}{PW(y)}$ have argument bounded away from 0 and $+\infty$ by the choice of Π_δ. Consequently, the conditions of Lemma 49 on g are automatically satisfied. We must now check the conditions on f.

For this we can think of $I(P,W)$ as a function of P only, and we write $\nabla I(P)$ and $\mathcal{H}(P)$ for the gradient vector and Hessian matrix correspondingly.

To check conditions on f in Lemma 49 it is sufficient to prove that for any $P^* \in \Pi$:

1. $\ker\mathcal{H}(P^*) = \ker W$. By $\ker W$ we understand all $|\mathcal{A}|$-vectors v such that $\sum_{x\in\mathcal{A}} v(x)W(y|x) = 0$; and
2. the largest non-zero eigenvalue of $\mathcal{H}(P^*)$ is negative and bounded away from zero uniformly in the choice of $P^* \in \Pi$.

We first show why these two conditions are sufficient. It is known that Π consists of all distributions P that satisfy two conditions: 1) $PW = P^*_Y$; and 2) $P(x) > 0$ only when $D(W_x||P^*_Y) = C$. Now take some $P' \notin \Pi$ and denote by P^* the projection of P' onto a compact Π. Then write

$$P' = P^* + v = P^* + v_0 + v_\perp\,, \tag{3.220}$$

where v_0 is projection of $v = (P' - P^*)$ onto $\ker W$ and $v_\perp$ is orthogonal to $\ker W$. Note that $d(P',\Pi) = ||v|| \leq \delta$. By Taylor's expansion we have

$$I(P') = I(P^*) + (\nabla I(P^*), v_0 + v_\perp) + \frac{1}{2}(\mathcal{H}(P^*)v_\perp, v_\perp) + o(||v||^2)\,. \tag{3.221}$$

Here we used the assumed fact that $(\mathcal{H}(P^*)v_0, v_0) = 0$. Since $v_0 \in \ker W$ but $P^* + \alpha v_0$ is not in Π for any $\alpha > 0$, we conclude that shifting along v_0 must involve inputs with $D(W_x||P^*_Y) < C$. But then $I(P,W)$ decays linearly along this direction, i.e. there is some constant $f_3 > 0$ such that

$$I(P^* + v_0) - I(P^*) = (\nabla I(P^*), v_0) \leq -f_3||v_0|| \leq -f_3||v_0||^2 \tag{3.222}$$

(the last inequality assumes $\delta \le 1$). Then, substituting this into expansion for $I(P')$ and upper-bounding $(\nabla I, v_\perp)$ by zero we obtain

$$I(P') - I(P^*) \le -f_3||v_0||^2 - \frac{1}{2}\lambda||v_\perp||^2 + o(||v||^2)\,, \tag{3.223}$$

where λ is the absolute value of the maximal non-zero eigenvalue of $\mathcal{H}(P^*)$. We will show that λ is uniformly bounded away from zero for any $P^* \in \Pi$. So we see that indeed $I(P,W)$ decays not slower than quadratically in $d(P,\Pi)$.

Now we need to prove the assumed facts about the Hessian $\mathcal{H}(P)$. The differentiation can be performed without complications since on Π_δ we always have $PW(y) > 0$. After some algebra we get

$$\mathcal{H}_{ij} \stackrel{\triangle}{=} \frac{\partial^2 I(P)}{\partial P(i)\partial P(j)} = -\log e \cdot \sum_y \frac{W(y|i)W(y|j)}{PW(y)}\,. \tag{3.224}$$

Thus, for any vector v we have

$$(\mathcal{H}v, v) = \sum v_i H_{ij} v_j = -\sum_y \frac{\left(\sum_i v_i W(y|i)\right)^2}{PW(y)} \le -\frac{||vW||^2}{(PW)_{max}}\,, \tag{3.225}$$

where we have denoted formally $vW = \sum_{x\in\mathcal{A}} v(x)W(y|x)$, which is a vector of dimension $|\mathcal{B}|$. From (3.225) we can see that indeed $(\mathcal{H}v, v) = 0$ if and only if $vW = 0$. In addition, the maximal non-zero eigenvalue of $\mathcal{H}(P)$ is always smaller than $\frac{\lambda_{min+}(WW^T)}{(PW)_{max}}$ for all $P \in \Pi$.

So Lemma 49 applies to (3.219), and thus

$$\log M^*_{P_0}(n,\epsilon) \le nC + \sqrt{nV_{min}}Q^{-1}(\alpha) + \frac{1}{2}\log n + O(1)\,. \tag{3.226}$$

This implies (3.203) if we note that $Q^{-1}(\alpha) = -Q^{-1}(\epsilon)$.

Case 3. The proof for this case is analogous to that for Case 1, except that when applying Lemma 49 we must choose $g^* = \sqrt{V_{max}}$ because the sign of $Q^{-1}(\alpha)$ is negative this time. An additional difficulty is that it might be possible that $V_{max} > 0$ but $V_{min} = 0$. In this case the bound (3.216) is no longer applicable. What needs to be done is to eliminate types inside Π_δ with small variance:

$$\Pi_V = \{P \in \Pi_\delta :\ V(P,W) < A\}\,. \tag{3.227}$$

Here A is chosen so that

$$\sqrt{\frac{2A}{1-\epsilon}} < -\sqrt{V_{max}}Q^{-1}(\epsilon)\,. \tag{3.228}$$

Since $\epsilon > 1/2$ it is possible to find such an A. Then, for types in Π_V we can apply the firm bound in Lemma 15. For remaining types in $\Pi_\delta \setminus \Pi_V$ the reasoning argument of Case 1 works, after V_{min} is replaced by A in (3.216).

Case 2. The idea is to apply Theorem 34, but this time we fix the output distribution to be $P_{Y^n} = (P^*_Y)^n$ for all types P_0 (before we chose $P_{Y^n} = (P_0W)^n$ different for each type P_0). It is well-known that

$$D(W||P^*_Y|P_0) \le D(W||P^*_Y|P^*) = C\,. \tag{3.229}$$

This fact is crucial for proving the bound.

Note that $V(W_x||P_Y^*)$ is defined and finite since all $W_x \ll P_Y^*$. Denote a special subset of *nonzero-variance inputs* as

$$\mathcal{A}_+ \stackrel{\triangle}{=} \{x \in \mathcal{A} : \; V(W_x||P_Y^*) > 0\} \,. \tag{3.230}$$

And also for every $P_0 \in \mathcal{P}_n$ denote $m(P_0) = nP_0(\mathcal{A}_+)$ which is the number of nonzero-variance letters in any $x \in T_{P_0}^n$. Also note that there are minimal and maximal variances $V_M \geq V_m > 0$ such that $V_m \leq V(W_x||P_Y^*) \leq V_M$ for all $x \in \mathcal{A}_+$.

Since P_{Y^n} is a product distribution, it is true that

$$\log M_{P_0}^*(n, \epsilon) \leq -\log \beta_\alpha^n(x, P_{Y^n}) \tag{3.231}$$

for all $x \in T_{P_0}^n$. We are going to apply Lemmas 14 and 15 and so need to compute D_n, V_n and an upper-bound on B_n. We have

$$D_n = D(W||P_Y^*|P_0) \text{ and } V_n = V(W||P_Y^*|P_0) \,. \tag{3.232}$$

To upper-bound B_n we must lower-bound V_n and upper-bound T_n. Note that

$$V(W||P_Y^*|P_0) \geq \frac{m(P_0)}{n} V_m \,. \tag{3.233}$$

For T_n, we can write

$$T_n \stackrel{\triangle}{=} \sum_{x,y} P_0(x)W(y|x) \left| \log \frac{W(y|x)}{P_Y^*(y)} - D(W_x||P_Y^*) \right|^3 = \sum_x P_0(x) T(x) \,. \tag{3.234}$$

Here, the $T(x)$'s are all finite and $T(x) = 0$ iff $x \notin \mathcal{A}_+$. Thus, for $x \in \mathcal{A}_+$ there is one maximal $T^* = \max_{x \in \mathcal{A}} T(x)$, and we have

$$T_n \leq \frac{m(P_0)}{n} T^* \,. \tag{3.235}$$

Then, we see that

$$B_n \stackrel{\triangle}{=} \frac{T_n}{V_n^{3/2}} \leq \sqrt{\frac{n}{m(P_0)}} \frac{T^*}{V_m^{3/2}} \stackrel{\triangle}{=} \sqrt{\frac{n}{m(P_0)}} M_B \,. \tag{3.236}$$

So we apply Lemma 14 with

$$\Delta = \sqrt{\frac{n}{m(P_0)}} \, (M_B + 1) - B_n \geq \sqrt{\frac{n}{m(P_0)}} \geq 1 \,. \tag{3.237}$$

Using (2.87) and lower-bounding $\log \Delta$ via the above bound yields

$$\log \beta_\alpha^n \geq -nD(W||P_Y^*|P_0) - \sqrt{nV(W||P_Y^*|P_0)} Q^{-1} \left(\alpha - \frac{M_B + 1}{\sqrt{m(P_0)}} \right) - \frac{1}{2} \log n \,. \tag{3.238}$$

Now, it is an elementary analytical fact that for any $\alpha \in (0, 1)$ it is possible to pick an $x_0 < \alpha$ and $f > 0$ such that

$$Q^{-1}(\alpha - x) \leq Q^{-1}(\alpha) + fx \,, \forall x \in [0, x_0] \tag{3.239}$$

(for $\alpha > 1/2$ just take f to be a slope of $Q^{-1}(y)$ at $y = \alpha$ for $\alpha \le 1/2$ simply join the point $(\alpha, Q^{-1}(\alpha))$ with any $(\delta, Q^{-1}(\delta))$ for small δ and set $x_0 = \alpha - \delta$). We now split types in $\mathcal{P}_n$ into two classes, $\mathcal{P}_A$ and $\mathcal{P}_B$:

$$P_0 \in \mathcal{P}_A \iff m(P_0) \ge m_* \,, \quad \mathcal{P}_B = \mathcal{P}_n \setminus \mathcal{P}_A \,. \tag{3.240}$$

Here m_* is chosen so that $\frac{M_B+1}{\sqrt{m_*}} \le x_0$. Then, for all types in $\mathcal{P}_A$ we have

$$Q^{-1}\left(\alpha - \frac{M_B+1}{\sqrt{m(P_0)}}\right) \le Q^{-1}(\alpha) + \frac{f'}{\sqrt{m(P_0)}} \,. \tag{3.241}$$

Notice also that with this choice of x_0 and m_*, the argument of Q^{-1} in (3.238) is positive and the bound is applicable to all types in $\mathcal{P}_A$. Substituting (3.229) we have, for any $P_0 \in \mathcal{P}_A$,

$$\log \beta_\alpha^n \ge -nC - \sqrt{nV(W||P_Y^*|P_0)}Q^{-1}(\alpha) - f'\sqrt{\frac{nV(W||P_Y^*|P_0)}{m(P_0)}} - \frac{1}{2}\log n \,. \tag{3.242}$$

Now notice that $Q^{-1}(\alpha) \ge 0$ (this is the key difference with Case 4) and also that

$$V(W||P_Y^*|P_0) \le \frac{m(P_0)}{n} V_M \,. \tag{3.243}$$

Finally, for $P_0 \in \mathcal{P}_A$ we have

$$\log M_{P_0}^*(n,\epsilon) \le nC + f'\sqrt{V_M} + \frac{1}{2}\log n \,. \tag{3.244}$$

Now for types in $\mathcal{P}_B$ we have $m(P_0) < m_*$ and thus,

$$nV(W||P_Y^*|P_0) \le m_* V_M \,. \tag{3.245}$$

So Lemma 15 yields

$$\log M_{P_0}^*(n,\epsilon) \le nC + \sqrt{\frac{2m_* V_M}{\alpha}} - \log\frac{\alpha}{2} \,. \tag{3.246}$$

In summary, we see that in both cases, $\mathcal{P}_A$ and $\mathcal{P}_B$, inequalities (3.244) and (3.246) imply (3.203) for $n \ge 1$.

Case 4. Fix a type $P_0 \in \mathcal{P}_n$ and use $P_{Y^n} = \prod(P_0 W)$. Then, a similar argument to that for Case 2 and Lemma 15 yield

$$\log M_{P_0}^*(n,\epsilon) \le nI(P_0,W) + \sqrt{\frac{2nV(P_0,W)}{\alpha}} + \log\frac{\alpha}{2} \tag{3.247}$$

for all $n \ge 1$. We need to maximize the right-hand side of this bound over $P_0 \in \mathcal{P}$. This can be done similarly to Lemma 49. The problem here, however, is that $V(P,W) = 0$ for $P \in \Pi$. Thus, even though $V(P,W)$ is differentiable in some neighborhood of Π, the $\sqrt{V(P,W)}$ is not. This is how a term of order $n^{1/3}$ can appear. Indeed, suppose that there is some direction h along which $I(P+\alpha h)$ decays quadratically, while $V(P+\alpha h)$ is linear. I.e.,

$$I(P+\alpha h) = C - f_1\alpha^2 + o(\alpha^2) \,, \text{and } V(P+\alpha h) = f_2\alpha h + o(\alpha) \,. \tag{3.248}$$

Then it is not hard to see that

$$\max_{\alpha}[nI(P+\alpha h)+\sqrt{nV(P+\alpha h)}]=nC+f_3 n^{1/3}+o(n^{1/3})\,. \tag{3.249}$$

Such a direction can only exist if all the conditions of the exotic DMC are satisfied. This can be proved by computing gradients of $I(P,W)$ and $V(P,W)$. ■

Some notes about the proof are of interest:

- As a by-product of Lemma 15 we can derive a certain converse bound for DMC. Indeed, denote $M_V=\max_P V(P,W)$. Then

$$M^*(\epsilon,n)\le \sum_{P_0\in\mathcal{P}_n}\frac{1}{\beta_\alpha(T^n_{P_0},P_0W)}\,, \tag{3.250}$$

  and

$$\log M^*(\epsilon,n)\le nC+\sqrt{\frac{2nM_V}{1-\epsilon}}+|\mathcal{A}|\log(n+1)-\log\frac{1-\epsilon}{2}\,, \tag{3.251}$$

  which are valid for all n and $\epsilon\in(0,1)$. The latter bound does not even yield the right sign of the $\sqrt{n}$ term. However, it holds for all $n\ge 1$ and also M_V can be upper-bounded so that it depends on $|\mathcal{A}|$ but not on W. The resulting bound is better than Fano's inequality.

- The method used in the proof of Case 2 is quite useful for symmetric channels. Indeed, if we take the BSC with parameter δ, then P_Y^* is equiprobable. As we have seen in Section 3.2.1, $\beta^n_\alpha(x,P^*_{Y^n})$ is the same for all $x\in\mathcal{A}^n$. So, we can lower-bound $M^*(n,\epsilon)$ directly without resorting to a type-by-type analysis:

$$\log M^*(n,\epsilon)\le -\log\beta^n_\alpha\,. \tag{3.252}$$

  Calculation of the parameters in Lemma 14 yields

$$D_n=C(\delta)\,,\quad V_n=v(\delta)\,,\text{and } T_n=t(\delta)\,, \tag{3.253}$$

  where $C(\delta),v(\delta)$ and $t(\delta)$ are the capacity, dispersion and third moment, respectively, defined as

$$C(\delta)=\log 2+\delta\log\delta+(1-\delta)\log(1-\delta)\,, \tag{3.254}$$

$$v(\delta)=\delta(1-\delta)\log^2(\delta^{-1}-1)\,, \tag{3.255}$$

$$t(\delta)=\delta(1-\delta)(\delta^2+(1-\delta)^2)\left|\log^3(\delta^{-1}-1)\right|\,. \tag{3.256}$$

  Substituting these expressions into Lemma 14 and using $\Delta=1$ we have the following firm bound for the BSC

$$\log M^*(n,\epsilon)\le nC(\delta)-\sqrt{nv(\delta)}Q^{-1}\left(\epsilon+\frac{t(\delta)+v(\delta)^{3/2}}{\sqrt{v(\delta)^3 n}}\right)+\frac{1}{2}\log n\,. \tag{3.257}$$

  The only requirement for validity of this bound is that the argument of Q^{-1} be less than 1. Note that we dropped the 6 in the definition of B_n, because the components are identically distributed; see the remark after the Berry-Esseen Theorem 13. Even though this bound gives the right asymptotic behavior of $\log M^*(n,\epsilon)$, it is much worse than the direct computation of β^n_α that we did in Section 3.2.1. Still it is much better than using Fano's inequality.

3.4.4 Asymptotic expansion

Proof of Theorem 45: Theorem 47 yields, by taking $P \in \mathcal{P}$ to be the distribution that achieves capacity and V_ϵ, the bound (3.112). Indeed, if $V_\epsilon = 0$ then the bound (3.133) applies and if $V_\epsilon > 0$ then we have $U(P,W) = V(P,W) > 0$ and thus (3.134) yields the needed result.

For the lower bound, assume that either DMC is not exotic or $\epsilon \le 1/2$ and take $n \ge N_0$ for N_0 from Theorem 50. Then any (n, M, ϵ) is composed of subcodes over types $T^n_{P_0}$ for $P_0 \in \mathcal{P}_n$. If we remove all codewords except those in $T^n_{P_0}$ and leave the decoding regions untouched, then we obtain an (n, M'_{P_0}, ϵ) code over $T^n_{P_0}$. But then Theorem 50 states that

$$\log M'_{P_0} \le \log M^*_{P_0}(n,\epsilon) \le nC - \sqrt{nV_\epsilon}Q^{-1}(\epsilon) + \frac{1}{2}\log n + F\,. \tag{3.258}$$

Since M is a sum of M'_{P_0} over all $P_0 \in \mathcal{P}_n$ and the cardinality of $\mathcal{P}_n$ is no more than $(n+1)^{|\mathcal{A}|-1}$, we conclude

$$\log M^*(n,\epsilon) \le nC - \sqrt{nV_\epsilon}Q^{-1}(\epsilon) + \left(|\mathcal{A}| - \frac{1}{2}\right)\log n + F'\,. \tag{3.259}$$

This completes the proof of (3.110).

To show (3.111) we use the traditional idea of dropping all codewords whose probability of error is above $\tau\epsilon$. In this way, we have

$$M^*_{avg}(n,\epsilon) \le \frac{1}{1-1/\tau} M^*(n,\tau\epsilon)\,. \tag{3.260}$$

Carefully following the proof of the converse we can conclude that the $O(\log n)$ term in the upper bound (3.259) does not have any singularities in a neighborhood of any $\epsilon \in (0,1)$. So we can claim that, for τ sufficiently close to 1, the expansion

$$\log M^*(n,\tau\epsilon) = nC - \sqrt{nV_\epsilon}Q^{-1}(\tau\epsilon) + O(\log n) \tag{3.261}$$

holds uniformly in τ. Now, setting $\tau_n = 1 + \frac{1}{\sqrt{n}}$, we obtain

$$\log M^*_{avg}(n,\epsilon) \le nC - \sqrt{nV_\epsilon}Q^{-1}\left(\epsilon + \frac{1}{\sqrt{n}}\right) + O(\log n)\,. \tag{3.262}$$

Expanding Q^{-1} in (3.262) by Taylor's formula and using the obvious lower bound $M^*_{avg} \ge M^*$ we obtain (3.111).

Finally, for the case of an exotic DMC and $\epsilon > 1/2$, bound (3.205) in Theorem 50 proves (3.113). Together, (3.110) and (3.113) cover all possible cases and by (2.25) prove that V_ϵ is indeed the ϵ-dispersion of the DMC.

Finally the claim that the order $n^{\frac{1}{3}}$ cannot be improved in general follows from the next result. ■

Theorem 51 *There exists an exotic DMC with*

$$W \approx \begin{pmatrix} 1/3 & 0 & 0 & 1/3 & 0.04 \\ 0 & 1/3 & 0 & 1/3 & 0.04 \\ 0 & 0 & 1/3 & 1/3 & 0.04 \\ 1/3 & 1/3 & 1/3 & 0 & 0.38 \\ 1/3 & 1/3 & 1/3 & 0 & 0.50 \end{pmatrix} \tag{3.263}$$

Moreover, for every $\epsilon > 1/2$ there exists a constant $F_\epsilon > 0$ such that

$$\log M^*(n, \epsilon) \geq nC + F_\epsilon n^{\frac{1}{3}}\,, \tag{3.264}$$

for all n sufficiently large, where $C = \log \frac{5}{3}$ is the capacity of DMC W.

Proof: First take

$$W' = \begin{pmatrix} 1/3 & 0 & 0 & 1/3 & 2/13 \\ 0 & 1/3 & 0 & 1/3 & 2/13 \\ 0 & 0 & 1/3 & 1/3 & 2/13 \\ 1/3 & 1/3 & 1/3 & 0 & 3/13 \\ 1/3 & 1/3 & 1/3 & 0 & 4/13 \end{pmatrix} \tag{3.265}$$

Now denote by x^* the unique negative root of the equation

$$(x-1)\left(\log \frac{13^7}{2^8 3^3} - 7\log(1-x)\right) + (6+7x)\log\frac{6+7x}{39} = -13\log 3\,. \tag{3.266}$$

Then, replace the last column of W' with the column

$$R = W'[0\ 0\ 0\ x^*\ 1-x^*]^T \approx [0.04\ 0.04\ 0.04\ 0.38\ 0.50]^T\,. \tag{3.267}$$

The resulting channel matrix W is of full rank, since W' is such and operation (3.267) preserves the rank. Therefore, the capacity achieving distribution is unique. A simple observation shows that equiprobable P_Y^* is achievable by taking

$$P^* = [1, 1, 1, 2, 0]^T/5\,. \tag{3.268}$$

Finally, the conditional entropies $H(Y|X=x)$ are all equal to $\log 3$. This is the consequence of the choice of x^* in (3.266). It follows that P^* is the capacity achieving distribution (unique). Moreover, we also have

$$V(P^*, W) = 0\,. \tag{3.269}$$

but at the same time

$$V(W_5||P_Y^*) > 0\,. \tag{3.270}$$

So this is an exotic channel.

It remains to show (3.264). On the simplex $\mathcal{P}$ we have differentiable functions $P_x = P(x), x = 1, \ldots, 5$ which satisfy

$$\sum_{x=1}^{5} P_x = 1\,, \tag{3.271}$$

and therefore their differentials satisfy

$$\sum_{x=1}^{5} dP_x = 0\,. \tag{3.272}$$

Differentiating the $I(P,W)$ as a function of P yields[10] :

$$dI(P,W) = \sum_{x\in\mathcal{A}} D(W_x||PW)dP_x\,. \tag{3.273}$$

For $P = P^*$ we obtain that $dI = 0$ and for the Hessian, according to (3.224) and since $P_Y^*(y) = \frac{1}{5}$, we get

$$\mathcal{H}(P^*) = -5\log e \cdot W^T W\,, \tag{3.274}$$

and therefore by Taylor's formula:

$$I(P,W) = C - 5\log e \cdot ||(P-P^*)W||^2 + o(||P-P^*||^2)\,. \tag{3.275}$$

For $U(P,W)$ we have

$$dU(P,W) = \sum_{x\in\mathcal{A}} \left(V(W_x||PW) + D^2(W_x||PW) - 2\log e \cdot D(P_{X|Y}||P|W_x)\right) dP_x\,, \tag{3.276}$$

where we have denoted by $P_{X|Y}$ the following conditional distribution (note that it is a function of P):

$$P_{X|Y}(a|b) \stackrel{\triangle}{=} \frac{P(a)W(b|a)}{PW(b)}\,. \tag{3.277}$$

Similarly, for $V(P,W)$ we have

$$dV(P,W) = \sum_{x\in\mathcal{A}} \left(V(W_x||PW) + 2\log e \cdot \left[D(W_x||PW|Q_x) - D(P_{X|Y}||P||W_x)\right]\right) dP_x\,, \tag{3.278}$$

where the distribution Q_x on $\mathcal{A}$ is given by

$$Q_x(a) = \sum_{y\in\mathcal{B}} P_{X|Y}(a|y)W(y|x)\,. \tag{3.279}$$

For the points where $V(P,W) = 0$ expressions (3.276) and (3.278) simplify to:

$$dU(P,W) = \sum_{x\in\mathcal{A}} \left(V(W_x||PW) + D^2(W_x||PW)\right) dP_x\,, \tag{3.280}$$

$$dV(P,W) = \sum_{x\in\mathcal{A}} V(W_x||PW)dP_x\,. \tag{3.281}$$

Finally, in the present case of W given by (3.263) and $P = P^*$ we also have $D(W_x||PW) = C$ for all $x \in \mathcal{A}$ and $V(W_x||PW) = 0$ for all $x \neq 5$. Therefore, we get (using (3.272) again):

$$dU(P^*,W) = V(W_5||P_Y^*)dP_5\,, \tag{3.282}$$

[10]Henceforth we use (3.272) to simplify the expressions for the differentials; equivalently we write the pullback of differentials from $\mathbb{R}^{|\mathcal{A}|}$ to $\mathcal{P}$.

and from Taylor's formula we obtain then

$$U(P,W) = V(W_5||P_Y^*)P(5) + o(||P - P^*||)\,. \tag{3.283}$$

Finally, we choose the following sequence of input distributions

$$P_n = P^* + \frac{\lambda}{n^{\frac{1}{3}}}[0,0,0,-1,1]^T\,, \tag{3.284}$$

where $\lambda > 0$ is to be determined. From (3.275) and (3.283) we get for some constant $F_1 > 0$

$$I(P_n,W) = C - F_1\lambda n^{-\frac{2}{3}} + o\left(n^{-\frac{2}{3}}\right)\,, \tag{3.285}$$

$$U(P_n,W) = V(W_5||P_Y^*)\lambda n^{-\frac{1}{3}} + o\left(n^{-\frac{1}{3}}\right)\,. \tag{3.286}$$

Next, from the definition (3.105) of $T_u(P,W)$ and (3.285) we get for some $F_2 > 0$

$$T_u(P_n,W) = \frac{F_2\lambda}{n^{\frac{1}{3}}} + o\left(n^{-\frac{1}{3}}\right)\,. \tag{3.287}$$

Now we proceed as in the proof of Theorem 47 except that instead of (3.138) we define B as a function of n:

$$B = \frac{T_u(P_n,W)}{[U(P_n,W)]^{\frac{3}{2}}} = F_3\sqrt{\lambda}n^{\frac{1}{6}} + o\left(n^{\frac{1}{6}}\right)\,, \tag{3.288}$$

where $F_3 = \frac{F_2}{[V(W_5||P_Y^*)]^{\frac{3}{2}}} > 0$. Then following the proof of Theorem 47 we can show that for all n such that

$$\epsilon - \left(\frac{2\log 2}{\sqrt{2\pi}} + 5B\right)\frac{1}{\sqrt{n}} > 0\,, \tag{3.289}$$

we have

$$\log M_{avg}^*(n,\epsilon) \geq nI(P_n,W) - \sqrt{nU(P_n,W)}Q^{-1}\left(\epsilon - \left(\frac{2\log 2}{\sqrt{2\pi}} + 5B\right)\frac{1}{\sqrt{n}}\right)\,. \tag{3.290}$$

Note that the argument of Q^{-1} according to (3.288) satisfies

$$\epsilon - \left(\frac{2\log 2}{\sqrt{2\pi}} + 5B\right)\frac{1}{\sqrt{n}} = \epsilon - 5F_3\sqrt{\lambda}n^{-\frac{1}{3}} + o\left(n^{-\frac{1}{3}}\right)\,. \tag{3.291}$$

and, therefore, we have

$$Q^{-1}\left(\epsilon - \left(\frac{2\log 2}{\sqrt{2\pi}} + 5B\right)\frac{1}{\sqrt{n}}\right) = Q^{-1}(\epsilon) + O\left(n^{-\frac{1}{3}}\right)\,. \tag{3.292}$$

Finally, continuing (3.290) we get

$$\log M_{avg}^*(n,\epsilon) \geq nI(P_n,W) - \sqrt{nU(P_n,W)}\left(Q^{-1}(\epsilon) + O\left(n^{-\frac{2}{3}}\right)\right) \tag{3.293}$$

$$= nC - \left(F_1\lambda + \sqrt{V(W_5||P_Y^*)\lambda}Q^{-1}(\epsilon)\right)n^{\frac{1}{3}} + O(1)\,, \tag{3.294}$$

where (3.293) follows from (3.292) and (3.294) from (3.285) and (3.286).

Now, observe that for $\epsilon > 1/2$ we have $Q^{-1}(\epsilon) < 0$ and therefore for sufficiently small λ the coefficient in front of $n^{\frac{1}{3}}$ becomes negative. Choosing such λ, we have therefore shown that for some constant $F' > 0$ and all n sufficiently large we have

$$\log M^*_{avg}(n,\epsilon) \geq nC + F' n^{\frac{1}{3}} \,. \tag{3.295}$$

Finally, as explained in the proof of Theorem 47, changing $M^*_{avg}(n,\epsilon)$ to $M^*(n,\epsilon)$ results in a $O(\log n)$ penalty factor, and therefore, for some constant $0 < F_\epsilon < F'$ and all n sufficiently large (3.264) holds. ■

As Theorem 51 demonstrates the conditions for exotic channels are quite hard to satisfy (especially, making $D(W_x||P^*_Y) = C$ but so that x does not participate in capacity achieving distributions); hence the name exotic.

3.4.5 Refined results on the $\log n$ term

We define the following quantity

$$\begin{aligned} V^r(P,W) &\stackrel{\triangle}{=} \mathrm{Var}[i(X;Y)|Y] && (3.296)\\ &= \sum_{x,y} P(x)W(y|x)\Bigg[\log^2\frac{W(y|x)}{PW(y)} \\ &\qquad - \left(\sum_{x'}\frac{W(y|x')P(x')}{PW(y)}\log\frac{W(y|x')}{PW(y)}\right)^2\Bigg]\,. && (3.297)\end{aligned}$$

Some of its properties relevant for this section are given below:

Lemma 52 *All of the following hold:*

$$\begin{aligned} V^r(P,W) = 0 \quad &\iff \quad \forall x,y,y' : W(y'|x) = W(y|x) \text{ or } P(x)W(y'|x) = 0 && (3.298)\\ V^r(P,W) > 0 \quad &\implies \quad U(P,W) > 0 && (3.299)\\ \forall x,y : W(y|x) > 0 \quad &\implies \quad V^r(P,W) > 0 \text{ or } I(P,W) = 0\,. && (3.300)\end{aligned}$$

In words, (3.298) gives a neccesary and sufficient condition for $V^r(P,W) = 0$ which means that restricted to columns with $P(x) > 0$ submatrix W has each row composed of two elements only: zero and a (row-specific) constant; (3.299) and (3.300) give simpler necessary and a sufficient conditions, respectively, for $V^r(P,W) > 0$.

Proof: Without loss of generality we may assume that all outputs $y \in \mathcal{B}$ are reachable from at least one input. Then to show (3.298) notice that $\mathrm{Var}[i(X;Y)|Y] = 0$ holds if and only if for all (x,y) with $P(x)W(y|x) > 0$ we have that $\log\frac{W(y|x)}{PW(y)}$ is a function of y only. This is precisely the condition (3.298). Next, (3.299) follows trivially from

$$V^r(P,W) \leq U(P,W)\,, \tag{3.301}$$

which in turn follows from the definition and

$$\mathrm{Var}[A|B] \leq \mathrm{Var}[A]\,. \tag{3.302}$$

Finally, for (3.300) notice that if $W(y|x) > 0$ for all x, y and we had $V^r(P,W) = 0$ then by (3.298) we should have $W(y'|x) = W(y|x)$ for all x with $P(x) > 0$. But then we have $D(W_x||PW) = 0$ and thus $I(P,W) = 0$. ■

Next we show that the RCU bound, Theorem 17, implies a stronger achievability bound that the one in Theorem 47 (for average probability of error formalism).

Theorem 53 *For any input distribution P with $V^r(P,W) > 0$ we have*

$$\log M^*_{avg}(n,\epsilon) \geq nI(P,W) - \sqrt{nU(P,W)}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,. \tag{3.303}$$

The immediate corollary of this result is

Corollary 54 *Suppose that there exists a distribution P achieving V_ϵ in* (3.109) *with $V^r(P,W) > 0$ then we have*

$$\log M^*_{avg}(n,\epsilon) \geq nC - \sqrt{nV_\epsilon}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,. \tag{3.304}$$

In particular, any channel with $W(y|x) > 0$ for all x, y satisfies (3.304) *unless $C = 0$.*

Proof of Theorem 53: We define random variables $(X^n, Y^n, \bar{X}^n)$ distributed as

$$P_{X^nY^n\bar{X}^n}(x^n, y^n, \bar{x}^n) = \prod_{j=1}^n P(x_j)W(y_j|x_j)P(\bar{x}_j)\,. \tag{3.305}$$

Then according to the RCU bound, Theorem 17 there exists an (n, M, ϵ') code with

$$\epsilon' \leq \mathbb{E}\left[\min\left\{1,\ M\mathbb{P}\left[\bar{I}_n \geq I_n \,\middle|\, X^n, Y^n\right]\right\}\right]\,, \tag{3.306}$$

where

$$\bar{I}_n \stackrel{\triangle}{=} \sum_{k=1}^n i(\bar{X}_k, Y_k)\,, \tag{3.307}$$

$$I_n \stackrel{\triangle}{=} \sum_{k=1}^n i(X_k, Y_k)\,, \tag{3.308}$$

$$i(x,y) \stackrel{\triangle}{=} \begin{cases} \log\frac{W(y|x)}{PW(y)}\,, & W(y|x) > 0\,, \\ -\infty\,, & W(y|x) = 0\,, \end{cases} \tag{3.309}$$

and thus $I_n > -\infty$ almost surely. Introduce the following function

$$f(t, y^n) \stackrel{\triangle}{=} \mathbb{P}\left[\bar{I}_n \geq t \middle| Y^n = y^n\right]\,. \tag{3.310}$$

Then since $\bar{X}^n$ is independent of X^n we have

$$\mathbb{P}\left[\bar{I}_n \geq I_n \,\middle|\, X^n, Y^n\right] = f(I_n, Y^n)\,. \tag{3.311}$$

Notice that for any realization $\bar{x}^n$ of $\bar{X}^n$ such that $\bar{I}_n > -\infty$ we have

$$\begin{aligned} \mathbb{P}[\bar{X}^n = \bar{x}^n | Y^n] &= \mathbb{P}[\bar{X}^n = x^n] && (3.312) \\ &= \mathbb{P}[X^n = x^n | Y^n] \exp\left\{ - \sum_{k=1}^{n} i(x_k, Y_k) \right\} && (3.313) \end{aligned}$$

and thus summing over all $\bar{x}^n$ such that $\bar{I}_n \geq t$ we obtain

$$\mathbb{P}\left[\bar{I}_n \geq t \,\middle|\, Y^n \right] = \mathbb{E}\left[\exp\{-I_n\} 1\{I_n \geq t\} \,|\, Y^n \right]. \tag{3.314}$$

Now observe that conditioned on Y^n random variable I_n is a sum of independent (non identically distributed) random variables; its variance is given by

$$\begin{aligned} \mathrm{Var}[I_n | Y^n = y^n] &= \sum_{k=1}^{n} S_{y_k}, && (3.315) \\ S_y &\stackrel{\triangle}{=} \mathrm{Var}[i(X_1, Y_1) | Y_1 = y]. && (3.316) \end{aligned}$$

Notice that

$$\mathbb{E}\left[S_{Y_1} \right] = V^r(P, W) > 0. \tag{3.317}$$

Denote by F the following event:

$$F \stackrel{\triangle}{=} \left\{ \sum_{k=1}^{n} S_{Y_k} \geq \frac{1}{2} V^r(P, W) \right\}. \tag{3.318}$$

Then, on the one hand by Chernoff bound we have for some $K_1 > 0$

$$\mathbb{P}[F^c] \leq \exp\{-K_1 n\}, \tag{3.319}$$

while on the other hand by Lemma 20 we have for some $K_2 > 0$ on the event F:

$$\mathbb{E}\left[\exp\{-I_n\} 1\{I_n \geq t\} \,|\, Y^n \right] \leq \frac{K_2}{\sqrt{n}} \exp\{-t\} \quad (\text{on } F). \tag{3.320}$$

Thus, we have

$$\begin{aligned} &\mathbb{E}\left[\min\left\{1, M f(I_n, Y^n)\right\} \right] && (3.321) \\ &\leq \mathbb{P}[F^c] + \mathbb{E}\left[\min\left\{ 1, \frac{MK_2}{\sqrt{n}} \exp\{-I_n\} \right\} \right] && (3.322) \\ &\leq \exp\{-K_1 n\} + \mathbb{E}\left[\min\{1, \frac{MK_2}{\sqrt{n}} \exp\{-I_n\} \right] && (3.323) \\ &\leq \exp\{-K_1 n\} + \mathbb{P}\left[I_n \leq \log \frac{MK_2}{\sqrt{n}} \right] \\ &\quad + \frac{MK_2}{\sqrt{n}} \mathbb{E}\left[1\left\{ I_n > \log \frac{MK_2}{\sqrt{n}} \right\} \exp\{-I_n\} \right] && (3.324) \\ &\leq \exp\{-K_1 n\} + \mathbb{P}\left[I_n \leq \log \frac{MK_2}{\sqrt{n}} \right] + \frac{K_3}{\sqrt{n}} && (3.325) \\ &\leq \exp\{-K_1 n\} + Q\left(\frac{nI(P,W) - \log \frac{MK_2}{\sqrt{n}}}{\sqrt{nU(P,W)}} \right) + \frac{K_3 + K_4}{\sqrt{n}}, && (3.326) \end{aligned}$$

where (3.322) is by (3.314) and (3.320), (3.323) is by (3.319), (3.324) simply expands the min, (3.325) holds for a suitable $K_3 > 0$ by Lemma 20 and (3.326) holds for a suitable $K_4 > 0$ by Berry-Esseen inequality (Theorem 13) both of which are applicable since $U(P,W) > 0$. Equating the right-hand side of (3.326) to ϵ and solving for M, we obtain for all sufficiently large n existence of an (n, M, ϵ) code with

$$\begin{aligned} \log M &= nI(P,W) - \sqrt{nU(P,W)}Q^{-1}\left(\epsilon - \frac{K_3+K_4}{\sqrt{n}} - \exp\{-K_1 n\}\right) \\ &\quad + \frac{1}{2}\log n - \log K_2 \qquad (3.327) \\ &= nI(P,W) - \sqrt{nU(P,W)}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,, \qquad (3.328) \end{aligned}$$

where (3.328) is by Taylor's expansion applied to $Q^{-1}(\cdot)$. This proves (3.303). ■

Remarks:

1. The estimate of $\frac{1}{2}\log n$ cannot be improved without further assumptions, as the example of the BSC shows; see Theorem 41.

2. According to (3.299) the right-hand side of (3.304) can never be of the form $nC + \frac{1}{2}\log n + O(1)$.

3. As we mentioned in the discussion of the BSC, Section 3.2.3, other bounds from Chapter 2 fail to achieve $\frac{1}{2}\log n$ term. In particular, using techniques from Chapter 2 it is not possible via our techniques to give an extension of Theorem 53 and Corollary 54 to maximal probability of error formalism (unless in some special cases). The reason we could give such result for the BSC is because of the appeal to random linear code method, which requires strong assumptions on the cardinalities of $|\mathcal{A}|$, $|\mathcal{B}|$ and the structure of W.

4. According to (3.298) the situation $V^r(P,W) = 0$ requires rather special structure of W and thus Corollary 54 holds for almost all channels (BEC being a notable exception).

5. In the case when $V^r(P,W) = 0$, we have

$$i(x,y) = \mathbb{E}\left[i(X_1,Y_1)|Y_1 = y\right] = -\log P[W(y|X_1) > 0]\,, \qquad (3.329)$$

which implies that I_n is a function of Y^n and $I_n = \bar{I}_n$ (unless $\bar{I}_n = -\infty$, but this case is irrelevant: see (3.314)). Thus, the RCU bound in this case yields

$$\epsilon' \le \mathbb{E}\left[\min\{1, M\exp\{-I_n\}\}\right] = \mathbb{P}[I_n \le \log M] + O\left(\frac{1}{\sqrt{n}}\right). \qquad (3.330)$$

Therefore, we can only achieve:

$$\log M^*_{avg}(n,\epsilon) \ge nI(P,W) - \sqrt{nU(P,W)}Q^{-1}(\epsilon) + O(1)\,, \qquad (3.331)$$

which is not interesting, since it is implied by (3.112).

6. To summarize, for the average probability of error Theorem 53 and Corollary 54 give the strongest possible results that can be derived using any bounds from Chapter 2.

We give a matching converse bound on $\log n$ term under an assumption of symmetry.

Definition 9 *A DMC W is called weakly input-symmetric if there exists an $x_0 \in \mathcal{A}$ and a random transformation $T_x : \mathcal{B} \to \mathcal{B}$ for each $x \in \mathcal{A}$ such that $T_x \circ W_{x_0} = W_x$ and $T_x \circ P_Y^*$, where P_Y^* is the capacity achieving output distribution.*

Note that the composition $T_x \circ P_Y$ with a distribution P_Y on $\mathcal{B}$, according to (2.2), is given by

$$(T_x \circ P_Y)(y) = \sum_{y' \in \mathcal{B}} T_x(y|y') P_Y(y') \,. \tag{3.332}$$

Thus, in other words, T_x is a stochastic matrix which upon multiplication by the column W_{x_0} yields the column W_x.

Examples:

1. Recall that Gallager [9, p. 94] defines the channel to be symmetric if the space of outputs $\mathcal{B}$ can be partitioned into disjoint subsets $\mathcal{B} = \bigcup_{j=1}^{d} \mathcal{B}_j$ such that each restriction of W to $\mathcal{A} \times \mathcal{B}_j$ has rows which are all permutations of each other and columns which are permutations of each other. It is easy to see that Gallager-symmetric channels are weakly input-symmetric.

2. However, not all weakly input-symmetric channels are Gallager-symmetric. Indeed, consider the following channel

$$W = \begin{pmatrix} 1/7 & 4/7 & 1/7 & 1/7 \\ 4/7 & 1/7 & 0 & 4/7 \\ 0 & 0 & 4/7 & 2/7 \\ 2/7 & 2/7 & 2/7 & 0 \end{pmatrix} . \tag{3.333}$$

   Since $\det W \neq 0$, the capacity achieving input distribution is unique. Since $H(Y|X = x)$ is independent of x and $P_X = [1/4, 1/4, 3/8, 1/8]$ achieves uniform P_Y^* it must be the unique optimum. Clearly any permutation T_x fixes a uniform P_Y^* and thus the channel is weakly input-symmetric. At the same time it is not Gallager-symmetric since no row is a permutation of another.

3. Allowing randomized (i.e. not induced by functional maps) kernels T_x in the Definition 9 is essential. Indeed, consider the channel

$$W = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & 0 & 1/4 \\ 0 & 1/2 & 1/4 \end{pmatrix} \tag{3.334}$$

   Clearly this channel is not Gallager-symmetric. To show it is weakly input-symmetric notice that by interchanging first two inputs (columns 1,2) and last two outputs (rows 2,3) we do not change the matrix. Thus $P_Y^*(2) = P_Y^*(3)$. Now take $x_0 = 1$ (corresp.,

first column), T_2 to be induced by the permutation that maps $1, 2, 3$ to $1, 3$ and 2, respectively. Finally, for T_3 we take the following matrix

$$T_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \tag{3.335}$$

To check that $T_3 \circ W(\cdot|1) = W(\cdot|3)$ simply write

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/4 \\ 1/4 \end{pmatrix} \tag{3.336}$$

Finally, we have shown that $P_Y^*(2) = P_Y^*(3)$ and thus $T_3 \circ P_Y^* = P_Y^*$ and $T_2 \circ P_Y^* = P_Y^*$ as required.

4. At the same time not all channels are weakly input-symmetric. Indeed, composition with a kernel yields a distribution dominated by the original: $T_x \circ W_{x_0} \le W_{x_0}$ (partial order is that of majorization). Thus to obtain a non weakly input-symmetric channel it is sufficient to take any channel without a column that dominates all others. A simpler example is the Z-channel:

$$W = \begin{pmatrix} 1-\delta & 0 \\ \delta & 1 \end{pmatrix}, \quad \delta \in (0, 1) \tag{3.337}$$

Indeed, clearly we cannot take input 1 as x_0 (since then T_2 would have to map both outputs to the second output and thus $T_2 \circ P_Y^* = [0\,1]$ – a contradiction). By computing P_Y^* it can be shown that taking $x_0 = 2$ there is no stochastic matrix T_1 that fixes P_Y^* and maps $[0, 1]$ to $[1-\delta, \delta]$.

Some of the crucial properties of weakly input-symmetric channels are summarized below:

Theorem 55 *For any weakly input-symmetric DMC W all of the following hold:*

1. *The capacity C satisfies:*
$$C = D(W_{x_0}||P_Y^*)\,, \tag{3.338}$$

2. *The ϵ-dispersion V_ϵ equals the dispersion V and satisfies*
$$\begin{aligned} V &= V(W_{x_0}||P_Y^*) && (3.339)\\ &= V(W_x||P_Y^*) \quad (\forall x : D(W_x||P_Y^*) = C)\,. && (3.340) \end{aligned}$$

3. *The following bound holds*[11]
$$\log M_{avg}^*(n, \epsilon) \le -\log \beta_{1-\epsilon}((W_{x_0})^n, (P_Y^*)^n)\,. \tag{3.341}$$

[11] In Section 6.5 we will also show that (3.341) holds (for each n and ϵ) even in the presence of instanteneous noiseless feedback. Consequently (3.342) (or (3.343)) is also valid even with feedback.

4. In particular, if $V > 0$ then as $n \to \infty$ we have

$$\log M^*_{avg}(n,\epsilon) \le nC - \sqrt{nV}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,. \tag{3.342}$$

If $V = 0$ then we have

$$\log M^*_{avg}(n,\epsilon) \le nC - \log(1-\epsilon)\,. \tag{3.343}$$

Proof: To show (3.338) notice that a transformation T_x maps pair of distributions (W_{x_0}, P^*_Y) to (W_x, P^*_Y) and therefore by the data processing for divergence we get

$$D(W_x||P^*_Y) \le D(W_{x_0}||P^*_Y)\,, \tag{3.344}$$

from which (3.338) follows via

$$C = \max_{x\in\mathcal{A}} D(W_x||P^*_Y)\,. \tag{3.345}$$

Similarly, for any distribution P_{X^n} on $\mathcal{A}^n$ we have

$$\beta_\alpha(P_{X^nY^n}, P_{X^n}(P^*_Y)^n) \ge \beta_\alpha((W_{x_0})^n, (P^*_Y)^n)\,. \tag{3.346}$$

Indeed, for each x^n define a random transformation $T_{x^n} : \mathcal{B}^n \to \mathcal{B}^n$ as follows:

$$T_{x^n}(z^n|y^n) = \prod_{k=1}^n T_{x_k}(z_k|y_k)\,. \tag{3.347}$$

Then T_{x^n} maps pair of distributions $(W^n_{x_0}, (P_{Y^*})^n)$ to $(P_{Y^n|X^n=x^n}, (P^*_Y)^n)$ and thus by the data-processing for β_α we obtain

$$\beta_\alpha(P_{Y^n|X^n=x^n}, (P^*_Y)^n) \ge \beta_\alpha(W^n_{x_0}, (P_{Y^*})^n)\,. \tag{3.348}$$

Therefore, (3.346) follows by Lemma 32 and convexity of β_α in α. Consequently, (3.341) then follows from Theorem 29 and (3.346), while (3.342) and (3.343) follow from (2.89) and (2.90), respectively.

Finally, to show (3.339) notice that by Lemma (14) we have for any $x \in \mathcal{A}$:

$$\log\beta_\alpha((W_x)^n, (P^*_Y)^n) = -nD(W_x||P^*_Y) - \sqrt{nV(W_x||P^*_Y)}Q^{-1}(\alpha) + o(\sqrt{n})\,. \tag{3.349}$$

But by (3.348) we must have

$$\log\beta_\alpha((W_x)^n, (P^*_Y)^n) \ge \log\beta_\alpha((W_{x_0})^n, (P^*_Y)^n)\,. \tag{3.350}$$

Now assuming that $x \in \mathcal{A}$ is such that $D(W_x||P^*_Y) = C$ and applying (3.349) to both sides of (3.350) for $\alpha > 1/2$ we obtain

$$V(W_x||P^*_Y) \ge V(W_{x_0}||P^*_Y)\,, \tag{3.351}$$

whereas taking $\alpha < 1/2$ we show

$$V(W_x||P^*_Y) \le V(W_{x_0}||P^*_Y)\,, \tag{3.352}$$

and consequently (3.339) follows. ■

Corollary 56 *Consider a weakly input-symmetric channel W with $V > 0$. If there exists a capacity achieving input distribution with $V^r(P, W) > 0$ then we have*

$$\log M^*_{avg}(n, \epsilon) = nC - \sqrt{nV}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,. \tag{3.353}$$

In particular, by (3.300) any Gallager-symmetric channel without zeros in W and positive capacity satisfies conditions of Corollary 56. Note that there exist weakly input-symmetric channels with $V > 0$ but $V^r(P, W) = 0$ (for example, the BEC or the q-ary erasure channel). For such channels, a converse bound different from Theorem 55 is likely to be needed in order to pin down the $\log n$ term.

Corollary 57 *Consider a weakly input-symmetric channel W with $V = 0$. Then we have*

$$nC - \log\frac{1}{\epsilon} \le \log M^*(n, \epsilon) \le \log M^*_{avg}(n, \epsilon) \le nC + \log\frac{1}{1-\epsilon} \tag{3.354}$$

Proof: Apply (3.133) and (3.343). ■

Somewhat unexpectedly, this corollary shows that we can obtain an almost exact value for $\log M^*(n, \epsilon)$ for some non-trivial channels, such as

$$W = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}\,. \tag{3.355}$$

3.4.6 Applications to other questions

In this section we discuss application of the methods developed in Chapter 2: the $\kappa\beta$ bound and the meta-converse – to some other questions of channel coding for the DMC.

As we have mentioned, the crucial fact for proving Theorem 45 was that $V(P, W)$ coincides with $U(P, W)$ for the capacity achieving distributions: the DT bound yields a lower bound on $\sqrt{n}$ term as $\sqrt{nU(P, W)}$, while Theorem 34 upper-bounds the $\sqrt{n}$ term by $\sqrt{nV(P, W)}$; since they coincide we are able to obtain the exact coefficient for the $\sqrt{n}$ term. The situation, however, is different if we want to obtain a lower-bound on $\log M^*_{P_0}(n, \epsilon)$ for the cardinality of the best constant-composition code, since in general $V(P_0, W) \neq U(P_0, W)$.

The problem is resolved by using the $\kappa\beta$ bound, Theorem 27. Here we briefly sketch how to apply Theorem 27. We choose $\mathsf{F}_n = T^n_{P_0}$ for a fixed type $P_0 \in \mathcal{P}_n$. We also choose $P_Y(y) = P_0 W$. Then, all the work reduces to lower-bounding $\kappa^n_\tau(\mathsf{A}, P_Y)$ and analyzing $\beta^n_\alpha(x, P_Y)$. The analysis of the latter has already been done in the proof of Theorem 50; specifically, we have shown

$$\log \beta^n_\alpha(x, P_Y) = -nI(P_0, W) + \sqrt{nV(P_0, W)}Q^{-1}(\epsilon) + O(\log n) \tag{3.356}$$

for any $x \in T^n_{P_0}$. For κ^n_τ the lower bound is obtained by applying (2.93):

$$\kappa^n_\tau(T^n_{P_0}, P_Y) \ge P_0\left[T^n_{P_0}\right]\,. \tag{3.357}$$

Since the right-hand side is polynomial in n, see [16], this bound is sufficient for our purposes. So, by $\kappa\beta$ bound we conclude that at least

$$\log M^*_{P_0}(n,\epsilon) \geq nI(P_0,W) - \sqrt{nV(P_0,W)}Q^{-1}(\epsilon) + O(\log n) \tag{3.358}$$

codewords from type $T^n_{P_0}$. Together with Theorem 50, we have then

$$\log M^*_{P_0}(n,\epsilon) = nI(P_0,W) - \sqrt{nV(P_0,W)}Q^{-1}(\epsilon) + O(\log n)\,. \tag{3.359}$$

So we see that $\kappa\beta$ bound gives the $\sqrt{n}$ term with $V(P_0,W)$ compared to a generally looser $U(P_0,W)$ for the DT bound. However, for the case when P_0 is capacity achieving (and $U(P_0,W) = V(P_0,W)$), the DT bound is tighter due to a better $\log n$ term.

An argument entirely similar to (3.359) applies to other cases of channels with input constraints, and shows that $\kappa\beta$ bound emerges as a natural candidate for the dispersion achievability bounds in such cases (see the treatment of Gaussian channels in Chapter 4 and Section 4.2 in particular).

Another application concerns the converse bound. Suppose that we are given information regarding the number of codewords inside each type $P_0 \in \mathcal{P}_n$, that is we know the value of $P_{X^n}(T_{P_0})$ for each $P_0 \in \mathcal{P}_n$, where P_{X^n} denotes the distribution on $\mathcal{A}^n$ induced by the code.[12] Can we upper-bound the average probability of error for such a code?

We aim to apply the meta-converse, Theorem 28, with the following Q-channel:

$$Q_{Y^n|X^n=x^n} = (P_{x^n}W)^n\,, \tag{3.360}$$

where P_{x^n} is the type of sequence x^n. Notice that the output distribution depends only on the type of the input. Therefore, we cannot distinguish more than $|\mathcal{P}_n|$ alternatives and we have:

$$\epsilon' \geq 1 - \frac{|\mathcal{P}_n|}{M}\,. \tag{3.361}$$

Then following the derivation of the Fano's inequality (2.251) in Section 2.7.3, we obtain

$$\frac{\log M}{n} \leq \frac{1}{1-\epsilon}\sum_{P_0\in\mathcal{P}_n} P_{X^n}(T_{P_0})I(P_0,W) + \frac{1}{n}\frac{h(\epsilon)}{1-\epsilon} + \frac{\log|\mathcal{P}_n|}{n}\,. \tag{3.362}$$

A standard bound on the cardinality $|\mathcal{P}_n|$, see [16], states

$$|\mathcal{P}_n| \leq (n+1)^{|\mathcal{A}|-1}\,. \tag{3.363}$$

Therefore, from (3.362) we obtain the following:

$$\frac{\log M}{n} \leq \frac{1}{1-\epsilon}\sum_{P_0\in\mathcal{P}_n} P_{X^n}(T_{P_0})I(P_0,W) + \frac{1}{n}\frac{h(\epsilon)}{1-\epsilon} + (|\mathcal{A}|-1)\frac{\log(n+1)}{n}\,. \tag{3.364}$$

Inequality (3.364) bounds possible parameters of (n,M,ϵ) codes with a given type-distribution of codewords. In particular, we see that the "capacity" for the codes with a given type-distribution is limited by

$$\tilde{C}(W) = \sum_{P_0\in\mathcal{P}_n} P_{X^n}(T_{P_0})I(P_0,W) \leq C\,. \tag{3.365}$$

[12]The motivation is clear: for linear codes over binary-input channels, $P_{X^n}(T_{P_0})$ is simply a weight distribution of the code, which is frequently known from algebraic considerations.

In other words, the code under consideration can have a small probability of error only for those channels $W : \mathcal{A} \to \mathcal{B}$ that have $\tilde{C}(W) \geq R$, where R is the rate of the code. This gives a rough estimate of what is possible with a given type-distribution.

Recall that in Section 2.7.3 we obtained stronger bounds (such as sphere packing) by replacing a simple data-processing lower-bound on β_α (that lead to Fano's inequality) with better bounds. Similarly, it is simple to obtain stronger versions of the bound (3.364).

3.5 Gilbert-Elliott channel (GEC)

In this section we discuss a binary symmetric channel, with crossover probability being determined by a binary Markov chain. This channel with memory is known as the Gilbert-Elliott channel (GEC); see [62, 63].

This channel is particularly interesting because the dynamics of the crossover probability can be viewed as a simplified model of a fading channel, where fading coefficients evolve as a Markov process. It is known [67], that for coherent channels behaving ergodically, channel capacity is independent of the fading dynamics since a sufficiently long codeword sees a channel realization whose empirical statistics have no randomness. The natural question arises: does the channel dispersion depend on channel dynamics? And if so, does it correctly predict the dependence of fundamental limits on channel dynamics? In this section we answer both questions affirmatively, thereby demonstrating how knowledge of channel dispersion can qualitatively extend our understanding of the channel, and enable the analysis of questions for which the capacity alone is insufficient.

3.5.1 Channel capacity

At blocklength n the channel $GEC(n, \tau, \delta_1, \delta_2, p_1)$, depending on $0 \leq \tau, \delta_1, \delta_2, p_1 \leq 1$ is defined as follows. The input and output alphabets are $\mathsf{A} = \mathsf{B} = \{0, 1\}^n$ and the channel acts on an input binary vector X^n by adding (modulo 2) the vector Z^n:

$$Y^n = X^n + Z^n \,, \tag{3.366}$$

where the distribution of Z^n is specified as follows.

Let $\{S_j\}_{j=1}^\infty$ be a homogeneous Markov process with states $\{1, 2\}$ and transition probabilities[13]

$$\mathbb{P}[S_2 = 1|S_1 = 1] = \mathbb{P}[S_2 = 2|S_1 = 2] = 1 - \tau \,, \tag{3.367}$$
$$\mathbb{P}[S_2 = 2|S_1 = 1] = \mathbb{P}[S_2 = 1|S_1 = 2] = \tau \,. \tag{3.368}$$

Now for $0 \leq \delta_1, \delta_2 \leq 1$ we define $\{Z_j\}_{j=1}^\infty$ as conditionally independent given $\{S_j\}_{j=1}^\infty$ and

$$\mathbb{P}[Z_j = 0|S_j = s] = 1 - \delta_s \,, \tag{3.369}$$
$$\mathbb{P}[Z_j = 1|S_j = s] = \delta_s \,. \tag{3.370}$$

[13]The results in this section can be readily generalized at the expense of more cumbersome expressions to Gilbert-Elliott channels with asymmetric Markov chains.

The description of the channel model is incomplete without specifying the distribution of S_1:

$$\mathbb{P}[S_1 = 1] = p_1, \tag{3.371}$$
$$\mathbb{P}[S_1 = 2] = p_2 = 1 - p_1\,. \tag{3.372}$$

In this way the Gilbert-Elliott channel is completely specified by the parameters $(\tau, \delta_1, \delta_2, p_1)$.

When $\tau > 0$ the chain S_1 is ergodic and for this reason we consider only the stationary case $p_1 = 1/2$. On the other hand, when $\tau = 0$ the mode of operation changes drastically (channel becomes non-ergodic). This special case (and arbitrary p_1) is considered in Section 3.6.

In addition, there are two practically possible scenarios: 1) the state sequence S^n is known perfectly at the receiver and 2) no state information. The capacity C_1 of a Gilbert-Elliott channel $\tau > 0$ and state S^n known perfectly at the receiver depends only on the stationary distribution P_{S_1} and is given by

$$C_1 = \log 2 - \mathbb{E}\left[h(\delta_{S_1})\right] \tag{3.373}$$
$$= \log 2 - \mathbb{P}[S_1 = 1]h(\delta_1) - \mathbb{P}[S_1 = 2]h(\delta_2)\,, \tag{3.374}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy function. In the symmetric-chain special case considered in this section, both states are equally likely and

$$C_1 = \log 2 - \frac{1}{2}h(\delta_1) - \frac{1}{2}h(\delta_2). \tag{3.375}$$

When $\tau > 0$ and state S^n is not known at the receiver, the capacity is given by [64]

$$C_0 = \log 2 - \mathbb{E}\left[h(\mathbb{P}[Z_0 = 1|Z_{-\infty}^{-1}])\right] \tag{3.376}$$
$$= \log 2 - \lim_{n\to\infty} \mathbb{E}\left[h(\mathbb{P}[Z_0 = 1|Z_{-n}^{-1}])\right]\,. \tag{3.377}$$

Throughout the section we use subscripts $\mathsf{1}$ and $\mathsf{0}$ for capacity and dispersion to denote the cases when the state S^n is known and is not known, respectively.

3.5.2 Asymptotic expansion

Theorem 58 *Suppose that the state sequence S^n is stationary, $\mathbb{P}[S_1 = 1] = 1/2$, and ergodic, $0 < \tau < 1$. Then the dispersion of the Gilbert-Elliott channel with state S^n known at the receiver is*

$$V_1 = \frac{1}{2}(V(\delta_1) + V(\delta_2)) + \frac{1}{4}\left(h(\delta_1) - h(\delta_2)\right)^2\left(\frac{1}{\tau} - 1\right)\,, \tag{3.378}$$

where $V(\delta)$ is the dispersion of the BSC; see (3.30)*. Furthermore, provided that $V_1 > 0$ and regardless of whether $0 < \epsilon < 1$ is a maximal or average probability of error we have*

$$\log M^*(n,\epsilon) = nC_1 - \sqrt{nV_1}Q^{-1}(\epsilon) + O(\log n)\,, \tag{3.379}$$

where C_1 is given in (3.375)*. Moreover,* (3.379) *holds even if the transmitter knows the full state sequence S^n in advance (i.e., non-causally).*

Note that the condition $V_1 > 0$ for (3.379) to hold excludes only some degenerate cases for which we have: $M^*(n,\epsilon) = 2^n$ (when both crossover probabilities are 0 or 1) or $M^*(n,\epsilon) = \lfloor \frac{1}{1-\epsilon} \rfloor$ (when $\delta_1 = \delta_2 = 1/2$).

To formulate the result for the case of no state information at the receiver, we define the following stationary process:

$$F_j = -\log P_{Z_j|Z_{-\infty}^{j-1}}(Z_j|Z_{-\infty}^{j-1})\,. \tag{3.380}$$

Theorem 59 *Suppose that $0 < \tau < 1$ and the state sequence S^n is started at the stationary distribution. Then the dispersion of the Gilbert-Elliott channel with no state information is*

$$V_0 = \mathrm{Var}\left[F_0\right] + 2\sum_{i=1}^{\infty} \mathbb{E}\left[(F_i - \mathbb{E}\left[F_i\right])(F_0 - \mathbb{E}\left[F_0\right])\right]\,. \tag{3.381}$$

Furthermore, provided that $V_0 > 0$ and regardless of whether ϵ is a maximal or average probability of error, we have

$$\log M^*(n,\epsilon) = nC_0 - \sqrt{nV_0}Q^{-1}(\epsilon) + o(\sqrt{n})\,, \tag{3.382}$$

where C_0 is given by (3.376)*.*

It can be shown that the process F_j has a spectral density $S_F(f)$, and that [68]

$$V_0 = S_F(0)\,, \tag{3.383}$$

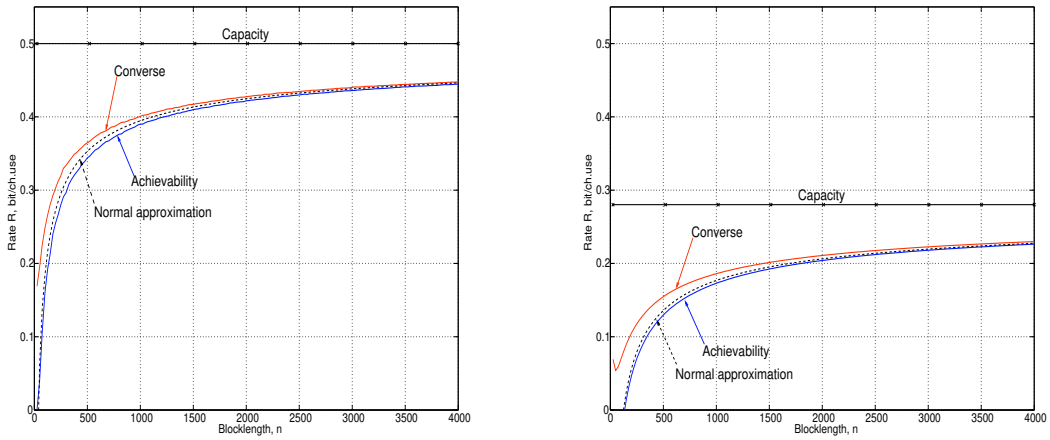
which provides a way of computing V_0 by Monte Carlo simulation paired with a spectral estimator. Alternatively, since the terms in the series (3.381) decay as $(1-2\tau)^j$, it is sufficient to compute only finitely many terms in (3.381) to achieve any prescribed approximation accuracy. In this regard note that each term in (3.381) can in turn be computed with arbitrary precision by noting that $P_{Z_j|Z_{-\infty}^{j-1}}[1|Z_{-\infty}^{j-1}]$ is a Markov process with a simple transition kernel.

Regarding the computation of C_0 it was shown in [64] that

$$\log 2 - \mathbb{E}\left[h(\mathbb{P}[Z_j = 1|Z^{j-1}])\right] \le C_0 \le \log 2 - \mathbb{E}\left[h(\mathbb{P}[Z_j = 1|Z^{j-1}, S_0])\right], \tag{3.384}$$

where the bounds are asymptotically tight as $j \to \infty$. The computation of the bounds in (3.384) is challenging because the distributions of $\mathbb{P}[Z_j = 1|Z_1^{j-1}]$ and $\mathbb{P}[Z_j = 1|Z_1^{j-1}, S_0]$ consist of 2^j atoms and therefore are impractical to store exactly. Rounding off the locations of the atoms to fixed quantization levels inside interval $[0,1]$, as proposed in [64], leads in general to unspecified precision. However, for the special case of $\delta_1, \delta_2 \le 1/2$ the function $h(\cdot)$ is monotonically increasing in the range of values of its argument and it can be shown that rounding down (up) the locations of the atoms shifts the locations of all the atoms on subsequent iterations down (up). Therefore, if rounding is performed this way, the quantized versions of the bounds in (3.384) are also guaranteed to sandwich C_0.

Proofs of Theorems 58 and 59 are given in Appendix E, respectively. Here we only make a few remarks:



(a) State S^n known at the receiver

(b) No state information

Figure 3.10: Rate-blocklength tradeoff at block error rate $\epsilon = 10^{-2}$ for the Gilbert-Elliott channel with parameters $\delta_1 = 1/2$, $\delta_2 = 0$ and state transition probability $\tau = 0.1$.

1. Proofs rely on the DT bound, Theorem 18, as their main achievability tool. Indeed, among the available achievability bounds, Gallager's random coding bound, Theorem 3, does not yield the correct dispersion term even for memoryless channels; Shannon's (or Feinstein's) bound, Theorem 2 is always weaker than the DT bound; and the RCU bound, Theorem 17, is harder to specialize to the channels considered in this section.

2. The converse bounds are given by the meta-converse, Theorem 28. It is interesting to notice that it is the generality of Theorem 28 (namely the fact that it holds for randomized encoders) that enables the extension to the case of state known at the transmitter.

3. We were unable to pin down the pre-log coefficient in (3.379). It is likely that doing so will require advances in the field of Berry-Esseen inequalities for the mixing processes.

3.5.3 Discussion and numerical comparisons

The natural application of expansions (2.22) is in approximating the maximal achievable rate according to (2.23). Unlike the BSC case, Theorem 41, the coefficient of the $\log n$ term (or "prelog") for the GEC is unknown. However, due to the fact that $\frac{1}{2}\log n$ in (3.59) is robust to variation in crossover probability, it is natural to conjecture that the unknown prelog for GEC is also $\frac{1}{2}$. With this choice, we arrive to the following approximation which will be used for numerical comparison:

$$\frac{1}{n}\log M^*(n,\epsilon) \approx C - \sqrt{\frac{V}{n}}Q^{-1}(\epsilon) + \frac{1}{2n}\log n\,, \tag{3.385}$$

Table 3.1: Capacity and dispersion for the Gilbert-Elliott channels in Fig. 3.10

State information	Capacity	Dispersion
known	0.5 bit	2.25 bit^2
unknown	0.280 bit	2.173 bit^2

Parameters: $\delta_1 = 1/2, \delta_2 = 0, \tau = 0.1$.

with $(C, V) = (C_1, V_1)$, when the state is known at the receiver, and $(C, V) = (C_0, V_0)$, when the state is unknown.

The approximation in (3.385) is obtained through new non-asymptotic upper and lower bounds on the quantity $\frac{1}{n} \log M^*(n, \epsilon)$, which are given in Appendix E for both cases. The asymptotic analysis of those bounds led to the approximation (3.385). It is natural to compare those bounds with the analytical two-parameter approximation (3.385). Such comparison is shown in Fig. 3.10. For the case of state known at the receiver, Fig. 3.10(a), the achievability bound is (E.48) and the converse bound is (E.65). For the case of unknown state, Fig. 3.10(b), the achievability bound is (E.102) and the converse is (E.118). The achievability bounds are computed for the maximal probability of error criterion, whereas the converse bounds are for the average probability of error. The values of capacity and dispersion, needed to evaluate (3.385), are summarized in Table 3.1.

Two main conclusions can be drawn from Fig. 3.10. First, we see that our bounds are tight enough to get an accurate estimate of $\frac{1}{n} \log M^*(n, \epsilon)$ even for moderate blocklengths n. Second, knowing only two parameters, capacity and dispersion, leads to approximation (3.385), which is precise enough for addressing the finite-blocklength fundamental limits even for rather short blocklengths. Both of these conclusions have already been observed in Sections 3.2.3 and 3.3.3 for memoryless channels.

In general, as $\tau \to 0$ the state availability at the receiver does not affect either the capacity or the dispersion too much as the following result demonstrates.

Theorem 60 *Assuming* $0 < \delta_1, \delta_2 \le 1/2$ *and* $\tau \to 0$ *we have*

$$C_0(\tau) \geq C_1 - O(\sqrt{-\tau \ln \tau}), \tag{3.386}$$

$$C_0(\tau) \leq C_1 - O(\tau), \tag{3.387}$$

$$V_0(\tau) = V_1(\tau) + O\left(\left[\frac{-\ln \tau}{\tau}\right]^{3/4}\right) \tag{3.388}$$

$$= V_1(\tau) + o(1/\tau) . \tag{3.389}$$

The proof is provided in Appendix E. Theorem 60 is useful for two related reasons. First, the evaluation of V_0 based on the definition (3.381) is quite challenging[14] , whereas the proof of Theorem 60 develops upper and lower bounds on V_1; see Lemma 123 in Appendix E. Second, Theorem 60 shows that for small values of τ one can approximate the unknown value of V_0 with V_1 given by (3.378) in closed form. Table 3.1 illustrates that such approximation happens to be rather accurate even for moderate values of τ.

[14]Observe that even analyzing $\mathbb{E}[F_j]$, the entropy rate of the hidden Markov process Z_j, is nontrivial; whereas V_0 requires the knowledge of the spectrum of the process F for zero frequency.

3.6 Non-ergodic mixture of BSCs

In this section we investigate the behavior of $\log M^*(n,\epsilon)$ for the case when the channel is a non-ergodic mixture of two memoryless channels. To keep the presentation simple, we focus on the example of a pair of BSCs, but similar to Section 3.4 the results in this section can be readily generalized to finite sums of arbitrary DMCs.

One way to motivate the interest in such channels is the following. Consider the Gilbert-Elliott channel with very small τ ("slow fading"). If the range of blocklengths of interest is much smaller than $\frac{1}{\tau}$, we cannot expect (3.379) or (3.382) to give a good approximation of $\log M^*(n,\epsilon)$. In fact, in this case, a model with $\tau = 0$ is intuitively much more suitable. Taking $\tau = 0$ the GEC becomes a mixture of a pair of BSCs.

To define a non-ergodic BSC, we define a state random variable S, which is generated before the start of the transmission according to:

$$\mathbb{P}[S=1] = 1 - \mathbb{P}[S=2] = p_1\,. \tag{3.390}$$

The input and output alphabets of the channel are binary, $\mathsf{A} = \mathsf{B} = \{0,1\}^n$, and the channel is defined as

$$P_{Y^n|X^n}(y^n|x^n) = \delta_S^{|y^n - x^n|}(1-\delta_S)^{n-|y^n-x^n|}\,, \tag{3.391}$$

where $|z^n|$ denotes the Hamming weight of the binary vector z^n and $0 < \delta_1, \delta_2 < 1/2$ are crossover probabilities. The function $M^*(n,\epsilon)$ is defined as usual.

For non-ergodic channels, the role of capacity is replaced by a more general ϵ-capacity, see (2.18), since the latter becomes a non-constant function of $\epsilon > 0$ (typically). In the case of the mixture of BSCs and regardless of the state knowledge at the transmitter or receiver, the ϵ-capacity is given by (assuming $h(\delta_1) > h(\delta_2)$)

$$C_\epsilon = \begin{cases} \log 2 - h(\delta_1)\,, & \epsilon < p_1\,, \\ \log 2 - h(\delta_2)\,, & \epsilon > p_1\,. \end{cases} \tag{3.392}$$

Other than the case of small $|\delta_2 - \delta_1|$, solved in [65], the value of the ϵ-capacity at the breakpoint $\epsilon = p_1$ is in general unknown (see also [26]).

3.6.1 Asymptotic expansion

Recall that the main idea behind the asymptotic expansion (2.22) is in approximating the distribution of an information density by a Gaussian distribution. For non-ergodic channels, it is natural to use an approximation via a mixture of Gaussian distributions. This motivates the next definition[15] .

Definition 10 *For a pair of channels with capacities C_1, C_2 and channel dispersions $V_1, V_2 > 0$ we define a normal approximation $R_{na}(n,\epsilon)$ of their non-ergodic sum with respective probabilities p_1, p_2 ($p_2 = 1 - p_1$) as the solution to*

$$p_1 Q\left((C_1 - R)\sqrt{\frac{n}{V_1}}\right) + p_2 Q\left((C_2 - R)\sqrt{\frac{n}{V_2}}\right) = \epsilon\,. \tag{3.393}$$

[15] This way of defining $R_{na}(n,\epsilon)$ has been suggested by S. Verdú.

Note that for any $n \geq 1$ and $0 < \epsilon < 1$ the solution exists and is unique; see Fig. 3.11 for an illustration. To understand better the behavior of $R_{na}(n,\epsilon)$ with n we assume $C_1 < C_2$ and then it can be shown easily that[16]

$$R_{na}(n,\epsilon) = \begin{cases} C_1 - \sqrt{\frac{V_1}{n}} Q^{-1}\left(\frac{\epsilon}{p_1}\right) + O(1/n)\,, & \epsilon < p_1 \\ C_2 - \sqrt{\frac{V_2}{n}} Q^{-1}\left(\frac{\epsilon - p_1}{1-p_1}\right) + O(1/n)\,, & \epsilon > p_1\,. \end{cases} \tag{3.394}$$

In fact, even more is true:

Lemma 61 *Assume $C_1 < C_2$ and $\epsilon \notin \{0, p_1, 1\}$. Then the following holds:*

$$R_{na}\left(n, \epsilon + O(1/\sqrt{n})\right) = R_{na}(n,\epsilon) + O(1/n)\,. \tag{3.395}$$

Proof: Denote

$$f_n(R) \stackrel{\triangle}{=} p_1 Q\left((C_1 - R)\sqrt{\frac{n}{V_1}}\right) + p_2 Q\left((C_2 - R)\sqrt{\frac{n}{V_2}}\right) \tag{3.396}$$

$$R_n \stackrel{\triangle}{=} R_{na}(n,\epsilon) = f_n^{-1}(\epsilon)\,. \tag{3.397}$$

It is clear that $f_n(R)$ is a monotonically increasing function, and that our goal is to show that

$$f_n^{-1}(\epsilon + O(1/\sqrt{n})) = R_n + O(1/n)\,. \tag{3.398}$$

Assume $\epsilon < p_1$; then for any $0 < \delta < (C_2 - C_1)$ we have $f_n(C_1 + \delta) \to p_1$ and $f_n(C_1 - \delta) \to 0$. Therefore,

$$R_n = C_1 + o(1)\,. \tag{3.399}$$

This implies, in particular, that for large enough n we have

$$0 \leq p_2 Q\left((C_2 - R_n)\sqrt{\frac{n}{V_2}}\right) \leq \frac{1}{\sqrt{n}}\,. \tag{3.400}$$

Then, from the definition of R_n we conclude that

$$\epsilon - \frac{1}{\sqrt{n}} \leq p_1 Q\left((C_2 - R_n)\sqrt{\frac{n}{V_2}}\right) \leq \epsilon\,. \tag{3.401}$$

After applying Q^{-1} to this inequality we get

$$Q^{-1}\left(\frac{\epsilon}{p_1}\right) \leq (C_2 - R_n)\sqrt{\frac{n}{V_2}} \leq Q^{-1}\left(\frac{\epsilon - 1/\sqrt{n}}{p_1}\right)\,. \tag{3.402}$$

By Taylor's formula we conclude

$$R_n = C_1 - \sqrt{\frac{V_1}{n}} Q^{-1}\left(\frac{\epsilon}{p_1}\right) + O(1/n)\,. \tag{3.403}$$

[16]See the proof of Lemma 61 below.

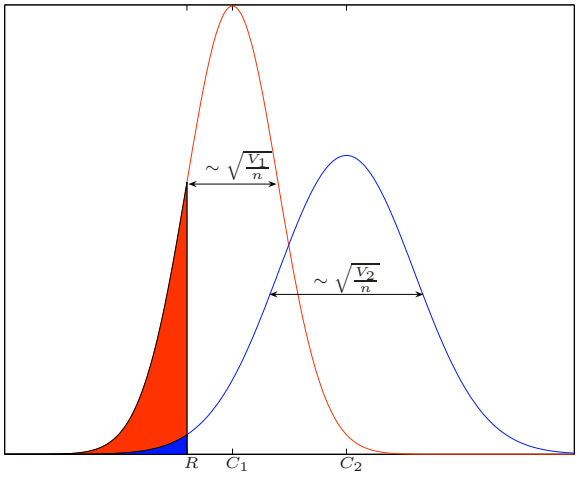


Figure 3.11: Illustration to the Definition 10: $R_{na}(n,\epsilon)$ is found as the unique point R at which the weighted sum of two shaded areas equals ϵ.

Note that the same argument works for ϵ that depends on n, provided that $\epsilon_n < p_1$ for all sufficiently large n. This is indeed the case when $\epsilon_n = \epsilon + O(1/\sqrt{n})$. Therefore, similarly to (3.403), we can show

$$\begin{aligned} f_n^{-1}(\epsilon + O(1/\sqrt{n})) &= C_1 - \sqrt{\frac{V_1}{n}} Q^{-1}\left(\frac{\epsilon + O(1/\sqrt{n})}{p_1}\right) + O(1/n)\,, && (3.404)\\ &= C_1 - \sqrt{\frac{V_1}{n}} Q^{-1}\left(\frac{\epsilon}{p_1}\right) + O(1/n)\,, && (3.405)\\ &= R_n + O(1/n)\,, && (3.406) \end{aligned}$$

where (3.405) follows by applying Taylor's expansion and (3.406) follows from (3.403). The case $\epsilon > p_1$ is treated similarly. ■

We now state our main result in this section.

Theorem 62 *Consider a non-ergodic BSC whose transition probability is $0 < \delta_1 < 1/2$ with probability p_1 and $0 < \delta_2 < 1/2$ with probability $1 - p_1$. Take $C_j = \log 2 - h(\delta_j)$, $V_j = V(\delta_j)$ and define $R_{na}(n,\epsilon)$ as the solution to (3.393). Then for $\epsilon \notin \{0, p_1, 1\}$ we have*

$$\log M^*(n,\epsilon) = nR_{na}(n,\epsilon) + \frac{1}{2}\log n + O(1) \qquad (3.407)$$

regardless of whether ϵ is a maximal or average probability of error, and regardless of whether the state S is known at the transmitter, receiver or both.

Proof: First of all, notice that $p_1 = 0$ and $p_1 = 1$ are treated by Theorem 41. So, everywhere below we assume $0 < p_1 < 1$.

Achievability: Since the proof of the achievability part closely follows the steps of the proof of Theorem 41, we adopt the notation used therein. In particular, from (3.35), we

have that for all n and M there exists an (n, M, p_e) code with

$$p_e \leq \sum_{k=0}^{n} \binom{n}{k} \left(p_1 \delta_1^k (1-\delta_1)^{n-k} + p_2 \delta_2^k (1-\delta_2)^{n-k} \right) \min\left\{1, M S_n^k\right\}, \tag{3.408}$$

where S_n^k is

$$S_n^k \stackrel{\triangle}{=} 2^{-n} \sum_{l=0}^{k} \binom{n}{l}. \tag{3.409}$$

Fix $\epsilon \notin \{0, p_1, 1\}$ and for each n select K as a solution to

$$p_1 Q\left(\frac{K - n\delta_1}{\sqrt{n\delta_1(1-\delta_1)}} \right) + p_2 Q\left(\frac{K - n\delta_2}{\sqrt{n\delta_2(1-\delta_2)}} \right) = \epsilon - \frac{G}{\sqrt{n}}, \tag{3.410}$$

where $G > 0$ is some constant. Application of the Berry-Esseen theorem shows that there exists a choice of G such that for all sufficiently large n we have

$$\mathbb{P}[W > K] \leq \epsilon, \tag{3.411}$$

where

$$W = \sum_{j=1}^{n} 1\{Z_j = 1\}. \tag{3.412}$$

The distribution of W is a mixture of two Bernoulli distributions:

$$\mathbb{P}[W = w] = \binom{n}{w} \left(p_1 \delta_1^w (1-\delta_1)^{n-w} + p_2 \delta_2^w (1-\delta_2)^{n-w} \right). \tag{3.413}$$

Repeating the steps (3.35)-(3.57) we can now prove that as $n \to \infty$ we have

$$\begin{aligned} \log M^*(n, \epsilon) &\geq -\log S_n^K & (3.414)\\ &\geq n \log 2 - nh\left(\frac{K}{n}\right) + \frac{1}{2}\log n + O(1), & (3.415) \end{aligned}$$

where h is the binary entropy function. Thus we only need to analyze the asymptotics of $h\left(\frac{K}{n}\right)$. First, notice that the definition of K as the solution to (3.410) is entirely analogous to the definition of $nR_{na}(n, \epsilon)$. Assuming without loss of generality $\delta_2 < \delta_1$ (the case of $\delta_2 = \delta_1$ is treated in Theorem 41), in parallel to (3.394) we have as $n \to \infty$

$$K = \begin{cases} n\delta_1 + \sqrt{n\delta_1(1-\delta_1)} Q^{-1}\left(\frac{\epsilon}{p_1}\right) + O(1), & \epsilon < p_1 \\ n\delta_2 + \sqrt{n\delta_2(1-\delta_2)} Q^{-1}\left(\frac{\epsilon - p_1}{p_2}\right) + O(1). & \epsilon > p_1. \end{cases} \tag{3.416}$$

From Taylor's expansion applied to $h\left(\frac{K}{n}\right)$ as $n \to \infty$ we get

$$nh\left(\frac{K}{n}\right) = \begin{cases} nh(\delta_1) + \sqrt{nV(\delta_1)} Q^{-1}\left(\frac{\epsilon}{p_1}\right) + O(1), & \epsilon < p_1 \\ nh(\delta_2) + \sqrt{nV(\delta_2)} Q^{-1}\left(\frac{\epsilon - p_1}{p_2}\right) + O(1), & \epsilon > p_1. \end{cases} \tag{3.417}$$

Comparing (3.417) with (3.394) we notice that for $\epsilon \neq p_1$ we have

$$n - nh\left(\frac{K}{n}\right) = nR_{na}(n,\epsilon) + O(1)\,. \tag{3.418}$$

Finally, after substituting (3.418) in (3.415) we obtain the required lower-bound of the expansion:

$$\log M^*(n,\epsilon) \geq nR_{na}(n,\epsilon) + \frac{1}{2}\log n + O(1)\,. \tag{3.419}$$

Before proceeding to the converse part we also need to specify the non-asymptotic bounds that have been used to numerically compute the achievability curves in Fig. 3.12 and 3.13. For this purpose we use Theorem 18 with equiprobable P_{X^n}. Without state knowledge at the receiver we have

$$\begin{aligned} i(X^n;Y^n) &= g_n(W)\,, & (3.420)\\ g_n(w) &= n\log 2 + \log\left(p_1\delta_1^w(1-\delta_1)^{n-w} + p_2\delta_2^w(1-\delta_2)^{n-w}\right)\,, & (3.421) \end{aligned}$$

where W is defined in (3.412). Theorem 18 guarantees that for every M there exists a code with (average) probability of error p_e satisfying

$$p_e \leq \mathbb{E}\left[\exp\left\{-\left[g_n(W) - \log\tfrac{M-1}{2}\right]^+\right\}\right]\,. \tag{3.422}$$

In addition, by application of the random linear code method, the same can be seen to be true for maximal probability of error, provided that $\log_2 M$ is an integer (see Appendix C). Therefore, the numerical computation of the achievability bounds in Fig. 3.12 and 3.13 amounts to finding the largest integer k such that right-hand side of (3.422) with $M = 2^k$ is still smaller than a prescribed ϵ.

With state knowledge at the receiver we can assume that the output of the channel is (Y^n, S_1) instead of Y^n. Thus, $i(X^n;Y^n)$ needs to be replaced by $i(X^n;Y^n,S_1)$ and then expressions (3.420), (3.421) and (3.413) become

$$\begin{aligned} i(X^n;Y^nS_1) &= g_n(W,S_1)\,, & (3.423)\\ g_n(w,s) &= n\log 2 + \log\left(\delta_s^w(1-\delta_s)^{n-w}\right)\,, & (3.424)\\ \mathbb{P}[W=w,S_1=s] &= p_s\binom{n}{w}\delta_s^w(1-\delta_s)^{n-w}\,. & (3.425) \end{aligned}$$

Again, in parallel to (3.422) Theorem 18 constructs a code with M codewords and probability of error p_e satisfying

$$p_e \leq \mathbb{E}\left[\exp\left\{-\left[g_n(W,S_1) - \log\tfrac{M-1}{2}\right]^+\right\}\right]\,. \tag{3.426}$$

Converse: In the converse part we will assume that the transmitter has access to the state realization S_1 and then generates X^n based on both the input message and S_1. Take the best such code with $M^*(n,\epsilon)$ codewords and average probability of error no greater than ϵ. We now propose to treat the pair (X^n, S_1) as a combined input to the channel (but the

S_1 part is independent of the input message) and the pair (Y^n, S_1) as a combined output, available to the decoder. Note that in this situation, the encoder induces a distribution $P_{X^nS_1}$ and is necessarily randomized, because the distribution of S_1 is not controlled by the input message and is given by

$$\mathbb{P}[S_1 = 1] = p_1 \,. \tag{3.427}$$

To apply Theorem 28 we select the auxiliary Q-channel as follows:

$$Q_{Y^nS_1|X^n}(y^n, s|x^n) = \mathbb{P}[S_1 = s]2^{-n} \quad \text{for all } y^n, s, x^n \,. \tag{3.428}$$

Then it is easy to see that under this channel, the output (Y^n, S_1) is independent of X^n. Hence, we have

$$1 - \epsilon' \le \frac{1}{M^*(n,\epsilon)} \,. \tag{3.429}$$

To compute $\beta_{1-\epsilon}(P_{X^nY^nS_1}, Q_{X^nY^nS_1})$ we need to find the likelihood ratio:

$$\begin{aligned} r(X^n; Y^n S_1) &\stackrel{\triangle}{=} \log \frac{P_{X^nY^nS_1}(X^n, Y^n, S_1)}{Q_{X^nY^nS_1}(X^n, Y^n, S_1)} && (3.430)\\ &= \log \frac{P_{Y^n|X^nS_1} P_{X^nS_1}}{Q_{Y^n|X^nS_1} Q_{X^nS_1}} && (3.431)\\ &= n \log 2 + \log P_{Y^n|X^nS_1}(Y^n|X^n S_1) && (3.432)\\ &= n \log 2(1 - \delta_{S_1}) - W \log \frac{1 - \delta_{S_1}}{\delta_{S_1}} \,, && (3.433) \end{aligned}$$

where (3.431) is because $P_{X^nS_1} = Q_{X^nS_1}$ (we omitted the obvious arguments for simplicity), (3.432) is by (3.428) and in (3.433) random variable W is defined in (3.412) and its distribution is given by (3.413).

Now, choose

$$\begin{aligned} R_n &= R_{na}\left(n, \epsilon + \frac{p_1 B_1 + p_2 B_2 + 1}{\sqrt{n}}\right) \,, && (3.434)\\ \gamma_n &= nR_n \,, && (3.435) \end{aligned}$$

where B_1 and B_2 are the Berry-Esseen constants for the sum of independent Bernoulli(δ_j) random variables. Then, we have

$$\begin{aligned} &\mathbb{P}[r(X^n; Y^n S_1) \le \gamma_n | S_1 = 1] \\ &= \mathbb{P}\left[n \log 2(1 - \delta_1) - W \log \frac{(1 - \delta_1)}{\delta_1} \le \gamma_n \,\middle|\, S_1 = 1\right] && (3.436)\\ &\ge Q\left(-\frac{\gamma_n - nC_1}{\sqrt{nV_1}}\right) - \frac{B_1}{\sqrt{n}} && (3.437)\\ &= Q\left((C_1 - R_n)\sqrt{\frac{n}{V_1}}\right) - \frac{B_1}{\sqrt{n}} \,, && (3.438) \end{aligned}$$

where (3.437) is by the Berry-Esseen theorem and (3.438) is just the definition of γ_n. Analogously, we have

$$\mathbb{P}[r(X^n; Y^n S_1) \le \gamma_n | S_1 = 2] \ge Q\left((C_2 - R_n)\sqrt{\frac{n}{V_2}}\right) - \frac{B_2}{\sqrt{n}} \,. \tag{3.439}$$

Together (3.438) and (3.439) imply

$$\begin{aligned}
&\mathbb{P}[r(X^n;Y^nS) \le \gamma_n] \\
&\quad \ge \quad p_1 Q\left((C_1-R_n)\sqrt{\frac{n}{V_1}}\right) + p_2 Q\left((C_2-R_n)\sqrt{\frac{n}{V_2}}\right) - \frac{p_1B_1+p_2B_2}{\sqrt{n}} && (3.440)\\
&\quad = \quad \epsilon + \frac{1}{\sqrt{n}}\,, && (3.441)
\end{aligned}$$

where (3.441) follows from (3.434). Then by using the bound (2.67) we obtain

$$\beta_{1-\epsilon}(P_{X^nY^nS_1}, Q_{X^nY^nS_1}) \ge \frac{1}{\sqrt{n}}\exp\{-\gamma_n\}\,. \tag{3.442}$$

Finally, by Theorem 28 and (3.429) we obtain

$$\begin{aligned}
\log M^*(n,\epsilon) \quad &\le \quad \log\frac{1}{\beta_{1-\epsilon}} && (3.443)\\
&\le \quad \gamma_n + \frac{1}{2}\log n && (3.444)\\
&= \quad nR_{na}\left(n, \epsilon + \frac{p_1B_1+p_2B_2+1}{\sqrt{n}}\right) + \frac{1}{2}\log n && (3.445)\\
&= \quad nR_{na}(n,\epsilon) + \frac{1}{2}\log n + O(1)\,, && (3.446)
\end{aligned}$$

where (3.446) is by Lemma 61. ■

As noted before, for $\epsilon = p_1$ even the capacity term is unknown. However, application of Theorem 28 with $Q_{Y|X} = BSC(\delta_{max})$ where $\delta_{max} = \max(\delta_1, \delta_2)$, yields the following upper bound:

$$C_{p_1} \le 1 - h(s^*)\,, \tag{3.447}$$

where s^* is found as the solution of

$$d(s^*||\delta_2) = d(s^*||\delta_1)\,. \tag{3.448}$$

To get (3.447), take any rate $R > 1 - h(\delta_{max})$ and apply a well-known above-the-capacity error estimate for the Q-channel [16]:

$$1 - \epsilon' \lesssim \exp\left(-nd(s||\delta_{max})\right)\,, \tag{3.449}$$

where $s < \delta_1$ satisfies $R = 1 - h(s)$. Then it is not hard to obtain that

$$\beta_{1-p_1}(P_{Y|X}, Q_{Y|X}) \sim \exp\left(-nd(s^*||\delta_{max})\right)\,. \tag{3.450}$$

The upper bound (3.447) then follows from Theorem 28 immediately. Note that the same upper-bound was derived in [65] (and there it was also shown to be tight in the special case of $|\delta_1 - \delta_2|$ being small enough), but the proof we have outlined above is more general since it also applies to the average probability of error criterion and various state-availability scenarios.

3.6.2 Discussion and numerical comparison

Comparing (3.407) and (3.394) we see that, on one hand, there is the usual $\frac{1}{\sqrt{n}}$ type of convergence to capacity. On the other hand, because the capacity in this case depends on ϵ, the argument of Q^{-1} has also changed accordingly. Moreover, we see that for $p_1/2 < \epsilon < p_1$ we have that capacity is equal to $1 - h(\delta_1)$ but the maximal rate approaches it *from above*. In other words, we see that in non-ergodic cases it is possible to communicate at rates above the ϵ-capacity at finite blocklength.

In view of (3.407) it is natural to choose the following expression as the normal approximation for the $\tau = 0$ case:

$$R_{na}(n, \epsilon) + \frac{1}{2n} \log n \,. \tag{3.451}$$

We compare converse and achievability bounds against the normal approximation (3.451) in Fig. 3.12 and Fig. 3.13. On the latter we also demonstrate numerically the phenomenon of the possibility of transmitting above capacity. The achievability bounds are computed for the maximal probability of error criterion using (3.422) with $i(X^n; Y^n)$ given by expression (3.420) in the case of no state knowledge at the receiver; and using (3.426) with $i(X^n; Y^n S_1)$ given by the (3.423) in the case when S_1 is available at the receiver. The converse bounds are computed using (3.443), that is for the average probability of error criterion, and with the assumption of state availability at both the transmitter and the receiver. Note that the "jaggedness" of the curves is a property of the respective bounds, and not of the computational precision.

On comparing the converse bound and the achievability bound in Fig. 3.13, we conclude that the maximal rate, $\frac{1}{n} \log M^*(n, \epsilon)$ cannot be monotonically increasing with blocklength. In fact, the bounds and approximation hint that it achieves a global maximum at around $n = 200$. We have already observed that for certain ergodic channels and values of ϵ, the supremum of $\frac{1}{n} \log M^*(n, \epsilon)$ need not be its asymptotic value. Although this conflicts with the principal teaching of the error exponent asymptotic analysis (the lower the required error probability, the higher the required blocklength), it does not contradict the fact that for a memoryless channel and any positive integer ℓ

$$\frac{1}{n\ell} \log M^*(n\ell, 1 - (1 - \epsilon)^\ell) \geq \frac{1}{n} \log M^*(n, \epsilon) \,, \tag{3.452}$$

since a system with blocklength $n\ell$ can be constructed by ℓ independent encoder/decoders with blocklength n.

The "typical sequence" approach fails to explain the behavior in Fig. 3.13, as it neglects the possibility that the two BSCs may be affected by an atypical number of errors. Indeed, typicality only holds asymptotically (and the maximal rate converges to the ϵ-capacity, which is equal to the capacity of the bad channel). In the short-run the stochastic variability of the channel is nonneglible, and in fact we see in Fig. 3.13 that atypically low numbers of errors for the bad channel (even in conjunction with atypically high numbers of errors for the good channel) allow a 20% decrease from the error probability (slightly more than 0.1) that would ensue from transmitting at a rate strictly between the capacities of the bad and good channels.

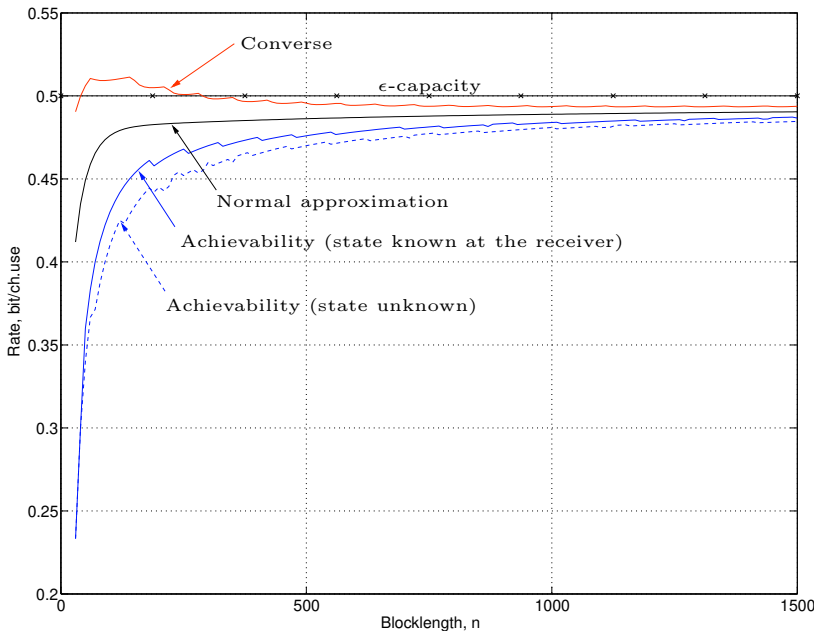


Figure 3.12: Rate-blocklength tradeoff at block error rate $\epsilon = 0.03$ for the non-ergodic BSC whose transition probability is $\delta_1 = 0.11$ with probability $p_1 = 0.1$ and $\delta_2 = 0.05$ with probability $p_2 = 0.9$.

Before closing this section, we also point out that Fano's inequality is very uninformative in the non-ergodic case. For example, for the setup of Fig. 3.12 we have

$$\limsup_{n\to\infty} \frac{\log M^*(n,\epsilon)}{n} \leq \limsup_{n\to\infty} \sup_{X^n} \frac{1}{n} \frac{I(X^n S_1; Y^n S_1) + \log 2}{1-\epsilon} \tag{3.453}$$

$$= \frac{\log 2 - p_1 h(\delta_1) - p_2 h(\delta_2)}{1-\epsilon} \tag{3.454}$$

$$= 0.71 \text{ bit} \tag{3.455}$$

which is a very loose bound.

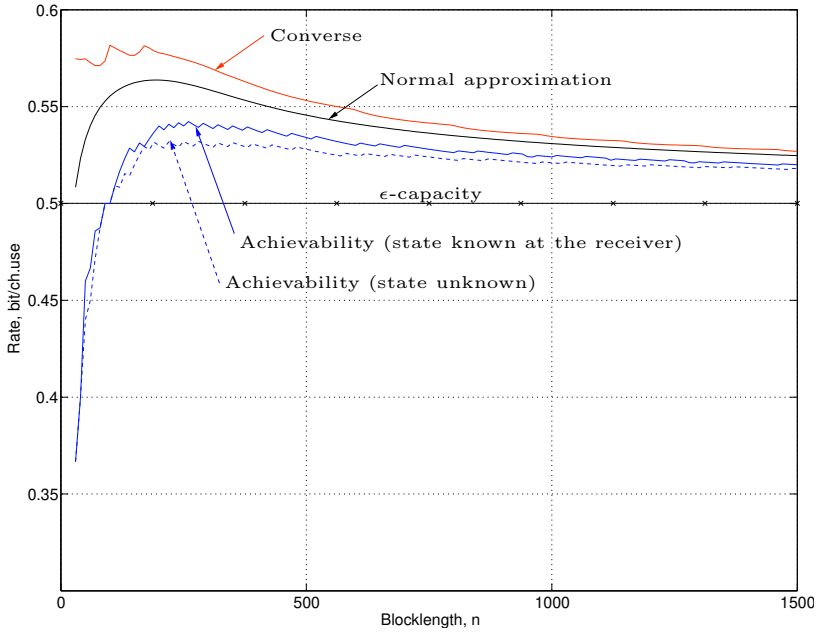


Figure 3.13: Rate-blocklength tradeoff at block error rate $\epsilon = 0.08$ for the non-ergodic BSC whose transition probability is $\delta_1 = 0.11$ with probability $p_1 = 0.1$ and $\delta_2 = 0.05$ with probability $p_2 = 0.9$.

Chapter 4

Gaussian channels

This chapter presents the main results regarding channels subject to Gaussian noise. Namely, a previous work on such channels is overviewed in Section 4.1. Particularization of the general bounds of Chapter 2 is undertaken in Section 4.2. Asymptotic analysis of these bounds in Section 4.3 derives a closed-form expression for the channel dispersion of the AWGN channel. Evaluation of both classical bounds and new bounds demonstrates (Section 4.4) that for the AWGN channel, the value of the fundamental limit $\log M^*(n,\epsilon)$ can be determined with good precision, and furthermore, the two-term approximation (2.23) involving the capacity and dispersion turns out to be surprisingly tight even at rather small blocklengths. Section 4.5 computes the channel dispersion for the parallel AWGN channel subject to a joint power constraint. Finally, the question of minimal achievable energy per bit over Gaussian channels, which has been previously addressed in the limit when then number of information bits goes to infinity, is studied non-asymptotically in Section 4.6. A significant improvement in energy efficiency is demonstrated in the presence of feedback. In addition, a method of feedback communication with zero-error and finite energy per bit is constructed. The material in this chapter has been presented in part in [32, 69–71]. The material of Section 4.3.3 – an asymptotic expansion and the formula for the ϵ-capacity of the AWGN under the average probability of error and average power constraint – appears here for the first time.

4.1 Previous work

4.1.1 Bounds

By the AWGN channel $AWGN(n,P)$ we understand a triple: two alphabets and a collection of conditional probability kernels $P_{Y^n|X^n}$. For blocklength n, we take the alphabets $\mathsf{A}=\mathbb{R}^n$ and $\mathsf{B}=\mathbb{R}^n$ as n-fold Cartesian products, their elements are denoted by x^n and y^n. The channel acts by adding a white Gaussian noise:

$$P_{Y^n|X^n=x^n}=\mathcal{N}(x^n,I_n)\,, \tag{4.1}$$

where I_n is the $n\times n$ identity covariance matrix. Finally, codewords are subject to one of three different power constraints:

1. The code satisfies an *equal-power constraint*, if for each codeword $c_i \in X^n, i = 1, \ldots, M$ we have
$$||c_i||^2 = nP\,. \tag{4.2}$$
The fundamental limit for equal-power constrained codes is defined as
$$M_e^*(n, \epsilon, P) = \max\{M : \exists(n, M, \epsilon)\text{-code satisfying (4.2)}\}\,. \tag{4.3}$$

2. The code satisfies a *maximum (per-codeword) power constraint*, if for each codeword $c_i \in X^n$ we have
$$||c_i||^2 \leq nP\,. \tag{4.4}$$
The fundamental limit for maximum power constrained codes is defined as
$$M_m^*(n, \epsilon, P) = \max\{M : \exists(n, M, \epsilon)\text{-code satisfying (4.4)}\}\,. \tag{4.5}$$

3. The code satisfies an *average power constraint*, if the codewords of the code $\{c_i, i = 1, \ldots, M\}$ satisfy
$$\frac{1}{M}\sum_{i=1}^{M} ||c_i||^2 \leq nP\,. \tag{4.6}$$
The fundamental limit for average power constrained codes is defined as
$$M_a^*(n, \epsilon, P) = \max\{M : \exists(n, M, \epsilon)\text{-code satisfying (4.6)}\}\,. \tag{4.7}$$

The capacity of the AWGN has been computed already in [2], where Shannon shows
$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log M^*(n, \epsilon, P) = \frac{1}{2}\log(1 + P)\,, \tag{4.8}$$
regardless of the power constraint.

Ever since Shannon computed the capacity of the AWGN channel (4.8), there has been some work devoted to the assessment of the penalty incurred by finite blocklength. Foremost, Shannon [4] provided the "cone-packing" bounds (both achievability and converse) that were numerically studied by Slepian [20] (cf. also [23, 61]). Recently, with the advent of sparse-graph codes, a number of works [11, 18, 19, 22] have studied the signal-to-noise ration (SNR) penalty as a function of blocklength in order to improve the assessment of the suboptimality of a given code with respect to the fundamental limit at that particular blocklength rather than the asymptotic limit embodied in the channel capacity. The bounds used for such analysis in all of the quoted work are given as follows [4]:

Theorem 63 (Shannon) *Let*
$$Y_i = x_i + Z_i \tag{4.9}$$
where Z_i are i.i.d. standard normal random variables. Assume that each codeword satisfies
$$\sum_{i=1}^{n} x_i^2 = nP\,. \tag{4.10}$$

Define for $0 \le \theta \le \pi/2$

$$q_n(\theta) = Q\left(\sqrt{nP}\right) + \frac{n-1}{\sqrt{\pi}} e^{-nP/2} \int_\theta^{\pi/2} (\sin\phi)^{n-2} f_n(\sqrt{nP}\cos\phi)\, d\phi \tag{4.11}$$

where

$$f_n(x) = \frac{1}{\Gamma((n+1)/2)} \int_0^\infty t^{n-1} e^{-t^2+\sqrt{2}tx}\, dt\,. \tag{4.12}$$

Then, any (n, M, ϵ) *code for equal-power constraint* (4.2) *satisfies*

$$q_n(\theta(M)) \le \epsilon \tag{4.13}$$

with $\theta(M)$ *defined as*

$$M\Omega_n(\theta(M)) = \Omega_n(\pi) \tag{4.14}$$

with

$$\Omega_n(\theta) = \frac{2\pi^{(n-1)/2}}{\Gamma((n-1)/2)} \int_0^\theta (\sin\phi)^{n-2}\, d\phi\,, \tag{4.15}$$

which is equal to the area of the unit sphere in $\mathbb{R}^n$ *cut out by a cone with semiangle* θ*. Furthermore, there exists an* (n, M, ϵ) *code satisfying equal-power constraint* (4.2) *with*

$$\begin{aligned} \epsilon &\le q_n(\theta(M)) - \frac{M}{\Omega_n(\pi)} \int_0^{\theta(M)} \Omega_n(\phi) \dot{q_n}(\phi)\, d\phi & (4.16)\\ &= \frac{\Gamma(n/2)M}{\sqrt{\pi}\Gamma((n-1)/2)} \int_0^{\theta(M)} q_n(\phi)(\sin\phi)^{n-2}\, d\phi\,. & (4.17) \end{aligned}$$

Computation of the bounds in Theorem 63 is challenging. Various methods [18–20] were proposed in the literature to address this problem; see also [21, 23, 61] for numerical evaluation.

Applying Theorem 4 to the AWGN channel with $P_{X^n} = \mathcal{N}(0, PI_n)$ and optimizing over r and λ, one obtains the following (see [9], Theorem 7.4.4).

Theorem 64 (Gallager, AWGN) *Consider the AWGN channel with noise power* 1*, and signal power* A*. Then for block length* n*, every* $0 \le R \le 1/2\log(1+P)$ *and every* $\delta \in (0, nP]$*, there exists an* $(\exp(nR), n, \epsilon)$ *code (maximal probability of error) satisfying maximal power constraint* (4.4) *with*

$$\epsilon \le \left(\frac{2e^{s(R)\delta}}{\mu(\delta)}\right)^2 e^{-nE_r(R)}\,, \tag{4.18}$$

where

$$E_r(R) = \frac{P}{4\beta}\left[(\beta+1) - (\beta-1)\sqrt{1+\frac{4\beta}{P(\beta-1)}}\right] + \tag{4.19}$$

$$\frac{1}{2}\log_e\left\{\beta - \frac{P(\beta-1)}{2}\left[\sqrt{1+\frac{4\beta}{P(\beta-1)}}-1\right]\right\} \qquad \textit{for } R\in[R_c, C]\,, \tag{4.20}$$

$$E_r(R) = 1-\beta+\frac{P}{2}+\frac{1}{2}\log_e\left(\beta-\frac{P}{2}\right)+\frac{1}{2}\log_e\beta - R\log_e 2\,, \qquad \textit{for } R\in[0,R_c]\,, \tag{4.21}$$

$$\beta = \exp(\max\{2R, 2R_c\})\,, \tag{4.22}$$

$$C = \frac{1}{2}\log(1+P)\,, \tag{4.23}$$

$$R_c = \frac{1}{2}\log\left(\frac{1}{2}+\frac{P}{4}+\frac{1}{2}\sqrt{1+\frac{P^2}{4}}\right)\,, \tag{4.24}$$

$$\mu(\delta) = \mathbb{P}\left[n-\frac{\delta}{P}\le\chi_n^2\le n\right] = \int_{n-\delta/P}^{n}\frac{(y/2)^{n/2-1}e^{-y/2}}{2\Gamma(n/2)}dy\,, \tag{4.25}$$

$$s(R) = \frac{\rho P}{2(1+\rho)^2\beta}\,, \tag{4.26}$$

$$\rho = \frac{P}{2\beta}\left[1+\sqrt{1+\frac{4\beta}{P(\beta-1)}}\right]-1\,. \tag{4.27}$$

Other bounds on the reliability function have appeared recently, e.g. [72]. However, those bounds provide an improvement only for rates well below capacity.

Regarding the asymptotic expansions, the following expansion

$$\frac{1}{n}\log M_m^*(n,\epsilon) = \frac{1}{2}\log(1+P) - \sqrt{\frac{P(P+2)}{2n(1+P)^2}}Q^{-1}(\epsilon)\log e + o\left(\frac{1}{\sqrt{n}}\right) \tag{4.28}$$

was conjectured by Baron *et al.* in [28] after analyzing error-exponent formulas of Shannon [4]. Similarly, analyzing the achievability bounds of Rice [73], [28] conjectures that the following lower bound must hold:

$$\frac{1}{n}\log M_m^*(n,\epsilon) \ge \frac{1}{2}\log(1+P) - \sqrt{\frac{P}{2n(1+P)}}Q^{-1}(\epsilon)\log e + o\left(\frac{1}{\sqrt{n}}\right)\,. \tag{4.29}$$

Clearly taking $\epsilon > 1/2$ the bound (4.29) contradicts the conjecture in (4.28) and thus both cannot be true. In this chapter we will prove a strengthening of (4.28) thus resolving the inconsistency with (4.29).

Moreover, we mention that Shannon's approach in [4] was based on the geometric approach that is hard to generalize to non-AWGN channels. We arrive at the same expansion using a quite general approach. Comparing with the discrete channels, we recall that Strassen [1] obtained (3.110) for the DMC by applying Feinstein's Theorem 1, which, as we demonstrate below, is not tight enough to obtain the expansion for the AWGN. In this case, however, the method of the $\kappa\beta$ bound, Theorem 27, succeeds.

For the parallel AWGN, the extensive study of the reliability function in different regimes has been given in [74], based on a number of previous publications (see the references therein).

4.1.2 Energy per bit

So far, we have considered power-constrained communication. In many practical situations, however, the communication engineer faces the problem of transmitting a certain message with the smallest possible energy per bit. In such situations the key parameters of the code are: the number of degrees of freedom n, the number of information bits k, the probability of block error ϵ and the total energy budget E. Of course, it is not possible to construct a code with arbitrary values of n, k, ϵ and E. Determining the boundary of the achievable (n, k, ϵ, E) is one of the most widely studied problems in information theory.

The first asymptotic result dates back to [30], where Shannon demonstrates that in the limit of $\epsilon \to 0$, $k \to \infty$, $n \to \infty$ and $\frac{k}{n} \to 0$ the smallest achievable energy per bit $E_{\mathsf{b}} \stackrel{\triangle}{=} \frac{E}{k}$ converges to

$$\left(\frac{E_{\mathsf{b}}}{N_0}\right)_{min} = \log_e 2 = -1.59 \text{ dB}\,, \tag{4.30}$$

where $\frac{N_0}{2}$ is the noise power per degree of freedom. The limit does not change if ϵ is fixed, if noiseless causal feedback is available at the encoder, or even if the modulation is suitably restricted.

Alternatively, if one fixes $\epsilon > 0$ and rate $\frac{k}{n} = R$ then as $k \to \infty$ and $n \to \infty$ as a consequence of (4.8) we have (e.g., [9]):

$$\frac{E_{\mathsf{b}}}{N_0} \to \frac{2^{2R} - 1}{2R}\,. \tag{4.31}$$

Thus in this case the minimum energy per bit becomes a function of R; this function being sensitive to modulation and fading scenarios; see [75].

Non-asymptotically, in the regime of fixed rate R and ϵ, the bounds on the minimum E_{b} follow immediately from Theorem 63 of Shannon [4]; such bounds were studied numerically in [14,20,21,23,61]. An optimal scheduling algorithm to minimize energy per bit is proposed in [76] for the purpose of transmitting a stream of buffered packets.

4.2 Computation of the bounds

Note that Theorem 63 yields a bound on $M_e^*(n, \epsilon)$, while practically we are more interested in $M_m^*(n, \epsilon)$. Following the ideas of Shannon [4] we can compare fundamental limits for different power constraints, as follows:

Lemma 65 *For any $0 < P < P'$ the following inequalities hold (maximal probability of error formalism):*

$$M_e^*(n, \epsilon, P) \le M_m^*(n, \epsilon, P) \le M_e^*(n+1, \epsilon, P) \tag{4.32}$$

$$M_m^*(n, \epsilon, P) \le M_a^*(n, \epsilon, P) \le \frac{1}{1 - P/P'} M_m^*(n, \epsilon, P')\,. \tag{4.33}$$

Moreover, in the average probability of error formalism (4.32) *holds without change, while* (4.33) *becomes*

$$M_m^*(n,\epsilon,P) \le M_a^*(n,\epsilon,P) \le \frac{1}{1-P/P'} M_m^*\left(n, \frac{\epsilon}{1-P/P'}, P'\right), \tag{4.34}$$

which holds provided that $\frac{\epsilon}{1-P/P'} \le 1$.

Proof: The left-hand bounds are obvious. The right-hand bound in (4.32) follows from the fact that we can always take the M_m^*-code and add an $(n+1)$-th coordinate to each codeword to equalize the total power to nP. The right-hand bound in (4.33) is obtained as follows. Take an arbitrary (n, M, ϵ) code satisfying average power constraint. By definition (4.6), we have

$$\frac{1}{M}\sum_{i=1}^{M} ||c_i||^2 \le nP\,, \tag{4.35}$$

which by Chebyshev inequality implies that the number of codewords with $||c_i||^2 > nP'$ is at most

$$\#\{i : ||c_i||^2 > nP'\} \le M\frac{P}{P'}\,, \tag{4.36}$$

and hence the remaining codewords constitute a subcode satisfying maximal power constraint (4.4) with power P'. This results in an inequality:

$$M\left(1-\frac{P}{P'}\right) \le M_m^*(n,\epsilon,P')\,. \tag{4.37}$$

Since every code satisfies (4.37), so does the one achieving $M_a^*(n,\epsilon,P)$.

Finally, to obtain (4.34) we proceed as in the proof of (4.33), but notice that the subcode obtained by dropping codewords in (4.36) might have a probability of error that is larger than ϵ. However, we have the following inequality:

$$\epsilon_1\frac{\#\{i : ||c_i||^2 > nP'\}}{M} + \epsilon_2\frac{\#\{i : ||c_i||^2 \le nP'\}}{M} \le \epsilon\,, \tag{4.38}$$

where ϵ_1 and ϵ_2 are the average probabilities of error for the subcode with codewords satisfying $||c_i||^2 > nP'$ and its complement, respectively. However, since $\epsilon_1 \ge 0$ and by the bound (4.36) we obtain

$$\epsilon_2 \le \frac{\epsilon}{1-P/P'}\,, \tag{4.39}$$

and then (4.34) follows. ■

By Lemma 65 it is enough to consider the fundamental limit $M_e^*(n,\epsilon,P)$, or the equal-power constraint. Since this is a per-codeword constraint, for each blocklength n there is a set F_n of permissible inputs, namely, the power sphere:

$$\mathsf{F}_n \stackrel{\triangle}{=} \{x^n : ||x^n||^2 = nP\} \subset \mathbb{R}^n\,. \tag{4.40}$$

This is a natural setting for the $\kappa\beta$ bound, Theorem 27 and the meta-converse, Theorem 34. To apply these bound to the AWGN channel we need to complete three steps: choose the auxiliary output distribution P_{Y^n} on $\mathbb{R}^n$, compute β_α and compute κ_τ. These steps are detailed below.

4.2.1 Choosing the output distribution

First of all, since in the end we aim to apply the central limit theorem (or Berry-Esseen inequality) it is necessary require that P_{Y^n} be a product distribution:

$$P_{Y^n} = P_Y \times \cdots \times P_Y \,. \tag{4.41}$$

On the other hand, because of the spherical symmetry of the problem, it is natural to require that P_{Y^n} be also spherically symmetric on $\mathbb{R}^n$, i.e. for any unitary $U : \mathbb{R}^n \to \mathbb{R}^n$ we must have:

$$P_{Y^n} \circ U^{-1} = P_{Y^n} \,. \tag{4.42}$$

From (4.42) we deduce that Y_1 and $-Y_1$ must be identically distributed. Thus the characteristic function of Y_1 must be positive, symmetric and upper-bounded by 1:

$$0 \le \psi(r) \stackrel{\triangle}{=} \mathbb{E}\left[e^{irY_1}\right] \le 1 \,. \tag{4.43}$$

Additionally, from isotropy it follows that $\gamma Y_1 + \beta Y_2$ must be distributed identically to Y_1, whenever $\gamma^2 + \beta^2 = 1$. Translated to $\psi(r)$ this implies that

$$\gamma^2 + \beta^2 = 1 \quad \Longrightarrow \quad \psi(r) = \psi(\gamma r)\psi(\beta r) \tag{4.44}$$

holds. In particular, suppose that for some r_0, $\psi(r_0) = 0$. Then on taking $\gamma = \beta = \frac{1}{\sqrt{2}}$ we find that $\psi\left(2^{-k/2} r_0\right) = 0$. But this is impossible since $\psi(0) = 1$ and ψ is continuous. Thus $\psi(r) > 0$ for all $r \in \mathbb{R}$.

Let us introduce $f(r) = \log \psi(\sqrt{r})$. Then, for any $\lambda \in [0, 1]$ and all $r \in \mathbb{R}_+$, we have from isotropy (4.44) that

$$f(r) = f(\lambda r) + f((1-\lambda) r) \,. \tag{4.45}$$

Or, equivalently,

$$f(x + y) = f(x) + f(y) \,. \tag{4.46}$$

This implies that for every rational $q \in \mathbb{Q}$ we have

$$f(qx) = q f(x) \,. \tag{4.47}$$

In particular, taking $x = 1$ we find from the continuity of $f(r)$ that[1]

$$f(r) = \text{const} \cdot r \,. \tag{4.48}$$

Thus the characteristic function of P_Y is $\psi(r) = \exp(-\text{const} \cdot r^2)$. This leaves the only possibility: choose P_{Y^n} to be Gaussian,

$$P_{Y^n} = \mathcal{N}(0, \sigma_Y^2 I_n) \tag{4.49}$$

with σ_Y^2 to be chosen later.

[1]Note that, condition (4.46) alone, without continuity, does not imply that $f(x)$ has the form (4.48). Indeed, for a counter-example, notice that over the rationals $\mathbb{Q}$, $\mathbb{R}$ is an infinite dimensional topological vector space; of course, on such space there exist discontinuous linear functions. Therefore, $\mathbb{Q}$-linearity (4.47) and $f(1) = 1$ alone do not imply that f should be of the form (4.48).

4.2.2 Computing β

Having defined necessary distributions we must now compute the function $\beta_\alpha^n(x, P_Y)$ defined in (2.60). For each $x \in \mathsf{F}_n$ the value of $\beta_\alpha^n(x, P_Y)$ can be found directly from (2.66).

A key observation here is that, due to spherical symmetry, the particular choice of $x \in \mathsf{F}_n$ does not affect the value of β_α and thus to simplify the notation we write everywhere below:

$$\beta_\alpha^n(x, P_Y) = \beta_\alpha^n\,, \quad \forall x \in \mathsf{F}_n\,. \tag{4.50}$$

To simplify calculations, we choose $x = x_0 = (\sqrt{P}, \sqrt{P}, \dots \sqrt{P})$.

The information density is given by

$$i(x_0, y^n) = \log \frac{dP_{Y^n|X^n=x_0}}{dP_{Y^n}}(y^n) = \frac{n}{2}\log\sigma_Y^2 + \frac{\log e}{2}\sum_{i=1}^n \left[\frac{y_i^2}{\sigma_Y^2} - (y_i - \sqrt{P})^2\right]\,. \tag{4.51}$$

It is convenient to define independent standard Gaussian variables $Z_i \sim \mathcal{N}(0,1)$. Then, under P_{Y^n}, the information density $i(x_0, Y^n)$ is distributed the same as

$$G_n = n\log\sigma_Y - n\frac{P}{2}\log e + \frac{1}{2}\log e\sum_1^n \left((1-\sigma_Y^2)Z_i^2 + 2\sqrt{P}\sigma_Y Z_i\right) \tag{4.52}$$

and under $P_{Y^n|X^n=x_0}$ it is distributed the same as

$$H_n = n\log\sigma_Y + n\frac{P}{2\sigma_Y^2}\log e + \frac{1}{2\sigma_Y^2}\log e\sum_1^n \left((1-\sigma_Y^2)Z_i^2 + 2\sqrt{P}Z_i\right)\,. \tag{4.53}$$

Note that distributions similar to that of G_n and H_n also appear in the application of the Feinstein's bound, Theorem 1, to the AWGN channel [13].

It is well known that asymptotically $\beta_\alpha^n = \exp\{-D(P_{Y^n|X^n=x_0}||P_{Y^n}) + o(n)\}$. Note that $D(P_{Y^n|X^n=x_0}||P_{Y^n}) = \mathbb{E}\,[H_n]$. Consequently, to have the tightest bound in (2.247) we want to choose[2] σ_Y^2 so as to maximize $\mathbb{E}\,[H_n]$. A simple exercise shows that

$$\sigma_{Y,opt}^2 = 1 + P\,. \tag{4.54}$$

Perhaps unsurprisingly, our P_{Y^n} distribution now coincides with the capacity achieving output distribution for the AWGN channel. Notice, however, that we have not invoked a maximization of mutual information argument.

With this choice of σ_Y^2 the equations for G_n and H_n become

$$G_n = \frac{n}{2}\log(1+P) - n\frac{P}{2}\log e - \frac{1}{2}\log e\sum_1^n \left(PZ_i^2 - 2\sqrt{P^2+P}Z_i\right) \tag{4.55}$$

and

$$H_n = \frac{n}{2}\log(1+P) + \frac{n}{2}\frac{P}{(1+P)}\log e - \frac{1}{2(1+P)}\log e\sum_1^n \left(PZ_i^2 - 2\sqrt{P}Z_i\right)\,. \tag{4.56}$$

Now by using the Neyman-Pearson lemma (2.66) we find that we have proved the following result.

[2] A question that we have not studied is whether the bounds can benefit if σ_Y is allowed to vary with n rather than being fixed to an asymptotically optimal value (4.54).

Theorem 66 *For any $0 \le \alpha \le 1$ we have*

$$\beta_\alpha^n = \beta_\alpha^n(x, P_Y) = \mathbb{P}[G_n \ge \gamma]\,, \tag{4.57}$$

where γ is chosen to satisfy

$$\mathbb{P}[H_n \ge \gamma] = \alpha\,. \tag{4.58}$$

Computation of the distributions of G_n and H_n is simplified by noting that both can be reduced to the form $C_1 + C_2 \sum (Z_i - \delta)^2$, so that they both have non-central χ^2 distributions. However, for large n the value of $P(G_n \ge \gamma)$ must be of the order of $\exp(-nC)$. For such a low quantile, traditional series expansions of the non-central χ^2 distribution do not work very well and a number of other techniques must be used to evaluate these probabilities, including Chernoff bounding and using (2.67) and (2.69).

Theorem 67 *For the AWGN channel with SNR P, for any n and ϵ we have*

$$M_m^*(n-1, \epsilon, P) \le M_e^*(n, \epsilon, P) \le \frac{1}{\mathbb{P}[G_n \ge \gamma_n]}\,, \tag{4.59}$$

regardless of whether ϵ is an average or maximal probability of error, where γ_n is chosen to satisfy

$$\mathbb{P}[H_n \ge \gamma_n] = 1 - \epsilon\,, \tag{4.60}$$

and the variables G_n and H_n are defined in (4.55) *and* (4.56)*.*

Proof: Applying Theorem 34 and Lemma 65 we reduce the problem to that of $\beta_{1-\epsilon}^n(x, P_{Y^n})$. However, Theorem 66 gives a precise value and no further lower-bounding is necessary. ■

Theorem 67 provides a basis for plotting a bound on $\log M_m^*(n, \epsilon, P)$ (see below).

4.2.3 Computing κ

According to the definition (2.92), we need to find the distribution $P_{Z|Y}^*$ which, for every $x \in \mathsf{F}_n$, satisfies

$$\int_{\mathsf{B}} P_{Z|Y}^*(1|y)\, P_{Y^n|X^n=x}(dy) \ge \tau \tag{4.61}$$

and which has the smallest possible value of

$$\int_{\mathsf{B}} P_{Z|Y}^*(1|y)\, P_{Y^n}(dy)\,. \tag{4.62}$$

In general this is a complex problem. In this case, however, the situation is greatly simplified by the spherical symmetry. Intuitively, we feel that the optimum in the definition of κ_τ^n should be spherically symmetric. Below we are going to establish this fact rigorously and also suggest how to find symmetries in other (non-AWGN) problems of interest.

We start by noting that any distribution $P_{Z|Y}$ is completely determined by defining a function $f : \mathsf{B} \mapsto [0, 1]$, namely,

$$f(y) = P_{Z|Y}(1|y)\,. \tag{4.63}$$

If we define the following class of functions on B

$$\mathcal{F}_\tau = \left\{ f : \begin{array}{l} f \text{measurable}\,, f(y) \in [0,1]\,, \\ \forall x \in \mathcal{A}^n : \int_{\mathcal{B}^n} f dP_{Y^n|X^n=x} \geq \tau \end{array} \right\}, \tag{4.64}$$

then

$$\kappa^n(\tau) = \inf_{f \in \mathcal{F}_\tau} \int_{\mathcal{B}^n} f dP_{Y^n}\,. \tag{4.65}$$

Now we define another class, the sub-class of spherically symmetric functions:

$$\mathcal{F}_\tau^{sym} = \left\{ h \in \mathcal{F}_\tau : h(y) = h_r\left(||y||^2\right) \text{ for some } h_r \right\}. \tag{4.66}$$

We can then state the following.

Lemma 68 *For every $0 \leq \tau \leq 1$ we have*

$$\kappa_\tau^n(\mathsf{F}_n, P_Y) = \inf_{h \in \mathcal{F}_\tau^{sym}} \int h dP_Y\,. \tag{4.67}$$

Proof: Since $\mathcal{F}_\tau^{sym} \subseteq \mathcal{F}_\tau$, the inequality

$$\kappa_\tau^n \leq \inf_{h \in \mathcal{F}_\tau^{sym}} \int h dP_Y \tag{4.68}$$

is obvious. It remains to show that

$$\kappa_\tau^n \geq \inf_{h \in \mathcal{F}_\tau^{sym}} \int h dP_Y\,. \tag{4.69}$$

We will show that for every $f \in \mathcal{F}_\tau$ there is a function $h \in \mathcal{F}_\tau^{sym}$ with $\int f\, dP_Y = \int h\, dP_Y$. The claim (4.69) then follows trivially.

Define G to be the isometry group of a unit sphere $\mathbb{S}^{n-1}$. Then $G = O(n)$, the orthogonal group. Define a function on $G \times G$ by

$$d(g, g') = \sup_{y \in \mathbb{S}^{n-1}} ||g(y) - g'(y)||\,. \tag{4.70}$$

Since $\mathbb{S}^{n-1}$ is compact, $d(g, g')$ is finite. Moreover, it defines a distance on G and makes G a topological group. The group action $H : G \times \mathbb{R}^n \mapsto \mathbb{R}^n$ defined as

$$H(g, y) = g(y) \tag{4.71}$$

is continuous in the product topology on $G \times \mathbb{R}^n$. Also, G is a separable metric space. Thus, as a topological space, it has a countable basis. Consequently, the Borel σ-algebra on $G \times \mathbb{R}^n$ coincides with the product of Borel σ-algebras on G and $\mathbb{R}^n$:

$$\mathcal{B}(G \times \mathbb{R}^n) = \mathcal{B}(G) \times \mathcal{B}(\mathbb{R}^n)\,. \tag{4.72}$$

Finally, $H(g, y)$ is continuous and hence is measurable with respect to $\mathcal{B}(G \times \mathbb{R}^n)$ and thus is also a measurable mapping with respect to a product σ-algebra.

It is also known that G is compact. On a compact topological group there exists a unique (right Haar) probability measure μ compatible with the Borel σ-algebra $\mathcal{B}(G)$, and such that

$$\mu(Ag) = \mu(A)\,, \quad \forall g \in G, A \in \mathcal{B}(G)\,. \tag{4.73}$$

Now take any $f \in \mathcal{F}_\tau$ and define an averaged function $h(y)$ as

$$h(y) \stackrel{\triangle}{=} \int_G (f \circ H)(g, y)\, \mu(dg)\,. \tag{4.74}$$

Note that as shown above $f \circ H$ is a positive measurable mapping $G \times \mathsf{B} \mapsto \mathbb{R}_+$ with respect to corresponding Borel σ-algebras. Then by Fubini's theorem, the function $h : \mathsf{B} \mapsto \mathbb{R}_+$ is also positive measurable. Moreover,

$$0 \le h(y) \le \int_G 1\, \mu(dg) = 1\,. \tag{4.75}$$

Define for convenience

$$Q_Y^x \stackrel{\triangle}{=} P_{Y|X=x}\,. \tag{4.76}$$

Then

$$\int_{\mathsf{B}} h(y)\, Q_Y^x(dy) = \int_{\mathsf{B}} Q_Y^x(dy) \int_G (f \circ H)(g, y)\, \mu(dg) = \int_G \mu(dg) \int_{\mathsf{B}} (f \circ H)(g, y) Q_Y^x(dy)\,. \tag{4.77}$$

Change of the order is possible by Fubini's theorem because $f \circ H$ is a bounded function. By the change of variable formula,

$$\int_G \mu(dg) \int_{\mathsf{B}} (f \circ g)(y) Q_Y^x(dy) = \int_G \mu(dg) \int_{\mathsf{B}} f\, \left(Q_Y^x \circ g^{-1}\right)(dy)\,. \tag{4.78}$$

By the definition of Q_Y^x we have, for every set E, $Q_Y^x(E) = Q_Y^0(E - x)$ and the measure Q_Y^0 is fixed under all isometries of $\mathbb{R}^n$:

$$\forall g \in G:\ Q_Y^0(F) = Q_Y^0(g(F))\,. \tag{4.79}$$

But then

$$\left(Q_Y^x \circ g^{-1}\right)(E) \stackrel{\triangle}{=} Q_Y^x\left(g^{-1}(E)\right) = Q_Y^0\left(g^{-1}(E) - x\right) = Q_Y^0\Big\{g^{-1}(E - g(x))\Big\} = Q_Y^{g(x)}(E)\,. \tag{4.80}$$

This proves that

$$Q_Y^x \circ g^{-1} = Q_Y^{g(x)}\,. \tag{4.81}$$

It is important that $x \in \mathsf{F}_n$ implies $g(x) \in \mathsf{F}_n$. In general terms, without AWGN specifics, the above argument shows that in the space of all measures on B the subset $\{Q_Y^x,\, x \in \mathsf{F}_n\}$ is invariant under the action of G.

But $f \in \mathcal{F}_\tau$ and thus $\int f\, dQ_Y^x \geq \tau$ for every $x \in \mathsf{F}_n$. So, from (4.78) and (4.81) we conclude

$$\int_{\mathsf{B}} h\, dQ_Y^x \geq \int_G \tau\, \mu(dg) = \tau\,. \tag{4.82}$$

Together with (4.75) this establishes that

$$h \in \mathcal{F}_\tau\,. \tag{4.83}$$

Now the P_Y measure is also fixed under any $g \in G$:

$$P_Y \circ g^{-1} = P_Y\,. \tag{4.84}$$

Then replacing Q_Y^x with P_Y in (4.78) we obtain

$$\int_{\mathsf{B}} h\, dP_Y = \int_G \mu(dg) \int_{\mathsf{B}} f(y)\, \left(P_Y \circ g^{-1}\right)(dy) = \int_{\mathsf{B}} f\, dP_Y\,. \tag{4.85}$$

The only thing that we have not shown yet is that $h \in \mathcal{F}_\tau^{sym}$. But, this is a simple consequence of the choice of μ. Indeed for any $g' \in G$,

$$(h \circ g')(y) = \int_G (f \circ H)(g, g'(y))\, \mu(dg) = \int_G (f \circ H)(gg', y)\, \mu(dg) = \int_G (f \circ H)(g'', y)\, \mu(dg'') = h(y)\,. \tag{4.86}$$

In the last equality we used a change of measure and invariance of μ under right translations. Thus, h must be constant on the orbits of G and hence, depends only on the norm of y. To summarize, we have shown that h belongs to $\mathcal{F}_\tau^{sym}$ and

$$\int h\, dP_Y = \int f\, dP_Y\,. \tag{4.87}$$

The statement of the lemma then follows. ■

Theorem 69 *For any $0 \leq \tau \leq 1$ we have*

$$\kappa_\tau^n(\mathsf{F}_n, P_Y) = P_0\left\{\frac{p_1(r)}{p_0(r)} \geq \gamma\right\}\,, \tag{4.88}$$

where p_0 and p_1 being are the probability density functions (PDF) of P_0 and P_1 to be defined, and γ is chosen to satisfy:

$$P_1\left\{\frac{p_1(r)}{p_0(r)} \geq \gamma\right\} = \tau\,. \tag{4.89}$$

Proof: By Lemma 68 we obtain a value of κ_τ^n by optimizing over spherically symmetric functions.

First, we will simplify the constraints on the functions in $\mathcal{F}_\tau^{sym}$. Define Q_Y^x and G as in the proof of Lemma 68. As we saw in that proof, each transformation $g \in G$ carries

one measure Q_Y^x into another $Q_Y^{x'}$. Also $x' = g(x)$ in this particular case, but this is not important. What is important, however, is that if $x \in \mathsf{F}_n$ then $x' \in \mathsf{F}_n$. If we define

$$\mathcal{Q} = \{Q_Y^x,\, x \in \mathsf{F}_n\} \tag{4.90}$$

then, additionally, the action of G on $\mathcal{Q}$ is transitive. This enables us to hope that the system of constraints on $h \in \mathcal{F}_\tau^{sym}$ might be overdetermined. Indeed, suppose that h satisfies

$$\int_{\mathsf{B}} h\, dQ_0 \geq \tau \tag{4.91}$$

for some $Q_0 \in \mathcal{Q}$. Then for any measure $Q \in \mathcal{Q}$ there is a transformation $g \in G$ such that

$$Q = Q_0 \circ g^{-1}\,. \tag{4.92}$$

But then

$$\int_{\mathsf{B}} h\, dQ = \int_{\mathsf{B}} h \circ g\, dQ_0 = \int_{\mathsf{B}} h\, dQ_0\,. \tag{4.93}$$

Here the last equality follows from the fact that all members of $\mathcal{F}_\tau^{sym}$ are spherically symmetric functions and as such are fixed under G: $h \circ g = h$.

That is, once a symmetric h satisfies

$$\int_{\mathsf{B}} h\, dP_{Y|X=x_0} \geq \tau \tag{4.94}$$

for one $x_0 \in \mathsf{F}_n$, it automatically satisfies the same inequality for all $x \in \mathsf{F}_n$. So we are free to check (4.94) at one arbitrary x_0 and then conclude that $h \in \mathcal{F}_\tau^{sym}$. For convenience we will choose x_0 to be

$$x_0 = \left(\sqrt{P}, \sqrt{P}, \ldots, \sqrt{P}\right)\,. \tag{4.95}$$

Since all functions in $\mathcal{F}_\tau^{sym}$ are spherically symmetric we will work with their radial parts:

$$h(y) = h_r\left(||y||^2\right)\,. \tag{4.96}$$

Note that P_Y induces a certain distribution on $R = ||Y||^2$, namely,

$$P_0 \sim \sum_1^n (1+P) Z_i^2 \tag{4.97}$$

(as above the Z_i's denote i.i.d. standard Gaussian random variables). Similarly, $P_{Y|X=x_0}$ induces a distribution on $R = ||Y||^2$, namely,

$$P_1 \sim \sum_1^n \left(Z_i + \sqrt{P}\right)^2\,. \tag{4.98}$$

Finally, we see that κ_τ^n is

$$\kappa_\tau^n = \inf_{\{h_r : \int h_r dP_1 \geq \tau\}} \int h_r\, dP_0 \tag{4.99}$$

– a randomized binary hypothesis testing problem with $P_1(\text{decide } P_1) \geq \tau$.

Finally, we are left to note that the existence of a unique optimal solution h_r^* is guaranteed by the Neyman-Pearson lemma. To conclude the proof we must show that the solution of (4.89) exists and thus that h_r^* is an indicator function (i.e., there is no "randomization on the boundary" of a likelihood ratio test). For that we need to show that for any γ the set

$$A_\gamma = \left\{ \frac{p_1(r)}{p_0(r)} = \gamma \right\} \tag{4.100}$$

satisfies $P_1(A_\gamma) = 0$.

To show this we will first show that each set $\{A_\gamma \cap [0, K]\}$ is finite. Then, the Lebesgue measure of $\{A_\gamma \cap [0, K]\}$ is zero. And since P_1 is absolutely continuous with respect to Lebesgue measure we conclude from monotone convergence theorem that

$$P_1(A_\gamma) = \lim_{K \to \infty} P_1(A_\gamma \cup [0, K]) = 0. \tag{4.101}$$

Note that the distribution P_0 is a scaled χ^2-distribution with n degrees of freedom; thus (e.g., (26.4.1) of [77]) the PDF of P_0 is

$$p_0(r) = \frac{r^{n/2-1} e^{-r/(2+2P)}}{(2+2P)^{n/2} \Gamma(n/2)}. \tag{4.102}$$

The distribution P_1 is the non-central χ^2-distribution with n degrees of freedom and non-centrality parameter, λ, equal to nP. Then (see (26.4.25) in [77]) we can write the PDF of P_1 as

$$p_1(r) = \frac{1}{2} e^{-(r+nP)/2} \left(\frac{r}{nP} \right)^{n/4-1/2} I_{n/2-1}(\sqrt{nPr}), \tag{4.103}$$

where $I_a(y)$ is a modified Bessel function of a first kind:

$$I_a(y) = (y/2)^a \sum_{j=0}^{\infty} \frac{(y^2/4)^j}{j! \Gamma(a+j+1)}. \tag{4.104}$$

Using these expressions we obtain

$$f(r) \stackrel{\triangle}{=} \frac{p_1(r)}{p_0(r)} = e^{-\mu r} \sum_{0}^{\infty} a_i r^i. \tag{4.105}$$

The coefficients a_i are such that the series converges for any $r < \infty$. Thus, we can extend $f(r)$ to be an analytic function $F(z)$ over the entire complex plane. Now fix a $K \in (0, \infty)$ and denote

$$S = A_\gamma \cap [0, K] = f^{-1}\{\gamma\} \cap [0, K]. \tag{4.106}$$

By the continuity of f the set S is closed. Thus S is compact. Suppose that S is infinite; then there is sequence $r_k \in S$ converging to some $r^* \in S$. But then from the uniqueness theorem of complex analysis, we conclude that $F(z) = \gamma$ over the entire disk $|z| \leq K$. Since $f(r)$ cannot be constant, we conclude that S is finite. This completes the proof. ■

We have attempted to make the proofs of this section as general as possible so that they can be applied to other situations as well. Indeed, as can be seen from the above

argument, in general one needs to find a group G of transformations of B that permutes elements of the family of measures $\{P_{Y|X=x}, x \in \mathsf{F}_n\}$ and that fixes P_Y. Then the optimum in the definition of κ_τ^n can be sought as a function $\mathsf{B} \mapsto [0,1]$ that is constant on the orbits of G (this was the class $\mathcal{F}_\tau^{sym}$). Moreover, if it happens that action of G on $\{P_{Y|X=x}\}$ is transitive, then a set of conditions on $h \in \mathcal{F}_\tau^{sym}$ can be replaced by just one:

$$\int h \, dP_{Y|X=x_0} \geq \tau \tag{4.107}$$

for any x_0, chosen to be the most convenient one. In this way, computing κ_τ^n is a matter of solving a single randomized binary hypothesis testing problem.

4.3 Asymptotic expansions

The results of applying Theorems 27 and 34 summarized for the AWGN channel give

$$\sup_{\tau \in (0,\epsilon)} \frac{\kappa_\tau^n}{\beta_{1-\epsilon+\tau}^n} \leq M_m^*(n,\epsilon,P) \leq \frac{1}{\beta_{1-\epsilon}^{n+1}} \tag{4.108}$$

with β_α^n and κ_τ^n given by Theorems 66 and 69. To analyze the asymptotic behavior of $\log M_m^*(n,\epsilon,P)$ with n we need to analyze the asymptotics of β_α^n and κ_τ^n. Since β_α^n, by Theorem 66, is found by solving a binary hypothesis testing problem between product distributions, its asymptotics is given by Lemma 14. For κ_τ^n a different approach is needed.

4.3.1 Asymptotic analysis of κ

We first return the index n to definitions (4.97) and (4.98) and will write $P_0^{(n)}$ and $P_1^{(n)}$. Then, from (4.97) we see that each term in the sum for $P_0^{(n)}$ has the characteristic function

$$\phi_0(t) = \mathbb{E}\left[\exp\left\{it(1+P)Z^2\right\}\right] = \frac{1}{\sqrt{1-2(1+P)it}} \tag{4.109}$$

with $\sqrt{z} : \mathbb{C} \to \mathbb{C}$ denoting the principal branch.

Analogously, for terms in $P_1^{(n)}$ we have

$$\phi_1(t) = \mathbb{E}\left[\exp\left\{it\left(Z+\sqrt{P}\right)^2\right\}\right] = \frac{\exp\left\{iPt/(1-2it)\right\}}{\sqrt{1-2it}}\,. \tag{4.110}$$

Define two new distributions, which are shifted and scaled versions of $P_{0,1}^{(n)}$:

$$Q_0^{(n)} \sim \frac{1}{\sqrt{n}}\left[\sum_1^n (1+P)Z_i^2 - n(1+P)\right], \tag{4.111}$$

$$Q_1^{(n)} \sim \frac{1}{\sqrt{n}}\left[\sum_1^n \left(Z_i+\sqrt{P}\right)^2 - n(1+P)\right]. \tag{4.112}$$

Note that we shifted both by the same amount, namely their mean $n(1+P)$. Since the transformation applied is the same for both, we conclude that binary hypothesis testing problem $P_0^{(n)}$ *vs.* $P_1^{(n)}$ is equivalent to $Q_0^{(n)}$ *vs.* $Q_1^{(n)}$.

From the central limit theorem,

$$\begin{aligned} Q_0^{(n)} &\to \mathcal{N}(0, V_0), \quad V_0 = \mathrm{Var}\left((1+P)Z_i^2\right), \text{ and} \\ Q_1^{(n)} &\to \mathcal{N}(0, V_1), \quad V_1 = \mathrm{Var}\left(\left[Z_i + \sqrt{P}\right]^2\right). \end{aligned} \tag{4.113}$$

The variances $V_{0,1}$ are, of course, trivially computed as

$$V_0 = 2(1+P)^2 \text{ and } V_1 = 2(1+2P). \tag{4.114}$$

It is not hard to see that both $|\phi_0(t)|^4$ and $|\phi_1(t)|^4$ are integrable on $(-\infty, \infty)$. But then the local limit theorem applies and a) both $Q_{0,1}^{(n)}$ have densities $q_{0,1}^n(r)$, and b) those densities converge uniformly on $r \in \mathbb{R}$:

$$q_0^n(r) \rightrightarrows g_{V_0}(r), \text{ and} \tag{4.115}$$

$$q_1^n(r) \rightrightarrows g_{V_1}(r), \tag{4.116}$$

where $g_{\sigma^2}(r)$ denotes the PDF of the normal distribution with zero mean and variance σ^2.

We summarize this result in the following lemma.

Lemma 70 *In the statement of Theorem 69, P_0, P_1 can be replaced with $Q_0^{(n)}, Q_1^{(n)}$ and densities p_0, p_1 with q_0^n, q_1^n defined by* (4.111) *and* (4.112)*. Additionally, limits* (4.113)-(4.116) *hold.*

Lemma 71 *Under the conditions of Theorem 69,*

$$\begin{aligned} \kappa_\tau^n &\to Q(-r_\tau^*) - Q(r_\tau^*), \text{ and} \\ r_\tau^* &= \sqrt{\frac{V_1}{V_0}} Q^{-1}\left(\frac{1-\tau}{2}\right). \end{aligned} \tag{4.117}$$

Proof: Denote

$$A_\gamma = \left\{\frac{q_1^n(r)}{q_0^n(r)} \geq \gamma\right\} \tag{4.118}$$

and suppose that we have computed two limits

$$T_0(\gamma) = \lim_n Q_0^{(n)}(A_\gamma), \text{ and} \tag{4.119}$$

$$T_1(\gamma) = \lim_n Q_1^{(n)}(A_\gamma). \tag{4.120}$$

Assume that both functions T_0 and T_1 are continuous and monotonically decreasing from 1 to 0 for $\gamma \in [0, \infty)$. Then, the inverse functions $T_i^{-1} : (0, 1] \to [0, \infty)$ are also continuous and decreasing.

Choose $\epsilon > 0$ and set γ to be

$$\gamma = T_1^{-1}(\tau + \epsilon). \tag{4.121}$$

Then for all sufficiently large n we have $Q_1^{(n)}(A_\gamma) \geq \tau + \epsilon/2$, and thus for such n

$$\kappa_\tau^n \leq \kappa_{\tau+\epsilon/2}^n \leq Q_0^{(n)}(A_\gamma)\,. \tag{4.122}$$

Taking the limit as $n \to \infty$ we find

$$\limsup_n \kappa_\tau^n \leq T_0(\gamma)\,. \tag{4.123}$$

In this last equation γ is a continuous function of ϵ. Then, on taking $\epsilon \to 0$, we have

$$\limsup_n \kappa_\tau^n \leq T_0(T_1^{-1}(\tau))\,. \tag{4.124}$$

Using the same argument for $\tau - \epsilon$, we get the same lower bound on $\liminf$. Thus, finally,

$$\lim_n \kappa_\tau^n = T_0(T_1^{-1}(\tau))\,. \tag{4.125}$$

The rest of the proof is devoted to finding T_0 and T_1. In view of Lemma 70, the sequence

$$f_n(r) = \frac{q_1^n(r)}{q_0^n(r)}\,, \quad n = 1, 2, \ldots \tag{4.126}$$

converges uniformly on compacts to

$$f_\infty(r) = \frac{g_{V_1}(r)}{g_{V_0}(r)} = \sqrt{\frac{V_0}{V_1}} e^{-\mu r^2} \tag{4.127}$$

where $\mu = (V_1^{-1} - V_0^{-1})/2$ is positive. We denote

$$D_\gamma = \{r : f_\infty(r) \geq \gamma\}\,. \tag{4.128}$$

Chose another $\epsilon > 0$ and denote $R_K = [-K, K]$. Then for sufficiently large n and all $r \in R_K$, we have

$$f_\infty(r) - \epsilon \leq f_n(r) \leq f_\infty(r) + \epsilon\,. \tag{4.129}$$

Consequently,

$$D_{\gamma+\epsilon} \cap R_K \;\subset\; \{f_n \geq \gamma\} \cap R_K \;\subset\; D_{\gamma-\epsilon} \cap R_K\,. \tag{4.130}$$

It follows that

$$Q_1^{(n)}[f_n \geq \gamma] \leq Q_1^{(n)}[R_K^c] + Q_1^{(n)}\left[D_{\gamma-\epsilon} \cap R_K\right]\,. \tag{4.131}$$

Now note that for K large enough $D_{\gamma-\epsilon} \subset R_K$. Using this and the central limit theorem for Q_1 we conclude

$$\limsup_n Q_1^{(n)}[f_n \geq \gamma] \leq 2Q(K) + T_1(\gamma - \epsilon)\,. \tag{4.132}$$

Here $T_1(\gamma)$ is defined as

$$T_1(\gamma) = Q\left(-\rho_\gamma V_1^{-1/2}\right) - Q\left(\rho_\gamma V_1^{-1/2}\right)\,, \tag{4.133}$$

$$\rho_\gamma = \sqrt{\frac{1}{\mu}} \ln\left[\sqrt{\frac{V_0}{V_1}} \frac{1}{\gamma}\right]\,. \tag{4.134}$$

In (4.132) we are also free to take $K \to \infty$ and $\epsilon \to 0$ due to continuity. We can also make the same argument for $D_{\gamma+\epsilon}$ and $\liminf Q_1^{(n)}$. Consequently,

$$\lim_n Q_1^{(n)}[f_n \geq \gamma] = T_1(\gamma)\,. \tag{4.135}$$

Similarly, for Q_0

$$\lim_n Q_0^{(n)}[f_n \geq \gamma] = T_0(\gamma) \tag{4.136}$$

with $T_0(\gamma)$ defined as

$$T_0(\gamma) = Q\left(-\rho_\gamma V_0^{-1/2}\right) - Q\left(\rho_\gamma V_0^{-1/2}\right)\,. \tag{4.137}$$

This proves assumptions (4.119) and (4.120).

Finally, to obtain (4.117) from (4.125) one must merely use the identity $Q(x)+Q(-x) = 1$. ■

In addition to precise asymptotic value, given by Lemma 70, we also will need a lower-bound that is uniform in τ.

Lemma 72 *For every $P > 0$ there are constants $C_1 > 0$ and $C_2 > 0$ such that for all sufficiently large n and all $\tau \in [0,1]$,*

$$\kappa_\tau^n \geq \frac{1}{C_1}\left(\tau - e^{-C_2 n}\right)\,. \tag{4.138}$$

Proof: Remember that κ_τ^n is determined by a binary hypothesis testing problem between $P_0^{(n)}$ and $P_1^{(n)}$, as defined by (4.97) and (4.98). We will omit indices (n) where it does not cause confusion. Also in this proof all exp exponents are to the base e. The argument consists of two steps.

Step 1. There is a $\delta > 0$ such that for all $n \geq 1$ the Radon-Nikodym derivative $\frac{dP_1^{(n)}}{dP_0^{(n)}}(r)$ is upper-bounded by a constant C_1 on the set

$$r \in R_n \stackrel{\triangle}{=} [n(1+P-\delta), n(1+P+\delta)]\,. \tag{4.139}$$

Step 2. Since the measures $P_1^{(n)}$ have mean $n(1+P)$, by the Chernoff bound there is a constant C_2 such that

$$P_1^{(n)}\left[\,\{R_n\}^c\,\right] \leq e^{-C_2 n}\,. \tag{4.140}$$

Now choose any set A such that $P_1(A) \geq \tau$. Then

$$P_1[A \cap R_n] \geq P_1(A) - P_1\left[\,\{R_n\}^c\,\right] \geq \tau - e^{-C_2 n}\,. \tag{4.141}$$

But then

$$\begin{aligned} P_0[A] \geq P_0[A \cap R_n] &= \int_{\mathbb{R}_+} 1_{A\cap R_n}\, dP_0 = \\ &= \int_{A\cap R_n} \frac{dP_0}{dP_1}\, dP_1 \geq \frac{1}{C_1}\int_{A\cap R_n} dP_1 \geq \frac{1}{C_1}\left(\tau - e^{-C_2 n}\right)\,. \end{aligned} \tag{4.142}$$

This establishes the required inequality. The rest is devoted to proving Step 1, namely,

$$f_n(r) \stackrel{\triangle}{=} \frac{dP_1}{dP_0} \leq C_1 \quad \text{on } R_n \ \ \forall n\,. \tag{4.143}$$

We have already discussed some properties of $f_n(r)$ in (4.105). Here, however, we will need a precise expression for it, easily obtainable via (4.102) and (4.103):

$$f_n(r) = (1+P)^{n/2} \exp\left\{-n\frac{P}{2} - r\frac{P}{2P+2}\right\} \times \\ \times (nPr)^{-n/4+1/2}\, 2^{n/2}\Gamma\left(\frac{n}{2}\right) I_{n/2-1}\left(\sqrt{nPr}\right)\,, \tag{4.144}$$

where $I_{n/2-1}(x)$ is the modified Bessel function of the first kind.

We will consider only the case in which n is even. This is possible because in [78] it is shown that

$$\mu > \nu \geq 0 \implies I_\mu(x) < I_\nu(x)\,, \tag{4.145}$$

for all $x > 0$. Thus if n is odd then an upper bound is obtained by replacing $I_{n/2-1}$ with $I_{n/2-3/2}$.

Now for integer index $k = n/2 - 1$ the following bound is shown in [79]:

$$I_k(z) \leq \sqrt{\frac{\pi}{8}} e^z \frac{1}{\sqrt{z}} \left(1 + \frac{k^2}{z^2}\right)^{-1/4} \exp\left\{-k \sinh^{-1}\frac{k}{z} + z\left(\sqrt{1+\frac{k^2}{z^2}} - 1\right)\right\}\,. \tag{4.146}$$

Note that we only need to establish the bound for r's that are of the same order as n, $r = O(n)$. Thus we will change the variable

$$r = nt \tag{4.147}$$

and seek an upper bound on $f_n(nt)$ for all t inside some interval containing $(1+P)$.

Using (4.146) and the expression

$$\ln\Gamma\left(\frac{n}{2}\right) = \frac{n-1}{2}\ln\frac{n}{2} - \frac{n}{2} + O(1), \tag{4.148}$$

$f_n(r)$ in (4.144) can be upper-bounded, after some algebra, as

$$f_n(nt) \leq \exp\left\{-\frac{n}{2}K(t,P) + O(1)\right\}\,. \tag{4.149}$$

Here the $O(1)$ term is uniform in t for all t on any finite interval not containing zero, and

$$K(t,P) = -\ln\left\{1+\sqrt{1+4Pt}\right\} + \sqrt{1+4Pt} + \ln(1+P) - P - \frac{Pt}{P+1} - 1 + \ln 2\,. \tag{4.150}$$

A straightforward exercise shows that a maximum of $K(t,P)$ is attained at $t^* = 1+P$ and

$$K_{max} = K(t^*, P) = 0\,. \tag{4.151}$$

Thus

$$f_n(nt) \le O(1)\,, \quad t \in [a,b]\,, a > 0\,. \tag{4.152}$$

In particular (4.143) holds if we take, for example, $a = (1+P)-1$ and $b = (1+P)+1$. ■

In fact, the Radon-Nikodym derivative is bounded for all r, not only $r \in R_n$ and, hence

$$\kappa_\tau^n \ge \frac{1}{C_1}\tau \tag{4.153}$$

instead of the weaker (4.138). But showing that this holds for all r complicates the proof unnecessarily.

4.3.2 Expansion for the additive white Gaussian noise (AWGN) channel

The main result of this section is the following:

Theorem 73 *For the AWGN channel with SNR P, $0 < \epsilon < 1$ and for equal-power, maximal-power and average-power constraints, the capacity and dispersion are given by*

$$C(P) = \frac{1}{2}\log(1+P)\,, \tag{4.154}$$

$$V(P) = \frac{P}{2}\frac{P+2}{(P+1)^2}\log^2 e\,, \tag{4.155}$$

respectively. Moreover, for any power constraint we have (maximal probability of error)

$$\log M^*(n,\epsilon,P) = nC - \sqrt{nV}Q^{-1}(\epsilon) + O(\log n)\,. \tag{4.156}$$

More precisely, for equal-power and maximal-power constraints, the $O(\log n)$ term in (4.156) can be bounded by

$$O(1) \le \log M_{e,m}^*(n,\epsilon,P) - \left[nC - \sqrt{nV}Q^{-1}(\epsilon)\right] \le \frac{1}{2}\log n + O(1)\,, \tag{4.157}$$

and this holds in both maximal and average probability of error formalism. For average-power constraint we have (only for maximal probability of error)

$$O(1) \le \log M_a^*(n,\epsilon,P) - \left[nC - \sqrt{nV}Q^{-1}(\epsilon)\right] \le \frac{3}{2}\log n + O(1)\,. \tag{4.158}$$

The proof of Theorem 73 depends on a number of results of independent interest.

Theorem 74 (Converse) *Consider the AWGN channel with SNR P and choose $\epsilon \in (0,1)$. Then there are $N_c(P,\epsilon)$ and $g_c(P,\epsilon)$ such that, for all $n > N_c(P,\epsilon)$, we have*

$$\log M_e^*(n,\epsilon,P) \le nC(P) - \sqrt{nV(P)}Q^{-1}(\epsilon) + \frac{1}{2}\log n + g_c(P,\epsilon) \tag{4.159}$$

regardless of whether ϵ is a maximal or average probability of error. Moreover, $N_c(P,\epsilon)$ and $g_c(P,\epsilon)$ are continuous functions of P and ϵ.

Proof: Denote $\alpha = 1 - \epsilon > 0$. As we have shown in Section 4.2.2 for any $x \in \mathsf{F}_n$ the distribution of $i(x, Y)$ is the same as that of H_n in (4.56). Thus, (2.67), for any $\gamma_n > 0$, we have

$$\inf_{x \in \mathsf{F}_n} \beta_\alpha^n(x) = \beta_\alpha^n(x) \geq \frac{1}{\gamma_n} \left(\alpha - \mathbb{P}[H_n \geq \log \gamma_n]\right) . \tag{4.160}$$

If we redefine $\gamma_n' = -(\log \gamma_n - nC(P))$, then

$$\mathbb{P}[H_n \geq \gamma_n] = \mathbb{P}\left[\sum h_i \leq \gamma_n'\right] \tag{4.161}$$

with

$$h_i = \frac{\log e}{2(1+P)} \left(PZ_i^2 - 2\sqrt{P} Z_i - P\right) \tag{4.162}$$

and the Z_i's are i.i.d. standard normal. Note that $\mathbb{E}\left[h_i\right] = 0$ and define

$$V(P) = \mathrm{Var}(h_i), \quad T(P) = \mathbb{E}\left[|h_i|^3\right], \quad \text{and} \quad B(P) = \frac{6T(P)}{V(P)^{3/2}} . \tag{4.163}$$

Explicit expressions for $T(P)$ and $B(P)$ are not important. We mention only that all are positive continuous functions of $P > 0$.

Set

$$N_c(P, \epsilon) = \left(\frac{2B(P)}{1-\epsilon}\right)^2 . \tag{4.164}$$

Then for $n > N_c(P, \epsilon)$ we have

$$\alpha_n = \alpha - \frac{2B(P)}{\sqrt{n}} > 0 . \tag{4.165}$$

For such n set

$$\gamma_n' = -\sqrt{nV(P)} Q^{-1}(\alpha_n) . \tag{4.166}$$

Then from the Berry-Esseen inequality, Theorem 13, we have

$$\left| \mathbb{P}\left[\sum h_i \leq \gamma_n'\right] - \alpha_n \right| \leq \frac{B(P)}{\sqrt{n}} . \tag{4.167}$$

Hence,

$$\mathbb{P}\left[\sum h_i \leq \gamma_n'\right] \leq \alpha_n + \frac{B(P)}{\sqrt{n}} \leq \alpha - \frac{B(P)}{\sqrt{n}} . \tag{4.168}$$

On substituting this bound into (4.160) we obtain

$$\beta_\alpha^n \geq \exp(\gamma_n' - nC(P)) \frac{B(P)}{\sqrt{n}} . \tag{4.169}$$

From Theorem 34 this then implies

$$\log M_e^*(n, \epsilon, P) \leq nC(P) - \gamma_n' + \frac{1}{2} \log n - \log B(P) . \tag{4.170}$$

From Taylor's theorem, for some $\theta \in \left[\alpha - \frac{2B(P)}{\sqrt{n}}, \alpha\right]$, we have

$$\gamma_n' = -\sqrt{nV(P)}Q^{-1}(\alpha) + 2B(P)\sqrt{V(P)}\frac{dQ^{-1}}{dx}(\theta)\,. \tag{4.171}$$

Without loss of generality, we assume that $\left[\alpha - \frac{2B(P)}{\sqrt{n}}, \alpha\right] \subset (0,1)$ for all $n > N_c(P,\epsilon)$ (otherwise just increase $N_c(P,\epsilon)$ until this is true).

Since $\frac{dQ^{-1}}{dx}$ is a continuous function on $(0,1)$, we can lower bound $\frac{dQ^{-1}}{dx}(\theta)$ by

$$g_1(P,\epsilon) = \min_{[\alpha_1,\alpha]} \frac{dQ^{-1}}{dx}\,, \tag{4.172}$$

where $\alpha_1 = \alpha - \frac{2B(P)}{\sqrt{N_c(P,\epsilon+1)}}$. Note that $g_1(P,\epsilon)$ is a continuous function of P and ϵ. This results in

$$\gamma_n' \geq -\sqrt{nV(P)}Q^{-1}(\alpha) + g_1(P,\epsilon)2B(P)\sqrt{V(P)}\,. \tag{4.173}$$

Substituting this bound into (4.170) and defining

$$g_c(P,\epsilon) = -2B(P)\sqrt{V(P)}g_1(P,\epsilon) - \log B(P) \tag{4.174}$$

we arrive at

$$\log M_e^*(n,\epsilon,P) \leq nC(P) + \sqrt{nV(P)}Q^{-1}(\alpha) + g_c(P,\epsilon)\,. \tag{4.175}$$

Trivial computation of $\mathrm{Var}(h_i)$ concludes the proof. ■

Corollary 75 *For the AWGN channel with SNR P and for each $\epsilon \in (0,1)$, we have (maximal probability of error)*

$$M_a^*(n,\epsilon,P) \leq nC(P) - \sqrt{nV(P)}Q^{-1}(\epsilon) + \frac{3}{2}\log n + O(1)\,. \tag{4.176}$$

Proof: Set

$$N(\epsilon,P) = \max_{P_1\in[P,2P]} N_c(\epsilon,P_1)\,, \tag{4.177}$$

$$g(\epsilon,P) = \max_{P_1\in[P,2P]} g_c(\epsilon,P_1)\,. \tag{4.178}$$

Now set $P_n = (1+1/n)P$ and use (4.33) in Lemma 65. Then for all $n > N(\epsilon,P)$ according to Theorem 74 we have

$$\begin{aligned}\log M_a^*(n,\epsilon,P) &\leq -\log\left(1-\frac{P}{P_n}\right) + \log M_m^*(n,\epsilon,P_n) \leq \\ &\leq \log(n+1) + \log M_e^*(n+1,\epsilon,P_n) \leq \\ &\leq (n+1)C(P_n) - \sqrt{(n+1)V(P_n)}Q^{-1}(\epsilon) + \frac{3}{2}\log(n+1) + g(\epsilon,P)\,.\end{aligned} \tag{4.179}$$

After repeated use of Taylor's theorem we can collect all $O(1)$, $O(1/n)$ and $O(1/\sqrt{n})$ terms into $O(\log n)$, and the statement of Corollary 75 follows. ■

Theorem 76 (Achievability) *For the AWGN channel with SNR P and for each $\epsilon \in (0,1]$, we have (maximal probability of error)*

$$\log M_e^*(n,\epsilon,P) \geq nC(P) - \sqrt{nV(P)}Q^{-1}(\epsilon) + O(1)\,. \tag{4.180}$$

Proof: We will use all the notation of the proof of Theorem 74, but redefine

$$\alpha_n = \alpha + \frac{2B(P)}{\sqrt{n}}\,. \tag{4.181}$$

Note that for n sufficiently large[3] $\alpha_n < 1$ and the definition of γ_n' in (4.166) is meaningful.

From the Berry-Esseen inequality (4.167) we conclude that

$$\mathbb{P}\left[\sum h_i \leq \gamma_n'\right] \geq \alpha_n - \frac{B(P)}{\sqrt{n}} \geq \alpha + \frac{B(P)}{\sqrt{n}}\,. \tag{4.182}$$

In other words, we have proven that, on setting

$$\log\gamma_n = nC(P) - \gamma_n' = nC(P) + \sqrt{nV(P)}Q^{-1}(\alpha_n)\,, \tag{4.183}$$

we obtain

$$P_{Y|X=x_0}\left[i(x_0,Y) \geq \log\gamma_n\right] = \mathbb{P}\left[\sum h_i \leq \gamma_n'\right] \geq \alpha + \frac{B(P)}{\sqrt{n}} \tag{4.184}$$

for sufficiently large n and any $x_0 \in \mathsf{F}_n$. Therefore, by setting

$$\tau_n \stackrel{\triangle}{=} \frac{B(P)}{\sqrt{n}}\,. \tag{4.185}$$

we have

$$\begin{aligned}
\log\beta^n_{1-\epsilon+\tau_n} &\leq P_{Y^n}\left[i(x^n;Y^n) \geq \log\gamma_n\right] && (4.186)\\
&= \mathbb{E}\left[\exp\{-i(x^n;Y^n)\}1_{\{i(x^n;Y^n)\geq\log\gamma_n\}}\,\middle|\, X^n = x^n\right] && (4.187)\\
&\leq -\log\gamma_n - \frac{1}{2}\log n + O(1) && (4.188)\\
&= -nC(P) + \gamma_n' - \frac{1}{2}\log n + O(1)\,, && (4.189)
\end{aligned}$$

where the (4.188) is by Lemma 20.

Finally, we use general Theorem 27 with $\tau = \tau_n$ to obtain

$$\log M_e^*(n,P,\epsilon) \geq \log\frac{\kappa^n_{\tau_n}}{\beta^n_{\alpha+\tau_n}}\,. \tag{4.190}$$

For the chosen τ_n Lemma 72 gives

$$\log\kappa^n_{\tau_n} \geq -\frac{1}{2}\log n + O(1)\,. \tag{4.191}$$

[3]The magnitude of n required for this to hold is practically unrealistic. Indeed, having $\alpha_n < 1$ requires taking $n > [2B(P)/\epsilon]^2$ which makes $n \sim 10^{12}$ for $\epsilon \sim 10^{-6}$. This implies that an expansion should be informative only for values of n irrelevant for practice. It is therefore somewhat unexpected how tight in reality the approximation (4.180) is; see Figs. 4.3 and 4.4 in Section 4.4. This observation suggests that Berry-Esseen inequality is too conservative to explain the tightness of the normal approximations.

This inequality, together with (4.189), yields

$$\log M_e^*(n,P,\epsilon) \geq nC(P) - \gamma_n' + O(1)\,. \tag{4.192}$$

It is easy to see that $Q^{-1}(\alpha_n) = Q^{-1}(\alpha) + O(1/\sqrt{n})$ and thus, for γ_n' we have

$$\gamma_n' = \sqrt{nV(P)}Q^{-1}(\epsilon) + O(1)\,. \tag{4.193}$$

This concludes the proof. ■

Obtaining this expansion with the term $o(\sqrt{n})$ is much easier (and amounts to the application of Lemma 71 to (4.108)). Refining the remaining term to $O(1)$ required application of Lemma 72. This is needed because the key difference in obtaining the $O(1)$ estimate is to set $\tau_n = O(1/\sqrt{n})$ instead of $\tau = O(1)$.

Proof of Theorem 73: Since (4.157) and (4.158) imply all other statements, it is sufficient to prove the former. Lower bounds in (4.157) and (4.158) follow from Theorem 76. Upper bound in (4.157) is a consequence of Theorem 74 and (4.32). Finally, the upper bound in (4.158) is given by Corollary 75. ■

As discussed in Section 4.1, the expression for channel dispersion of the AWGN was conjectured by Baron *et al.* in [28], see (4.28), after analyzing asymptotic formulas of Shannon [4]. The reasoning given in [28] relied on expressions (9) and (73) in [4]. However, the latter are not directly applicable here because they are asymptotic, $n \to \infty$, equivalence relations under a fixed rate R, whereas in Theorem 73 the rate is changing with n. Similarly, an asymptotic expansion up to the $o(\sqrt{n})$ term is put forward in [80]. The proof given there reduces the AWGN problem to that of the cost-constrained DMC by a method of fine quantization of the input/output alphabets. However, the proof of the DMC case given there implicitly assumes a fixed alphabet size and makes heuristic appeals to the central-limit theorem.

4.3.3 A special case: average power constraint and average probability of error

The only case not covered by Theorem 73 is the case of the average power constraint and average probability of error. In this Section we demonstrate that this case is drastically different from all the rest. The main result is the following:

Theorem 77 *For the AWGN channel with SNR P and average probability of error $0 < \epsilon < 1$ we have*

$$\log M_a^*(n,\epsilon,P) = \frac{n}{2}\log\left(1+\frac{P}{1-\epsilon}\right) + O\left(n^{\frac{2}{3}}\right)\,, \tag{4.194}$$

as $n \to \infty$. In other words, in the setup of average power constraint (4.6) *and average probability of error, the strong converse does not hold and the ϵ-capacity of the AWGN channel is given by*

$$C_\epsilon = \frac{1}{2}\log\left(1+\frac{P}{1-\epsilon}\right)\,. \tag{4.195}$$

Proof:

First, we show the upper (converse) bound. By Lemma 65 from (4.34) we have:

$$\log M_a^*(n,\epsilon,P) \le \log \frac{1}{1-P/P'} + \log M_m^*(n,\epsilon',P')\,, \tag{4.196}$$

where $\epsilon' = \frac{\epsilon}{1-P/P'}$ we chose P' so that

$$\begin{aligned}
\epsilon' &= 1 - n^{-\frac{1}{3}}\,, && (4.197)\\
P' &= \frac{P}{1-\epsilon} + O\left(n^{-\frac{1}{3}}\right)\,, && (4.198)\\
C(P') &= \frac{1}{2}\log\left(1+\frac{P}{1-\epsilon}\right) + O\left(n^{-\frac{1}{3}}\right)\,, && (4.199)\\
V(P') &= V\left(\frac{P}{1-\epsilon}\right) + O\left(n^{-\frac{1}{3}}\right)\,, && (4.200)
\end{aligned}$$

where (4.199) and (4.200) are possibly by Taylor's expansion applied to smooth functions $C(P)$ and $V(P)$, as defined in (4.154) and (4.155), resp. Then, as in the proof of Theorem 67 we have

$$\begin{aligned}
\log M_m^*(n,\epsilon',P') &\le -\log \beta_{1-\epsilon'}^n && (4.201)\\
&\le nC(P') + \sqrt{\frac{nV(P')}{1-\epsilon'}} - \log\frac{1-\epsilon'}{2} && (4.202)\\
&= n\left[C\left(\frac{P}{1-\epsilon}\right) + O\left(n^{-\frac{1}{3}}\right)\right] + n^{\frac{2}{3}}\left[\sqrt{V\left(\frac{P}{1-\epsilon}\right)} + O\left(n^{-\frac{1}{3}}\right)\right]\\
&\quad + O(\log n) && (4.203)\\
&= nC\left(\frac{P}{1-\epsilon}\right) + O\left(n^{\frac{2}{3}}\right)\,, && (4.204)
\end{aligned}$$

where (4.202) follows by Lemma 15, and (4.203) holds by the Taylor's expansion.

Next we show the lower (achievability) bound. Denote

$$M_1 \stackrel{\triangle}{=} M_m^*\left(n, 2n^{-\frac{1}{3}}, \frac{P}{1-\epsilon}\left(1-2n^{-\frac{1}{3}}\right)\right)\,, \tag{4.205}$$

and assume that

$$\log M_1 \ge nC\left(\frac{P}{1-\epsilon}\right) + O\left(n^{\frac{2}{3}}\right)\,. \tag{4.206}$$

Denote

$$M = M_1\frac{1-2n^{-\frac{1}{3}}}{1-\epsilon}\,. \tag{4.207}$$

For sufficiently large n we know that $M > M_1$. Then consider a code with M_1 codewords chosen from the maximal power constraint code achieving M_m^* in the definition (4.205), and $(M-M_1)$ all-zero codewords. According to (4.207), the average probability of error of such a code is upper-bounded by

$$2n^{-\frac{1}{3}}\cdot\frac{M_1}{M} + 1\cdot\frac{M-M_1}{M} \le \epsilon \tag{4.208}$$

as required. At the same time, the average power is given by

$$\frac{P}{1-\epsilon}\left(1-2n^{-\frac{1}{3}}\right)\cdot\frac{M_1}{M}+0\frac{M-M_1}{M}\le P\,. \tag{4.209}$$

Therefore, we constructed an (n,M,ϵ) code satisfying average power constraint (4.6) for the power P, which together with (4.206) and (4.207) implies

$$\log M_a^*(n,\epsilon,P)\ge nC\left(\frac{P}{1-\epsilon}\right)+O\left(n^{\frac{2}{3}}\right)\,. \tag{4.210}$$

We are left to prove (4.206). Choosing $\tau_n=n^{-\frac{1}{3}}$ in (4.108) we get together with Lemma 72

$$\log M_1\ge-\log\beta^n_{1-n^{-\frac{1}{3}}}+O(\log n)\,, \tag{4.211}$$

where β^n_α is given by Theorem 66. A simple upper-bound on β^n_α sufficient for proving (4.206) will follow from (2.69) if we can show that

$$\mathbb{P}\left[H_n<nC\left(\frac{P}{1-\epsilon}\right)-n^{\frac{2}{3}}\right]\le n^{-\frac{1}{3}}\,, \tag{4.212}$$

where H_n is defined in Theorem 66. Since according to (4.56), H_n is the sum of i.i.d. random variables, then [81, Theorem 3.7.1] implies that

$$\mathbb{P}\left[H_n<nC\left(\frac{P}{1-\epsilon}\right)-n^{\frac{2}{3}}\right]=O\left(e^{-\frac{1}{2V(\frac{P}{1-\epsilon})}n^{\frac{1}{3}}}\right)\,, \tag{4.213}$$

which in particular means (4.212) holds for all sufficiently large n. Thus, (4.206) has been shown[4] . ■

4.4 Numerical comparison

In this section our goal is to compare achievability and converse bounds on $\log M_m^*(n,\epsilon)$ (maximal power constraint). As before, we plot the converse bounds for the average probability of error criterion and achievability bounds for the maximal. The results are found on Figs. 4.1 and 4.2. Let us first explain how each bound was computed:

1. Converse bound is Theorem 67. Note that in [4] Shannon gives another converse bound (4.13). However, in this case both bounds numerically coincide almost exactly, while Theorem 67 is slightly easier to compute. For these reasons only the new one is plotted.

[4] In fact this argument can be trivially changed to show that in (4.206) the residual term $O\left(n^{\frac{2}{3}}\right)$ can be replaced with $O\left(n^{\frac{1}{2}+\delta}\right)$ for any $\delta>0$. Similarly, based on the $\kappa\beta$ bound and [81, Theorem 3.7.1] we can easily extend to the AWGN channel the results on the moderate deviations shown in [82] for the DMC; see Section 5.6.

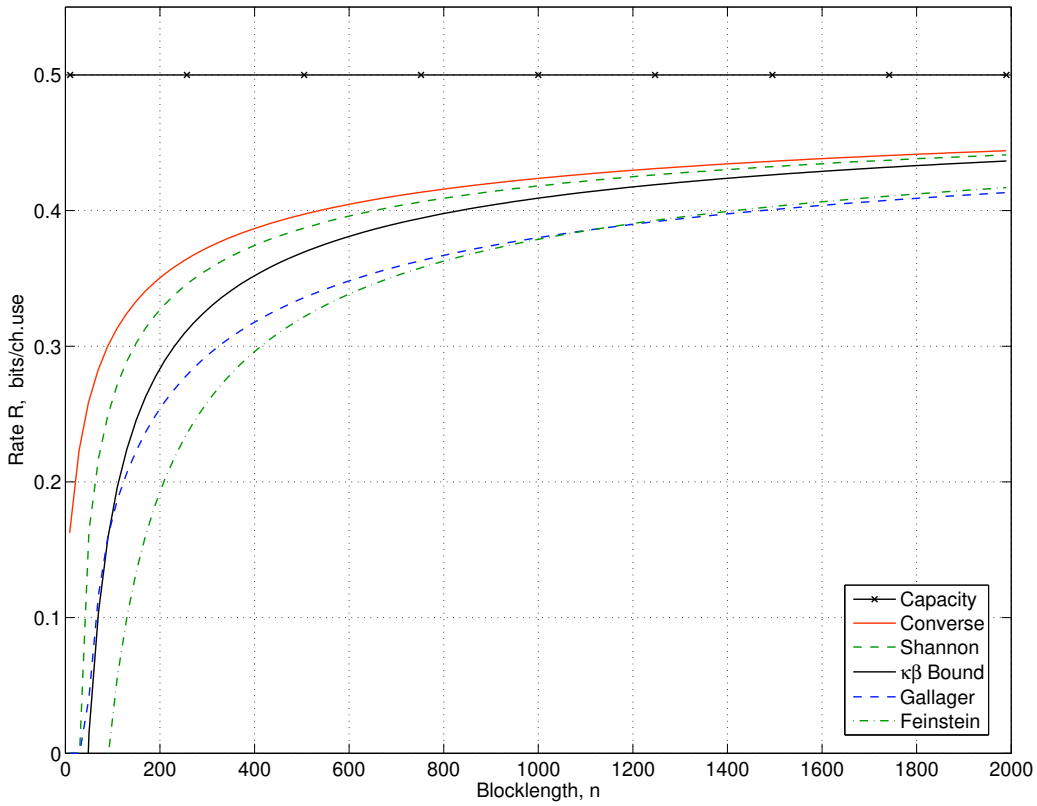


Figure 4.1: Bounds for the AWGN channel, $SNR = 0$ dB, $\epsilon = 10^{-3}$.

2. Shannon bound is Theorem 63. As suggested by [20] we first integrate [4, (20)] by parts and then calculate $Q(\theta)$ via [4, (17)]. In this way computation is reduced to evaluation of the non-central t-distribution and numerical integration.

   We also need to convert the codebook with a given *average* probability of error to the codebook with a prescribed *maximal* probability of error; for the BSC and BEC we used the random linear code method, which is not applicable to the case of the AWGN channel. Instead, we applied the following well-known method: if there exists an $(M, \tau\epsilon)$-code for average probability then there must exist a $(\tau M, \epsilon)$-subcode for maximal probability. Consequently, if $M_S(n, \epsilon)$ is the maximal cardinality of the codebook guaranteed by the Shannon bound, then instead we plot

$$M_S^{max}(n, \epsilon) = \max_{\tau \in [0,1)} (1 - \tau) M_S(n, \tau\epsilon)\,. \tag{4.214}$$

3. Feinstein's bound is the strengthening of Feinstein's lemma as given by Corollary 26 with

$$\mathsf{F}_n = \left\{x^n : ||x^n||^2 \le nP\right\} \tag{4.215}$$

   and $P_{X^n} = \mathcal{N}(0, PI_n)$.

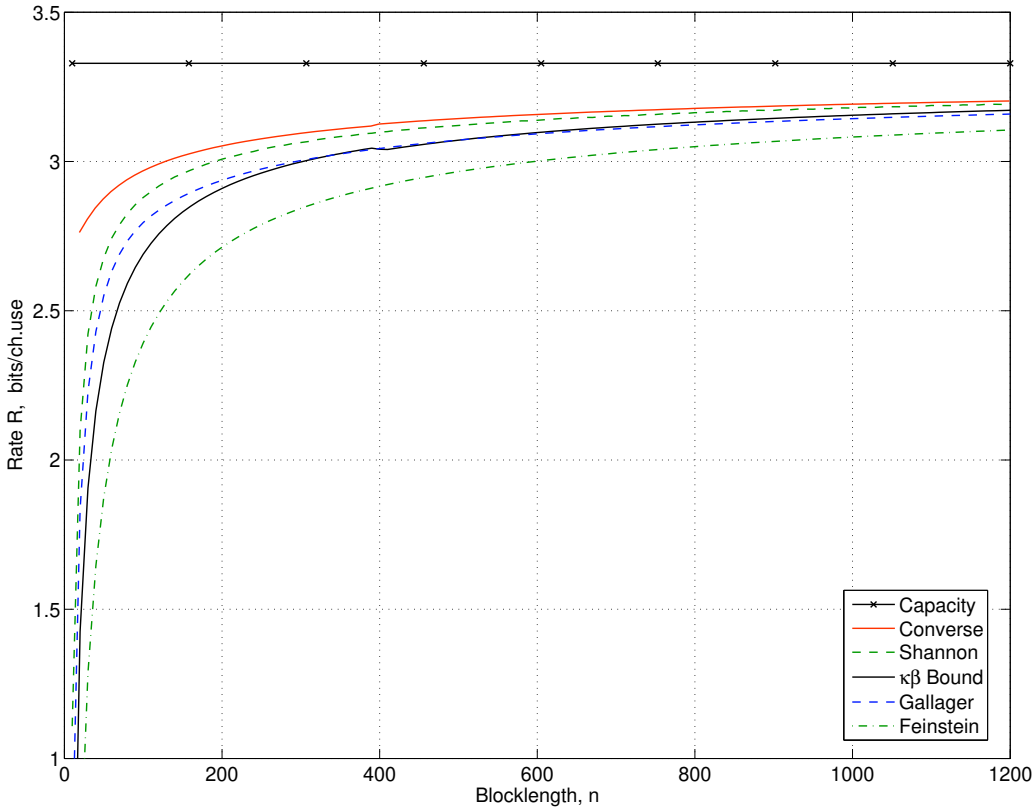


Figure 4.2: Bounds for the AWGN channel, $SNR = 20$ dB, $\epsilon = 10^{-6}$.

4. Gallager's bound is Theorem 64, where we optimize the choice of δ for each R, and then select the largest R that still keeps the bound (4.18) below the required ϵ.

5. $\kappa\beta$ bound is an application of Theorem 27 with β_α and κ_τ given by Theorems 66 and 69. Lemma 71 is a significant help in computations. Experimentally, we have observed that convergence in (4.117) is very fast. For example, for $P = 1, n = 10$ and $\tau \in [10^{-6}, 10^{-1}]$,
$$|\kappa_\tau^n - \kappa_\tau^\infty| \le 5 \cdot 10^{-3} \kappa_n(\tau)\,. \tag{4.216}$$
Note also that κ_τ^n affects the rate $\frac{\log M}{n}$ only as $\frac{\log \kappa_\tau^n}{n}$. So, numerically we can safely replace κ_τ^n with its asymptotic value. In that way for every n we must solve only one binary hypothesis testing problem: the one that yields β_α^n. A small downward jump can be seen occurring for $n \approx 400$ on Fig. 4.2. That is not a property of the $\kappa\beta$ bound, rather this happens because the value of β_α^n becomes so small that a precise computation needs to be replaced by a numerically stable bound.

Note that Feinstein's lemma generates codewords inside the power sphere, Gallager's codebook is in the thin layer around the power sphere, while Shannon's codebook and that of $\kappa\beta$ bound are both precisely on the power sphere.

As we can see, Shannon's bound is the clear winner on both Figs. 4.1 and 4.2. It comes very close to the converse bound: for example, on Fig. 4.1 the gap between the bounds on

$\log M$ is smaller than 6 bits, uniformly across the blocklengths depicted. This illustrates, that the methods of information theory allow computation of the fundamental limits of the AWGN channel within a few bits.

The $\kappa\beta$ bound, although slightly looser than Shannon's, has the advantage of being more general (Shannon's bound is purely AWGN specific and is based on geometric arguments), easier to compute and, most importantly, easier to analyze asymptotically. Indeed, it is $\kappa\beta$ bound that was used in the proof of Theorem 73. As we can see on Figs. 4.1 and 4.2 it is also quite competitive for finite n.

Regarding the classical bound of Feinstein, we can see that, as shown analytically, the $\kappa\beta$ bound is uniformly better than the Feinstein bound (even in the stronger form given by Corollary 26). Comparison with Gallager's bound demonstrates that for large n, the $\kappa\beta$ bound is always more advantageous. However, for small n Gallager's bound can yield better performance. Informally speaking, this happens because Gallager upper-bounds the performance of the optimal (maximum-likelihood) decoder, while in the $\kappa\beta$ bound we analyze a suboptimal hypothesis-testing based decoder, but we do not use further bounds. Numerical comparison demonstrates that for small n it is crucial to use a maximum-likelihood decoder. The effect is more pronounced as we lower the target probability of error, see Fig. 4.2. In general we observe that Gallager's bound improves as the channel becomes better and as ϵ gets smaller. On the other hand, the new $\kappa\beta$ bound is more uniform over both SNR and ϵ.

Let us compare the behavior of the bounds in terms of their asymptotic expansions. As shown by Theorem 67, the converse bound on $\log M_m^*$ has the behavior

$$nC - \sqrt{nV}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,, \tag{4.217}$$

as $n \to \infty$. Gallager's and Feinstein's' bounds achieve a correct linear (capacity) term, but not the $\sqrt{n}$-term. Both the $\kappa\beta$ and Shannon bounds yield a correct $\sqrt{n}$ term. Unfortunately, no bound achieves a $\frac{1}{2}\log n$ term (for maximal probability of error), and for this reason, the true value of the constant in front of $\log n$ in Theorem 73 remains unknown.

According to Theorem 73, we expect the following normal approximation to be describe the behavior of the fundamental limit $\log M_m^*(n, \epsilon, P)$ non-asymptotically:

$$\log M_m^*(n, \epsilon, P) \approx \frac{n}{2}\log(1+P) - \sqrt{n\frac{P}{2}\frac{P+2}{(P+1)^2}\log e}Q^{-1}(\epsilon) + \frac{1}{2}\log n\,. \tag{4.218}$$

Although, Theorem 73 does not provide an exact value of the coefficient in $\log n$ term, based on the bounds given there and on empirical evidence we conjecture that the true coefficient is indeed $\frac{1}{2}$. This is reflected in the approximation (4.218).

On Fig. 4.3 and Fig. 4.4 we compare this approximation with the converse and the best achievability (Shannon's) bounds on $M_m^*(n, \epsilon, P)$. It is quite clear that the approximation (4.218) is indeed very tight.

4.4.1 A remark on the $\kappa\beta$ bound

As we noted in Section 3.4, see the discussion after Theorem 47, the $\kappa\beta$ bound is a natural choice for the situations with cost constraints. The detailed application of the $\kappa\beta$ bound

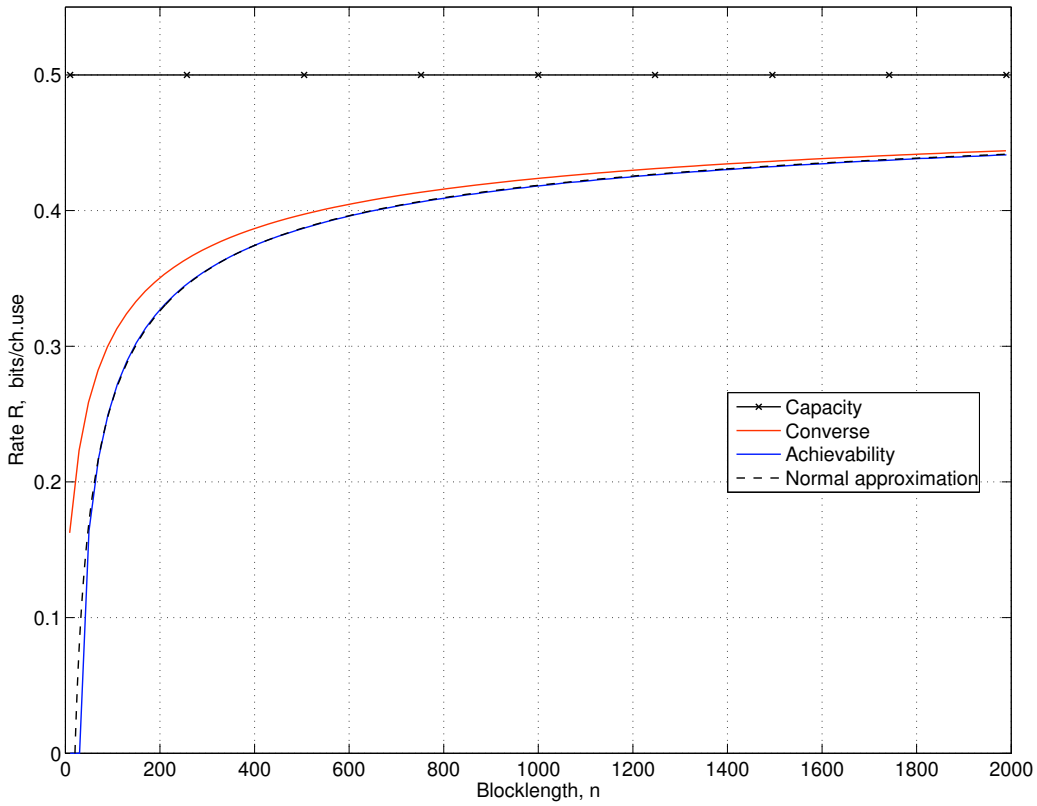


Figure 4.3: Normal approximation for the AWGN channel, $SNR = 0$ dB, $\epsilon = 10^{-3}$.

in this chapter, suggests the following observation. According to (4.108), the behavior of $\log M^*(n,\epsilon)$ is determined by the exponents of β_α^n and κ_τ^n. However, for memoryless channels whenever we choose $P_{Y^n} = P_Y \times \cdots \times P_Y$, it is well-known that

$$\beta_\alpha(P_{Y^n|X^n=x^n}, P_{Y^n}) \sim e^{-nE_\beta(P_Y)}\,, \tag{4.219}$$

where the exponent $E_\beta(P_Y) = \frac{1}{n} D(P_{Y^n|X^n=x^n}||P_{Y^n})$.

In deriving expressions for the converse, we selected $P_Y = P_Y^*$ so as to *minimize* the exponent

$$E_\beta(P_Y^*) = E_{\beta,min} \tag{4.220}$$

and thereby obtain the tightest converse.

When considering achievability, however, it seems that our goal should be different and we must have chosen P_Y to *maximize* E_β to obtain a better bound. However, for different choice of P_Y, with a higher E_β, from the achievability Theorem 27 we immediately conclude that $\kappa_\tau^n(\mathsf{F}_n, P_Y)$ should then decay exponentially $\kappa_\tau^n \sim \exp(-nE_\kappa)$ with exponent such that

$$E_\beta(P_Y) - E_\kappa(P_Y) \le E_{\beta,min}\,. \tag{4.221}$$

Otherwise achievability would contradict the converse. What makes the $\kappa\beta$ bound useful is that, for the choice of $P_Y = P_Y^*$, it happens that $E_\kappa = 0$. And then, of course, the

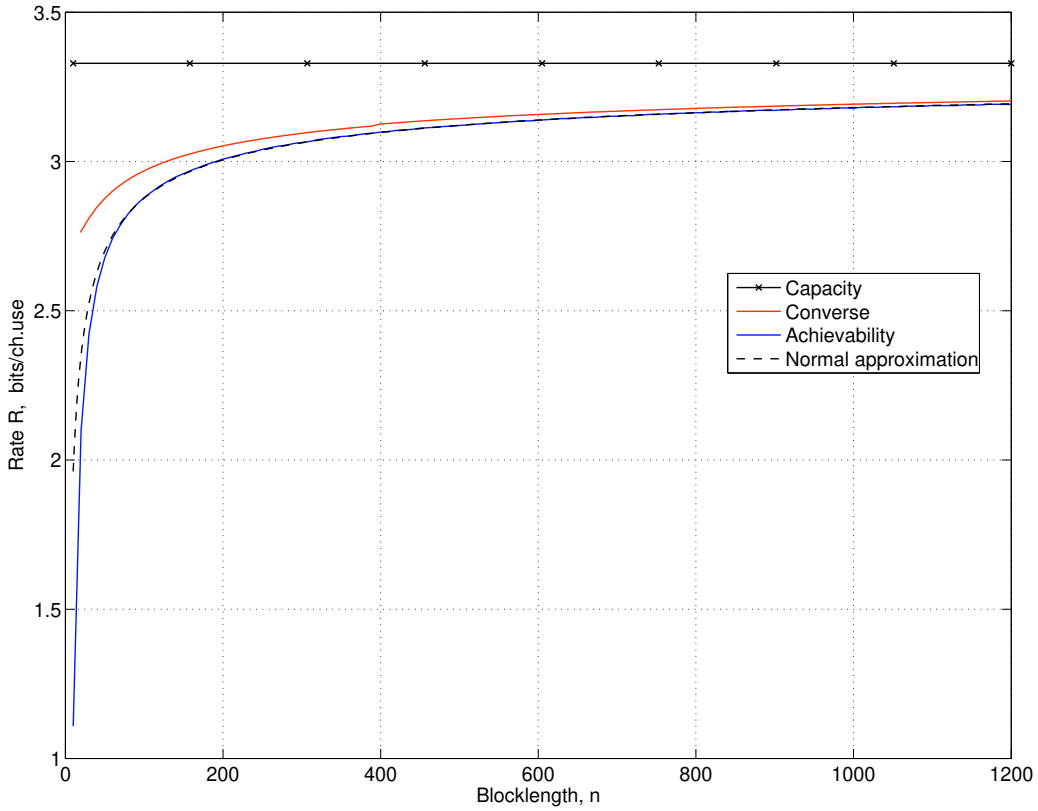


Figure 4.4: Normal approximation for the AWGN channel, $SNR = 20$ dB, $\epsilon = 10^{-6}$.

converse and achievability bounds meet in the exponent, giving way to the capacity and strong converse theorems:

$$C_\epsilon = C = E_{\beta,min}\,. \tag{4.222}$$

Thus, the quest for proving capacity (and the strong converse) is, in our language, a quest for attaining $E_\kappa = 0$. Indeed, if we rewrite (4.221) as

$$E_\kappa(P_Y) \geq E_\beta(P_Y) - E_{\beta,min} \tag{4.223}$$

then we can immediately see that whenever $E_\kappa(P_Y) = 0$ it must be that $E_\beta(P_Y) = E_{\beta,min}$. It might happen that for the cases where we cannot directly minimize $E_\beta(P_Y)$ this indirect characterization will help.

4.5 Parallel AWGN channel

For the real-valued L-parallel AWGN channel we set $\mathsf{A} = \mathbb{R}^{L\times n}$, $\mathsf{B} = \mathbb{R}^{L\times n}$ and $P_{Y|X}$ is defined by

$$Y_{i,j} = X_{i,j} + \sigma_i Z_{i,j}\,, \quad i = 1\ldots L,\ j = 1\ldots n\,, \tag{4.224}$$

where $Z_{i,j}$ are independent $\mathcal{N}(0,1)$ random variables. Also, codewords $\mathbf{c}$ are subject to a (maximal) power constraint:

$$||\mathbf{c}||^2 = \sum_{j=1}^{n}\sum_{i=1}^{L}|c_{i,j}|^2 \le nP\,. \tag{4.225}$$

As usual we define

$$M^*(n,\epsilon,P) = \sup\{M:\ \exists \text{ an } (n,M,\epsilon,P)-\text{code satisfying (4.225)}\}\,. \tag{4.226}$$

Theorem 78 *For a parallel AWGN channel and $\epsilon \in (0,1)$ we have*

$$\log M^*(n,\epsilon,P) = nC_L(P) - \sqrt{nV_L(P)}Q^{-1}(\epsilon) + O(\log n)\,, \tag{4.227}$$

regardless of whether ϵ is a maximal or average probability of error, where[5]

$$C_L(P) = \sum_{i=1}^{L} C\left(\frac{W_i}{\sigma_i^2}\right)\,, \quad \textit{and} \tag{4.229}$$

$$V_L(P) = \sum_{i=1}^{L} V\left(\frac{W_i}{\sigma_i^2}\right)\,, \tag{4.230}$$

where C and V are the capacity and dispersion of the AWGN channel, see (4.154) *and* (4.155)*, and $\{W_j\}$ are the usual waterfilling powers*

$$W_i = \left[\lambda - \sigma_i^2\right]^+ \tag{4.231}$$

and λ is the solution of

$$\sum_{i=1}^{L} W_i = P\,. \tag{4.232}$$

4.5.1 Converse bound

Similarly to (4.32) in Lemma 65, by replacing n with $n+1$ if needed, we can assume that each codewords $\mathbf{x}$ satisfies the power constraint (4.225) with equality, that is each $\mathbf{x}$ belongs to the set:

$$\mathsf{F}'_n = \{\mathbf{x} \in \mathbb{R}^{L\times n} : ||\mathbf{x}||^2 = nP\}\,. \tag{4.233}$$

To each codeword $\mathbf{x} \in \mathsf{F}'_n$ we associate a *power allocation vector*

$$\mathbf{v}(\mathbf{x}) \in \mathbb{R}^L: \quad v_j(\mathbf{x}) = \frac{1}{n}||\mathbf{x}_{j,\cdot}||^2 = \frac{1}{n}\sum_{i=1}^{n} x_{j,i}^2\,. \tag{4.234}$$

[5]Note the following expression for $V_L(P)$:

$$V_L(P) = 2\left(\frac{\log e}{2}\right)^2 \sum_{j=1}^{L}\left[1-\left(\frac{\sigma_j^2}{T}\right)^2\right]^+\,. \tag{4.228}$$

Therefore $\mathbf{v}$ maps F'_n to the simplex

$$\mathcal{V}(P) \stackrel{\triangle}{=} \{\mathbf{v} : \sum_{i=1}^{L} v_i = P\} \subset \mathbb{R}^L \,. \tag{4.235}$$

To prove the converse, we apply Theorem 33 with the following Q-channel:

$$Q_{\mathbf{Y}|\mathbf{X}=\mathbf{x}} = \prod_{i=1}^{n} \prod_{j=1}^{L} Q_{Y_{j,i}|\mathbf{X}=\mathbf{x}} \,, \tag{4.236}$$

where

$$Q_{Y_{j,i}|\mathbf{X}=\mathbf{x}} = \mathcal{N}\left(0, \sigma_j^2 + v_j(\mathbf{x})\right) \,. \tag{4.237}$$

We need to compute the asymptotics of the following quantity:

$$\beta_\alpha^n(\mathbf{x}) \stackrel{\triangle}{=} \beta_\alpha^n(P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}, Q_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}) \,. \tag{4.238}$$

By spherical symmetry $\beta_\alpha^n(\mathbf{x})$ depends on $\mathbf{x}$ only through $\mathbf{v}(\mathbf{x})$. The Radon-Nikodym derivative between $P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$ and $Q_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$ is distributed under $P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}$ as

$$\log \frac{dP_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}}{dQ_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}} \sim \sum_{i=1}^{n} \sum_{j=1}^{L} C\left(\frac{v_j(\mathbf{x})}{\sigma_j^2}\right) + \frac{\log e}{2} \frac{v_j(\mathbf{x}) + 2\sqrt{v_j(\mathbf{x})}\sigma_j Z_{j,i} - v_j(\mathbf{x}) Z_{j,i}^2}{v_j(\mathbf{x}) + \sigma_j^2} \,, \tag{4.239}$$

where the $Z_{j,i}$'s are i.i.d. standard Gaussian. Then, from Lemma 14 we conclude:

$$\log \beta_\alpha^n(\mathbf{x}) = -n \sum_{i=1}^{L} C\left(\frac{v_i(\mathbf{x})}{\sigma_i^2}\right) - \sqrt{n \sum_{i=1}^{L} V\left(\frac{v_i(\mathbf{x})}{\sigma_i^2}\right)} Q^{-1}(\alpha) - \frac{1}{2}\log n + O(1) \,. \tag{4.240}$$

An important observation is that $O(1)$ term is bounded uniformly in $\mathbf{x} \in \mathsf{F}'_n$ as $n \to \infty$. To establish this technical result by (2.87) and (2.88) we need only to prove that B_n defined there can be uniformly bounded over $\mathbf{x} \in \mathsf{F}'_n$. But by (4.239), we have

$$B_n = 6 \frac{\mathbb{E}\left[|J|^3\right]}{\left(\mathbb{E}\left[J^2\right]\right)^{\frac{3}{2}}} \,, \tag{4.241}$$

where

$$J = \sum_{j=1}^{L} C\left(\frac{v_j}{\sigma_j^2}\right) + \frac{\log e}{2} \frac{v_j + 2\sqrt{v_j}\sigma_j Z_j - v_j Z_j^2}{v_j + \sigma_j^2} \tag{4.242}$$

and the vector $\mathbf{v} = (v_1, \ldots, v_L) \in \mathcal{V}(P)$. By denoting

$$\sigma_{min}^2 = \min_j \sigma_j^2, \tag{4.243}$$

$$\sigma_{max}^2 = \max_j \sigma_j^2, \tag{4.244}$$

we have

$$|J| \leq \sum_{1}^{L} \frac{1}{2} \log \left(1 + \frac{P}{\sigma_{min}^2}\right) + \frac{1}{2} \frac{\log e}{\sigma_{min}^2} \left(P + 2\sqrt{P}\sigma_{max}|Z_j| + PZ_j^2\right), \tag{4.245}$$

where we have used the fact that $v_j \leq P$. Now we see that the right-hand side of (4.245) is independent of the choice of $\mathbf{v}$. Thus there are constants $\zeta_1 \geq 0$ and $\zeta_2 \geq 0$ such that

$$\mathbb{E}\left[|J|^2\right] \leq \zeta_1, \tag{4.246}$$
$$\mathbb{E}\left[|J|^3\right] \leq \zeta_2, \tag{4.247}$$

for any choice of $\mathbf{v} \in \mathcal{V}(P)$. Similarly, since the variance of J is

$$\mathbb{E}\left[J^2\right] = \sum_{j=1}^{L} V\left(\frac{v_j}{\sigma_j^2}\right), \tag{4.248}$$

and since $\sum v_j = P$, we must have at least one $v_j > \frac{P}{L}$, and therefore,

$$\mathbb{E}\left[J^2\right] \geq 2\left(\frac{\log e}{2}\right)^2 \frac{\left(\frac{P}{L}\right)^2 + 2\left(\frac{P}{L}\right)\sigma_{min}^2}{(P + \sigma_{max}^2)^2}, \tag{4.249}$$

where we observe again the right-hand side does not depend on $\mathbf{v}$. By the Lyapunov inequality, $\left(\mathbb{E}\left[|X|^2\right]\right)^{1/2} \leq \left(\mathbb{E}\left[|X|^3\right]\right)^{1/3}$, (4.249) implies a lower bound on $\mathbb{E}\left[|J|^3\right]$ as well. Thus, there exist constants $\zeta_3 > 0$ and $\zeta_4 > 0$ such that

$$\mathbb{E}\left[|J|^2\right] \geq \zeta_3, \tag{4.250}$$
$$\mathbb{E}\left[|J|^3\right] \geq \zeta_4, \tag{4.251}$$

for any choice of $\mathbf{v} \in \mathcal{V}(P)$. Hence, (4.246), (4.247), (4.250) and (4.251) applied to (4.241) imply that

$$0 < \inf_{\mathbf{x} \in \mathsf{F}_n'} B_n \leq \sup_{\mathbf{x} \in \mathsf{F}_n'} B_n < \infty. \tag{4.252}$$

Thus, we have demonstrated that $O(1)$ term in (4.240) is uniform in $\mathbf{x} \in \mathsf{F}_n'$. In particular, we have

$$\inf_{\mathbf{x} \in \mathsf{F}_n'} \log \beta_\alpha^n(\mathbf{x}) = -\sup_{\mathbf{v} \in \mathcal{V}(P)} f_n(\mathbf{v}) - \frac{1}{2}\log n + O(1), \tag{4.253}$$

where

$$f_n(\mathbf{v}) \stackrel{\triangle}{=} n \sum_{i=1}^{L} C\left(\frac{v_i(\mathbf{x})}{\sigma_i^2}\right) + \sqrt{n \sum_{i=1}^{L} V\left(\frac{v_i(\mathbf{x})}{\sigma_i^2}\right)} Q^{-1}(\alpha). \tag{4.254}$$

Since the unique maximizer of $\sum_{i=1}^{L} C\left(\frac{v_i(\mathbf{x})}{\sigma_i^2}\right)$ is given by the waterfilling $v_j = W_j$, Lemma 49 implies

$$\sup_{\mathbf{v} \in \mathcal{V}(P)} f_n(\mathbf{v}) = nC_L(P) + \sqrt{nV_L(P)}Q^{-1}(\alpha) + O(1), \tag{4.255}$$

where C_L and V_L are defined in (4.229) and (4.229), respectively. Therefore, from (4.253) we have

$$\inf_{\mathbf{x}\in\mathsf{F}_n'} \log \beta_\alpha^n(\mathbf{x}) = -nC_L(P) - \sqrt{nV_L(P)}Q^{-1}(\alpha) - \frac{1}{2}\log n + O(1)\,. \tag{4.256}$$

To complete the application of Theorem 33, we need to establish a converse bound for the Q-channel. The following result serves this purpose:

Lemma 79 *There exists a constant $K_3 > 0$ such that for any code with M codewords the maximal probability of error ϵ' over a Q-channel satisfies*

$$1 - \epsilon' \le \frac{K_3 n^{L/2}}{M}\,. \tag{4.257}$$

Assuming Lemma 79, we have from Theorem 33:

$$\inf_{\mathbf{x}\in\mathsf{F}_n'} \beta_\alpha^n(\mathbf{x}) \;\le\; 1 - \epsilon' \tag{4.258}$$

$$\le\; \frac{K_3 n^{L/2}}{M}\,. \tag{4.259}$$

Taking the logarithms of both sides and applying (4.256) we obtain

$$-nC_L(P) - \sqrt{nV_L(P)}Q^{-1}(\alpha) - \frac{1}{2}\log n + O(1) \le -\log M + \frac{L}{2}\log n + O(1)\,, \tag{4.260}$$

which after the rearrangement of terms completes the proof of the following:

Theorem 80 *For the parallel AWGN channel and arbitrary $0 < \epsilon < 1$, we have (maximal probability of error)*

$$\log M^*(n,\epsilon,P) \le nC_L(P) - \sqrt{nV_L(P)}Q^{-1}(\alpha) + \frac{L+1}{2}\log n + O(1)\,. \tag{4.261}$$

Proof of Lemma 79: According to (4.237) the output $\mathbf{Y}$ depends only on $\mathbf{V} = \mathbf{v}(\mathbf{X})$ and moreover $\mathbf{U} = \mathbf{v}(\mathbf{Y})$ is a sufficient statistic of $\mathbf{Y}$ for $\mathbf{X}$. Therefore, an equivalent channel $Q_{\mathbf{U}|\mathbf{V}}$ is defined as

$$U_i = (\sigma_i^2 + V_i)\frac{1}{n}\sum_{j=1}^{n} Z_{i,j}^2\,, \quad i = 1,\ldots,L\,, \tag{4.262}$$

where $Z_{i,j} \sim \mathcal{N}(0,1)$. Note that $\mathbf{V}$ is required to belong to a certain ball in $\mathbb{R}^L$, and that up to probability of order $O\left(e^{-\mathrm{const}\cdot n}\right)$, $\mathbf{U}$ belongs to a slightly larger ball. Therefore, we can assume that the output space has a bounded Lebesgue measure K_4. Then at least for one codeword $\mathbf{v}_0$ the decoding set D_0 must have a Lebesgue measure smaller than $\frac{K_4}{M}$:

$$\mathrm{Leb}[D_0] \le \frac{K_4}{M}\,. \tag{4.263}$$

But $Q_{\mathbf{U}|\mathbf{V}=\mathbf{v}}$ is a product of L copies of a χ^2-distribution and we can show that its density is bounded everywhere by $K_5 n^{L/2}$. Hence, we have

$$1 - \epsilon' \le Q_{\mathbf{U}|\mathbf{V}=\mathbf{v}_0}[D_0] \le K_5 n^{L/2}\mathrm{Leb}[D_0] \le \frac{K_4 K_5 n^{L/2}}{M}\,. \tag{4.264}$$

■

4.5.2 Achievability bound

We plan to apply the $\kappa\beta$ bound, Theorem 27, with the following constraint set:

$$\mathsf{F}_n \stackrel{\triangle}{=} \{\mathbf{x} : v_j(\mathbf{x}) = W_j\}\,, \tag{4.265}$$

where $v_j(\cdot)$ are the coordinates of the power allocation vector $\mathbf{v}(\mathbf{x})$ defined in (4.234) and W_j are the waterfilling powers (4.231). We choose the following output distribution on B:

$$P_{\mathbf{Y}} = \prod_{i=1}^{n}\prod_{j=1}^{L} P_{Y_{j,i}}\,, \tag{4.266}$$

where

$$P_{Y_{j,i}} = \mathcal{N}\left(0, \sigma_j^2 + W_j\right)\,, \tag{4.267}$$

Notice that $P_{\mathbf{Y}} = Q_{\mathbf{Y}|\mathbf{X}=\mathbf{x}_0}$ for some (and any) $\mathbf{x}_0$ with $v_j(\mathbf{x}_0) = W_j$, where Q-channel was defined in (4.236). This motivates the choice of $P_{\mathbf{Y}}$ and also yields by (4.240):

$$\log \beta_\alpha^n(P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}, P_{\mathbf{Y}}) = -nC_L(P) - \sqrt{nV_L(P)}Q^{-1}(\alpha) - \frac{1}{2}\log n + O(1) \tag{4.268}$$

for all $\mathbf{x} \in \mathsf{F}_n$.

To analyze $\kappa_\tau(\mathsf{F}_n, P_{\mathbf{Y}})$ notice that by the spherical symmetry of all measures in each $Y_{j,\cdot}$ sub-component, we can apply the same argument as in Section 4.3.1 to show that κ_τ^n is determined by a test between two distributions on $\mathbb{R}_+^L$:

$$P_0 \sim \left(||\sigma_1\mathbf{Z}_1 + \sqrt{W_1}\mathbf{e}||^2, \ldots, ||\sigma_L\mathbf{Z}_L + \sqrt{W_L}\mathbf{e}||^2\right) \tag{4.269}$$

$$P_1 \sim \left((\sigma_1^2 + W_1)||\mathbf{Z}_1||^2, \ldots, (\sigma_1^2 + W_1)||\mathbf{Z}_L||^2\right)\,, \tag{4.270}$$

where $\mathbf{Z}_j, j = 1, \ldots, L$ are Gaussian vectors with zero mean and covariance matrix equal to the $n \times n$ identity, $\mathbf{Z}_j \sim \mathcal{N}(0, \mathbf{I}_n)$; in (4.269) $\mathbf{e}$ denotes a vector of all 1's of dimension n. Therefore, κ_τ^n can be found as

$$\kappa_\tau^n = \inf_{P_{Z|Y}: P_0[Z=1] \geq \tau} P_1[Z = 1]\,. \tag{4.271}$$

Then similarly to Lemma 71 it can be shown that

$$\kappa_\tau^n \to \kappa_\tau^\infty > 0\,, \quad n \to \infty\,, \tag{4.272}$$

and similarly to Lemma 72 it can be shown that for some constants $C_1 > 0$ and $C_2 > 0$, for all sufficiently large n and for all $\tau \in [0, 1]$ we have

$$\kappa_\tau^n \geq \frac{1}{C_1}\left(\tau - e^{-C_2 n}\right)\,. \tag{4.273}$$

Finally, by Theorem 27 we obtain:

$$\log M^*(n, \epsilon, P) \geq \log \kappa_\tau^n - \log \beta_{1-\epsilon+\tau}^n(\bar{R})\,. \tag{4.274}$$

Choosing $\tau = \frac{1}{\sqrt{n}}$ we get from (4.268) and (4.273):

$$\log M^*(n,\epsilon,P) \geq nC_L(P) - \sqrt{nV_L(P)}Q^{-1}\left(\epsilon - \frac{1}{\sqrt{n}}\right) + O(1) \tag{4.275}$$

$$= nC_L(P) - \sqrt{nV_L(P)}Q^{-1}(\epsilon) + O(1)\,, \tag{4.276}$$

where (4.276) follows by applying Taylor's expansion to Q^{-1}. Therefore, we have shown the following result:

Theorem 81 *For the parallel AWGN channel and arbitrary $0 < \epsilon < 1$, we have (maximal probability of error)*

$$\log M^*(n,\epsilon,P) \geq nC_L(P) - \sqrt{nV_L(P)}Q^{-1}(\alpha) + O(1)\,. \tag{4.277}$$

4.5.3 Proof of the main theorem

Proof of Theorem 78: After applying Theorems 80 and 81, the only remaining claim is to show that the converse bound

$$\log M^*(n,\epsilon,P) \leq nC_L(P) - \sqrt{nV_L(P)}Q^{-1}(\epsilon) + O(\log n) \tag{4.278}$$

holds for the average probability of error formalism. This is done completely as in the proof of (3.111). ■

Before concluding the section on the parallel AWGN, we mention that in the maximal probability of error formalism, the expansion (4.227) still holds for the average power constraint. The method of deriving it is the same as in the AWGN case before: the residual $O(\log n)$ term in the converse should be studied and shown to not have any singularities for any $P > 0$ (cf. the proof of Theorem 74). Then, the upper bound for the average power constraint follows along the same lines as the proof of Corollary 75 since the generalization of (4.33) in Lemma 65 is straightforward.

4.5.4 Deviations from the optimal allocation in the low-power regime

Suppose that we only have a very small power budget P and want to assess the penalty incurred by the power allocations different from the optimal (water-filling) solution. In this section we show that in the low power regime there is virtually no penalty, provided that the available (tiny) drop of power is distributed around the "bottom" of the noise spectrum.

Formally, suppose that $\sigma_1 = \sigma_2 = \ldots = \sigma_L$ and all other $\sigma_j > \sigma_1$. Then, fix an arbitrary vector $\alpha_j, j = 1,\ldots,L,\ldots$, with $\alpha_j = 0$ for $j > L$ and $\sum \alpha_j = 1$. Then it turns out that the capacity under such power allocation is approximately independent of α_j's, namely,

$$\left.\frac{dC(P\cdot\alpha)}{dP}\right|_{P=0} = \frac{\log e}{2\sigma_1^2}\,, \tag{4.279}$$

where

$$C(\mathbf{P}) \stackrel{\triangle}{=} \sum_{j=1}^{\infty} \frac{1}{2}\log\left(1 + \frac{P_j}{\sigma_j^2}\right)\,. \tag{4.280}$$

Consequently, from (4.279) we conclude that the power can be distributed arbitrarily as far as the capacity is concerned. In other words, for very large blocklengths no penalty is incurred by choosing α different from the $\alpha_j^* = \frac{1}{L}1\{j \le L\}$.

Using results of Theorem 78 we can argue that the same conclusion holds for the practical blocklengths as well. Indeed, a straightforward computation shows that the second-order term is also unaffected:

$$\left.\frac{dV(P\cdot\alpha)}{dP}\right|_{P=0} = \frac{\log^2 e}{\sigma_1^2}\,, \tag{4.281}$$

where

$$V(\mathbf{P}) = \sum_{i=j}^{\infty} V_1\left(\frac{P_j}{\sigma_j^2}\right)\,, \tag{4.282}$$

and $V_1(\cdot)$ is the dispersion of the scalar AWGN channel (4.155). Since the right-hand side of (4.281) does not depend on α we see that indeed the conclusion regarding allocation-tolerance of the fundamental limits holds not only in the capacity-term but also in the dispersion term.

4.6 Minimum energy per bit with and without feedback

In this section we investigate the minimum energy per bit E_b required to deliver a k-bit message with probability of error $\epsilon \ge 0$ over an AWGN channel with noise level $\frac{N_0}{2}$ per degree of freedom.

The analysis is made for two different regimes. First, the regime of an apriori fixed rate $R = \frac{k}{n}$ and finite k is considered. Note that in the limit $k \to \infty$, the minimum energy per bit is given by (4.31). Non-asymptotically we use the bounds discussed in Section 4.4 to get an estimate and a tight approximation via (4.218).

Second, we further generalize the problem by dropping the rate restriction. In other words, we consider the minimal achievable energy per bit in the regime of infinite degrees of freedom $n = \infty$, fixed $\epsilon \ge 0$ and finite k. Equivalently, we determine the maximal number of bits of information that can be transmitted with a fixed (non-asymptotic) energy budget and an error probability constraint, but without any limitation on the number of degrees of freedom used. Note that asymptotically, $k \to \infty$, the answer in this case is given by (4.30). Our treatment is different from [30] in that we do not take $k \to \infty$, and from the regime 1 (and from [14, 61]) in that we do not restrict the rate $\frac{k}{n}$. Even though the asymptotic value (4.30) can be obtained from (4.31) (i.e. from the regime of restricted rate) by taking $R \to 0$, such an argument is not possible for finite k. This approach can be viewed as a non-asymptotic extension of [83] in which we also explicitly allow infinitely long codewords (which only serves to strengthen the applicability of our converse bound, since our achievability constructions only use finite-length codewords).

Interestingly, for the second case we will demonstrate how feedback coding can dramatically improve the energy efficiency.

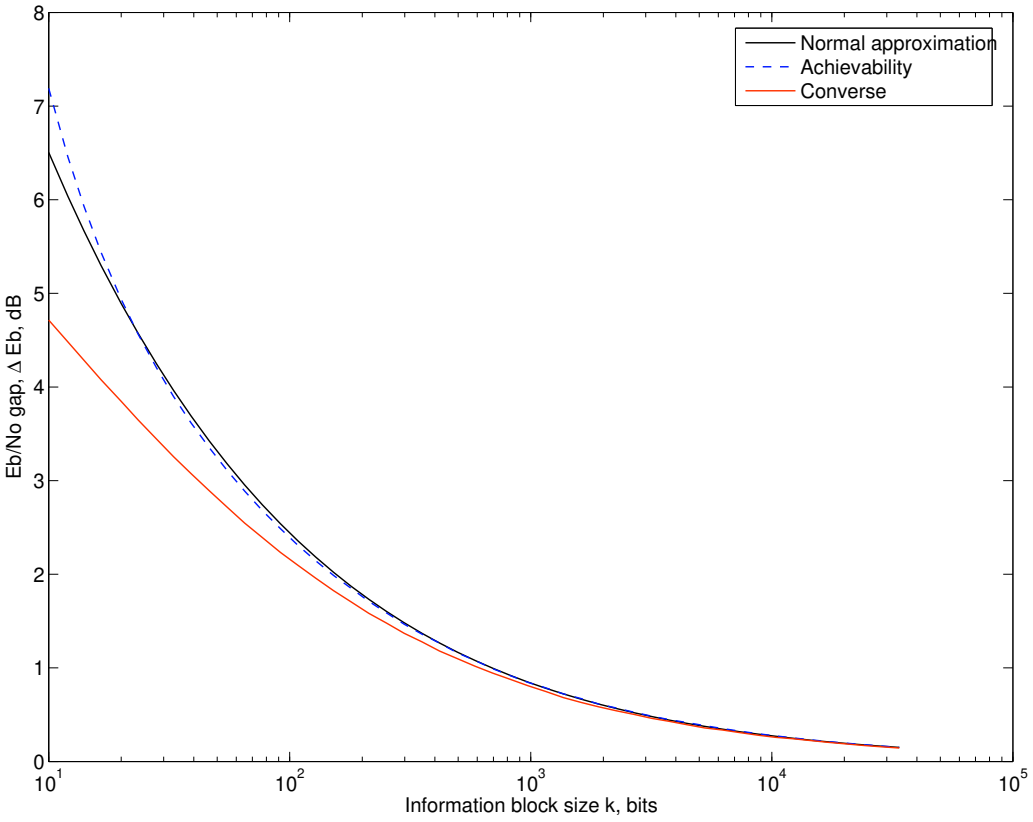


Figure 4.5: Normal approximation for the E_b/N_0 gap for the AWGN channel, $R = 1/2, \epsilon = 10^{-4}$

4.6.1 Fixed rate

In this Section we consider the following setup. A k-bit message is to be communicated with rate $R > 0$ (bits per degree of freedom) over an AWGN channel with $n = \frac{k}{R}$ degrees of freedom and noise level $\frac{N_0}{2}$ per degree of freedom. We are interested in the minimum energy per bit (as function of k) achievable via the best possible coding scheme. Since as $k \to \infty$ the answer is given by (4.31), an interesting figure of merit is the excess energy per bit, $\Delta E_\mathsf{b}(k)$, over that predicted by channel capacity incurred by finiteness of k. ΔE_b as a function of k for a given required bit rate and block error rate ϵ is given by

$$\Delta E_\mathsf{b}(k, R) = 10 \log_{10} \frac{P(n, R, \epsilon)}{\exp(2R) - 1}\,, \tag{4.283}$$

where $P(n, R, \epsilon)$ is the smallest SNR required to achieve block error ϵ at blocklength $n = \frac{k}{R}$ and rate R. The results of Section 4.4 suggest that a tight estimate of $P(n, R, \epsilon)$ can be found from the normal approximation (4.218); namely, $P(n, R, \epsilon)$ is the solution to

$$C(P) - \sqrt{\frac{V(P)}{n}} Q^{-1}(\epsilon) + \frac{1}{2n} \log n = R\,, \tag{4.284}$$

and C and V are as in Theorem 73.

Figure 4.5 gives a representative computation of (4.283)–(4.284) along with the corresponding lower[6] and upper bounds obtained from (4.59) and (4.214) respectively. We note a good precision of the simple approximation (4.283), e.g., for $k = 100$ bits the gap to the achievability bound is only 0.04 dB. A similar comparison (without the normal approximation, of course) for rate 2/3 is presented in [14, Fig. 8].

Note that for the rate 1/2 the asymptotic minimal energy per bit is equal to $0dB$, see (4.31), and hence the reference level for ΔE_{b} in the denominator of (4.284) is simply equal to 1. In this way, on Fig. 4.5 the y-axis represents both the gap and the optimal $\frac{E_{\mathsf{b}}}{N_0}$ at the same time.

4.6.2 No rate constraint

To obtain the *bona fide* energy-information tradeoff we must drop the restriction of the rate, which was made in Section 4.6.1. Indeed, adding additional assumptions increases the required energy and therefore, does not help in computing the absolutely minimal energy needed to convey k bits to the destination.

For clarity, we first introduce complete definitions for this special case. The AWGN channel acts between input space $\mathsf{A} = \mathbb{R}^\infty$ and output space $\mathsf{B} = \mathbb{R}^\infty$ by addition:

$$\mathbf{y} = \mathbf{x} + \mathbf{z}\,, \tag{4.285}$$

where $\mathbb{R}^\infty$ is the vector space of real valued sequences[7] $(x_1, x_2, \ldots, x_n, \ldots)$, $\mathbf{x} \in \mathsf{A}$, $\mathbf{y} \in \mathsf{B}$ and $\mathbf{z}$ is a random vector with i.i.d. Gaussian components $Z_k \sim \mathcal{N}(0, N_0/2)$ independent of $\mathbf{x}$.

Definition 11 *An (E, M, ϵ) code is a list of codewords $(\mathbf{c}_1, \ldots, \mathbf{c}_M) \in \mathsf{A}^M$, satisfying*

$$||\mathbf{c}_j||^2 \le E\,, j = 1, \ldots, M\,, \tag{4.286}$$

and a decoder $g : \mathsf{B} \to \{1, \ldots, M\}$ satisfying

$$\mathbb{P}[g(\mathbf{y}) \ne W] \le \epsilon\,, \tag{4.287}$$

where $\mathbf{y}$ is the response to $\mathbf{x} = \mathbf{c}_W$, and W is the message which is equiprobable on $\{1, \ldots, M\}$. The fundamental energy-information tradeoff is given by

$$M^*(E, \epsilon) = \max\{M : \exists (E, M, \epsilon)\text{-}code\}\,. \tag{4.288}$$

Equivalently, we define the minimum energy per bit:

$$E_{\mathsf{b}}^*(k, \epsilon) = \frac{1}{k} \inf\{E : \exists (E, 2^k, \epsilon)\text{-}code\}\,. \tag{4.289}$$

[6] Another lower bound is given in [61, Fig. 3] which shows [4, (15)].

[7] In this section, boldface letters $\mathbf{x}$, $\mathbf{y}$ etc. denote the infinite dimensional vectors with coordinates X_k, Y_k etc., correspondingly.

Although, we are interested in (4.289), $M^*(E,\epsilon)$ is more suitable for expressing our results and (4.289) is the solution to

$$k = \log M^*(E_{\mathsf{b}}^*(k,\epsilon),\epsilon)\,. \tag{4.290}$$

Note that (4.285) also models an infinite-bandwidth continuous-time Gaussian channel without feedback observed over an interval $[0,T]$, in which each component corresponds to a different tone in an orthogonal frequency division representation. Then, E corresponds to the allowed power P times T and $\frac{N_0}{2}$ is the power spectral density of the white Gaussian noise.

Definition 12 *An (E,M,ϵ) code with feedback is a sequence of encoder functions $\{f_k\}_{k=1}^{\infty}$ determining the channel input as a function of the message W and the past channel outputs,*

$$X_k = f_k(W, Y_1^{k-1})\,, \tag{4.291}$$

satisfying

$$\mathbb{E}\left[||\mathbf{x}||^2|W=j\right] \le E\,, j=1,\ldots,M\,, \tag{4.292}$$

and a decoder $g:\mathsf{B}\to\{1,\ldots,M\}$ satisfying (4.287). The fundamental energy-information tradeoff with feedback is given by

$$M_{\mathsf{f}}^*(E,\epsilon) = \max\{M : \exists (E,M,\epsilon)\text{-code with feedback}\} \tag{4.293}$$

and the minimum energy per bit by

$$E_{\mathsf{f}}^*(k,\epsilon) = \frac{1}{k}\inf\{E : \exists (E,2^k,\epsilon)\text{-code with feedback}\}\,. \tag{4.294}$$

Similarly to the approach we have taken for studying $\log M^*(n,\epsilon)$ as a function of n, in this section we concentrate on obtaining upper and lower bounds on $\log M^*(E,\epsilon)$ and $\log M_{\mathsf{f}}^*(E,\epsilon)$ and corresponding asymptotics for fixed ϵ and $E\to\infty$.

Theorem 82 *For every $M>0$ there exists an (E,M,ϵ) code for the channel (4.285) with*

$$\epsilon \;=\; \mathbb{E}\left[\min\left\{MQ\left(\sqrt{\frac{2E}{N_0}}+Z\right),1\right\}\right]\,, \tag{4.295}$$

and $Z\sim\mathcal{N}(0,1)$. Conversely, any (E,M,ϵ) code without feedback satisfies

$$\frac{1}{M} \ge Q\left(\sqrt{\frac{2E}{N_0}}+Q^{-1}(1-\epsilon)\right)\,. \tag{4.296}$$

Proof: To prove (4.295), notice that a codebook with M orthogonal codewords under a maximum likelihood decoder has probability of error equal to

$$P_e = 1-\frac{1}{\sqrt{\pi N_0}}\int_{-\infty}^{\infty}\left[1-Q\left(\sqrt{\frac{2}{N_0}}z\right)\right]^{M-1} e^{-\frac{(z-\sqrt{E})^2}{N_0}}dz\,. \tag{4.297}$$

A change of variables $x = \sqrt{\frac{2}{N_0}}z$ and application of the bound $1-(1-y)^{M-1} \le \min\{My, 1\}$ weakens (4.297) to (4.295).

To prove (4.296) fix an arbitrary codebook $(\mathbf{c}_1, \ldots, \mathbf{c}_M)$ and a decoder $g : \mathsf{B} \to \{1, \ldots, M\}$. We denote the measure $P^j = P_{\mathbf{y}|\mathbf{x}=\mathbf{c}_j}$ on $\mathsf{B} = \mathbb{R}^\infty$ as the infinite dimensional Gaussian distribution with mean $\mathbf{c}_j$ and independent components with individual variances equal to $\frac{N_0}{2}$; i.e.,

$$P^j = \prod_{k=1}^{\infty} \mathcal{N}\left(c_{j,k}, \frac{N_0}{2}\right), \quad n = 1, 2, \ldots \tag{4.298}$$

where $c_{j,k}$ is the k-th coordinate of the vector $\mathbf{c}_j$. We also define an auxiliary measure

$$\Phi = \prod_{k=1}^{\infty} \mathcal{N}\left(0, \frac{N_0}{2}\right), \quad n = 1, 2, \ldots \tag{4.299}$$

Assume for now that the following holds for each j and event $F \in \mathcal{B}^\infty$:

$$P^j(F) \ge \alpha \Longrightarrow \Phi(F) \ge \beta_\alpha(E), \tag{4.300}$$

where the right-hand side of (4.296) is denoted by

$$\beta_\alpha(E) = Q\left(\sqrt{\frac{2E}{N_0}} + Q^{-1}(\alpha)\right). \tag{4.301}$$

From (4.300) we complete the proof of (4.296):

$$\begin{aligned}
\frac{1}{M} &= \frac{1}{M}\sum_{j=1}^{M} \Phi(g^{-1}(j)) && (4.302)\\
&\ge \frac{1}{M}\sum_{j=1}^{M} \beta_{P^j(g^{-1}(j))}(E) && (4.303)\\
&\ge \beta_{1-\epsilon}(E), && (4.304)
\end{aligned}$$

where (4.302) follows because $g^{-1}(j)$ partitions the space B, (4.303) follows from (4.300), and (4.304) follows since the function $\alpha \to \beta_\alpha(E)$ is non-decreasing convex for any E and

$$\frac{1}{M}\sum_{j=1}^{M} P^j(g^{-1}(j)) \ge 1 - \epsilon \tag{4.305}$$

is equivalent to (4.287), which holds for every (E, M, ϵ) code.

To prove (4.300) we compute the Radon-Nikodym derivative

$$\log_e \frac{dP^j}{d\Phi}(\mathbf{y}) = \sum_{k=1}^{\infty} \left(-\tfrac{1}{2}c_{j,k}^2 + c_{j,k}Y_k\right), \tag{4.306}$$

and hence $\log_e \frac{dP^j}{d\Phi}$ is distributed as

$$\log_e \frac{dP^j}{d\Phi}(\mathbf{y}) \sim \mathcal{N}\left(\frac{||\mathbf{c}_j||^2}{2}, N_0\frac{||\mathbf{c}_j||^2}{2}\right) \tag{4.307}$$

if $\mathbf{y} \sim P^j$ and as

$$\log_e \frac{dP^j}{d\Phi}(\mathbf{y}) \sim \mathcal{N}\left(-\frac{||\mathbf{c}_j||^2}{2}, N_0\frac{||\mathbf{c}_j||^2}{2}\right) \tag{4.308}$$

if $\mathbf{y} \sim \Phi$. Then, (4.300) follows by the Neyman-Pearson lemma since $||\mathbf{c}_j||^2 \leq E$ for all $j \in \{1, \ldots, M\}$. This method of proving a converse result is in the spirit of the meta-converse, Theorem 28. ■

Theorem 83 *In the absence of feedback, the number of bits that can be transmitted with energy E and error probability $0 < \epsilon < 1$ behaves as*

$$\log M^*(E, \epsilon) = \frac{E}{N_0}\log e - \sqrt{\frac{2E}{N_0}}Q^{-1}(\epsilon)\log e + \frac{1}{2}\log\frac{E}{N_0} + O(1) \tag{4.309}$$

as $E \to \infty$.

Proof: To obtain (4.309) fix $0 < \epsilon < 1$ and denote

$$x^* = \sqrt{\frac{2E}{N_0}} + Q^{-1}\left(1 - \epsilon + \sqrt{\frac{2N_0}{E}}\right). \tag{4.310}$$

We now choose $M = \frac{1}{Q(x^*)}$ and observe that we have

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \min(MQ(x), 1)e^{-\frac{1}{2}\left(x-\sqrt{\frac{2E}{N_0}}\right)^2}dx && (4.311)\\
&= 1 - Q\left(x^* - \sqrt{\frac{2E}{N_0}}\right) + \frac{M}{\sqrt{2\pi}}\int_{x^*}^{+\infty} Q(x)e^{-\frac{1}{2}\left(x-\sqrt{\frac{2E}{N_0}}\right)^2}dx && (4.312)\\
&= \epsilon - \sqrt{\frac{2N_0}{E}} + \frac{M}{\sqrt{2\pi}}\int_{x^*}^{+\infty} Q(x)e^{-\frac{1}{2}\left(x-\sqrt{\frac{2E}{N_0}}\right)^2}dx && (4.313)\\
&\leq \epsilon - \sqrt{\frac{2N_0}{E}} + \frac{M}{2\pi x^*}\int_{x^*}^{+\infty} e^{-\frac{1}{2}\left(x-\sqrt{\frac{2E}{N_0}}\right)^2 - \frac{x^2}{2}}dx && (4.314)\\
&= \epsilon - \sqrt{\frac{2N_0}{E}} + \frac{e^{-\frac{E}{2N_0}}Q\left(\sqrt{2}x^* - \sqrt{\frac{E}{N_0}}\right)}{2\sqrt{\pi}x^*Q(x^*)} && (4.315)\\
&= \epsilon - \sqrt{\frac{2N_0}{E}} + \sqrt{\frac{N_0}{E}}(1 + o(1)), && (4.316)
\end{aligned}$$

as $E \to \infty$, where in (4.314) and (4.316) we used [84, (3.53)]

$$Q(x) \leq \frac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}} \tag{4.317}$$

and

$$\log Q(x) = -\frac{x^2\log e}{2} - \log x - \frac{1}{2}\log 2\pi + o(1), x \to \infty \tag{4.318}$$

respectively. Therefore, for E large enough (4.316) falls below ϵ and, consequently, by (4.295) there exists an (M, E, ϵ) code for the chose $M = \frac{1}{Q(x^*)}$. In other words, for such E we have demonstrated

$$\log M^*(E,\epsilon) \geq -\log Q\left(\sqrt{\frac{2E}{N_0}} + Q^{-1}\left(1-\epsilon+\sqrt{\frac{2N_0}{E}}\right)\right). \tag{4.319}$$

Using the expansion (4.318) in (4.319) and (4.296), we obtain (4.309). ■

As discussed above, Theorems 82 and 83 may be interpreted in the context of the infinite-bandwidth continuous-time Gaussian channel with noise spectral density $\frac{N_0}{2}$. Indeed, denote by $M_c^*(T,\epsilon)$ the maximum number of messages that is possible to communicate over such a channel over the time interval $[0,T]$ with probability of error ϵ and power-constraint P. According to Shannon [30] we have

$$\lim_{T\to\infty} \frac{1}{T}\log M_c^*(T,\epsilon) = \frac{P}{N_0}\log e\,. \tag{4.320}$$

Theorem 83 sharpens (4.320) to

$$\log M_c^*(T,\epsilon) = \frac{PT}{N_0}\log e - \sqrt{\frac{2PT}{N_0}}Q^{-1}(\epsilon)\log e + \frac{1}{2}\log\frac{PT}{N_0} + O(1) \tag{4.321}$$

as $T\to\infty$. Furthermore, Theorem 82 provides tight non-asymptotic bounds on $\log M_c^*(T,\epsilon)$.

We now proceed to the case of feedback coding.

Theorem 84 *Let $0 \leq \epsilon < 1$. Any (E, M, ϵ) code with feedback for the channel* (4.285) *must satisfy*

$$d\left(1-\epsilon||\tfrac{1}{M}\right) \leq \frac{E}{N_0}\log e\,, \tag{4.322}$$

where $d(x||y) = x\log\frac{x}{y} + (1-x)\log\frac{1-x}{1-y}$ is the binary relative entropy.

Note that in the special case $\epsilon = 0$ (4.322) reduces to

$$\log M \leq \frac{E}{N_0}\log e\,. \tag{4.323}$$

Theorem 85 *For any $E > 0$ and positive integer M there exists an (E, M, ϵ) code with feedback for the channel* (4.285) *satisfying*

$$\epsilon \leq \inf\left\{1-\alpha+(M-1)\beta\right\}, \tag{4.324}$$

where the infimum is over all $0 < \beta < \alpha \leq 1$ satisfying

$$d(\alpha||\beta) = \frac{E}{N_0}\log e\,. \tag{4.325}$$

Moreover, there exists an (E, M, ϵ) decision feedback code, which uses the feedback link only once to send a "ready-to-decode" signal; its probability of error is bounded by (4.324) *with $\alpha = 1$, namely,*

$$\epsilon \leq (M-1)e^{-\frac{E}{N_0}}. \tag{4.326}$$

Proofs of Theorems 84 and 85 may be found in Appendix D.

The asymptotic behavior with feedback is given by

Theorem 86 *In the presence of feedback, the number of bits that can be transmitted with energy E and error probability $0 < \epsilon < 1$ behaves as*

$$\log M_{\mathsf{f}}^*(E,\epsilon) = \frac{E}{N_0}\frac{\log e}{1-\epsilon} + O\left(\log \frac{E}{N_0}\right) \tag{4.327}$$

as $E \to \infty$. More precisely, we have

$$\begin{aligned} \frac{E}{N_0}\frac{\log e}{1-\epsilon} - \log\frac{E}{N_0} + O(1) &\leq \log M_{\mathsf{f}}^*(E,\epsilon) &(4.328)\\ &\leq \frac{E}{N_0}\frac{\log e}{1-\epsilon} + \frac{h(\epsilon)}{1-\epsilon}\,, &(4.329)\end{aligned}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy function.

Proof: Bound (4.328) follows from (4.324) by taking

$$\alpha = 1 - \epsilon + \frac{N_0}{E}\,. \tag{4.330}$$

Alternatively, (4.328) can be achieved by sending the all-zero codeword with probability $1 - \frac{(1-\epsilon)E}{E-1}$ and otherwise using an $(\frac{E-1}{1-\epsilon}, \frac{1}{E}, M)$ decision feedback code guaranteed to exist by Theorem 85 and (4.326). Bound (4.329) follows from (4.322) and

$$d(\alpha||\beta) \geq \alpha\log\frac{1}{\beta} - h(\alpha)\,. \tag{4.331}$$

■

Note that as $\epsilon \to 0$, the leading term in (4.327) coincides with the leading term in (4.309). As we know, in the regime of arbitrarily reliable communication (and therefore $k \to \infty$) feedback does not help.

At first sight it may be plausible that infinite bandwidth may allow finite energy per bit when zero-error is required. However, a simple consequence of [85] is that without feedback

$$\log M^*(E,0) = 0 \tag{4.332}$$

for all $E > 0$. With noiseless feedback the situation changes.

Theorem 87 *For any positive integer k and $E > kN_0$ there exists an $(E, 2^k, 0)$-code with feedback. Equivalently, for all positive integers k we have*

$$E_{\mathsf{f}}^*(k,0) \leq N_0\,. \tag{4.333}$$

Proof: An $(E_1, M_1, 0)$ code and an $(E_2, M_2, 0)$ code can be combined into an $(E_1 + E_2, M_1M_2, 0)$ code by using the first one on odd channel inputs and the second one on even. This also shows that function $E_{\mathsf{f}}^*(\cdot, 0)$ is non-increasing. Therefore, to prove the theorem, it is sufficient to prove that for any $E > N_0$ there exists an $(E, 2, 0)$ code with feedback. To

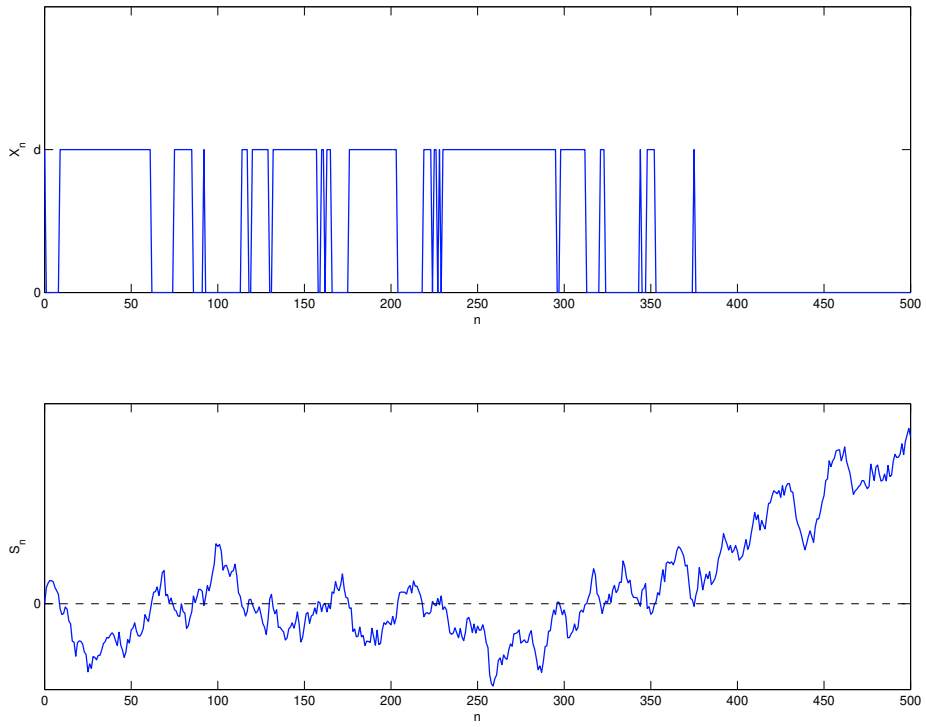


Figure 4.6: Illustration of the zero-error feedback code of Theorem 87, conditioned on $W = +1$.

this end, we construct the following binary communication scheme. Fix an arbitrary $d > 0$, assume $W = \pm 1$ and consider the following code with feedback:

$$f_n(W, Y_1^{n-1}) = \begin{cases} Wd, & i(W; Y^{n-1}) \leq i(-W; Y^{n-1}) \,, \\ 0, & \text{otherwise} \end{cases} \tag{4.334}$$

where we have defined information densities

$$i(w; y_1^k) = \sum_{j=1}^{k} \log \frac{P_{Y_j|X_j}(y_j|f_j(w; y_1^{j-1}))}{P_{Y_j|Y_1^{j-1}}(y_j|y_1^{j-1})} \,. \tag{4.335}$$

Since the alternative in (4.334) depends on the difference of the information densities, it is convenient to define

$$\begin{aligned} S_n &= \log \frac{\mathbb{P}[W = +1|Y^n]}{\mathbb{P}[W = -1|Y^n]} && (4.336) \\ &= i(+1; Y_1^n) - i(-1; Y_1^n) \,. && (4.337) \end{aligned}$$

The main observation is that assuming $W = +1$ and regardless of the alternative in (4.334) we have for each $n > 1$

$$S_n = S_{n-1} + \tfrac{1}{2}d^2 + d\, Z_n \,. \tag{4.338}$$

Fig. 4.6 represents the typical joint behavior of the channel input X_n and the process S_n when the encoder is sending $W = +1$. From (4.338) we see that under $W = +1$, S_n is a submartingale drifting towards $+\infty$. Since the transmitter outputs $X_n = +d$ only when $S_n < 0$ and otherwise outputs $X_n = 0$, the positive drift of S_n implies that only finitely many X_n's will be different from zero with probability one. Another conclusion is that the measures $P_{Y^\infty|W=+1}$ and $P_{Y^\infty|W=-1}$ are mutually singular and therefore W can be recovered from Y^∞ with zero error.

To finish the proof, we need to compute the average energy spent by our scheme. It is easy to see that (again conditioning on $W = +1$)

$$||\mathbf{x}||^2 = \sum_{j=1}^{\infty} ||X_j||^2 = \sum_{j=1}^{\infty} d^2 1\{S_j \le 0\}\,. \tag{4.339}$$

To simplify the computation of $\mathbb{E}\left[||\mathbf{x}||^2\right]$, we replace $d\,Z_n$ in (4.338) with $W_{nd^2} - W_{(n-1)d^2}$, where W_t is a standard Wiener process. In this way, we can write

$$S_n = \left(\frac{s}{2} + \sqrt{\tfrac{N_0}{2}} W_s\right)|_{s=nd^2}\,, \tag{4.340}$$

i.e. S_n is just a sampling of W_t on a d^2-spaced grid. According to (4.339), $||\mathbf{x}||^2$ is a total number of negative samples multiplied by a grid step. Since every realization of W_t is continuous, as $d \to 0$ the $||\mathbf{x}||^2$ tends to the total time the Brownian motion $\frac{t}{2} + \sqrt{\frac{N_0}{2}} W_t$ spends below zero:

$$\lim_{d\to 0} ||X||^2 = T = \int_0^\infty 1_{\left\{\frac{t}{2} + \sqrt{\frac{N_0}{2}} W_t \le 0\right\}} dt\,. \tag{4.341}$$

Then, taking expectations we get that the average energy spent to transmit 1 bit is

$$\begin{aligned}
\mathbb{E}\left[T\right] &= \int_0^\infty \mathbb{P}\left[\frac{t}{2} + \sqrt{\tfrac{N_0}{2}} W_t \le 0\right] dt && (4.342)\\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2}} \int_0^\infty 1_{\left\{x > \sqrt{\frac{t}{2N_0}}\right\}} dx dt && (4.343)\\
&= \frac{2N_0}{\sqrt{2\pi}} \int_0^\infty x^2 e^{-\frac{x^2}{2}} dx = N_0\,. && (4.344)
\end{aligned}$$

Hence, $M_{\mathsf{f}}^*(E, 0) \ge 2$ for any $E > N_0$, as required. ■

The weaker result that $M_{\mathsf{f}}^*(E, 0) \ge 2$ for sufficiently large E follows from [86, Lemma 4.2], which analyzes a modification of an original method of Zigangirov [87]. In contrast, our method is motivated by the Brownian motion analysis and antipodal signaling arising in the achievability proof of Theorem 85. At the expense of a significantly more involved analysis, the bound in Theorem 87 can be further improved by using multidimensional constellations. It remains to be seen whether such a method could close the gap with the upper bound in (4.323).

Before concluding this section, we discuss implications of the results shown. As the number of information bits, k, goes to infinity, the minimum energy per bit required for arbitrarily reliable communication is equal to -1.59 dB with or without feedback. However,

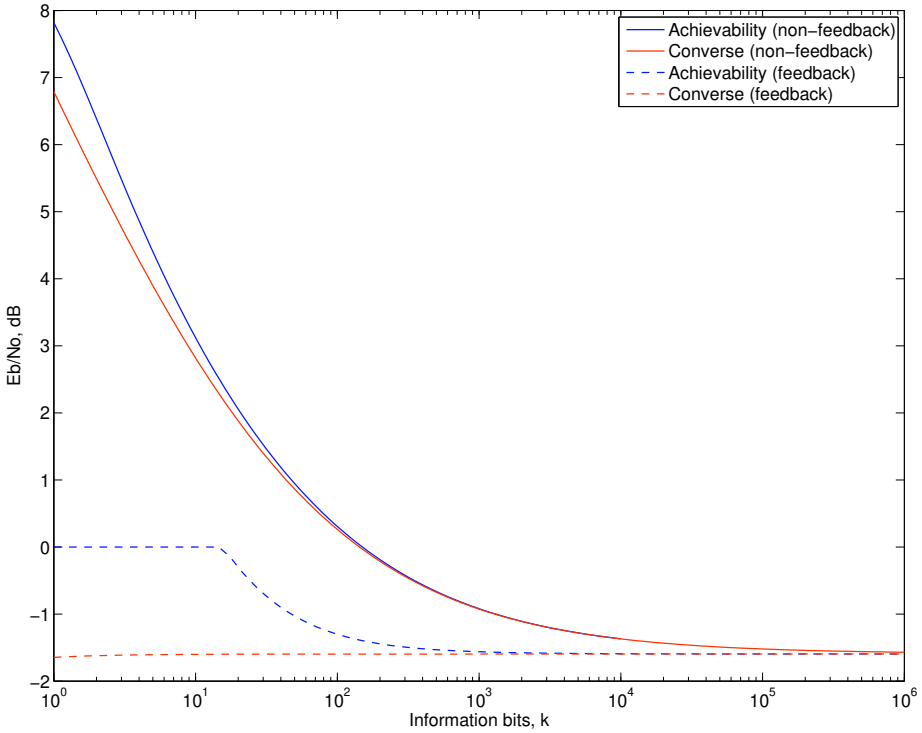


Figure 4.7: Bounds on the minimum energy per bit as a function of the number of information bits with and without feedback; block error rate $\epsilon = 10^{-3}$.

in the non-asymptotic regime, in which the block error probability is set to ϵ, the minimum energy per bit may substantially reduced thanks to the availability of feedback. Comparing Theorems 83 and 86, we observe a double benefit: feedback reduces the leading term in the minimum energy by a factor of $1 - \epsilon$, and the penalty due to the second-order term in (4.309) disappears. Moreover, Theorem 87 has demonstrated that thanks to availability of an infinite number of degrees of freedom, feedback enables zero-error transmission of any number of bits with finite energy per bit. This complements the famous result of Schalkwijk and Kailath [88], that in the fixed rate setup one achieves significantly better reliabilities over the AWGN with feedback.

Our bounds enable a quantitative analysis of the dependence of the required energy on the number of information bits. In Fig. 4.7 we take $\epsilon = 10^{-3}$ and compare the bounds on $E_{\mathsf{b}}^*(k, \epsilon)$ and $E_{\mathsf{f}}^*(k, \epsilon)$ given by Theorem 82 and Theorems 84, 85 and 87, respectively. The non-feedback upper (4.295) and lower (4.296) bounds are tight enough to conclude that for messages of size $k \sim 100$ bits the minimum $\frac{E_{\mathsf{b}}}{N_0}$ is 0.20 dB, whereas the Shannon limit of -1.59 dB is only approachable at $k \sim 10^5 - 10^6$ bits. In contrast, with feedback the upper bound, which is the best of (4.324) and (4.333), and the lower bound (4.322) demonstrate the significant advantages of using feedback with practical values of k; e.g., with feedback, -1.5 dB is achievable already at $k \sim 200$.

Surprisingly, our results demonstrate that virtually all the benefits of feedback are realized by codes that use the feedback link only to send a single "stop transmission" signal

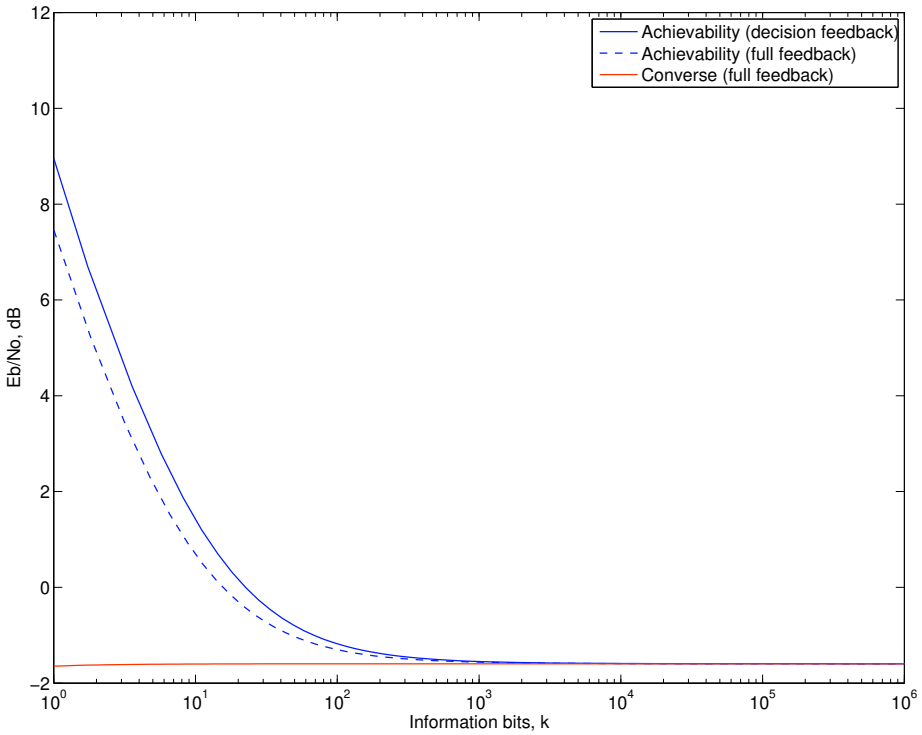


Figure 4.8: Comparison of the achievability bounds on the minimum energy per bit as a function of the number of information bits with decision feedback and full feedback; block error rate $\epsilon = 10^{-3}$.

(as opposed to requiring a full noiseless feedback available at the transmitter). Indeed, the proof of Theorem 86 demonstrates that the asymptotic expansion (4.327) does not change if we restrict attention to decision feedback codes. Moreover, in Fig. 4.8 we compare the decision feedback achievability bound (4.326) against the bound (4.324) which requires full feedback. It can be seen that numerically the difference between the two is insignificant compared to the gain with respect to the non-feedback codes; see Fig. 4.7. In this way, the results of Theorems 85 and 86 extend to noisy and/or finite capacity feedback links.

Finally, it is interesting to investigate the difference in minimum energy per bit between the setup analyzed in this section and that of Section 4.6.1. In Fig. 4.9 we compare the normal approximation curve taken from[8] Fig. 4.5 against the non-feedback achievability-converse bounds of Theorem 82, as computed in Fig. 4.7. This comparison explicitly demonstrates how much energy per bit is lost non-asymptotically due to restricting the rate to 1/2. As rate goes to zero, the rate-constrained curve approaches the ultimate fundamental limit given by (the bounds of) Theorem 82. This is a non-asymptotic version of the argument when one takes the limit of $R \to 0$ in (4.31) to obtain (4.30).

[8]Note that a slight discrepancy with Fig. 4.5 is explained by the change of ϵ compared to the setup of Fig. 4.9.

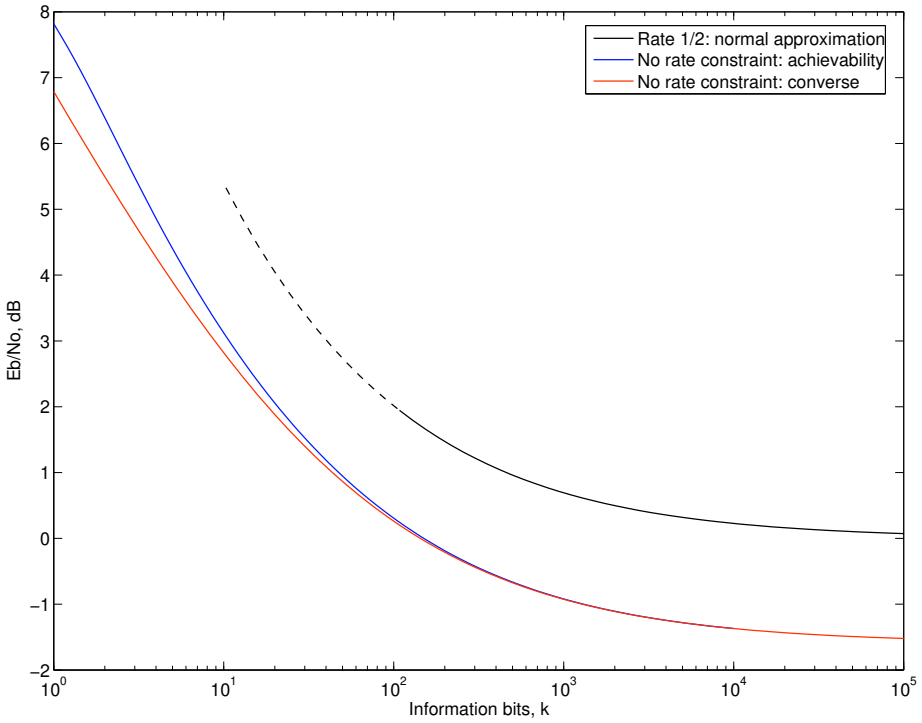


Figure 4.9: Comparison of the minimum achievable energy per bit (without feedback) as a function of the number of information bits k in two regimes: fixed rate $R = 1/2$ and no rate constraints; block error probability is $\epsilon = 10^{-3}$.

Chapter 5

Normal approximation

In previous chapters it was demonstrated how channel dispersion captures the behavior of fundamental limits near capacity. This chapter considers various ramifications of this fact. The question of attaining a given fraction of the capacity is discussed in Section 5.1. It is demonstrated that the normal approximation (2.23) results in a significantly more precise prediction of the coding blocklength than does error-exponent based analysis. In Section 5.2 a few practical families of codes are compared against the fundamental limits on the AWGN channel and the BSC. The dispersion of a parallel DMC is considered in Section 5.3 and in particular the performance loss due to coding separately on each constituent DMC is quantified. Section 5.4 proves an order-optimal bound on the dispersion of the DMC in terms of its alphabet sizes. The material in Sections 5.1-5.4 has been presented in part in [32]; the discussion of classical binary codes over the BSC in Section 5.2 and the result about the dispersion of parallel channels, Section 5.3, are new.

Based on the dispersion of the Gilbert-Elliott channel obtained before it is shown in Section 5.5 that when the channel evolves dynamically, the capacity considerations alone may lead to completely wrong design decisions. In such questions taking dispersion of the channel into account becomes crucial. This material has appeared in [57]. Finally, in Section 5.6 we apply our methods to characterize the optimal decay of the probability of error when the rate converges to the capacity slower than $\frac{1}{\sqrt{n}}$. The investigation of such questions has been recently initiated by Altug and Wagner [82] (for the case of the DMC). Results in Section 5.6 are published here for the first time.

5.1 Comparison to the error-exponent approximation

As we discussed in the introduction, the main purpose of computing asymptotic quantities (such as capacity and dispersion) is to obtain non-asymptotic approximations. Following this rationale, in the previous chapters we have shown how knowledge of the capacity-dispersion pair (and of a $\log n$ term, sometimes) yields an approximation (2.23), which compares very favorably with the (bounds on the) true value; see Sections 3.2.3, 3.3.3, 3.5.3, 3.6.2 and 4.4.

Having a tight approximation to the value of the fundamental limit $\log M^*(n, \epsilon)$ opens many practical applications. For example, we may estimate the minimal blocklength n

Table 5.1: Bounds on the minimal blocklength n needed to achieve $R = 0.9C$

Channel	Converse	RCU	DT or $\kappa\beta$	Error-exp.	Norm. Ap.
BEC(0.5), $\epsilon = 10^{-3}$	$n \geq 899$	$n \leq 1021$	$n \leq 991$	$n \approx 1380$	$n \approx 955$
BSC(0.11), $\epsilon = 10^{-3}$	$n \geq 2985$	$n \leq 3106$	$n \leq 3548$	$n \approx 4730$	$n \approx 3150$
AWGN(0dB), $\epsilon = 10^{-3}$	$n \geq 2550$	$n \leq 2814$	$n \leq 3400$	$n \approx 4120$	$n \approx 2750$
AWGN(20dB), $\epsilon = 10^{-6}$	$n \geq 147$	$n \leq 188$	$n \leq 296$	$n \approx 220$	$n \approx 190$

needed to achieve a fraction η of capacity, see (2.24):

$$n \gtrsim \left(\frac{Q^{-1}(\epsilon)}{1-\eta}\right)^2 \frac{V}{C^2}\,. \tag{5.1}$$

Recall that the reliability function $E(R)$ for the rate $0 < R < C$ is defined as (provided that the limit exists):

$$E(R) = \lim_{n\to\infty} -\frac{1}{n}\log \epsilon^*(n, 2^{nR})\,, \tag{5.2}$$

where

$$\epsilon^*(n, M) = \inf\{\epsilon : \exists(n, M, \epsilon)\text{-code}\}\,, \tag{5.3}$$

i.e. a functional inverse of the fundamental limit $\epsilon \to M^*(n, \epsilon)$. For some memoryless channels $E(R)$ is known, at least in the region $R_{cr} \leq R < C$, where R_{cr} is a so-called critical rate of the channel.

The rationale of the definition of $E(R)$ is in obtaining an *error-exponent approximation:*

$$\epsilon^*(n, M) \approx \exp\left(-nE\left(\frac{\log M}{n}\right)\right)\,. \tag{5.4}$$

Therefore, according to the (5.4) the minimal blocklength n needed to achieve a fraction η of capacity with a given probability of error ϵ should be approximately:

$$n \gtrsim -\frac{1}{E(\eta C)}\log \epsilon\,. \tag{5.5}$$

Which of approximations (5.1) and (5.5) is better? To answer this question in the Table 5.1 we show the numerical results for the blocklength required by the converse, guaranteed by the achievability and predicted by error-exponents and normal approximation[1] for achieving rate $R = 0.9C$.

Clearly we can see that the normal approximation is superior in this regime. Together with the extensive comparisons (e.g., Fig. 3.3, 3.8, 4.3 and 4.4) Table 5.1 demonstrates that the asymptotic analysis undertaken in this thesis, such as needed for the proof of (2.22), is not just a mathematical curiosity, but rather a tool especially useful for a practically important range $n \sim 10^3$, $\epsilon \sim 10^{-3}$.

As another observation, notice that according to (5.1) the quantity $\frac{V}{C^2}$ is important for determining the "coding horizon" of a channel. Interestingly, for all (families of) channels

[1]For the BSC and the AWGN channel we use the approximation formula (3.59) which has an additional $\frac{1}{2}\log n$ term. For the AWGN channel the DT bound is replaced by the $\kappa\beta$ bound.

considered in this thesis – including the AWGN channel and the BSC – the fraction $\frac{V}{C^2}$ blows up when the noise level increases without bound. The meaning of this fact is clear: to achieve a fraction of a low-capacity link it is necessary to code over a large blocklength.

To give a possible reason why error-exponent asymptotics is less important for finite blocklength recall that obtaining the value of $E(R)$ requires a pair of bounds, for example, Gallager's random-coding and Shannon-Gallager-Berlekamp sphere-packing:

$$\exp(-nE(R-o_1(1))+o_2(n)) \le \epsilon^*(n, \exp\{nR\}) \le 4\exp(-nE(R))\,. \tag{5.6}$$

For the BSC the expression for $o_1(1)$ and $o_2(n)$ can be taken from [9, (5.6.41),(5.8.21)]. Even neglecting the presence of $o_1(1)$ we see that ratio of the upper bound to the lower bound is approximately

$$\frac{\text{upper-bound on } \epsilon^*}{\text{lower-bound on } \epsilon^*} \approx \frac{4\sqrt{8n}}{\delta}\,. \tag{5.7}$$

That is, if we take moderate $n \sim 10^3$ and $\delta = 0.11$ we get that this fraction is $\approx 3\cdot 10^3$. Although the sub-exponential factor $\frac{4}{\sqrt{8n}}$ is completely irrelevant for the asymptotics, for the regime of $\epsilon \approx 10^{-3}$ the effect of such a factor is huge. Consequently, near the capacity such sub-exponential (and rate dependent!) factors do more harm to the approximation (5.4) than sub-logarithmic factors do to (2.22).

5.2 Practical codes

It is interesting to compare performance of the codes actually used in practice against the finite blocklength fundamental limits. One such comparison is given in Fig. 5.1 where the lower curve depicts the performance of a certain family of multi-edge low-density parity-check (ME-LDPC) codes decoded via a low-complexity belief-propagation decoder [89]. We notice that in the absence of the non-asymptotic finite-blocklength curves, one has to compare the performance against the capacity alone. Such comparison leads to an incorrect conclusion that a given family of codes becomes closer to optimal with increasing blocklength. In reality we see that the relative gap to the finite blocklength fundamental limit is approximately constant. In other words, the fraction $\frac{\log M_{LDPC}(n,\epsilon,P)}{\log M^*(n,\epsilon,P)}$ seems to be largely blocklength independent.

This observation leads us to a natural way of comparing two different codes over a given channel. Over the AWGN channel the codes have traditionally been compared in terms of E_b/N_0. Such comparison, although justified for a low-rate codes, unfairly penalizes higher rate codes. Instead, we define a normalized rate of a code with M codewords as (this can be extended to discrete channels parametrized by a scalar in a natural way)

$$R_{norm}(\epsilon) = \frac{\log M}{\log M^*(n,\epsilon,\gamma_{min}(\epsilon))}\,, \tag{5.8}$$

where $\gamma_{min}(\epsilon)$ is the smallest SNR at which the code still admits decoding with probability of error below ϵ. Here and below ϵ is chosen to be a fixed number representing the required reliability level. Since the true value of M^* in (5.8) is in general unknown, instead of one value for R_{norm} we get an interval (by bounding M^*). However, as Fig. 5.1 demonstrates

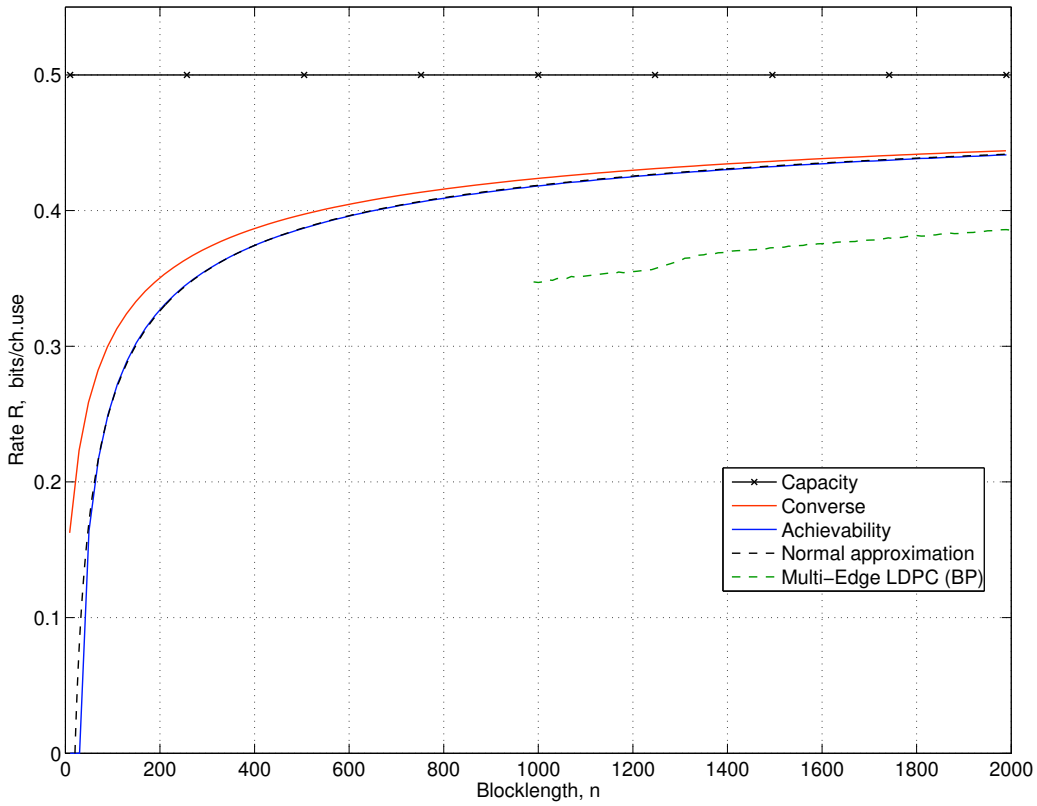


Figure 5.1: Normal approximation for the AWGN channel, $SNR = 0$ dB, $\epsilon = 10^{-3}$. The LDPC curve demonstrates the performance achieved by a particular family of multi-edge LDPC codes (designed by T. Richardson).

typically, instead of hard to compute bounds we may simply use the normal approximation (4.218) to get an approximation to (5.8) virtually no loss of precision for blocklength as low as 100.

The evolution of the coding schemes from 1980s (Voyager) to 2009 in terms of the normalized rate $R_{norm}(10^{-4})$ is presented on Fig. 5.2. ME-LDPC is the same family as in Fig. 4.3 [89] and the rest of the data is taken from [61]. A comparison of certain turbo codes to Feinstein's bound and Shannon's converse can also be found on Fig. 6 and 7 of [14].

Of course, the definition of R_{norm} can be extended to any other single-parameter family of channels (ordered by degradation [90]), such as BSCs or BECs.

For the BSC a sample of popular algebraic and state-of-the-art LDPC codes is compared in terms of $R_{norm}(10^{-3})$ in Fig. 5.3. One difference from Fig. 5.2 is that to approximate M^* in the definition of R_{norm} instead of the normal approximation (3.59) we have used the value of the sphere packing bound, Theorem 40. This was necessary since on the BSC many moderate-length algebraic codes approach $R_{norm} \approx 1$ and the precision of the normal approximation becomes insufficient.

Another subtlety is that performance of algebraic codes was evaluated via Poltyrev bound, Theorem 36, where the weight distribution was taken from [91–95]. Regarding how

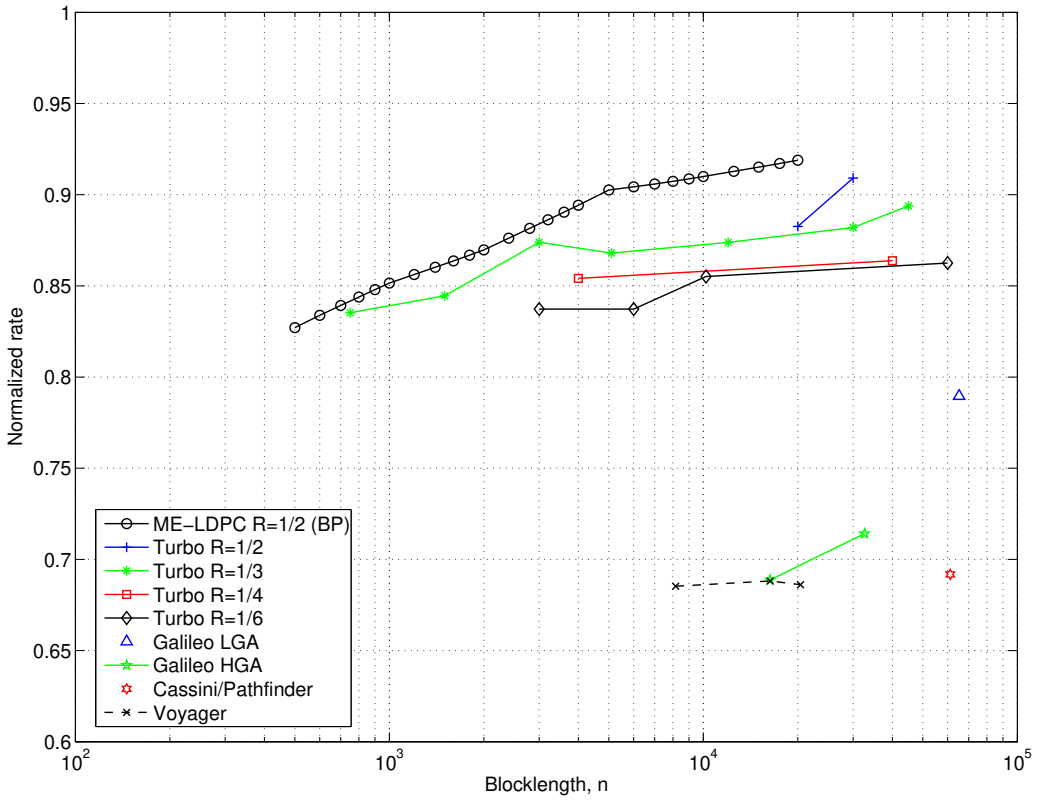


Figure 5.2: Normalized rates for various practical codes over AWGN, probability of block error $\epsilon = 10^{-4}$.

tight Poltyrev bound might be (compared to the true probability of error under maximum likelihood decoding) see [18, 55, 96]. All in all, however, taking the upper bound on ϵ (via Poltyrev) and an upper bound on $M^*(n, \epsilon)$ (via sphere packing) in (5.8) we guarantee that so obtained approximation to R_{norm} is a provable lower bound for the true value. This allows to appreciate performance of some of the algebraic codes in Fig. 5.3 even better.

Unsurprisingly, perfect binary codes, i.e. Hamming 1-error correcting and Golay, have $R_{norm} = 1$. Interestingly, however, that other points in Fig. 5.3 with large value of R_{norm} also correspond to high-rate codes (e.g., $BCH(255, 239)$, $BCH(255, 215)$, $BCH(127, 113)$, their extended versions, 4-th order Reed-Muller $(64, 57)$, etc.). At the same time, the points at the bottom part of the graph correspond to the low rate codes such as extended $BCH(64, 7)$ or first order Reed-Muller $(64, 7)$. However tempting, we cannot yet conclude that achieving fundamental limits at low-rates (equivalently, high noise levels) is harder, since tightness of the sphere packing bound in that regime is unclear.

To conclude, for the BSC known algebraic codes approach fundamental limits very closely both for low and high noise levels. This is in contrast with the situation for the AWGN.

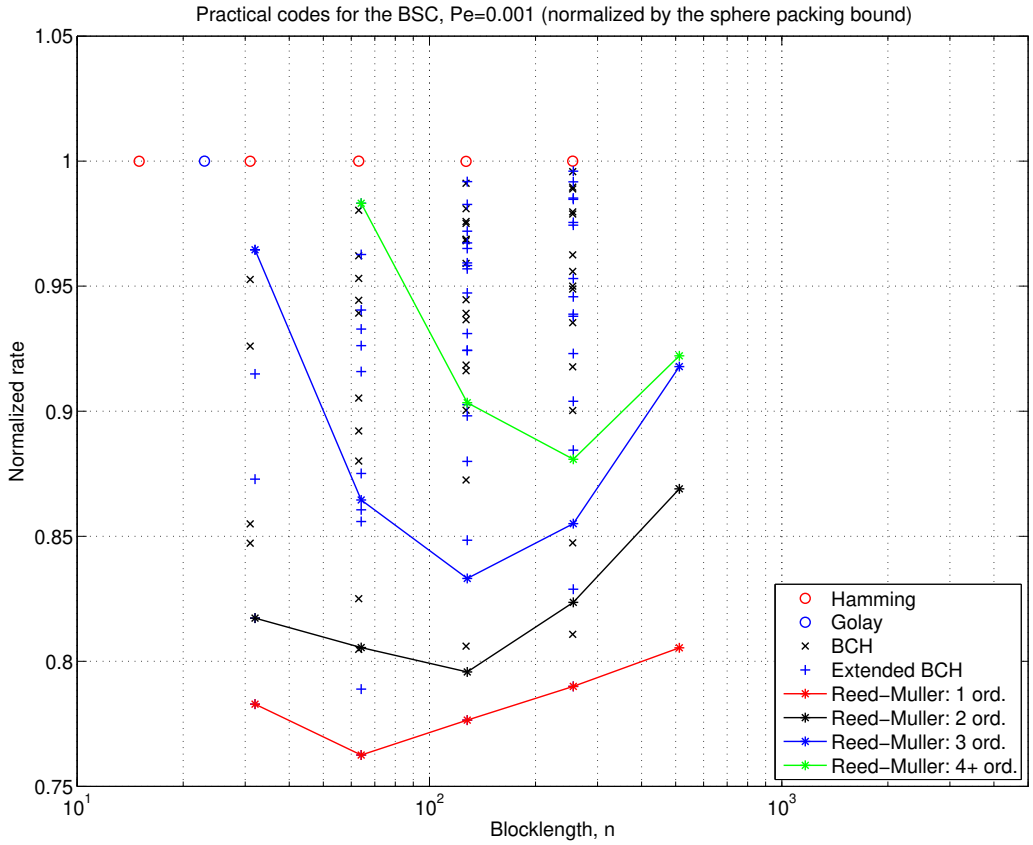


Figure 5.3: Normalized rates for various practical codes over BSC, probability of block error $\epsilon = 10^{-3}$.

5.3 Dispersion of parallel channels

In Section 4.5 we have observed that the capacity C_L and the dispersion V_L of the L-parallel AWGN channel satisfies:

$$C_L = \sum_{j=1}^{L} C\left(\frac{W_j}{\sigma_j^2}\right), \tag{5.9}$$

$$V_L = \sum_{j=1}^{L} V\left(\frac{W_j}{\sigma_j^2}\right), \tag{5.10}$$

where $C(\cdot)$ and $V(\cdot)$ are capacity and dispersion for the scalar AWGN (as functions of the SNR), and $(W_1, \ldots, W_L)$ is a water-filling power allocation; see Theorem 78.

Notably, we see that the dispersion of the parallel AWGN is a sum of dispersions of constituent AWGNs. In fact, it is easy to see that a similar conclusion holds also for the parallel DMC (with no input constraints):

Theorem 88 *Consider two DMCs $(\mathcal{A}_1, \mathcal{B}_1, W_1)$ and $(\mathcal{A}_2, \mathcal{B}_2, W_2)$. Then capacity C and ϵ-dispersion V_ϵ of the parallel DMC $(\mathcal{A}_1 \times \mathcal{A}_2, \mathcal{B}_1 \times \mathcal{B}_2, W_1 \times W_2)$, see Definition 3, is given by*

$$C = C_1 + C_2 \,, \tag{5.11}$$
$$V_\epsilon = V_{1,\epsilon} + V_{2,\epsilon} \,, \tag{5.12}$$

where $(C_1, V_{1,\epsilon})$ and $(C_2, V_{2,\epsilon})$ are the capacity and the ϵ-dispersion of the DMC W_1 and DMC W_2, respectively.

Proof: Since (5.11) is self-evident, we concentrate on (5.12). We denote the input of the parallel DMC by (X_1, X_2) and its output by (Y_1, Y_2), where $X_j \in \mathcal{A}_j$, $Y_j \in \mathcal{B}_j$ and

$$P_{Y_1Y_2|X_1X_2}(b_1, b_2|a_1a_2) \stackrel{\triangle}{=} W_1(b_1|a_1)W_2(b_2|a_2) \,. \tag{5.13}$$

According to Theorem 45 and following its notation, it is sufficient to prove that for a capacity achieving $P_{X_1X_2}$ we have

$$V_{1,min} + V_{2,min} \le V(P_{X_1X_2}, W_1 \times W_2) \le V_{1,max} + V_{2,max} \,. \tag{5.14}$$

Indeed, the lower bounds in (5.14) is achievable by taking $P_{X_1X_2} = P_{X_1}P_{X_2}$ where P_{X_j} is a distribution achieving $V_{j,min}$ and capacity of W_j, $j = 1, 2$. Similarly, the upper bound in (5.14) is also achievable.

To prove (5.14) observe that since capacity achieving *output* distribution is unique, it must be a product distribution $P_{Y_1Y_2} = P_{Y_1}P_{Y_2}$, where P_{Y_j} is the unique capacity achieving output distribution of W_j, $j = 1, 2$. Therefore, we have

$$\begin{aligned}
C_1 + C_2 &= I(P_{X_1X_2}, W_1 \times W_2) && (5.15)\\
&= \mathbb{E}\left[\log \frac{W_1(Y_1|X_1)}{P_{Y_1}(Y_1)}\right] + \mathbb{E}\left[\log \frac{W_2(Y_2|X_2)}{P_{Y_2}(Y_2)}\right] && (5.16)\\
&\le C_1 + \mathbb{E}\left[\log \frac{W_2(Y_2|X_2)}{P_{Y_2}(Y_2)}\right] && (5.17)\\
&= C_1 + \sum_{a_1 \in \mathcal{A}_1} P_{X_1}(a_1) D(W_2||P_{Y_2}|P_{X_2|X_1=a_1}) && (5.18)\\
&\le C_1 + \sum_{a_1 \in \mathcal{A}_1} P_{X_1}(a_1) C_2 \,, && (5.19)
\end{aligned}$$

where (5.17) follows since C_1 is the capacity of W_1, (5.18) is a consequence of expanding the conditional expectation $\mathbb{E}[\cdot|X_1]$ for the second term, (5.19) follows since for any distribution P on $\mathcal{A}_2$ we have $D(W_2||P_{Y_2}|P_0) \le C_2$.

Since inequality in (5.19) is in fact an equality, we must have P_{X_1} to be capacity achieving and

$$\begin{aligned}
D(W_1(\cdot|a_1)||P_{Y_1}) &= C_1 \,, && (5.20)\\
D(W_2||P_{Y_2}|P_{X_2|X_1=a_1}) &= C_2 \,, && (5.21)
\end{aligned}$$

for P_{X_1}-almost all $a_1 \in \mathcal{A}_1$. By symmetry, P_{X_2} is also capacity achieving.

Finally, for the divergence variance we have

$$\begin{aligned}
V(P_{X_1X_2}, W_1 \times W_2) = & \\
&= \mathbb{E}\left[\left(\log\frac{W_1(Y_1|X_1)}{P_{Y_1}(Y_1)} + \log\frac{W_2(Y_2|X_2)}{P_{Y_2}(Y_2)} - D(W_1(\cdot|X_1)W_2(\cdot|X_2)||P_{Y_1}P_{Y_2})\right)^2\right] && (5.22)\\
&= \mathbb{E}\left[\left(\log\frac{W_1(Y_1|X_1)}{P_{Y_1}(Y_1)} + \log\frac{W_2(Y_2|X_2)}{P_{Y_2}(Y_2)} - C_1 - C_2\right)^2\right] && (5.23)\\
&= V(P_{X_1}, W_1) + V(P_{X_2}, W_2) \\
&\quad + 2\mathbb{E}\left[\left(\log\frac{W_1(Y_1|X_1)}{P_{Y_1}(Y_1)} - C_1\right)\left(\log\frac{W_2(Y_2|X_2)}{P_{Y_2}(Y_2)} - C_2\right)\right] && (5.24)\\
&= V(P_{X_1}, W_1) + V(P_{X_2}, W_2)\,, && (5.25)
\end{aligned}$$

where (5.23) is by (5.20) and (5.21), (5.24) is by (3.99), and (5.25) follows from

$$\begin{aligned}
\mathbb{E}\left[\log\frac{W_2(Y_2|X_2)}{P_{Y_2}(Y_2)} - C_2\middle|Y_1X_1\right] &= \mathbb{E}\left[\log\frac{W_2(Y_2|X_2)}{P_{Y_2}(Y_2)} - C_2\middle|X_1\right] && (5.26)\\
&= D(W_2||P_{Y_2}|P_{X_2|X_1}) - C_2 && (5.27)\\
&= 0\,, && (5.28)
\end{aligned}$$

Finally, by the definition of V_{min} and V_{max} in (3.108) and (3.107), respectively, we have

$$V_{j,min} \le V(P_{X_j}, W_j) \le V_{j,max}\,, j = 1, 2\,, \tag{5.29}$$

and therefore, (5.25) implies (5.14). ■

Together with the normal approximation (2.23), (5.10) (for the parallel AWGN channel) and (5.12) (for the parallel DMC) immediately highlight the benefit obtained from joint coding on parallel channels simultaneously. Indeed, suppose we were using independent codes on each of the L channels with capacity-dispersion pairs $(C_j, V_j), j = 1, \ldots, L$. Then if we took each of the L best (n, M_j, ϵ)-codes we obtain an $(n, \prod M_j, 1-(1-\epsilon)^L)$-code for the parallel channel. Assuming that ϵ is small we have approximately

$$1 - (1-\epsilon)^L \approx L\epsilon\,. \tag{5.30}$$

Thus, the performance of the best possible independent coding scheme is

$$\begin{aligned}
\log M^*_{ind}(n, L\epsilon) &= \sum_{j=1}^{L} \log M^*_j(n, \epsilon) && (5.31)\\
&\approx n\left(\sum_{j=1}^{L} C_j\right) - Q^{-1}(\epsilon)\sqrt{n}\left(\sum_{j=1}^{L}\sqrt{V_j}\right)\,. && (5.32)
\end{aligned}$$

Now if ϵ and L are both sufficiently small, then $Q^{-1}(\epsilon) \approx Q^{-1}(L\epsilon)$ and we can see that

gaps to capacity are compared as

$$\text{gap for independent coding} \quad \sim \quad \frac{1}{\sqrt{n}} \sum_{j=1}^{L} \sqrt{V_j} \tag{5.33}$$

$$\text{gap for joint coding} \quad \sim \quad \frac{1}{\sqrt{n}} \sqrt{\sum_{j=1}^{L} V_j} \tag{5.34}$$

Therefore the loss in the maximum achievable rate due to using independent coding at blocklength n and small ϵ is proportional to

$$\frac{1}{\sqrt{n}} \left(\sum_{j=1}^{L} \sqrt{V_j} - \sqrt{\sum_{j=1}^{L} V_j} \right) > 0 \,. \tag{5.35}$$

5.4 Dispersion and alphabet size

Because of the importance of channel dispersion, we note the following upper-bound (see also [9, Exercise 5.23]):

Theorem 89 *For the DMC with* $\min\{|\mathcal{A}|, |\mathcal{B}|\} > 2$ *we have*

$$V \le 2 \log^2 \min\{|\mathcal{A}|, |\mathcal{B}|\} - C^2 \,. \tag{5.36}$$

For the DMC with $\min\{|\mathcal{A}|, |\mathcal{B}|\} = 2$ *we have*

$$V \le 1.2 \log^2 e - C^2 \,. \tag{5.37}$$

The estimate (5.36) is order-optimal for $\min\{|\mathcal{A}|, |\mathcal{B}|\} \to \infty$. Indeed, consider a channel with additive noise $\mathcal{A} = \mathcal{B} = \mathbb{Z}/n\mathbb{Z}$:

$$Y = X + Z \mod n \,, \tag{5.38}$$

where $\mathbb{P}[Z = 0] = \frac{1}{2}$ and $P[Z = 1] = \cdots = P[Z = n-1] = \frac{1}{2(n-1)}$. The capacity and dispersion of such a channel are

$$C_n = \frac{1}{2} \log n + O(1) \,, \tag{5.39}$$

$$V_n = \frac{1}{4} \log^2 n + O(\log n) \,. \tag{5.40}$$

Thus, the estimate of Theorem 89

$$V = O(\log^2 \min\{|\mathcal{A}|, |\mathcal{B}|\}) \,, \quad \min\{|\mathcal{A}|, |\mathcal{B}|\} \to \infty \tag{5.41}$$

cannot be generally improved.

Comparing (5.36) with Theorem 88, we notice that product channels posses untypically small dispersion.

Since the typical blocklength needed to achieve capacity is governed by V/C^2, it is natural to ask whether for very small capacities the upper-bound in (5.36) can be improved to prevent the blowing up of $\frac{V}{C^2}$. Such a bound is not possible over all W with fixed alphabet sizes, since such a collection of DMCs always includes all of the BSCs for which we know that $\frac{V}{C^2} \to \infty$ as $C \to 0$.

Theorem 89 is a simple consequence of the following (see Section 3.4 for the notation):

Lemma 90 *For functions V and U defined on $\mathcal{P}$, the following inequality holds:*

$$V(P,W) \le U(P,W) \le 2g(\min\{|\mathcal{A}|,|\mathcal{B}|\}) - [I(P,W)]^2\,, \tag{5.42}$$

where

$$g(n) = \begin{cases} 0.6\log^2 e\,, & n = 2\,, \\ \log^2 n\,, & n \ge 3\,. \end{cases} \tag{5.43}$$

Proof: consider the following chain of inequalities:

$$\begin{aligned} U(P,W) + [I(P,W)]^2 &\stackrel{\triangle}{=} \sum_{x\in\mathcal{A}}\sum_{y\in\mathcal{B}} P(x)W(y|x)\Big[\log^2 W(y|x) + \log^2 PW(y) && (5.44)\\ &\quad - 2\log W(y|x)\cdot\log PW(y)\Big] && (5.45)\\ &\le \sum_{x\in\mathcal{A}}\sum_{y\in\mathcal{B}} P(x)W(y|x)\left[\log^2 W(y|x) + \log^2 PW(y)\right] && (5.46)\\ &= \sum_{x\in\mathcal{A}} P(x)\left[\sum_{y\in\mathcal{B}} W(y|x)\log^2 W(y|x)\right] && (5.47)\\ &\quad + \left[\sum_{y\in\mathcal{B}} PW(y)\log^2 PW(y)\right] && (5.48)\\ &\le \sum_{x\in\mathcal{A}} P(x)g(|\mathcal{B}|) + g(|\mathcal{B}|) && (5.49)\\ &= 2g(|\mathcal{B}|)\,, && (5.50) \end{aligned}$$

where (5.46) is because $2\log W(y|x)\cdot\log PW(y)$ is always non-negative, and (5.49) follows because each term in square-brackets can be upper-bounded using the following optimization problem:

$$g(n) \stackrel{\triangle}{=} \sup_{a_j\ge 0:\sum_{j=1}^n a_j = 1} \sum_{j=1}^n a_j \log^2 a_j\,. \tag{5.51}$$

Since the $x\log^2 x$ has unbounded derivative at the origin, the solution of (5.51) is always in the interior of $[0,1]^n$. Then it is straightforward to show that for $n > e$ the solution is actually $a_j = \frac{1}{n}$. For $n = 2$ it can be found directly that $g(2) = 0.5629 < 0.6$. Finally, because of the symmetry, a similar argument can be made with $|\mathcal{B}|$ replaced by $|\mathcal{A}|$ and hence in (5.42) we are free to choose the best bound. ■

5.5 Communication rate and channel state dynamics

So far we have dominantly discussed the value of the normal approximation (2.23) for memoryless channels where the transition kernel $P_{Y_j|X_j}$ remains fixed (static) throughout transmission. Although valid in many cases, such static assumption frequently is not satisfied in practice where the wave propagation conditions change dynamically. One example of such a model involving dynamics is a Gilbert-Elliott channel, analyzed in Section 3.5[2]

In Section 3.5 we have demonstrated (theoretically and by numerical computations) that the normal approximation

$$\frac{1}{n}\log M^*(n,\epsilon) \approx C - \sqrt{\frac{V}{n}}Q^{-1}(\epsilon) + \frac{1}{2n}\log n \tag{5.52}$$

is very tight. The capacity and dispersion pair (C, V) is equal to (C_1, V_1), see Theorem 58, for the case when state sequence S^n is known at the receiver; and to (C_0, V_0), see Theorem 59, for the case of no state knowledge.

Let us discuss two practical applications of (5.52). First, for the state-known case, the capacity C_1 is independent of the state transition probability τ. However, according to Theorem 58, the channel dispersion V_1 does indeed depend on τ. Therefore, according to (5.53), the minimal blocklength needed to achieve a fraction of capacity behaves as

$$n \gtrsim \left(\frac{Q^{-1}(\epsilon)}{1-\eta}\right)^2 \frac{V}{C^2}\,, \tag{5.53}$$

or as $O\left(\frac{1}{\tau}\right)$ when $\tau \to 0$, since according to (3.378):

$$V_1 = O\left(\frac{1}{\tau}\right)\,. \tag{5.54}$$

This has an intuitive explanation: to achieve the full capacity of a Gilbert-Elliott channel we need to wait until the influence of the random initial state "washes away". Since transitions occur on average every $\frac{1}{\tau}$ channel uses, the blocklength should be $O\left(\frac{1}{\tau}\right)$ as $\tau \to 0$. Comparing (3.30) and (3.378) we can ascribe a meaning to each of the two terms in (3.378): the first one gives the dispersion due to the usual BSC noise, whereas the second one is due to memory in the channel. In particular, knowledge of channel dispersion and (5.53) allows us to interpret the quantity $\frac{1}{\tau}$ as a natural "time constant" of the channel. According to Theorem 60, similar conclusion holds for the case of no state knowledge, since (5.54) holds for V_0 as well.

Next, consider the case in which the state is not known at the decoder. As shown in [64], when the state transition probability τ decreases to 0 the capacity $C_0(\tau)$ increases to C_1. This is sometimes interpreted as implying that if the state is unknown at the receiver slower dynamics are advantageous. Our refined analysis, however, shows that this is true only up to a point.

Indeed, fix a rate $R < C_0(\tau)$ and an $\epsilon > 0$. In view of the tightness of (5.52), the minimal blocklength, as a function of state transition probability τ needed to achieve rate

[2]For the remainder of this section we use the notation introduced in Section 3.5.

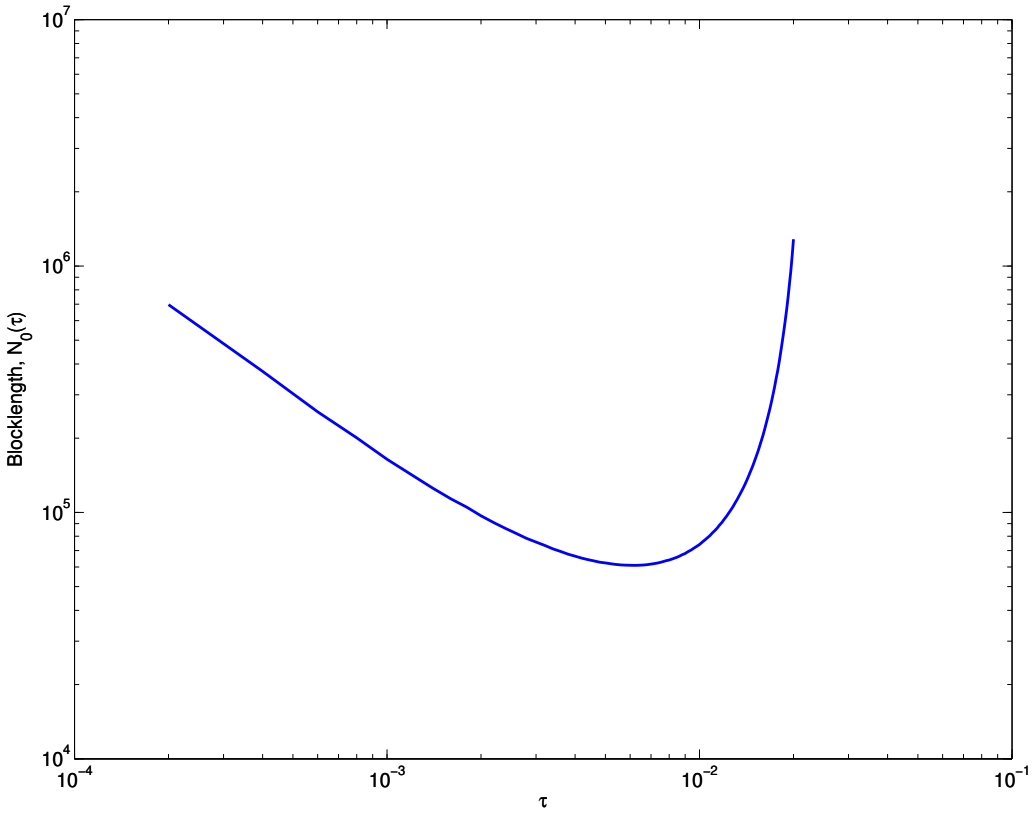


Figure 5.4: Minimal blocklength needed to achieve $R = 0.4$ bit and $\epsilon = 0.01$ as a function of state transition probability τ. The channel is the Gilbert-Elliott with no state information at the receiver, $\delta_1 = 1/2, \delta_2 = 0$.

R is approximately given by

$$N_0(\tau) \approx V_0(\tau) \left(\frac{Q^{-1}(\epsilon)}{C_0(\tau) - R} \right)^2 . \tag{5.55}$$

When the state transition probability τ decreases we can predict the current state better; on the other hand, we also have to wait longer until the chain "forgets" the initial state. The trade-off between these two effects is demonstrated in Fig. 5.4, where we plot $N_0(\tau)$ for the setup of Fig. 3.10(b).

The same effect can be demonstrated by analyzing the maximal achievable rate as a function of τ. In view of the tightness of the approximation in (5.52) for large n we may replace $\frac{1}{n} \log M^*(n, \epsilon)$ with (5.52). The result of such analysis for the setup in Fig. 3.10(b) and $n = 3 \cdot 10^4$ is shown as a solid line in Fig. 5.5, while a dashed line corresponds to the capacity $C_0(\tau)$. Note that at $n = 30000$ (5.52) is indistinguishable from the upper and lower bounds. We can see that once the blocklength n is fixed, the fact that capacity $C_0(\tau)$ grows when τ decreases does not imply that we can actually transmit at a higher rate. In fact we can see that once τ falls below some critical value, the maximal rate

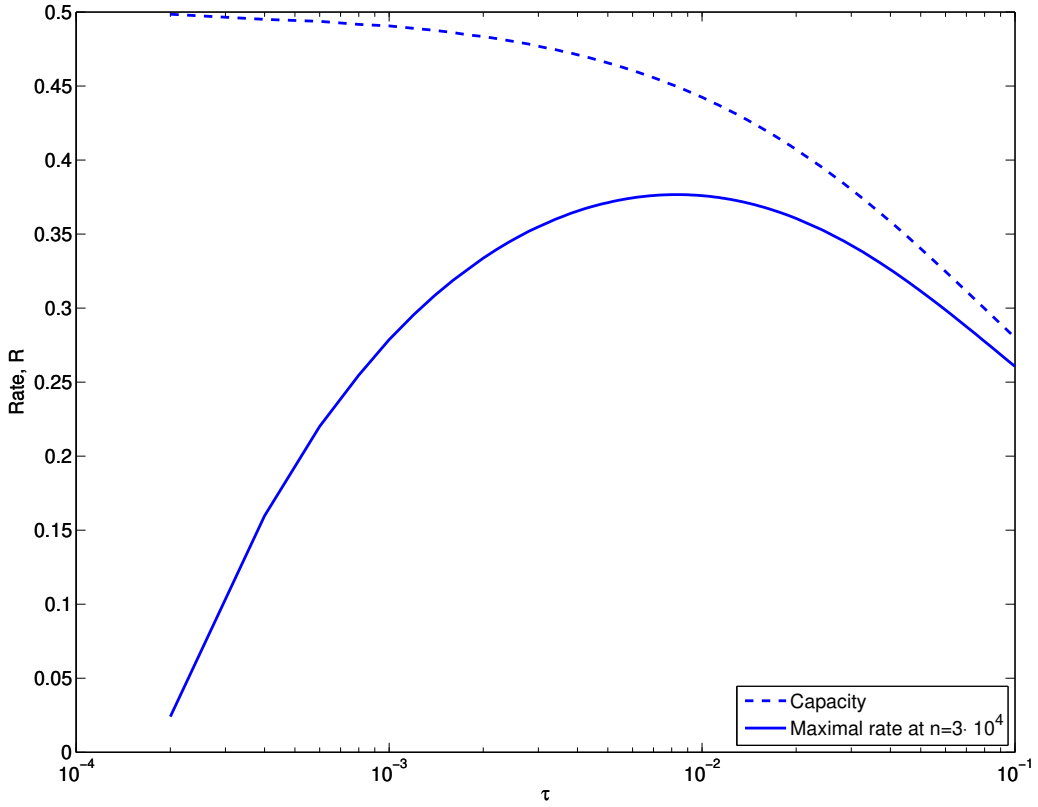


Figure 5.5: Comparison of the capacity and the maximal achievable rate $\frac{1}{n}\log M^*(n,\epsilon)$ at blocklength $n = 3 \cdot 10^4$ as a function of the state transition probability τ for the Gilbert-Elliott channel with no state information at the receiver, $\delta_1 = 1/2, \delta_2 = 0$; probability of block error is $\epsilon = 0.01$.

drops steeply with decreasing τ. This situation exemplifies the drawbacks of neglecting the second term in (2.22). Note that, according to Theorem 60 the value of $N_0(\tau)$ for small τ is approximated by replacing $V_0(\tau)$ with $V_1(\tau)$ in (5.55). Since the latter admits a very simple expression (3.378), this method helps to quickly isolate the extremum of $N_0(\tau)$, cf. Fig. 5.4.

In summary, for the Gilbert-Elliott channel the capacity term in (5.52) fails to adequately describe the effect of channel dynamics on the fundamental limits. At the same time the refinement provided by the channel dispersion resolves this difficulty.

5.6 Moderate deviations

In [82] authors raised the question of the best possible behavior of the probability of error when the coding rate approaches capacity slower than $1/\sqrt{n}$. In [82] the question is answered for a certain subset of the DMCs (which excludes, for example the BEC). We show how a refinement of their result can be simply derived using methods developed in Chapter 2

and Section 3.4. Our contribution is in deriving necessary and sufficient conditions for the moderate deviation property to hold, thereby extending the subset of DMCs to the maximal possible one. Additionally, we prove a similar result for the AWGN.

The moderate deviation property (MDP) is formulated as follows. Consider a sequence of channels indexed by the blocklength n and define

$$\begin{aligned} \epsilon^*(n,M) &= \inf\{\epsilon : \exists (n,M,\epsilon)\text{-code (maximal probability of error)}\} \quad (5.56)\\ \epsilon^*_{avg}(n,M) &= \inf\{\epsilon : \exists (n,M,\epsilon)\text{-code (average probability of error)}\}\,. \quad (5.57)\end{aligned}$$

Definition 13 *A sequence of channels with capacity C is said to satisfy MDP with constant V if for any sequence of integers M_n such that*

$$\log M_n = nC - n\rho_n\,, \quad (5.58)$$

where $\rho_n > 0$, $\rho_n \to 0$ and $n\rho_n^2 \to \infty$, we have

$$\lim_{n\to\infty} \frac{1}{n\rho_n^2}\log \epsilon^*(n,M_n) = \lim_{n\to\infty} \frac{1}{n\rho_n^2}\log \epsilon^*_{avg}(n,M_n) = -\frac{1}{2V}\,. \quad (5.59)$$

5.6.1 Discrete memoryless channels

Below in this section we use the notation of Section 3.4; in particular, recall the definitions of $\mathcal{A}, \mathcal{B}, W, I(P,W), V(P,W)$ and V_{min}.

Apart from analyzing the limit of ϵ^*_{avg} the result of [82] can be states as follows:

Theorem 91 (Altug-Wagner) *Consider a DMC W. If $W(y|x) > 0$ for all $x \in \mathcal{A}, y \in \mathcal{B}$ and $V_{min} > 0$ then DMC W satisfies MDP with the constant V_{min}.*

The main result of this section is:

Theorem 92 *The DMC W satisfies MDP if and only if $V_{min} > 0$, in which case V_{min} is the MDP constant of the DMC.*

Note that V_{min} is precisely the channel dispersion of the DMC, see Theorem 45.

Theorem 93 *Consider a DMC W and a sequence ρ_n such that $\rho_n > 0$, $\rho_n \to 0$ and $\rho_n^2 n \to \infty$. If $V_{min} > 0$ then there exists a sequence of $(n, \exp\{nC - n\rho_n\}, \epsilon_n)$ codes (maximal probability of error) with*

$$\limsup \frac{1}{n\rho_n^2}\log \epsilon_n \le -\frac{1}{2V_{min}}\,. \quad (5.60)$$

On the other hand, when $V_{min} = 0$ there exists a sequence of $(n, \exp\{nC - n\rho_n\}, \epsilon_n)$ codes (maximal probability of error) with

$$\epsilon_n \le 2\exp\{-n\rho_n\}\,, \quad (5.61)$$

so that the channel cannot satisfy MDP.

Proof: Denote by P the capacity achieving distribution that also achieves V_{min}. According to the DT bound, Theorem 18, there exist an $(n, 2\exp\{nC-n\rho_n\}, \epsilon'_n)$ code (average probability of error) such that

$$\epsilon'_n \le \mathbb{E}\left[\exp\left\{-|i(X^n,Y^n)-nC+n\rho_n|^+\right\}\right], \tag{5.62}$$

where

$$i(x^n,y^n) \stackrel{\triangle}{=} \sum_{j=1}^{n} \log\frac{W(y_j|x_j)}{PW(y_j)}. \tag{5.63}$$

And therefore, by a standard "purging" method, there also exists an $(n, \exp\{nC-n\rho_n\}, \epsilon_n)$ code (maximal probability of error) with $\epsilon_n = 2\epsilon'_n$, or

$$\epsilon_n \le 2\mathbb{E}\left[\exp\left\{-|i(X^n,Y^n)-nC+n\rho_n|^+\right\}\right]. \tag{5.64}$$

If $V_{min}=0$ then $i(X^n,Y^n)=nC$ and (5.61) follows trivially.

Assume $V_{min}>0$, fix arbitrary $\lambda<1$ and observe a chain of obvious inequalities:

$$\begin{aligned}
&\exp\left\{-|i(X^n,Y^n)-nC+n\rho_n|^+\right\} && (5.65)\\
&\le\ 1\{i(X^n,Y^n)\le nC-\lambda n\rho_n\} && (5.66)\\
&\quad +\exp\left\{-|i(X^n,Y^n)-nC+n\rho_n|^+\right\}1\{i(X^n,Y^n)>nC-\lambda n\rho_n\} && (5.67)\\
&\le\ 1\{i(X^n,Y^n)\le nC-\lambda n\rho_n\}+\exp\{-(1-\lambda)n\rho_n\}. && (5.68)
\end{aligned}$$

By [81, Theorem 3.7.1] we have

$$\limsup\frac{1}{n\rho_n^2}\log\mathbb{P}[i(X^n,Y^n)\le nC-\lambda n\rho_n] \le -\frac{\lambda^2}{2V_{min}}. \tag{5.69}$$

Therefore, by taking the expectation in (5.68) and by conditions on ρ_n the second term is asymptotically dominated by the first and we obtain:

$$\limsup\frac{1}{n\rho_n^2}\log\mathbb{E}\left[\exp\left\{-|i(X^n,Y^n)-nC+n\rho_n|^+\right\}\right] \le -\frac{\lambda^2}{2V_{min}}. \tag{5.70}$$

Since $\lambda<1$ was arbitrary we can take $\lambda\to 1$ to obtain (5.60). ■

The main analytic tool required in proving the converse bound in this section is a tight non-asymptotic lower bound for the probability of a large deviation of a random variable from its mean. This question has been addressed by many authors working in probability and statistics, starting from Kolmogorov [97]. Currently, one of the most general such results belongs to Rozovsky [98, 99]. The following is a weakening of [98, Theorem 1] which plays the same role as Berry-Esseen inequality in the previous analysis[3] :

Theorem 94 (Rozovsky) *There exist universal constants $A_1>0$ and $A_2>0$ with the following property. Let X_k, $k=1,\ldots,n$ be independent with finite third moments:*

$$\mu_k=\mathbb{E}[X_k],\ \sigma_k^2=\mathrm{Var}[X_k],\ \text{and } t_k=\mathbb{E}[|X_k-\mu_k|^3]. \tag{5.71}$$

[3]Similar to well-known extensions of the Berry-Esseen inequality to the case of random variables without a third absolute moment, Rozovsky does not require that $\mathbb{E}|X_k|^3$ be bounded. However, we only will need this weaker result.

Denote $V = \sum_{k=1}^{n} \sigma_k^2$ *and* $T = \sum_{k=1}^{n} t_k$. *Whenever* $x \geq 1$ *we have*

$$\mathbb{P}\left[\sum_{k=1}^{n}(X_k - \mu_k) > x\sqrt{V}\right] \geq Q(x)e^{-\frac{A_1 T}{V^{3/2}}x^3}\left(1 - \frac{A_2 T}{V^{3/2}}x\right). \tag{5.72}$$

Theorem 95 *Consider a DMC* W *and a sequence of* (n, M_n, ϵ_n) *codes (average probability of error) with*

$$\log M_n \geq nC - n\rho_n\,, \tag{5.73}$$

where $\rho_n > 0$, $\rho_n \to 0$ *and* $\rho_n^2 n \to \infty$. *If* $V_{min} > 0$ *then we have*

$$\liminf_{n\to 0} \frac{1}{n\rho_n^2} \log \epsilon_n \geq -\frac{1}{2V_{min}}\,. \tag{5.74}$$

Proof: Replacing the encoder with an optimal deterministic one, we only reduce the average probability of error. Next, if we have an (n, M_n, ϵ_n) code (average probability of error) with a deterministic encoder, then a standard argument shows that there exists an $(n, \frac{1}{2}M_n, 2\epsilon_n)$ subcode (maximal probability of error). Replacing $M_n \to \frac{1}{2}M_n$ and $\epsilon_n \to 2\epsilon_n$, without loss of generality we may assume the code to have a deterministic encoder and a maximal probability of error ϵ_n.

Now for each n denote by $P_n \in \mathcal{P}_n$ the n-type containing the largest number of codewords. A standard type-counting argument shows that then there exists an (n, M_n', ϵ_n) constant composition P_n subcode with

$$\log M_n' \geq nC - n\rho_n - |\mathcal{A}|\log(n+1)\,. \tag{5.75}$$

By compactness of $\mathcal{P}$, the simplex of distributions on $\mathcal{A}$, the sequence P_n has an accumulation point P^*. Without loss of generality, we may assume $P_n \to P^*$.

Now for each n define the following probability distribution Q_Y^n on $\mathcal{B}^n$:

$$Q_Y^n(y^n) = \prod_{j=1}^{n} P_n W(y_j)\,. \tag{5.76}$$

According to the Theorem 34 we have

$$\beta_{1-\epsilon_n}(P_{X^nY^n}||P_{X^n}Q_Y^n) \leq \frac{1}{M_n'}\,, \tag{5.77}$$

where here and below P_{X^n} is the distribution induced by the encoder on $\mathcal{A}^n$.

Applying (2.67) we get that for any γ we have:

$$\epsilon_n \geq \mathbb{P}\left[\log\frac{W(Y^n|X^n)}{Q_Y^n(Y^n)} < \gamma\right] - \exp\{\gamma - \log M_n'\}\,. \tag{5.78}$$

We now fix arbitrary $\lambda > 1$ and take $\gamma = nC - \lambda n\rho_n$ to obtain:

$$\epsilon_n \geq \mathbb{P}\left[\log\frac{W(Y^n|X^n)}{Q_Y^n(Y^n)} < nC - \lambda n\rho_n\right] - \exp\{-n\rho_n(\lambda - 1) + |\mathcal{A}|\log(n+1)\}\,. \tag{5.79}$$

Notice that since the code has constant composition P_n, the distribution of $\log \frac{W(Y^n|X^n)}{Q_Y^n(Y^n)}$ given $X^n = x^n$ is the same for all x^n. Therefore, assuming such conditioning we have

$$\log \frac{W(Y^n|X^n)}{Q_Y^n(Y^n)} \sim \sum_{j=1}^{n} Z_j \,, \tag{5.80}$$

where Z_j are independent and

$$\begin{aligned} \sum_{j=1}^{n} \mathbb{E}\left[Z_j\right] &= nI(P_n, W)\,, & (5.81)\\ \sum_{j=1}^{n} \mathrm{Var}[Z_j] &= nV(P_n, W)\,, & (5.82)\\ \sum_{j=1}^{n} \mathbb{E}\left[|Z_j - \mathbb{E}\left[Z_j\right]|^3\right] &= nT(P_n, W)\,, & (5.83) \end{aligned}$$

where we used the notation introduced in Section 3.4. In terms of Z_j the bound in (5.79) asserts

$$\epsilon_n \geq \mathbb{P}\left[\sum_{j=1}^{n} Z_j < nC - \lambda n\rho_n\right] - \exp\{-n\rho_n(\lambda - 1) + |\mathcal{A}|\log(n+1)\}\,. \tag{5.84}$$

First, suppose that $I(P^*, W) < C$. Then a simple Chernoff-bound implies that the right-hand side of (5.79) converges to 1 and (5.74) holds.

Next, assume $I(P^*, W) = C$. Since $I(P_n, W) < C$ we have from (5.84):

$$\epsilon_n \geq \mathbb{P}\left[\sum_{j=1}^{n} Z_j - nI(P_n, W) < -\lambda n\rho_n\right] - \exp\{-n\rho_n(\lambda - 1) + |\mathcal{A}|\log(n+1)\}\,. \tag{5.85}$$

Note that by continuity of $V(P, W)$ we have

$$V(P_n, W) \to V(P^*, W) \geq V_{min} > 0\,, \tag{5.86}$$

where $V(P^*, W) \geq V_{min}$ since P^* is capacity-achieving. Therefore, by Theorem 94 we obtain:

$$\begin{aligned} &\mathbb{P}\left[\sum_{j=1}^{n} Z_j - nI(P_n, W) < -\lambda n\rho_n\right] & (5.87)\\ &\geq \quad Q\left(\frac{\lambda}{\sqrt{V(P_n, W)}}\sqrt{n\rho_n^2}\right) e^{-\frac{\lambda^3 A_1 T(P_n, W)}{V^3(P_n, W)} n\rho_n^3} \left(1 - \frac{\lambda A_2 T(P_n, W)}{V^2(P_n, W)}\rho_n\right)\,, & (5.88) \end{aligned}$$

since $T(P_n, W)$ is a continuous and bounded function, we see that the term in parentheses is $1 + o(1)$ by conditions on ρ_n. Therefore,

$$\liminf_{n\to\infty} \frac{1}{n\rho_n^2} \log \mathbb{P}\left[\sum_{j=1}^{n} Z_j - nI(P_n, W) < -\lambda n\rho_n\right] \tag{5.89}$$

$$\geq \lim_{n\to\infty} \frac{1}{n\rho_n^2} \log Q\left(\frac{\lambda}{\sqrt{V(P_n, W)}}\sqrt{n\rho_n^2}\right) + \lim_{n\to\infty} \frac{1}{n\rho_n^2}\left(-\frac{\lambda^3 A_1 T(P_n, W)}{V^3(P_n, W)} n\rho_n^3\right) \tag{5.90}$$

$$= -\frac{\lambda^2}{2V(P_*, W)} \tag{5.91}$$

$$\geq -\frac{\lambda^2}{2V_{min}}\,. \tag{5.92}$$

Finally, it is easy to see that the second term in (5.85) is asymptotically dominated by the first term according to (5.92) and $n\rho_n \gg n\rho_n^2$. Thus, from (5.92) we conclude that

$$\liminf_{n\to\infty} \frac{1}{n\rho_n^2} \log \epsilon_n \geq -\frac{1}{2V_{min}}\,. \tag{5.93}$$

■

Proof of Theorem 92: Apply Theorems 93 and 95. ■

5.6.2 AWGN

Theorem 96 *AWGN channel with SNR P satisfies the MDP with constant $V(P)$, which is the channel dispersion of the AWGN given by* (4.155)*.*

Proof: Converse: Consider a sequence of (n, M_n, ϵ_n) codes (average probability of error) with

$$M_n = \exp\{nC - n\rho_n\}\,, \tag{5.94}$$

where $\rho_n > 0$, $\rho_n \to 0$ and $\rho_n^2 n \to \infty$. Following the method of [4] we can assume without loss of generality that every codeword $\mathbf{C}_j \in \mathbb{R}^n, j = 1, \ldots, M_n$ lies on a power-sphere:

$$||\mathbf{C}_j||^2 = nP\,. \tag{5.95}$$

We apply the meta-converse bound, Theorem 29 with Q_{Y^n} chosen as in Section 4.2.1:

$$Q_{Y^n} = \prod_{j=1}^{n} \mathcal{N}(0, 1 + P)\,, \tag{5.96}$$

to obtain

$$\beta_{1-\epsilon_n}(P_{X^nY^n}, P_{X^n}Q_{Y^n}) \leq \exp\{-nC + n\rho_n\}\,, \tag{5.97}$$

where P_{X^n} is the distribution induced by the encoder on $\mathbb{R}^n$. As explained in Section 4.2.2, we have the equality

$$\beta_{1-\epsilon_n}(P_{X^nY^n}, P_{X^n}Q_{Y^n}) = \beta_{1-\epsilon_n}(P_{Y^n|X^n=x}, Q_{Y^n})\,, \tag{5.98}$$

where $x = [\sqrt{P}, \ldots, \sqrt{P}]^T$. Now applying (2.67) $\beta_{1-\epsilon_n}(P_{Y^n|X^n=x}, Q_{Y^n})$ with $\gamma = nC - \lambda n\rho_n$, where $\lambda > 1$ is arbitrary we obtain

$$\epsilon_n \geq \mathbb{P}\left[\frac{1}{2(1+P)} \log e \sum_1^n \left(P(1-Z_i^2) + 2\sqrt{P}Z_i\right) < -\lambda n\rho_n\right] - \exp\{-n\rho_n(\lambda-1)\}\,, \quad (5.99)$$

where we have written the distribution of $\log \frac{P_{Y^n|X^n=x}}{Q_{Y^n}}$ explicitly in terms of i.i.d. $Z_j \sim \mathcal{N}(0,1)$; see (4.53). According to [81, Theorem 3.7.1], the first term dominates the second and we have

$$\liminf_{n\to\infty} \frac{1}{n\rho_n^2} \log \epsilon_n \geq -\frac{\lambda^2}{2V(P)}\,, \quad (5.100)$$

and taking $\lambda \searrow 1$ we obtain

$$\liminf_{n\to\infty} \frac{1}{n\rho_n^2} \log \epsilon_n \geq -\frac{1}{2V(P)}\,. \quad (5.101)$$

Achievability: Just like in Section 4.2 we apply the $\kappa\beta$ bound with $\mathcal{F}$ chosen to be the power sphere and Q_{Y^n} as in (5.96). Using the identity (5.98) and the lower bound on $\kappa_\tau(\mathcal{F}, Q_{Y^n})$ given by Lemma 72 we show that for all $0 < \epsilon < 1$ and $0 < \tau < \epsilon$ there exists an (n, M, ϵ) code (maximal probability of error) with

$$M \geq \frac{1}{C_1} \frac{\tau - e^{-C_2 n}}{\beta_{1-\epsilon+\tau}(P_{Y^n|X^n=x}, Q_{Y^n})}\,, \quad (5.102)$$

where $x = [\sqrt{P}, \ldots, \sqrt{P}]^n \in \mathbb{R}^n$ is a vector on the the power sphere. We now take $\tau = \frac{\epsilon}{2}$ and apply the upper-bound on β from (2.69) to obtain the statement: For any γ there exists and (n, M, ϵ) code (maximal probability of error) with

$$M \geq \frac{\epsilon - 2e^{-C_2 n}}{2C_1} \exp\{\gamma\} \quad (5.103)$$

and

$$\epsilon = 2\mathbb{P}\left[\log \frac{dP_{Y^n|X^n=x}}{Q_{Y^n}} \leq \gamma\right]\,. \quad (5.104)$$

Now take $\gamma_n = nC - \lambda n\rho_n$, where $\lambda < 1$ is arbitrary. By [81, Theorem 3.7.1] we obtain a sequence of codes with

$$\log M_n \geq nC - n\rho_n \quad (5.105)$$

for all n sufficiently large and

$$\limsup_{n\to\infty} \frac{1}{n\rho_n^2} \log \epsilon_n \leq -\frac{\lambda^2}{2V(P)}\,. \quad (5.106)$$

In particular,

$$\limsup_{n\to\infty} \frac{1}{n\rho_n^2} \log \epsilon^*(n, \exp\{nC - n\rho_n\}) \leq \frac{-\lambda^2}{2V(P)}\,, \quad (5.107)$$

and since $\lambda < 1$ is arbitrary we can take $\lambda \nearrow 1$ to finish the proof. ■

Chapter 6

Communication with feedback

Without feedback, the backoff from capacity due to non-asymptotic blocklength can be quite substantial for blocklengths and error probabilities of interest in many practical applications. In this chapter, novel achievability bounds are used to demonstrate that in the non-asymptotic regime, the maximal achievable rate improves dramatically thanks to variable-length coding with feedback. Section 6.1 reviews the previous work, including Burnashev's derivation of the closed-form expression for the error-exponent. In Section 6.2 various notions of codes with feedback are defined formally and related to each other. A digression regarding a natural generalization of the concept of channel for situations with feedback, *a synchronized channel*, is undertaken in Section 6.3. Before considering more complex feedback systems, automatic repeat request (ARQ) system is considered in Section 6.4. In the paradigm of fixed-blocklength coding feedback is useless even non-asymptotically – a result shown in Section 6.5. Variable-length codes without feedback already exhibit a novel symptom: the ϵ-capacity becomes a function of ϵ (Section 6.6). Next, in the main part of this chapter, Section 6.7, the the variable-length codes with feedback are constructed and shown to extremely improve upon the best possible fixed-blocklength ones. Furthermore, virtually all the advantages of noiseless feedback are shown to be achievable with decision-feedback only. For example, for the binary symmetric channel with capacity 1/2 the blocklength required to achieve 90% of the capacity (with decision feedback) is smaller than 200, compared to at least 3100 for the best fixed-blocklength code (even with noiseless feedback).

For a practically motivated variation of the problem, coding with a *termination symbol*, the zero-error communication scheme is constructed and evaluated numerically and asymptotically in Section 6.8. Regarding the delay constraint, it is shown in Section 6.9 that restricting the excess of the delay results in feedback being almost useless; that is non-feedback, fixed-blocklength codes achieve virtually the same performance. Finally, Section 6.10 summarizes our main findings. The material in this chapter has been presented in part in [100, 101].

6.1 Previous work

For a given channel, the fundamental limit of traditional coding with fixed blocklength and no feedback is given by the function $M^*(n,\epsilon)$. We have demonstrated in previous chapters that for several channels the behavior of this function at fixed ϵ and moderate n is tightly characterized by the

$$\log M^*(n,\epsilon) = nC - \sqrt{nV}Q^{-1}(\epsilon) + O(\log n)\,, \tag{6.1}$$

where C is the channel capacity, V is the channel dispersion.

In the context of fixed blocklength communication, Shannon showed [31] that noiseless feedback does not increase the capacity of memoryless channels but can increase the zero-error capacity. For a class of symmetric DMCs, Dobrushin demonstrated [102] that the sphere-packing bound holds even in the presence of noiseless feedback. Similarly, it will be shown in Section 6.5 that for such channels the expansion (6.1) still holds with feedback as long as blocklength is not allowed to depend on feedback.

Nevertheless, it is known that feedback can be very useful provided that we allow variable-length codes. In his ground-breaking contribution, Burnashev [103] demonstrated that the error exponent improves in this setting and admits a particularly simple expression:

$$E(R) = \frac{C_1}{C}(C-R)\,, \tag{6.2}$$

for all rates $0 < R < C$, where C is the capacity of the channel and C_1 is the maximal relative entropy between output distributions. Moreover, zero-error capacity may improve from zero to the Shannon capacity (as in the case of the BEC) if variable-length is allowed. Furthermore, since existing communication systems with feedback (such as ARQ) have variable-length, in the analysis of fundamental limits for channels with feedback, it is much more relevant and interesting to allow codes whose length is allowed to depend on the channel behavior.

We mention a few extensions of Burnashev's work [103,104] relevant to this chapter. Yamamoto and Itoh proposed a simple and conceptually important two-phase coding scheme, attaining the optimal error exponent [105]. Using the notion of Goppa's [106] empirical mutual information (EMI) several authors have constructed universal coding schemes attaining rates arbitrarily close to capacity with small probability of error [107, 108], exponentially decaying probability of error [109] and even attaining the optimal Burnashev exponent [110, 111] simultaneously for a collection of channels. An extension to arbitrary varying channels with full state information available at the decoder has been recently proposed as well [112].

The error exponent analysis focused on fixed rate, rather than fixed probability of error as in (6.1). Another aspect that was not previously addressed in the literature is the following. In practice, control information, marking the beginning and the end of a packet, is rarely handled by the physical layer code. This contrasts with Burnashev's setting in which control layer signaling is modeled on the same noisy channel as the physical layer one. Moreover, as (6.2) shows the error exponent is, in fact, limited by the reliability with which the termination information is conveyed to the receiver through the DMC. To address this issue, we propose a simple modification of the (forward) channel model through

the introduction of a "use-once" termination symbol whose transmission disables further communication.

6.2 Channels and codes with feedback

In this chapter we consider a restricted (compared to Definition 1) class of channels. A non-anticipatory channel consists of a pair of input and output alphabets $\mathcal{A}$ and $\mathcal{B}$ together with a collection of conditional probability kernels $\{P_{Y_i|X_1^iY_1^{i-1}}\}_{i=1}^{\infty}$. Such channel is called (stationary) memoryless if

$$P_{Y_i|X_1^iY_1^{i-1}} = P_{Y_i|X_i} = P_{Y_1|X_1}\,, \quad \forall i \geq 1 \tag{6.3}$$

and if $\mathcal{A}$ and $\mathcal{B}$ are finite, it is known as a DMC.

Definition 14 *An (ℓ, M, ϵ) variable-length feedback (VLF) code, where ℓ is a positive real, M is a positive integer and $0 \leq \epsilon \leq 1$, is defined by:*

1. *A space $\mathcal{U}$ with[1] $|\mathcal{U}| \leq 3$ and a probability distribution P_U on it, defining a random variable U which is revealed to both transmitter and receiver before the start of transmission; i.e. U acts as common randomness used to initialize the encoder and the decoder before the start of transmission.*

2. *A sequence of encoders $f_n : \mathcal{U} \times \{1, \ldots, M\} \times \mathcal{B}^{n-1} \to \mathcal{A}$, $n \geq 1$, defining channel inputs*
$$X_n = f_n(U, W, Y^{n-1})\,, \tag{6.4}$$
where $W \in \{1, \ldots, M\}$ is the equiprobable message.

3. *A sequence of decoders $g_n : \mathcal{U} \times \mathcal{B}^n \to \{1, \ldots, M\}$ providing the best estimate of W at time n.*

4. *A non-negative integer-valued random variable τ, a stopping time of the filtration $\mathcal{G}_n = \sigma\{U, Y_1, \ldots, Y_n\}$, which satisfies*
$$\mathbb{E}\,[\tau] \;\leq\; \ell\,. \tag{6.5}$$

The final decision $\hat{W}$ is computed at the time instant τ:

$$\hat{W} = g_\tau(U, Y^\tau)\,, \tag{6.6}$$

and must satisfy

$$\mathbb{P}[\hat{W} \neq W] \;\leq\; \epsilon\,. \tag{6.7}$$

The fundamental limit of channel coding with feedback is given by the following quantity:

$$M_f^*(\ell, \epsilon) = \max\{M : \exists(\ell, M, \epsilon)\text{-}VLF\ code\}\,. \tag{6.8}$$

[1]The bound on the cardinality of $\mathcal{U}$ is to be justified shortly; see Theorem 97 below.

Those codes that do not require the availability of U, i.e. the ones with $|\mathcal{U}| = 1$, are called *deterministic* codes. Although from a practical viewpoint there is hardly any motivation to allow for non-deterministic codes, they simplify the analysis and expressions just like randomized tests do in hypothesis testing. Also similar to the latter, the difference in performance between the deterministic and non-deterministic codes is negligible for any practically interesting M and ℓ.

In a VLF code the decision about stopping transmission is taken solely upon observation of channel outputs in a causal manner. This is the setup investigated by Burnashev [103]. Note that since τ is computed at the decoder, it is not necessary to specify the values of $g_n(Y^n)$ for $n \neq \tau$. In this way the decoder is a map $g : \mathcal{B}^\infty \to \{1, \ldots, M\}$ measurable with respect to $\mathcal{G}_\tau$.

Definition 15 *An (ℓ, M, ϵ) variable-length feedback code with termination (VLFT), where ℓ is a positive real, M is a positive integer and $0 \leq \epsilon \leq 1$, is defined similarly to VLF codes with an exception that condition 4) in the Definition 14 is replaced by*

4') A non-negative integer-valued random variable τ, a stopping time of the filtration $\mathcal{G}_n = \sigma\{W, U, Y_1, \ldots, Y_n\}$, which satisfies

$$\mathbb{E}[\tau] \quad \leq \quad \ell\,. \tag{6.9}$$

The fundamental limit of channel coding with feedback and termination is given by the following quantity:

$$M_{\mathsf{t}}^*(\ell, \epsilon) = \max\{M : \exists(\ell, M, \epsilon)\text{-}VLFT\ code\}\,. \tag{6.10}$$

In a VLFT code, "termination" is used to indicate the fact that the practical realization of such a coding scheme requires a method of sending a reliable end-of-packet signal by means other than using the $\mathcal{A} \to \mathcal{B}$ channel (e.g., by cutting off a carrier). As we discussed in the introduction, timing (including termination) is usually handled by a different layer in the protocol. The following are examples of VLFT codes:

1. VLF codes are a special case in which the stopping time τ is determined autonomously by the decoder; due to availability of the feedback, τ is also known to the encoder so that transmission can be cut off at τ.

2. *decision feedback codes* are a special case of VLF codes where encoder functions $\{f_n\}_{n=1}^\infty$ satisfy:
$$f_n(U, W, Y^{n-1}) = f_n(U, W)\,. \tag{6.11}$$
Such codes require very limited communication over feedback: only a single signal to stop the transmission once the decoder is ready to decode.

3. variable-length codes (without feedback), or *VL codes*, defined in [16, Problem 2.1.25] and [113], are VLFT codes required to satisfy two additional requirements: τ is a function of (W, U) and the encoder is not allowed to use feedback, i.e. (6.11) holds. The fundamental limit and the ϵ-capacity of variable-length codes are given by
$$M_v^*(\ell, \epsilon) \quad = \quad \max\{M : \exists(\ell, M, \epsilon)\text{-VL code}\}\,, \tag{6.12}$$
$$[\![C_\epsilon]\!] \quad = \quad \lim_{\ell \to \infty} \frac{1}{\ell} \log M_v^*(\ell, \epsilon)\,. \tag{6.13}$$

4. An (n, M, ϵ) *fixed-blocklength feedback code* is an (n, M, ϵ) VLF code with $\tau = n$. The fundamental limit of fixed-blocklength feedback codes is given by

$$M_b^*(n, \epsilon) = \max\{M : \exists(n, M, \epsilon) \text{ fixed-length feedback code}\}. \tag{6.14}$$

5. fixed-to-variable codes, or *FV codes*, defined in [113] are also required to satisfy (6.11), while the stopping time is[2]

$$\tau = \inf\{n \geq 1 : g_n(U, Y^n) = W\}, \tag{6.15}$$

and therefore, such codes are zero-error VLFT codes. Of course, not all zero-error VLFT codes are FV codes, since in general condition (6.11) does not necessarily hold.

6. automatic repeat request (ARQ) are yet a more restricted class of deterministic FV codes, where a single fixed-blocklength, non-feedback code is used repeatedly until the decoder produces a correct estimate.

The next result shows that restriction on the cardinality of $\mathcal{U}$ in the Definitions 14 and 15 does not entail loss of generality.

Theorem 97 *Consider an (ℓ, M, ϵ) VLFT code possibly violating the cardinality requirement $|\mathcal{U}| \leq 3$. Then there exists an (ℓ, M, ϵ) VLFT code with $|\mathcal{U}| \leq 3$.*

Proof: Denote by G_k the following subsets of $\mathbb{R}^2$:

$$G_k \stackrel{\triangle}{=} \{(\ell', \epsilon') : \exists(\ell', M, \epsilon')\text{-code with } |\mathcal{U}| \leq k\}, k = 1, 2, \ldots, \tag{6.16}$$

and

$$G_\infty \stackrel{\triangle}{=} \{(\ell', \epsilon') : \exists(\ell', M, \epsilon')\text{-code}\}. \tag{6.17}$$

Notice that G_∞ is a convex hull of G_1 since by taking a general code and conditioning on U we obtain a deterministic code. By Caratheodory's theorem we then know that $G_3 = G_\infty$. Since by assumption $(\ell, \epsilon) \in G_\infty$ then $(\ell, \epsilon) \in G_3$. ■

The main goal of this chapter is to analyze the behavior of $\log M_f^*(\ell, \epsilon)$ and $\log M_{\mathtt{t}}^*(\ell, \epsilon)$ and compare them with the behavior of the fundamental limit without feedback, $\log M^*(n, \epsilon)$. Regarding the behavior of $\log M_f^*(\ell, \epsilon)$ Burnashev's result (6.2) can be restated as

$$\log M_f^*(\ell, \exp\{-E\ell\}) = \ell C \left(1 - \frac{E}{C_1}\right) + o(\ell), \tag{6.18}$$

for any $0 < E < C_1$. Although (6.18) does not imply any statement about the expansion of $\log M_f^*(\ell, \epsilon)$ for a fixed ϵ, it still demonstrates that in the regime of very small probability of error, the parameter C_1 emerges as an important quantity.

[2]As explained in [113], this model encompasses fountain codes in which the decoder can get a highly reliable estimate of τ autonomously without the need for a termination symbol.

6.3 Synchronized channels

In the previous section we have defined the notion of feedback codes. For the sake of clarity, however, we defined those only for a non-anticipatory channel. Although for the remainder of this chapter only non-anticipatory channels are considered, for completeness we also include a general treatment.

Notice that our definition of the channel, see Definition 1, is too general for studying the questions involving the notions of time and causality. For this reason we have introduced the concept of a non-anticipatory channel. This, however, can be done in a more abstract way.

Consider a random transformation $(\mathsf{A}, \mathsf{B}, P_{Y|X})$ in the sense of Definition 1. We denote σ-algebras on A and B by $\mathcal{F}$ and $\mathcal{G}$, resp. Recall that a transition probability kernel $P_{\mathbf{Y}|\mathbf{X}}$ from $(\mathsf{A}, \mathcal{F})$ to $(\mathsf{B}, \mathcal{G})$ is required to satisfy two conditions:

1. for a fixed $\mathbf{x} \in \mathsf{A}$, $P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}(\cdot)$ is a probability measure on $(\mathsf{B}, \mathcal{G})$, and
2. for a fixed $E \in \mathcal{G}$ the function $\mathbf{x} \mapsto P_{\mathbf{Y}|\mathbf{X}=\mathbf{x}}(E)$ is $\mathcal{F}$-measurable.

Definition 16 *A synchronized channel is a random transformation* $(\mathsf{A}, \mathsf{B}, P_{\mathbf{Y}|\mathbf{X}})$ *with filtrations* $\mathcal{F}_n$ *and* $\mathcal{G}_n$ *on* A *and* B*, resp., and the requirement that* $P_{Y|X}$ *be a transition probability kernel from* $(\mathsf{A}, \mathcal{F}_n)$ *to* $(\mathsf{B}, \mathcal{G}_n)$ *for each* $n \geq 0$*.*

To draw a parallel with Section 6.2 and also for notational simplicity we may assume that there have been pre-selected two sequences of functions on A and B, such that

$$\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}\,, \text{ and } \mathcal{G}_n = \sigma\{Y_1, \ldots, Y_n\}\,. \tag{6.19}$$

Consider examples of synchronized channels:

1. Any "single-letter" channel $(\mathcal{A}, \mathcal{B}, P_{Y|X})$ can be extended memorylessly to a synchronized channel by the following construction. We take $\mathsf{A} = \mathcal{A}^\infty$ and X_j as the usual projections onto j-th coordinate; similarly we construct $\mathsf{B} = \mathcal{B}^\infty$ and Y_j. The kernel $P_{\mathbf{Y}|\mathbf{X}}$ is defined as an extension of the following sequence of finite dimensional kernels:

$$P_{Y^n|X^n=(x_1,\ldots x_n)} = \prod_{j=1}^{n} P_{Y|X=x_1}\,, \tag{6.20}$$

   where the product is the product of measures on $\mathcal{B}$. A synchronized channel obtained in this way starting from finite spaces $\mathcal{A}$ and $\mathcal{B}$ is called discrete memoryless (i.e., a DMC).

2. Any non-anticipatory channel is defined by "single-letter" spaces $\mathcal{A}$, $\mathcal{B}$ and a collection of conditional distributions $P_{Y_i|X_1{}^iY_1{}^{i-1}}$; it defines a synchronized channel by a natural extension of the product construction discussed previously.

3. Any non-anticipatory channel, in particular any DMC, can be extended to a channel with termination symbol by the following construction:

$$\begin{aligned} \mathcal{A}' &= \mathcal{A} \cup \{T\}\,, && (6.21)\\ \mathcal{B}' &= \mathcal{B} \cup \{T\}\,, && (6.22)\\ P'_{Y_i|X_1^i Y_1^{i-1}}(b_i|a_1^i b_1^{i-1}) &= \begin{cases} P_{Y_i|X_1^i Y_1^{i-1}}(b_i|a_1^i b_1^{i-1})\,, & a_1 \neq T, \ldots, a_i \neq T \\ 1\{b_i = T\}\,, & \text{otherwise}\,. \end{cases} && (6.23) \end{aligned}$$

Notice that the so-extended channel is also non-anticipatory.

These examples demonstrate that definition of a synchronized channel generalizes the concept of non-anticipatory channel by dropping the requirement that all X_j's and Y_j's in (6.19) take values in the same spaces $\mathcal{A}$ and $\mathcal{B}$, respectively. The next definition shows how the concept of a VLF code generalizes to synchronized channels.

Definition 17 *An (ℓ, M, ϵ) code variable-length code with feedback (VLF code) for a synchronized channel $(\mathsf{A}, \mathsf{B}, P_{Y|X}, \mathcal{F}_n, \mathcal{G}_n)$ is defined by:*

1. *A space $\mathcal{U}$ with $|\mathcal{U}| \leq 3$ and a probability distribution P_U on it, defining a random variable U. On the space $\mathsf{B} \times \mathcal{U}$ we define a filtration*

$$\mathcal{G}'_n = \sigma\mathcal{U} \times \mathcal{G}_n\,. \tag{6.24}$$

2. *an encoder mapping $f : \{1, \ldots, M\} \times \mathcal{U} \times \mathsf{B} \to \mathsf{A}$ satisfying causality constraint*

$$f^{-1}\mathcal{F}_n \subset \mathcal{H}_M \times \mathcal{G}'_{n-1}\,, \tag{6.25}$$

where $\mathcal{H}_M$ is a σ-algebra of all subsets of $\{1, \ldots, M\}$;

3. *a stopping time $\tau \geq 0$ of the filtration $\mathcal{G}'_n$, satisfying*

$$\mathbb{E}\,[\tau] \leq \ell\,, \tag{6.26}$$

4. *a decoder mapping $g : \mathcal{U} \times \mathsf{B} \to \{1, \ldots, M\}$ measurable with respect to $\mathcal{G}'_\tau$ and satisfying*

$$\mathbb{P}[g(U, \mathbf{Y}) \neq W] \leq \epsilon\,. \tag{6.27}$$

To complete the definition we need to specify the probability space which is used for $\mathbb{E}$ and $\mathbb{P}$ in (6.26) and (6.27). This space is taken to be

$$\Omega = \{1, \ldots, M\} \times \mathcal{U} \times \mathsf{A} \times \mathsf{B}\,, \tag{6.28}$$

with σ-algebra $\mathcal{H}_M \times \sigma\mathcal{U} \times \mathcal{F} \times \mathcal{G}$. The projection $\Omega \to \{1, \ldots, M\}$ is denoted by W; the projection $\Omega \to \mathcal{U}$ is denoted U; projections X_j and Y_j, $j = 1, \ldots, \infty$ are defined according to (6.19). The probability distribution on Ω is defined recursively:

1. the distribution of (W, U) is taken to be:

$$P_{WU}(w, u) = \frac{1}{M} P_U(u)\,. \tag{6.29}$$

2. the measure on (W, U, X_1) is defined as a push-forward along $f : \mathcal{H}_M \times \sigma\mathcal{U} \to \mathcal{F}_1$;

3. the measure on (W, U, X_1, Y_1) is defined by applying the kernel $P_{Y_1|X_1}$;

4. once the measure on (W, U, X^{n-1}, Y^{n-1}) has been defined we again use push-forward along $f : \mathcal{H}_M \times \sigma\mathcal{U} \times \mathcal{G}_{n-1} \to \mathcal{F}_n$ to extend the measure to (W, Y^{n-1}, X^n) and then we apply the kernel $P_{Y^n|X^nY^{n-1}}$ to extend to (W, X^n, Y^n).

5. the restriction of such measure to $\mathcal{H}_M \times \sigma\mathcal{U} \times \mathcal{F}_k \times \mathcal{G}_k$, $k < n$ coincides with the measure defined on k-th step; therefore, this sequence of measures (by Kolmogorov's theorem) extends to a measure on $\mathcal{H}_M \times \sigma\mathcal{U} \times \mathcal{F} \times \mathcal{G}$.

Similarly, we can define a VLFT code for the synchronized channel. Notice that VLFT code for the non-anticipatory channel can be equivalently viewed as a VLF code for a different non-anticipatory channel (6.23).

6.4 Automatic repeat request (ARQ)

In this section we consider a simple zero-error VLFT code, known as ARQ, in which a packet (protected by a forward error correcting code) is retransmitted until the receiver acknowledges successful decoding (which the receiver determines using a variety of known highly reliable hashing methods). Typically, the size k of the information packets is determined by the particular application, and both the blocklength n and the block error probability ϵ are degrees of freedom. In this section we focus on the average data rate (error-free) that ARQ is able to deliver to the destination. For a discussion focused more on energy efficiency see [76].

Given an $(n, 2^k, \epsilon)$ block code, and assuming that decoding errors are independent for different retransmissions, the average number of channel uses is given by

$$\mathbb{E}[\tau] = \frac{n}{1-\epsilon}\,. \tag{6.30}$$

Therefore, to maximize the rate $\frac{k}{\mathbb{E}[\tau]}$ we have to solve the following optimization problem:

$$T(k) = \max_{n,\epsilon} \frac{k}{n}(1-\epsilon)\,, \tag{6.31}$$

where the maximization is over those (n, ϵ) such that

$$\log_2 M^*(n, \epsilon) = k\,. \tag{6.32}$$

As we have demonstrated in previous chapters, for many channels the normal approximation (2.23) is tight; therefore, equivalently we can maximize

$$\tilde{T}(k) = \max_n \frac{k}{n}\left[1 - Q\left(\frac{nC-k}{\sqrt{nV}}\right)\right]\,, \tag{6.33}$$

where C and V are the channel capacity and channel dispersion, respectively. For the BSC with $\delta = 0.11$ we show the results of the optimization in (6.33) in Fig. 6.1, where the

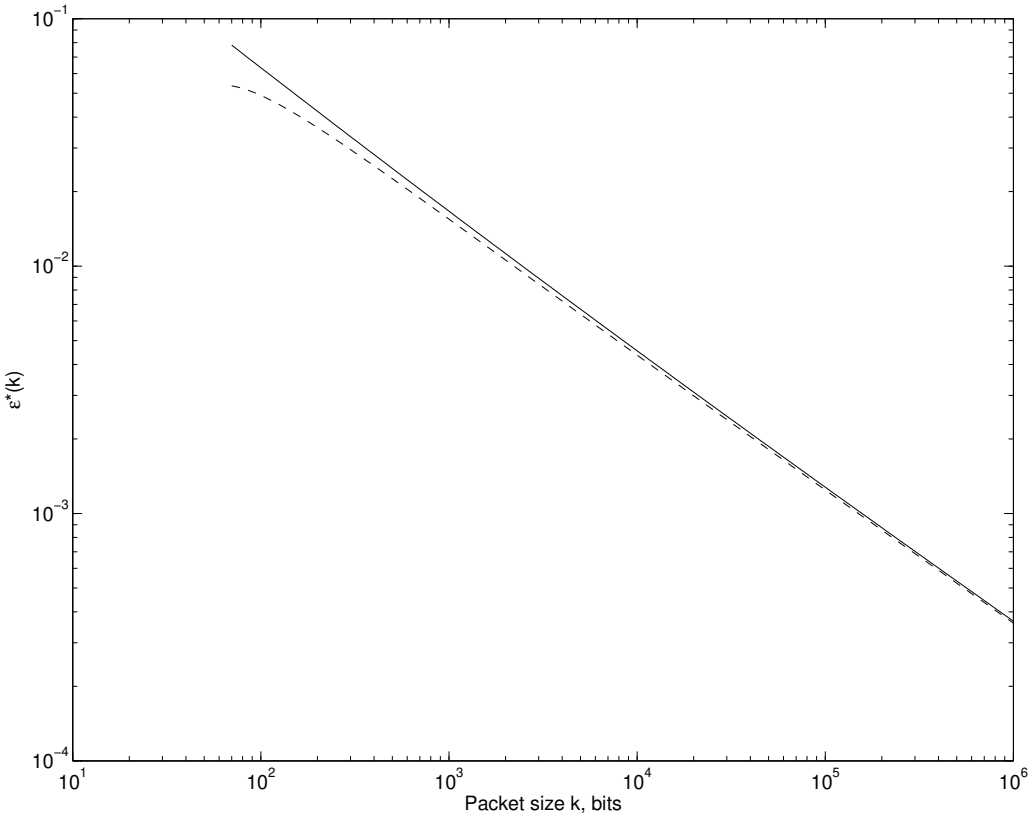


Figure 6.1: Optimal block error rate $\epsilon^*(k)$ maximizing average throughput under ARQ feedback for the BSC with $\delta = 0.11$. Solid curve is obtained by using normal approximation, dashed curve is an asymptotic formula (6.34).

optimal block error rate, $\epsilon^*(k)$ is shown, and Fig. 6.2, where the optimal coding rate $\frac{k}{n^*(k)}$ is shown. Table 6.1 shows the results of the optimization for the channel examples we have used throughout the chapter. Of particular note is that for 1000 information bits, and a capacity-1/2 BSC, the optimal block error rate is as high as 0.0167. Similar observations regarding the optimal block error rate have also been made in [114].

The tight approximation to the optimal error probability as a function of k in Figure 6.1 is the function

$$\tilde{\epsilon}(k) = \left(\frac{kC}{V}\ln\frac{kC}{2\pi V}\right)^{-1/2}\left(1 - \frac{1}{\ln\frac{kC}{2\pi V}}\right) \tag{6.34}$$

obtained by retaining only the dominant terms in the asymptotic solution as $k \to \infty$.

6.5 Fixed-blocklength codes with feedback

In the case of the BEC, the tightest converse bound, Theorem 43, has been proved in Section 3.3 for fixed-blocklength codes already under assumption of availability of feedback.

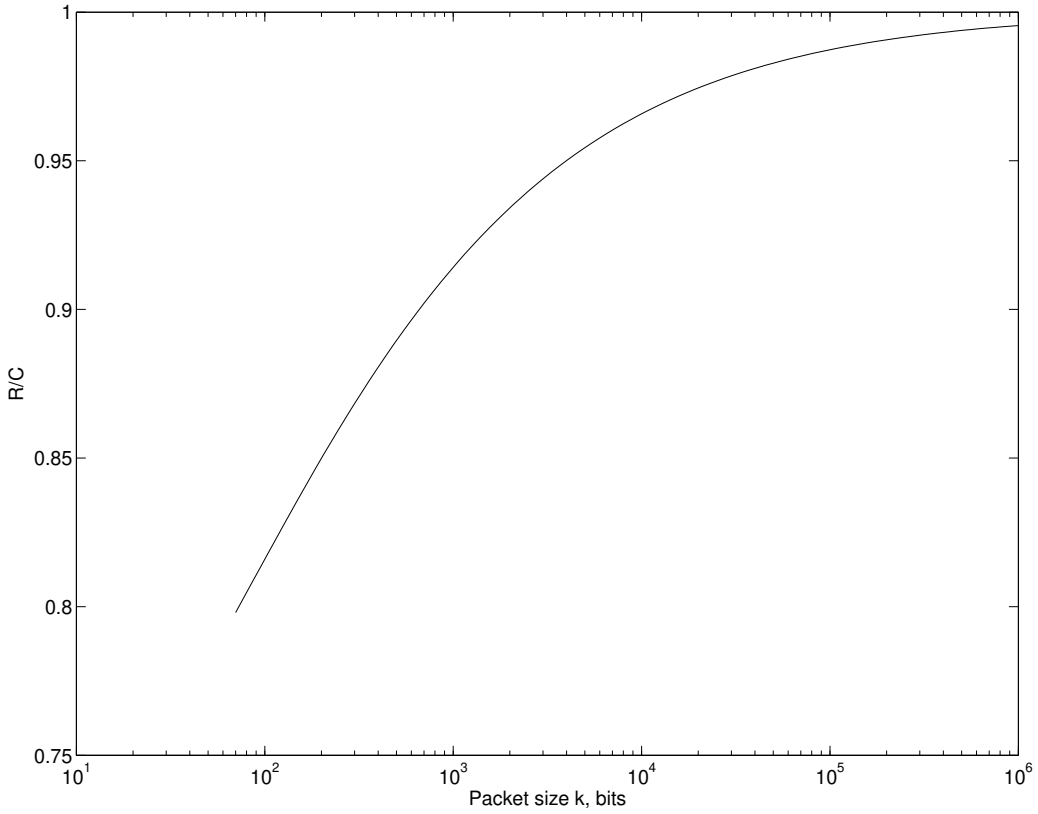


Figure 6.2: Optimal rate of a constituent block code, that maximizes the average throughput under ARQ feedback for the BSC with $\delta = 0.11$. Solid curve is obtained using normal approximation.

Therefore, the proof of the asymptotic expansion, Theorem 44, automatically applies to the feedback case and we have:

Theorem 98 *For the BEC we have*

$$\log M_b^*(n,\epsilon) = nC - \sqrt{nV}Q^{-1}(\epsilon) + O(1)\,, \tag{6.35}$$

where C and V are the capacity and the dispersion of the BEC.

Therefore, we see that in this case the feedback is unable to improve the penalty $\sqrt{n}$-term. In fact, much more is true. The numerical comparison of the converse and achievability bounds for the BEC, see Section 3.3.3, has demonstrated that the converse bound (which holds even for feedback codes) can be approached extremely closely by the non-feedback block codes. Namely, it was shown that non-feedback codes exist that achieve values of $\log_2 M$ within 2-3 bits of the converse bound for all blocklengths $n \gtrsim 10$. Consequently, it implies that the potential benefit of feedback is limited to enlarging $\log_2 M$ by those 2-3 bits at most.

The same conclusion holds for a wide class of weakly-input symmetric channels (including the BSC), see Section 3.4.5 for relevant definitions.

Table 6.1: Optimal block error rate for packet size $k = 1000$ bits

Channel	Optimal $\epsilon^*(k)$	Optimal R/C	Optimal throughput
BEC(0.5)	$8.1 \cdot 10^{-3}$	0.95	0.94
BSC(0.11)	$16.7 \cdot 10^{-3}$	0.91	0.90
AWGN, SNR = 0dB	$15.5 \cdot 10^{-3}$	0.92	0.90
AWGN, SNR = 20dB	$6.2 \cdot 10^{-3}$	0.96	0.95

Theorem 99 *Consider a weakly input-symmetric DMC with capacity C and dispersion V. Then $M_b^*(n,\epsilon)$ satisfies the non-feedback bound* (3.341)*:*

$$\log M_b^*(n,\epsilon) \le -\log \beta_{1-\epsilon}((P_{Y|X=x_0})^n, (P_Y^*)^n)\,, \tag{6.36}$$

where P_Y^* and x_0 are as defined in Definition 9. Consequently, we have

$$\log M_b^*(n,\epsilon) \le nC - \sqrt{nV}Q^{-1}(\epsilon) + \frac{1}{2}\log n + O(1)\,, \tag{6.37}$$

if $V > 0$ and

$$\log M_b^*(n,\epsilon) \le nC - \log(1-\epsilon)\,, \tag{6.38}$$

if $V = 0$.

This result is not surprising, in view of the classical result of Dobrushin [102] that for certain symmetric channels the sphere-packing bound on the error-exponent holds in the presence of feedback (although, the class of weakly input symmetric channels is much larger; see Section 3.4.5). To illuminate the non-asymptotic nature of the bound (6.36), notice that for the BSC the β_α in the right-hand side of (6.36) coincides with β_α^n in the sphere-packing converse, Theorem 40. Therefore, according to the results of Section 3.2.3 the bound (6.36) is achievable to within a 3-4 bits by non-feedback block codes (for a wide range of n). Therefore, for the BSC and such n, feedback codes can improve the value of $\log M$ compared to the non-feedback ones by at most 3-4 bits (!).

Proof: Fix an (n, M, ϵ) fixed-blocklength feedback code. Its encoder defines a transition probability kernel $P_{Y^n|W}$ from the input space

$$\mathcal{D}_M \stackrel{\triangle}{=} \{1, \ldots, M\} \tag{6.39}$$

to the output space $\mathcal{B}^n$. We can view then the triplet $(\mathcal{D}_M, \mathcal{B}^n, P_{Y^n|W})$ as a random transformation for which we have a usual (M,ϵ) code in the sense of Definition 2. For such a code Theorem 29 shows

$$M \le \frac{1}{\beta_{1-\epsilon}(P_{WY^n}, P_W Q_{Y^n})}\,, \tag{6.40}$$

where P_W is the equiprobable distribution on $\mathcal{D}_M$ and Q_{Y^n} is a product distribution

$$Q_{Y^n} \stackrel{\triangle}{=} (P_Y^*)^n\,. \tag{6.41}$$

Therefore, the proof of (6.36) will be complete if we can show

$$\beta_\alpha(P_{WY^n}, P_W Q_{Y^n}) \ge \beta_\alpha((P_{Y|X=x_0})^n, Q_{Y^n}) \tag{6.42}$$

By Lemma 32 to show (6.42) it is enough to prove that for any $j \in \{1, \ldots, M\}$ we have

$$\beta_\alpha(P_{Y^n|W=j}, Q_{Y^n}) \geq \beta_\alpha((P_{Y|X=x_0})^n, Q_{Y^n})\,. \tag{6.43}$$

Fix arbitrary $j \in \{1, \ldots, M\}$ and $x_0 \in \mathcal{A}$. The sequence of encoder functions $f_k, k = 1, \ldots, n$ defines the measure $P_{Y^n|W=j}$ as follows:

$$P_{Y^n|W=j}(y^n) = \prod_{k=1}^{n} P_{Y|X}(y_k|f_k(j, y^{k-1}))\,. \tag{6.44}$$

Since the channel is weakly input-symmetric, to each $x \in \mathcal{A}$ there exists a transformation $T_x : \mathcal{B} \to \mathcal{B}$ such that

$$P_{Y|X=x} = T_x \circ P_{Y|X=x_0}\,, \tag{6.45}$$

where the composition is understood as in (2.2) (see also (3.332)). We will now define a transformation $T_j : \mathcal{B}^n \to \mathcal{B}^n$ as follows

$$T_j(z^n|y^n) = \prod_{k=1}^{n} T_{f_k(j,y^{k-1})}(z_k|y_k)\,. \tag{6.46}$$

Then according to this construction and (6.44), on the one hand we have

$$T_j \circ (P_{Y|X=x_0})^n = P_{Y^n|W=j}\,, \tag{6.47}$$

whereas on the other hand, since each T_x preserves P_Y^*, we also have

$$T_j \circ Q_{Y^n} = (P_Y^*)^n\,. \tag{6.48}$$

Then it follows that

$$\begin{aligned} \beta_\alpha(P_{Y^n|W=j}, Q_{Y^n}) &= \beta_\alpha(T_j \circ (P_{Y|X=x_0})^n, T_j \circ Q_{Y^n}) && (6.49)\\ &\geq \beta_\alpha((P_{Y|X=x_0})^n, Q_{Y^n})\,, && (6.50) \end{aligned}$$

where (6.49) follows by (6.47) and (6.48), and (6.50) follows by data-processing form β_α (i.e. simultaneous application of T_j to both measures cannot improve the value of β_α). This completes the proof of (6.36). Expansions (6.37) and (6.38) follow from (2.89) and (2.90), respectively, after applying (3.339). ■

6.6 Variable-length codes (without feedback)

The next result shows that under variable-length coding allowing a non-vanishing error probability ϵ boosts the ϵ-capacity by a factor of $\frac{1}{1-\epsilon}$ even in the absence of feedback:

Theorem 100 *For any non-anticipatory channel with capacity C that satisfies the strong converse for fixed-blocklength codes (without feedback), the ϵ-capacity under variable-length coding without feedback, cf.* (6.13)*, is*

$$[\![C_\epsilon]\!] = \frac{C}{1-\epsilon}\,, \epsilon \in (0, 1)\,. \tag{6.51}$$

In general, it is known [113, Theorem 16] that the VL capacity, $[\![C]\!] = \lim_{\epsilon\to 0}[\![C_\epsilon]\!]$, is equal to the conventional fixed-blocklength capacity without feedback, C, for any non-anticipatory channel (not necessarily satisfying the strong converse). On the other hand, the capacity of FV codes for state-dependent non-ergodic channels can be larger than C [113].

Proof: Fix $\epsilon' < \epsilon$ and a large n. Then there exists a fixed-blocklength code without feedback with blocklength n, probability of error ϵ' and number of messages M satisfying:

$$\log M \geq nC + o(n)\,. \tag{6.52}$$

Consider the following variable-length code (without feedback): with probability $\frac{1-\epsilon}{1-\epsilon'}$ encoder sends a codeword of length n, otherwise it sends nothing. It is easy to see that the probability of decoding error is upper-bounded by ϵ whereas the average transmission time is equal to $\ell = \frac{1-\epsilon}{1-\epsilon'}n$, and therefore the average transmission rate is

$$R \stackrel{\triangle}{=} \frac{\log M}{\ell} \geq C\frac{1-\epsilon'}{1-\epsilon} + o(1)\,. \tag{6.53}$$

By taking the limit $n\to\infty$ we obtain

$$[\![C_\epsilon]\!] \geq C\frac{1-\epsilon'}{1-\epsilon}\,. \tag{6.54}$$

Since ϵ' is arbitrary we can achieve any rate close to $\frac{C}{1-\epsilon}$.

For the converse recall that a channel is said to satisfy strong converse if its fixed-blocklength no feedback fundamental limit $\log M^*(n,\epsilon)$ satisfies

$$\log M^*(n,\epsilon) = nC + o(n)\,, n\to\infty\,, \quad \forall \epsilon\in(0,1)\,. \tag{6.55}$$

Now, consider an (ℓ, M, ϵ) variable-length code. Define the following quantities for each $n\geq 0$ and $u\in\mathcal{U}$:

$$\epsilon(n,u) \;\; = \;\; \mathbb{P}[\hat{W}\neq W|\tau = n, U = u]\,, \tag{6.56}$$

which satisfy, of course,

$$\mathbb{E}\left[\epsilon(\tau,U)\right] \leq \epsilon\,. \tag{6.57}$$

Fix u and notice that conditioned on $U=u$, τ is a function of W, and therefore $M\mathbb{P}[\tau = n|U=u]$ is an integer. Then the condition $\tau = n$ defines an $(n, M\mathbb{P}[\tau=n|U=u], \epsilon(n,u))$ fixed blocklength subcode. Therefore, we have for each $n\geq 0$:

$$\mathbb{P}[\tau = n|U=u]M \leq M^*(n,\epsilon(n,u))\,. \tag{6.58}$$

We now fix arbitrary $N\geq 0$ and $\epsilon' > 0$ and sum (6.58) for all $n\leq N$ such that $\epsilon(n,u)\leq\epsilon'$:

$$\begin{aligned} M\mathbb{P}[\tau\leq N, \epsilon(\tau,u)\leq\epsilon'|U=u] \;\; &\leq \;\; \sum_{n=0}^{N} M^*(n,\epsilon(n,u))1\{\epsilon(n,u)\leq\epsilon'\}\,, && (6.59)\\ &\leq \;\; \sum_{n=0}^{N} M^*(n,\epsilon')\,, && (6.60)\\ &\leq \;\; NM^*(N,\epsilon')\,, && (6.61)\end{aligned}$$

where (6.60) follows since by definition $M^*(n,\epsilon)$ is a non-decreasing function of ϵ, and (6.61) follows because for a non-anticipatory channel $M^*(n,\epsilon)$ is also a non-decreasing function of n. By taking the expectation of (6.61) with respect to U we obtain

$$M\mathbb{P}[\tau \leq N, \epsilon(\tau, U) \leq \epsilon'] \leq NM^*(N, \epsilon')\,. \tag{6.62}$$

On the other hand, by the Chebyshev inequality we have

$$\begin{aligned} \mathbb{P}[\tau \leq N, \epsilon(\tau, U) \leq \epsilon'] &\geq 1 - \frac{\mathbb{E}\,[\tau]}{N} - \frac{\mathbb{E}\,[\epsilon(\tau, U)]}{\epsilon'} && (6.63)\\ &\geq 1 - \frac{\ell}{N} - \frac{\epsilon}{\epsilon'}\,. && (6.64) \end{aligned}$$

Finally, we choose $\epsilon' > \epsilon$ and take

$$N = \frac{\ell + 1}{1 - \epsilon/\epsilon'}\,. \tag{6.65}$$

Now from (6.62), (6.64) and (6.65) we obtain

$$\begin{aligned} \log M &\leq \log M^*\left(\frac{\ell+1}{1-\epsilon/\epsilon'}, \epsilon'\right) + 2\log\frac{\ell+1}{1-\epsilon/\epsilon'} && (6.66)\\ &= C\frac{\ell+1}{1-\epsilon/\epsilon'} + o(\ell)\,, && (6.67) \end{aligned}$$

where (6.67) follows from strong converse (6.55). Dividing both sides of (6.67) by ℓ we have proven that the rate of any (ℓ, M, ϵ) variable-length code must satisfy:

$$\frac{\log M}{\ell} \leq \frac{C}{1-\epsilon/\epsilon'} + o(1)\,, \tag{6.68}$$

or in other words, for any $\epsilon' > \epsilon$ we have

$$[\![C_\epsilon]\!] \leq \frac{C}{1-\epsilon/\epsilon'}\,. \tag{6.69}$$

Taking $\epsilon' \to 1$ completes the proof. ■

6.7 Variable-length codes with feedback

Our main result is the following:

Theorem 101 *For an arbitrary DMC with capacity C we have for any $0 < \epsilon < 1$*

$$\begin{aligned} \log M_f^*(\ell, \epsilon) &= \frac{\ell C}{1-\epsilon} + O(\log \ell)\,, && (6.70)\\ \log M_{\mathtt{t}}^*(\ell, \epsilon) &= \frac{\ell C}{1-\epsilon} + O(\log \ell)\,. && (6.71) \end{aligned}$$

More precisely, we have

$$\begin{aligned} \frac{\ell C}{1-\epsilon} - \log \ell + O(1) &\leq \log M_f^*(\ell, \epsilon) \leq \frac{\ell C}{1-\epsilon} + O(1)\,, && (6.72)\\ \log M_f^*(\ell, \epsilon) &\leq \log M_{\mathtt{t}}^*(\ell, \epsilon) \leq \frac{\ell C + \log \ell}{1-\epsilon} + O(1)\,. && (6.73) \end{aligned}$$

A consequence of Theorem 101 is that for DMCs, feedback (even in the setup of VLFT codes) does not increase the ϵ-capacity, namely,

$$\lim_{\ell\to\infty}\frac{1}{\ell}\log M_t^*(\ell,\epsilon)=[\![C_\epsilon]\!]\,, \tag{6.74}$$

where $[\![C_\epsilon]\!]$ is defined in (6.13) and given by Theorem 100.

However, while in the absence of feedback and within the paradigm of fixed-length coding, the backoff from ϵ-capacity (equal to capacity for DMCs) is governed by the $\frac{1}{\sqrt{n}}$ term (6.1), variable-length coding with feedback completely eliminates that penalty. Thus, the capacity is attainable at a much smaller (average) blocklength. Furthermore, the achievability (lower) bound in (6.72) is obtained via decision feedback codes that use feedback only to let the encoder know that the decoder has made its final decision; namely, the encoder maps f_n satisfy (6.11). As (6.72) demonstrates, such a sparing use of feedback does not lead to any significant loss in rate even non-asymptotically. Naturally, such a strategy is eminently practical in many applications, unlike those strategies that require full, noiseless, instantaneous feedback. In the particular case of the BSC, a lower bound (6.72) with a weaker $\log n$ term has been claimed in [109].

The proof of Theorem 101 is an application of a general achievability bound:

Theorem 102 *Fix a real number $\gamma>0$, a channel $\{P_{Y_i|X_1^iY_1^{i-1}}\}_{i=1}^\infty$ and an arbitrary process $X=(X_1,X_2,\ldots,X_n,\ldots)$ taking values in $\mathcal{A}$. Define a probability space with finite-dimensional distributions given by*

$$P_{X^nY^n\bar{X}^n}(a^n,b^n,c^n)=P_{X^n}(a^n)P_{\bar{X}^n}(c^n)\prod_{j=1}^n P_{Y_j|X_1^jY_1^{j-1}}(a_j|b^j,a^{j-1})\,, \tag{6.75}$$

i.e. X and $\bar{X}$ are independent copies of the same process and Y is the output of the channel when X is its input. For the joint distribution (6.75) *define a sequence of information density functions $\mathcal{A}^n\times\mathcal{B}^n\to\bar{\mathbb{R}}$*

$$i(a^n;b^n)=\log\frac{dP_{Y^n|X^n}(b^n|a^n)}{dP_{Y^n}(b^n)}\,, \tag{6.76}$$

and a pair of hitting times:

$$\tau = \inf\{n\ge 0: i(X^n;Y^n)\ge\gamma\}\,, \tag{6.77}$$
$$\bar{\tau} = \inf\{n\ge 0: i(\bar{X}^n;Y^n)\ge\gamma\}\,. \tag{6.78}$$

Then for any M there exists an (ℓ,M,ϵ) VLF code with

$$\ell \le \mathbb{E}[\tau] \tag{6.79}$$
$$\epsilon \le (M-1)\mathbb{P}[\bar{\tau}\le\tau]\,. \tag{6.80}$$

Furthermore, for any M there exists a deterministic (ℓ',M,ϵ) VLF code with ϵ satisfying (6.80) *and*

$$\ell'\le \operatorname{esssup}\mathbb{E}[\tau|X]\,. \tag{6.81}$$

Remarks:

1. Loosening the bound to (6.81) is advantageous, since for symmetric channels we have $\mathbb{E}[\tau|X] = \mathbb{E}[\tau]$ and thus the second part of Theorem 102 guarantees the existence of a deterministic code without any sacrifice in performance.

2. Theorem 102 is a natural extension of the DT bound, Theorem 18, since (6.80) corresponds to the second term in (2.117), whereas the first term in (2.117) is missing because the information density corresponding to the true message eventually crosses any level γ with probability one.

3. Interestingly, pairing a fixed stopping rule with a random-coding argument has been already discovered from a different perspective: in the context of universal variable-length codes [107–111], stopping rules based on a sequentially computed EMI were shown to be optimal in several different asymptotic senses. Although invaluable for universal coding, EMI-based decoders are hard to evaluate non-asymptotically and their analysis relies on inherently asymptotic methods, such as type-counting, cf. [111].

Proof: To define a code we need to specify (U, f_n, g_n, τ). First we define a random variable U as follows:

$$\mathcal{U} \stackrel{\triangle}{=} \underbrace{\mathcal{A}^\infty \times \cdots \times \mathcal{A}^\infty}_{M \text{ times}} \tag{6.82}$$

$$P_U \stackrel{\triangle}{=} \underbrace{P_{X^\infty} \times \cdots \times P_{X^\infty}}_{M \text{ times}}, \tag{6.83}$$

where P_{X^∞} is the distribution of the process X. Note that even for $|\mathcal{A}| = 2$, $\mathcal{U}$ will have the cardinality of a continuum. However, in view of Theorem 97 this can always be reduced to 3.

The realization of U defines M infinite dimensional vectors $\mathbf{C}_j \in \mathcal{A}^\infty, j = 1, \ldots, M$. Our encoder and decoder will depend on U implicitly through $\{\mathbf{C}_j\}$. The coding scheme consists of a sequence of encoders f_n that map a message j to an infinite sequence of inputs $\mathbf{C}_j \in \mathcal{A}^\infty$ without any regard to feedback:

$$f_n(w) = (\mathbf{C}_w)_n\,, \tag{6.84}$$

where $(\mathbf{C}_j)_n$ is the n-th coordinate of the vector $\mathbf{C}_j$. Obviously, such encoder satisfies (6.11).

At time instant n the decoder computes M information densities:

$$S_{j,n} \stackrel{\triangle}{=} i(\mathbf{C}_j(n); Y^n)\,, \qquad j = 1, \ldots, M, \tag{6.85}$$

where $\mathbf{C}_j(n)$ is the restriction of $\mathbf{C}_j$ to the first n symbols. The decoder also defines M stopping times:

$$\tau_j \stackrel{\triangle}{=} \inf\{n \geq 0 : S_{j,n} \geq \gamma)\}\,. \tag{6.86}$$

The final decision is made by the decoder at the stopping time τ^*:

$$\tau^* \stackrel{\triangle}{=} \min_{j=1,\ldots M} \tau_j\,. \tag{6.87}$$

I.e. τ^* is the moment of the first γ-upcrossing among all S_j. The output of the encoder is

$$g(Y^{\tau^*}) = \max\{j : \tau_j = \tau^*\}. \tag{6.88}$$

We are left with the problem of choosing $\mathbf{C}_j, j = 1, \dots M$.

This will be done by generating $\mathbf{C}_j$ randomly, independently of each other and distributed according to a distribution P_{X^∞} on $\mathcal{A}^\infty$.

We give an interpretation for our decoding scheme in the special case of a memoryless channel with $P_{X^\infty} = P_X^\infty$, i.e. X_k are independent and identically distributed with a single-letter distribution P_X. In this case, the decoder observes M random walks S_j one of which has a positive drift $I(X;Y)$ (the true message) and $(M-1)$ have negative drifts $-D(P_X P_Y || P_{XY})$, a quantity known as lautum information $L(X;Y)$, see [115]. The goal of the decoder, of course, is to detect the one with positive drift.

The average length of transmission satisfies:

$$\begin{aligned}
\mathbb{E}[\tau^*] &\leq \frac{1}{M}\sum_{j=1}^{M} \mathbb{E}[\tau_j | W = j] && (6.89)\\
&= \mathbb{E}[\tau_1 | W = 1] && (6.90)\\
&= \mathbb{E}[\tau], && (6.91)
\end{aligned}$$

where (6.90) is by symmetry and (6.91) is by the definition of τ in (6.77). Analogously, the average probability of error satisfies

$$\begin{aligned}
\mathbb{P}[g(Y^{\tau^*}) \neq W] &\leq \mathbb{P}[g(Y^{\tau^*}) \neq 1 | W = 1] && (6.92)\\
&\leq \mathbb{P}[\tau_1 \geq \tau^* | W = 1] && (6.93)\\
&\leq \mathbb{P}\left[\bigcup_{j=2}^{M} \{\tau_j \leq \tau_1\} \middle| W = 1\right] && (6.94)\\
&\leq (M-1)\mathbb{P}[\tau_2 \leq \tau_1 | W = 1], && (6.95)
\end{aligned}$$

where (6.92) is by (6.88), (6.94) is by the definition (6.87), and (6.95) is by a union bound and symmetry. Finally, notice that conditioned on $W = 1$ the joint distribution of $(S_{1,n}, S_{2,n}, \tau_1, \tau_2)$ is exactly the same as that of $(i(X^n;Y^n), i(\bar{X}^n;Y^n), \tau, \bar{\tau})$ defined in the formulation of the theorem and (6.77), thus we have proved (6.79) and (6.80).

To prove (6.81) simply notice that similarly to (6.91) we have almost surely:

$$\mathbb{E}[\tau^* | U] \leq \operatorname{esssup} \mathbb{E}[\tau | X], \tag{6.96}$$

and thus the bound (6.81) is automatically satisfied for every realization U. On the other hand, because of (6.95) there must exist a realization u_0 of U such that

$$\mathbb{P}[g(Y^{\tau^*}) \neq W | U = u_0] \leq (M-1)\mathbb{P}[\bar{\tau} \leq \tau], \tag{6.97}$$

which therefore defines a deterministic code with the sought-after performance (6.80) and (6.81). ■

The converse parts of Theorem 101 follow from the following result:

Theorem 103 *Consider an arbitrary DMC with capacity C. Then any (ℓ, M, ϵ) VLF code with $0 \le \epsilon < 1$ satisfies*

$$\log M \le \frac{C\ell + h(\epsilon)}{1-\epsilon}, \tag{6.98}$$

whereas each (ℓ, M, ϵ) VLFT code with $0 \le \epsilon < 1$ satisfies

$$\begin{aligned} \log M &\le \frac{C\ell + h(\epsilon) + (\ell+1)h\left(\frac{1}{\ell+1}\right)}{1-\epsilon} && (6.99)\\ &\le \frac{C\ell + \log(\ell+1) + h(\epsilon) + \log e}{1-\epsilon}, && (6.100)\end{aligned}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$ is the binary entropy function.

Proof: The inequality (6.98) is contained essentially in Lemmas 1 and 2 of [103]. Thus we focus on (6.99) only briefly mentioning how to obtain (6.98). First we give an informal argument. According to the Fano inequality

$$\begin{aligned} (1-\epsilon)\log M &\le I(W; Y^\tau, \tau) + h(\epsilon) && (6.101)\\ &= I(W; Y^\tau) + I(W; \tau | Y^\tau) + h(\epsilon) && (6.102)\\ &\le I(W; Y^\tau) + H(\tau) + h(\epsilon) && (6.103)\\ &\le I(W; Y^\tau) + (\ell+1)h\left(\frac{1}{\ell+1}\right) + h(\epsilon) && (6.104)\\ &\le C\ell + (\ell+1)h\left(\frac{1}{\ell+1}\right) + h(\epsilon), && (6.105)\end{aligned}$$

where in (6.104) we have upper-bounded $H(\tau)$ by solving a simple optimization problem for an integer valued non-negative random variable τ:

$$\max_{\tau: \mathbb{E}[\tau] \le \ell} H(\tau) = (\ell+1)h\left(\frac{1}{\ell+1}\right), \tag{6.106}$$

and in (6.105) we used the result of Burnashev [103]:

$$I(W; Y^\tau) \le C\,\mathbb{E}[\tau] \le C\ell. \tag{6.107}$$

Clearly (6.105) is equivalent to (6.99). The case of VLF codes is even simpler since τ is a function of Y^τ and thus $I(W; Y^\tau, \tau) = I(W; Y^\tau)$.

Unfortunately, the random variables (Y^τ, τ) and Y^τ are not well-defined and thus a different proof is required. Nevertheless, the main idea still pivots on the fact that because of the restriction on expectation, τ cannot convey more than $O(\log \ell)$ bits of information about the message.

Initially, we will assume that the code is deterministic and $|U| = 1$. Consider a triplet (f_n, g_n, τ) defining a given code. For a VLFT code, τ is a stopping moment of the filtration $\sigma\{W, Y^k\}_{k=0}^\infty$. To get rid of dependence of τ on W we introduce an extended channel

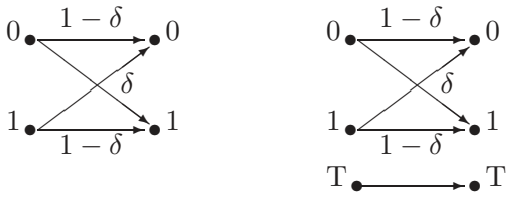


Figure 6.3: Illustration of the channel extension in the proof of Theorem 103.

$(\hat{\mathcal{A}}, \hat{\mathcal{B}}, P_{\hat{Y}|\hat{X}})$ as follows:

$$\hat{\mathcal{A}} = \mathcal{A} \cup \{T\}, \tag{6.108}$$

$$\hat{\mathcal{B}} = \mathcal{B} \cup \{T\}, \tag{6.109}$$

$$P_{\hat{Y}|\hat{X}}(\hat{y}|\hat{x}) = \begin{cases} P_{Y|X}(\hat{y}|\hat{x}), & \hat{x} \neq T, \\ 1\{\hat{y} = T\}, & \hat{x} = T. \end{cases} \tag{6.110}$$

In other words, the channel $P_{\hat{Y}|\hat{X}}$ has an additional input T conveyed noiselessly to the output. If $P_{Y|X}$ is a BSC with crossover probability δ then the extended channel has transition diagram as represented on Fig. 6.3. We also assume that the original and extended channels are defined on the same probability space where they are coupled in such a way that whenever $\hat{X} = X$ we have $\hat{Y} = Y$.

Next, we convert the given code (τ, f_n, g_n) to the code $(\hat{\tau}, \hat{f}_n, \hat{g}_n)$ for the extended channel as follows:

$$\hat{f}_n(W, \hat{Y}^{n-1}) = \begin{cases} f_n(W, \hat{Y}^{n-1}), & \tau \geq n, \\ T, & \tau < n, \end{cases} \tag{6.111}$$

$$\hat{\tau} = \tau + 1 = \inf\{n : \hat{Y}_n = T\}, \tag{6.112}$$

$$\hat{g}_n(\hat{Y}^n) = \begin{cases} g_n(\hat{Y}^n), & \hat{\tau} > n, \\ g_n(\hat{Y}^{\hat{\tau}-1}) & \hat{\tau} \leq n, \end{cases}. \tag{6.113}$$

Note that by definition $\tau \geq n$ can be decided by knowing W and Y^{n-1} only and hence $\hat{f}_n$ is indeed a function of $(W, \hat{Y}^{n-1})$; also notice that $\hat{Y}^{n-1} \in \mathcal{A}^{n-1}$ whenever $\tau \geq n$, and therefore the expression $f_n(W, \hat{Y}^{n-1})$ is meaningful.

Since $\hat{\tau}$ is a stopping time of the filtration

$$\mathcal{F}_n \stackrel{\triangle}{=} \sigma\{\hat{Y}^j\}_{j=1}^n \tag{6.114}$$

the triplet $(\hat{f}_n, \hat{g}_n, \hat{\tau})$ forms an $(\ell+1, M, \epsilon)$ VLF code for the extended channel (6.110). This code satisfies an additional constraint: input symbol T is used only once and it terminates the transmission. Now we prove that any such code must satisfy a certain upper bound on its cardinality M. To do so, consider the space $\{1, \ldots, M\} \times \hat{\mathcal{A}}^\infty$ and two measures on it: $P_{W\hat{Y}^\infty}$ and $P_W \times P_{\hat{Y}^\infty}$, where $P_{W\hat{Y}^\infty}$ is the joint distribution of random variables W and $\hat{Y}^\infty$ induced by the code $(\hat{f}_n, \hat{g}_n, \hat{\tau})$. Consider a measurable function

$$\phi : \{1, \ldots, M\} \times \hat{\mathcal{A}}^\infty \to \{0, 1\} \tag{6.115}$$

defined as

$$\phi = 1\{\hat{g}_{\hat{\tau}}(Y^{\hat{\tau}}) = W\}\,. \tag{6.116}$$

Notice that under measure $P_{W\hat{Y}^\infty}$ we have

$$P_{W\hat{Y}^\infty}[\phi = 1] \geq 1 - \epsilon\,, \tag{6.117}$$

due to the requirement (6.7). On the other hand, since under $P_W \times P_{\hat{Y}^\infty}$ $\hat{g}_{\hat{\tau}}$ is independent of W, we have

$$P_W \times P_{\hat{Y}^\infty}[\phi = 1] = \frac{1}{M}\,. \tag{6.118}$$

Therefore by the data-processing inequality we must have

$$D(P_{W\hat{Y}^\infty}||P_W P_{\hat{Y}^\infty}) \geq d(1-\epsilon||\tfrac{1}{M})\,, \tag{6.119}$$

where $d(x||y) = x\log\frac{x}{y} + (1-x)\log\frac{1-x}{1-y}$ is the binary relative entropy. After trivial manipulations in (6.119) we obtain

$$(1-\epsilon)\log M \leq I(W;\hat{Y}^\infty) + h(\epsilon)\,. \tag{6.120}$$

Although, (6.120) is just the Fano inequality, inclusion of the complete derivation illustrates the similarity with the meta-converse approach in Theorem 28 (see also Section 2.7.3). Another important observation is that for small ℓ, the bound can be tightened by replacing the step of data-processing (6.119) with an exact non-asymptotic solution of the Wald's sequential hypothesis testing problem.

We proceed to upper bound $I(W;\hat{Y}^\infty)$[3] . To do so we define a sequence of random variables:

$$Z_k = \log\frac{P_{\hat{Y}_k|W\hat{Y}^{k-1}}(\hat{Y}_k|W,\hat{Y}^{k-1})}{P_{\hat{Y}_k|\hat{Y}^{k-1}}(\hat{Y}_k|\hat{Y}^{k-1})}\,, \tag{6.121}$$

which are relevant to $I(W;\hat{Y}^\infty)$ because by simple telescoping we have

$$I(W;\hat{Y}^\infty) = \sum_{k=1}^{\infty}\mathbb{E}\left[Z_k\right]. \tag{6.122}$$

For Z_k we have the following property:

$$\mathbb{E}\left[Z_k|\mathcal{F}_{k-1}\right] = I(W;\hat{Y}_k|\mathcal{F}_{k-1})\,, \tag{6.123}$$

where $I(\cdot;\cdot|\mathcal{F})$ denotes mutual information, conditioned on $\mathcal{F}$ (i.e. it is an $\mathcal{F}$-measurable random variable). Specifically, for discrete random variables A and B we have

$$I(A;B|\mathcal{F}) \stackrel{\triangle}{=} \sum_{a,b}\mathbb{P}[A=a,B=b|\mathcal{F}]\log\frac{\mathbb{P}[A=a,B=b|\mathcal{F}]}{\mathbb{P}[A=a|\mathcal{F}]\mathbb{P}[B=b|\mathcal{F}]}\,, \tag{6.124}$$

where summation is over the alphabets of A and B. Similarly we can define $I(A;B|C,\mathcal{F})$ and other information measures.

[3]Notice that $\hat{Y}^\infty$ formalizes the idea of viewing (Y^τ,τ) as a random variable.

We define yet another process adapted to filtration $\mathcal{F}_n$, cf. (6.114),

$$V_n \stackrel{\triangle}{=} 1\{\hat{\tau} \le n\}\,. \tag{6.125}$$

With this notation we have:

$$\begin{aligned} I(W;\hat{Y}_k|\mathcal{F}_{k-1}) &= I(W;\hat{Y}_k V_k|\mathcal{F}_{k-1}) && (6.126)\\ &= I(W;V_k|\mathcal{F}_{k-1}) + I(W;\hat{Y}_k|V_k,\mathcal{F}_{k-1}) && (6.127)\\ &\le H(V_k|\mathcal{F}_{k-1}) + I(W;\hat{Y}_k|V_k,\mathcal{F}_{k-1}) && (6.128)\\ &\le H(V_k|\mathcal{F}_{k-1}) + I(\hat{X}_k;\hat{Y}_k|V_k,\mathcal{F}_{k-1})\,, && (6.129) \end{aligned}$$

where (6.126) follows because V_k is a function of $\hat{Y}_k$, (6.127) is the usual chain rule and (6.129) is obtained by applying the data-processing lemma to the Markov relation $W - \hat{X}_k - \hat{Y}_k - V_k$, which holds almost surely when conditioned on $\mathcal{F}_{k-1}$. We now upper-bound the second term in (6.129) as follows

$$I(\hat{X}_k;\hat{Y}_k|V_k,\mathcal{F}_{k-1}) \le 0\cdot\mathbb{P}[V_k=1|\mathcal{F}_{k-1}] + \mathbb{P}[V_k=0|\mathcal{F}_{k-1}]C\,, \tag{6.130}$$

because when $V_k = 1$ we must have $\hat{X}_k = \hat{Y}_k = T$ and the mutual information is zero, while when $V_k = 0$ we are computing the mutual information acquired on the $P_{\hat{Y}|\hat{X}}$ channel over a distribution $P_{\hat{X}_k|V_k\neq 0}$ which has a zero mass on the symbol T, and thus

$$\sup_{P_{\hat{X}}:P_{\hat{X}}(T)=0} I(\hat{X};\hat{Y}) = C\,. \tag{6.131}$$

Overall, from (6.123), (6.129) and (6.130) it follows:

$$\mathbb{E}\left[Z_k|\mathcal{F}_{k-1}\right] \le H(V_k|\mathcal{F}_{k-1}) + \mathbb{P}[V_k=0|\mathcal{F}_{k-1}]C\,. \tag{6.132}$$

Finally, we obtain

$$\begin{aligned} I(W;\hat{Y}^\infty) &= \sum_{k=1}^{\infty}\mathbb{E}\left[\mathbb{E}\left[Z_k|\mathcal{F}_{k-1}\right]\right] && (6.133)\\ &\le \sum_{k=1}^{\infty} H(V_k|\hat{Y}^{k-1}) + \mathbb{P}[V_k=0]C && (6.134)\\ &= \sum_{k=1}^{\infty} H(V_k|\hat{Y}^{k-1}) + C\,\mathbb{E}\left[\tau\right] && (6.135)\\ &\le \sum_{k=1}^{\infty} H(V_k|V^{k-1}) + C\,\mathbb{E}\left[\tau\right] && (6.136)\\ &= H(V_1,V_2,\ldots) + C\,\mathbb{E}\left[\tau\right] && (6.137)\\ &= H(\hat{\tau}) + C\,\mathbb{E}\left[\tau\right] && (6.138)\\ &= H(\tau) + C\,\mathbb{E}\left[\tau\right] && (6.139) \end{aligned}$$

where (6.133) follows from (6.122), (6.134) results from (6.132), (6.135) follows by taking an expectation of the obvious identity

$$\sum_{k=1}^{\infty} 1\{V_k = 0\} = \sum_{k=1}^{\infty} 1\{\hat{\tau} > k\} = \hat{\tau} - 1\,, \tag{6.140}$$

and recalling that $\hat{\tau} - 1 = \tau$, (6.136) follows because V^{k-1} is a function of $\hat{Y}^{k-1}$, (6.137) is obtained by the entropy chain rule, (6.139) follows since $(V_1, V_2, \ldots, V_n, \ldots)$ is an invertible function of $\hat{\tau}$, and finally (6.139) follows since $\hat{\tau} = \tau + 1$.

Together (6.120), (6.139) and (6.106) prove (6.99) in the case of a deterministic code with $|U| = 1$. For the case of $|U| > 1$ the above argument has shown that we have

$$(1 - \mathbb{P}[W \neq \hat{W}|U]) \log M \leq C\,\mathbb{E}\,[\tau|U] + H(\tau|\sigma U) + h(\mathbb{P}[W \neq \hat{W}|U])\,, \quad \text{a.s.}, \tag{6.141}$$

where $\hat{W} = g_\tau(Y^\tau)$ is the output message estimate of the decoder. By taking the expectation of both sides of (6.141) and applying the Jensen's inequality to the binary entropy terms we obtain

$$(1 - \mathbb{P}[W \neq \hat{W}|U]) \log M \leq C\,\mathbb{E}\,[\tau] + H(\tau|U) + h(\epsilon)\,, \tag{6.142}$$

and then (6.99) follows since by (6.106) we have

$$H(\tau|U) \leq H(\tau) \leq (\ell + 1) h\left(\frac{1}{\ell + 1}\right)\,. \tag{6.143}$$

Notice that in the case of VLF codes, the first term in (6.135) disappears because V_k is a function of $\hat{Y}^{k-1}$ thus leading to the tighter bound (6.98). ■

Proof of Theorem 101: The upper bounds in (6.70) and (6.71) follow from Theorem 103. For the lower bound (6.70), suppose that for each ℓ' there exists an $\left(\ell', M, \frac{1}{\ell'}\right)$-VLF code with

$$\log M = C\ell' - \log \ell' - a_0\,, \tag{6.144}$$

where a_0 is some constant. To see that (6.144) implies the lower bound in (6.70) consider the code which terminates without any channel uses, i.e. $\tau = 0$, with probability $\frac{\ell'\epsilon - 1}{\ell' - 1}$ and uses the $\left(\ell', M, \frac{1}{\ell'}\right)$-VLF code otherwise[4] . Such a code has probability of error ϵ and average length $\ell = \frac{\ell'^2(1-\epsilon)}{\ell' - 1}$ and, therefore, using (6.144) we have

$$\begin{aligned} \log M^*(\ell, \epsilon) &\geq C\ell' - \log \ell' - a_0 && (6.145)\\ &= \frac{\ell C}{1 - \epsilon} - \log \ell + O(1)\,, && (6.146) \end{aligned}$$

as required.

To prove (6.144), we apply Theorem 102 with the process $\{X_n\}_{n=1}^{\infty}$ chosen to be independent and identically distributed (i.i.d.) with a marginal distribution P_X – a capacity achieving distribution. To analyze (6.80) it is convenient to define a pair of random walks

$$\begin{aligned} S_n &\stackrel{\triangle}{=} i(X^n; Y^n)\,, && (6.147)\\ \bar{S}_n &\stackrel{\triangle}{=} i(\bar{X}^n; Y^n)\,. && (6.148) \end{aligned}$$

[4]Note that due to availability of a decision feedback such a randomization can be realized on the decoder side only, i.e. without requiring any common randomness, U. Thus if $\left(\ell', M, \frac{1}{\ell'}\right)$-VLF code exists with $|U| = 1$ then the overall coding scheme constructed to achieve (6.70) also has $|U| = 1$.

First notice that for any (measurable) function f we have

$$\mathbb{E}\left[f(\bar{X}^n, Y^n)\right] = \mathbb{E}\left[f(X^n, Y^n)\exp\{-S_n\}\right], \tag{6.149}$$

because $S_n = \log\frac{dP_{X^nY^n}}{dP_{\bar{X}^nY^n}}$. Therefore, we have

$$\begin{aligned}
\mathbb{P}[\bar{\tau} \le \tau] &\le \mathbb{P}[\bar{\tau} < \infty] && (6.150)\\
&= \mathbb{E}\left[\exp\{-S_\tau\}1\{S_\tau \ge \gamma_1\}\right. && (6.151)\\
&\le \exp\{-\gamma\}, && (6.152)
\end{aligned}$$

where (6.150) is because $\tau < \infty$ almost surely, and (6.152) is by the definition of τ in (6.77). Since the sequence $S_n - nI(X;Y) = S_n - nC$ is a martingale we obtain from Wald's identity

$$\begin{aligned}
C\,\mathbb{E}[\tau] &= \mathbb{E}[S_\tau] && (6.153)\\
&\le \gamma + a_0, && (6.154)
\end{aligned}$$

where a_0 is an upper-bound on S_1. The existence of an $\left(\ell', M, \frac{1}{\ell'}\right)$-VLF code with M satisfying (6.144) now follows by taking $\gamma = C\ell' - a_0$ and using (6.154) and (6.152) in (6.79) and (6.80), respectively. ■

We note in passing that while the codes with encoders utilizing full noiseless feedback can achieve the Burnashev exponent (6.2), it was noted in [109, 111] that the lower error exponent

$$E_1(R) = C - R \tag{6.155}$$

is achievable at all rates $R < C$ with decision feedback codes (6.11). Indeed, this property easily follows from (6.152) and (6.154).

A numerical comparison of the upper and lower bounds for the BSC with crossover probability $\delta = 0.11$ and $\epsilon = 10^{-3}$ is given in Fig. 6.4, where the upper bound is (6.98) and the lower bound is Theorem 102 (evaluated with various γ depending on the average blocklength). Note that for $BSC(\delta)$ the $i(X^n;Y^n)$ becomes a random walk taking steps $\log 2\delta$ and $\log(2-2\delta)$ with probabilities δ and $1-\delta$, i.e.,

$$i(X^n;Y^n) = n\log(2-2\delta) + \log\frac{\delta}{1-\delta}\sum_{k=1}^{n} Z_k, \tag{6.156}$$

where Z_k are independent Bernoulli $\mathbb{P}[Z_k = 1] = 1 - \mathbb{P}[Z_k = 0] = \delta$. The evaluation of (6.80) is simplified by using (6.149) to get rid of the process $i(\bar{X}^n;Y^n)$, which in this case is independent of (X^n, Y^n):

$$\epsilon \le (M-1)\mathbb{E}\left[f(\tau)\right], \tag{6.157}$$

where

$$f(n) \stackrel{\triangle}{=} \mathbb{E}\left[1\{\tau \le n\}\exp\{-i(X^\tau;Y^\tau)\}\right]. \tag{6.158}$$

The dashed line in Fig. 6.4 is the approximate fundamental limit for fixed blocklength codes without feedback given by the equation (6.1) with $O(\log n)$ substituted by $\frac{1}{2}\log n$; see Theorem 41.

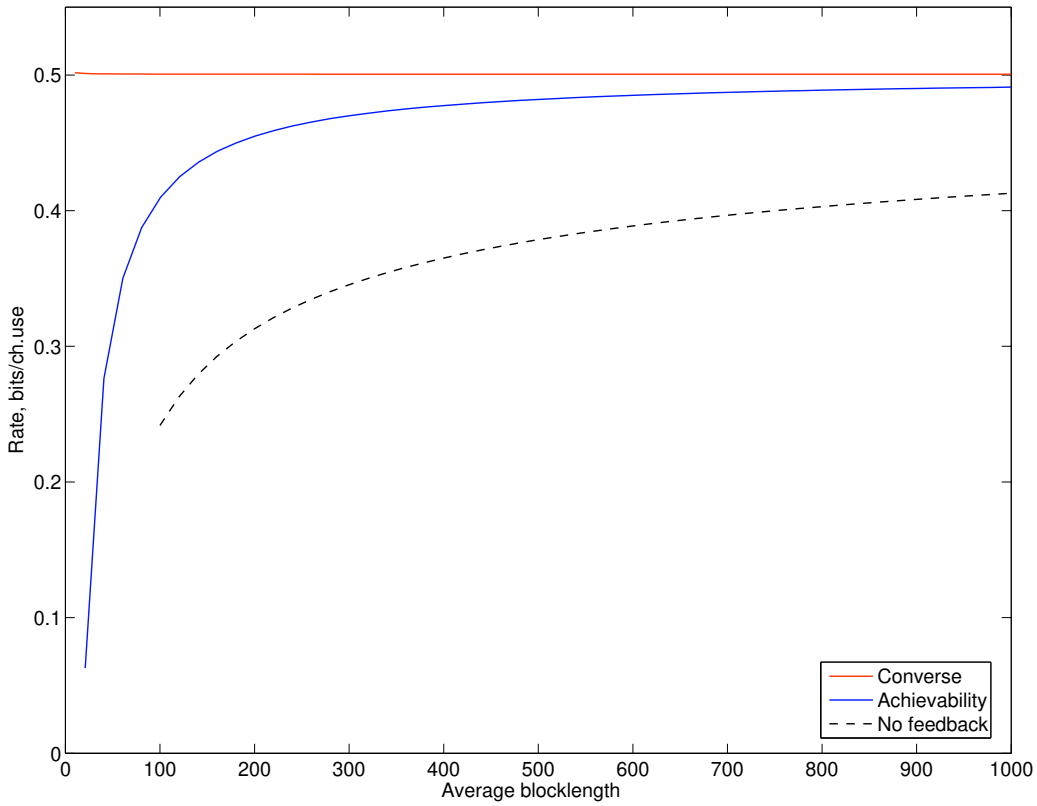


Figure 6.4: Comparison of upper and lower bounds for the BSC(0.11) with variable-length and feedback; probability of error $\epsilon = 10^{-3}$.

Theorem 104 *For a $BEC(\delta)$ and $\epsilon \in [0,1)$ we have*

$$\log_2 M_f^*(\ell,\epsilon) = \frac{\ell C}{1-\epsilon} + O(1)\,, \tag{6.159}$$

where $C = 1-\delta$ bit. More precisely,

$$\left\lfloor \frac{\ell C}{1-\epsilon} \right\rfloor \le \log_2 M_f^*(\ell,\epsilon) \le \frac{\ell C}{1-\epsilon} + \frac{h(\epsilon)}{1-\epsilon}\,. \tag{6.160}$$

Proof: The upper bound in Theorem 101 holds even for $\epsilon = 0$, so we need only to prove a lower bound. First, we assume $\epsilon = 0$ and take arbitrary k. Consider the strategy that simply retransmits each of k bits until it gets through the channel unerased. More formally, we define a stopping time as

$$\tau_0 = \inf\{n \ge 1 :\ \text{there are } k \text{ unerased symbols in } Y_1, \dots Y_n\}\,. \tag{6.161}$$

It is easy to see that

$$\mathbb{E}\left[\tau_0\right] = \frac{k}{1-\delta}\,. \tag{6.162}$$

Hence for any ℓ we have shown

$$\log_2 M_f^*(\ell, 0) \geq \lfloor \ell C \rfloor \,. \tag{6.163}$$

For $\epsilon > 0$ we make use of the randomization to construct a transmission scheme that stops at time 0 with probability ϵ and otherwise proceeds as above. We define a stopping time

$$\tau_\epsilon = \tau_0 1\{U \geq \epsilon\} \,, \tag{6.164}$$

where U is uniform on $[0, 1]$ and measurable with respect to $\mathcal{G}_0$. It is clear that using such a strategy we obtain a probability of error upper-bounded by ϵ and

$$\mathbb{E}\left[\tau_\epsilon\right] = \frac{k}{1-\delta}(1-\epsilon) \,. \tag{6.165}$$

Hence we are able to achieve

$$\log_2 M_f^*(\ell, \epsilon) \geq \left\lfloor \frac{\ell C}{1-\epsilon} \right\rfloor \,. \tag{6.166}$$

■

The result of Theorem 104 suggests that to improve the expansion (6.70) to the order $O(1)$, it is likely that we need to go beyond encoders satisfying (6.11).

6.8 Zero-error communication

The general achievability bound, Theorem 102, applies only to $\epsilon > 0$. What can be said about $\epsilon = 0$?

6.8.1 Without a termination symbol (VLF codes)

Recall that C_1 in (6.2) is defined as

$$C_1 = \max_{a_1, a_2 \in \mathcal{A}} D(P_{Y|X=a_1} || P_{Y|X=a_2}) \,. \tag{6.167}$$

Burnashev [103] showed that if $C_1 = \infty$, then as $\ell \to \infty$ we have for some $a > 0$

$$\log M_f^*(\ell, 0) \geq C\ell - a\sqrt{\ell \log \ell} + O(\log \ell) \,. \tag{6.168}$$

For this reason, for such channels zero-error VLF capacity is equal to the conventional capacity. However, the penalty bound $\sqrt{\ell \log \ell}$ is rather loose, as the following result demonstrates.

Theorem 105 *For a BEC(δ) with capacity $C = 1 - \delta$ bit we have*

$$\log_2 M_f^*(\ell, 0) = \ell C + O(1) \,. \tag{6.169}$$

Proof: Theorem 104 applied with $\epsilon = 0$. ■

Regarding any channel with $C_1 < \infty$ (e.g. the BSC), the following negative result holds:

Theorem 106 *For any DMC with $C_1 < \infty$ we have*

$$\log M_f^*(\ell, 0) = 0 \tag{6.170}$$

for all $\ell \geq 0$.

Proof: We show that when $C_1 < \infty$ no $(\ell, 2, 0)$ VLF code exists. Indeed, assume that (U, f_n, g_n, τ) is such a code. For zero-error codes, randomization does not help and hence, without loss of generality we assume $|\mathcal{U}| = 1$. Then, conditioning on $W = 1$ and $W = 2$ gives two measures P_1 and P_2 on B, which are mutually singular when considered on the σ-algebra $\mathcal{G}_\tau$, where $\mathcal{G}_n = \sigma\{Y_1, \ldots, Y_n\}$ is a filtration on B, with respect to which τ is a stopping time. Define a process, adapted to the same filtration:

$$R_n = \log \left. \frac{dP_1}{dP_2} \right|_{\mathcal{G}_n}, \tag{6.171}$$

where $\left. \frac{dP_1}{dP_2} \right|_{\mathcal{G}_n}$ denotes the Radon-Nikodym derivative between P_1 and P_2 considered as measures on the space B with σ-algebra $\mathcal{G}_n$. Then, by memorylessness we have

$$R_n - R_{n-1} = \log \frac{P_{Y|X}(Y_n | f_n(1, Y^{n-1}))}{P_{Y|X}(Y_n | f_n(2, Y^{n-1}))}. \tag{6.172}$$

From (6.172) and $C_1 < \infty$ it follows that there exists a constant $a_1 > 0$ such that

$$R_n - R_{n-1} \geq -a_1, \tag{6.173}$$

and, consequently,

$$R_n \geq -n a_1. \tag{6.174}$$

On the other hand, taking the conditional expectation of (6.172) with respect to P_1 we obtain from the definition of C_1 in (6.167):

$$\mathbb{E}[R_n | \mathcal{G}_{n-1}] \leq R_{n-1} + C_1 < \infty, \tag{6.175}$$

where here and in the remainder of this proof the expectation $\mathbb{E}$ is taken with respect to measure P_1. Thus (6.175) implies that under P_1 the process $R_n - nC_1$ is a supermartingale. By Doob's stopping theorem for any integer $k \geq 0$ we have then

$$\mathbb{E}[R_{\min\{\tau,k\}}] \leq C_1 \mathbb{E}[\min\{\tau, k\}] \leq C_1 \mathbb{E}[\tau] < \infty. \tag{6.176}$$

At the same time from (6.174) we have

$$R_{\min\{\tau,k\}} \geq -a_1 \min\{\tau, k\} \geq -a_1 \tau, \tag{6.177}$$

and since $\mathbb{E}[\tau] < \infty$ we can apply Fatou's lemma to (6.176) to obtain

$$\mathbb{E}[R_\tau] = \mathbb{E}[\liminf_{k \to \infty} R_{\min\{\tau,k\}}] \leq C_1 \mathbb{E}[\tau] < \infty. \tag{6.178}$$

On the other hand,

$$D(P_1|_{\mathcal{G}_\tau} || P_2|_{\mathcal{G}_\tau}) = \mathbb{E}[R_\tau] < \infty, \tag{6.179}$$

thus implying that P_1 and P_2 cannot be mutually singular on $\mathcal{G}_\tau$ – a contradiction. ■

6.8.2 With a termination symbol (VLFT codes)

The shortcoming of VLF coding found in Theorem 106 is overcome in the paradigm of VLFT coding. Our main tool is the following achievability bound.

Theorem 107 *Fix an arbitrary channel* $\{P_{Y_i|X_1^iY_1^{i-1}}\}_{i=1}^{\infty}$ *and a process* $X = (X_1, X_2, \ldots, X_n, \ldots)$ *with values in* $\mathcal{A}$. *Then for every positive integer* M *there exists an* $(\ell, M, 0)$ *VLFT code with*

$$\ell \le \sum_{n=0}^{\infty} \mathbb{E}\left[\min\left\{1, (M-1)\mathbb{P}[i(X^n;Y^n) \le i(\bar{X}^n;Y^n)|X^nY^n]\right\}\right], \tag{6.180}$$

where X^n, $\bar{X}^n$, Y^n *and* $i(\cdot;\cdot)$ *are defined in* (6.75) *and* (6.76). *Moreover, this is an FV code which is deterministic and uses feedback only to compute the stopping time, i.e.* (6.11) *holds.*

Proof: To construct a deterministic code we need to define a triplet (f_n, g_n, τ). Consider a collection of M infinite $\mathcal{A}$-strings $\{\mathbf{C}_1, \ldots, \mathbf{C}_M\}$. The sequence of the encoder functions is defined as

$$f_n(w) = (\mathbf{C}_w)_n\,, \tag{6.181}$$

where $(\mathbf{C}_j)_n$ is the n-th coordinate of the vector $\mathbf{C}_j$. Recall that in the paradigm of VLFT codes it is allowable for the stopping rule τ to depend on the true message W, so we may define

$$\tau = \inf\{n \ge 0 : i(\mathbf{C}_W(n);Y^n) > \max_{u\neq W} i(\mathbf{C}_u(n);Y^n)\}\,, \tag{6.182}$$

where as before $\mathbf{C}_j(n) \in \mathcal{A}^n$ is a restriction of $\mathbf{C}_j$ to the first n coordinates. Definition (6.182) means that if the true message is j then the transmitter stops at the first time instant n when $i(\mathbf{C}_j(n);Y^n)$ is strictly larger than any other $i(\mathbf{C}_u(n);Y^n), u \neq j)$. Finally, the sequence of decoder functions is defined as

$$g_n(y^n) = \begin{cases} k, & \text{if } \forall j \neq k : i(\mathbf{C}_k(n);y^n) > i(\mathbf{C}_j(n);y^n) \\ 1, & \text{otherwise}\,. \end{cases} \tag{6.183}$$

Upon receiving a stop signal, the decoder outputs the index of the unique message corresponding to the maximal information density, thus we have

$$g_\tau(Y^\tau) = W\,, \tag{6.184}$$

and the constructed code is indeed a zero-error VLFT code for any selection of M strings $\mathbf{C}_j, j = 1, \ldots, M$. We need to provide an estimate only of the expected length of communication $\mathbb{E}[\tau]$.

The result is proved by applying a random coding argument with each $\mathbf{C}_j$ generated independently with probability distribution P_{X^∞}, corresponding to the fixed input process X. Averaging over all realizations of $\{\mathbf{C}_j, j = 1, \ldots, M\}$ we obtain the following estimate:

$$\begin{aligned} \mathbb{P}[\tau > n] &= \mathbb{P}[\tau > n|W = 1] && (6.185)\\ &\le \mathbb{P}\left[\bigcup_{j=2}^{M}\{i(\mathbf{C}_1(n);Y^n) \le i(\mathbf{C}_j(n);Y^n)\}\middle| W = 1\right], && (6.186) \end{aligned}$$

where (6.185) follows from symmetry and (6.186) simply states that if $\tau > n$ and $W = 1$ then at least one information density should not be smaller than $i(\mathbf{C}_1(n); Y^n)$. We can proceed from (6.186) as in the RCU bound, Theorem 17:

$$\mathbb{P}[\tau > n] \le \mathbb{E}\left[\min\left\{1, (M-1)\mathbb{P}[i(X^n;Y^n) \le i(\bar{X}^n;Y^n)|X^nY^n]\right\}\right], \tag{6.187}$$

where we have additionally noted that conditioned on $W = 1$ the joint distribution of $(\mathbf{C}_1(n), \mathbf{C}_j(n), Y^n)$ coincides with that of $(X^n, \bar{X}^n, Y^n)$ for every $j \neq 1$. Summing (6.187) over all n from 0 to ∞ we obtain

$$\mathbb{E}[\tau] = \sum_{n=0}^{\infty} \mathbb{P}[\tau > n] \le \mathbb{E}\left[\min\left\{1, (M-1)\mathbb{P}[i(X^n;Y^n) \le i(\bar{X}^n;Y^n)|X^nY^n]\right\}\right]. \tag{6.188}$$

Thus, there must exist a realization of $\{\mathbf{C}_j, j = 1, \ldots, M\}$ achieving (6.180). ■

Theorem 108 *For an arbitrary DMC we have*

$$\log M_{\mathsf{t}}^*(\ell, 0) = \ell C + O(\log \ell). \tag{6.189}$$

More specifically we have

$$\log M_{\mathsf{t}}^*(\ell, 0) \le \ell C + \log \ell + O(1), \tag{6.190}$$
$$\log M_{\mathsf{t}}^*(\ell, 0) \ge \ell C + O(1). \tag{6.191}$$

Furthermore, the encoder achieving (6.191) *uses feedback to calculate the stopping time only, i.e. it is an FV code.*

Proof: The upper bound (6.190) follows from (6.100). To prove a lower bound, we will apply Theorem 107 with the process X selected as i.i.d. with a capacity-achieving marginal distribution. We first weaken the bound (6.180) to a form that is easier to analyze:

$$\begin{aligned}
&\mathbb{E}\left[\min\left\{1, (M-1)\mathbb{P}[i(X^n;Y^n) \le i(\bar{X}^n;Y^n)|X^nY^n]\right\}\right] && (6.192)\\
&\quad\le \mathbb{E}\left[\min\left\{1, M\mathbb{P}[i(X^n;Y^n) \le i(\bar{X}^n;Y^n)|X^nY^n]\right\}\right] && (6.193)\\
&\quad= \mathbb{E}\left[\min\left\{1, M\mathbb{P}[i(X^n;Y^n) \le i(\bar{X}^n;Y^n)|X^nY^n]\right\} 1\{i(X^n;Y^n) \le \log M\}\right]\\
&\qquad\quad + \mathbb{E}\left[\min\left\{1, M\mathbb{P}[i(X^n;Y^n) \le i(\bar{X}^n;Y^n)|X^nY^n]\right\} 1\{i(X^n;Y^n) > \log M\}\right] && (6.194)\\
&\quad\le \mathbb{P}[i(X^n;Y^n) \le \log M] + M\mathbb{P}[i(\bar{X}^n;Y^n) > \log M] && (6.195)\\
&\quad= \mathbb{E}\left[\exp\left\{-[i(X^n;Y^n) - \log M]^+\right\}\right], && (6.196)
\end{aligned}$$

where (6.195) is obtained from (6.194) by upper-bounding min by 1 in the first term and by $M\mathbb{P}[i(\bar{X}^n;Y^n) > \log M]$ in the second term, and (6.196) is an application of (6.149).

In view of (6.196), Theorem 107 guarantees the existence of an $(\ell, M, 0)$ VLFT code with[5]

$$\ell \le \mathbb{E}\left[\sum_{n=0}^{\infty} \exp\left\{-[i(X^n;Y^n) - \log M]^+\right\}\right]. \tag{6.197}$$

[5] $i(X^0;Y^0) = 0$ by convention.

We now define the filtration $\mathcal{F}$ as

$$\mathcal{F}_n = \sigma\{X^n, \bar{X}^n, Y^n\}, \quad n = 0, 1, \dots \tag{6.198}$$

Notice that $i(X^n; Y^n)$ is a random walk adapted to $\mathcal{F}$ with bounded jumps and positive drift equal to the capacity C:

$$\mathbb{E}\left[i(X^n; Y^n)\right] = nC\,, \tag{6.199}$$

whereas the process $i(\bar{X}^n; Y^n)$ is also a random walk with bounded jumps, but with a negative drift equal to the lautum information [115]:

$$\mathbb{E}\left[i(\bar{X}^n; Y^n)\right] = -nD(P_X P_Y || P_{XY}) = -nL(X; Y)\,. \tag{6.200}$$

Define a stopping time of the filtration $\mathcal{F}$ as follows:

$$\tau = \inf\{n \geq 0 : i(X^n; Y^n) \geq \log M\}\,. \tag{6.201}$$

With this definition we have

$$\mathbb{E}\left[\sum_{n=0}^{\infty} \exp\left\{-[i(X^n; Y^n) - \log M]^+\right\}\right] = \mathbb{E}\left[\tau + \sum_{k=0}^{\infty} \exp\left\{-[i(X^{k+\tau}; Y^{k+\tau}) - \log M]^+\right\}\right]. \tag{6.202}$$

Because $i(X^\tau; Y^\tau) \geq \log M$ we have

$$\begin{aligned}
[i(X^{k+\tau}; Y^{k+\tau}) - \log M]^+ &= [i(X^{k+\tau}; Y^{k+\tau}) - i(X^\tau; Y^\tau) + i(X^\tau; Y^\tau) - \log M]^+ && (6.203)\\
&\geq [i(X^{k+\tau}; Y^{k+\tau}) - i(X^\tau; Y^\tau)]^+\,. && (6.204)
\end{aligned}$$

Application of (6.204) gives

$$\mathbb{E}\left[\sum_{k=0}^{\infty} \exp\left\{-[i(X^{k+\tau}; Y^{k+\tau}) - \log M]^+\right\}\right] \leq \mathbb{E}\left[\sum_{k=0}^{\infty} \exp\left\{-[i(X^{k+\tau}; Y^{k+\tau}) - i(X^\tau; Y^\tau)]^+\right\}\right]. \tag{6.205}$$

By the strong Markov property of the random walk, conditioned on $\mathcal{F}_\tau$ the distribution of the process $i(X^{k+\tau}; Y^{k+\tau}) - i(X^\tau; Y^\tau)$ is the same as that of the process $i(X^k; Y^k)$. Thus, (6.202) and (6.205) imply

$$\mathbb{E}\left[\sum_{n=0}^{\infty} \exp\left\{-[i(X^n; Y^n) - \log M]^+\right\}\right] \leq \mathbb{E}\left[\tau\right] + \mathbb{E}\left[\sum_{k=0}^{\infty} \exp\left\{-[i(X^k; Y^k)]^+\right\}\right]. \tag{6.206}$$

To estimate the second term, notice that for some constants $a_1, a_2 > 0$ we have

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{-[i(X^k; Y^k)]^+\right\}\right] && (6.207)\\
&= \mathbb{P}[i(X^k; Y^k) \leq 0] + \mathbb{E}\left[\exp\left\{-i(X^k; Y^k)\right\} 1\{i(X^k; Y^k > 0\}\right] && (6.208)\\
&= \mathbb{P}[i(X^k; Y^k) \leq 0] + \mathbb{P}[i(\bar{X}^k; Y^k) > 0] && (6.209)\\
&\leq a_2 \exp\{-a_1 k\}\,, && (6.210)
\end{aligned}$$

where (6.209) is an application of (6.149), and (6.210) follows from Chernoff bound since both $i(X^k;Y^k)$ and $i(\bar{X}^k;Y^k)$ are sums of k i.i.d. random variables with positive expectation C and negative expectation $L(X;Y)$, respectively. Summing (6.210) over all non-negative integers k we obtain that for some constant $a_3 > 0$ we have

$$\mathbb{E}\left[\sum_{k=0}^{\infty} \exp\left\{-[i(X^k;Y^k)]^+\right\}\right] \le a_3\,. \tag{6.211}$$

Finally, by the boundedness of jumps of $i(X^n;Y^n)$ there is a constant $a_4 > 0$ such that

$$i(X^\tau;Y^\tau) - \log M \le a_4\,. \tag{6.212}$$

Since $i(X^n;Y^n) - nC$ is a martingale with bounded increments we have from Doob's stopping theorem:

$$\mathbb{E}\left[i(X^\tau;Y^\tau)\right] = C\,\mathbb{E}\left[\tau\right], \tag{6.213}$$

but on the other hand from (6.212) we have

$$\mathbb{E}\left[i(X^\tau;Y^\tau)\right] \le \log M + a_4 \tag{6.214}$$

and thus

$$\mathbb{E}\left[\tau\right] \le \frac{\log M}{C} + a_4\,. \tag{6.215}$$

Together (6.215), (6.211) imply via (6.206) and (6.197) the required lower bound (6.191). ■

Theorem 108 suggests that VLFT codes may achieve capacity even at very short blocklengths. To illustrate this numerically we first notice that Theorem 107 particularized to the BSC with i.i.d. input process X and an equiprobable marginal distribution yields the following result:[6]

Corollary 109 *For the BSC with crossover probability δ and for every positive integer M there exists an $(\ell, M, 0)$ VLFT code satisfying*

$$\ell \le \sum_{n=0}^{\infty}\sum_{t=0}^{n}\binom{n}{t}\delta^t(1-\delta)^{n-t}\min\left\{1,\ M\sum_{k=0}^{t}\binom{n}{k}2^{-n}\right\}. \tag{6.216}$$

A comparison of (6.216) and the upper bound (6.100) is given in Fig. 6.5. We see that despite the requirement of zero probability of error, VLFT codes attain the capacity of the BSC at blocklengths as short as 30. As in Theorem 104 the convergence to capacity is very fast. Additionally, we have depicted the (approximate) performance of the best non-feedback code paired with the simple ARQ strategy, see Section 6.4. The comparison on Fig. 6.5 suggests that even having access to the best possible block codes the ARQ is considerably suboptimal. It is interesting to note in this regard, that a Yamamoto-Itoh [105] strategy also pairs the best block code with a noisy version of ARQ (therefore, it is a VLF achievability bound). Consequently, we expect a similar gap in performance.

[6] This expression is to be compared with the (almost) optimal non-feedback achievability bound for the BSC, see (3.10).

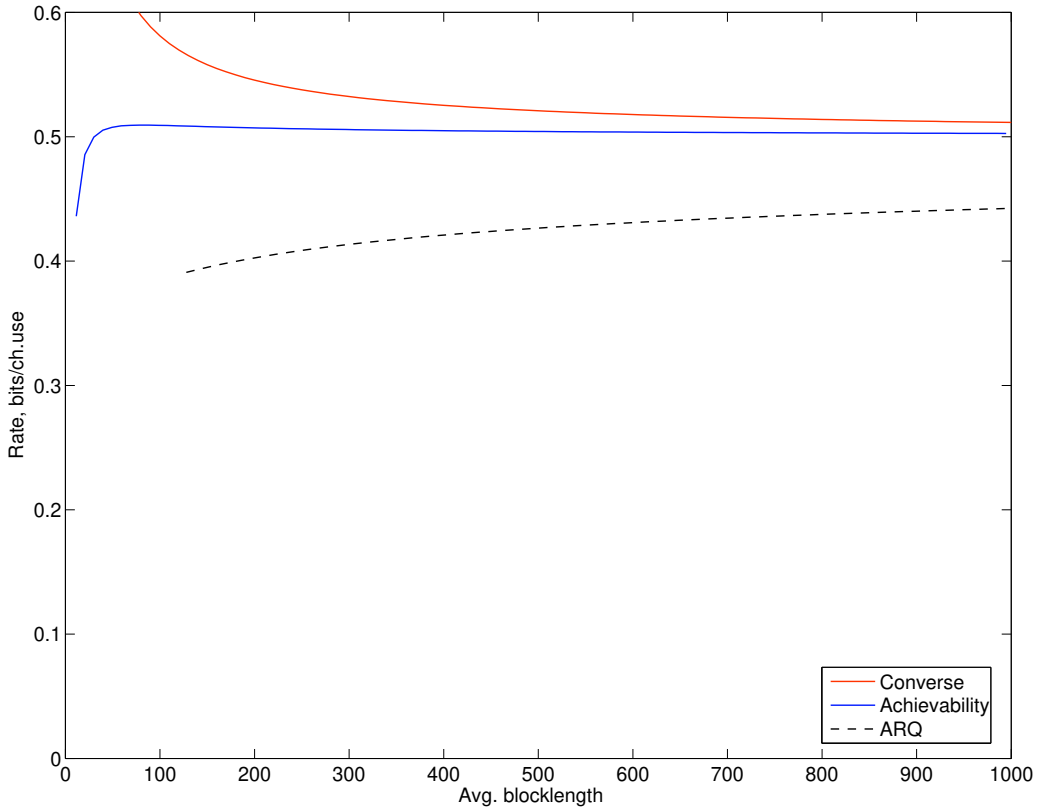


Figure 6.5: Zero-error communication over the BSC(0.11) with a termination symbol. The lower bound is (6.216); the upper-bound is (6.100).

Another property of VLFT codes is that the maximal achievable rate for very small blocklengths may be noticeably above capacity. This can be seen as an artifact of the model which provides for an error-free termination symbol. Ordinarily, the overhead required in a higher layer to provide much higher reliability than the individual physical-layer symbols would not make short blocklengths attractive.

This point is best demonstrated by computing the following specialized achievability bound for the BEC, which improves the general Theorem 107 in this particular case.

Theorem 110 *For the BEC with erasure probability δ and any positive integer M there exists an $(\mu(M), M, 0)$ VLFT code, where function $\mu : \mathbb{Z}_+ \to \mathbb{R}_+$ is the solution to*

$$\begin{aligned}\mu(M) &= \frac{1}{M}\cdot 0 + \frac{M-1}{M}\cdot 1 \\ &\quad + (1-\delta)\cdot\frac{1}{M}\left[\left\lceil\frac{M-1}{2}\right\rceil \mu\left(\left\lceil\frac{M-1}{2}\right\rceil\right) + \left\lfloor\frac{M-1}{2}\right\rfloor \mu\left(\left\lfloor\frac{M-1}{2}\right\rfloor\right)\right] \\ &\quad + \delta\cdot\frac{1}{M}(M-1)\mu(M-1)\,, \end{aligned} \tag{6.217}$$

initialized by $\mu(1) = 0$.

Proof: If we need to transmit only one message, $M = 1$, then we can simply set $\tau = 0$. Therefore, we have

$$\mu(1) = 0\,. \tag{6.218}$$

If we need to transmit an arbitrary $M > 1$ number of messages than we do the following. First, all M messages are split into three groups. The first group consists of a single message and the remaining $M - 1$ messages are split according to

$$M - 1 = \left\lceil \frac{M-1}{2} \right\rceil + \left\lfloor \frac{M-1}{2} \right\rfloor . \tag{6.219}$$

Second, if W is equal to the special message, then the encoder terminates the communication by setting $\tau = 0$. If W belongs to one of $\left\lceil \frac{M-1}{2} \right\rceil$ messages the the encoder sets $f_1 = 0$, and $f_1 = 1$, otherwise. Third, upon passing through the channel one of the possibilities can happen: the digit was erased or was delivered correctly. In the case of correct delivery we reiterate the algorithm with either $M' = \left\lceil \frac{M-1}{2} \right\rceil$ or $M' = \left\lfloor \frac{M-1}{2} \right\rfloor$, depending on the group W belonged to. In the case of erasure we reiterate with $M' = M - 1$ since the special message was ruled out. A careful analysis of the probabilities of each case yields the recursion (6.217). ■

The first few values of the μ-function are

$$\mu(1) = 0\,, \tag{6.220}$$

$$\mu(2) = 1/2\,, \tag{6.221}$$

$$\mu(3) = \frac{1}{3}(2+\delta)\,, \tag{6.222}$$

$$\mu(4) = 1 + \frac{1}{4}(\delta + \delta^2)\,. \tag{6.223}$$

Numerical comparison of the achievability bound[7] of Theorem 110 against the converse bound (6.100) is given on Fig. 6.6 for the case of $\delta = 0.5$. We notice that indeed for small ℓ (and, equivalently, M) the availability of the termination symbol allows the rate to exceed the capacity slightly. Also, the horizontal capacity line coincides with the "traditional" achievability bound for the BEC, as given by Theorem 104 with $\epsilon = 0$, which does not take advantage of the additional degree of freedom enabled in the VLFT paradigm (i.e., a termination symbol).

6.9 Excess delay constraints

Quantifying the notion of delay for variable-length coding with feedback has proven to be notoriously hard, see, for example, [116] for a related discussion. While for fixed-blocklength

[7]Since it is not possible to compute $\mu(2^{500})$ directly, the following idea was used. Fix some k_{max} and compute $\mu(2^k)$ for all $k \le k_{max}$ via (6.217). For $k > k_{max}$ we can use a strategy of simply retransmitting each of the first $k - k_{max}$ bits until it is delivered unerased, and then use the recursive strategy to transmit the remaining k_{max} bits. This gives the bound

$$\mu(2^k) \le \frac{k - k_{max}}{1-\delta} + \mu\left(2^{k_{max}}\right)\,. \tag{6.224}$$

As k_{max} increases, the upper bound improves. Experimentation shows that there is no visible improvement once $k_{max} \gtrsim 10$.

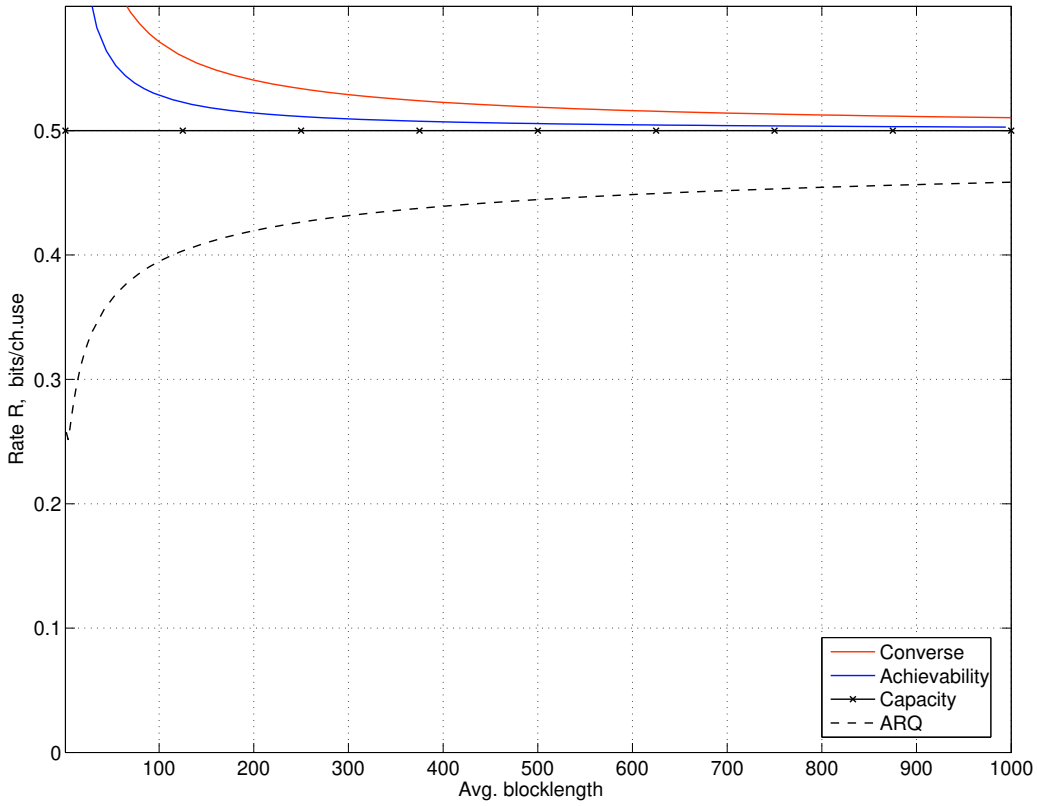


Figure 6.6: Zero-error communication over the BEC(0.5) with a termination symbol.

codes delay is naturally associated with the blocklength, in the variable-length setup, however, the usage of average blocklength $\mathbb{E}[\tau]$ as a proxy for delay is not appropriate in real-time applications with hard delay constraints. On the contrary, the definition of rate as $\frac{\log M}{\mathbb{E}[\tau]}$ is very natural, since by the law of large numbers the ratio of bits to channel uses will approach $\frac{\log M}{\mathbb{E}[\tau]}$ for a repeated usage of the same code.

An advantage of feedback is the ability to terminate transmission early on favorable noise realizations thereby reducing average blocklength. However, it remains to be seen whether under a constraint on the probability of excess delay, variable-length coding offers any advantage. Our testbed will be the BEC, as the channel for which feedback achievability strategies are understood best.

The first formulation is to consider an arbitrary VLF code and define the error event differently from (6.7). Namely, fix a delay d and define the probability of error as

$$\epsilon = \mathbb{P}[\{\hat{W} \neq W\} \cup \{\tau > d\}]\,. \tag{6.225}$$

The question is then: what is the maximum M compatible with a chosen d and ϵ? The answer is obvious: since in such formulation the encoder has no incentive to terminate before the delay d, we might as well fix blocklength to be d and force the decoder to take the decision at time d. This, however, is no different from fixed-blocklength coding with feedback, which we have considered in Section 6.5. In particular, we have demonstrated

that for the BEC (and the BSC, and many other symmetric channels) the feedback does not affect the $\sqrt{n}$ term in the expansion (6.1).

The second formulation applies to zero-error VLFT codes for which we define ϵ-delay as

$$D_\epsilon = \min\{n : \mathbb{P}[\tau > n] \le \epsilon\}\,, \quad \epsilon \in [0,1]. \tag{6.226}$$

Thus a zero-error VLFT code with $D_\epsilon \le d$ is a code which is guaranteed to deliver the data error-free, and does so in less than d channel uses in all except ϵ-portion of the cases. The question arises: for a fixed ϵ, what is the maximum M compatible with a given ϵ-delay requirement d:

$$M_z^*(d,\epsilon) = \max\{M : \exists\,\text{zero-error VLFT code with } D_\epsilon \le d\}\,? \tag{6.227}$$

The obvious achievability bound is to simply pair a fixed-blocklength non-feedback (n, M, ϵ) with $n = d$ code with an ARQ retransmission strategy to achieve zero error. We have thus

$$M_z^*(d,\epsilon) \ge M^*(d,\epsilon) = dC - \sqrt{dV}Q^{-1}(\epsilon) + O(\log d)\,, \tag{6.228}$$

where $M^*(d,\epsilon)$ denotes the performance of the best non-feedback, fixed-blocklength code and is thus given by (6.1).

Can we improve the crucial $\sqrt{d}$-penalty term in (6.228)? The answer is negative, at least for the BEC:

Theorem 111 *For the BEC, we have*

$$\log M_z^*(d,\epsilon) \le dC - \sqrt{dV}Q^{-1}(\epsilon) + \log d + O(1)\,, \tag{6.229}$$

where C and V are the capacity and the dispersion of the BEC.

Proof: Let E_j be the i.i.d. process corresponding to erasures: $\mathbb{P}[E_j = 0] = 1 - \mathbb{P}[E_j = 1] = \delta$, where δ is the erasure probability of the BEC. Then the total number of unerased symbols by time n is given by

$$N_n = \sum_{j=1}^{n} E_j\,. \tag{6.230}$$

Following the steps of the proof of the converse theorem for the BEC, Theorem 43, we can see that by time n the total number of messages distinguishable at the receiver is upper-bounded by $\sum_{j=0}^{n} 2^{N_j}$ (summation corresponds to the fact that a VLFT code has the freedom of sending a termination symbol at any time). Therefore, since the code achieves zero-error we have

$$\mathbb{P}[\tau \le n] \le \frac{1}{M}\mathbb{E}\left[\min\left\{\sum_{j=0}^{n} 2^{N_j}\,, M\right\}\right]\,. \tag{6.231}$$

Since N_t is a monotonically non-decreasing it follows that

$$\begin{aligned}
\sum_{j=0}^{n} 2^{N_j} &\le \sum_{t=0}^{N_n} 2^t + (n - N_n)2^{N_n} && (6.232)\\
&\le 2^{N_n}(n + 2 - N_n) && (6.233)\\
&\le (n+2)2^{N_n}\,. && (6.234)
\end{aligned}$$

Although the bound (6.233) is useful for numerical evaluation, the bound (6.234) is more convenient for the analysis. Indeed, we have from (6.231) and (6.234):

$$\begin{aligned} \mathbb{P}[\tau \le n] &\le \frac{1}{M}\mathbb{E}\left[\min\left\{(n+2)2^{N_n}, M\right\}\right] && (6.235)\\ &= \frac{n+2}{M}\mathbb{E}\left[\min\left\{2^{N_n}, \frac{M}{n+2}\right\}\right]. && (6.236) \end{aligned}$$

Recall now that for the non-feedback case, Theorem 43 can be restated as

$$1-\epsilon \le \frac{1}{M}\mathbb{E}\left[\min\left\{2^{N_n}, M\right\}\right]. \tag{6.237}$$

The analysis of the bound (6.237) in the proof of Theorem 44 (asymptotic expansion for the BEC), has shown that (6.237) implies

$$\log M \le nC - \sqrt{nV}Q^{-1}(\epsilon) + O(1), \tag{6.238}$$

as $n \to \infty$, where C and V are the capacity and the dispersion of the BEC. Comparing (6.237) and (6.236) we see that M is replaced by $\frac{M}{n+2}$. Therefore, the same argument as the one leading from (6.237) to (6.238) when applied to (6.236) must give

$$\log M \le nC - \sqrt{nV}Q^{-1}(\epsilon) + \log(n+2) + O(1), \tag{6.239}$$

which implies (6.229). ■

6.10 Discussion of the results

We have demonstrated that by allowing variable length, even a modicum of feedback is enough to considerably speed up convergence to capacity. For illustration purposes we can see in Fig. 6.4 that we have constructed a decision feedback code, that achieves, for example, 90% of the capacity of the BSC with crossover probability $\delta = 0.11$ and probability of error $\epsilon = 10^{-3}$ at blocklength 200; see Fig. 6.4. In contrast, to obtain the same performance with fixed-blocklength codes requires a blocklength of at least 3100 even if full noiseless feedback is available at the transmitter. This practical benefit of VLF codes opens the possibility of utilizing the full capacity of the link without the complexity required to implement coding of very long data packets.

A major ingredient of the achievability bounds in this chapter is the idea of terminating early on favorable noise realizations. Although, we have applied this idea to the codes with codewords with unbounded durations, it is clear that without any significant effect on probability of error we could also assume that the transmission forcibly terminates after a time which is a few times the average blocklength ℓ. Consequently, it can be shown that any point on the achievability curve of Fig. 6.4 can be realized by pairing some linear block code with the stopping rule (6.87). In other words, even traditional fixed-blocklength linear codes can be decoded with significantly less (average) delay if used in the variable-length setting. It is important, thus, to investigate whether traditionally good codes (such as LDPC codes) are also competitive in this setting.

Theoretically, the benefit of feedback is manifested by the absence of the $\sqrt{\ell}$ term in the expansions (6.70) and (6.71), whereas this term is crucial to determine the non-asymptotic performance without feedback. Equivalently, we have demonstrated that for variable-length codes with feedback the channel dispersion is zero. To intuitively explain this phenomenon, we note that without feedback the main effect governing the $\sqrt{n}$ behavior was the stochastic variation of information density around its mean, which is tightly characterized by the central limit theorem. In the variable-length setup with feedback the main idea is that of Wald-like stopping once the information density of some message is large enough. Therefore, there is virtually no stochastic variation (besides a negligible overshoot) and this explains the absence of any references to the central limit theorem.

We have also analyzed a modification of the coding problem by introducing a termination symbol (VLFT codes), which is practically motivated in many situations in which control signals are sent over a highly reliable upper layer. We have shown that in this setup, in addition to the absence of $\sqrt{\ell}$ term, the principal new effect is that the zero-error capacity increases to the full Shannon capacity of the channel. Although availability of a "use-once" termination symbol is immaterial asymptotically, the transient behavior is significantly improved. Analytically, this effect is predicted by the absence of not only the $\sqrt{\ell}$ term but also of the $\log \ell$ term in the achievability bound (6.191). Furthermore, our codes with termination have a particularly convenient structure: the encoder uses the feedback link only to choose the time when to stop the transmission (by sending the termination symbol), and otherwise simply sends a fixed message-dependent codeword. The codes with such structure have been called fixed-to-variable (FV), or fountain, codes in [113]. Thus, in short, we have demonstrated that fountain codes can achieve 90% of the capacity of the BSC with crossover probability $\delta = 0.11$ at average blocklength < 20 and with zero probability of error. Practically, of course, "zero-error" should be understood as the reliability being essentially the probability with which the termination symbol is correctly detected.

Finally, we have discussed some questions regarding communication of real-time data. We have demonstrated that constraints on the excess delay nullify the advantage of feedback (and variable-length), i.e. the improvement in performance of the best feedback code can be marginal at best compared to non-feedback, fixed-blocklength codes. This, of course, contrasts sharply with the results regarding the average length.

Appendix A

Asymptotic behavior of β

This appendix provides proofs of Lemmas 14 and 15.

Proof of Lemma 14: We will simply apply the Berry-Esseen Theorem 13 twice. We start from the lower bound. Applying (2.67) we get

$$\beta_\alpha \geq \frac{1}{\gamma_n}\left(\alpha - P\left[\log\frac{dP}{dQ} \geq \log\gamma_n\right]\right) \tag{A.1}$$

for $\gamma_n > 0$. Now set

$$\alpha_n = \alpha - \frac{B_n + \Delta}{\sqrt{n}}\,. \tag{A.2}$$

This quantity is positive by requirement that the argument of Q^{-1} in (2.87) be positive. Therefore, choose

$$\log\gamma_n = nD_n + \sqrt{nV_n}Q^{-1}(\alpha_n)\,. \tag{A.3}$$

Then since $\log\frac{dP}{dQ}$ is a sum of independent random variables, Berry-Esseen Theorem 13 applies and

$$\left|P\left[\log\frac{dP}{dQ} \geq \log\gamma_n\right] - \alpha_n\right| \leq \frac{B_n}{\sqrt{n}}\,. \tag{A.4}$$

Consequently,

$$P\left[\log\frac{dP}{dQ} \geq \log\gamma_n\right] \leq \alpha - \frac{\Delta}{\sqrt{n}}\,. \tag{A.5}$$

Substituting this bound into (A.1) we obtain (2.87).

From Neyman-Pearson lemma and by the monotonicity of β_α, cf. derivation of (2.68), we have

$$\beta_\alpha \leq Q\left[\log\frac{dP}{dQ} \geq \log\gamma_n\right]\,, \tag{A.6}$$

whenever γ_n satisfies

$$P\left[\log\frac{dP}{dQ} \geq \log\gamma_n\right] \geq \alpha\,. \tag{A.7}$$

Again, set

$$\alpha_n = \alpha + \frac{B_n}{\sqrt{n}}\,. \tag{A.8}$$

This $\alpha_n < 1$ by the requirement that the argument of Q^{-1} in (2.88) be below 1. Also choose

$$\log \gamma_n = nD + \sqrt{nV} Q^{-1}(\alpha_n) \,. \tag{A.9}$$

From the Berry-Esseen bound, we have

$$\left| P\left[\log \frac{dP}{dQ} \geq \log \gamma_n \right] - \alpha_n \right| \leq \frac{B_n}{\sqrt{n}} \,. \tag{A.10}$$

Consequently,

$$P\left[\log \frac{dP}{dQ} \geq \log \gamma_n \right] \geq \alpha \,. \tag{A.11}$$

Thus, this choice of γ_n indeed satisfies (A.7).

Finally, (2.88) follows from (A.6):

$$\begin{aligned} \beta_\alpha &\leq Q\left[\log \frac{dP}{dQ} \geq \log \gamma_n \right] && \text{(A.12)} \\ &= \mathbb{E}_P \left[\exp\left\{ -\log \frac{dP}{dQ} \right\} 1\left\{ \log \frac{dP}{dQ} \geq \log \gamma_n \right\} \right] && \text{(A.13)} \\ &\leq \frac{1}{\sqrt{n}\gamma_n} \left(\frac{2\log 2}{\sqrt{2\pi V_n}} + 4B_n \right) \frac{1}{\sqrt{n}} && \text{(A.14)} \end{aligned}$$

where (A.14) is by Lemma 20. ■

Proof of Lemma 15: Just as in the above argument, we start by writing

$$\beta_\alpha \geq \frac{1}{\gamma_n} \left(\alpha - P\left[\log \frac{dP}{dQ} \geq \log \gamma_n \right] \right) \,. \tag{A.15}$$

We notice that

$$nD_n = \mathbb{E}_P \left[\log \frac{dP}{dQ} \right] \,, \quad nV_n = \mathbb{E}_P \left[\left(\log \frac{dP}{dQ} - nD_n \right)^2 \right] \,. \tag{A.16}$$

Thus, if we set

$$\log \gamma_n = nD_n + \sqrt{\frac{2nV_n}{\alpha}} \,, \tag{A.17}$$

then

$$\begin{aligned} P\left[\log \frac{dP}{dQ} \geq \log \gamma_n \right] &= P\left[\log \frac{dP}{dQ} - nD_n \geq \sqrt{\frac{2nV_n}{\alpha}} \right] \leq && \text{(A.18)} \\ &\leq P\left[\left(\log \frac{dP}{dQ} - nD_n \right)^2 \geq \frac{2nV_n}{\alpha} \right] \leq \frac{\alpha}{2} \,. && \text{(A.19)} \end{aligned}$$

The last step is by Chebyshev inequality. Putting this into (A.15) we obtain the required result after taking the logarithm. ■

Appendix B

$\kappa\beta$ bound and deterministic hypothesis tests

In this appendix we will formulate the analog of Theorem 27 that constructs a non-randomized decoder.

Similarly to (2.60), a non-randomized test between distributions $P_{Y|X=x}$ and Q_Y is defined by a critical set $E \subseteq \mathsf{B}$, where $y \in E$ indicates that the test chooses $P_{Y|X=x}$. The best performance achievable among those randomized tests is given by

$$\tilde{\beta}_\alpha(x, P_Y) = \inf_{E: P_{Y|X=x}(E) \geq \alpha} P_Y(E)\,. \tag{B.1}$$

For β_α we have a Neyman-Pearson lemma that guarantees that the optimal randomized test does exist. We used this fact extensively in the proof of Theorem 27. So we must show that infimum in (B.1) is in fact a minimum. We could not find this result in the literature and give the proof below.

Recall that measurable spaces X and Y are called isomorphic if there is a bijective measurable map $f : X \mapsto Y$ with measurable inverse f^{-1}. The space X is *standard* if it is isomorphic to a Borel subset D of real line (with the inherited σ-algebra). A motivating result is that every Polish space (in particular complete separable metric space) with Borel σ-algebra is standard. Another deep result is that every standard space is isomorphic to either $[0,1]$, $\{0, \frac{1}{1}, \ldots, \frac{1}{n}\}$ or $\{0, \frac{1}{n}, n \in \mathbb{N}\}$. Thus in the definition above we can restrict D to be a closed subset of $[0,1]$.

Theorem 112 *Given a standard measurable space X, two finite measures P and Q on X, and any $\alpha \in [0, Q(X)]$ there exists a measurable set E^*_α such that $Q(E^*_\alpha) \geq \alpha$ and*

$$P(E^*_\alpha) = \inf_{E:\, Q(E) \geq \alpha} P(E)\,. \tag{B.2}$$

The proof of this theorem can be found at the end of this appendix.

Theorem 113 (Achievability) *Suppose that B is a standard measurable space. Then for any $0 < \epsilon < 1$, there exists an (M, ϵ) code with codewords chosen from $\mathsf{F} \subset \mathsf{A}$, satisfying*

$$M \geq \sup_{0<\tau<\epsilon} \sup_{Q_Y} \frac{\kappa_\tau(\mathsf{F}, Q_Y)}{\sup_{x \in \mathsf{F}} \tilde{\beta}_{1-\epsilon+\tau}(x, Q_Y)}\,. \tag{B.3}$$

Proof: The proof of the Theorem 27 applies with the only change that random transformations $P^*_{Z_x|Y}$ are replaced by deterministic ones. More specifically,

$$P^*_{Z_x|Y}(\cdot|y) \to 1\{y \in B^*_x\}\,, \tag{B.4}$$

where B^*_x is the set that achieves infimum in the (B.1). Such B^*_x does exist by Theorem 112. ■

Theorem 113 gives an achievability bound by constructing a codebook and a non-randomized decoder. The drawback is that it is not straightforward to calculate the non-randomized $\tilde{\beta}_\alpha$ because we do not have a Neyman-Pearson lemma for this case. A possible workaround is to find the connection between β_α and $\tilde{\beta}_\alpha$. Such a connection is established in the next two lemmas.

Assume that we have two probability measures P_Y and Q_Y on B. We define the performance of the best deterministic and best randomized test, correspondingly, as

$$\tilde{\beta}_\alpha = \inf_{\{E:\, Q_Y(E)\geq\alpha\}} P_Y(E)\,, \tag{B.5}$$

$$\beta_\alpha = \min_{\substack{P_{Z|Y}:\\ \sum_{y\in\mathsf{B}} Q_Y(y)P_{Z|Y}(1|y)\geq\alpha}} \sum_{y\in\mathsf{B}} P_Y(y)P_{Z|Y}(1|y)\,. \tag{B.6}$$

Note that $f(y) = P_{Z|Y}(1|y)$ is a function of y such that $f(y) \in [0,1]$. Thus, we can rewrite the definition of β_α as

$$\beta_\alpha = \inf_{\{f:\, \int f dQ_Y\geq\alpha\}} \int f\, dP_Y \tag{B.7}$$

with f satisfying $f(y) \in [0,1]$.

Lemma 114 *Consider P_Y and Q_Y such that*[1] *$Q_Y \ll P_Y$. Then the following holds:*

$$\beta_\alpha \leq \tilde{\beta}_\alpha \leq \beta_\alpha + \sup_{y\in\mathsf{B}} P_Y(y)\,. \tag{B.8}$$

Proof: The left-hand inequality is obvious and we will concentrate on the right-hand one. From the Neyman-Pearson lemma we know that, the inf in (B.7) is achieved by some function f. Moreover, the optimal f takes only three values: $0, 1, \tau$, where $\tau \in (0,1)$. Denote

$$E_1 = f^{-1}\{1\} \text{ and } \quad E_2 = f^{-1}\{\tau\}\,. \tag{B.9}$$

Then

$$\beta_\alpha = \int f\, dP_Y = P_Y(E_1) + \tau P_Y(E_2)\,. \tag{B.10}$$

If $P_Y(E_2) = 0$ then $Q_Y(E_2) = 0$ (by absolute continuity). Consequently, we can then take the non-randomized test to be $E = E_1$. Indeed,

$$Q_Y(E) = Q_Y(E_1) = \int f\, dQ_Y \geq \alpha \tag{B.11}$$

[1]This assumption can be dropped, in which case in the proof we must consider only tests f such that $f(N) = 1$. Here N is the set such that $P_Y(N) = 0$ but $Q_Y(N)$ is the maximum possible value. Existence of such a set is well-known. Then on N^c we have $Q_Y \ll P_Y$.

and $P_Y(E) = P_Y(E_1) = \beta_\alpha$.

Assume that $P_Y(E_2) > 0$. Since B is standard we may assume that it is a subset of $[0,1]$ with Borel σ-algebra. Define a non-decreasing right-continuous function

$$g(t) = P_Y(E_2 \cap [0,t])\,. \tag{B.12}$$

Note that $g(t)$ grows from $g(0)$ to $P_Y(E_2)$. Then define

$$\lambda = \inf\{t:\ g(t) > \tau P_Y(E_2)\}\,. \tag{B.13}$$

Since the height of any jump in $g(t)$ is bounded by $\sup_y P_Y(y)$, we can see that

$$\tau P_Y(E_2) \le g(\lambda) \le \tau P_Y(E_2) + \sup_y P_Y(y)\,. \tag{B.14}$$

Now define a non-randomized test

$$E = E_1 \cup (E_2 \cap [0,\lambda])\,. \tag{B.15}$$

Then, we have

$$Q_Y(E) = Q_Y(E_1) + Q_Y(E_2 \cap [0,\lambda]) = Q_Y(E_1) + g(\lambda)\frac{Q_Y(E_2)}{P_Y(E_2)}\,. \tag{B.16}$$

Here we used another property of the optimal randomized test f, namely, that $\frac{dQ_Y}{dP_Y}$ is constant on E_2. Then lower bounding $g(\lambda)$ by (B.14) we conclude that E is a valid deterministic test with $Q_Y(E) \ge \alpha$. But by the upper bound in (B.14), we have

$$P_Y(E) = P_Y(E_1) + g(\lambda) \le P_Y(E_1) + \tau P_Y(E_2) + \sup_y P_Y(y)\,. \tag{B.17}$$

This establishes the lemma. ■

Sometimes, it is more convenient to quantify the difference between $\tilde{\beta}_\alpha$ and β_α in terms of Q_Y. This is the content of the next lemma.

Lemma 115 *In the conditions of previous lemma take $E_2 = f^{-1}(0,1)$ to be as defined in the above proof. Then, if $P_Y(E_2) = 0$ we have $\tilde{\beta}_\alpha = \beta_\alpha$. Otherwise, on E_2, the derivative $\frac{dQ_Y}{dP_Y}$ is a constant $\gamma > 0$ (P_Y almost surely) and*

$$\beta_\alpha \le \tilde{\beta}_\alpha \le \beta_\alpha + \frac{1}{\gamma}\sup_y Q_Y(y)\,. \tag{B.18}$$

Proof: The case in which $P_Y(E_2) = 0$ was taken care of in the proof above. The fact that $\frac{dQ_Y}{dP_Y}$ is P_Y-a.s. constant is known from the Neyman-Pearson lemma. That constant $\gamma > 0$ because if γ were equal to 0 then we could improve the test by setting $f(y) = 0$ on E_2. This would leave the integral with respect to Q_Y the same but reduce the integral with respect to P_Y by $\tau P_Y(E_2)$.

Finally, as in the proof above, there is a λ such that

$$\tau Q_Y(E_2) \le Q_Y(E_2 \cap [0,\lambda]) \le \tau Q_Y(E_2) + \sup_y Q_Y(y)\,. \tag{B.19}$$

Thus,

$$P_Y(E_2 \cap [0,\lambda]) = Q_Y(E_2 \cap [0,\lambda])\frac{1}{\gamma} \le \tau Q_Y(E_2)\frac{1}{\gamma} + \frac{1}{\gamma}\sup_y Q_Y(y) = \tau P_Y(E_2) + \frac{1}{\gamma}\sup_y Q_Y(y)\,. \tag{B.20}$$

■

It is not too hard to see that, due to the last two lemmas, the asymptotic behavior of $\tilde{\beta}_\alpha^n$ and β_α^n coincide. Thus, the achievability part of normal approximation Theorems 45 and 73 can be established with deterministic decoders. Moreover, the difference between $\tilde{\beta}_\alpha^n$ and β_α^n is so insignificant, that all the numerical plots remain the same for the deterministic encoders.

Proof of Theorem 112: We wish to show the existence of the minimizing set E_α^* in the optimization problem:

$$\beta_\alpha = \inf_{E:Q(E)\ge\alpha} P(E)\,. \tag{B.21}$$

We denote the underlying measurable space X, its σ-algebra $\mathcal{F}$ and assume that all sets appearing below (and in optimization problem above) belong to $\mathcal{F}$.

First, write the Lebesgue decomposition of P as

$$P(A) = \int_A \frac{dP}{dQ}\,dQ + P(A \cap N)\,, \quad \forall A \in \mathcal{F} \tag{B.22}$$

and $Q(N) = 0$. Note that every test E can be improved by replacing it with $E \cap N^c$. Indeed, such change does not affect its Q-measure while reducing its P-measure. Thus we can restrict the optimization only to sets inside N^c. Equivalently, we can assume from now on that

$$P \ll Q\,. \tag{B.23}$$

Now write the decomposition of Q as

$$Q(A) = \int_A \frac{dQ}{dP}\,dP + Q(A \cap M)\,, \quad \forall A \in \mathcal{F} \tag{B.24}$$

and $P(M) = 0$. Adding M to any test E can not change its P-measure, while it increases its Q-measure "for free". Thus we can restrict our attention only to tests $E \supseteq M$ and replace α with $\alpha - Q(M)$. If $\alpha - Q(M) \le 0$ then a solution is found: $E_\alpha^* = M$. Otherwise, $\alpha - Q(M) > 0$ and we can ignore the singular part of Q. Equivalently, from now on we assume

$$Q \ll P \text{ and } P \ll Q \quad (\text{i.e., } P \sim Q)\,. \tag{B.25}$$

Because X is a standard space, singletons $\{x\}$ are measurable and we can split every measure into purely atomic and diffuse (non-atomic) parts:

$$P = P_a + P_d\,,\ \ Q = Q_a + Q_d\,. \tag{B.26}$$

Denote the set of all atoms of P by

$$D = \{x \in X : P(x) > 0\} \tag{B.27}$$

which is at most countable since P is finite. Note that because of condition (B.25) atoms of P and Q must coincide. In other words, measures P_a and Q_a are restrictions of P and Q to D:

$$P_a(A) = P(A \cap D) \text{ and } Q_a(A) = Q(A \cap D)\,. \tag{B.28}$$

Measures P_d and Q_d are restrictions to D^c and also

$$P_a \perp P_d,\ \ Q_a \perp Q_d\,, \text{and } P_a \sim Q_a\,,\ P_d \sim Q_d\,. \tag{B.29}$$

Our general direction will now be as follows. We can see that each test P vs. Q is also a test P_a vs Q_a and P_d vs Q_d. This idea allows us to separately treat cases of purely atomic and purely diffuse measures. Indeed, if we will find optimal test for P_a vs Q_a and P_d vs Q_d then their sum will be an optimal test for P vs Q with β's and α's added (because $P_a \perp P_d$ and $Q_a \perp Q_d$). Let us make this idea precise.

Denote

$$\mathcal{A}_\alpha = \{E \in \mathcal{F} : Q(E) \geq \alpha\}\,. \tag{B.30}$$

Then, set $t_0 = (\alpha - Q_a(X)) \vee 0$ and $t_1 = Q_d(X)$. Every set $E \in \mathcal{A}_\alpha$ has $Q_d(E) \in [t_0, t_1]$. Conversely, for every $t \in [t_0, t_1]$ there is at least one set $E \in \mathcal{A}_\alpha$ such that $Q_d(E) = t$. This is because X is standard and Q_d diffuse (this can be shown using the method in the proof of Lemma 114). Define for any $t \in [t_0, t_1]$

$$\mathcal{B}_t = \{E \in \mathcal{A}_\alpha : Q_d(E) = t\}\,. \tag{B.31}$$

As we have shown each $\mathcal{B}_t$ is non-empty and

$$\mathcal{A}_\alpha = \bigcup_{t \in [t_0, t_1]} \mathcal{B}_t\,. \tag{B.32}$$

But then

$$\beta_\alpha \stackrel{\triangle}{=} \inf_{E \in \mathcal{A}_\alpha} P(E) = \inf_{t \in [t_0, t_1]} g(t)\,, \tag{B.33}$$

where

$$g(t) = \inf_{E \in \mathcal{B}_t} P(E)\,. \tag{B.34}$$

Note that the following two claims imply the statement of the theorem.

Claim A. The infimum in the definition of $g(t)$ is achievable by some set E_t^*.

Claim B. $g(t)$ is lower semi-continuous.

Indeed, then $\inf_{t \in [t_0, t_1]} g(t)$ must be achieved for some t^* and we can take $E_{t^*}^*$ as the set E_α^*.

We now proceed to show those two claims. Define two new functions:

$$f_d(r) = \inf_{F:\, Q_d(F) \geq r} P_d(F)\,, \tag{B.35}$$

$$f_a(r) = \inf_{G:\, Q_a(G) \geq r} P_a(G)\,. \tag{B.36}$$

Suppose, that we were able to show the following four facts:

F1. For every $r \in [0, Q_d(X)]$ there is a set $F^* \in \mathcal{F}$ such that $P_d(F^*) = f_d(r)$ and $Q_d(F^*) = r$.

F2. Function $r \mapsto f_d(r)$ is continuous on $[0, Q_d(x)]$.

F3. For every $r \in [0, Q_a(X)]$ there is a set $G^* \in \mathcal{F}$ such that $P_a(G^*) = f_a(r)$.

F4. Function $r \mapsto f_a(r)$ is left-continuous on $[0, Q_a(X)]$.

Note that because Q_d and P_d live on D^c, and Q_a and P_a live on D we can safely assume that

$$F^* \subseteq D^c\,,\; G^* \subseteq D\,, \text{and } F^* \cap G^* = \emptyset\,. \tag{B.37}$$

Claim A follows from F1 and F3. Indeed, fix $t \in [t_0, t_1]$ and take F^* corresponding to $r = t$ in F1. Then we have $Q_d(F^*) = t$. Also take G^* corresponding to $r = \alpha - t$ in F3. Then we have $Q_a(G^*) \geq \alpha - t$. Thanks to conditions (B.37) we have

$$P(F^* \cup G^*) = f_d(t) + f_a(\alpha - t)\,, \tag{B.38}$$

$$Q(F^* \cup G^*) \geq \alpha\,. \tag{B.39}$$

Thus $F^* \cup G^* \in \mathcal{B}_t$, and that means that

$$g(t) \leq f_d(t) + f_a(\alpha - t)\,. \tag{B.40}$$

On the other hand, definition of $\mathcal{B}_t$ can be rewritten as

$$\mathcal{B}_t = \{E \in \mathcal{F} :\; Q_d(E) = t,\, Q_a(E) \geq \alpha - t\}\;. \tag{B.41}$$

Thus,

$$\begin{aligned} g(t) = \inf_{\mathcal{B}_t}[P_d(E)+P_a(E)] \geq \inf_{\mathcal{B}_t} P_d(E)+\inf_{\mathcal{B}_t} P_a(E) = \inf_{E:\, Q_d(E)=t} P_d(E)+\inf_{E:\, Q_a(E)\geq\alpha-t} P_a(E) \geq \\ \geq \inf_{E:\, Q_d(E)\geq t} P_d(E) + f_a(\alpha - t) = f_d(t) + f_a(\alpha - t)\,. \end{aligned} \tag{B.42}$$

Together with (B.40) this gives

$$g(t) = f_d(t) + f_a(\alpha - t)\,. \tag{B.43}$$

And we have shown that value $f_d(t) + f_a(\alpha - t)$ is achieved by $F^* \cup G^*$.

Claim B follows from (B.43) and F2 and F4. Indeed, take $t_n \to t_\infty$. Then $f_d(t_n) \to f_d(t_\infty)$ by F2. The function $f_a(x)$ is non-decreasing and left-continuous by F4. Denote $x_n = \alpha - t_n \to x_\infty = \alpha - t_\infty$. We must show

$$\liminf_{n\to\infty} f_a(x_n) \geq f_a(x_\infty)\,. \tag{B.44}$$

in order to establish lower semi-continuity of $g(t)$.

Define $y_n = \inf_{k\geq n} x_k$. Then $x_n \geq y_n$ and, by the non-decreasing property of f_a, we have

$$f_a(x_n) \geq f_a(y_n)\,. \tag{B.45}$$

On the other hand, $y_n \nearrow x_\infty$ and by left-continuity $f_a(y_n) \nearrow f_a(x_\infty)$. Taking the lim inf in (B.45) gives (B.44).

So, *Claims A and B* were shown to be implied by F1-F4. We are now left to prove F1-F4. Note that facts F1 and F2 are statements about non-randomized tests between two diffuse measures. The Neyman-Pearson lemma yields a randomized test and we know (see proof of Lemma 114) that in the case of diffuse measures (and standard space) it is possible to construct a non-randomized test of the same performance. The ROC curve for the randomized test is continuous and coincides in this case with the ROC curve f_d of the non-randomized test. Thus facts F1 and F2 are apparent.

Proof of facts F3 and F4 is given as a lemma below. Concluding the proof of this theorem we want to emphasize that essentially all work is relegated to that lemma since in this proof we have just shown that only atoms can give problems. ■

Lemma 116 *Let P and Q be two finite measures on the space $X = \mathbb{Z}_+$ of positive integers, and choose the σ-algebra to be $\mathcal{F} = 2^{\mathbb{Z}_+}$. Define*

$$\beta(r) = \inf_{E:\, Q(E) \geq r} P(E)\,. \tag{B.46}$$

Then for each $r \in [0, Q(X)]$ there is a set E^* achieving the infimum in the above, and the function $\beta(r)$ is left-continuous.

Proof: We are going to turn $\mathcal{F}$ into a metric space by defining a distance between sets:

$$d(A, B) = Q(A \,\triangle\, B)\,. \tag{B.47}$$

This quantity satisfies all the properties of a metric except that $d(A, B) = 0$ may not imply that $A = B$. The standard workaround for this difficulty is to define an equivalence relation,

$$A \sim B :\; Q(A \,\triangle\, B) = 0\,, \tag{B.48}$$

and consider a quotient space $\tilde{\mathcal{F}} = \mathcal{F}/\!\sim$ of classes of equivalence. However, in this particular case we can simply drop all elements $x \in X$ that satisfy $Q(x) = 0$. In other words, we can restrict the space X to positive atoms of Q. The corrected X might now be finite. In any case every non-empty set A has positive Q-measure and the only case when $Q(A \,\triangle\, B) = 0$ is if $A = B$. So, $\mathcal{F}$ is a metric space. Note that this automatically makes $P \ll Q$.

Now, notice that functional $Q : \mathcal{F} \mapsto \mathbb{R}_+$ is continuous in this metric simply because $Q(E) = d(E, \emptyset)$ and a metric is always continuous. What is more pleasing is that $P(E)$ is also continuous. To see this, simply write

$$P(A) = \int_A \frac{dP}{dQ}\, dQ\,. \tag{B.49}$$

Then since P is finite, $\frac{dP}{dQ}$ is an integrable function. Thus it is also uniformly integrable. In other words any shrinking A_n makes $P(A_n)$ converge to zero. Formally,

$$Q(A_n) \to 0 \quad \Longrightarrow \quad P(A_n) \to 0\,. \tag{B.50}$$

Now take any $E_n \to E$. Then

$$P(E) - P(E \vartriangle E_n) \leq P(E_n) \leq P(E) + P(E \vartriangle E_n)\,. \tag{B.51}$$

Denote $A_n = E \vartriangle E_n$. Then $Q(A_n) \to 0$ and thus $P(A_n) \to 0$ and $P(E)$ is continuous by (B.51).

The metric space $\mathcal{F}$ is complete. Indeed, take any Cauchy sequence E_n and notice that condition $Q(E_n \vartriangle E_m) \to 0$ for $n, m \to \infty$ is equivalent to

$$\mathbb{E}\left[|1_{E_n} - 1_{E_m}|\right] \to 0\,. \tag{B.52}$$

Or, in other words, a sequence of random variables 1_{E_n} is Cauchy in the L_1 sense for measure Q. But then, it has a measurable limit, denote it h:

$$1_{E_n} \xrightarrow{L_1} h\,. \tag{B.53}$$

Then, using well-known implications, $1_{E_n} \to h(x)$ in measure and thus, there is a subsequence of 1_{E_n} that converges to h almost surely. Denote this subsequence by 1_{F_k}, $F_k = E_{n_k}$. Then, since

$$1_{F_k} \to h \quad Q\text{-a.s.} \tag{B.54}$$

we can conclude that the limit $h(x) = \lim 1_{F_k}(x)$ is equal to zero or one almost surely. Taking $E_\infty = \{x : h(x) = 1\}$ we conclude that $\mathbb{E}\left[|1_{E_n} - 1_{E_\infty}|\right] \to 0$. Thus, $E_n \to E_\infty$ in the Q-metric.

Note that nothing we have done so far used the fact that X is a discrete space. Indeed, if we worked with equivalence classes and postulated $P \ll Q$ in the statement we could still have that $P(E)$ and $Q(E)$ are continuous in the Q-metric and the space is complete. However, what $\mathcal{F}$ misses in the general case is total boundedness. But for X discrete we can show this.

So finally, $\mathcal{F}$ is totally bounded. Indeed, first assume that atoms of Q are sorted in decreasing order, that is $Q(1) \geq Q(2) \geq Q(3) \geq \cdots$. This is possible because Q is finite. Then choose some $\epsilon > 0$ and select the set

$$G_\epsilon = \{1, 2, \ldots, N_\epsilon\}\,, \tag{B.55}$$

where N_ϵ chosen such that

$$\sum_{n > N_\epsilon} Q(n) \leq \epsilon\,. \tag{B.56}$$

Now, generate all subsets of G_ϵ

$$\mathcal{C}_\epsilon = 2^{G_\epsilon}\,, |\mathcal{C}_\epsilon| = 2^{N_\epsilon}\,. \tag{B.57}$$

Then, for any element $E \in \mathcal{F}$ we can choose a set $C \in \mathcal{C}_\epsilon$ such that

$$Q(E \vartriangle C) \leq \epsilon\,. \tag{B.58}$$

Simply take $C = E \cap G_\epsilon$. Thus, we have shown that for every $\epsilon > 0$ there is a finite ϵ-cover of the space $\mathcal{F}$. Thus by definition $\mathcal{F}$ is totally bounded. Finally, we have established that $\mathcal{F}$ is a compact metric space[2] .

To conclude the proof we are left to note that constraint $Q(E) \geq \alpha$ can be restated as

$$E \in Q^{-1}[\alpha, +\infty)\,. \tag{B.60}$$

Here Q is a continuous function. Thus a pre-image of a closed $[\alpha, +\infty)$ is closed. Consequently, $Q^{-1}[\alpha, +\infty)$ is compact and in (B.46) we are minimizing a continuous function $P(E)$ over a compact set. A famous result guarantees the existence of a minimizing set E^*.

Now we will show that $\beta(r)$ is left-continuous. Fix some $r_0 > 0$. First of all, because $\beta(r)$ is non-decreasing

$$\beta(r_0) \geq \beta(r_0-)\,. \tag{B.61}$$

Second, take a sequence $r_n \nearrow r_0$. For every r_n there is a set E_n^* with $P(E_n^*) = \beta(r_n)$ and $Q(E_n^*) \geq r_n$. Now, since the entire $\mathcal{F}$ is compact there is a convergent subsequence $E_{n_k}^* \to E^*$. Then, since Q and P are both continuous, we can see that

$$Q(E^*) = \lim_{k\to\infty} Q(E_{n_k}^*) \geq \lim_{k\to\infty} r_{n_k} = r_0\,. \tag{B.62}$$

Thus E^* satisfies condition $Q(E^*) \geq r_0$. But on the other hand

$$P(E^*) = \lim_{k\to\infty} P(E_{n_k}^*) = \beta(r_0-)\,. \tag{B.63}$$

Thus $\beta(r_0) \leq \beta(r_0-)$. Together with (B.61) this concludes the proof of the lemma. ■

Another approach for the latter lemma could be the following. Note that each set is a mapping $X \mapsto \{0,1\}$. Thus sets are elements of $\{0,1\}^X$. By Tychonoff's theorem any product of compact spaces (and $\{0,1\}$ is compact) is itself compact. Thus $\{0,1\}^X$ is compact (and Hausdorff) in the product topology. Then $P(E)$ and $Q(E)$ are continuous functionals $\{0,1\}^X \mapsto \mathbb{R}_+$ because of bounded convergence theorem (product topology means pointwise convergence). This approach fails in the general case (when X is non-discrete) for a subtle reason: $\mathcal{F}$ is not a closed subset of $\{0,1\}^X$. In fact the closure of $\mathcal{F}$ would be the entire $\{0,1\}^X$ which is a power set 2^X. Thus $\mathcal{F}$ can not be compact in this topology.

[2]For completeness we give a counter example for the general case. Take $X = [0,1]$ with $\mathcal{F}$ a Borel σ-algebra and Q a Lebesgue measure. Then define a collection of sets

$$F_k = \bigcup_{j=1}^{k} [(2j-2)2^{-k},\, (2j-1)2^{-k}]\,. \tag{B.59}$$

It is easy to see that $Q(F_k \,\triangle\, F_m) = 1/2$ whenever $k \neq m$. But if $\mathcal{F}$ with this metric were totally bounded then it would be compact and then F_k would contain a convergent subsequence. This is impossible because that subsequence is not Cauchy.

Appendix C

Bounds via linear codes

The goal of this appendix is to illustrate how Theorems 17 and 18, which give an upper bound on average probability of error, can also be used to derive an upper bound on maximal probability of error. To that end, we first notice that in both proofs we relied only on pairwise independence between randomly chosen codewords. So, the average probability of error for any other ensemble of codebooks with this property and whose marginals are identical and equal to P_X will still satisfy bounds of Theorems 17 and 18. In particular, for the BSC and the BEC we can generate an ensemble with equiprobable P_X by using a linear code with entries in its generating matrix chosen equiprobably on $\{0,1\}$. Then, Theorems 17 and 18 guarantee the existence of the codebook, whose probability of error under maximum likelihood (ML) decoding is small. Note that this is only possible if $M = 2^k$ for some integer k. A question arises: for these structured codebooks are there randomized ML decoders whose maximal probability of error coincides with the average? This question is answered by the following result.

Theorem 117 *Suppose that* A *is a group and suppose that there is a collection of measurable mappings* $T_x : \mathsf{B} \mapsto \mathsf{B}$ *for each* $x \in \mathsf{A}$ *such that*

$$P_{Y|X=x' \circ x} = P_{Y|X=x'} \circ (T_x)^{-1}\,, \quad \forall x' \in \mathsf{A}\,. \tag{C.1}$$

Then any code $\mathcal{C}$ *that is a subgroup of* A *has a maximum likelihood decoder whose maximal probability of error coincides with the average probability of error.*

Note that condition (C.1) can be reformulated as

$$\mathbb{E}\left[g(Y)\,|\,X = x' \circ x\right] = \mathbb{E}\left[g(T_x(Y))\,|\,X = x'\right] \tag{C.2}$$

for all bounded measurable $g : \mathsf{B} \mapsto \mathsf{B}$ and all $x' \in \mathsf{A}$.

Proof: Define P_X to be a measure induced by the codebook $\mathcal{C}$:

$$P_X(E) = \frac{1}{M}|E \cap \mathcal{C}|\,. \tag{C.3}$$

Note that in this case P_Y induced by this P_X dominates all of $P_{Y|X=x}$ for $x \in \mathcal{C}$:

$$P_{Y|X=x} \ll P_Y\,, \quad \forall x \in \mathcal{C}\,. \tag{C.4}$$

Thus, we can introduce densities

$$f_{Y|X}(y|x) \stackrel{\triangle}{=} \frac{dP_{Y|X=x}}{dP_Y}\,. \tag{C.5}$$

Observe that for any bounded measurable g we have

$$\mathbb{E}\left[g(Y)\right] = \mathbb{E}\left[g(T_x(Y))\right], \quad \forall x \in \mathcal{C}\,. \tag{C.6}$$

Indeed,

$$\begin{aligned}
\mathbb{E}\left[g(T_x(Y))\right] &= \sum_{x'\in\mathcal{C}} \frac{1}{M}\mathbb{E}\left[g(T_x(Y)\,|\,X=x'\right] && \text{(C.7)}\\
&= \sum_{x'\in\mathcal{C}} \frac{1}{M}\mathbb{E}\left[g(Y)\,|\,X=x'\circ x\right] && \text{(C.8)}\\
&= \mathbb{E}\left[g(Y)\right], && \text{(C.9)}
\end{aligned}$$

where (C.8) follows from (C.2). Also for any $x, x' \in \mathcal{C}$ we have

$$f_{Y|X}(y|x') = f_{Y|X}(T_x(y)|x'\circ x) \quad P_Y\text{-a.s.}\,. \tag{C.10}$$

Indeed, denote

$$E_1 = \{y : f_{Y|X}(y|x') < f_{Y|X}(T_x(y)|x'\circ x)\} \tag{C.11}$$

and assume that $P_Y(E_1) = P_Y(T_x^{-1}E_1) > 0$. Then, on one hand

$$\begin{aligned}
P_{Y|X}[T_x^{-1}E_1\,|\,x'] &= \int_{\mathsf{B}} P_Y(dy)1_{\{T_x(y)\in E_1\}}f_{Y|X}(y|x') && \text{(C.12)}\\
&< \int P_Y(dy)1_{\{T_x(y)\in E_1\}}f_{Y|X}(T_x(y)|x'\circ x) && \text{(C.13)}\\
&= \int P_Y(dy)1_{\{y\in E_1\}}f_{Y|X}(y|x'\circ x) && \text{(C.14)}\\
&= P_{Y|X}[E_1|x'\circ x]\,, && \text{(C.15)}
\end{aligned}$$

where (C.14) follows from (C.6). But (C.15) contradicts (C.1) and hence $P_Y(E_1) = 0$ and (C.10) is proved.

We proceed to define a decoder by the following rule: upon reception of y compute $f_{Y|X}(y|x)$ for each $x \in \mathcal{C}$; choose equiprobably among all the codewords that achieve the maximal $f_{Y|X}(y|x)$. Obviously, such decoder is maximum likelihood. We now analyze the conditional probability of error given that the true codeword is x. Define two collections of functions of y, parametrized by $x \in \mathcal{C}$:

$$A_x(y) = \min\left\{1, \sum_{x'\in\mathcal{C}} 1\{f_{Y|X}(y|x') > f_{Y|X}(y|x)\}\right\} \tag{C.16}$$

$$N_x(y) = \sum_{x'\in\mathcal{C}} 1\{f_{Y|X}(y|x') = f_{Y|X}(y|x)\}\,. \tag{C.17}$$

It is easy to see that

$$\epsilon_x \stackrel{\triangle}{=} \mathbb{P}[\text{error} \,|\, X = x] \tag{C.18}$$

$$= \mathbb{E}\left[A_x(Y) + 1\{A_x(Y) = 0\}\frac{N_x(Y) - 1}{N_x(Y)} \,\middle|\, X = x\right]. \tag{C.19}$$

If we denote the unit element of $\mathcal{X}$ by x_0, then by (C.10) it is clear that

$$A_x \circ T_x = A_{x_0} \tag{C.20}$$

$$N_x \circ T_x = N_{x_0}\,. \tag{C.21}$$

But then, by (C.19) we have

$$\epsilon_x = \mathbb{E}\left[A_x(Y) + 1\{A_x(Y) = 0\}\frac{N_x(Y) - 1}{N_x(Y)} \,\middle|\, X = x_0 \circ x\right] \tag{C.22}$$

$$= \mathbb{E}\left[A_x(T_x(Y)) + 1\{A_x(T_x(Y)) = 0\}\frac{N_x(T_x(Y)) - 1}{N_x(T_x(Y))} \,\middle|\, X = x_0\right] \tag{C.23}$$

$$= \mathbb{E}\left[A_{x_0}(Y) + 1\{A_{x_0}(Y) = 0\}\frac{N_{x_0}(Y) - 1}{N_{x_0}(Y)} \,\middle|\, X = x_0\right] \tag{C.24}$$

$$= \epsilon_{x_0}\,, \tag{C.25}$$

where (C.22) follows because x_0 is a unit of A, (C.23) is by (C.2), and (C.24) is by (C.20) and (C.21). ■

The construction of T_x required in Theorem 117 is feasible for a large class of channels. For example, for an L-ary phase-shift-keying (PSK) modulated complex AWGN channel with soft decisions, we can assume that the input alphabet is $\{e^{j\frac{2\pi k}{L}}, k = 0, L-1\}$; then

$$T_x(y) = yx \tag{C.26}$$

satisfies the requirements because $P_{Y|X}(y|x')$ depends only on $|y{-}x'|$ and $|yx{-}x'x| = |y{-}x'|$.

We give a general result for constructing T_x.

Theorem 118 *Suppose that* B *is a monoid,* $\mathsf{A} \subset \mathsf{B}$ *is a group (in particular* A *consists of only invertible elements of* B*) and the channel is*

$$Y = N \circ X \tag{C.27}$$

with $N \in \mathsf{B}$ *being independent of* $X \in \mathsf{A}$*. If each* $T_x(y) = y \circ x$ *is measurable, then this family satisfies the conditions of Theorem 117.*

Proof: Indeed, for any $E \subset \mathsf{B}$ we have

$$T_x^{-1}E = E \circ x^{-1}\,. \tag{C.28}$$

Then, on the one hand

$$P_{Y|X=x'\circ x}[E] = P_N[E \circ (x' \circ x)^{-1}]\,, \tag{C.29}$$

but on the other hand,

$$\begin{aligned} P_{Y|X=x'}[T_x^{-1}E] &= P_{Y|X=x'}[E \circ x^{-1}] && \text{(C.30)} \\ &= P_N[E \circ x^{-1} \circ x'^{-1}]\,. && \text{(C.31)} \end{aligned}$$

■

It is easy to see that if we take $\mathcal{A} = \mathbb{Z}_2$ and $\mathsf{A} = \mathcal{A}^n$ then the BSC (even if the noise has memory) satisfies the conditions of Theorem 118. For the BEC we take $\mathcal{A} = \{-1, 1\}$ and $\mathcal{B} = \{-1, 0, 1\}$, and the usual multiplication of reals converts $\mathcal{B}$ to a monoid; taking the usual product – $\mathsf{A} = \mathcal{A}^n$ and $\mathsf{B} = \mathcal{B}^n$ – we see that the BEC (even with memory) also satisfies the conditions of Theorem 118. Similar generalizations are possible for any additive noise channel with erasures.

Appendix D

Energy efficient codes with feedback

Proof of Theorem 84: Consider an arbitrary (E, M, ϵ) code with feedback, namely a sequence of encoder functions $\{f_n\}_{n=1}^{\infty}$ and a decoder map $g : \mathsf{B} \to \{1, \ldots, M\}$. The "meta-converse" part of the proof proceeds step by step as in the non-feedback case (4.298)-(4.305), with the exception that measures $P^j = P_{\mathbf{y}|W=j}$ on B are defined as

$$P^j = \prod_{k=1}^{\infty} \mathcal{N}(f_k(j, Y_1^{k-1}), \tfrac{1}{2}N_0) \tag{D.1}$$

and β_α is replaced by $\tilde{\beta}_\alpha$, which is a unique solution $\tilde{\beta} < \alpha$ of

$$\tilde{\beta}_\alpha : \quad d(\alpha||\tilde{\beta}) = \frac{E}{N_0} \log e \,. \tag{D.2}$$

We need only to show that (4.300) holds with these modifications, i.e. for any $\alpha \in [0, 1]$

$$\inf_{F \subset \mathsf{B}: P^j(F) \geq \alpha} \Phi(F) \geq \tilde{\beta}_\alpha \,. \tag{D.3}$$

Once $W = j$ is fixed, channel inputs X_k become functions on the space B defined as $X_k = f_k(j, Y_1^{k-1})$. To find the critical set F achieving the infimum in the hypothesis testing problem (D.3) we compute the Radon-Nikodym derivative:

$$R \stackrel{\triangle}{=} \log_e \frac{dP^j}{d\Phi} = \sum_{k=1}^{\infty} X_k Y_k - \tfrac{1}{2} X_k^2 \,. \tag{D.4}$$

In general, the infimum in (D.3) depends on the choice of functions X_k. However, to prove a lower bound we may optimize over the choice of X_k in addition to optimizing over the choice of the critical region F. The key simplification comes from identifying the noise random variables Z_k with increments of the Wiener process.

Formally, define a standard Wiener process W_t with the filtration $\{\mathcal{F}_t\}_{t \in [0,\infty)}$ and two Brownian motions:

$$B_t = \frac{t}{2} + \sqrt{\tfrac{N_0}{2}} W_t \,, \tag{D.5}$$

$$\bar{B}_t = -\frac{t}{2} + \sqrt{\tfrac{N_0}{2}} W_t \,. \tag{D.6}$$

Then we can see that under P^j we have $Y_k = X_k + Z_k$ and hence we can assume

$$X_k Y_k - \tfrac{1}{2} X_k^2 = B_{\tau_k} - B_{\tau_{k-1}}\,, \tag{D.7}$$

where we have denoted the random instants

$$\tau_k = \sum_{m=1}^{k} X_m^2\,. \tag{D.8}$$

Then the distribution of R, under P^j is

$$R \sim B_\tau\,, \tag{D.9}$$

where the random variable τ is defined as

$$\tau = \sum_{k=1}^{\infty} X_k^2\,. \tag{D.10}$$

Similarly, under Φ we have

$$R \sim \bar{B}_\tau\,. \tag{D.11}$$

Note that without loss of generality $X_k \neq 0$ since having $X_k = 0$ does not help in discriminating P^j vs. Φ. Then each Y_k can be recovered from $X_k Y_k - \frac{1}{2} X_k^2$ since X_k is known. Consequently, each X_k is a function of only the past observations $(B_0, B_{\tau_1}, \ldots, B_{\tau_{k-1}})$. This implies that each τ_k, and thus τ, is a stopping time of the filtration $\mathcal{F}_t$ satisfying (under P^j)

$$\mathbb{E}_{P^j}[\tau] \leq E \tag{D.12}$$

by the energy constraint. In other words, as far as the problem (D.3) is concerned, the choice of a sequence of functions $X_k = f_k(Y_1^{k-1})$ satisfying an average power constraint amounts to specifying an increasing collection of stopping times τ_k of a Brownian motion B_t such that the limit τ satisfies (D.12).

To summarize, the encoder maps $\{f_n\}_{n=1}^{\infty}$ and the minimizing set F in (D.3) define a sequential hypothesis test, namely a stopping time τ and a decision region $F \in \mathcal{F}_\tau$, for discriminating between a Brownian motion with a positive drift B_t (under P) and a Brownian motion with a negative drift $\bar{B}_t$ (under Φ). According to Shiryaev [117, Section 4.2], among all (τ, F) satisfying (D.12) and having $P(F) \geq \alpha$ there exists an optimal one achieving[1]

$$\Phi(F) = \tilde{\beta}_\alpha\,, \tag{D.13}$$

where $\tilde{\beta}_\alpha$ is defined in (D.2). Any other test (τ, F) has a larger value of $\Phi(F)$, which proves (D.3). ■

[1] If instead of (4.292) we impose the maximum energy constraint: $||\mathbf{x}||^2 \leq E$ (a.s.), then $\tau \leq E$ and hence instead of $F \in \mathcal{F}_\tau$ we would have $F \in \mathcal{F}_E$, thus obtaining a usual, fixed-sample-size, binary hypothesis test. It is not hard to see that then $\tilde{\beta}_\alpha$ should be replaced with β_α from (4.301). Consequently, such an energy constraint renders feedback useless because the non-feedback converse (4.322) then holds. This parallels the result of Wyner [118] on block coding for the AWGN channel with fixed rate.

Proof of Theorem 85: Fix a list of elements $(\mathbf{c}_1, \ldots, \mathbf{c}_M) \in \mathsf{A}^M$ to be chosen later; $||\mathbf{c}_j||^2$ need not be finite. Upon receiving channel outputs $Y_1, \ldots, Y_n$ the decoder computes the likelihood $S_{j,n}$ for each codeword $j = 1, \ldots, M$, cf. (4.306) and (D.4):

$$S_{j,n} = \sum_{k=1}^{n} C_{j,k} Y_k - \tfrac{1}{2} C_{j,k}^2 \,, \quad j = 1, \ldots, M \,. \tag{D.14}$$

Fix two scalars $\gamma_0 < 0 < \gamma_1$ and define M stopping times

$$\tau_j = \inf\{n > 0 : S_{j,n} \notin (\gamma_0, \gamma_1)\} \,. \tag{D.15}$$

The decoder output $\hat{W}$ is the index j of the process $S_{j,n}$ that is the first to upcross γ_1 without having downcrossed γ_0 previously. The encoder conserves energy by transmitting only up until time τ_j (when the true message $W = j$):

$$X_n \stackrel{\triangle}{=} f_n(j, Y_1^{n-1}) = C_{j,n} 1\{\tau_W \geq n\} \,. \tag{D.16}$$

To complete the construction of the encoder-decoder pair we need to choose $(\mathbf{c}_1, \ldots, \mathbf{c}_M)$. This is done by a random-coding argument. Fix $d > 0$ and generate each $\mathbf{c}_j$ independently with equiprobable antipodal coordinates:

$$\mathbb{P}[C_{j,k} = +d] = \mathbb{P}[C_{j,k} = -d] = \frac{1}{2} \,, \; j = 1, \ldots, M. \tag{D.17}$$

We now upper-bound the probability of error P_e averaged over the choice of the codebook. By symmetry it is sufficient to analyze the probability $\mathbb{P}[\hat{W} \neq 1 | W = 1]$. We then have

$$\mathbb{P}[\hat{W} \neq 1 | W = 1] \leq \mathbb{P}[S_{1,\tau_1} \leq \gamma_0 | W = 1] + \sum_{j=2}^{M} \mathbb{P}[S_{j,\tau_j} \geq \gamma_1, \tau_j \leq \tau_1 | W = 1] \,, \tag{D.18}$$

because there are only two error mechanisms: S_1 downcrosses γ_0 before upcrossing γ_1, or some other S_j upcrosses γ_1 before S_1. Notice that in computing probabilities $\mathbb{P}[S_{1,\tau_1} \leq \gamma_0 | W = 1]$ and $\mathbb{P}[S_{2,\tau_2} \geq \gamma_1, \tau_2 \leq \tau_1 | W = 1]$ on the right-hand side of (D.18) we are interested only in time instants $0 \leq n \leq \tau_1$. For all such moments $X_n = C_{j,n}$. Therefore, below for simplicity of notation we will assume that $X_n = C_{j,n}$ for all n (whereas in reality $X_n = 0$ for all $n > \tau_1$, which is relevant only for calculating the total energy spent).

We define B_t and $\bar{B}_t$ as in (D.5) and (D.6); then conditioned on $W = 1$ the process S_1 can be rewritten as

$$S_{1,n} = B_{nd^2} \,, \tag{D.19}$$

because according to (D.18) we are interested only in $0 \leq n \leq \tau_1$ and thus $X_k = C_{j,k}$. The stopping time τ_1 then becomes

$$d^2 \tau_1 = \inf\{t > 0 : B_t \notin (\gamma_0, \gamma_1) \,, t = nd^2, n \in \mathbb{Z}\} \,. \tag{D.20}$$

If we now define

$$\tau = \inf\{t > 0 : B_t \notin (\gamma_0, \gamma_1)\} \,, \tag{D.21}$$

$$\bar{\tau} = \inf\{t > 0 : \bar{B}_t \notin (\gamma_0, \gamma_1)\} \,, \tag{D.22}$$

then the path-continuity of B_t implies that

$$d^2\tau_1 \searrow \tau \text{ as } d \to 0\,. \tag{D.23}$$

Similarly, still under the condition $W = 1$ we can rewrite

$$S_{2,n} = d^2 \sum_{k=1}^{n} L_k + \bar{B}_{nd^2}\,, \tag{D.24}$$

where L_k are i.i.d., independent of $\bar{B}_t$ and

$$\mathbb{P}[L_k = +1] = \mathbb{P}[L_k = -1] = \frac{1}{2}\,. \tag{D.25}$$

Extending (D.23), we will show below that as $d \to 0$ we have

$$\mathbb{P}[S_{1,\tau_1} \le \gamma_0 | W = 1] \;\; \to \;\; 1 - \alpha(\gamma_0, \gamma_1)\,, \tag{D.26}$$

$$\mathbb{P}[S_{2,\tau_2} \ge \gamma_1, \tau_2 < \infty | W = 1] \;\; \to \;\; \beta(\gamma_0, \gamma_1)\,, \tag{D.27}$$

where $\alpha(\gamma_0, \gamma_1)$ and $\beta(\gamma_0, \gamma_1)$ are

$$\alpha(\gamma_0, \gamma_1) \;\; = \;\; \mathbb{P}[B_\tau = \gamma_1]\,, \tag{D.28}$$

$$\beta(\gamma_0, \gamma_1) \;\; = \;\; \mathbb{P}[\bar{B}_{\bar{\tau}} = \gamma_1, \bar{\tau} < \infty]\,, \tag{D.29}$$

i.e. the probabilities of exiting the interval (γ_0, γ_1) through the upper-boundary by B_t and $\bar{B}_t$, respectively[2] . Thus, the interval (γ_0, γ_1) determines the boundaries of the sequential probability ratio test. As shown by Shiryaev [117, Section 4.2], α and β satisfy

$$d(\alpha(\gamma_0, \gamma_1)||\beta(\gamma_0, \gamma_1)) = \frac{\log e}{N_0}\mathbb{E}\,[\tau]\,. \tag{D.30}$$

Assuming (D.26) and (D.27) as $d \to 0$ the probability of error is upper-bounded by (D.18):

$$\mathbb{P}[\hat{W} \ne 1 | W = 1] \le 1 - \alpha(\gamma_0, \gamma_1) + (M-1)\beta(\gamma_0, \gamma_1)\,. \tag{D.31}$$

At the same time, the average energy spent by our scheme is

$$\lim_{d\to 0} \mathbb{E}\,[||\mathbf{x}||^2] = \lim_{d\to 0} \mathbb{E}\,[d^2\tau_1] = \mathbb{E}\,[\tau]\,, \tag{D.32}$$

because of (D.23).

Finally, comparing (4.325) and (D.30) it follows that optimizing (D.31) over all $\gamma_0 < 0 < \gamma_1$ satisfying $\mathbb{E}\,[\tau] = E$ we obtain (4.324). To prove (4.326) simply notice that when $\alpha = 1$ we have $\gamma_0 = -\infty$, and hence the decision is taken by the decoder the first time any S_j upcrosses γ_1. Therefore, in the encoder rule (D.16) the time τ_j, whose computation requires the full knowledge of Y_k, can be replaced with the time of decoder decision, which requires sending only a single signal. Obviously, this modification will not change the probability

[2]The condition $\bar{\tau} < \infty$ is required for handling the special case $\gamma_0 = -\infty$.

of error and will conserve energy even more (since under $\gamma_0 = -\infty$, τ_j cannot occur before the decision time).

We now prove (D.26) and (D.27). By (D.19) and (D.23) we have

$$S_{1,\tau_1} = B_{d^2\tau_1} \to B_\tau \,, \tag{D.33}$$

because of the continuity of B_t. From (D.33) we obtain (D.26) after noticing that again due to continuity

$$\mathbb{P}[B_\tau \le \gamma_0] = 1 - \mathbb{P}[B_\tau \ge \gamma_1] = 1 - \mathbb{P}[B_\tau = \gamma_1] \,. \tag{D.34}$$

The proof of (D.27) requires a slightly more intricate argument for which it is convenient to introduce a probability space denoted by $(\Omega, \mathcal{H}, \mathbb{P})$ which is the completion of the probability space generated by $\{\bar{B}_t\}_{t=0}^{\infty}$ and $\{L_k\}_{k=1}^{\infty}$ defined in (D.6) and (D.25), respectively. For each $0 < d \le 1$ we define the following random variables, where their explicit dependence on d is omitted for brevity

$$\begin{aligned}
D_t &= d^2 \sum_{k \le \lfloor t/d^2 \rfloor} L_k \,, && \text{(D.35)}\\
\Sigma_t &= D_t + \bar{B}_{d^2 \left\lfloor \frac{t}{d^2} \right\rfloor} \,, && \text{(D.36)}\\
\tau_2 &= \inf\{t > 0 : \Sigma_t \notin (\gamma_0, \gamma_1)\} \,, && \text{(D.37)}\\
\bar{\tau} &= \inf\{t > 0 : B_t \notin (\gamma_0, \gamma_1)\} \,. && \text{(D.38)}
\end{aligned}$$

In comparison with the random variables appearing in (D.27) Σ_{nd^2} and τ_2 take the role of $S_{2,n}$ and $d^2\tau_2$, respectively; and also $\mathbb{P}$ henceforth is already normalized by the conditioning on $W = 1$. Thus in the new notation we need to prove

$$\lim_{d \to 0} \mathbb{P}[\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty] = \mathbb{P}[\bar{B}_{\bar{\tau}} = \gamma_1, \bar{\tau} < \infty] \,. \tag{D.39}$$

We define the following subsets of Ω:

$$\begin{aligned}
E_0 &= \{\omega \in \Omega : \exists T < \infty \ \forall t > T : \sup_{0 < d \le 1} \Sigma_t < 0\} \,, && \text{(D.40)}\\
E_1 &= \{\bar{\tau} = \infty\} \cup \{\bar{\tau} < \infty, \forall \epsilon > 0 \ \exists t_1, t_2 \in (0, \epsilon) \text{ s.t. } \bar{B}_{\bar{\tau}+t_1} > \bar{B}_{\bar{\tau}}, \bar{B}_{\bar{\tau}+t_2} < \bar{B}_{\bar{\tau}}\} \,, && \text{(D.41)}\\
E_2 &= \{\omega \in \Omega : \lim_{d \to 0} D_t = 0 \text{ uniformly on compacts}\} \,, && \text{(D.42)}\\
E &= E_0 \cap E_1 \cap E_2 \,. && \text{(D.43)}
\end{aligned}$$

According to Lemma 119 the sets in (D.40)-(D.43) belong to $\mathcal{H}$ and have probability 1.

The next step is to show

$$\{\bar{B}_{\bar{\tau}} = \gamma_1, \bar{\tau} < \infty\} \cap E \subset \liminf_{d \to 0} \{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\} \cap E \,. \tag{D.44}$$

To that end select an arbitrary element $\omega \in \{\bar{B}_{\bar{\tau}} = \gamma_1, \bar{\tau} < \infty\} \cap E$. Since $\bar{B}_t$ is continuous it must attain its minimum b_0 on $[0; \bar{\tau}]$; of course, $b_0 > \gamma_0$. Again, due to continuity of $\bar{B}_t$ at $t = \bar{\tau}$ there must exist an $\epsilon_1 > 0$ such that

$$b_0' \stackrel{\triangle}{=} \min_{0 \le t \le \bar{\tau} + \epsilon_1} \bar{B}_t > \gamma_0 \,. \tag{D.45}$$

On the other hand, because $\omega \in E_1$ we have

$$b_1 \stackrel{\triangle}{=} \max_{0 \le t \le \bar{\tau}+\epsilon_1} \bar{B}_t > \gamma_1 \,. \tag{D.46}$$

Moreover, since $\omega \in E_2$ we have $D_t \to 0$ uniformly on $[0; \bar{\tau} + \epsilon_1]$; therefore, there exists a $d_1 > 0$ such that for all $d \le d_1$ we have

$$\sup_{t \in [0; \bar{\tau}+\epsilon_1]} |D_t| \le \epsilon_2 \,, \tag{D.47}$$

where

$$\epsilon_2 = \frac{1}{3} \min(b_1 - \gamma_1, b_0' - \gamma_0) > 0 \,. \tag{D.48}$$

If we denote by t_1 the point at which $B_{t_1} = b_1$, then by continuity of B_t at t_1 there exists a $\delta > 0$ such that

$$\forall t \in (t_1 - \delta; t_1 + \delta) : B_t > b_1 - \epsilon_2 \,. \tag{D.49}$$

Then for every $d < \sqrt{\delta}$ we have

$$\max_{t \in [0, \bar{\tau}+\epsilon_1]} \bar{B}_{d^2 \left\lfloor \frac{t}{d^2} \right\rfloor} > b_1 - \epsilon_2 \,. \tag{D.50}$$

Finally, for every $d \le \min(\sqrt{\delta}, d_1)$ we have

$$\sup_{t \in [0, \bar{\tau}+\epsilon_1]} \Sigma_t \ge b_1 - 2\epsilon_2 > \gamma_1 \tag{D.51}$$

and

$$\inf_{t \in [0, \bar{\tau}+\epsilon_1]} \Sigma_t \ge b_0' - \epsilon_2 > \gamma_0 \tag{D.52}$$

by (D.45), (D.46) (D.48) and (D.50). Then of course, (D.51) and (D.52) prove that $\tau_2 \le \bar{\tau} + \epsilon_1$ and $\{\Sigma_{\tau_2} \ge \gamma_1\}$ holds for all $d \le \min(\sqrt{\delta}, d_1)$. Equivalently,

$$\omega \in \liminf_{d \to 0} \{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\} \,, \tag{D.53}$$

proving (D.44).

Next, we show

$$\limsup_{d \to 0} \{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\} \cap E \subset \{\bar{B}_{\bar{\tau}} = \gamma_1, \bar{\tau} < \infty\} \cap E \,. \tag{D.54}$$

Indeed, take $\omega \in \limsup_{d \to 0} \{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\} \cap E$, that is a point in the sample space for which there exists a subsequence $d_l \to 0$ such that $\Sigma_{\tau_2} \ge \gamma_1$ for every l. Since $\omega \in E_0$ we know that for all d we have $\tau_2(\omega) \le T < \infty$. First, we show

$$b_1 \stackrel{\triangle}{=} \max_{0 \le t \le T} \bar{B}_t \ge \gamma_1 \,. \tag{D.55}$$

Indeed, assuming otherwise and repeating with minor changes the argument leading from (D.46) to (D.51), we can show that in this case

$$\sup_{t \in [0,T]} \Sigma_t < \gamma_1 \tag{D.56}$$

for all sufficiently small d. This contradicts the choice of ω.

We denote

$$t_1 = \inf\{t > 0 : \bar{B}_t = b_1\}\,. \tag{D.57}$$

Then (D.55) and continuity of $\bar{B}_t$ imply

$$\bar{\tau} \le t_1 < \infty\,. \tag{D.58}$$

We are only left to show that $\bar{B}_{\bar{\tau}} = \gamma_0$ is impossible. If it were so, then $\bar{\tau} < t_1 < T$. Moreover because $\omega \in E_2$ there must exist an $\epsilon_1 > 0$ (similar to (D.45) and (D.46)) such that

$$b_0' \stackrel{\triangle}{=} \min_{0\le t\le \bar{\tau}+\epsilon_1} \bar{B}_t < \gamma_0\,, \tag{D.59}$$

and

$$b_1' \stackrel{\triangle}{=} \max_{0\le t\le \bar{\tau}+\epsilon_1} \bar{B}_t < \gamma_1\,. \tag{D.60}$$

Thus, by repeating the argument behind (D.51) and (D.52) we can show that for all sufficiently small d we have

$$\sup_{t\in[0,\bar{\tau}+\epsilon_1]} \Sigma_t < \gamma_1\,, \tag{D.61}$$

and

$$\inf_{t\in[0,\bar{\tau}+\epsilon_1]} \Sigma_t < \gamma_0\,, \tag{D.62}$$

which contradicts the assumption that $\omega \in \limsup_{d\to 0}\{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\}$.

Together (D.44) and (D.54) prove that

$$\begin{aligned}\{\bar{B}_{\bar{\tau}} = \gamma_1, \bar{\tau} < \infty\} \cap E \subset \liminf_{d\to 0}\{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\} \cap E \subset \\ \limsup_{d\to 0}\{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\} \cap E \subset \{\bar{B}_{\bar{\tau}} = \gamma_1, \bar{\tau} < \infty\} \cap E\,,\end{aligned} \tag{D.63}$$

which implies that all three sets are equal. By Lemma 119 and completeness of $\mathcal{H}$ both sets $\liminf_{d\to 0}\{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\}$ and $\limsup_{d\to 0}\{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\}$ are measurable and computing their probabilities is meaningful. Finally, we have

$$\begin{aligned}\lim_{d\to 0} \mathbb{P}[\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty] &= \lim_{d\to 0} \mathbb{P}[\{\Sigma_{\tau_2} \ge \gamma_1, \tau_2 < \infty\} \cap E] &\text{(D.64)}\\ &= \mathbb{P}[\bar{B}_{\bar{\tau}} = \gamma_1, \bar{\tau} < \infty]\,, &\text{(D.65)}\end{aligned}$$

where (D.64) is by Lemma 119 and (D.65) by (D.63) and bounded convergence theorem. ■

Lemma 119 *Set E defined in* (D.43) *is $\mathcal{H}$-measurable and*

$$\mathbb{P}[E] = 1\,. \tag{D.66}$$

Proof: By completeness of $\mathcal{H}$ it is sufficient to prove that all sets E_0, E_1 and E_2 contain a measurable subset of probability 1. To prove

$$\mathbb{P}[E_0] = 1\,, \tag{D.67}$$

notice that

$$\sup_{0<d\le 1} D_t = t \sup_{N\ge t} \frac{1}{N}\sum_{k=1}^{N} L_k\,, \tag{D.68}$$

and therefore, by the Chernoff bound,

$$\begin{aligned}\mathbb{P}\left[\sup_{0<d\le 1} D_t > \frac{t}{4}\right] &\le \sum_{N\ge t} O\left(e^{-a_1 N}\right) &\text{(D.69)}\\ &= O\left(e^{-a_1 t}\right)\,, &\text{(D.70)}\end{aligned}$$

for some constant $a_1 > 0$. Hence, for an arbitrary t we have an estimate

$$\begin{aligned}\mathbb{P}[\bar{B}_t + \sup_{0<d\le 1} D_t \ge -1] &\le \mathbb{P}\left[\bar{B}_t \ge -1 - \frac{t}{4}\right] + \mathbb{P}\left[\sup_{0<d\le 1} D_t > \frac{t}{4}\right] &\text{(D.71)}\\ &\le O\left(e^{-a_1 t}\right)\,, &\text{(D.72)}\end{aligned}$$

where (D.72) is because $\bar{B}_t \sim \mathcal{N}\left(-\frac{t}{2}, \frac{tN_0}{2}\right)$ and (D.70).

Next, denote

$$\begin{aligned}\delta_j &= \frac{1}{\sqrt{j}}\,, &\text{(D.73)}\\ t_n &= \sum_{j=1}^{n}\delta_j\,, &\text{(D.74)}\\ M_j &= \max_{t_j\le t\le t_{j-1}} W_t - W_{t_j}\,, &\text{(D.75)}\end{aligned}$$

where $W_t = t/2 + \sqrt{\frac{2}{N_0}}\bar{B}_t$ is the standard Wiener process; cf. (D.6).

Since $t_n \sim 2\sqrt{n}$ and the series $\sum_{n=1}^{\infty} e^{-a_1\sqrt{n}}$ converges, we can apply the Borel-Cantelli lemma via (D.72) to show that

$$F_1 = \left\{\{B_{t_n} + \sup_{0<d\le 1} D_{t_n} \ge -1\} \text{ –infinitely often}\right\} \tag{D.76}$$

has measure zero. Similarly, since $M_j \sim |W_{\delta_j}|$ we have

$$\sum_{j=1}^{\infty}\mathbb{P}[M_j > (2N_0)^{-1}] = \sum_{j=1}^{\infty} 2Q\left(\frac{1}{2N_0\sqrt{\delta_j}}\right) \le a_3\sum_{j=1}^{\infty} e^{-a_2\sqrt{n}} < \infty\,, \tag{D.77}$$

for some positive constants a_2, a_3. And therefore,

$$F_2 = \left\{M_j > (2N_0)^{-1} \text{ –infinitely often}\right\} \tag{D.78}$$

also has measure zero. Finally we show that

$$F_1^c \cap F_2^c \subset E_0\,. \tag{D.79}$$

Indeed, for all $t \in [t_j; t_j + \delta_j)$ we have

$$\bar{B}_t + D_t \le \bar{B}_{t_j} + D_{t_j} + \sqrt{\frac{N_0}{2}} M_j + 2\delta_j \,, \tag{D.80}$$

because, from the definition of D_t,

$$|D_{s_1} - D_{s_2}| \le 2|s_1 - s_2| \,, \tag{D.81}$$

for all $d > 0$. From (D.80) for any $\omega \in F_1^c \cap F_2^c$ we have for all sufficiently large t

$$\sup_{0<d\le 1} \bar{B}_t + D_t \le -1 + \frac{1}{2} + 2\delta_j \,, \tag{D.82}$$

where j denotes the index of the unique interval $t \in [t_j; t_{j+1})$. Therefore, for all sufficiently large t we have shown

$$\sup_{0<d\le 1} \Sigma_t \le \sup_{0<d\le 1} \bar{B}_t + D_t < 0 \,, \tag{D.83}$$

completing the proof of (D.79) and, hence, of (D.67).

To show $\mathbb{P}[E_1] = 1$ notice that by the strong Markov property of Brownian motion for any finite stopping time σ according to Blumenthal's zero-one law [119] for

$$F_\sigma = \{\forall \epsilon > 0 \; \exists t_1, t_2 \in (0, \epsilon) \text{ s.t. } \bar{B}_{\sigma+t_1} > \bar{B}_\sigma, \bar{B}_{\sigma+t_2} < \bar{B}_\sigma\} \tag{D.84}$$

we have

$$\mathbb{P}[F_\sigma] = 1 \,. \tag{D.85}$$

Since $\sigma_n = \min(\bar{\tau}, n)$ are finite stopping times and $\sigma_n \nearrow \bar{\tau}$, we have

$$E_1 \supset \bigcap_{n=1}^{\infty} F_{\sigma_n} \,. \tag{D.86}$$

Therefore, $\mathbb{P}[E_1] = 1$ since $\mathbb{P}[F_{\sigma_n}] = 1$ for all $n \ge 1$.

To show

$$\mathbb{P}[E_2] = 1 \tag{D.87}$$

it is sufficient to show for every integer $K > 0$

$$\mathbb{P}[\lim_{d\to 0} D_t = 0 \text{ uniformly on} [0; K]] = 1 \tag{D.88}$$

and take intersection of such sets over all $K \in \mathbb{Z}_+$. To prove (D.88) notice that

$$\begin{aligned}
\mathbb{P}[\limsup_{d\to 0} \sup_{0\le t\le K} |D_t| \ge \epsilon] &= \mathbb{P}\left[\limsup_{d\to 0} d^2 \max_{0\le n\le \frac{K}{d^2}} \left|\sum_{k=0}^{n} L_k\right| \ge \epsilon\right] && \text{(D.89)}\\
&= \mathbb{P}\left[\limsup_{m\to\infty} \frac{K}{m} \max_{0\le n\le m} \left|\sum_{k=1}^{n} L_k\right| \ge \epsilon\right] && \text{(D.90)}\\
&\le \mathbb{P}\left[\frac{1}{n}\sum_{k=1}^{n} L_k \ge \frac{\epsilon}{K} \text{ –i.o.}\right] + \mathbb{P}\left[\frac{1}{n}\sum_{k=1}^{n} L_k \le -\frac{\epsilon}{K} \text{ –i.o.}\right], && \text{(D.91)}
\end{aligned}$$

where "i.o." stands for infinitely often. By the strong law of large numbers both probabilities in (D.91) are zero and we obtain

$$\limsup_{d\to 0} \sup_{0\le t\le K} |D_t| = 0 \quad \text{a.s.}\,, \tag{D.92}$$

which is equivalent to (D.88). ■

Appendix E

Gilbert-Elliott channel: proofs

In this appendix we prove Theorems 58, 59 and 60 stated in Section 3.5.

Proof of Theorem 58: Achievability: We choose P_{X^n} – equiprobable. To model the availability of the state information at the receiver, we assume that the output of the channel is (Y^n, S^n). Thus we need to write down the expression for $i(X^n; Y^n S^n)$. To do that we define an operation on $\mathbb{R} \times \{0, 1\}$:

$$a^{\{b\}} = \begin{cases} 1-a\,, & b=0\,, \\ a\,, & b=1 \end{cases}\,. \tag{E.1}$$

Then we obtain

$$\begin{aligned} i(X^n; Y^n S^n) &= \log \frac{P_{Y^n|X^n S^n}(Y^n|X^n, S^n)}{P_{Y^n|S^n}(Y^n|S^n)} && \text{(E.2)} \\ &= n \log 2 + \sum_{j=1}^{n} \log \delta_{S_j}^{\{Z_j\}}\,, && \text{(E.3)} \end{aligned}$$

where (E.2) follows since $P_{S^n|X^n}(s^n|x^n) = P_{S^n}(s^n)$ by independence of X^n and S^n, (E.3) is because under equiprobable X^n we have that $P_{Y^n|S^n}$ is also equiprobable, while $P_{Y_j|X_j S_j}(Y_j|X_j, S_j)$ is equal to $\delta_{S_j}^{\{Z_j\}}$ with Z_j defined in (3.369). Using (E.3) we find

$$\mathbb{E}\,[i(X^n; Y^n S^n)] = nC_1\,. \tag{E.4}$$

The next step is to compute $\mathrm{Var}[i(X^n; Y^n S^n)]$. For convenience we write

$$h_a = \frac{1}{2}[h(\delta_1) + h(\delta_2)] \tag{E.5}$$

and

$$\Theta_j = \log \delta_{S_j}^{\{Z_j\}}\,. \tag{E.6}$$

Therefore

$$\begin{aligned}
\sigma_n^2 &\stackrel{\triangle}{=} \mathrm{Var}[i(X^n;Y^nS^n)] && \text{(E.7)}\\
&= \mathbb{E}\left[\left(\sum_{j=1}^n \Theta_j\right)^2\right] - n^2h_a^2 && \text{(E.8)}\\
&= \sum_{j=1}^n \mathbb{E}\left[\Theta_j^2\right] + 2\sum_{i<j}\mathbb{E}\left[\Theta_i\Theta_j\right] - n^2h_a^2 && \text{(E.9)}\\
&= n\mathbb{E}\left[\Theta_1^2\right] + 2\sum_{k=1}^n (n-k)\mathbb{E}\left[\Theta_1\Theta_{1+k}\right] - n^2h_a^2 && \text{(E.10)}\\
&= n(\mathbb{E}\left[\Theta_1^2\right] - h_a^2)\\
&\quad +2\sum_{k=1}^n (n-k)\mathbb{E}\left[h\left(\delta_{S_1}\right)h\left(\delta_{S_{1+k}}\right) - h_a^2\right], && \text{(E.11)}
\end{aligned}$$

where (E.10) follows by stationarity and (E.11) by conditioning on S^n and regrouping terms.

Before proceeding further we define an α-mixing coefficient of the process (S_j, Z_j) as

$$\alpha(n) = \sup|\mathbb{P}[A,B] - \mathbb{P}[A]\mathbb{P}[B]|\,, \qquad \text{(E.12)}$$

where the supremum is over $A \in \sigma\{S_{-\infty}^0, Z_{-\infty}^0\}$ and $B \in \sigma\{S_n^\infty, Z_n^\infty\}$; by $\sigma\{\cdots\}$ we denote a σ-algebra generated by a collection of random variables. Because S_j is such a simple Markov process it is easy to show that for any $a, b \in \{1,2\}$ we have

$$\frac{1}{2} - \frac{1}{2}|1-2\tau|^n \le \mathbb{P}[S_n = a|S_0 = b] \le \frac{1}{2} + \frac{1}{2}|1-2\tau|^n\,, \qquad \text{(E.13)}$$

and, hence,

$$\alpha(n) \le |1-2\tau|^n\,. \qquad \text{(E.14)}$$

By Lemma 1.2 of [68] for any pair of bounded random variables U and V measurable with respect to $\sigma\{S_j, j \le m\}$ and $\sigma\{S_j, j \ge m+n\}$, respectively, we have

$$|\mathbb{E}\left[UV\right] - \mathbb{E}\left[U\right]\mathbb{E}\left[V\right]| \le 16\alpha(n)\cdot \text{ess sup}\,|U|\cdot\text{ess sup}\,|V|\,. \qquad \text{(E.15)}$$

Then we can conclude that since $|h\left(\delta_{S_1}\right)| \le \log 2$ we have for some constant B_3

$$\begin{aligned}
&\left|\sum_{k=1}^n k\mathbb{E}\left[h\left(\delta_{S_1}\right)h\left(\delta_{S_{1+k}}\right) - h_a^2\right]\right|\\
&\quad\le \sum_{k=1}^n k\mathbb{E}\left[\left|h\left(\delta_{S_1}\right)h\left(\delta_{S_{1+k}}\right) - h_a^2\right|\right] && \text{(E.16)}\\
&\quad\le \sum_{k=1}^n 16k\alpha(k)\log^2 2 && \text{(E.17)}\\
&\quad\le B_3\sum_{k=1}^\infty k(1-2\tau)^k && \text{(E.18)}\\
&\quad= O(1)\,, && \text{(E.19)}
\end{aligned}$$

where (E.17) is by (E.15) and (E.18) is by (E.30). On the other hand,

$$n\left|\sum_{k=n+1}^{\infty}\mathbb{E}\left[h\left(\delta_{S_1}\right)h\left(\delta_{S_{1+k}}\right)-h_a^2\right]\right| \tag{E.20}$$

$$\leq 16n\sum_{k=n+1}^{\infty}\alpha(k)\log^2 2 \tag{E.21}$$

$$\leq 16Kn\sum_{k=n+1}^{\infty}(1-2\tau)^k\log^2 2 \tag{E.22}$$

$$= O(1)\,. \tag{E.23}$$

Therefore, we have proved that

$$\sum_{k=1}^{n}(n-k)\mathbb{E}\left[h\left(\delta_{S_1}\right)h\left(\delta_{S_{1+k}}\right)-h_a^2\right] \tag{E.24}$$

$$= n\sum_{k=1}^{n}\mathbb{E}\left[h\left(\delta_{S_1}\right)h\left(\delta_{S_{1+k}}\right)-h_a^2\right]+O(1) \tag{E.25}$$

$$= n\sum_{k=1}^{\infty}\mathbb{E}\left[h\left(\delta_{S_1}\right)h\left(\delta_{S_{1+k}}\right)-h_a^2\right]+O(1)\,, \tag{E.26}$$

A straightforward calculation reveals that

$$\sum_{k=1}^{\infty}\mathbb{E}\left[h\left(\delta_{S_1}\right)h\left(\delta_{S_{1+k}}\right)-h_a^2\right] \tag{E.27}$$

$$= \frac{1}{4}\left(h\left(\delta_1\right)-h\left(\delta_2\right)\right)^2\left[\frac{1}{2\tau}-1\right]. \tag{E.28}$$

Therefore, using (E.26) and (E.28) in (E.11), we obtain after some algebra that

$$\sigma_n^2=\mathrm{Var}[i(X^n;Y^nS^n)]=nV_1+O(1)\,. \tag{E.29}$$

By (E.3) we see that $i(X^n;Y^nS^n)$ is a sum over an α-mixing process. For such sums the following theorem of Tikhomirov [120] serves the same purpose in this section as the Berry-Esseen inequality does in Sections 3.2.2, 3.3.2 and 3.4.4.

Theorem 120 *Suppose that a stationary zero-mean process $X_1, X_2, \ldots$ is α-mixing and for some positive K, β and γ we have*

$$\alpha(k) \leq Ke^{-\beta k}\,, \tag{E.30}$$

$$\mathbb{E}\left[|X_1|^{4+\gamma}\right] < \infty \tag{E.31}$$

$$\sigma_n^2 \to \infty\,, \tag{E.32}$$

where

$$\sigma_n^2=\mathbb{E}\left[\left(\sum_1^n X_j\right)^2\right]. \tag{E.33}$$

Then, there is a constant B, depending on K, β and γ, such that

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left[\sum_1^n X_j \geq x\sqrt{\sigma_n^2}\right] - Q(x)\right| \leq \frac{B\log n}{\sqrt{n}}\,. \tag{E.34}$$

Application of Theorem 120 to $i(X^n;Y^nS^n)$ proves that

$$\left|\mathbb{P}\left[i(X^n;Y^nS^n) \geq nC_1 + \sqrt{\sigma_n^2}x\right] - Q(x)\right| \;\leq\; \frac{B\log n}{\sqrt{n}}\,. \tag{E.35}$$

But then for arbitrary λ there exists some constant $B_2 > B$ such that we have

$$\begin{aligned}
&\left|\mathbb{P}\left[i(X^n;Y^nS^n) \geq nC_1 + \sqrt{nV_1}\lambda\right] - Q(\lambda)\right| && \text{(E.36)}\\
&= \left|\mathbb{P}\left[i(X^n;Y^nS^n) \geq nC_1 + \sqrt{\sigma_n^2}\sqrt{\frac{nV_1}{\sigma_n^2}}\lambda\right] - Q(\lambda)\right| && \text{(E.37)}\\
&\leq \frac{B\log n}{\sqrt{n}} + \left|Q(\lambda) - Q\left(\lambda\sqrt{\frac{nV_1}{\sigma_n^2}}\right)\right| && \text{(E.38)}\\
&= \frac{B\log n}{\sqrt{n}} + |Q(\lambda) - Q\left(\lambda + O(1/n)\right)| && \text{(E.39)}\\
&\leq \frac{B\log n}{\sqrt{n}} + O(1/n) && \text{(E.40)}\\
&\leq \frac{B_2\log n}{\sqrt{n}}\,, && \text{(E.41)}
\end{aligned}$$

where (E.38) is by (E.35), (E.39) is by (E.29) and (E.40) is by Taylor's theorem.

Now, we state the following extension of Lemma 20, to be proved later:

Lemma 121 *Let $X_1, X_2, \ldots$ be a process satisfying the conditions of Theorem 120; then for any constant A*

$$\mathbb{E}\left[\exp\left\{-\sum_{j=1}^n X_j\right\}\cdot 1\left\{\sum_{j=1}^n X_j > A\right\}\right] \leq 2\left(\frac{\log 2}{\sqrt{2\pi\sigma_n^2}} + \frac{2B\log n}{\sqrt{n}}\right)\exp\{-A\}\,, \tag{E.42}$$

where B is the constant in (E.34).

Observe that there exists some $B_1 > 0$ such that

$$\begin{aligned}
2\left(\frac{\log 2}{\sqrt{2\pi\sigma_n^2}} + \frac{2B\log n}{\sqrt{n}}\right) &= 2\left(\frac{\log 2}{\sqrt{2\pi(nV + O(1))}} + \frac{2B\log n}{\sqrt{n}}\right) && \text{(E.43)}\\
&\leq \frac{B_1\log n}{\sqrt{n}}\,, && \text{(E.44)}
\end{aligned}$$

where σ_n^2 is defined in (E.7) and (E.43) follows from (E.29). Therefore, from (E.44) we conclude that there exists a constant B_1 such that for any A

$$\mathbb{E}\left[\exp\{-i(X^n;Y^nS^n) + A\}\cdot 1\{i(X^n;Y^nS^n) \geq A\}\right] \leq \frac{B_1\log n}{\sqrt{n}}\,, \tag{E.45}$$

Finally, we set

$$\log \frac{M-1}{2} = nC - \sqrt{nV}Q^{-1}(\epsilon_n)\,, \tag{E.46}$$

where

$$\epsilon_n = \epsilon - \frac{(B_1+B_2)\log n}{\sqrt{n}}\,. \tag{E.47}$$

Then, by Theorem 18 we know that there exists a code with M codewords and average probability of error p_e bounded by

$$\begin{aligned} p_e &\le \mathbb{E}\left[\exp\left\{-\left[i(X^n;Y^nS^n) - \log\frac{M-1}{2}\right]^+\right\}\right] && \text{(E.48)}\\ &\le \mathbb{P}\left[i(X^n;Y^nS^n) \le \log\frac{M-1}{2}\right] + \frac{B_1}{\sqrt{n}} && \text{(E.49)}\\ &\le \epsilon_n + \frac{(B_1+B_2)\log n}{\sqrt{n}} && \text{(E.50)}\\ &\le \epsilon\,, && \text{(E.51)} \end{aligned}$$

where (E.49) is by (E.45) with $A = \log\frac{M-1}{2}$, (E.50) is by (E.41) and (E.46), and (E.51) is by (E.47). Therefore, invoking Taylor's expansion of Q^{-1} in (E.46) we have

$$\log M^*(n,\epsilon) \ge \log M \ge nC - \sqrt{nV}Q^{-1}(\epsilon) + O(\log n)\,. \tag{E.52}$$

This proves the achievability bound with the average probability of error criterion.

However, as explained in Appendix C, Theorem 18 by using a random linear code method can be strengthened to guarantee a codebook with a prescribed maximal probability of error. In this way, the above argument actually applies to both average and maximal error criteria after replacing $\log M$ by $\lfloor \log M \rfloor$, which is asymptotically immaterial.

Converse: In the converse part we will assume that the transmitter has access to the full state sequence S^n and then generates X^n based on both the input message and S^n. Take the best such code with $M^*(n,\epsilon)$ codewords and average probability of error no greater than ϵ. We now propose to treat the pair (X^n, S^n) as a combined input to the channel (but the S^n part is independent of the message) and the pair (Y^n, S^n) as a combined output, available to the decoder. Note that in this situation, the encoder induces a distribution $P_{X^nS^n}$ and is necessarily randomized because the distribution of S^n is not controlled by the input message and is given by the output of the Markov chain.

To apply Theorem 28 we choose the auxiliary channel which passes S^n unchanged and generates Y^n equiprobably:

$$Q_{Y^n|X^nS^n}(y^n, s^n|x^n) = 2^{-n} \quad \text{for all } x^n, y^n, s^n\,. \tag{E.53}$$

Note that by the constraint on the encoder, S^n is independent of the message W. Moreover, under Q-channel the Y^n is also independent of W and we clearly have

$$\epsilon' \ge 1 - \frac{1}{M^*}\,. \tag{E.54}$$

Therefore by Theorem 28 we obtain

$$\beta_{1-\epsilon}\left(P_{X^nY^nS^n}, Q_{X^nY^nS^n}\right) \le \frac{1}{M^*}\,. \tag{E.55}$$

To lower bound $\beta_{1-\epsilon}(P_{X^nY^nS^n}, Q_{X^nY^nS^n})$ via (2.67) we notice that

$$\begin{aligned}
\log \frac{P_{X^nY^nS^n}(x^n,y^n,s^n)}{Q_{X^nY^nS^n}(x^n,y^n,s^n)} &= \log \frac{P_{Y^n|X^nS^n}(y^n|x^n,s^n)P_{X^nS^n}(x^n,s^n)}{Q_{Y^n|X^nS^n}(y^n|x^n,s^n)Q_{X^nS^n}(x^n,s^n)} && \text{(E.56)}\\
&= \log \frac{P_{Y^n|X^nS^n}(y^n|x^n,s^n)}{Q_{Y^n|X^nS^n}(y^n|x^n,s^n)} && \text{(E.57)}\\
&= i(x^n;y^ns^n)\,, && \text{(E.58)}
\end{aligned}$$

where (E.57) is because $P_{X^nS^n} = Q_{X^nS^n}$ and (E.58) is simply by noting that $P_{Y^n|S^n}$ in the definition (E.2) of $i(X^n;Y^nS^n)$ is also equiprobable and, hence, is equal to $Q_{Y^n|X^nS^n}$. Now set

$$\log\gamma = nC - \sqrt{nV}Q^{-1}(\epsilon_n)\,, \tag{E.59}$$

where this time

$$\epsilon_n = \epsilon + \frac{B_2\log n}{\sqrt{n}} + \frac{1}{\sqrt{n}}\,. \tag{E.60}$$

By (2.67) we have for $\alpha = 1-\epsilon$ that

$$\begin{aligned}
\beta_{1-\epsilon} &\geq \frac{1}{\gamma}\left(1-\epsilon-\mathbb{P}\left[\log\frac{P_{X^nY^nS^n}(X^n,Y^n,S^n)}{Q_{X^nY^nS^n}(X^n,Y^n,S^n)}\geq\log\gamma\right]\right) && \text{(E.61)}\\
&= \frac{1}{\gamma}\left(1-\epsilon-\mathbb{P}\left[i(X^n;Y^nS^n)\geq\log\gamma\right]\right) && \text{(E.62)}\\
&\geq \frac{1}{\gamma}\left(1-\epsilon-(1-\epsilon_n)-\frac{B_2\log n}{\sqrt{n}}\right) && \text{(E.63)}\\
&= \frac{1}{\sqrt{n}\gamma}\,, && \text{(E.64)}
\end{aligned}$$

where (E.62) is by (E.58), (E.63) is by (E.41) and (E.64) is by (E.60).

Finally,

$$\begin{aligned}
\log M^*(n,\epsilon) &\leq \log\frac{1}{\beta_{1-\epsilon}} && \text{(E.65)}\\
&\leq \log\gamma + \frac{1}{2}\log n && \text{(E.66)}\\
&= nC - \sqrt{nV}Q^{-1}(\epsilon_n) + \frac{1}{2}\log n && \text{(E.67)}\\
&= nC - \sqrt{nV}Q^{-1}(\epsilon) + O(\log n)\,, && \text{(E.68)}
\end{aligned}$$

where (E.65) is just (E.55), (E.66) is by (E.64), (E.67) is by (E.59) and (E.68) is by Taylor's formula applied to Q^{-1} using (E.60) for ϵ_n.

■

Proof of Lemma 121: By Theorem 120 for any z we have that

$$\begin{aligned}
&\mathbb{P}\left[z \le \sum_{j=1}^{n} X_j < z + \log 2\right] \\
&\quad \le \int_{z/\sigma_n}^{(z+\log 2)/\sigma_n} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt + \frac{2B \log n}{\sqrt{n}} \,. \qquad \text{(E.69)} \\
&\quad \le \frac{\log 2}{\sigma_n \sqrt{2\pi}} + \frac{2B \log n}{\sqrt{n}} \,. \qquad \text{(E.70)}
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{-\sum_{j=1}^{n} X_j\right\} \cdot 1\left\{\sum_{j=1}^{n} X_j > A\right\}\right] \\
&\quad \le \sum_{l=0}^{\infty} \exp\{-A - l\log 2\}\, \mathbb{P}\left[A + l\log 2 \le \sum_{j=1}^{n} X_j < A + (l+1)\log 2\right] \,. \qquad \text{(E.71)}
\end{aligned}$$

Using (E.70) we get (E.42) after noting that

$$\sum_{l=0}^{\infty} 2^{-l} = 2 \,. \qquad \text{(E.72)}$$

■

We proceed to the case of no state knowledge and prove Theorems 59 and 60. For convenience, we begin by summarizing the definitions and some of the well-known properties of the processes used in the remainder of this appendix.

$$\begin{aligned}
R_j &= \mathbb{P}[S_{j+1} = 1 | Z_1^j] \,, && \text{(E.73)} \\
Q_j &= \mathbb{P}[Z_{j+1} = 1 | Z_1^j] = \delta_1 R_j + \delta_2 (1 - R_j) \,, && \text{(E.74)} \\
R_j^* &= \mathbb{P}[S_{j+1} = 1 | Z_1^j, S_0] \,, && \text{(E.75)} \\
G_j &= -\log P_{Z_j | Z_1^{j-1}}(Z_j | Z_1^{j-1}) = -\log Q_{j-1}^{\{Z_j\}} \,, && \text{(E.76)} \\
\Psi_j &= \mathbb{P}[S_{j+1} = 1 | Z_{-\infty}^j] \,, && \text{(E.77)} \\
U_j &= \mathbb{P}[Z_{j+1} = 1 | Z_{-\infty}^j] = \delta_1 \Psi_j + \delta_2 (1 - \Psi_j) \,, && \text{(E.78)} \\
F_j &= -\log P_{Z_j | Z_{-\infty}^{j-1}}(Z_j | Z_{-\infty}^{j-1}) = -\log U_{j-1}^{\{Z_j\}} \,, && \text{(E.79)} \\
\Theta_j &= \log P_{Z_j | S_j}(Z_j | S_j) = \log \delta_{S_j}^{\{Z_j\}} \,, && \text{(E.80)} \\
\Xi_j &= F_j + \Theta_j \,. && \text{(E.81)}
\end{aligned}$$

With this notation, the entropy rate of the process Z_j is given by

$$\begin{aligned}
\mathcal{H} &= \lim_{n\to\infty} \frac{1}{n} H(Z^n) && \text{(E.82)} \\
&= \mathbb{E}\,[F_0] && \text{(E.83)} \\
&= \mathbb{E}\,[h(U_0)] \,. && \text{(E.84)}
\end{aligned}$$

Define two functions $T_{0,1} : [0,1] \mapsto [\tau, 1-\tau]$:

$$T_0(x) = \frac{x(1-\tau)(1-\delta_1) + (1-x)\tau(1-\delta_2)}{x(1-\delta_1) + (1-x)(1-\delta_2)}, \tag{E.85}$$

$$T_1(x) = \frac{x(1-\tau)\delta_1 + (1-x)\tau\delta_2}{x\delta_1 + (1-x)\delta_2}. \tag{E.86}$$

Applying Bayes' formula to the conditional probabilities in (E.73), (E.75) and (E.77) yields[1]

$$R_{j+1} = T_{Z_{j+1}}(R_j), j \geq 0, \text{ a.s.} \tag{E.87}$$

$$R^*_{j+1} = T_{Z_{j+1}}(R^*_j), j \geq -1, \text{ a.s.} \tag{E.88}$$

$$\Psi_{j+1} = T_{Z_{j+1}}(\Psi_j), j \in \mathbb{Z}, \text{ a.s.} \tag{E.89}$$

where we start R_j and R^*_j as follows:

$$R_0 = 1/2, \tag{E.90}$$

$$R^*_0 = (1-\tau)1\{S_0 = 1\} + \tau 1\{S_0 = 2\}. \tag{E.91}$$

In particular, R_j, R^*_j, Q_j, Ψ_j and U_j are Markov processes.

Because of (E.89) we have

$$\min(\tau, 1-\tau) \leq \Psi_j \leq \max(\tau, 1-\tau). \tag{E.92}$$

For any pair of points $0 < x, y < 1$ denote their projective distance (as defined in [121]) by

$$d_P(x,y) = \left| \ln \frac{x}{1-x} - \ln \frac{y}{1-y} \right|. \tag{E.93}$$

As shown in [121] operators T_0 and T_1 are contracting in this distance (see also Section V.A of [122]):

$$d_P(T_a(x), T_a(y)) \leq |1-2\tau| d_P(x,y). \tag{E.94}$$

Since the derivative of $\ln \frac{x}{1-x}$ is lower-bounded by 4 we also have

$$|x-y| \leq \frac{1}{4} d_P(x,y), \tag{E.95}$$

which implies for all $a \in \{0,1\}$ that

$$|T_a(x) - T_a(y)| \leq \frac{1}{4}|1-2\tau| d_P(x,y). \tag{E.96}$$

Applying (E.96) to (E.87)-(E.89) and in the view of (E.90) and (E.92) we obtain

$$|R_j - \Psi_j| \leq \frac{1}{4}\left| \ln \frac{\tau}{1-\tau} \right| |1-2\tau|^{j-1} \qquad j \geq 1, \tag{E.97}$$

$$|Q_j - U_j| \leq \frac{|\delta_1 - \delta_2|}{4}\left| \ln \frac{\tau}{1-\tau} \right| |1-2\tau|^{j-1} \qquad j \geq 1. \tag{E.98}$$

[1]Since all conditional expectations are defined only up to almost sure equivalence, the qualifier "a.s." will be omitted below when dealing with such quantities.

Proof of Theorem 59: Achievability: In this proof we demonstrate how a central-limit theorem (CLT) result for the information density implies the $o(\sqrt{n})$ expansion. Otherwise, the proof is a repetition of the proof of Theorem 58. In particular, with equiprobable P_{X^n}, the expression for the information density $i(X^n;Y^n)$ becomes

$$\begin{aligned} i(X^n;Y^n) &= n\log 2 + \log P_{Z^n}(Z^n)\,, && \text{(E.99)}\\ &= n\log 2 + \sum_{j=1}^{n} G_j\,. && \text{(E.100)}\end{aligned}$$

One of the main differences with the proof of Theorem 58 is that the process G_j need not be α-mixing. In fact, for a range of values of δ_1, δ_2 and τ it can be shown that all (Z_j, G_j), $j = 1 \dots n$ can be reconstructed by knowing G_n. Consequently, α-mixing coefficients of G_j are all equal to $1/4$, hence G_j is not α-mixing and Theorem 120 is not applicable. At the same time G_j is mixing and ergodic (and Markov) because the underlying time-shift operator is Bernoulli.

Nevertheless, Theorem 2.6 in [68] provides a CLT extension of the classic Shannon-MacMillan-Breiman theorem. Namely it proves that the process $\frac{1}{\sqrt{n}}\log P_{Z^n}(Z^n)$ is asymptotically normal with variance V_0. Or, in other words, for any $\lambda \in \mathbb{R}$ we can write

$$\mathbb{P}\left[i(X^n;Y^n) > nC_0 + \sqrt{nV_0}\lambda\right] \to Q(\lambda)\,. \tag{E.101}$$

Conditions of Theorem 2.6 in [68] are fulfilled because of (E.14) and (E.98). Note that Appendix I.A of [122] also establishes (E.101) but with an additional assumption $\delta_1, \delta_2 > 0$.

By Theorem 18 we know that there exists a code with M codewords and average probability of error p_e bounded as

$$\begin{aligned} p_e &\le \mathbb{E}\left[\exp\left\{-\left[i(X^n;Y^n) - \log\frac{M-1}{2}\right]^+\right\}\right] && \text{(E.102)}\\ &\le \mathbb{E}\left[\exp\left\{-\left[i(X^n;Y^n) - \log M\right]^+\right\}\right] && \text{(E.103)}\end{aligned}$$

where (E.103) is by monotonicity of $\exp\{-[i(X^n;Y^n)-a]^+\}$ with respect to a. Furthermore, notice that for any random variable U and $a, b \in \mathbb{R}$ we have[2]

$$\mathbb{E}\left[\exp\left\{-\left[U-a\right]^+\right\}\right] \le \mathbb{P}[U \le b] + \exp\{a-b\}\,. \tag{E.104}$$

Fix some $\epsilon' > 0$ and set

$$\log\gamma_n = nC_0 - \sqrt{nV_0}Q^{-1}(\epsilon - \epsilon')\,. \tag{E.105}$$

Then continuing from (E.103) we obtain

$$\begin{aligned} p_e &\le \mathbb{P}[i(X^n;Y^n) \le \log\gamma_n] + \exp\{\log M - \log\gamma_n\} && \text{(E.106)}\\ &= \epsilon - \epsilon' + o(1) + \frac{M}{\gamma_n}\,, && \text{(E.107)}\end{aligned}$$

[2]This upper-bound reduces (E.102) to the usual Feinstein Lemma.

where (E.106) follows by applying (E.104) and (E.107) is by (E.101). If we set $\log M = \log \gamma_n - \log n$ then the right-hand side of (E.107) for sufficiently large n falls below ϵ. Hence we conclude that for n large enough we have

$$\begin{aligned} \log M^*(n,\epsilon) &\geq \log \gamma_n - \log n && \text{(E.108)} \\ &\geq nC_0 - \sqrt{nV_0}Q^{-1}(\epsilon - \epsilon') - \log n\,, && \text{(E.109)} \end{aligned}$$

but since ϵ' is arbitrary,

$$\log M^*(n,\epsilon) \geq nC_0 - \sqrt{nV_0}Q^{-1}(\epsilon) + o(\sqrt{n})\,. \tag{E.110}$$

Converse: To apply Theorem 28 we choose the auxiliary channel $Q_{Y^n|X^n}$ which simply outputs an equiprobable Y^n independent of the input X^n:

$$Q_{Y^n|X^n}(y^n|x^n) = 2^{-n}\,. \tag{E.111}$$

Similarly to the proof of Theorem 58 we get

$$\beta_{1-\epsilon}(P_{X^nY^n}, Q_{X^nY^n}) \leq \frac{1}{M^*}\,, \tag{E.112}$$

and also

$$\begin{aligned} \log \frac{P_{X^nY^n}(X^n,Y^n)}{Q_{X^nY^n}(X^n,Y^n)} &= n\log 2 + \log P_{Z^n}(Z^n) && \text{(E.113)} \\ &= i(X^n;Y^n)\,. && \text{(E.114)} \end{aligned}$$

We choose $\epsilon' > 0$ and set

$$\log \gamma_n = nC_0 - \sqrt{nV_0}Q^{-1}(\epsilon + \epsilon')\,. \tag{E.115}$$

By (2.67) we have, for $\alpha = 1 - \epsilon$,

$$\begin{aligned} \beta_{1-\epsilon} &\geq \frac{1}{\gamma_n}\left(1 - \epsilon - \mathbb{P}\left[i(X^n;Y^n) \geq \log \gamma_n\right]\right) && \text{(E.116)} \\ &= \frac{1}{\gamma_n}(\epsilon' + o(1))\,, && \text{(E.117)} \end{aligned}$$

where (E.117) is from (E.101). Finally, from (E.112) we obtain

$$\begin{aligned} \log M^*(n,\epsilon) &\leq \log \frac{1}{\beta_{1-\epsilon}} && \text{(E.118)} \\ &= \log \gamma_n - \log(\epsilon' + o(1)) && \text{(E.119)} \\ &= nC_0 - \sqrt{nV_0}Q^{-1}(\epsilon + \epsilon') + O(1) && \text{(E.120)} \\ &= nC_0 - \sqrt{nV_0}Q^{-1}(\epsilon) + o(\sqrt{n})\,. && \text{(E.121)} \end{aligned}$$

■

Proof of Theorem 60: Without loss of generality, we assume everywhere throughout the remainder of the appendix

$$0 < \delta_2 \leq \delta_1 \leq 1/2\,. \tag{E.122}$$

The bound (3.386) follows from Lemma 122: (3.387) follows from (E.126) after observing that when $\delta_2 > 0$ the right-hand side of (E.126) is $O(\tau)$ when $\tau \to 0$. Finally, by (E.127) we have

$$B_0 = O\left(\sqrt{-\tau \ln \tau}\right) \tag{E.123}$$

which implies that

$$\frac{B_1}{B_0} = O\left(\frac{-\ln^{3/4} \tau}{\tau^{1/4}}\right) . \tag{E.124}$$

Substituting these into the definition of Δ in Lemma 123, see (E.149), we obtain

$$\Delta = O\left(\sqrt{\frac{-\ln^3 \tau}{\tau}}\right) \tag{E.125}$$

as $\tau \to 0$. Then (3.388) follows from Lemma 123 and (3.378). ■

Lemma 122 *For any $0 < \tau < 1$ the difference $C_1 - C_0$ is lower bounded as*

$$C_1 - C_0 \geq h(\delta_1 \tau_{max} + \delta_2 \tau_{min}) - \tau_{max} h(\delta_1) - \tau_{min} h(\delta_2) , \tag{E.126}$$

where $\tau_{max} = \max(\tau, 1-\tau)$ and $\tau_{min} = \min(\tau, 1-\tau)$. Furthermore, when $\tau \to 0$ we have

$$C_1 - C_0 \leq O\left(\sqrt{-\tau \ln \tau}\right) . \tag{E.127}$$

Proof: First, notice that

$$C_1 - C_0 = \mathcal{H} - H(Z_1|S_1) = \mathbb{E}\left[\Xi_1\right] , \tag{E.128}$$

where $\mathcal{H}$ and Ξ_j were defined in (E.82) and (E.81), respectively. On the other hand we can see that

$$\mathbb{E}\left[\Xi_1 | Z_{-\infty}^0\right] \quad = \quad f(\Psi_0) , \tag{E.129}$$

where f is a non-negative, concave function on $[0,1]$, which attains 0 at the endpoints; explicitly,

$$f(x) = h(\delta_1 x + \delta_2 (1-x)) - x h(\delta_1) - (1-x) h(\delta_2) . \tag{E.130}$$

Since we know that Ψ_0 almost surely belongs to the interval between τ and $1-\tau$ we obtain after trivial algebra

$$f(x) \geq \min_{t \in [\tau_{min}, \tau_{max}]} f(t) = f(\tau_{max}) , \quad \forall x \in [\tau_{min}, \tau_{max}] . \tag{E.131}$$

Taking expectation in (E.129) and using (E.131) we prove (E.126).

On the other hand,

$$\begin{aligned} C_1 - C_0 &= \mathcal{H} - H(Z_1|S_1) && \text{(E.132)} \\ &= \mathbb{E}\left[h(\delta_1 \Psi_0 + \delta_2 (1-\Psi_0)) - h(\delta_1 1\{S_1 = 1\} + \delta_2 1\{S_1 = 2\})\right] . && \text{(E.133)} \end{aligned}$$

Because $\delta_2 > 0$ we have

$$B = \max_{x\in[0,1]} \left| \frac{d}{dx} h(\delta_1 x + \delta_2(1-x)) \right| < \infty \,. \tag{E.134}$$

So we have

$$\begin{aligned} \mathbb{E}\left[\Xi_1\right] &\leq B\mathbb{E}\left[|\Psi_0 - 1\{S_1 = 1\}|\right] && \text{(E.135)} \\ &\leq B\sqrt{\mathbb{E}\left[(\Psi_0 - 1\{S_1 = 1\})^2\right]}\,, && \text{(E.136)} \end{aligned}$$

where (E.136) follows from the Lyapunov inequality. Notice that for any estimator $\hat{A}$ of $1\{S_1 = 1\}$ based on $Z_{-\infty}^0$ we have

$$\mathbb{E}\left[(\Psi_0 - 1\{S_1 = 1\})^2\right] \leq \mathbb{E}\left[(\hat{A} - 1\{S_1 = 1\})^2\right], \tag{E.137}$$

because $\Psi_0 = \mathbb{E}\left[1\{S_1 = 1\}|Z_{-\infty}^0\right]$ is a minimal mean square error estimate.

We now take the following estimator:

$$\hat{A}_n = 1\left\{ \sum_{j=-n+1}^{0} Z_j \geq n\delta_a \right\}, \tag{E.138}$$

where n is to be specified later and $\delta_a = \frac{\delta_1+\delta_2}{2}$. We then have the following upper bound on its mean square error:

$$\begin{aligned} \mathbb{E}\left[(\hat{A}_n - 1\{S_1 = 1\})^2\right] &= \mathbb{P}[1\{S_1 = 1\} \neq \hat{A}_n] && \text{(E.139)} \\ &\leq \mathbb{P}[\hat{A}_n \neq 1\{S_1 = 1\}, S_1 = \cdots = S_{-n+1}] \\ &\quad + 1 - \mathbb{P}[S_1 = \cdots = S_{-n+1}] && \text{(E.140)} \\ &= \frac{1}{2}(1-\tau)^n \left(\mathbb{P}[B(n,\delta_1) < n\delta_a] + \mathbb{P}[B(n,\delta_2) \geq n\delta_a]\right) \\ &\quad + 1 - (1-\tau)^n\,, && \text{(E.141)} \end{aligned}$$

where $B(n,\delta)$ denotes the binomially distributed random variable. Using Chernoff bounds we can find that for some E_1 we have

$$\mathbb{P}[B(n,\delta_1) < n\delta_a] + \mathbb{P}[B(n,\delta_2) \geq n\delta_a] \leq 2e^{-nE_1}\,. \tag{E.142}$$

Then we have

$$\mathbb{E}\left[(\hat{A}_n - 1\{S_1 = 1\})^2\right] \leq 1 - (1-\tau)^n(1 - e^{-nE_1})\,. \tag{E.143}$$

If we denote

$$\beta = -\ln(1-\tau)\,. \tag{E.144}$$

and choose

$$n = \left\lceil -\frac{1}{E_1} \ln \frac{\beta}{E_1} \right\rceil, \tag{E.145}$$

we obtain that

$$\mathbb{E}\left[(\hat{A}_n - 1\{S_1 = 1\})^2\right] \leq 1 - (1-\tau)\cdot e^{-\frac{\beta}{E_1}\ln\frac{\beta}{E_1}} \left(1 - \frac{\beta}{E_1}\right). \tag{E.146}$$

When $\tau \to 0$ we have $\beta = \tau + o(\tau)$ and then it is not hard to show that

$$\mathbb{E}\left[(\hat{A}_n - 1\{S_1 = 1\})^2\right] \leq \frac{\tau}{E_1} \ln \frac{\tau}{E_1} + o(\tau \ln \tau)\,. \tag{E.147}$$

From (E.136), (E.137), and (E.147) we obtain (E.127). ■

Lemma 123 *For any $0 < \tau < 1$ we have*

$$|V_0 - V_1| \leq 2\sqrt{V_1 \Delta} + \Delta\,, \tag{E.148}$$

where Δ satisfies

$$\Delta \leq B_0 + \frac{B_0}{2(1 - \sqrt{|1-2\tau|})} \ln \frac{eB_1}{B_0}\,, \tag{E.149}$$

$$B_0 = \frac{d_2(\delta_1||\delta_2)}{d(\delta_1||\delta_2)} |C_0 - C_1|\,, \tag{E.150}$$

$$B_1 = \sqrt{\frac{B_0}{|1-2\tau|}} \left(d(\delta_1||\delta_2) \left| \ln \frac{\tau}{1-\tau} \right| + \frac{h(\delta_1) - h(\delta_2)}{2|1-2\tau|} \right)\,, \tag{E.151}$$

$$d_2(a||b) = a \log^2 \frac{a}{b} + (1-a) \log^2 \frac{1-a}{1-b} \tag{E.152}$$

and $d(a||b) = a \log \frac{a}{b} + (1-a) \log \frac{1-a}{1-b}$ is the binary divergence.

Proof: First denote

$$\Delta = \lim_{n \to \infty} \frac{1}{n} \operatorname{Var} \left[\sum_{j=1}^{n} \Xi_j \right]\,, \tag{E.153}$$

where Ξ_j was defined in (E.81); the finiteness of Δ is to be proved below.

By (E.81) we have

$$F_j = -\Theta_j + \Xi_j\,. \tag{E.154}$$

In the course of the proof of Theorem 58 we have shown that

$$\mathbb{E}\left[\Theta_j\right] = C_1 - \log 2\,, \tag{E.155}$$

$$\operatorname{Var} \left[\sum_{j=1}^{n} \Theta_j \right] = nV_1 + O(1)\,. \tag{E.156}$$

Essentially, Ξ_j is a correction term, compared to the case of state known at the receiver, which we expect to vanish as $\tau \to 0$. By definition of V_0 we have

$$V_0 = \lim_{n \to \infty} \frac{1}{n} \operatorname{Var} \left[\sum_{j=1}^{n} F_j \right] \tag{E.157}$$

$$= \lim_{n \to \infty} \operatorname{Var} \left[-\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Theta_j + \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Xi_j \right]\,. \tag{E.158}$$

Now (E.148) follows from (E.153), (E.156) and by an application of the Cauchy-Schwarz inequality to (E.158).

We are left to prove (E.149). First, notice that

$$\Delta = \mathrm{Var}[\Xi_0] + 2\sum_{j=1}^{\infty} \mathrm{cov}(\Xi_0, \Xi_j)\,. \tag{E.159}$$

The first term is bounded by Lemma 124

$$\mathrm{Var}[\Xi_j] \le \mathbb{E}\left[\Xi_j^2\right] \le B_0\,. \tag{E.160}$$

Next, set

$$N = \left\lceil \frac{2\ln\frac{B_0}{B_1}}{\ln|1-2\tau|} \right\rceil\,. \tag{E.161}$$

We have then

$$\begin{aligned}
\sum_{j=1}^{\infty} \mathrm{cov}[\Xi_0, \Xi_j] &\le (N-1)B_0 + B_1 \sum_{j\ge N} |1-2\tau|^{j/2} && \text{(E.162)}\\
&\le \frac{\ln\frac{B_0}{B_1}}{\ln\sqrt{|1-2\tau|}} B_0 + \frac{B_0}{1-\sqrt{|1-2\tau|}} && \text{(E.163)}\\
&\le \frac{B_0}{1-\sqrt{|1-2\tau|}} \ln\frac{eB_1}{B_0}\,, && \text{(E.164)}
\end{aligned}$$

where in (E.162) for $j < N$ we used Cauchy-Schwarz inequality and (E.160), for $j \ge N$ we used Lemma 125; (E.163) follows by definition of N and (E.164) follows by $\ln x \le x - 1$. Finally, (E.149) follows now by applying (E.160) and (E.164) to (E.159). ■

Lemma 124 *Under the conditions of Lemma 123, we have*

$$\mathrm{Var}[\Xi_j] \le \mathbb{E}\left[\Xi_j^2\right] \le B_0\,. \tag{E.165}$$

Proof: First notice that

$$\begin{aligned}
\mathbb{E}\left[\Xi_1 | Z_{-\infty}^0\right] &= \Psi_0 d(\delta_1||\delta_1\Psi_0 + \delta_2(1-\Psi_0))\\
&\quad + (1-\Psi_0) d(\delta_2||\delta_1\Psi_0 + \delta_2(1-\Psi_0))\,, && \text{(E.166)}\\
\mathbb{E}\left[\Xi_1^2 | Z_{-\infty}^0\right] &= \Psi_0 d_2(\delta_1||\delta_1\Psi_0 + \delta_2(1-\Psi_0))\\
&\quad + (1-\Psi_0) d_2(\delta_2||\delta_1\Psi_0 + \delta_2(1-\Psi_0))\,. && \text{(E.167)}
\end{aligned}$$

Below we adopt the following notation

$$\bar{x} = 1 - x\,. \tag{E.168}$$

Applying Lemma 126 twice (with $a = \delta_1\,, b = \delta_1 x + \delta_2\bar{x}$ and with $a = \delta_2\,, b = \delta_1 x + \delta_2\bar{x}$) we obtain

$$\begin{aligned}
&x d_2(\delta_1||\delta_1 x + \delta_2\bar{x}) + \bar{x} d_2(\delta_2||\delta_1 x + \delta_2\bar{x})\\
&\quad \le \frac{d_2(\delta_1||\delta_2)}{d(\delta_1||\delta_2)} \left(x d(\delta_1||\delta_1 x + \delta_2\bar{x}) + \bar{x} d(\delta_2||\delta_1 x + \delta_2\bar{x})\right)\,. && \text{(E.169)}
\end{aligned}$$

If we substitute $x = \Psi_0$ here, then by comparing (E.166) and (E.167) we obtain that

$$\mathbb{E}\left[\Xi_1^2 | Z_{-\infty}^0\right] \leq \frac{d_2(\delta_1||\delta_2)}{d(\delta_1||\delta_2)} \mathbb{E}\left[\Xi_1 | Z_{-\infty}^0\right]. \tag{E.170}$$

Averaging this we obtain[3]

$$\mathbb{E}\left[\Xi_1^2\right] \leq \frac{d_2(\delta_1||\delta_2)}{d(\delta_1||\delta_2)} (C_1 - C_0). \tag{E.172}$$

■

Lemma 125 *Under the conditions of Lemma 123, we have*

$$\mathrm{cov}[\Xi_0, \Xi_j] \leq B_1 |1 - 2\tau|^{j/2}. \tag{E.173}$$

Proof: From the definition of Ξ_j we have that

$$\mathbb{E}\left[\Xi_j | S_{-\infty}^0, Z_{-\infty}^{j-1}\right] = f(\Psi_{j-1}, R_{j-1}^*), \tag{E.174}$$

where

$$f(x,y) = y d(\delta_1||\delta_1 x + \delta_2(1-x)) + (1-y) d(\delta_2||\delta_1 x + \delta_2(1-x)). \tag{E.175}$$

Notice the following relationship:

$$\frac{d}{d\lambda} H(\bar{\lambda}Q + \lambda P) = D(P||\bar{\lambda}Q + \lambda P) - D(Q||\bar{\lambda}Q + \lambda P) + H(P) - H(Q). \tag{E.176}$$

This has two consequences. First it shows that the function

$$D(P||\bar{\lambda}Q + \lambda P) - D(Q||\bar{\lambda}Q + \lambda P) \tag{E.177}$$

is monotonically decreasing with λ (since it is a derivative of a concave function). Second, we have the following general relation for the excess of the entropy above its affine approximation:

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} [H((1-\lambda)Q + \lambda P) - (1-\lambda)H(Q) - \lambda H(P)] = D(P||Q), \tag{E.178}$$

$$\left.\frac{d}{d\lambda}\right|_{\lambda=1} [H((1-\lambda)Q + \lambda P) - (1-\lambda)H(Q) - \lambda H(P)] = -D(Q||P). \tag{E.179}$$

Also it is clear that for all other λ's the derivative is in between these two extreme values.

[3]Note that it can also be shown that

$$\mathbb{E}\left[\Xi_1^2\right] \geq \frac{d_2(\delta_2||\delta_1)}{d(\delta_2||\delta_1)} (C_1 - C_0), \tag{E.171}$$

and therefore (E.172) cannot be improved significantly.

Applying this to the binary case we have

$$\max_{x,y\in[0,1]}\left|\frac{df(x,y)}{dy}\right| = \max_{x\in[0,1]}|d(\delta_1||\delta_1 x+\delta_2(1-x))-d(\delta_2||\delta_1 x+\delta_2(1-x))| \quad \text{(E.180)}$$

$$= \max(d(\delta_1||\delta_2), d(\delta_2||\delta_1)) \quad \text{(E.181)}$$

$$= d(\delta_1||\delta_2)\,, \quad \text{(E.182)}$$

where (E.181) follows because the function in the right side of (E.180) is decreasing and (E.182) is because we are restricted to $\delta_2 \le \delta_1 \le \frac{1}{2}$. On the other hand, we see that

$$f(x,x) = h(\delta_1 x+\delta_2(1-x)) - xh(\delta_1) - (1-x)h(\delta_2) \ge 0\,. \quad \text{(E.183)}$$

Comparing with (E.178) and (E.179), we have

$$\max_{x\in[0,1]}\left|\frac{df(x,x)}{dx}\right| = \max(d(\delta_1||\delta_2), d(\delta_2||\delta_1)) \quad \text{(E.184)}$$

$$= d(\delta_1||\delta_2)\,. \quad \text{(E.185)}$$

By the properties of f we have

$$\left|f(\Psi_{j-1}, R^*_{j-1}) - f(\Psi_{j-1},\Psi_{j-1})\right| \le d(\delta_1||\delta_2)|R^*_{j-1}-\Psi_{j-1}| \quad \text{(E.186)}$$

$$\le B_2|1-2\tau|^{j-1}\,, \quad \text{(E.187)}$$

where for convenience we denote

$$B_2 = \frac{1}{2}d(\delta_1||\delta_2)\left|\ln\frac{\tau}{1-\tau}\right|\,. \quad \text{(E.188)}$$

Indeed, (E.186) is by (E.182) and (E.187) follows by observing that

$$\Psi_{j-1} = T_{Z_{j-1}}\circ\cdots\circ T_{Z_1}(\Psi_0)\,, \quad \text{(E.189)}$$

$$R^*_{j-1} = T_{Z_{j-1}}\circ\cdots\circ T_{Z_1}(R^*_0) \quad \text{(E.190)}$$

and applying (E.96). Consequently, we have shown

$$\left|\mathbb{E}\left[\Xi_j|S^0_{-\infty}, Z^{j-1}_{-\infty}\right] - f(\Psi_{j-1},\Psi_{j-1})\right| \le B_2|1-2\tau|^{j-1}\,, \quad \text{(E.191)}$$

or, after a trivial generalization,

$$\left|\mathbb{E}\left[\Xi_j|S^k_{-\infty}, Z^{j-1}_{-\infty}\right] - f(\Psi_{j-1},\Psi_{j-1})\right| \le B_2|1-2\tau|^{j-1-k}\,. \quad \text{(E.192)}$$

Notice that by comparing (E.183) with (E.166) we have

$$\mathbb{E}\left[f(\Psi_{j-1},\Psi_{j-1})\right] = \mathbb{E}\left[\Xi_j\right]. \quad \text{(E.193)}$$

Next we show that

$$\left|\mathbb{E}\left[\Xi_j|S^0_{-\infty}, Z^0_{-\infty}\right] - \mathbb{E}\left[\Xi_j\right]\right| \le |1-2\tau|^{\frac{j-1}{2}}\left[2B_2+B_3\right]\,, \quad \text{(E.194)}$$

where

$$B_3 = \frac{h(\delta_1) - h(\delta_2)}{2|1-2\tau|}\,. \tag{E.195}$$

Denote

$$t(\Psi_k, S_k) \stackrel{\triangle}{=} \mathbb{E}\left[f(\Psi_{j-1}, \Psi_{j-1})|S_{-\infty}^k Z_{-\infty}^k\right]. \tag{E.196}$$

Then because of (E.185) and since Ψ_k affects only the initial condition for Ψ_{j-1} when written as (E.189), we have for arbitrary $x_0 \in [\tau, 1-\tau]$,

$$|t(\Psi_k, S_k) - t(x_0, S_k)| \le B_2|1-2\tau|^{j-k-1}\,. \tag{E.197}$$

On the other hand, as an average of $f(x,x)$ the function $t(x_0, s)$ satisfies

$$0 \le t(x_0, S_k) \le \max_{x\in[0,1]} f(x,x) \le h(\delta_1) - h(\delta_2)\,. \tag{E.198}$$

From here and (E.13) we have

$$\left|\mathbb{E}\left[t(x_0, S_k)|S_{-\infty}^0 Z_{-\infty}^0\right] - \mathbb{E}\left[t(x_0, S_k)\right]\right| \le \frac{h(\delta_1) - h(\delta_2)}{2}|1-2\tau|^k\,, \tag{E.199}$$

or, together with (E.197),

$$\left|\mathbb{E}\left[t(\Psi_k, S_k)|S_{-\infty}^0 Z_{-\infty}^0\right] - \mathbb{E}\left[t(x_0, S_k)\right]\right| \le \frac{h(\delta_1) - h(\delta_2)}{2}|1-2\tau|^k + B_2|1-2\tau|^{j-k-1}\,. \tag{E.200}$$

This argument remains valid if we replace x_0 with a random variable $\tilde{\Psi}_k$, which depends on S_k but conditioned on S_k is independent of $(S_{-\infty}^0, Z_{-\infty}^0)$. Having made this replacement and assuming $P_{\tilde{\Psi}_k|S_k} = P_{\Psi_k|S_k}$ we obtain

$$\left|\mathbb{E}\left[t(\Psi_k, S_k)|S_{-\infty}^0 Z_{-\infty}^0\right] - \mathbb{E}\left[t(\Psi_k, S_k)\right]\right| \le \frac{h(\delta_1) - h(\delta_2)}{2}|1-2\tau|^k + B_2|1-2\tau|^{j-k-1}\,. \tag{E.201}$$

Summing together (E.192), (E.193), (E.196), (E.197) and (E.201) we obtain that for arbitrary $0 \le k \le j-1$ we have

$$\left|\mathbb{E}\left[\Xi_j|S_{-\infty}^0 Z_{-\infty}^0\right] - \mathbb{E}\left[\Xi_j\right]\right| \le \frac{h(\delta_1) - h(\delta_2)}{2}|1-2\tau|^k + 2B_2|1-2\tau|^{j-k-1}\,. \tag{E.202}$$

Setting here $k = \lfloor j - 1/2 \rfloor$ we obtain (E.194).

Finally, we have

$$\begin{aligned}
\mathrm{cov}[\Xi_0, \Xi_j] &= \mathbb{E}\left[\Xi_0\Xi_j\right] - \mathbb{E}^2[\Xi_0] && \text{(E.203)}\\
&= \mathbb{E}\left[\Xi_0 \mathbb{E}\left[\Xi_j|S_{-\infty}^0, Z_{-\infty}^0\right]\right] - \mathbb{E}^2[\Xi_0] && \text{(E.204)}\\
&\le \mathbb{E}\left[\Xi_0 \mathbb{E}\left[\Xi_j\right]\right] + \mathbb{E}\left[|\Xi_0|(2B_2+B_3)|1-2\tau|^{\frac{j-1}{2}}\right] - \mathbb{E}^2[\Xi_0] && \text{(E.205)}\\
&= \mathbb{E}\left[|\Xi_0|\right](2B_2+B_3)|1-2\tau|^{\frac{j-1}{2}} && \text{(E.206)}\\
&\le \sqrt{\mathbb{E}\left[\Xi_0^2\right]}(2B_2+B_3)|1-2\tau|^{\frac{j-1}{2}} && \text{(E.207)}\\
&= \sqrt{B_0}(2B_2+B_3)|1-2\tau|^{\frac{j-1}{2}}\,, && \text{(E.208)}
\end{aligned}$$

where (E.205) is by (E.194), (E.207) is a Lyapunov's inequality and (E.208) is Lemma 124. ■

Lemma 126 *Assume that $\delta_1 \geq \delta_2 > 0$ and $\delta_2 \leq a, b \leq \delta_1$; then*

$$\frac{d(a||b)}{d_2(a||b)} \geq \frac{d(\delta_1||\delta_2)}{d_2(\delta_1||\delta_2)}\,. \tag{E.209}$$

Proof: While inequality (E.209) can be easily checked numerically, its rigorous proof is somewhat lengthy. Since the base of the logarithm cancels in (E.209), we replace log by ln below. Observe that the lemma is trivially implied by the following two statements:

$$\forall \delta \in [0, 1/2]: \quad \frac{d(a||\delta)}{d_2(a||\delta)} \quad \text{is a non-increasing function of } a \in [0, 1/2]\,; \tag{E.210}$$

and

$$\frac{d(\delta_1||b)}{d_2(\delta_1||b)} \quad \text{is a non-decreasing function of } b \in [0, \delta_1]\,. \tag{E.211}$$

To prove (E.210) we show that the derivative of $\frac{d_2(a||\delta)}{d(a||\delta)}$ is non-negative. This is equivalent to showing that

$$\begin{cases} f_a(\delta) \leq 0\,, & \text{if } a \leq \delta\,, \\ f_a(\delta) \geq 0\,, & \text{if } a \geq \delta\,, \end{cases} \tag{E.212}$$

where

$$f_a(\delta) = 2d(a||\delta) + \ln\frac{a}{\delta} \cdot \ln\frac{1-a}{1-\delta}\,. \tag{E.213}$$

It is easy to check that

$$f_a(a) = 0\,, f_a'(a) = 0\,. \tag{E.214}$$

So it is sufficient to prove that

$$f_a(\delta) = \begin{cases} \text{convex}\,, & 0 \leq \delta \leq a\,, \\ \text{concave}\,, & a \leq \delta \leq 1/2\,. \end{cases} \tag{E.215}$$

Indeed, if (E.215) holds then an affine function $g(\delta) = 0\delta + 0$ will be a lower bound for $f_a(\delta)$ on $[0, a]$ and an upper bound on $[a, 1/2]$, which is exactly (E.212). To prove (E.215) we analyze the second derivative of f_a:

$$f_a''(\delta) = \frac{2a}{\delta^2} + \frac{2\bar{a}}{\bar{\delta}^2} - \frac{1}{\delta^2}\ln\frac{\bar{\delta}}{\bar{a}} - \frac{2}{\delta\bar{\delta}} - \frac{1}{\bar{\delta}^2}\ln\frac{\delta}{a}\,. \tag{E.216}$$

In the case $\delta \geq a$ an application of the bound $\ln x \leq x - 1$ yields

$$\begin{aligned} f_a''(\delta) &\leq \frac{2a}{\delta^2} + \frac{2\bar{a}}{\bar{\delta}^2} - \frac{1}{\delta^2}\left(\frac{\bar{\delta}}{\bar{a}} - 1\right) - \frac{2}{\delta\bar{\delta}} - \frac{1}{\bar{\delta}^2}\left(\frac{\delta}{a} - 1\right) && \text{(E.217)} \\ &\leq 0\,. && \text{(E.218)} \end{aligned}$$

Similarly, in the case $\delta \leq a$ an application of the bound $\ln x \geq 1 - \frac{1}{x}$ yields

$$\begin{aligned} f_a''(\delta) &\geq \frac{2a}{\delta^2} + \frac{2\bar{a}}{\bar{\delta}^2} - \frac{1}{\delta^2}\left(1 - \frac{\bar{a}}{\delta}\right) - \frac{2}{\delta\bar{\delta}} - \frac{1}{\bar{\delta}^2}\left(1 - \frac{a}{\delta}\right) && \text{(E.219)} \\ &\geq 0\,. && \text{(E.220)} \end{aligned}$$

This proves (E.215) and, therefore, (E.210).

To prove (E.211) we take the derivative of $\frac{d(\delta_1||b)}{d_2(\delta_1||b)}$ with respect to b; requiring it to be non-negative is equivalent to

$$2(1-2b)\left(\delta\ln\frac{\delta}{b}\right)\left(\bar{\delta}\ln\frac{\bar{\delta}}{\bar{b}}\right)+(\delta\bar{b}+\bar{\delta}b)\left(\delta\ln^2\frac{\delta}{b}-\bar{\delta}\ln^2\frac{\bar{\delta}}{\bar{b}}\right)\geq 0\,. \tag{E.221}$$

It is convenient to introduce $x=b/\delta\in[0,1]$ and then we define

$$f_\delta(x)=2(1-2\delta x)\delta\bar{\delta}\ln x\cdot\ln\frac{1-\delta x}{\bar{\delta}}+\delta(1+x(1-2\delta))\left(\delta\ln^2 x-\bar{\delta}\ln^2\frac{1-\delta x}{\bar{\delta}}\right)\,, \tag{E.222}$$

for which we must show

$$f_\delta(x)\geq 0\,. \tag{E.223}$$

If we think of $A=\ln x$ and $B=\ln\frac{1-\delta x}{\delta}$ as independent variables, then (E.221) is equivalent to solving

$$2\gamma AB+\alpha A^2-\beta B^2\geq 0\,, \tag{E.224}$$

which after some manipulation (and observation that we naturally have a requirement $A<0<B$) reduces to

$$\frac{A}{B}\leq-\frac{\gamma}{\alpha}-\frac{1}{\alpha}\sqrt{\gamma^2+\alpha\beta}\,. \tag{E.225}$$

After substituting the values for A,B,α,β and γ we get that (E.221) will be shown if we can show for all $0<x<1$ that

$$\frac{\ln\frac{1}{x}}{\ln\frac{1-\delta x}{\delta}}\geq\frac{1-2\delta x}{1+x(1-2\delta)}\frac{\bar{\delta}}{\delta}+\left(\left(\frac{1-2\delta x}{1-2\delta x+x}\right)^2\left(\frac{\bar{\delta}}{\delta}\right)^2+\frac{\bar{\delta}}{\delta}\right)^{1/2}\,. \tag{E.226}$$

To show (E.226) we are allowed to upper-bound $\ln x$ and $\ln\frac{1-\delta x}{\delta}$. We use the following upper bounds for $\ln x$ and $\ln\frac{1-\delta x}{\delta}$, correspondingly:

$$\ln x \leq (x-1)-(x-1)^2/2+(x-1)^3/3-(x-1)^4/4+(x-1)^5/5\,, \tag{E.227}$$

$$\ln y \leq (y-1)-(y-1)^2/2+(y-1)^3/3\,, \tag{E.228}$$

particularized to $y=1-\frac{\delta x}{\delta}$; both bounds follow from the fact that the derivative of $\ln x$ of the corresponding order is always negative. Applying (E.227) and (E.228) to the left side of (E.226) and after some tedious algebra, we find that (E.226) is implied by the

$$\frac{\delta^2(1-x)^3}{(1-\delta)^5}P_\delta(1-x)\geq 0\,, \tag{E.229}$$

where

$$\begin{aligned}P_\delta(x) = &-(4\delta^2-1)(1-\delta)^2/12\\ &+(1-\delta)(4-5\delta+4\delta^2-24\delta^3+24\delta^4)x/24\\ &+(8-20\delta+15\delta^2+20\delta^3-100\delta^4+72\delta^5)x^2/60\\ &-(1-\delta)^3(11-28\delta+12\delta^2)x^3/20\\ &+(1-\delta)^3(1-2\delta)^2x^4/5\,.\end{aligned} \tag{E.230}$$

Assume that $P_\delta(x_0) < 0$ for some x_0. For all $0 < \delta \leq 1/2$ we can easily check that $P_\delta(0) > 0$ and $P_\delta(1) > 0$. Therefore, there must be a root x_1 of P_δ in $(0, x_0)$ and a root x_2 in $(x_0, 1)$ by continuity. It is also easily checked that $P'_\delta(0) > 0$ for all δ. But then we must have at least one root of P'_δ in $[0, x_1)$ and at least one root of P'_δ in $(x_2, 1]$.

Now, $P'_\delta(x)$ is a cubic polynomial such that $P'_\delta(0) > 0$. So it must have at least one root on the negative real axis and two roots on $[0, 1]$. But since $P''_\delta(0) > 0$, it must be that $P''_\delta(x)$ also has two roots on $[0, 1]$. But $P''_\delta(x)$ is a quadratic polynomial, so its roots are algebraic functions of δ, for which we can easily check that one of them is always larger than 1. So, $P'_\delta(x)$ has at most one root on $[0, 1]$. And therefore we arrive at a contradiction and $P_\delta \geq 0$ on $[0, 1]$, which proves (E.229). ■

References

[1] V. Strassen, "Asymptotische Abschätzungen in Shannon's Informationstheorie," in *Trans. 3d Prague Conf. Inf. Theory*, Prague, 1962, pp. 689–723.

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

[3] J. Wolfowitz, "The coding of messages subject to chance errors," *Illinois J. Math.*, vol. 1, pp. 591–606, 1957.

[4] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.

[5] P. Elias, "Coding for two noisy channels," in *Proc. 3d London Symp. Inf. Theory*, Washington, DC, Sep. 1955, pp. 61–76.

[6] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inf. Theory*, vol. 11, no. 1, pp. 3–18, 1965.

[7] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels i," *Inf. Contr.*, vol. 10, pp. 65–103, 1967.

[8] A. Barg and A. McGregor, "Distance distribution of binary codes and the error probability of decoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4237–4226, Dec. 2005.

[9] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[10] R. A. Costa, M. Langberg, and J. Barros, "One-shot capacity of discrete channels," in *Proc. 2010 IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, TX, USA, Jun. 2010.

[11] J. Shi and R. D. Wesel, "A study on universal codes with finite block lengths," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3066–3074, 2007.

[12] S. J. MacMullan and O. M. Collins, "A comparison of known codes, random codes and the best codes," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 3009–3022, 1998.

[13] J. N. Laneman, "On the distribution of mutual information," in *Proc. 2006 Workshop Inf. Theory and Appl.*, San Diego, CA, Feb. 2006.

[14] D. Buckingham and M. Valenti, "The information-outage probability of finite-length codes over AWGN channels," in *Proc. 2008 Conf. Inf. Sci. and Syst. (CISS)*, Princeton, NJ, Mar. 2008.

[15] A. Feinstein, "A new basic theorem of information theory," *IRE Trans. Inform. Theory*, vol. 4, no. 4, pp. 2–22, 1954.

[16] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[17] ——, "Graph decomposition: a new key to coding theorems," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 5–12, 1981.

[18] A. Valembois and M. P. C. Fossorier, "Sphere-packing bounds revisited for moderate block lengths," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 2998–3014, 2004.

[19] G. Wiechman and I. Sason, "An improved sphere-packing bound for finite-length codes over symmetric memoryless channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1962–1990, 2008.

[20] D. Slepian, "Bounds on communication," *Bell Syst. Tech. J.*, vol. 42, pp. 681–707, 1963.

[21] E. N. Gilbert, "Information theory after 18 years," *Science*, vol. 152, no. 3720, pp. 320–326, Apr. 1966.

[22] D. Lazic, T. Beth, and S. Egner, "Constrained capacity of the AWGN channel," in *Proc. 1998 IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, 1998.

[23] C. Salema, *Microwave Radio Links: From Theory to Design*. New York: Wiley, 2002.

[24] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1284–1292, 1994.

[25] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1570–1579, 2002.

[26] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, 1994.

[27] L. Weiss, "On the strong converse of the coding theorem for symmetric channels without memory," *Quart. Appl. Math.*, vol. 18, no. 3, 1960.

[28] D. Baron, M. A. Khojastepour, and R. G. Baraniuk, "How quickly can we approach channel capacity?" in *Proc. 38th Asilomar Conf. Signals, Syst., and Comput.*, Pacific Grove, CA, Nov. 2004.

[29] R. L. Dobrushin, "Mathematical problems in the Shannon theory of optimal coding of information," in *Proc. 4th Berkeley Symp. Mathematics, Statistics, and Probability*, vol. 1, 1961, pp. 211–252.

[30] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.

[31] ——, "The zero error capacity of a noisy channel," *IRE Trans. Inform. Theory*, vol. 2, no. 3, pp. 8–19, Sep. 1956.

[32] 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.

[33] ——, "New channel coding achievability bounds," in *Proc. 2008 IEEE Int. Symp. Inf. Theory (ISIT)*, Toronto, Canada, Jul. 2008.

[34] C. E. Shannon, "Certain results in coding theory for noisy channels," *Inf. Contr.*, vol. 1, pp. 6–25, 1957.

[35] A. J. Thomasian, "Error bounds for continuous channels," in *Proc. 4th London Symp. Inf. Theory*, C. Cherry, Ed., Washington, DC, 1961.

[36] R. Ash, *Information Theory.* New York: Interscience Publishers, 1965.

[37] R. J. McEliece, *The Theory of Information and Coding: A Framework for Communication.* Reading, MA: Addison-Wesley, 1977.

[38] E. A. Haroutunian, "Lower bounds for error probability in channels with feedback," *Prob. Peredachi Inf.*, vol. 13, no. 2, pp. 36–44, 1977.

[39] J. Wolfowitz, *Coding Theorems of Information Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1962.

[40] C. E. Shannon, "Certain results in coding theory for noisy channels," *Inf. Contr.*, vol. 1, no. 1, pp. 6–25, 1957.

[41] H. V. Poor and S. Verdú, "A lower bound on the error probability in multihypothesis testing," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1992–1993, 1995.

[42] J. Wolfowitz, "Notes on a general strong converse," *Inf. Contr.*, vol. 12, pp. 1–4, 1968.

[43] M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1753–1768, 2003.

[44] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4947–4966, Nov. 2009.

[45] R. E. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inf. Theory*, vol. 20, no. 4, pp. 405–417, 1974.

[46] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.

[47] S. Verdú, *EE528–Information Theory, Lecture Notes.* Princeton, NJ: Princeton Univ., 2007.

[48] T. Kadota, "Generalization of Feinstein's fundamental lemma," *IEEE Trans. Inf. Theory*, vol. 16, no. 6, pp. 791–792, 1970.

[49] F. Liese and I. Vajda, "On divergences and informations in statistics and information theory," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4394–4412, 2006.

[50] I. Csiszár, "Information-type measures of difference of probability distributions and indirect observation," *Studia Sci. Math. Hungar.*, vol. 2, pp. 229–318, 1967.

[51] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. New York: Wiley, 1971, vol. II.

[52] P. V. Beeck, "An application of Fourier methods to the problem of sharpening the Berry-Esseen inequality," *Z. Wahrscheinlichkeitstheorie und Verw. Geb.*, vol. 23, pp. 187–196, 1972.

[53] L. Wang, R. Colbeck, and R. Renner, "Simple channel coding bounds," in *Proc. 2009 IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jul. 2009.

[54] M. Donsker and S. Varadhan, "Asymptotic evaluation of certain markov process expectations for large time. i. ii." *Comm. Pure Appl. Math.*, vol. 28, no. 1, pp. 1–47, 1975.

[55] O. Keren and S. Litsyn, "A lower bound on the probability of decoding error over a BSC channel," in *Electrical and ELectronic Engineers in Israel. The 21st IEEE Convention of the*, 2000, pp. 271–273.

[56] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Dispersion of the Gilbert-Elliott channel," in *Proc. 2009 IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jul. 2009.

[57] ——, "Dispersion of the Gilbert-Elliott channel," *IEEE Trans. Inf. Theory*, to appear.

[58] A. Barg and G. D. Forney, "Random codes: minimum distances and error exponents," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2568–2573, 2002.

[59] T. Helleseth, T. Klove, and V. I. Levenshtein, "On the information function of an error-correcting code," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 549–557, 1997.

[60] A. E. Ashikhmin, personal communication, 2009.

[61] S. Dolinar, D. Divsalar, and F. Pollara, "Code performance as a function of block size," *JPL TDA Progress Report*, vol. 42, no. 133, 1998, Jet Propulsion Laboratory, Pasadena, CA.

[62] E. N. Gilbert, "Capacity of burst-noise channels," *Bell Syst. Tech. J.*, vol. 39, pp. 1253–1265, Sep. 1960.

[63] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, Sep. 1963.

[64] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert-Elliott channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1277–1290, 1989.

[65] J. C. Kieffer, "Epsilon-capacity of binary symmetric averaged channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 288–303, 2007.

[66] C.-G. Esseen, "On the concentration function of a sum of independent random variables," *Z. Wahrscheinlichkeitstheorie und Verw. Geb.*, vol. 9, no. 4, pp. 290–308, 1968.

[67] E. Biglieri, J. Proakis, , and S. S. (Shitz), "Fading channels: Information-theoretic and communication aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[68] I. A. Ibragimov, "Some limit theorems for stationary processes," *Theor. Probability Appl.*, vol. 7, no. 4, 1962.

[69] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Dispersion of Gaussian channels," in *Proc. 2009 IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jul. 2009.

[70] ——, "Minimum energy to send k bits with and without feedback," in *Proc. 2010 IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, TX, Jul. 2010.

[71] ——, "Minimum energy to send k bits with and without feedback," *IEEE Trans. Inf. Theory*, Mar. 2010, submitted for publication.

[72] A. E. Ashikhmin, A. Barg, and S. N. Litsyn, "A new upper bound on the reliability function of the Gaussian channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 1945–1961, 2000.

[73] S. O. Rice, "Communication in the presence of noise – probability of error for two encoding schemes," *Bell Syst. Tech. J.*, vol. 29, pp. 60–93, 1950.

[74] P. Ebert, "Error bounds for parallel communication channels," MIT, Tech. Rep. RLE-TR-448, Aug. 1966.

[75] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[76] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Trans. Networking*, vol. 10, no. 4, pp. 487–499, Aug. 2002.

[77] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th ed. New York: Dover, 1972.

[78] A. L. Jones, "An extension of inequality involving modified Bessel functions," *J. Math. and Phys.*, vol. 47, pp. 220–221, 1968.

[79] A. V. Prokhorov, "Inequalities for Bessel functions of a purely imaginary argument," *Theor. Probability Appl.*, vol. 13, pp. 496–501, 1968.

[80] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4947–4966, Nov. 2009.

[81] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*. New York: Springer Verlag, 2009.

[82] Y. Altug and A. B. Wagner, "Moderate deviation analysis of channel coding: Discrete memoryless case," in *Proc. 2010 IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, TX, USA, Jun. 2010.

[83] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, Sep. 1990.

[84] ——, *Multiuser Detection*. Cambridge, UK: Cambridge Univ. Press, 1998.

[85] L. A. Shepp, "Distinguishing a sequence of random variables from a translate of itself," *Ann. Math. Stat.*, vol. 36, no. 4, pp. 1107–1112, 1965.

[86] R. G. Gallager and B. Nakiboglŭ, "Variations on a theme by Schalkwijk and Kailath," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 6–17, 2010.

[87] K. S. Zigangirov, "Upper bounds for the error probability for channels with feedback," *Prob. Peredachi Inform.*, vol. 6, no. 2, pp. 87–92, 1970.

[88] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback," *IEEE Trans. Inf. Theory*, vol. 12, no. 2, pp. 172–182, Apr. 1966.

[89] T. Richardson, personal communication, 2009.

[90] 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.

[91] T. Fujiwara and T. Kasami, "The weight distribution of (256, k) extended binary primitive BCH codes with $k \leq 63$ and $k \geq 207$," IEICE, Tech. Rep. IT97-46, Sep. 1997.

[92] M. Sugino, Y. Ienaga, N. Tokura, and T. Kasami, "Weight distribution of (128, 64) Reed-Muller code," *IEEE Trans. Inf. Theory*, vol. 17, no. 5, pp. 627–628, 1971.

[93] T. Sugita, T. Kasami, and T. Fujiwara, "The weight distribution of the third-order Reed-Muller code of length 512," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1622–1625, Sep. 1996.

[94] Y. Desaki, T. Fujiwara, and T. Kasami, "The weight distributions of extended binary primitive BCH codes of length 128," *IEEE Trans. Inf. Theory*, vol. 43, no. 4, pp. 1364–1371, Jul. 1997.

[95] M. Terada, J. Asatani, and T. Koumoto. (2010, Apr.) Weight distribution of BCH and Reed-Muller codes. [Online]. Available: http://www.infsys.cne.okayama-u.ac.jp/~kusaka/wd/index.html

[96] A. Cohen and N. Merhav, "Lower bounds on the error probability of block codes based on improvements on de Caen's inequality," *IEEE Trans. Inf. Theory*, vol. 50, no. 2, pp. 290–310, Feb. 2004.

[97] A. N. Kolmogorov, "Über das Gesets des iterierten Logarithmus," *Math. Ann.*, vol. 101, pp. 126–135, 1929.

[98] L. V. Rozovsky, "Estimate from below for large-deviation probabilities of a sum of independent random variables with finite variances (in Russian)," *Zapiski Nauchn. Sem. POMI*, vol. 260, pp. 218–239, 1999.

[99] ——, "Estimate from below for large-deviation probabilities of a sum of independent random variables with finite variances," *J. Math. Sci.*, vol. 109, no. 6, pp. 2192–2209, May 2002.

[100] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Feedback in the non-asymptotic regime," *IEEE Trans. Inf. Theory*, Apr. 2010, submitted for publication.

[101] ——, "Variable-length coding with feedback in the non-asymptotic regime," in *Proc. 2010 IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, TX, Jul. 2010.

[102] R. L. Dobrushin, "Asymptotic bounds on error probability for transmission over DMC with symmetric transition probabilities," *Theor. Probability Appl.*, vol. 7, pp. 283–311, 1962.

[103] M. V. Burnashev, "Data transmission over a discrete channel with feedback. random transmission time," *Prob. Peredachi Inform.*, vol. 12, no. 4, pp. 10–30, 1976.

[104] ——, "Sequential discrimination of hypotheses with control of observations," *Math. USSR, Izvestia*, vol. 15, no. 3, pp. 419–440, 1980.

[105] H. Yamamoto and K. Itoh, "Asymptotic performance of a modified Schalkwijk-Barron scheme for channels with noiseless feedback," *IEEE Trans. Inf. Theory*, vol. 25, no. 6, pp. 729–733, Nov. 1979.

[106] V. D. Goppa, "Nonprobabilistic mutual information with memory," *Probl. Contr. Inf. Theory*, vol. 4, pp. 97–102, 1975.

[107] N. Shulman, "Communication over an unknown channel via common broadcasting," Ph.D. dissertation, Tel-Aviv Univ., Tel-Aviv, Israel, 2003.

[108] S. C. Draper, B. J. Frey, and F. R. Kschischang, "Efficient variable length channel coding for unknown DMCs," in *Proc. 2004 IEEE Int. Symp. Inf. Theory (ISIT)*, Chicago, IL, USA, Jun. 2004.

[109] A. Tchamkerten and E. Telatar, "A feedback strategy for binary symmetric channels," in *Proc. 2002 IEEE Int. Symp. Inf. Theory (ISIT)*, Lausanne, Switzerland, Jun. 2004.

[110] ——, "Optimal feedback schemes over unknown channels," in *Proc. 2004 IEEE Int. Symp. Inf. Theory (ISIT)*, Chicago, IL, USA, Jun. 2004.

[111] ——, "Variable length coding over an unknown channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2126–2145, May 2006.

[112] S. C. Draper, F. R. Kschischang, and B. Frey, "Rateless coding for arbitrary channel mixtures with decoder channel state information," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4119–4133, Sep. 2009.

[113] S. Verdú and S. Shamai, "Variable-rate channel capacity," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2651–2667, Jun. 2010.

[114] D. Baron, S. Sarvotham, and R. Baraniuk, "Coding vs. packet retransmission over noisy channels," in *Proc. 2006 40th Annual Conf. Inf. Sci. and Syst.*, 2006, pp. 537–541.

[115] D. P. Palomar and S. Verdú, "Lautum information," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 964–975, Mar. 2008.

[116] A. Sahai, "Why do block length and delay behave differently if feedback is present?" *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1860–1886, May 2008.

[117] A. N. Shiryaev, *Optimal Stopping Rules*. New York: Springer, 1978.

[118] A. D. Wyner, "On the Schalkwijk-Kailath coding scheme with a peak energy constraint," *IEEE Trans. Inf. Theory*, vol. 14, no. 1, pp. 129–134, Jan. 1968.

[119] R. M. Blumenthal, "An extended Markov property," *Trans. of AMS*, vol. 85, no. 1, pp. 52–72, May 1957.

[120] A. N. Tikhomirov, "On the convergence rate in the central limit theorem for weakly dependent random variables," *Theor. Probability Appl.*, vol. 25, no. 4, 1980.

[121] G. Birkhoff, "Extensions of Jentzsch's theorem," *Trans. of AMS*, vol. 85, pp. 219–227, 1957.

[122] T. Holliday, A. Goldsmith, and P. Glynn, "Capacity of finite state channels based on Lyapunov exponents of random matrices," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3509–3532, Aug. 2006.