Abstract

Title of Dissertation: Compression and Transmission of Information at Multiple Resolutions

Angelos Kanlis, Doctor of Philosophy, 1998

Dissertation directed by: Professor Prakash Narayan
Department of Electrical Engineering and
Institute for Systems Research

This dissertation addresses several problems concerning the compression and transmission of information at multiple resolutions. A multiresolution encoder provides a hierarchically compressed description of the source information. The first stage typically contains a low-quality or coarse description of the source. The subsequent stages contain additional information which serves to enhance the quality of the received or decoded signal.

Some general results regarding the properties of good (i.e., nearly-optimal) block codes set the stage for the subsequent discussion. It is established that the empirical distribution of most codewords of a good code is close to the output distribution induced by the rate-distortion achieving channel. Furthermore, the probability of the source sequences which are not encoded by such codewords is shown to decay exponentially with the block-length.

Next, we characterize the performance of a multiresolution code by defining a rate-distortion function and obtain error exponents for this multistage description. The performance of such a code is also assessed by the rate at which the expected distortion redundancy converges to the theoretical optimum.

The next contribution is related to the problem of hierarchically compressed data over a network with constrained link capacities. The multistage distortion-rate function is used as a measure of the quality of service. An algorithm is presented which ensures fairness by minimizing the maximum distortion among all recipients. A variant of this algorithm serves as a heuristic for the problem of minimizing the average cost over all recipients. Simulation results demonstrate that this heuristic algorithm yields solutions close to the optimal.

Finally, the problem of evaluating the performance of a mismatched block code is addressed in the context of two-stage universal vector quantization. In the first stage of such a scheme, a quantized version of the source statistics is produced; in the second stage, a fixed-rate code matched to the quantized statistics is used. The problem of optimally apportioning a specified bit rate among the two stages of the universal vector quantizer is studied.

Compression and Transmission of Information at Multiple Resolutions

by

Angelos Kanlis

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland at College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1998

Advisory Committee:

Professor Prakash Narayan, Chairman/Advisor
Professor Benjamin Kedem
Associate Professor Thomas E. Fuja
Associate Professor Adrian Papamarcou
Assistant Professor Leandros Tassiulas

Dedication

To my family and friends, who make life worth living

Acknowledgements

I am indebted to my research advisor, Prof. Prakash Narayan, for his constant support, help, and guidance throughout the course of this work. He has been an invaluable source of expert advice, assistance and moral support. His contribution to this dissertation has been immeasurable. He has spent numerous hours working with me on the problems studied in this dissertation and reviewing my thesis. He has been a great inspiration to me and my tenure with him has helped me forge a rigorous scientific style. I feel truly fortunate to have had him as my advisor and I enjoyed having worked under his supervision.

I wish to acknowledge Dr. Leandros Tassiulas, for several discussions on aspects of communications networks and Dr. Tamás Linder for comments on universal coding. I also wish to thank my colleagues and good friends Dr. Sanjeev Khudanpur, Dr. Nikolaos Sidiropoulos, Deepak Ayyagari and Anastassios Michail for their willingness to review parts of my dissertation.

I greatfully acknowledge the financial support provided in the form of a graduate stipend by the Bodosaki Foundation, Greece, in the form of a graduate fellowship by IBM Corporation, and in the form of a graduate research assistantship by the National Scientific Foundation, under grants NSFD CDR-8803012 and

EEC-9402384, NASA, under grant NAGW-27775, and ARL through Lockheed Sanders, under grant QK-9931, in the form of a graduate research assistantship.

I would not have managed to overcome the hardships I encountered during my stay in the United States without the friendship and emotional support of my brother, Nikolaos, and all my close friends. They have contributed each in their own way, helped me maintain my perspective, and constituted an ideal environment for me to work and live in.

Most of all, I wish to express my deepest appreciation and gratitude to my beloved parents, Athanasios and Krystallia. They have instilled their persistence, perseverance and philosophy of life in me, and provided me with all the necessary ingredients to pursue the highest degree in education. Their patience, understanding and love were invaluable assets for me.

Contents

List of Tables

List of Figures

Compression and Transmission
of Information at Multiple Resolutions

Angelos Kanlis

February 3, 1998

This comment page is not part of the dissertation.

Chapter 1

Introduction

1.1 Hierarchical Encoding and Multicasting

Hierarchical encoding is a data compression technique which is widely used in many present-day applications for the purposes of information representation, storage and transmission. Its main feature which renders it appealing is its flexibility, which results from its ability to represent the encoded information in various degrees of quality. The output of a hierarchical encoder comprises an ordered series of blocks (levels) of encoded information. The first block contains a low-quality description and carries only some "basic information." The next levels contain additional information which, when combined with that of all the previous levels, improves the quality of the received/decoded signal. It is at the discretion of the receiver/decoder to decide how many levels of encoded information to use, based on bandwidth limitations, decoding complexity, desired quality or any other performance criterion, as also the nature of the receiving equipment. This is a feature not supported by conventional data compression schemes, where the only option for the receiver is to either accept all the infor-

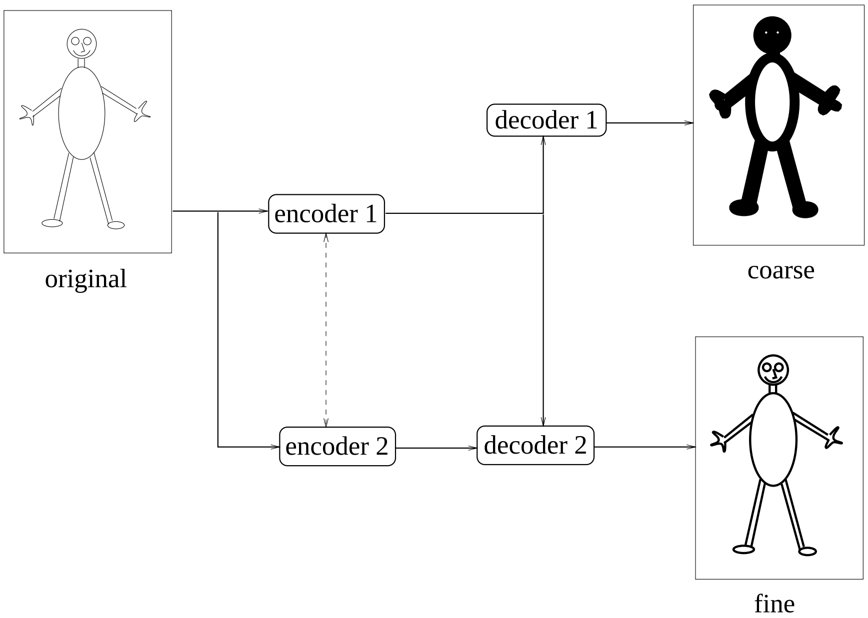

Figure 1.1: A simple example of two-level hierarchical compression.

mation sent by the transmitter or discard it *in toto*, since in this case partial reception of information does not correspond to partial reconstruction.

Figure 1.1 illustrates a simple case of two-level hierarchical image compression, where compression of the original image is performed by encoders 1 and 2. Initially, decoder 1 provides a coarse description of the original image using only the output of encoder 1. Subsequently, if a refined description is desired, decoder 2 is used, which combines the outputs of encoders 1 and 2 (levels 1 and 2 in our hierarchical compression scheme, respectively) and provides a finer representation of the original image. Note that the second encoder does not provide a full high-quality compression. Instead, it only encodes the additional information needed, which, when combined with the output of the first encoder, yields a high-quality description of the original image.

As is to be expected, the flexibility of hierarchical encoding does not come without a price. In most cases, we may have to sacrifice compressibility, namely more information bits are needed in order to achieve this multilevel description. It has been shown, however, that for some of the most common probabilistic source models, we may be able to successively refine the source information (*c.f.*, *e.g.*, [16]). The term "successive refinement" refers to the situation in which it is possible to provide multilevel descriptions of the source information with effectively the same number of bits as with any conventional single-level data compression scheme so that there is no loss in the degree of compressibility, while flexibility is added. Nevertheless, this observation is true only in the limiting case when the block-length of the code becomes extremely large, corresponding to a very high encoding/decoding complexity. In practice the code operates with a relatively small block-length, so that it becomes necessary to examine the rates of convergence of the performance of hierarchical codes to their theoretical limits.

There are several performance measures of interest for which we can study "rates of convergence." The most common is, perhaps, the expected induced distortion, which has been extensively examined for the single-level case. Another performance measure is the so-called "error exponent," *i.e.*, the exponential rate at which the probability that a given code does not have an acceptable behaviour decays to zero. At this point, the notion of an error exponent is rather vague; a rigorous definition is given later in this dissertation. It is important to investigate the various rates of convergence of the performance measures of a given code to their theoretical limits, since they provide significant insights into the way practical data compression schemes should be devised for satisfactory performance.

Hierarchical data compression has numerous applications. Information retrieval from a database is a typical example. A user looking for a specific kind of information can browse through low-quality descriptions of the database contents in order to initially identify items of interest; subsequently, the user can ask for more detail, by requesting additional description levels. Another area in which hierarchical encoding is gaining ground is information transmission over a packet-switched communication network, where the source sends out packets containing multiple descriptions of the source signal. The router at each node encountered by a packet on its way to the destination must then decide whether to send the entire packet through or discard it, depending upon capacity constraints, cost, delay or quality of service.

Multicasting is a network application where multilevel source encoding is usually a desirable feature. The term "multicasting," as opposed to the unicast point-to-point model where the source transmits information to a single user, is the network equivalent of broadcasting and refers to the situation where the source transmits information of the same nature to a number of users or subscribers over a communication network. Since the population of users, in general, is heterogeneous, not all of them desire or are capable of receiving high-quality descriptions. Using a hierarchical encoding representation, the source avoids having to retransmit the same information at a smaller rate. Instead, it is up to the network routers and the subscribers to decide how many levels of information should be delivered to a destination. It is, therefore, essential to investigate the problem of multicasting of hierarchically compressed data and develop efficient techniques for devising optimal routing and rate selection policies. Here, optimal routing deals with the selection of a network link path from the source to the destinations, over which transmission of the encoded information results in the optimization of some performance criterion, *e.g.*, transmission cost. Optimal rate allocation refers to the selection of the compression rate of each level, so as to render the compression scheme more flexible and appropriate for the particular network environment.

A final aspect of data compression, studied in this dissertation, is universal fixed-rate source coding. Many source coding techniques used for both single- and multilevel data representation assume that the encoder has some prior knowledge of the exact source statistics, so that the code is designed to "match" these statistics. However, if the source model used is not accurate, it is possible that the performance of the code designed for this model will, in general, perform poorly under the "true" source statistics. For this reason, several researchers have focused attention on codes which perform almost equally well for a variety of sources, without any effective loss in compressibility. The design of such universal data compression codes is typically carried out in two stages. In the first stage, the encoder attempts to estimate the underlying statistics which describe the information to be encoded and quantizes the statistics parameters. In the second stage, it quantizes (encodes) the source information using a vector-quantizer (code) matched to the quantized estimated statistics. Clearly, the first stage consumes a fraction of the overall available encoding rate which, if the true statistics were known, could be used to represent the source information. On the other hand, it is necessary to allocate a portion of the available representation rate to the first stage; else, we run the risk of contending with a poor description of the source model. This trade-off motivates an interesting problem of apportioning optimally (in a suitable sense) a given number of bits among the two stages. In order to characterize the optimality of a particular rate allocation scheme, a performance criterion is needed, such as the expected distortion induced by the optimal code for these allocated rates. Once such a criterion has been established, the determination of the optimal rate for the first stage (and consequently the optimal rate for the second stage) follows by a straightforward optimization procedure. However, in most cases, it is not possible to characterize the optimal rate for the first stage in closed form; we resort to numerical methods in order to solve this optimization problem.

1.2 Overview of Main Results

This dissertation examines various aspects of hierarchical data compression from an information-theoretic point of view. It assumes that the data to be compressed obey a probabilistic law, and provides theoretical bounds for the expected performance of multilevel codes, namely average performance over all possible source sequences with respect to the assumed probability measure. The following paragraphs summarize the results obtained in the course of this work. A modicum of technical notation and terminology are also used in this introduction, but the statement of results is descriptive and qualitative, rather than quantitative.

1.2.1 Properties of a Good Rate-Distortion Code

We begin by examining some properties of rate-distortion codes which are designed for a known source model (*e.g.*, with known probability mass function (pmf) P) and whose performance converges to the optimal theoretically achievable performance as their complexity becomes arbitrarily large. These results are essential for our subsequent study of hierarchical coding problems. An important benchmark in source coding is the *rate-distortion function*, denoted $R(P, \Delta)$, which provides the smallest rate or number of bits per symbol in a code which represents a typical source output within some given distortion Δ, and provides a characterization of the trade-off between rate and quality of description. In the definition of the rate-distortion function, we minimize the mutual information between the source and reproduction random variables over all channels (conditional probability distributions) connecting them, subject to certain conditions. For most sources and distortion measures, the minimum is achieved by a single channel, which will be referred to as the *rate-distortion achieving channel*, W^*.

We consider a good code for a discrete memoryless source with a specified distortion level to be one whose rate is close to the corresponding rate-distortion function and which, with large probability, reproduces the source within the allowed distortion level. We show that for any good code, the number of codewords whose empirical distribution is close to $P \cdot W^*$, the output distribution induced by the rate-distortion achieving channel W^*, is exponentially large in the block-length of the code with effectively the same rate (exponent). Thus, for most codewords, the relative frequency of an output symbol in a codeword is close to the probability of that symbol under the output distribution $P \cdot W^*$.

We also show that the probability of the set of source sequences which are mapped to codewords not endowed with the previous property (*i.e.*, empirical distributions close to $P \cdot W^*$) is exponentially decaying in the block-length of the code. Similarly, an exponential decay is established for the probability of the set of source sequences whose joint empirical distribution with their corresponding codewords is far from PW^*, the joint distribution induced by the source pmf P and the minimizing channel W^*.

Furthermore, the output distribution induced by a good code is asymptotically singular with respect to the independent and identically distributed (i.i.d) output distribution induced by the rate-distortion achieving channel. However, the normalized (Kullback-Leibler) divergence between these output distributions converges to the conditional entropy of the output under the rate-distortion achieving channel. In other words, although each codeword looks as if it were produced by random selection using $P \cdot W^*$, the sequence of (individual) output symbols does not resemble an i.i.d. sequence; clearly, only some of the set of *all* possible sequences, namely the codewords, are seen; these are "typical," *i.e.*, have composition close to $P \cdot W^*$.

1.2.2 Error Exponents for Hierarchical Coding

As stated previously, a good code must have the desirable property of being able to reproduce the source within an acceptable distortion level with high probability. Therefore, we can define the *error event* for such a code to be the event that the distortion induced by the code for a source sequence exceeds the acceptable level. It has been shown (*e.g.*, [52, 13]) that, depending on the rate of the code, the probability of this error event will converge either to zero exponentially fast with the block-length or to one, as the block-length grows larger. The exponential rate at which the error decays to zero, also known as *error exponent*, is a performance measure for rate-distortion codes, and has been completely characterized for the single-level case (*c.f.*, [40]). The higher the error exponent, the faster the probability with which the error event goes to zero;

consequently, codes with higher error exponents require smaller block-lengths in order to achieve the same error probability, and hence have smaller complexity.

Since this work focuses on multilevel rate-distortion coding, we wish to define the same quantities for this case as well. The rate-distortion characterization of multilevel coding has effectively been carried out by other researchers (*e.g.*, [15, 46]). Based on their results, we define the multilevel rate-distortion function which, for two levels of compression, is formally defined as follows : Given a discrete memoryless source (DMS) with probability mass function (pmf) P, we seek first an asymptotically optimal description of the source with distortion not exceeding Δ_1, followed by an asymptotically optimal refined description with distortion not exceeding $\Delta_2 < \Delta_1$. The two-level rate-distortion function, denoted $R(P, \Delta_1, \Delta_2)$, is the overall optimal rate of these descriptions obtained via a two-level coding process.

We provide a "covering lemma" for two-level coding, namely we establish the existence of two codeword sets which "cover" the set of the "most likely" source sequences within distortions Δ_1 and Δ_2 and whose cardinalities are exponential in the block-length with exponents (rates) effectively equal to the one- and two-level rate-distortion functions. In particular, we show that there exists a codeword set $\mathcal{C}_1$ of rate arbitrarily close to $R(P, \Delta_1)$ with the property that for every source sequence with empirical distribution close to P, there exists a codeword from $\mathcal{C}_1$ with distortion not exceeding Δ_1. Furthermore, there exists a second codeword set $\mathcal{C}_2$ of rate arbitrarily close to $R(P, \Delta_1, \Delta_2)$, which is a collection of smaller disjoint codeword sets, i.e.,

$$\mathcal{C}_2 = \bigcup_{\mathbf{c}_1 \in \mathcal{C}_1} \mathcal{C}_2(\mathbf{c}_1).$$

The sets $\mathcal{C}_2(\mathbf{c}_1)$, $\mathbf{c}_1 \in \mathcal{C}_1$, possess the property that each of them covers the set of source sequences (in the same sense in which $\mathcal{C}_1$ covers the set of source sequences) which are encoded by $\mathbf{c}_1$, (*i.e.*, are closer to $\mathbf{c}_1$ than to any other codeword from $\mathcal{C}_1$) at distortion Δ_2.

Furthermore, we determine the error exponents for this two-level coding process, namely the asymptotic normalized loglikelihoods of the event that the distortion in either level exceeds its prespecified acceptable value, and of the conditional event that the distortion in the second level exceeds the prespecified value given the rate and distortion of the code for the first level.

In general, the minimum rate required in order to describe the DMS in two levels (given by the two-level rate-distortion function) is no less than that required for the single-level description (given by the conventional, single-level rate-distortion function). As has been shown by other researchers, the two required minimum rates (rate-distortion functions) are identical for some of the most common source-distortion measure combinations, *e.g.*, for a discrete source with the Hamming distortion measure and for a Gaussian source with the squared difference distortion measure.

We provide a simple example of a DMS with Hamming distortion measure showing that even when the rate-distortion functions for one- and two-level coding coincide, the error exponent in the former case may exceed those in the latter. Thus, for a given distortion level, although the sizes of the codewords for one- and two-level coding may be the same, an error event is more likely in the latter situation.

1.2.3 Distortion Redundancy of Hierarchical Codes

The error exponent is an important performance measure for block codes. Another measure of interest is the rate at which the expected distortion of a block code converges to its optimal value, namely the distortion-rate function.

First, we prove the existence of a single-level block code, based on the work of Pilc in [45], where the existence was shown of a code whose expected distortion converges to the distortion-rate function as $O(\frac{\log n}{n})$, with n being the block-length. We provide a clearer presentation of Pilc's proof, which is based on the notion of types (*c.f.*, *e.g.*, Csiszár and Körner [13]). This makes it easier to extend the proof to the multilevel case, for which it is also shown that the rate of convergence is again $A\frac{\log n}{n}$, but with an expectedly larger constant A.

1.2.4 Multicasting of Hierarchical Information

Next, we examine the problem of the multicasting of a hierarchically compressed source signal over a network taking into account link capacity constraints. In this communication protocol, a source node transmits the same information via network links to a number of user nodes who have requested it. We examine a situation in which, every node of the network is a receiver and, at the same time, a router which retransmits the received information to its neighboring nodes further down in the multicast tree.

As has been pointed out previously, the hierarchical compression of source information is an important feature for multicasting in a heterogeneous network. Based on the network topology, the source has to decide on the routes over which to send the various compression levels (which need not be sent over the same route, as is the case in the previously published work by N. Shacham [49]) and the compression rate to use for each level so that the levels reach as many users as possible.

In order to evaluate the quality of service provided by a particular routing and rate selection policy, we adopt as a performance measure, the multilevel distortion-rate function for hierarchical compression. According to the number of levels each node receives, the corresponding value of the multilevel distortion-rate function indicates the quality (in fact, the "poorness") of the received information. We address two different cases : the minimax and the average cost problems. In the former situation, the optimal policy is determined by trying to minimize the distortion at the node with the maximum value of the distortion-rate function. Thus, fairness is guaranteed to all users by providing them, in effect, with the same amount of information. In the latter case, the average quality of reception is maximized over all users, *i.e.*, we try to minimize the averaged value (over all users) of the distortion-rate function.

We present an iterative algorithm which solves the minimax cost problem. Each iteration comprises two simple steps : First, the maximum capacity tree (MCT) from the source to the users is determined for each level. Then, the rate of each level is set equal to the minimum capacity of the routing tree of each level, and this rate is subtracted from the available link capacities for the links used heretofore. These two steps are repeated for all the subsequent levels.

The average cost problem is much harder, and is amenable only to NP-complete algorithms. Instead, we propose a variant of the algorithm for the minimax problem, which serves as a heuristic for this case. Instead of examining all possible routing trees, the algorithm restrictively uses the MCT's for the routing of the various levels, and looks through all possible compression rates to determine the rate selection which induces minimal average distortion.

In order to exhibit the efficiency of our algorithms, we provide simulation results for randomly generated networks. We compare these results with those of the algorithm proposed in [49], which assumes that the various levels are routed over the same paths in the network. The simulation results also show for the average distortion problem that our proposed heuristic algorithm yields solutions close to the optimal.

1.2.5 Universal Rate-Distortion Coding

In the last part of this dissertation, we address the issue of optimal rate allocation between the two stages of a universal vector quantizer, where a quantized description of the source statistics is produced in the first stage, and the data sequence is compressed in the second stage according to the output of the first stage.

First, we investigate the problem of performance evaluation of a mismatched fixed-rate code, namely a code with fixed block-length and rate, which is designed for source statistics different from the true statistics. This problem has already been addressed in the variable-length noiseless coding context, where the excess average length due to mismatch is given by the Kullback-Leibler divergence between the assumed and the true source pmf's. We introduce the notion of the rate-distortion discrepancy between two pmf's and show that it is a lower bound to the expected excess rate due to the mismatch for every fixed-rate code. We also conjecture that the expected excess rate, under the true pmf, of any "good code" which is matched to a different pmf is bounded above by the same quantity. This is followed by the demonstration of a code whose performance is close to the optimal under the “wrongly assumed” statistics, and asymptotically approaches the mismatched optimal performance under the true statistics.

Next, based on the previous set of ideas, we study a two-stage universal vector quantization scheme. In the first stage of such a scheme, a quantized version of the source statistics is produced; in the second stage, a fixed-rate code matched to the quantized statistics is used. We define the expected excess rate for this coding scheme and address the problem of optimal rate allocation between the first and second stages. It is shown that the optimal first-stage rate should grow linearly with the logarithm of the block-length, which formally validates similar assertions made by other researchers. Finally, we consider as an example the simple case of a binary source with the Hamming distance distortion measure, and numerically verify our theoretical assertions.

Chapter 2

Typicality of a Good Rate-Distortion Code

2.1 Introduction

In this chapter, we establish some properties of a "good" code for lossy data compression. In particular, the composition of a good code and its probabilistic characteristics, are studied. These properties, which are technical in nature, will prove useful in some of the subsequent chapters. The contents of Sections 2.3 and 2.4, which have been published in [28], resulted from a collaborative effort which includes S. Khudanpur.

A good code for a discrete memoryless source (DMS) with a specified distortion level is one whose rate is close to the corresponding rate-distortion function and which, with large probability, reproduces the source within the allowed distortion level. The Covering Lemma of Rate-Distortion Theory (*c.f.*,*e.g.*, [13, Lemma 4.1, p. 150]) asserts the existence of a good code, all of whose codewords are typical with respect to the *optimal* distribution on the reproduction alphabet, namely that induced by the rate-distortion achieving channel. We show that *any* good code must contain an exponentially large set of codewords which are *all* typical (in the same sense as above), and is of effectively the same rate. Next, this source code gives rise to a deterministic channel between the spaces of all the input and output sequences, which in turn induces an output distribution on the latter. Since the output distribution corresponding to this code restricts its mass to the set of codewords, it differs significantly from the independent and identically distributed (i.i.d.) measure induced by the optimal output distribution. It is shown that these two measures are asymptotically singular, except in trivial cases. Furthermore, under an additional assumption of the "minimality" of a good code, the normalized (Kullback-Leibler) divergence between these output distributions converges to the conditional entropy of the output under the rate-distortion achieving channel. As is to be expected, this behavior of a good source code is in contrast with that of a good *channel* code for which, as shown by Han-Verdú [25], the normalized divergence between the corresponding output distribution and the optimal output distribution (induced by the capacity-achieving input distribution) vanishes asymptotically.

Next, we sharpen our result concerning the typicality of the codewords of a good code, by showing that the probability of the source sequences which are mapped to nontypical codewords decays exponentially with the block length of the code. Similarly, exponential decay is established for the probability of source sequences which are jointly nontypical with their corresponding codewords, where joint typicality is with respect to the joint distribution induced by the source pmf and the rate distortion-achieving channel.

2.2 A Few Comments on Notation

For the sake of easy reference, we summarize below some useful notation which recurs throughout this dissertation.

Let P be a pmf on the finite set $\mathcal{X}$. Let $\mathcal{Y}$ be the (finite) reproduction alphabet, and $W : \mathcal{X} \mapsto \mathcal{Y}$ a stochastic matrix (channel). We define $P \cdot W$ to be the pmf on $\mathcal{Y}$ induced by P and W, *i.e.*,

$$(P \cdot W)(y) \triangleq \sum_{x \in \mathcal{X}} P(x)W(y|x), \quad y \in \mathcal{Y}.$$

We also define PW to be the joint pmf on $\mathcal{X} \times \mathcal{Y}$, *i.e.*,

$$(PW)(y) \triangleq P(x)W(y|x), \quad x \in \mathcal{X}, y \in \mathcal{Y}.$$

Now, let (X, Y) be a $(\mathcal{X} \times \mathcal{Y})$-valued random variable with joint pmf PW, and let $d : \mathcal{X} \times \mathcal{Y} \mapsto \mathbf{R}^+$ be a nonnegative-valued mapping. We define the expected distortion $d(P, W)$ as

$$d(P, W) \triangleq \mathbf{E}d(X, Y) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x)W(y|x)d(x, y).$$

We also define the joint entropy $H(P, W)$ and mutual information $I(P, W)$ in a similar manner, as

$$\begin{aligned} H(P, W) &\triangleq H(X, Y) &= -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x)W(y|x) \log(P(x)W(y|x)), \\ I(P, W) &\triangleq I(X \wedge Y) &= \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P(x)W(y|x) \log \frac{W(y|x)}{(P \cdot W)(y)}. \end{aligned}$$

Let $\mathbf{x} \in \mathcal{X}^n$ be a sequence of length n. We shall denote its *empirical pmf* by $T_{\mathbf{x}}$, *i.e.*,

$$T_{\mathbf{x}}(x) \triangleq \frac{1}{n} \sum_{t=1}^{n} \mathbf{1}(x_t = x), \quad x \in \mathcal{X},$$

where x_t is the tth component of the sequence $\mathbf{x}$, and $\mathbf{1}(\cdot)$ is the indicator function. We shall refer to $T_{\mathbf{x}}$ as the *type* of $\mathbf{x}$.

Let Q be a pmf on $\mathcal{X}$ such that $nQ(x)$ is an integer for every $x \in \mathcal{X}$. Then Q is a possible type on $\mathcal{X}$ and we shall denote by $\mathcal{T}_Q^n$ the set of all sequences of length n with empirical pmf Q, *i.e.*,

$$\mathcal{T}_Q^n \triangleq \{\mathbf{x} \in \mathcal{X}^n : T_{\mathbf{x}} = Q\}.$$

We shall also use $\mathcal{P}_n$ to denote the set of all possible types on $\mathcal{X}$. It has been shown (*cf.*, [13, Lemma 2.2, p. 29, Lemma 2.3, p. 30]) that the size of $\mathcal{P}_n$ is polynomial in n, whereas the size of $\mathcal{T}_Q^n$ is exponential in n.

The set $\mathcal{T}_Q^n$ is called the *Q-type set*. We also define the *Q-typical set* $\mathcal{T}_{[Q]}^n$ to be the set of all sequences whose type is close to Q, *i.e.*,

$$\mathcal{T}_{[Q]}^n \triangleq \{\mathbf{x} \in \mathcal{X}^n : |T_{\mathbf{x}}(x) - Q(x)| \leq \delta, x \in \mathcal{X}\}.$$

The "thickness" δ of the typical set usually follows the "delta convention" (*c.f.*, [13, eq. (1.2.9), p.34]), where δ is chosen to decay "suitably fast" to zero as n grows to infinity.

Let $V : \mathcal{X} \mapsto \mathcal{Y}$ be a channel from the input alphabet $\mathcal{X}$ to the output alphabet $\mathcal{Y}$, and let $\mathbf{x} \in \mathcal{X}^n$ be a n-length input sequence. We define the *V-shell* of $\mathbf{x}$ to be the set of output n-length sequences, whose joint empirical distribution with $\mathbf{x}$ is $T_{\mathbf{x}}V$, *i.e.*,

$$\mathcal{T}_V^n(\mathbf{x}) \triangleq \{\mathbf{y} \in \mathcal{Y}^n : T_{\mathbf{xy}} = T_{\mathbf{x}}V\}.$$

We can also define the V-typical set $\mathcal{T}_{[V]}^n(\mathbf{x})$ in a manner similar to the definition of the Q-typical set $\mathcal{T}_{[Q]}^n$.

For a more detailed presentation of types and a summary of the main results the reader is referred to [13, Chapter 1.2].

2.3 Preliminaries and Main Results

We have adopted the terminology of Csiszár-Körner [13]. In particular, all logarithms and exponentials are taken with respect to the base 2.

Let $\{X_t\}_{t=1}^{\infty}$ be a DMS with finite alphabet $\mathcal{X}$, *i.e.*, an independent and identically distributed (i.i.d.) process, with common probability mass function (pmf) P, where $P(x) > 0,\ x \in \mathcal{X}$. Let $\mathcal{Y}$ be a finite reproduction alphabet. Let $d : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}^+$ be a nonnegative-valued mapping with $\min_{x,y} d(x,y) = 0$ and $d_{max} \triangleq \max_{x,y} d(x,y) < \infty$. This mapping induces a distortion measure on $\mathcal{X}^n \times \mathcal{Y}^n$ according to

$$d(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{t=1}^{n} d(x_t, y_t), \quad \mathbf{x} \in \mathcal{X}^n,\ \mathbf{y} \in \mathcal{Y}^n.$$

A n-length block code consists of two mappings: An encoder $f_n : \mathcal{X}^n \mapsto \{1, \ldots, M_n\}$ and a decoder $\phi_n : \{1, \ldots, M_n\} \mapsto \mathcal{Y}^n$. The rate of this code is $R_n = \frac{1}{n} \log M_n$.

For $\Delta > 0$, the rate distortion function, $R(P, \Delta)$, characterizing the minimum achievable rate for a distortion Δ, is well known and given by

$$R(P, \Delta) = \min_{W : d(P,W) \leq \Delta} I(P, W), \tag{2.1}$$

where W ranges over all stochastic matrices $W : \mathcal{X} \mapsto \mathcal{Y}$. It will be assumed throughout that the minimizing W in (2.1), denoted W^*, is unique.

We shall hereafter be interested only in a "good" code, *i.e.*, a code which with large probability reproduces sequences from $\mathcal{X}^n$ with distortion no greater than a specified $\Delta > 0$ and has a rate arbitrarily close to $R(P, \Delta)$. Precisely, a *good code* is defined as a sequence of n-length block codes (f_n, ϕ_n) with the following two properties: For every $\epsilon > 0$, $\alpha > 0$, and n sufficiently large (depending on

$\epsilon, \alpha)$,

$$P^n \left(\mathbf{x} \in \mathcal{X}^n \ : \ d(\mathbf{x}, \phi_n(f_n(\mathbf{x}))) \leq \Delta \right) \geq 1 - \epsilon, \tag{2.2}$$

where

$$P^n(\mathbf{x}) = \prod_{t=1}^{n} P(x_t), \quad \mathbf{x} \in \mathcal{X}^n,$$

and

$$R(P, \Delta) < R_n \leq R(P, \Delta) + \alpha. \tag{2.3}$$

We shall use $\mathcal{C}_n = \{\mathbf{y}_i \in \mathcal{Y}^n : \mathbf{y}_i = \phi_n(i),\ 1, \ldots, M_n\}$ to denote the codewords of the n-length code (f_n, ϕ_n). These codewords induce a partition $\{\mathcal{A}_i\}_{i=1}^{M_n}$ of $\mathcal{X}^n$, where

$$\mathcal{A}_i = \{\mathbf{x} \in \mathcal{X}^n \ : \ \phi_n(f_n(\mathbf{x})) = \mathbf{y}_i\}, \quad i = 1, \ldots, M_n.$$

Of particular interest is the family of (disjoint) subsets $\{\mathcal{B}_i\}_{i=1}^{M_n}$ of $\mathcal{X}^n$ defined by

$$\mathcal{B}_i = \{\mathbf{x} \in \mathcal{A}_i \ : \ d(\mathbf{x}, \mathbf{y}_i) \leq \Delta\}, \quad i = 1, \ldots, M_n.$$

By virtue of property (2.2) of a good code, and the fact that $\{\mathcal{A}_i\}_{i=1}^{M_n}$ is a partition of $\mathcal{X}^n$, note that for every $\epsilon > 0$,

$$P^n \left(\bigcup_{i=1}^{M_n} \mathcal{B}_i \right) \geq 1 - \epsilon, \tag{2.4}$$

for all n sufficiently large (depending on ϵ).

We recall from [13] that the *type* of $\mathbf{x} \in \mathcal{X}^n$ is a pmf $T_{\mathbf{x}}$ on $\mathcal{X}$ where $T_{\mathbf{x}}(x)$ is the relative frequency of the symbol x in $\mathbf{x}$. For any type Q on $\mathcal{X}$, let $\mathcal{T}_Q^{(n)}$ denote the set of all sequences $\mathbf{x} \in \mathcal{X}^n$ with $T_{\mathbf{x}} = Q$. (A similar notation will be used for types on $\mathcal{Y}$.) For any pmf Q on $\mathcal{X}$, pmf S on $\mathcal{Y}$, and $\Delta > 0$, define

$$\tilde{R}(Q, S, \Delta) = \min_{W : d(Q,W) \leq \Delta,\ Q \cdot W = S} I(Q, W),$$

where W ranges over stochastic matrices $W : \mathcal{X} \mapsto \mathcal{Y}$. Note that $\tilde{R}(Q, S, \Delta)$ is well-defined since $I(Q, W)$ is a convex function of W and $\{W : d(Q, W) \leq \Delta,\ Q \cdot W = S\}$ is a convex compact set.

We present below a technical result needed to prove our main result in the subsequent Proposition 2.1. A refined version of this result can be found in the earlier work of Zhang-Yang-Wei [63, Lemma 3]. Let $|\cdot|$ denote cardinality of sets and $H(\cdot)$ denote entropy.

Lemma 2.1 *Let Q be any type on $\mathcal{X}$. Then, for $i = 1, \ldots, M_n$, it holds that*

$$\left|\mathcal{B}_i \bigcap \mathcal{T}_Q^{(n)}\right| \leq (n+1)^{|\mathcal{X}||\mathcal{Y}|} \exp\left[n\left(H(Q) - \tilde{R}(Q, T_{\mathbf{y}_i}, \Delta)\right)\right]. \tag{2.5}$$

Proof: For each conditional type V on $\mathcal{X}$ given $\mathbf{y}_i \in \mathcal{Y}^n$, let $\mathcal{T}_V^{(n)}(\mathbf{y}_i)$ denote the V-shell of $\mathbf{y}_i$ (*c.f.*, *e.g.*, [13]). Then for $i = 1, \ldots, M_n$,

$$\mathcal{B}_i \bigcap \mathcal{T}_Q^{(n)} = \bigcup_{V:\ d(T_{\mathbf{y}_i}, V) \leq \Delta,\ T_{\mathbf{y}_i} \cdot V = Q} \left(\mathcal{B}_i \bigcap \mathcal{T}_V^{(n)}(\mathbf{y}_i)\right),$$

where $d(T_{\mathbf{y}_i}, V) = \sum_{x \in \mathcal{X},\ y \in \mathcal{Y}} T_{\mathbf{y}_i}(y) V(x|y) d(x, y)$ and $T_{\mathbf{y}_i} \cdot V$ denotes a pmf on $\mathcal{X}$ given by

$$T_{\mathbf{y}_i} \cdot V(x) = \sum_{y \in \mathcal{Y}} T_{\mathbf{y}_i}(y) V(x|y), \quad x \in \mathcal{X}.$$

Consequently, using standard bounds on types (*c.f.*, *e.g.*, [13]) we obtain

$$\begin{aligned}
&\left|\mathcal{B}_i \bigcap \mathcal{T}_Q^{(n)}\right| \\
&= \sum_{V:\ d(T_{\mathbf{y}_i}, V) \leq \Delta,\ T_{\mathbf{y}_i} \cdot V = Q} \left|\mathcal{B}_i \bigcap \mathcal{T}_V^{(n)}(\mathbf{y}_i)\right| \\
&\leq \sum_{V:\ d(T_{\mathbf{y}_i}, V) \leq \Delta,\ T_{\mathbf{y}_i} \cdot V = Q} \left|\mathcal{T}_V^{(n)}(\mathbf{y}_i)\right| \\
&\leq (n+1)^{|\mathcal{X}||\mathcal{Y}|} \max_{V:\ d(T_{\mathbf{y}_i}, V) \leq \Delta,\ T_{\mathbf{y}_i} \cdot V = Q} \left|\mathcal{T}_V^{(n)}(\mathbf{y}_i)\right| \\
&\leq (n+1)^{|\mathcal{X}||\mathcal{Y}|} \max_{V:\ d(T_{\mathbf{y}_i}, V) \leq \Delta,\ T_{\mathbf{y}_i} \cdot V = Q} \exp\left[nH(V|T_{\mathbf{y}_i})\right]
\end{aligned}$$

$$= (n+1)^{|\mathcal{X}||\mathcal{Y}|} \max_{V:\ d(T_{\mathbf{y}_i},V)\le\Delta,\ T_{\mathbf{y}_i}\cdot V=Q} \exp\left[n\left(H(T_{\mathbf{y}_i}\cdot V) - I(T_{\mathbf{y}_i},V)\right)\right]$$
$$= (n+1)^{|\mathcal{X}||\mathcal{Y}|} \max_{W:\ d(Q,W)\le\Delta,\ Q\cdot W=T_{\mathbf{y}_i}} \exp\left[n\left(H(Q) - I(Q,W)\right)\right]$$

from which the assertion of Lemma 2.1 follows. ■

At this juncture it is useful to recall two distance measures for pmf's on $\mathcal{X}$ (or $\mathcal{Y}$). The *Kullback-Leibler divergence* between the pmf's P and Q on $\mathcal{X}$ is defined as:

$$D(P\|Q) = \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)},$$

where the usual convention applies that $0\log 0 = 0$, $0\log\left(\frac{0}{0}\right) = 0$ and $t\log\left(\frac{t}{0}\right) = +\infty,\ t > 0$. The *variational distance* between P and Q is defined as

$$\|P - Q\| = \sum_{x\in\mathcal{X}} |P(x) - Q(x)|.$$

Proposition 1 below states that of the roughly $\exp[nR(P,\Delta)]$ codewords of a good code, an exponentially large subset, of effectively the same rate, has codewords whose types asymptotically approach $P\cdot W^*$ in variational distance. Clearly, not all codewords of a good code need possess this property.

Let $\mathcal{C}_n^T$ denote the subset of codewords which are $P\cdot W^*$-typical. Precisely, given $\gamma > 0$, define $\mathcal{C}_n^T = \mathcal{C}_n^T(\gamma)$ by

$$\mathcal{C}_n^T = \left\{\mathbf{y}\in\mathcal{C}_n : \|T_{\mathbf{y}} - P\cdot W^*\| \le \gamma\right\}. \tag{2.6}$$

Proposition 2.1 *For every* $\gamma > 0$

$$\lim_{n\to\infty}\frac{1}{n}\log\left|\ \mathcal{C}_n^T(\gamma)\right| = R(P,\Delta).$$

Proof: For any $\delta > 0$ and $\zeta > 0$, it holds for all n sufficiently large (depending on δ,ζ) that

$$P^n\left(\mathcal{T}_{[P]}^{(n)}\right) \ge 1-\zeta,$$

where

$$\mathcal{T}_{[P]}^{(n)} = \bigcup_{Q:\|P-Q\|\leq\delta} \mathcal{T}_Q^{(n)},$$

with Q denoting types on $\mathcal{X}$. A suitable choice of δ and ζ will be made later.

In conjunction with (2.4), and using standard bounds for types, we get for all n sufficiently large (depending on ϵ, δ, ζ) that

$$\begin{aligned}
1-\epsilon-\zeta &\leq P^n\left(\mathcal{T}_{[P]}^{(n)} \bigcap \left(\bigcup_{i=1}^{M_n} \mathcal{B}_i\right)\right) \\
&= \sum_{i=1}^{M_n} P^n\left(\mathcal{T}_{[P]}^{(n)} \bigcap \mathcal{B}_i\right) \\
&= \sum_{i=1}^{M_n} \sum_{Q:\|P-Q\|\leq\delta} P^n\left(\mathcal{T}_Q^{(n)} \bigcap \mathcal{B}_i\right) \\
&= \sum_{i=1}^{M_n} \sum_{Q:\|P-Q\|\leq\delta} \left|\mathcal{T}_Q^{(n)} \bigcap \mathcal{B}_i\right| \times \exp\left[-n\left(H(Q)+D(Q\|P)\right)\right] \\
&\leq \sum_{i=1}^{M_n} \sum_{Q:\|P-Q\|\leq\delta} (n+1)^{|\mathcal{X}||\mathcal{Y}|} \exp\left[n\left(H(Q)-\tilde{R}(Q,T_{\mathbf{y}_i},\Delta)\right)\right] \\
&\qquad\qquad \times \exp\left[-n\left(H(Q)\right)\right], \quad \text{by (2.5)} \\
&= \sum_{i=1}^{M_n} \sum_{Q:\|P-Q\|\leq\delta} (n+1)^{|\mathcal{X}||\mathcal{Y}|} \exp\left[-n\left(\tilde{R}(Q,T_{\mathbf{y}_i},\Delta)\right)\right] \\
&\leq \sum_{i=1}^{M_n} (n+1)^{|\mathcal{X}|(|\mathcal{Y}|+1)} \exp\left[-n\left(\min_{Q:\|P-Q\|\leq\delta} \tilde{R}(Q,T_{\mathbf{y}_i},\Delta)\right)\right]. \qquad (2.7)
\end{aligned}$$

Given any $\eta > 0$, observe by the continuity of $\tilde{R}(Q, T_{\mathbf{y}_i}, \Delta)$ in Q that $\delta > 0$ can be chosen to satisfy

$$\min_{Q:\|P-Q\|\leq\delta(\eta)} \tilde{R}(Q,T_{\mathbf{y}_i},\Delta) \geq \tilde{R}(P,T_{\mathbf{y}_i},\Delta) - \eta.$$

Continuing the bounding in (2.7), we obtain

$$1-\epsilon-\zeta \leq (n+1)^{|\mathcal{X}|(|\mathcal{Y}|+1)} \sum_{i=1}^{M_n} \exp\left[-n\left(\tilde{R}(P,T_{\mathbf{y}_i},\Delta)-\eta\right)\right]. \qquad (2.8)$$

Consider first the sum in (2.8) over those indices i for which $\mathbf{y}_i \in \mathcal{C}_n^T$. Note that $\tilde{R}(P, T_{\mathbf{y}_i}, \Delta) \geq R(P, \Delta)$ and therefore

$$\sum_{i:\mathbf{y}_i\in\mathcal{C}_n^T} \exp\left[-n\left(\tilde{R}(P,T_{\mathbf{y}_i},\Delta)-\eta\right)\right] \leq |\ \mathcal{C}_n^T\ | \exp\left[-n\left(R(P,\Delta)-\eta\right)\right].$$

Next, the sum in (2.8) over those indices i for which $\mathbf{y}_i \notin \mathcal{C}_n^T$ can be bounded above as follows:

$$
\begin{aligned}
&\sum_{i:\mathbf{y}_i \notin \mathcal{C}_n^T} \exp\left[-n\left(\tilde{R}(P, T_{\mathbf{y}_i}, \Delta) - \eta\right)\right] \\
&\leq \exp\left[n\left(R(P,\Delta) + \alpha\right)\right] \exp\left[-n \min_{i:\mathbf{y}_i \notin \mathcal{C}_n^T} \left(\tilde{R}(P, T_{\mathbf{y}_i}, \Delta) - \eta\right)\right], \quad \text{by (2.3)} \\
&= \exp\left[-n\left(\min_{i:\mathbf{y}_i \notin \mathcal{C}_n^T} \tilde{R}(P, T_{\mathbf{y}_i}, \Delta) - R(P,\Delta) - \alpha - \eta\right)\right]. \qquad (2.9)
\end{aligned}
$$

Next, observe that $\tilde{R}(P, T_{\mathbf{y}_i}, \Delta) \geq R(P,\Delta)$ with equality iff $T_{\mathbf{y}_i} = P \cdot W^*$. Together with the condition $\|T_{\mathbf{y}_i} - P \cdot W^*\| > \gamma$ and the assumed uniqueness of W^*, there exists $\theta = \theta(\gamma) > 0$ such that

$$
\min_{i:\mathbf{y}_i \notin \mathcal{C}_n^T} \tilde{R}(P, T_{\mathbf{y}_i}, \Delta) - R(P,\Delta) > \theta.
$$

Therefore, given $\gamma > 0$, we can pick $\alpha = \alpha(\gamma)$, $\eta(\gamma)$ such that $\alpha + \eta < \frac{\theta}{3}$. Then the right side in (2.9) is bounded above by $\exp\left[-n\frac{2\theta}{3}\right]$ for all n sufficiently large (depending on γ). Using this in (2.8), we obtain

$$
1 - \epsilon - \zeta \leq (n+1)^{|\mathcal{X}|(|\mathcal{Y}|+1)} \left|\mathcal{C}_n^T\right| \exp\left[-n\left(R(P,\Delta) - \eta\right)\right] + \exp\left[-n\frac{\theta}{3}\right],
$$

which can be rearranged as

$$
R(P,\Delta) - \xi \leq \frac{1}{n} \log \left| \mathcal{C}_n^T \right|,
$$

for all n sufficiently large (depending on γ), where

$$
\xi = |\mathcal{X}|(|\mathcal{Y}|+1)\frac{\log(n+1)}{n} + \eta + \frac{1}{n}\log\left(1 - \epsilon - \zeta - \exp\left(-n\frac{\theta}{3}\right)\right).
$$

Observe that ξ can be made arbitrarily small, for all n sufficiently large (depending on ξ, γ, ζ) by choosing η to be, say, the smaller of $\frac{\xi}{2}$ and $\frac{\theta(\gamma)}{4}$. Combining this with (2.3) yields the assertion of the Proposition. ■

The sequence of n-length block codes (f_n, ϕ_n) constituting a good code induces a (deterministic) channel $\{\tilde{W}^{(n)} : \mathcal{X}^n \mapsto \mathcal{Y}^n\}_{n=1}^{\infty}$, given by

$$\tilde{W}^{(n)}(\mathbf{y}|\mathbf{x}) = 1(\mathbf{y} = \phi_n(f_n(\mathbf{x}))), \quad \mathbf{x} \in \mathcal{X}^n, \ \mathbf{y} \in \mathcal{Y}^n,$$

where $1(\cdot)$ denotes the indicator function. As a consequence of Proposition 1, we shall characterize how "distant" the sequence of pmf's $P^n \cdot \tilde{W}^{(n)}$ on $\mathcal{Y}^n$, induced by a good code, can be from the sequence of pmf's $(P \cdot W^*)^n$ on $\mathcal{Y}^n$, induced by the rate distortion achieving stochastic matrix W^*. We note that the pmf's under consideration are given by

$$\begin{aligned}
\left(P^n \cdot \tilde{W}^{(n)}\right)(\mathbf{y}) &= \sum_{\mathbf{x} \in \mathcal{X}^n} \left(P^n(\mathbf{x}) \tilde{W}^{(n)}\right)(\mathbf{y}|\mathbf{x}) \\
&= \begin{cases} P^n(\mathcal{A}_i) & \text{if } \mathbf{y} = \mathbf{y}_i \text{ for some } i \in \{1, \ldots, M_n\} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

and

$$(P \cdot W^*)^n(\mathbf{y}) = \prod_{t=1}^{n} (P \cdot W^*)(y_t).$$

Clearly, if W^* is $\{0,1\}$-valued, then a good code can be chosen accordingly, *i.e.*, $\tilde{W}^{(n)} = (W^*)^n$ for all n. On the other hand, if W^* is not $\{0,1\}$-valued, then the sequence of pmf's $P^n \cdot \tilde{W}^{(n)}$ and $(P \cdot W^*)^n$ on $\mathcal{Y}^n$ asymptotically become mutually singular. To see this, note that

$$\left(P^n \cdot \tilde{W}^{(n)}\right)(\mathcal{C}_n) = 1.$$

Also,

$$(P \cdot W^*)^n(\mathcal{C}_n) = (P \cdot W^*)^n(\mathcal{C}_n^T) + (P \cdot W^*)^n(\mathcal{C}_n \setminus \mathcal{C}_n^T).$$

The first term, by (2.6), is bounded above by a quantity that is of the order of

$$|\mathcal{C}_n| \exp[-nH(P \cdot W^*)] \approx \exp[-nH(W^*|P)],$$

which decays to zero iff $H(W^*|P) > 0$. The second term is the $(P \cdot W^*)^n$-probability of codewords which are not $P \cdot W^*$-typical and, hence, can be made arbitrarily small by choosing n sufficiently large. Consequently,

$$\lim_{n\to\infty} \|P^n \cdot \tilde{W}^{(n)} - (P \cdot W^*)^n\| = 2,$$

and

$$\lim_{n\to\infty} D(P^n \cdot \tilde{W}^{(n)} \| (P \cdot W^*)^n) = \infty. \tag{2.10}$$

In the next section, we investigate the behavior of the normalized version of the divergence in (2.10).

At this juncture it is interesting to recall the following result of Han-Verdú ([25], Theorem 5) for the channel coding problem: The normalized divergence between the output distribution induced by a good channel code (*i.e.*, a code with rate close to channel capacity and arbitrarily small probability of error) and the optimal output distribution (induced by the capacity achieving input distribution) vanishes asymptotically. Not surprisingly, a similar result does not hold for a good source code, as will be seen in Proposition 2 below.

2.4 A Minimal Good Code

The assertion of Proposition 1 can be rephrased as follows: Given $\xi > 0$, $\gamma > 0$, and for all n sufficiently large (depending on ξ, γ), the set $\mathcal{C}_n^T$, as defined in (2.6), satisfies

$$\left| \mathcal{C}_n^T \right| \geq \exp\left[n \left(R(P, \Delta) - \xi \right) \right],$$

i.e., $\mathcal{C}_n^T$ is an exponentially large subset of $\mathcal{C}_n$, of effectively the same rate, whose codewords are $P \cdot W^*$-typical. This does not, however, exclude the possibility of $\mathcal{C}_n \setminus \mathcal{C}_n^T$ being exponentially large with the same rate. But we know, from the Covering Lemma (*c.f.*,*e.g.*, [13], Lemma 4.1, p. 150) that there exists a good code, all of whose codewords are $P \cdot W^*$-typical. This motivates us to define a *minimal good code* as a good code with the following additional property: There exists a fixed $\omega > 0$ such that

$$\left|\mathcal{C}_n \setminus \mathcal{C}_n^T(\gamma)\right| \leq \exp\left[n\left(R(P,\Delta) - \omega\right)\right], \tag{2.11}$$

for all n sufficiently large (depending on ω, γ).

Claim 2.1 *Given any $\kappa > 0$, a minimal good code satisfies*

$$P^n\left(\bigcup_{i\,:\,\mathbf{y}_i \in \mathcal{C}_n \setminus \mathcal{C}_n^T} \mathcal{A}_i\right) \leq \kappa, \tag{2.12}$$

for all n sufficiently large (depending on κ, ω).

Proof: Since the pmf $P^n \cdot \tilde{W}^{(n)}$ on $\mathcal{Y}^n$ assigns full mass to $\mathcal{C}_n$ and has entropy arbitrarily close to the rate of $\mathcal{C}_n$ (see (2.16)), it follows from the continuity of the entropy function that $\mathcal{C}_n$ contains a subset, $\mathcal{C}_n^U$, of large $P^n \cdot \tilde{W}^{(n)}$-probability with roughly equiprobable codewords. Precisely, any $\upsilon > 0$, $\nu > 0$, there exists a subset $\mathcal{C}_n^U \subseteq \mathcal{C}_n$ given by

$$\mathcal{C}_n^U =$$

$$\left\{\mathbf{y} \in \mathcal{C}_n : \exp\left[-n\left(R(P,\Delta) + \nu\right)\right] \leq P^n \cdot \tilde{W}^{(n)}(\mathbf{y}) \leq \exp\left[-n\left(R(P,\Delta) - \nu\right)\right]\right\},$$

with

$$\left(P^n \cdot \tilde{W}^{(n)}\right)(\mathcal{C}_n^U) > 1 - \upsilon,$$

for all n sufficiently large (depending on υ, ν). Consequently,

$$\sum_{i\,:\,\mathbf{y}_i \in \mathcal{C}_n \setminus \mathcal{C}_n^T} P^n(\mathcal{A}_i)$$

$$
\begin{aligned}
&= \sum_{\mathbf{y}_i \in \mathcal{C}_n \setminus \mathcal{C}_n^T} \left(P^n \cdot \tilde{W}^{(n)}\right)(\mathbf{y}_i) \\
&= \sum_{\mathbf{y}_i \in (\mathcal{C}_n \setminus \mathcal{C}_n^T) \bigcap \mathcal{C}_n^U} \left(P^n \cdot \tilde{W}^{(n)}\right)(\mathbf{y}_i) + \sum_{\mathbf{y}_i \in (\mathcal{C}_n \setminus \mathcal{C}_n^T) \bigcap (\mathcal{C}_n \setminus \mathcal{C}_n^U)} \left(P^n \cdot \tilde{W}^{(n)}\right)(\mathbf{y}_i) \\
&\le \mid \mathcal{C}_n \setminus \mathcal{C}_n^T \mid \times \exp\left[-n\left(R(P,\Delta) - \nu\right)\right] + \left(P^n \cdot \tilde{W}^{(n)}\right)(\mathcal{C}_n \setminus \mathcal{C}_n^U) \\
&\le \exp\left[n\left(R(P,\Delta) - \omega\right)\right] \exp\left[-n\left(R(P,\Delta) - \nu\right)\right] + \upsilon \\
&= \exp\left[-n\left(\omega - \nu\right)\right] + \upsilon, \qquad (2.13)
\end{aligned}
$$

which can be made smaller than any $\kappa > 0$ for all n sufficiently large (depending on κ, ω), since $\omega > 0$, by choosing, say, $\nu = \frac{\omega}{2}$ and $\upsilon = \frac{\kappa}{2}$. ■

In the following proposition, we show for a minimal good code that the normalized (Kullback-Leibler) divergence has a finite limit.

Proposition 2.2 *For any minimal good code, it holds that*

$$\lim_{n\to\infty} \inf \frac{1}{n} D(P^n \cdot \tilde{W}^{(n)} \| (P \cdot W^*)^n) \ge H(W^*|P). \qquad (2.14)$$

Furthermore, if W^* is such that for every $y \in \mathcal{Y}$, there is an $x \in \mathcal{X}$ with $W^*(y|x) > 0$, then

$$\lim_{n\to\infty} \frac{1}{n} D(P^n \cdot \tilde{W}^{(n)} \| (P \cdot W^*)^n) = H(W^*|P).$$

Proof: First observe that the normalized divergence may be written as

$$
\begin{aligned}
&\frac{1}{n} D(P^n \cdot \tilde{W}^{(n)} \| (P \cdot W^*)^n) \\
&\quad = -\frac{1}{n} H(P^n \cdot \tilde{W}^{(n)}) \\
&\qquad + \frac{1}{n} \sum_{i:\mathbf{y}_i \in \mathcal{C}_n^T} P^n(\mathcal{A}_i) \log \frac{1}{(P \cdot W^*)^n(\mathbf{y}_i)} \\
&\qquad + \frac{1}{n} \sum_{i:\mathbf{y}_i \in \mathcal{C}_n \setminus \mathcal{C}_n^T} P^n(\mathcal{A}_i) \log \frac{1}{(P \cdot W^*)^n(\mathbf{y}_i)}. \qquad (2.15)
\end{aligned}
$$

Since $P^n \cdot \tilde{W}^{(n)}$ assigns full mass to a set of M_n codewords and since, by (2.3), $\frac{1}{n} \log M_n \le R(P,\Delta) + \alpha$, it follows that

$$H(P^n \cdot \tilde{W}^{(n)}) \le n(R(P,\Delta) + \alpha).$$

Also, the third term on the right hand side of (2.15) is nonnegative, so that

$$\begin{aligned}
&\frac{1}{n} D(P^n \cdot \tilde{W}^{(n)} \| (P \cdot W^*)^n) \\
&\geq \; -R(P,\Delta) - \alpha + \frac{1}{n} \sum_{i:\mathbf{y}_i \in \mathcal{C}_n^T} P^n(\mathcal{A}_i) \log \frac{1}{\exp\left[-n\left(H(P \cdot W^*) - \lambda(\gamma)\right)\right]} \\
&= \; -R(P,\Delta) - \alpha + \left(H(P \cdot W^*) - \lambda(\gamma)\right) \sum_{i:\mathbf{y}_i \in \mathcal{C}_n^T} P^n(\mathcal{A}_i) \\
&\geq \; -R(P,\Delta) - \alpha + \left(H(P \cdot W^*) - \lambda(\gamma)\right)(1-\kappa), \quad \text{by (2.12)} \\
&\geq \; H(W^*|P) - \mu,
\end{aligned}$$

for every $\mu > 0$ and for all n sufficiently large (depending on μ, ω). This establishes the inequality in (2.14).

To show the second part of the Proposition, note first that

$$\begin{aligned}
d(P^n, \tilde{W}^{(n)}) &= \sum_{i=1}^{M_n} \sum_{\mathbf{x} \in \mathcal{B}_i} P^n(\mathbf{x}) d(\mathbf{x}, \mathbf{y}_i) \; + \sum_{\mathbf{x} \notin \left(\bigcup_{i=1}^{M_n} \mathcal{B}_i\right)} P^n(\mathbf{x}) d(\mathbf{x}, \phi_n(f_n(\mathbf{x}))) \\
&\leq \sum_{i=1}^{M_n} \sum_{\mathbf{x} \in \mathcal{B}_i} P^n(\mathbf{x}) \Delta \; + \sum_{\mathbf{x} \notin \left(\bigcup_{i=1}^{M_n} \mathcal{B}_i\right)} P^n(\mathbf{x}) d_{max} \\
&= P^n\left(\bigcup_{i=1}^{M_n} \mathcal{B}_i\right) \Delta \; + \left[1 - \left(\bigcup_{i=1}^{M_n} \mathcal{B}_i\right)\right] d_{max} \\
&\leq \Delta + \epsilon d_{max}.
\end{aligned}$$

Hence,

$$\begin{aligned}
H(P^n \cdot \tilde{W}^{(n)}) &= I(P^n, \tilde{W}^{(n)}) \\
&\geq \min_{W^{(n)}: d(P^n, W^{(n)}) \leq \Delta + \epsilon d_{max}} I(P^n, W^{(n)}) \\
&= nR(P, \Delta + \epsilon d_{max}) \\
&\geq n\left(R(P,\Delta) - \beta\right), \qquad (2.16)
\end{aligned}$$

for all n sufficiently large (depending on β), where the last inequality uses the continuity of $R(P, \Delta)$ in Δ. Therefore, the first term in (2.15) is bounded above by $-R(P,\Delta)+\beta$. The second term can easily be seen to be bounded above by $H(P \cdot W^*) + \lambda(\gamma)$ similarly as above. Finally, since

$$(P \cdot W^*)^n(\mathbf{y}) \geq \left(\min_{y \in \mathcal{Y}} (P \cdot W^*)(y)\right)^n > 0,$$

the third term on the right hand side of (2.15) is bounded above by

$$\kappa \log \left(\frac{1}{\min_{y \in \mathcal{Y}} (P \cdot W^*)(y)} \right),$$

and hence for any $\mu > 0$,

$$\frac{1}{n} D(P^n \cdot \tilde{W}^{(n)} \| (P \cdot W^*)^n) \leq H(W^*|P) + \mu,$$

for all n sufficiently large (depending on μ, ω). Hence,

$$\lim_{n \to \infty} \sup \frac{1}{n} D(P^n \cdot \tilde{W}^{(n)} \| (P \cdot W^*)^n) \leq H(W^*|P),$$

completing the proof of the Proposition. ■

We remark that for a good code which is not minimal in the sense of (2.11), it is not difficult to see that the limiting value of the normalized divergence in (2.15) can be bounded away from $H(W^*|P)$.

2.5 Probability of Nontypical Codewords

Proposition 2.1 above establishes that the number of *typical* codewords in a good code is of the same order as the total number of codewords, namely $\exp\{nR(P,\Delta)\}$. We shall now provide a stronger version of Proposition 2.1, by establishing that the probability of the set of source sequences which are mapped to nontypical codewords is exponentially decaying in the block length n.

Without any essential loss of generality, we can assume that a good code consists of a sequence of *nearest-neighbour codes* (f_n, ϕ_n) with codeword set $\mathcal{C}_n \subset \mathcal{Y}^n$, *i.e.*,

$$g_n(\mathbf{x}) \stackrel{\triangle}{=} \phi_n(f_n(\mathbf{x})) = \arg\min_{\mathbf{y}\in\mathcal{C}_n} d(\mathbf{x}, \mathbf{y}),$$

where $d(\cdot, \cdot)$ is a distortion measure defined on $\mathcal{X}^n \times \mathcal{Y}^n$ and induced by a single-letter distortion measure $d : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}_+$.

We shall assume that our good code satisfies a slightly stronger version of property (2.3). Thus, we now require that

$$R(P, \Delta) \leq R_n \leq R(P, \Delta) + \alpha_n, \tag{2.17}$$

$$P^n \left(d(X^n, g_n(X^n)) > \Delta\right) \leq \epsilon, \tag{2.18}$$

where α_n is a positive sequence which decays to 0, and $0 < \epsilon < 1$. Let $\mathcal{C}_n^T(\gamma)$ denote the set of typical codewords given by (2.6). Let $\mathcal{C}_n^N(\gamma)$ denote the set of *nontypical codewords*, *i.e.*,

$$\mathcal{C}_n^N(\gamma) \stackrel{\triangle}{=} \mathcal{C}_n \setminus \mathcal{C}_n^T(\gamma).$$

The codewords in $\mathcal{C}_n^N(\gamma)$ induce a partition $\{A_n(\mathbf{y}),\ \mathbf{y} \in \mathcal{C}_n\}$ of $\mathcal{X}^n$, where

$$A_n(\mathbf{y}) \stackrel{\triangle}{=} g_n^{-1}(\mathbf{y}) = \{\mathbf{x} \in \mathcal{X}^n : g_n(\mathbf{x}) = \mathbf{y}\}, \quad \mathbf{y} \in \mathcal{C}_n.$$

Lemma 2.2 *For each $0 < \epsilon < 1$ and $\gamma > 0$, there exists $\delta = \delta(\epsilon, \gamma) > 0$ such that*

$$P^n \left(\bigcup_{\mathbf{y}\in\mathcal{C}_n^N(\gamma)} A_n(\mathbf{y}) \right) \leq e^{-n\delta}, \tag{2.19}$$

for all $n \geq N(\gamma, \epsilon)$.

The proof of the lemma is based on the following decomposition of the set $\mathcal{P}_n$ of types on $\mathcal{X}^n$:

$$\mathcal{P}_n = \{Q \in \mathcal{P}_n : \|Q - P\| \le b\} \bigcup \{Q \in \mathcal{P}_n : \|Q - P\| > b\},$$

where $b > 0$ is a constant to be chosen appropriately. We can then write

$$\begin{aligned}
P^n \left(\bigcup_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} A_n(\mathbf{y}) \right) &= \sum_{Q \in \mathcal{P}_n} P^n \left\{ \bigcup_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \left(A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \right) \right\} \\
&\le \sum_{Q : \|Q-P\| \le b} \left[\sum_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \left| A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \right| \right] e^{-n[H(Q)+D(Q\|P)]} \\
&\quad + \sum_{Q : \|Q-P\| > b} |T_Q^n| e^{-n[H(Q)+D(Q\|P)]}. \qquad (2.20)
\end{aligned}$$

The second sum can be easily shown to decay as e^{-nb^2}; so it remains to obtain a suitable bound for the first sum.

Before proceeding with the proof of Lemma 2.2, we need a few additional definitions. To this end, define

$$\begin{aligned}
D_n(\mathbf{y}, \Delta) &\triangleq \{\mathbf{x} \in \mathcal{X}^n : d(\mathbf{x}, \mathbf{y}) \le \Delta\}, \\
F_n(\Delta) &\triangleq \{\mathbf{x} \in \mathcal{X}^n : d(\mathbf{x}, g_n(\mathbf{x})) > \Delta\}, \\
\bar{F}_n(\Delta) &\triangleq \mathcal{X}^n \setminus F_n(\Delta) = \bigcup_{\mathbf{y} \in \mathcal{C}_n} D_n(\mathbf{y}, \Delta),
\end{aligned}$$

and

$$h_n(\Delta) \triangleq -\frac{1}{n} \log P^n \left(F_n(\Delta) \right).$$

In other words, $D_n(\mathbf{y}, \Delta)$ is a "ball" with center $\mathbf{y}$ and radius Δ, $F_n(\Delta)$ is the set of "unacceptably encoded" source sequences, and $h_n(\Delta)$ is the associated "error exponent," where the "error" is defined as the event that a sequence is mapped into a codeword resulting in a distortion which exceeds Δ.

We next state an important result provided by Zhang *et al* [63], which will be used in our proof of Lemma 2.2.

Lemma 2.3 ([63, Lemma 3]) *For every sequence* $\mathbf{y} \in \mathcal{C}_n$ *and* $\Delta > 0$

$$\left| D_n(\mathbf{y}, \Delta) \bigcap \mathcal{T}_Q^n \right| = \exp\left\{ n \left[H(Q) - \tilde{R}(Q, T_{\mathbf{y}}, \Delta) - o(1) \right] \right\} \tag{2.21}$$

We are now in a position to proceed with the proof of Lemma 2.2.

Proof: We shall first prove the following

Claim 2.2 *There exists* $b > 0$ *such that for every type* $Q \in \mathcal{P}_n$ *with* $\|Q - P\| \leq b$, *we have*

$$\sum_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \left| A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \right| \leq \exp\left\{ n \left[H(Q) - \eta \right] \right\}, \tag{2.22}$$

for $n > N(b, \gamma, \epsilon)$, *where* $\eta > 0$ *depends on* b, γ, ϵ.

Let $\Delta_0 = \Delta + \ell \cdot d_{max}$, where $\ell > 0$ is to be chosen appropriately later. We break the sum on the left side of (2.22) as follows :

$$\begin{aligned}
\sum_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \left| A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \right| &= \sum_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \left| A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \bigcap F_n(\Delta_0) \right| \\
&\quad + \sum_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \left| A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \bigcap \bar{F}_n(\Delta_0) \right| \\
&\stackrel{(a)}{\leq} \left| \mathcal{T}_Q \bigcap F_n(\Delta_0) \right| \\
&\quad + \sum_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \exp\left\{ n \left[H(Q) - \tilde{R}(Q, T_{\mathbf{y}}, \Delta_0) - o(1) \right] \right\} \\
&\stackrel{(b)}{\leq} e^{n[H(Q) + D(Q\|P) - h_n(\Delta_0)]} \\
&\quad + e^{n\left[H(Q) + R(P,\Delta) - \min_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \tilde{R}(Q, T_{\mathbf{y}}, \Delta_0) + \alpha_n - o(1) \right]}.
\end{aligned}$$

Here, (a) results since

$$\begin{aligned}
\sum_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \left| A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \bigcap F_n(\Delta_0) \right| &= \left| \left(\bigcup_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} A_n(\mathbf{y}) \right) \bigcap \mathcal{T}_Q^n \bigcap F_n(\Delta_0) \right| \\
&\leq \left| \left(\bigcup_{\mathbf{y} \in \mathcal{C}_n} A_n(\mathbf{y}) \right) \bigcap \mathcal{T}_Q^n \bigcap F_n(\Delta_0) \right| \\
&= \left| \mathcal{T}_Q^n \bigcap F_n(\Delta_0) \right|,
\end{aligned}$$

and $\mathbf{x} \in A_n(\mathbf{y}) \bigcap \bar{F}_n(\Delta_0)$ implies that

$$g_n(\mathbf{x}) = \mathbf{y} \text{ and } d(\mathbf{x}, g_n(\mathbf{x})) = d(\mathbf{x}, \mathbf{y}) \leq \Delta_0.$$

Thus,

$$A_n(\mathbf{y}) \bigcap \bar{F}_n(\Delta_0) \subset D(y, \Delta_0),$$

which implies that

$$\begin{aligned} \left| A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \bigcap \bar{F}_n(\Delta_0) \right| &\leq \left| D_n(\mathbf{y}, \Delta_0) \bigcap \mathcal{T}_Q^n \right| \\ &= \exp\left\{ n \left[H(Q) - \tilde{R}(Q, T_{\mathbf{y}}, \Delta_0) - o(1) \right] \right\}, \end{aligned}$$

by (2.21).

Further, (b) holds since

$$\begin{aligned} \left| \mathcal{T}_Q^n \bigcap F_n(\Delta_0) \right| \cdot e^{-n[H(Q)+D(Q\|P)]} &= P^n \left(\mathcal{T}_Q^n \bigcap F_n(\Delta_0) \right) \\ &\leq P^n \left(F_n(\Delta_0) \right) \\ &= \exp\left\{ -n h_n(\Delta_0) \right\}, \end{aligned}$$

which implies that

$$\left| \mathcal{T}_Q^n \bigcap F_n(\Delta_0) \right| \leq \exp\left\{ n \left[H(Q) + D(Q\|P) - h_n(\Delta_0) \right] \right\}$$

and

$$\begin{aligned} \left| \mathcal{C}_n^N(\gamma) \right| &\leq \left| \mathcal{C}_n \right| \\ &= \exp\left(n \cdot R_n \right) \\ &\leq \exp\left\{ n \left[R(P, \Delta) + \alpha_n \right] \right\}, \end{aligned}$$

by (2.17).

It remains to find suitable bounds on $D(Q\|P)$, $h_n(\Delta_0)$ and the difference $R(P,\Delta) - \min_{\mathbf{y}\in\mathcal{C}_n^N(\gamma)} \tilde{R}(Q, T_{\mathbf{y}}, \Delta_0)$. To this end, for any set $A \subset \mathcal{X}^n$, define its *ℓ-neighbourhood* to be

$$[A]_\ell \triangleq \left\{\mathbf{x} \in \mathcal{X}^n : \min_{\hat{x}\in A} d_H(\mathbf{x}, \hat{\mathbf{x}}) \le \ell\right\}, \tag{2.23}$$

where d_H is the normalized Hamming distance defined by

$$d_H(\mathbf{x}, \hat{\mathbf{x}}) \triangleq \frac{1}{n}\sum_{t=1}^{n} \mathbf{1}\left(x_t \neq \hat{x}_t\right).$$

We shall show for this choice of Δ_0, that

$$[F_n(\Delta_0)]_\ell \subset F_n(\Delta). \tag{2.24}$$

Let $\mathbf{x} \in [F_n(\Delta_0)]_\ell$. Then, from (2.23), there exists $\hat{\mathbf{x}} \in F_n(\Delta_0)$ such that $d_H(\mathbf{x}, \hat{\mathbf{x}}) \le \ell$; hence

$$\begin{aligned}
\Delta_0 &\overset{(\alpha)}{<} d\left(\hat{\mathbf{x}}, g_n(\hat{\mathbf{x}})\right) \\
&\overset{(\beta)}{\le} d\left(\hat{\mathbf{x}}, g_n(\mathbf{x})\right) \\
&= d\left(\mathbf{x}, g_n(\mathbf{x})\right) + \sum_{t:x_t\neq\hat{x}_t} \left[d\left(\hat{x}_t, g_n(\mathbf{x})_t\right) - d\left(x_t, g_n(\mathbf{x})_t\right)\right] \\
&\le d\left(\mathbf{x}, g_n(\mathbf{x})\right) + \left|\{t : x_t \neq \hat{x}_t\}\right| \cdot d_{max} \\
&= d\left(\mathbf{x}, g_n(\mathbf{x})\right) + d_H\left(\mathbf{x}, \hat{\mathbf{x}}\right) \cdot d_{max} \\
&\le d\left(\mathbf{x}, g_n(\mathbf{x})\right) + \ell \cdot d_{max}.
\end{aligned}$$

Here, (α) follows from the fact that $\hat{\mathbf{x}} \in F_n(\Delta_0)$ and from the definition of $F_n(\Delta_0)$ while (β) is due to the fact that (f_n, ϕ_n) is a sequence of nearest-neighbour codes, so that

$$d\left(\hat{\mathbf{x}}, g_n(\hat{\mathbf{x}})\right) \le d\left(\hat{\mathbf{x}}, \mathbf{y}\right),\ \mathbf{y} \in \mathcal{C}_n.$$

Consequently,

$$d\left(\mathbf{x}, g_n(\mathbf{x})\right) \ge \Delta_0 - \ell \cdot d_{max} = \Delta,$$

so that, $\mathbf{x} \in F_n(\Delta)$.

By the Blowing-Up Lemma (*c.f.*, final remark in [41, p. 446]), we have

$$h_n(\Delta_0) = -\frac{1}{n} \log P^n \left(F_n(\Delta_0)\right) \geq \ell^2 \left[1 - P^n \left([F_n(\Delta_0)]_\ell\right)\right]^2, \ \ \forall \ell > 0.$$

But from (2.24) and (2.18),

$$P^n \left([F_n(\Delta_0)]_\ell\right) \leq P^n \left(F_n(\Delta)\right) \leq \epsilon,$$

so that

$$h_n(\Delta_0) \geq \ell^2 (1-\epsilon)^2. \tag{2.25}$$

Also, since Q satisfies $\|Q - P\| \leq b$, applying the standard χ^2-bound (*c.f.*, *e.g.*, [11, Problem 12.2, p. 333]) on the Kullback-Leibler distance, we get

$$D(Q\|P) \leq \sum_{x \in \mathcal{X}} \frac{|Q(x) - P(x)|^2}{Q(x)} \leq |\mathcal{X}| \frac{b^2}{p_{\min} - b}, \tag{2.26}$$

where

$$p_{\min} \stackrel{\triangle}{=} \min_{x \in \mathcal{X}} P(x) > 0.$$

Finally, it remains to find a lower bound on the difference $\tilde{R}(Q, T_{\mathbf{y}}, \Delta_0) - R(P, \Delta)$ for $\mathbf{y} \in \mathcal{C}_n^N(\gamma)$. To this end, note that

$$\begin{aligned} \min_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \tilde{R}(Q, T_{\mathbf{y}}, \Delta_0) - R(P, \Delta) &= \min_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \tilde{R}(Q, T_{\mathbf{y}}, \Delta_0) - \tilde{R}(P, P \cdot W^*, \Delta) \\ &= C_1 + C_2 + C_3, \end{aligned} \tag{2.27}$$

where

$$\begin{aligned} C_1 &\stackrel{\triangle}{=} \min_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \tilde{R}(Q, T_{\mathbf{y}}, \Delta_0) - \min_{\mathbf{y} \in \mathcal{C}_n^N(\gamma)} \tilde{R}(Q, T_{\mathbf{y}}, \Delta) \\ &\geq \tilde{R}(Q, S_{Q,\Delta_0}, \Delta_0) - \tilde{R}(Q, S_{Q,\Delta_0}, \Delta) \\ &\geq -c_1(\ell), \end{aligned}$$

$$
\begin{aligned}
C_2 &\stackrel{\triangle}{=} \min_{\mathbf{y}\in\mathcal{C}_n^N(\gamma)} \tilde{R}(Q,T_{\mathbf{y}},\Delta) - \min_{\mathbf{y}\in\mathcal{C}_n^N(\gamma)} \tilde{R}(P,T_{\mathbf{y}},\Delta)\\
&\geq \tilde{R}(Q,S_{Q,\Delta},\Delta) - \tilde{R}(P,S_{Q,\Delta},\Delta)\\
&\geq -c_2(b),
\end{aligned}
$$

by the lower semi-continuity of $\tilde{R}(\cdot,\cdot,\cdot)$ and

$$
\begin{aligned}
C_3 &\stackrel{\triangle}{=} \min_{\mathbf{y}\in\mathcal{C}_n^N(\gamma)} \tilde{R}(P,T_{\mathbf{y}},\Delta) - \tilde{R}(P,P\cdot W^*,\Delta)\\
&\geq \tilde{R}(P,S_{P,\Delta},\Delta) - \tilde{R}(P,P\cdot W^*,\Delta)\\
&\geq c_3(\gamma),
\end{aligned}
$$

since $\tilde{R}(P,S,\Delta)$ takes its minimum value at $S=P\cdot W^*$ and is lower-semicontinous with respect to S. In the previous inequalities, $c_1(\ell), c_2(b), c_1(\gamma)$ are positive lower semi-continuous functions of ℓ, b and γ, respectively, and

$$
S_{Q,\Delta} \stackrel{\triangle}{=} \arg\min_{T_{\mathbf{y}}:\mathbf{y}\in\mathcal{C}_n^N(\gamma)} \tilde{R}(Q,T_{\mathbf{y}},\Delta).
$$

Combining (2.23) with (2.25),(2.26) and (2.27), and observing that $\alpha_n = o(1)$, we get

$$
\begin{aligned}
&\sum_{\mathbf{y}\in\mathcal{C}_n^N(\gamma)} \left| A_n(\mathbf{y}) \bigcap \mathcal{T}_Q^n \right|\\
&\leq o(1) e^{nH(Q)} e^{-n\min\left[\ell^2(1-\epsilon)^2 - |\mathcal{X}|\frac{b^2}{p_{\min}-b}, c_3(\gamma)-c_1(\ell)-c_2(b)-o(1)\right]}. \qquad (2.28)
\end{aligned}
$$

For a given γ, if we choose ℓ to satisfy

$$
c_1(\ell) < c_3(\gamma),
$$

and then pick b small enough so that

$$
c_2(b) < c_3(\gamma) - c_1(\ell),
$$

and

$$|\mathcal{X}|\frac{b^2}{p_{\min}-b} < \ell^2(1-\epsilon)^2,$$

then the minimum in the second exponential term on the right side of (2.28)is strictly positive, *i.e.*, there is an $\eta > 0$ (depending on γ and ϵ) such that (2.22) is satified. In particular, choosing

$$\ell = c_1^{-1}\left(\frac{c_3(\gamma)}{4}\right)$$

and

$$b = \min\left\{\frac{p_{\min}}{2}, c_2^{-1}\left(\frac{c_3(\gamma)}{4}\right), c_1^{-1}\left(\frac{c_3(\gamma)}{4}\right)\frac{(1-\epsilon)}{2}\sqrt{\frac{p_{\min}}{|\mathcal{X}|}}\right\}, \tag{2.29}$$

enables us to define $\eta > 0$ as :

$$\eta = \min\left\{\frac{c_3(\gamma)}{4}, \frac{c_1^{-1}\left(\frac{c_3(\gamma)}{4}\right)^2(1-\epsilon)^2}{2}\right\}. \tag{2.30}$$

■

Returning to the proof of Lemma 2.2, note that (2.20) can be written as

$$\begin{aligned} P^n\left(\bigcup_{\mathbf{y}\in\mathcal{C}_n^N(\gamma)} A_n(\mathbf{y})\right) &\leq \sum_{Q:\|Q-P\|\leq b} e^{-n[\eta+D(Q\|P)]} + \sum_{Q:\|Q-P\|>b} e^{-nD(Q\|P)} \\ &\leq (n+1)^{|\mathcal{X}|}e^{-n\eta} + (n+1)^{|\mathcal{X}|}e^{-n\frac{b^2}{2\log 2}}, \end{aligned}$$

by the standard bounds on the number of types (*c.f.*, *e.g.*, [11, Section 16.4, p. 490]) and the size of the type sets and by invoking Pinsker's inequality (*c.f.*, *e.g.*, [13, Problem 1.3.17, p. 58])

$$D(Q\|P) \geq \frac{\|Q-P\|^2}{2\ln 2}.$$

The existence of a positive δ such that (2.19) holds is then straightforward. Observe from (2.29)that η is a function of γ and ϵ, for b given by (2.29). Thus, we can take

$$\delta = \min\left\{\frac{\eta(\gamma,\epsilon)}{2}, \frac{b(\gamma,\epsilon,p_{\min})^2}{4\ln 2}\right\},$$

where $\eta(\gamma,\epsilon)$ and $b(\gamma,\epsilon,p_{\min})$ are given by (2.30) and (2.29), respectively.

2.6 Probability of Joint Nontypicality

In this section, we show that the set of source sequences in $\mathcal{X}^n$ which, together with their codewords, are not jointly typical with respect to PW^*, has an exponentially small probability.

Define $\mathcal{N}_n(\gamma)$ to be the set of sequences, which together with their corresponding codewords, are not jointly typical with respect to the joint pmf PW^* on $\mathcal{X}\times\mathcal{Y}$, *i.e.*,

$$\mathcal{N}_n(\gamma) = \left\{\mathbf{x}\in\mathcal{X}^n : \left\|T_{\mathbf{x}g_n(\mathbf{x})} - PW^*\right\| > \gamma\right\},$$

where $T_{\mathbf{xy}}$ denotes the joint type of the sequences $\mathbf{x}\in\mathcal{X}$ and $\mathbf{y}\in\mathcal{Y}$. Also, let $\mathcal{N}_n^M(\gamma)$ denote the subset of $\mathcal{X}^n$ for which the corresponding codewords are nontypical with respect to $P\cdot W^*$, the $\mathcal{Y}$-marginal of PW^*, namely

$$\mathcal{N}_n^M(\gamma) \triangleq \bigcup_{\mathbf{y}\in\mathcal{C}_n^N(\gamma)} A_n(\mathbf{y}) = \left\{\mathbf{x}\in\mathcal{X}^n : \left\|T_{g_n(\mathbf{x})} - P\cdot W^*\right\| > \gamma\right\}.$$

It is easy to see that

$$\mathcal{N}_n^M(\gamma) \subset \mathcal{N}_n(\gamma). \tag{2.31}$$

Indeed, for every $\mathbf{x}\in\mathcal{N}_n^M(\gamma)$,

$$\begin{aligned}
\left\|T_{\mathbf{x}g_n(\mathbf{x})} - (PW^*)\right\| &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\left|T_{\mathbf{x}g_n(\mathbf{x})}(x,y) - (PW^*)(x,y)\right| \\
&\ge \sum_{y\in\mathcal{Y}}\left|\sum_{x\in\mathcal{X}}\left[T_{\mathbf{x}g_n(\mathbf{x})}(x,y) - (PW^*)(x,y)\right]\right|
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{y\in\mathcal{Y}} \left| T_{g_n(\mathbf{x})}(y) - (P\cdot W^*)(y) \right| \\
&= \left\| T_{g_n(\mathbf{x})} - P\cdot W^* \right\| \\
&> \gamma.
\end{aligned}
$$

Thus, $\mathbf{x} \in \mathcal{N}_n(\gamma)$, and (2.31) follows.

We wish to obtain an upper bound for $P^n\left(\mathcal{N}_n(\gamma)\right)$. In the previous section, we have shown that

$$
P^n\left(\mathcal{N}_n^M(\gamma)\right) \le \exp\left\{-n\delta(\gamma)\right\},
$$

and so it remains to find a bound for $P^n\left(\mathcal{N}_n(\gamma)\setminus\mathcal{N}_n^M(\gamma)\right)$, namely, the probability of that portion of $\mathcal{N}_n(\gamma)$ for which the $\mathcal{Y}$-marginal is close to $P\cdot W^*$.

At this point, note that for a ball $D(\mathbf{y},\Delta)$ with center $\mathbf{y}$ and radius Δ, it holds that

$$
\left| D(\mathbf{y},\Delta)\bigcap \mathcal{T}_Q^n \right| = \sum_{V:d(T_{\mathbf{y}},V)\le\Delta, T_{\mathbf{y}}\cdot V=Q} \left|\mathcal{T}_V^n(\mathbf{y})\right|.
$$

Thus,

$$
(n+1)^{-|\mathcal{X}||\mathcal{Y}|} e^{n[H(Q)-\min I(T_{\mathbf{y}},V)]} \le \left| D(\mathbf{y},\Delta)\bigcap \mathcal{T}_Q^n \right| \le (n+1)^{|\mathcal{X}||\mathcal{Y}|} e^{n[H(Q)-\min I(T_{\mathbf{y}},V)]},
$$

where the minimum is taken over all V-shells of $\mathbf{y}$ that satisfy

$$
d(T_{\mathbf{y}},V) \le \Delta \text{ and } T_{\mathbf{y}}\cdot V = Q.
$$

Let V^* be the minimizing inverse channel. Then, only those V-shells with V close to V^* will contribute significantly to the cardinality of the ball. Thus,

$$
\frac{\left| D(\mathbf{y},\Delta)\cap \mathcal{T}_Q^n \cap \mathcal{T}_{[V^*]}^n(\mathbf{y}) \right|}{\left| D(\mathbf{y},\Delta)\cap \mathcal{T}_Q^n \right|} \ge 1 - e^{-n\eta},
$$

where η will depend on the "width" of $\mathcal{T}_{[V^*]}^n(\mathbf{y})$. Now, for Q close to P, and $T_{\mathbf{y}}$ close to $P\cdot W^*$ (as is the case for the codewords of the sequences in $\mathcal{N}_n(\gamma)\setminus$

$\mathcal{N}_n^M(\gamma))$, $T_{\mathbf{y}}V^*$ will be close to PW^*, which will lead to an exponentially decaying upper bound on the P^n probability of $\mathcal{N}_n(\gamma) \setminus \mathcal{N}_n^M(\gamma)$.

Concluding, we have outlined the proof to the following

Lemma 2.4 *There exists $\delta = \delta(\epsilon, \gamma) > 0$, such that*

$$P^n\left(\mathcal{N}_n(\gamma)\right) \leq e^{-n\delta}.$$

for all n sufficiently large

Chapter 3

Error Exponents for Successive Refinement by Partitioning

3.1 Introduction

The problem of successive refinement of information by partitioning, also referred to variously as "hierarchical lossy data compression" and "sequential approximations," has received much attention over the years (Koshélev [33, 32], Equitz-Cover [16], Yamamoto [58], Rimoldi [46]). Related problems include those of "multiple descriptions" (*c.f.*, [1, 15, 44, 55, 56, 62]) and that of determining the achievable rate region for cascade communication systems [58]. Given a discrete memoryless source (DMS) with probability mass function (pmf) P, and a suitable distortion measure, suppose that we first seek to describe the source with distortion not exceeding Δ_1. The (asymptotically) minimum rate of coding is, of course, given by the rate distortion function $R(P, \Delta_1)$. Subsequently, if a better (finer) description is required, say with distortion $\Delta_2 < \Delta_1$, additional information at rate ΔR can be provided, so that the overall augmented rate is $R(P, \Delta_1) + \Delta R$. Clearly, $R(P, \Delta_1) + \Delta R \geq R(P, \Delta_2)$. It is of interest to determine the rate-distortion function $R(P, \Delta_1, \Delta_2)$ for this two-level coding process, and find conditions under which it coincides with $R(P, \Delta_2)$.

The condition under which these two rate-distortion functions coincide was determined independently by Koshélev [33] and Equitz-Cover [16], and subsequently given a geometrical interpretation by Rimoldi [46]. This condition requires that the source random variable (rv) and the two reproduction rv's satisfy a Markov property. Rimoldi [46] also provided a complete characterization of the achievable rate region for two-level coding.

In this chapter, we determine the error exponents for the two-level coding process. It is then shown that even under the Markov condition, when the two rate-distortion functions coincide, the performance of the two-level coding process – as measured by its error exponents – may be inferior to that of one-level coding.

3.2 Preliminaries

Let $\mathcal{X}$ be a finite set. Let $\{X_t\}_{t=1}^{\infty}$ be a $\mathcal{X}$-valued discrete memoryless source (DMS), *i.e.*, an independent and identically distributed (i.i.d.) process, with (common) probability mass function (pmf) P. Let $\mathcal{Y}_1$ be a finite reproduction alphabet. Let $d_1 : \mathcal{X} \times \mathcal{Y}_1 \mapsto \mathbf{R}^+$ be a nonnegative-valued mapping with $\min_{x \in \mathcal{X}, y \in \mathcal{Y}_1} d_1(x, y) = 0$. This mapping induces a distortion measure on $\mathcal{X}^n \times \mathcal{Y}_1^n$ according to

$$d_1(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{t=1}^{n} d_1(x_t, y_t), \quad \mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}_1^n.$$

An n-length block code consists of two mappings : An encoder $f_1^{(n)} : \mathcal{X}^n \mapsto \mathcal{M}_1 = \{1, \cdots, M_1\}$ and a decoder $\phi_1^{(n)} : \mathcal{M}_1 \mapsto \mathcal{Y}_1^n$. The rate of this code is

$R_1 = \frac{1}{n}\log M_1$. All logarithms and exponentials are with respect to the base 2.

For $R_1 > 0, \Delta_1 > 0$, we say that the pair (R_1, Δ_1) is achievable if for every $\epsilon > 0, \delta > 0$ and n sufficiently large, there exists an n-length block code $(f_1^{(n)}, \phi_1^{(n)})$ of rate not exceeding $R_1 + \delta$ such that

$$\Pr\left\{d_1(X^n, \phi_1^{(n)}(f_1^{(n)}(X^n))) \leq \Delta_1\right\} \geq 1 - \epsilon.$$

The corresponding rate-distortion function, $R(P, \Delta_1)$, characterizing the minimum achievable rate for a distortion Δ_1, is well-known and given by

$$R(P, \Delta_1) = \inf_{P_X = P,\ \mathbf{E}d_1(X, Y_1) \leq \Delta_1} I(X \wedge Y_1),$$

where $\mathbf{E}$ denotes expectation.

Let $\mathcal{Y}_2$ be a (refining) finite reproduction alphabet. A refined description of the source $\{X_t\}_{t=1}^{\infty}$ can be provided by means of a n-length *refining block code* $(f_2^{(n)}, \phi_2^{(n)})$, specified by an encoder $f_2^{(n)} : \mathcal{X}^n \mapsto \mathcal{M}_2 = \{1, \cdots, M_2\}$ and a decoder $\phi_2^{(n)} : \mathcal{M}_1 \times \mathcal{M}_2 \mapsto \mathcal{Y}_2^n$. The rate of the refining code $(f_2^{(n)}, \phi_2^{(n)})$ is defined as $R_2 = \frac{1}{n}\log M_1 M_2$. Let $d_2 : \mathcal{X} \times \mathcal{Y}_2 \mapsto \mathbf{R}^+$ be a nonnegative-valued mapping with $\min_{x \in \mathcal{X}, y \in \mathcal{Y}_2} d_2(x, y) = 0$, which induces a distortion measure on $\mathcal{X}^n \times \mathcal{Y}_2^n$ according to

$$d_2(\mathbf{x}, \mathbf{y}) = \frac{1}{n}\sum_{t=1}^{n} d_2(x_t, y_t), \quad \mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}_2^n.$$

Definition 3.1 (Rimoldi [46]) *For the DMS $\{X_t\}_{t=1}^{\infty}$ with pmf P and distortion measures d_1, d_2, the quadruple $(R_1, R_2, \Delta_1, \Delta_2), R_1 > 0, R_2 > 0, \Delta_1 > 0, \Delta_2 > 0$, is achievable if for every $\epsilon > 0, \delta > 0$ and n sufficiently large, there exist*

- *a n-length code $(f_1^{(n)}, \phi_1^{(n)})$ such that*

$$\frac{1}{n}\log M_1 \leq R_1 + \delta$$

and

$$\Pr\left\{d_1(X^n, \phi_1^{(n)}(f_1^{(n)}(X^n))) \le \Delta_1\right\} \ge 1 - \epsilon;$$

- *a n-length refining code* $(f_2^{(n)}, \phi_2^{(n)})$ *such that*

$$\frac{1}{n}\log M_1 M_2 \le R_2 + \delta$$

and

$$\Pr\Big\{d_1(X^n, \phi_1^{(n)}(f_1^{(n)}(X^n))) \le \Delta_1, \\ d_2(X^n, \phi_2^{(n)}(f_1^{(n)}(X^n), f_2^{(n)}(X^n))) \le \Delta_2\Big\} \ge 1 - \epsilon.$$

A characterization of the set of achievable quadruples $(R_1, R_2, \Delta_1, \Delta_2)$ has been provided independently by Koshélev [32] and Rimoldi [46] and also follows as a combination of the results of El Gamal-Cover [15] and Yamamoto [58].

Theorem 3.1 ([32, 15, 58, 46]) *Consider a DMS* $\{X_t\}_{t=1}^{\infty}$ *with pmf* P *and distortion measures* d_1, d_2*. The quadruple* $(R_1, R_2, \Delta_1, \Delta_2)$ *is achievable iff there exists a pmf* $P_{XY_1Y_2}$ *on* $\mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2$ *with marginal* P *on* $\mathcal{X}$ *satisfying the following inequalities :*

$$I(X \wedge Y_1) \le R_1, \qquad I(X \wedge Y_1Y_2) \le R_2,$$

$$\mathbf{E}d_1(X, Y_1) \le \Delta_1, \qquad \mathbf{E}d_2(X, Y_2) \le \Delta_2,$$

where (X, Y_1, Y_2) *is a* $\mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2$*-valued rv with pmf* $P_{XY_1Y_2}$*.*

Suppose that the coding in the first level was done at rate $R_1 > R(P, \Delta_1)$. If the DMS is now to be described in a second level at distortion $\Delta_2 < \Delta_1$, let $R(P, R_1, \Delta_1, \Delta_2)$ be the minimum rate R_2 of the refining code such that $(R_1, R_2, \Delta_1, \Delta_2)$ is achievable. This minimal rate $R(P, R_1, \Delta_1, \Delta_2)$ is characterized by the following Corollary to Theorem 3.1.

Corollary 3.1 *For $\Delta_1 > 0$, $\Delta_2 > 0$ and $R_1 > R(P, \Delta_1)$, we have*

$$R(P, R_1, \Delta_1, \Delta_2) = \inf_{\substack{P_X = P \\ \mathbf{E}d_1(X,Y_1) \leq \Delta_1 \\ \mathbf{E}d_2(X,Y_2) \leq \Delta_2 \\ I(X \wedge Y_1) \leq R_1}} I(X \wedge Y_1 Y_2). \tag{3.1}$$

Suppose next that the coding in the first level at distortion Δ_1 was done optimally, *i.e.*, $R_1 \simeq R(P, \Delta_1)$. If the DMS is now to be described in a second level at distortion $\Delta_2 < \Delta_1$, what is the smallest possible amount of additional information required for this purpose?

The following definition is a special case of Definition 3.1 for the situation wherein the first level is optimal.

Definition 3.2 *For the DMS $\{X_t\}_{t=1}^{\infty}$ with pmf P, the rate R_2 is (Δ_1, Δ_2) refinement achievable if for every $\epsilon > 0, \delta > 0$ and n sufficiently large, there exist*

- *a n-length code $(f_1^{(n)}, \phi_1^{(n)})$ such that*

$$\frac{1}{n} \log M_1 \leq R(P, \Delta_1) + \delta$$

and

$$\Pr\left\{ d_1(X^n, \phi_1^{(n)}(f_1^{(n)}(X^n))) \leq \Delta_1 \right\} \geq 1 - \epsilon;$$

- *a n-length refining code $(f_2^{(n)}, \phi_2^{(n)})$ such that*

$$\frac{1}{n} \log M_1 M_2 \leq R_2 + \delta$$

and

$$\Pr\Big\{ d_1(X^n, \phi_1^{(n)}(f_1^{(n)}(X^n))) \leq \Delta_1, \\ d_2(X^n, \phi_2^{(n)}(f_1^{(n)}(X^n), f_2^{(n)}(X^n))) \leq \Delta_2 \Big\} \geq 1 - \epsilon.$$

Let $R(P, \Delta_1, \Delta_2)$ denote the infimum over the set of (Δ_1, Δ_2) refinement achievable rates. It constitutes the rate-distortion function for the refining code and is given by the following Corollary to Theorem 3.1.

Corollary 3.2 *For $\Delta_1 > 0, \Delta_2 > 0$, we have*

$$R(P, \Delta_1, \Delta_2) = R(P, R(P, \Delta_1), \Delta_1, \Delta_2) = \inf_{\substack{P_X = P \\ \mathbf{E}d_1(X, Y_1) \leq \Delta_1 \\ \mathbf{E}d_2(X, Y_2) \leq \Delta_2 \\ I(X \wedge Y_1) = R(P, \Delta_1)}} I(X \wedge Y_1 Y_2). \quad (3.2)$$

Remarks

1. For $d_1 = d_2, \Delta_1 = \Delta_2$ and $\mathcal{Y}_1 = \mathcal{Y}_2$, we have that $R(P, \Delta_1, \Delta_2) = R(P, \Delta_1)$, the minimum achievable rate for one-level coding.

2. For two-level coding with $\Delta_2 < \Delta_1$, and $R_1 > R(P, \Delta_1)$, clearly

$$R(P, \Delta_2) \leq R(P, R_1, \Delta_1, \Delta_2) \leq R(P, \Delta_1, \Delta_2).$$

   Koshélev [33] has provided a sufficient condition for the inequalities above to hold as equalities. Cover-Equitz [16] have independently shown this condition to be both necessary and sufficient (see Theorem 3.2 below).

3. It follows from the observation of Equitz-Cover [16, p. 271] in the context of Gray's work [22] on conditional rate-distortion function that *if* $\{Y_{1,t}\}_{t=1}^{\infty}$ is an i.i.d. process, then

$$R(P, \Delta_1, \Delta_2) = R(P, \Delta_1) + R_{X|Y_1}(\Delta_2),$$

   where the conditional rate-distortion function $R_{X|Y_1}(\Delta_2)$ [22] is given by

$$R_{X|Y_1}(\Delta_2) = \inf_{P_{Y_2|XY_1} : \sum_{x, y_1, y_2} P_{Y_2|XY_1}(y_2|x, y_1) P_{XY_1}(x, y_1) d_2(x, y_2) \leq \Delta_2} I(X \wedge Y_2 | Y).$$

However, $R_{X|Y_1}(\Delta_2)$ is defined only when $\{Y_{1,t}\}_{t=1}^{\infty}$ is an i.i.d. process, which, of course, does not hold in general.

Theorem 3.2 (Cover-Equitz [16]) *For the DMS $\{X_t\}_{t=1}^{\infty}$ with pmf P, given distortion measures $d_1 = d_2 = d$, reproduction alphabet $\mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{Y}$, and for $R_1 > R(P, \Delta_1)$, we have $R(P, \Delta_1, \Delta_2) = R(P, R_1, \Delta_1, \Delta_2) = R(P, \Delta_2)$ iff there exists a pmf $P_{XY_1Y_2}$ on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Y}$ with $P_X = P$, and*

$$R(P, \Delta_1) = I(X \wedge Y_1), \quad Ed(X, Y_1) \leq \Delta_1,$$

$$R(P, \Delta_2) = I(X \wedge Y_2), \quad Ed(X, Y_2) \leq \Delta_2,$$

where (X, Y_1, Y_2) is a $\mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2$-valued rv with pmf $P_{XY_1Y_2}$. Furthermore, $P_{XY_1Y_2}$ must satisfy a Markov condition, namely

$$P_{XY_1Y_2}(x, y_1, y_2) = P_X(x) P_{Y_2|X}(y_2|x) P_{Y_1|Y_2}(y_1|y_2), \quad x \in \mathcal{X}, y_1, y_2 \in \mathcal{Y}, \tag{3.3}$$

i.e., *(X, Y_2, Y_1) form a Markov chain.*

3.3 The Distortion-Rate Function

In analogy with the case of one-level coding, we define the distortion-rate function for two-level coding, $\Delta(P, R_1, R_2)$, as follows :

$$\Delta(P, R_1, R_2) = \inf_{\substack{P_X = P \\ \mathbf{E}d_1(X,Y_1) \leq \Delta(P,R_1) \\ I(X \wedge Y) \leq R_1 \\ I(X \wedge Y_1 Y_2) \leq R_2}} Ed_1(X, Y_2), \quad 0 \leq R_1 \leq R_2 \leq R(P, \Delta(P, R_1), 0),$$

where (X, Y_1, Y_2) is a $\mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2$-valued rv with pmf $P_{XY_1Y_2}$.

The two-level distortion-rate function is a quantity which will prove very useful in the subsequent chapters in evaluating the performance of a block code. In this section, we briefly summarize its properties.

Lemma 3.1 *For fixed P,*

1. $\Delta(P, R_1, R_2)$ *is jointly convex in* (R_1, R_2);
2. $\Delta(P, R_1, R_2)$ *is jointly continuous in* (P, R_1, R_2);
3. $\Delta(P, R_1, R_2)$ *is strictly decreasing in* $R_2, R_1 \leq R_2 \leq R(P, \Delta(P, R_1), 0)$.

Proof

The proof is standard (cf. [13]). ■

Furthermore, the two-level distortion-rate function is the "inverse" of the rate-distortion function. To see this, define

$$R(\delta) \triangleq R(P, \Delta_1, \delta), \quad 0 \leq \delta \leq \Delta_1,$$

and

$$\Delta(r) \triangleq \Delta(P, R(P, \Delta_1), r), \quad R(P, \Delta_1) \leq r \leq R(P, \Delta_1, 0).$$

Lemma 3.2 *For* $0 \leq \delta \leq \Delta_1$, $R(P, \Delta_1) \leq r \leq R(P, \Delta_1, 0)$, *it holds that*

$$R(\Delta(r)) = r, \qquad \Delta(R(\delta)) = \delta$$

Proof

We know that $R(\delta)$ and $\Delta(r)$ are strictly decreasing. Suppose that $P^*_{XY_1Y_2}$ achieves the minimum for $R(\delta)$, *i.e.*,

$$\begin{aligned}
I(X^* \wedge Y_1^*) &\leq R(P, \Delta_1), \\
d_1(X^*, Y_1^*) &\leq \Delta_1, \\
d_2(X^*, Y_2^*) &\leq \delta, \\
I(X \wedge Y_1^* Y_2^*) &= R(\delta),
\end{aligned}$$

where (X^*, Y_1^*, Y_2^*) is distributed according to $P^*_{XY_1Y_2}$. Since $P^*_{XY_1Y_2}$ satisfies the first, second and fourth inequalities above, it holds that

$$\delta' \triangleq \Delta(R(\delta)) \leq d_2(X^*, Y_2^*) \leq \delta,$$

using the third inequality above and the definition of $\Delta(r)$.

In an analogous manner, we obtain $R(\Delta(r)) \leq r$. Thus, $R(\delta') \leq R(\delta)$ follows by substituting $r = R(\delta)$. This, in turn, implies that $\delta' \geq \delta$, since $R(\delta)$ is nonincreasing. Hence,

$$\Delta(R(\delta)) = \delta.$$

Working in a similar manner, we can prove that $R(\Delta(r)) = r$. ■

Finally, the operational significance of the distortion-rate function for two-level coding, $\Delta(P, R_1, R_2)$ is provided by the following

Theorem 3.3 *Consider a DMS $\{X_t\}_{t=1}^{\infty}$ with pmf P and distortion measures d_1, d_2. Fix $R_1 > 0$ and $R_2 \geq R_1$. Then for every $\epsilon > 0$, there exist*

- *a n-length code $(f_1^{(n)}, \phi_1^{(n)})$ such that*

  $$\frac{1}{n} \log M_1 \leq R_1 + \delta$$

  and

  $$\Pr\left\{d_1(X^n, \phi_1^{(n)}(f_1^{(n)}(X^n))) \leq \Delta(P, R_1)\right\} \geq 1 - \epsilon;$$

- *a n-length refining code $(f_2^{(n)}, \phi_2^{(n)})$ such that*

  $$\frac{1}{n} \log M_1 M_2 \leq R_2 + \delta$$

  and

  $$\Pr\Big\{d_1(X^n, \phi_1^{(n)}(f_1^{(n)}(X^n))) \leq \Delta(P, R_1),$$
  $$d_2(X^n, \phi_2^{(n)}(f_1^{(n)}(X^n), f_2^{(n)}(X^n))) \leq \Delta(P, R_1, R_2)\Big\} \geq 1 - \epsilon.$$

Proof

The proof is standard and relies on Theorem 3.1 ■

3.4 A Covering Lemma for Two-Level Coding

In this section we provide a covering lemma for two-level coding, which is a straightforward extension of the corresponding lemma for one-level coding [13, Lemma 2.4.1] and is essential for the proof of our main result in Theorem 3.4.

Lemma 3.3 *For distortion measures d_1 on $\mathcal{X} \times \mathcal{Y}_1$ and d_2 on $\mathcal{X} \times \mathcal{Y}_2$, type $P = P^{(n)}$ on $\mathcal{X}$ and numbers $\Delta_1 > \Delta_2 > 0$, $R_1 > R(P, \Delta_1)$, $\delta_1 > \delta_2 > 0$, there exist*

- *a set $\mathcal{B}_1 \subset \mathcal{Y}_1^n$, such that*

$$\frac{1}{n} \log \mid \mathcal{B}_1 \mid \leq R_1 + \delta_1 \tag{3.4}$$

and $\mathcal{B}_1$ covers $\mathcal{T}_P^n$

$$\bigcup_{\mathbf{y}_1 \in \mathcal{B}_1} \mathcal{N}_1(\mathbf{y}_1) = \mathcal{T}_P^n,$$

where

$$\mathcal{N}_1(\mathbf{y}_1) = \{\mathbf{x} \in \mathcal{T}_P^n : d_1(\mathbf{x}, \mathbf{y}_1) \leq \Delta_1\}, \quad \mathbf{y}_1 \in \mathcal{Y}_1^n;$$

- *sets $\mathcal{B}_2(\mathbf{y}_1) \subset \mathcal{Y}_2^n$, $\quad \mathbf{y}_1 \in \mathcal{B}_1$, such that*

$$\frac{1}{n} \log \left(\sum_{\mathbf{y}_1 \in \mathcal{B}_1} \mid \mathcal{B}_2(\mathbf{y}_1) \mid \right) \leq R(P, R_1, \Delta_1, \Delta_2) + \delta_2 \tag{3.5}$$

and $\mathcal{B}_2(\mathbf{y}_1)$ covers $\mathcal{N}_1(\mathbf{y}_1)$, i.e.,

$$\bigcup_{\mathbf{y}_2 \in \mathcal{B}_2(\mathbf{y}_1)} \mathcal{N}_2(\mathbf{y}_2) \supset \mathcal{N}_1(\mathbf{y}_1), \quad \mathbf{y}_1 \in \mathcal{B}_1,$$

where

$$\mathcal{N}_2(\mathbf{y}_2) = \{\mathbf{x} \in \mathcal{T}_P^n : d_2(\mathbf{x}, \mathbf{y}_2) \leq \Delta_2\}, \quad \mathbf{y}_2 \in \mathcal{Y}_2^n,$$

provided that $n \geq N(d_1, d_2, \delta_1, \delta_2)$.

Proof: Let $P = P^{(n)}$ be a fixed but arbitrary type of sequences in $\mathcal{X}^n$. Let $\mathcal{U}(\mathcal{B}_1)$ and $\mathcal{U}(\mathcal{B}_2(\mathbf{y}_1))$ respectively denote the subsets of $\mathcal{T}_P^n$ not covered by $\mathcal{B}_1$ and $\mathcal{B}_2(\mathbf{y}_1)$, *i.e.*,

$$\mathcal{U}(\mathcal{B}_1) = \mathcal{T}_P^n \setminus \bigcup_{\mathbf{y}_1 \in \mathcal{B}_1} \mathcal{N}_1(\mathbf{y}_1) = \mathcal{T}_P^n \bigcap \left(\bigcap_{\mathbf{y}_1 \in \mathcal{B}_1} \mathcal{N}_1(\mathbf{y}_1)^c \right)$$

and

$$\begin{aligned} \mathcal{U}(\mathcal{B}_2(\mathbf{y}_1)) &= \mathcal{N}_1(\mathbf{y}_1) \setminus \bigcup_{\mathbf{y}_2 \in \mathcal{B}_2(\mathbf{y}_1)} \mathcal{N}_2(\mathbf{y}_2) \\ &= \mathcal{N}_1(\mathbf{y}_1) \bigcap \left(\bigcap_{\mathbf{y}_2 \in \mathcal{B}_2(\mathbf{y}_1)} \mathcal{N}_2(\mathbf{y}_2)^c \right), \quad \mathbf{y}_1 \in \mathcal{B}_1. \end{aligned}$$

Consider a $\mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2$-valued rv (X, Y_1, Y_2) with $P_X = P$ and

$$\mathbf{E}d_1(X, Y_1) = |\, \Delta_1 - \eta_1 \,|^+, \ \mathbf{E}d_2(X, Y_2) \leq |\, \Delta_2 - \eta_2 \,|^+, \tag{3.6}$$

where $\eta_1 > 0, \eta_2 > 0$ will be specified later.

Let $\mathcal{T}_{[Y_1]}^n$ (resp. $\mathcal{T}_{[Y_2|Y_1]}^n(\mathbf{y}_1)$) denote the set of Y_1-typical sequences $\mathbf{y}_1 \in \mathcal{Y}_1^n$ (resp. $Y_2|Y_1$-typical sequences $\mathbf{y}_2 \in \mathcal{Y}_2^n$ with respect to $\mathbf{y}_1 \in \mathcal{Y}_1^n$) [13, Definitions 1.2.8, 1.2.9, pp. 33-34]. Let $\mathcal{F}_m$ denote the set of all (not necessarily distinct) collections of m elements of $\mathcal{T}_{[Y_1]}^n$. Similarly, for $i = 1, \ldots, m$, let $\mathcal{G}_{n_i}(\mathbf{y}_1)$ denote the set of all (not necessarily distinct) collections of n_i elements of $\mathcal{T}_{[Y_2|Y_1]}^n(\mathbf{y}_1)$.

Let $Z^m = (Z_1, \ldots, Z_m)$ be a rv uniformly distributed on $\mathcal{F}_m$. Next, for $i = 1, \ldots, m$, given that $Z_i = \mathbf{z}_i$, let $W_i^{n_i} = (W_{i1}, \ldots, W_{in_i})$, be a rv distributed (conditionally) uniformly on $\mathcal{G}_{n_i}(\mathbf{z}_i)$. In other words, Z_i's are i.i.d. with

$$\Pr\{Z_i = \mathbf{z}_i\} = \frac{\mathbf{1}(\mathbf{z}_i \in \mathcal{T}_{[Y_1]}^n)}{|\, \mathcal{T}_{[Y_1]}^n \,|}, \quad i = 1, \ldots, m$$

and the W_{ij}'s are conditionally i.i.d., with

$$\Pr\{W_{ij} = \mathbf{w}_{ij} | Z_i = \mathbf{z}_i\} = \frac{\mathbf{1}(\mathbf{w}_{ij} \in \mathcal{T}^n_{[Y_2|Y_1]}(\mathbf{z}_i))}{|\mathcal{T}^n_{[Y_2|Y_1]}(\mathbf{z}_i)|}, j = 1, \ldots, n_i, i = 1, \ldots, m.$$

We must show the existence of sets $\mathcal{B}_1 \subset \mathcal{Y}_1^n$ and $\mathcal{B}_2(\mathbf{y}_1) \subset \mathcal{Y}_2^n, \mathbf{y}_1 \in \mathcal{B}_1$, such that

$$\left| \mathcal{U}(\mathcal{B}_1) \bigcup \left(\bigcup_{\mathbf{y}_1 \in \mathcal{B}_1} \mathcal{U}(\mathcal{B}_2(\mathbf{y}_1)) \right) \right| = 0.$$

To this end, it suffices to show that

$$\mathbf{E} \left| \mathcal{U}(Z^m) \bigcup \left(\bigcup_{i=1}^m \mathcal{U}(W_i^{n_i}) \right) \right| < 1.$$

Now,

$$\begin{aligned} &\mathbf{E} \left| \mathcal{U}(Z^m) \bigcup \left(\bigcup_{i=1}^m \mathcal{U}(W_i^{n_i}) \right) \right| \\ &= \sum_{\mathbf{x} \in \mathcal{T}_P^n} \Pr \left\{ \mathbf{x} \in \mathcal{U}(Z^m) \bigcup \left(\bigcup_{i=1}^m \mathcal{U}(W_i^{n_i}) \right) \right\} \\ &\leq \sum_{\mathbf{x} \in \mathcal{T}_P^n} \Pr\{\mathbf{x} \in \mathcal{U}(Z^m)\} + \sum_{\mathbf{x} \in \mathcal{T}_P^n} \sum_{i=1}^m \Pr\{\mathbf{x} \in \mathcal{U}(W_i^{n_i})\}. \end{aligned} \tag{3.7}$$

We consider first the first term on the right side of the inequality above. By choosing

$$\exp\left[n\left(I(X \wedge Y_1) + \frac{2}{3}\delta_1\right)\right] \leq m \leq \exp\left[n\left(I(X \wedge Y_1) + \frac{3}{4}\delta_1\right)\right],$$

we get that

$$\sum_{\mathbf{x} \in \mathcal{T}_P^n} \Pr\{\mathbf{x} \in \mathcal{U}(Z^m)\} \leq \exp\left[n \log |\mathcal{X}| - \exp\frac{n\delta_1}{6}\right] < \frac{1}{2}, \tag{3.8}$$

for all $n \geq N_1(\delta_1, d_1, \eta_1)$ (*c.f.*, *e.g.*, [13, p. 152]). Next, to bound the second term on the right side of (3.7), for each $\mathbf{x} \in \mathcal{T}_P^n$, let

$$\mathcal{A}(\mathbf{x}) \triangleq \left\{\mathbf{z} \in \mathcal{T}^n_{[Y_1]} : \mathbf{x} \in \mathcal{N}_1(\mathbf{z})\right\} = \left\{\mathbf{z} \in \mathcal{T}^n_{[Y_1]} : d_1(\mathbf{x}, \mathbf{z}) \leq \Delta_1\right\}.$$

Assume for the time being, the following

Claim 3.1 *For any $\delta' > 0$, we have*

$$\mid \mathcal{A}(\mathbf{x}) \mid \leq \exp\left[n\left(H(Y_1) - R(P, \Delta_1) + \delta'\right)\right]$$

for all $n \geq N_2(d_1, \eta_1, \delta')$.

Then, the remainder of the proof of the Lemma is straightforward. The second term on the right side of (3.7) is

$$\begin{aligned}
&\Pr\left\{\mathbf{x} \in \mathcal{U}(W_i^{n_i})\right\} \\
&= \Pr\left\{\mathbf{x} \in \mathcal{N}_1(Z_i) \bigcap \left(\bigcap_{j=1}^{n_i} \mathcal{N}_2(W_{ij})^c\right)\right\} \\
&= \sum_{\mathbf{z} \in \mathcal{T}_{[Y_1]}^n} \Pr\left\{Z_i = \mathbf{z}\right\} \Pr\left\{\mathbf{x} \in \mathcal{N}_1(Z_i) \bigcap \left(\bigcap_{j=1}^{n_i} \mathcal{N}_2(W_{ij})^c\right) \middle| Z_i = \mathbf{z}\right\} \\
&= \frac{1}{\mid \mathcal{T}_{[Y_1]}^n \mid} \cdot \sum_{\mathbf{z} \in \mathcal{A}(\mathbf{x})} \Pr\left\{\mathbf{x} \in \bigcap_{j=1}^{n_i} \mathcal{N}_2(W_{ij})^c \middle| Z_i = \mathbf{z}\right\}.
\end{aligned}$$

Further,

$$\begin{aligned}
&\Pr\left\{\mathbf{x} \in \bigcap_{j=1}^{n_i} \mathcal{N}_2(W_{ij})^c \middle| Z_i = \mathbf{z}\right\} \\
&= \prod_{j=1}^{n_i} \Pr\left\{\mathbf{x} \in \mathcal{N}_2(W_{ij})^c \middle| Z_i = \mathbf{z}\right\} \\
&= \prod_{j=1}^{n_i} \left(1 - \Pr\left\{\mathbf{x} \in \mathcal{N}_2(W_{ij}) \middle| Z_i = \mathbf{z}\right\}\right) \\
&\leq \prod_{j=1}^{n_i} \left(1 - \Pr\left\{W_{ij} \in \mathcal{T}_{[Y_2/X]}^n(\mathbf{x}) \middle| Z_i = \mathbf{z}\right\}\right),
\end{aligned}$$

where the previous inequality follows as in [13, p. 151]. We continue the bounding according to

$$\begin{aligned}
&\Pr\left\{\mathbf{x} \in \bigcap_{j=1}^{n_i} \mathcal{N}_2(W_{ij})^c \middle| Z_i = \mathbf{z}\right\} \\
&\leq \prod_{j=1}^{n_i} \left(1 - \Pr\left\{W_{ij} \in \mathcal{T}_{[Y_2/X,Y_1]}^n(\mathbf{x}, \mathbf{z}) \middle| Z_i = \mathbf{z}\right\}\right)
\end{aligned}$$

for all $n \geq N_3(d_2, \eta_2)$ (*c.f.*, [13, p. 151]). Furthermore, the right side above is bounded above by

$$\left\{1 - \exp\left[-n\left(I(X \wedge Y_2/Y_1) + \frac{\delta_2'}{2}\right)\right]\right\}^{n_i}$$

with $\delta_2' \triangleq \delta_2 - \delta_1$, for all $n \geq N_4(d_2, \eta_2, \delta_1, \delta_2)$, which in turn does not exceed

$$\exp\left[-\exp\frac{n\delta_2'}{6}\right],$$

provided that

$$\exp\left\{n\left[I(X \wedge Y_2|Y_1) + \frac{2}{3}\delta_2'\right]\right\} \leq n_i \leq \exp\left\{n\left[I(X \wedge Y_2|Y_1) + \frac{3}{4}\delta_2'\right]\right\}.$$

Hence, we can finally bound the second term on the right side of (3.7) as

$$\begin{aligned} &\sum_{\mathbf{x}\in\mathcal{T}_P^n}\sum_{i=1}^{m}\Pr\left\{\mathbf{x}\in\mathcal{U}(W_i^{n_i})\right\} \\ &\leq\ m\exp\left\{n\log\mid\mathcal{X}\mid - \exp\frac{n\delta_2'}{6}\right\}\frac{1}{\mid\mathcal{T}_{[Y_1]}^n\mid}\mid\mathcal{A}(\mathbf{x})\mid. \end{aligned}$$

Using the facts

$$m \leq \exp\left\{n\left[I(X \wedge Y_1) + \frac{3}{4}\delta_1\right]\right\}, \quad \left|\mathcal{T}_{[Y_1]}^n\right| \geq \exp\left\{n\left[H(Y_1) - \frac{\delta_1}{4}\right]\right\}$$

and Claim 3.1, the right side above is easily seen to be bounded by

$$\exp\left\{n\left[2\log\mid\mathcal{X}\mid + 2\delta_1\right] - \exp\frac{n\delta_2'}{6}\right\} < \frac{1}{2} \tag{3.9}$$

for all $n \geq N_5(\delta_1, \delta_2')$ (because $R_1 - R(P, \Delta_1) \leq \log|\mathcal{X}|$). Finally, we obtain from (3.8) and (3.9) that

$$\mathbf{E}\left|\mathcal{U}(Z^m)\bigcup\left(\bigcup_{i=1}^{m}\mathcal{U}(W_i^{n_i})\right)\right| < 1$$

for all $n \geq N(d_1, d_2, \eta_1, \eta_2, \delta_1, \delta_2) = \max\{N_1, \ldots, N_5\}$. It then follows for all n suitably large that there exist sets $\mathcal{B}_1 \subset \mathcal{Y}_1^n$, $\mathcal{B}_2(\mathbf{y}_i) \subset \mathcal{Y}_2^n, \mathbf{y}_i \in \mathcal{B}_1$, such that

$$|\mathcal{B}_1| \leq m, \quad |\mathcal{B}_2(\mathbf{y}_i)| \leq n_i, \quad i = 1, \ldots, m$$

so that

$$\begin{aligned}
\sum_{i=1}^{m} |\mathcal{B}_2(\mathbf{y}_i)| &\le \sum_{i=1}^{m} n_i \\
&\le m \exp\left\{ n \left[I(X \wedge Y_2|Y_1) + \frac{3}{4}\delta_2' \right] \right\} \\
&\le \exp\left\{ n \left[I(X \wedge Y_1 Y_2) + \frac{3}{4}\delta_2 \right] \right\}.
\end{aligned}$$

Equivalently,

$$\begin{aligned}
\frac{1}{n}\log|\mathcal{B}_1| &\le I(X \wedge Y_1) + \frac{3}{4}\delta_1, \\
\frac{1}{n}\log|\mathcal{B}_2| &\le I(X \wedge Y_1 Y_2) + \frac{3}{4}\delta_2. \qquad (3.10)
\end{aligned}$$

Assume that $I(X \wedge Y_1) \le R_1$. Then by the uniform continuity of $R(P, R_1, \Delta_1, \Delta_2)$ (which follows analogously as in [13, Lemma 2.2.2, pp. 124-125]), the desired inequalities (3.4), (3.5) follow for η_1, η_2 sufficiently small (*c.f.*, (3.6)).

It remains to establish the Claim. To this end, observe that $\mathbf{z} \in \mathcal{A}(\mathbf{x})$ iff

$$\mathbf{z} \in \mathcal{T}^n_{[Y_1]} \quad \text{and} \quad d_1(\mathbf{x}, \mathbf{z}) = \mathbf{E}_{P_\mathbf{x} V}[d_1(X, \tilde{Y})] \le \Delta_1$$

where $(X, \tilde{Y})$ is a $\mathcal{X} \times \mathcal{Y}_1$-valued r.v. with joint pmf equal to the joint type $P_\mathbf{x} V$ of $(\mathbf{x}, \mathbf{z})$. Further,

$$\begin{aligned}
H(V|P_\mathbf{x}) &= H(P_\mathbf{z}) - I(X \wedge \tilde{Y}) \\
&\le H(Y_1) + \frac{\delta'}{3} - \min I(X \wedge \tilde{Y}) \\
&= H(Y_1) + \frac{\delta'}{3} - R(P, \Delta_1), \qquad (3.11)
\end{aligned}$$

where the inequality follows for all n large (depending on δ') from [13, Lemma 1.2.7] and the minimum is taken over all rv's $(X, \tilde{Y})$ such that $P_X = P_\mathbf{x}$ and $\mathbf{E} d_1(X, \tilde{Y}) \le \Delta_1$.

Hence, finally,

$$\mathcal{A}(\mathbf{x}) = \left(\bigcup_{V: \mathbf{E}_{P_{\mathbf{x}}V}\left[d_1(X,\tilde{Y})\right] \le \Delta_1} \mathcal{T}_V^n(\mathbf{x}) \right) \bigcap \mathcal{T}_{[Y_1]}^n,$$

so that,

$$|\mathcal{A}(\mathbf{x})| \le (1+n)^{|\mathcal{X}||\mathcal{Y}_1|} \exp\left\{ n \left[H(Y_1|X) + \frac{2\delta'}{3} \right] \right\},$$

for all $n \ge N_2(d_1, \eta_1, \delta')$, where the inequality above is a consequence of (3.11).

3.5 The Error Exponents

In this section, we shall characterize the error exponents for successive refinement by partitioning. Corollaries 3.2 and 3.1 imply that for $\Delta_2 \le \Delta_1$ and for numbers R_1, R_2 with $R_1 < R_2$ and $R_1 > R(P, \Delta_1)$, $R_2 > R(P, R_1, \Delta_1, \Delta_2)$, there exists a sequence of n-length block codes $(f_1^{(n)}, \phi_1^{(n)}), (f_2^{(n)}, \phi_2^{(n)})$ such that

$$\lim_n \frac{1}{n} \log \parallel f_1^{(n)} \parallel \ = \ R_1 \qquad (3.12)$$

$$\lim_n \frac{1}{n} \log \parallel f_1^{(n)} \parallel \parallel f_2^{(n)} \parallel \ = \ R_2 \qquad (3.13)$$

and

$$\lim_n \Pr\left(d_1(X^n, \phi_1(f_1(X^n))) > \Delta_1 \text{ or } d_2(X^n, \phi_2(f_1(X^n), f_2(X^n))) > \Delta_2 \right) = 0,$$

where $\parallel f_1^{(n)} \parallel$ (resp. $\parallel f_2^{(n)} \parallel$) denotes the cardinality of the domain of $f_1^{(n)}$ (resp. $f_2^{(n)}$). Our objective is to characterize the rate of convergence to zero of the previous probability.

Definition 3.3 *For given distortion measures d_1, d_2, positive numbers $\Delta_1 > \Delta_2 > 0$, n-length block code (f_1, ϕ_1) and n-length refining block code (f_2, ϕ_2), we define the error function $e(P, (f_1, \phi_1), (f_2, \phi_2), \Delta_1, \Delta_2)$ as the probability that the* *source sequence X^n in $\mathcal{X}^n$ of the DMS with (common) pmf P is not reproduced within distortion Δ_1 in the first level of coding, or within distortion Δ_2 by the refining code. Thus the error function for two-level coding is defined as*

$$e(P,(f_1,\phi_1),(f_2,\phi_2),\Delta_1,\Delta_2) \triangleq \tag{3.14}$$
$$\triangleq \Pr\left(d_1(X^n,\phi_1(f_1(X^n))) > \Delta_1 \text{ or } d_2(X^n,\phi_2(f_1(X^n),f_2(X^n))) > \Delta_2\right).$$

Definition 3.4 *The conditional error function $e(P,(f_1,\phi_1),(f_2,\phi_2),\Delta_2|R_1,\Delta_1)$ is the probability that the source sequence X^n in $\mathcal{X}^n$ is not reproduced within distortion Δ_2 by the refining code, given that (f_1,ϕ_1),* i.e., *the code for the first level, has rate that does not exceed R_1 and distortion that does not exceed Δ_1. Thus,*

$$e(P,(f_1,\phi_1),(f_2,\phi_2),\Delta_2|R_1,\Delta_1) \triangleq$$
$$\triangleq \Pr\left(d_2(X^n,\phi_2(f_1(X^n),f_2(X^n))) > \Delta_2\right). \tag{3.15}$$

We show below for suitable two-level n-length block codes with rates converging to R_1 and R_2, that $e(P,(f_1^{(n)},\phi_1^{(n)}),(f_2^{(n)},\phi_2^{(n)}),\Delta_1,\Delta_2)$ converges to zero exponentially rapidly with rate given by the error exponent

$$F(P,R_1,R_2,\Delta_1,\Delta_2) \triangleq \inf_{\substack{Q:\quad R(Q,\Delta_1)>R_1 \\ or\quad R(Q,R_1,\Delta_1,\Delta_2)>R_2}} D(Q\|P), \tag{3.16}$$

provided $R_1 > R(P,\Delta_1)$ and $R_2 > R(P,R_1,\Delta_1,\Delta_2)$. This provides an extension of the result of Marton [40] (*c.f.*, also Csiszár-Körner [13]) on the error-exponent for one-level coding. Our approach is along the lines of Csiszár-Körner [13, section 2.4].

It follows from (3.14) that upon setting $d_1 = d_2, \Delta_1 = \Delta_2, \mathcal{Y}_1 = \mathcal{Y}_2$ and $\phi_2(m_1,m_2) = \phi_1(m_1)$, $m_1 \in \mathcal{M}_1, m_2 \in \mathcal{M}_2$ that

$$e(P,(f_1,\phi_1),(f_2,\phi_2),\Delta_1,\Delta_2) = e(P,(f_1,\phi_1),\Delta_1),$$

where the term on the right side above corresponds to the error function for one-level coding which converges to zero exponentially rapidly with rate given by the error exponent

$$F(P, R_1, \Delta_1) \triangleq \inf_{Q: R(Q, \Delta_1) > R_1} D(Q \| P) \tag{3.17}$$

provided $R_1 > R(P, \Delta_1)$ (*c.f.*, Marton [40], Csiszár-Körner [13]).

We further show that for suitable two-level n-length block codes with rates converging to R_1 and R_2, that $e(P, (f_1, \phi_1), (f_2, \phi_2), \Delta_2 | R_1, \Delta_1)$, *i.e.*, the conditional error function converges to zero exponentially rapidly with rate given by the error exponent

$$F_c(P, R_2, \Delta_2 | R_1, \Delta_1) \triangleq \inf_{Q: R(Q, R_1, \Delta_1, \Delta_2) > R_2} D(Q || P).$$

Theorem 3.4 (Two-level coding error exponent) *Let $\{X_t\}_{t=1}^{\infty}$ be a DMS with pmf P. For every $R_1 < R_2 < \log \mid \mathcal{X} \mid$ and distortion measures d_1 on $\mathcal{X} \times \mathcal{Y}_1$ and d_2 on $\mathcal{X} \times \mathcal{Y}_2$, there exists a sequence of n-length block codes for two-level coding such that*

- $\lim_n \frac{1}{n} \log \parallel f_1^{(n)} \parallel = R_1$ *and* $\lim_n \frac{1}{n} \log \parallel f_1^{(n)} \parallel \parallel f_2^{(n)} \parallel = R_2$; (3.18)

- *for every pmf P on $\mathcal{X}$, $\Delta_1 > \Delta_2 \geq 0$ and $\delta_1 > 0, \delta_2 > 0$*

$$\frac{1}{n} \log e(P, (f_1^{(n)}, \phi_1^{(n)}), \Delta_1) \leq -F(P, R_1, \Delta_1) + \delta_1,$$

$$\frac{1}{n} \log e(P, (f_1^{(n)}, \phi_1^{(n)}), (f_2^{(n)}, \phi_2^{(n)}), \Delta_1, \Delta_2) \leq -F(P, R_1, R_2, \Delta_1, \Delta_2) + \delta_2,$$

and

$$\frac{1}{n} \log e(P, (f_1^{(n)}, \phi_1^{(n)}), (f_2^{(n)}, \phi_2^{(n)}), \Delta_2 | R_1, \Delta_1)$$
$$\leq -F_c(P, R_2, \Delta_2 | R_1, \Delta_1) + \delta_3,$$

whenever $n \geq N(\mid \mathcal{X} \mid, d_1, d_2, \delta_1, \delta_2, \delta_3)$.

Further, for every sequence of codes satisfying (10) and every distribution P on $\mathcal{X}$, the following hold :

$$\liminf_{n\to\infty} \frac{1}{n} \log e(P, (f_1^{(n)}, \phi_1^{(n)}), \Delta_1) \geq -F(P, R_1, \Delta_1),$$

$$\liminf_{n\to\infty} \frac{1}{n} \log e(P, (f_1^{(n)}, \phi_1^{(n)}), (f_2^{(n)}, \phi_2^{(n)}), \Delta_1, \Delta_2) \geq -F(P, R_1, R_2, \Delta_1, \Delta_2),$$

and

$$\liminf_{n\to\infty} \frac{1}{n} \log e(P, (f_1^{(n)}, \phi_1^{(n)}), (f_2^{(n)}, \phi_2^{(n)}), \Delta_2 | R_1, \Delta_1) \geq -F_c(P, R_2, \Delta_2 | R_1, \Delta_1).$$

Remark: Note that the two-level code in the statement of the forward part of Theorem 3.4 does not rely on a knowledge of the pmf P of the DMS. Hence, this code is *universal* in that it is applicable to any $\mathcal{X}$-valued DMS. ■

As is to be expected, the error exponent for two-level coding cannot exceed that for one-level coding. This is obvious from (3.16) by observing for $d_1 = d_2 = d$ and $\mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{Y}$, that

$$F(P, R_1, R_2, \Delta_1, \Delta_2) = \min \left\{ F(P, R_1, \Delta_1), F_c(P, R_2, \Delta_2 | R_1, \Delta_1) \right\}. \tag{3.19}$$

Even with the Markov condition in effect so that the rate-distortion functions for one- and two-level coding coincide, *i.e.*,

$$R(P, R_1, \Delta_1, \Delta_2) = R(P, \Delta_2), \quad \Delta_2 < \Delta_1,$$

note that if

$$F(P, R_1, \Delta_1) < F_c(P, R_2, \Delta_2 | R_1, \Delta_1), \tag{3.20}$$

then

$$F(P, R_1, R_2, \Delta_1, \Delta_2) < F(P, R_2, \Delta_2). \tag{3.21}$$

This is illustrated in the following example.

Example

Let $\{X_t\}_{t=1}^{\infty}$ be a DMS with pmf P, where $P(x) > 0, \;\; x \in \mathcal{X}$. Let $\mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{Y}$, and $d_1 = d_2 = d$ where d denotes Hamming distortion measure. Thus,

$$d(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{t=1}^{n} \mathbf{1}(x_t \neq y_t), \quad \mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}^n.$$

It was shown in [16] that this set-up admits the Markov condition so that

$$R(P, R_1, \Delta_1, \Delta_2) = R(P, \Delta_2).$$

Next, note that

$$R(P, R_1, \Delta_1, \Delta_2) = R(P, \Delta_2) = H(P) - \Delta_2 \log(\mid \mathcal{X} \mid -1) - h_b(\Delta_2), \;\; \Delta_1 \geq \Delta_2,$$

where $H(P)$ is the entropy of P. It then follows from (3.17) and (3.16) that the exponents for one- and two-level coding are :

$$F(P, R_2, \Delta_2) = \min_{Q: H(Q) \geq c(R_2, \Delta_2)} D(Q \| P), \tag{3.22}$$

$$F(P, R_1, R_2, \Delta_1, \Delta_2) = \min_{Q: H(Q) \geq \max\{c(R_1, \Delta_1), c(R_2, \Delta_2)\}} D(Q \| P), \tag{3.23}$$

where

$$c(R, \Delta) \;\triangleq\; \Delta \log(\mid \mathcal{X} \mid -1) + h_b(\Delta) + R.$$

Let

$$\mathcal{Q}_\alpha \triangleq \{Q : H(Q) > \alpha\}, \;\; \alpha \geq 0.$$

Now, the necessary and sufficient condition for (3.20) for (3.21) to hold can be expressed as $\mathcal{Q}_{c(R_1, \Delta_1)} \supset \mathcal{Q}_{c(R_2, \Delta_2)}$, or, equivalently,

$$c(R_1, \Delta_1) < c(R_2, \Delta_2),$$

which is the same as

$$R(P, \Delta_1) < R_1 < R_2 - [R(P, \Delta_2) - R(P, \Delta_1)].$$

The condition above says that, with the Markov condition in effect, the error exponent (3.23) for two-level coding is worse than that for one-level coding iff

$$R_2 - R(P, \Delta_2) > R_1 - R(P, \Delta_1). \tag{3.24}$$

Note that this condition is given by (3.22) and (3.23) as $D(\mathcal{Q}_2 \| P) < D(\mathcal{Q}_1 \| P)$, where

$$D(\mathcal{Q}_i \| P) \triangleq \min_{Q: H(Q) \geq c(R_i, \Delta_i)} D(Q \| P), \quad i = 1, 2.$$

This means that *in the case of the Hamming distortion measures*, (divergence) distances in the space of pmf's on $\mathcal{X}$ correspond to distances (differences) between the actual rates and the corresponding values of the rate-distortion functions.

Clearly, if R_1, R_2 are chosen to violate inequality (3.24), *e.g.*, with R_1 large enough, the two-level coding process will no longer suffer from the disadvantage of a smaller error exponent. ■

Proof of Theorem 3.4 : We commence with the existence part of the proof. It is convenient to define the following quantities : For a pmf Q on $\mathcal{X}$ and $R_2 > R_1 > 0$, let

$$\Delta(Q, R_1) \triangleq \inf_{P_X = Q,\, I(X \wedge Y_1) \leq R_1} \mathbf{E} d_1(X, Y_1)$$

and

$$\Delta(Q, R_1, \Delta_1, R_2) \triangleq \inf_{\substack{P_X = Q \\ I(X \wedge Y_1) \leq R_1 \\ I(X \wedge Y_1 Y_2) \leq R_2 \\ \mathbf{E} d_1(X, Y_1) \leq \Delta_1}} \mathbf{E} d_2(X, Y_2).$$

Consider the sets

$$\mathcal{U}_1^{(n)} \triangleq \bigcup_{Q: R(Q, \Delta_1) > R_1} \mathcal{T}_Q^n$$

and

$$\mathcal{U}_2^{(n)} \triangleq \bigcup_{Q:R(Q,R_1,\Delta_1,\Delta_2)>R_2} \mathcal{T}_Q^n = \bigcup_{Q:R(Q,R_1,\Delta_1,\Delta_2)>R_2} \mathcal{T}_Q^n.$$

Obviously,

$$\begin{aligned} P^n(\mathcal{U}_1^{(n)}) &\leq (n+1)^{|\mathcal{X}|} \exp\left\{-nF(P,R_1,\Delta_1)\right\} \\ &\leq \exp\left\{-n\left[F(P,R_1,\Delta_1)-\delta_1\right]\right\}, \end{aligned}$$

for all n large (*c.f.*, [13, Lemma 1.2.6]), and

$$\begin{aligned} P^n(\mathcal{U}_2^{(n)}) &\leq (n+1)^{|\mathcal{X}|} \exp\left\{-nF_c(P,R_2,\Delta_2|R_1,\Delta_1)\right\} \\ &\leq \exp\left\{-n\left[F_c(P,R_2,\Delta_2|R_1,\Delta_1)-\delta_2\right]\right\}, \end{aligned}$$

for all n large. Next, by Lemma 3.3, there exist sequences $\epsilon_1^{(n)}$ and $\epsilon_2^{(n)}$ with $\lim_n \epsilon_1^{(n)} = 0$ and $\lim_n \epsilon_2^{(n)} = 0$, such that for every type Q of sequences in $\mathcal{X}^n$, there exist sets $\mathcal{B}_{Q,1} \subset \mathcal{Y}_1^n$ and $\mathcal{B}_{Q,2}(\mathbf{y}_1) \subset \mathcal{Y}_2^n, \mathbf{y}_1 \in \mathcal{B}_{Q,1}$, satisfying

$$\frac{1}{n}\log|\mathcal{B}_{Q,1}| \leq R_1 + \epsilon_1^{(n)},$$

$$\frac{1}{n}\log\left(\sum_{\mathbf{y}_1\in\mathcal{B}_{Q,1}} |\mathcal{B}_{Q,2}(\mathbf{y}_1)|\right) \leq R_2 + \epsilon_2^{(n)}$$

and for every $\mathbf{x} \in \mathcal{T}_Q^k$,

$$d_1(\mathbf{x},\mathcal{B}_{Q,1}) \triangleq \min_{\mathbf{y}_1\in\mathcal{B}_{Q,1}} d_1(\mathbf{x},\mathbf{y}_1) \leq \Delta(Q,R_1),$$

and

$$d_2(\mathbf{x},\mathcal{B}_{Q,2}(\mathbf{y}_1)) \triangleq \min_{\mathbf{y}_2\in\mathcal{B}_{Q,2}(\mathbf{y}_1)} d_2(\mathbf{x},\mathbf{y}_2) \leq \Delta(Q,R_1,\Delta_1,R_2),\ \mathbf{y}_1 \in \mathcal{B}_{Q,1},$$

for $\mathbf{x}$ such that $d_1(\mathbf{x},\mathbf{y}_1) \leq \Delta_1$. Next, we set

$$\begin{aligned} \mathcal{B}_1 &\triangleq \bigcup_Q \mathcal{B}_{Q,1}, \\ \mathcal{B}_2(\mathbf{y}_1) &\triangleq \mathcal{B}_{T_{\mathbf{y}_1},2}(\mathbf{y}_1),\quad \mathbf{y}_1 \in \mathcal{B}_1, \end{aligned}$$

where $T_{\mathbf{y}_1}$ is the type of $\mathbf{y}_1$. Then, the Type Counting Lemma [13, Lemma 1.2.2] yields that

$$\frac{1}{n}\log \mid \mathcal{B}_1 \mid \leq R_1 + \eta_1^{(n)},$$

$$\frac{1}{n}\log\left(\sum_{\mathbf{y}_1 \in \mathcal{B}_1} \mid \mathcal{B}_2(\mathbf{y}_1) \mid\right) \leq R_2 + \eta_2^{(n)},$$

where $\lim_n \eta_1^{(n)} = \lim_n \eta_2^{(n)} = 0$. Furthermore, $R(Q, \Delta_1) \leq R_1$ implies in a standard manner that

$$\Delta(Q, R_1) \leq \Delta_1 \tag{3.25}$$

and $R(Q, R_1, \Delta_1, \Delta_2) \leq R_2$ implies

$$\Delta(Q, R_1, R_2) \leq \Delta_2, \tag{3.26}$$

which is seen as follows.

Let $P^*_{XY_1Y_2}$ achieve $R(Q, R_1, \Delta_1, \Delta_2) = I(X \wedge Y_1Y_2) \leq R_2$, where the rv (X, Y_1, Y_2) is distributed according to $P^*_{XY_1Y_2}$. Then,

$$\mathbf{E}d_1(X, Y_1) \leq \Delta_1, \mathbf{E}d_2(X, Y_2) \leq \Delta_2, I(X \wedge Y_1) \leq R_1.$$

Hence, $\Delta(Q, R_1, R_2) \leq \mathbf{E}d(X, Y_2) \leq \Delta_2$, since

$$I(X \wedge Y_1) \leq R_1, I(X \wedge Y_1Y_2) \leq R_2, \mathbf{E}d_1(X, Y_1) \leq \Delta_1 \leq \Delta(Q, R_1).$$

As a consequence of (3.25) and (3.26), we have

$$d_1(\mathbf{x}, \mathcal{B}_1) \leq \Delta_1, \quad \mathbf{x} \in \mathcal{X}^n \setminus \mathcal{U}_1^{(n)},$$

$$d_2(\mathbf{x}, \mathcal{B}_2(\mathbf{y}_1)) \leq \Delta_2, \quad \mathbf{y}_1 \in \mathcal{B}_1, \mathbf{x} \in \mathcal{X}^n \setminus \mathcal{U}_2^{(n)}.$$

These inequalities establish the direct (existence) part of the proof.

Turning to the converse part, consider any pmf Q on $\mathcal{X}$ with the property that

$$R(Q, R_1, \Delta_1, \Delta_2) > R_2 + \delta_2. \tag{3.27}$$

The remainder of the proof relies on the following Claim which constitutes a strong converse to Theorem 3.1.

Claim 3.2 *For any sequence of codes* $(f_1^{(n)}, \phi_1^{(n)})$, $(f_2^{(n)}, \phi_2^{(n)})$ *with*

$$\frac{1}{n}\log\|f_1^{(n)}\| \leq R_1 + \delta_1',$$

$$\frac{1}{n}\log\|f_1^{(n)}\|\|f_2^{(n)}\| \leq R_2 + \delta_2',$$

for all n large, it holds that

$$e(Q, (f_1^{(n)}, \phi_1^{(n)}), (f_2^{(n)}, \phi_2^{(n)}), \Delta_2|R_1, \Delta_1) \geq \frac{1}{2},$$

whenever $n \geq N(d_1, d_2, \delta_1', \delta_2')$.

The proof of the Claim is obtained by mimicking the proof of [13, Theorem 2.2.3], and is, therefore, omitted.

Then, by [13, Corollary 1.1.2], it holds for all n large enough that

$$e(P, (f_1^{(n)}, \phi_1^{(n)}), (f_2^{(n)}, \phi_2^{(n)}), \Delta_2|R_1, \Delta_1) \geq \exp\left\{-n\left[D(Q\|P) + \delta_1\right]\right\}.$$

Note that the case $R(Q, \Delta_1) > R_1 + \delta_1$ alone was shown ([13, Theorem 2.4.5]) to imply that

$$e(P, (f_1^{(n)}, \phi_1^{(n)}), \Delta_1) \geq \exp\left\{-n\left[D(Q\|P) + \delta_1\right]\right\}.$$

Since δ_1, δ_2 were arbitrary as was Q subject to (3.27), the desired converse follows.

It is obvious that since

$$\begin{aligned}
&\max\left\{e(P,(f_1^{(n)},\phi_1^{(n)}),\Delta_1), e(P,(f_1^{(n)},\phi_1^{(n)}),(f_2^{(n)},\phi_2^{(n)}),\Delta_2|R_1,\Delta_1)\right\} \\
&\leq \ e(P,(f_1^{(n)},\phi_1^{(n)}),(f_2^{(n)},\phi_2^{(n)}),\Delta_2|R_1,\Delta_1) \\
&\leq \ e(P,(f_1^{(n)},\phi_1^{(n)}),\Delta_1) + e(P,(f_1^{(n)},\phi_1^{(n)}),(f_2^{(n)},\phi_2^{(n)}),\Delta_2|R_1,\Delta_1),
\end{aligned}$$

the overall error exponent will be the greater of the two. Hence, (3.16) follows.

3.6 Discussion

We have determined the error exponent for the problem of successive refinement by partitioning for a DMS. As expected, it is generally smaller than the error exponent for one-level coding. It is interesting to note that even when the Markov condition (3.3) holds so that the rate-distortion functions coincide for one-level and two-level coding, it may hold that the error exponents for the latter are strictly smaller than that for the former.

For a DMS with Hamming distance distortion measure, a simple necessary and sufficient condition for the error exponents to differ in the presence of Markov condition, can be expressed in terms of the coding rates R_1, R_2 and the one-level rate-distortion function. An extension of this result to arbitrary distortion measures remains unresolved.

Finally, it can be shown that the error function for two-level coding (*c.f.*, (3.14)) goes to one exponentially fast, whenever $R_2 < R(P,\Delta_1,\Delta_2)$. To this end, we conclude from Rimoldi ([46, Theorem 1, converse part]) that if $R_2 < R(P,R_1,\Delta_1,\Delta_2)$, then for every sequence of n-length block codes satisfying

(3.12) and (3.13), we have

$$\lim_n e(P,(f_1^{(n)},\phi_1^{(n)}),(f_2^{(n)},\phi_2^{(n)}),\Delta_1,\Delta_2)=1.$$

Indeed, in analogy with one-level coding (*c.f.*, Csiszár-Körner [13, section 2.4]), this convergence occurs exponentially fast. To see this, define

$$e^{(n)}(P,R_1,R_2,\Delta_1,\Delta_2)\triangleq \min e(P,(f_1^{(n)},\phi_1^{(n)}),(f_2^{(n)},\phi_2^{(n)}),\Delta_1,\Delta_2),$$

where the minimum is taken over all codes satisfying (3.12) with $R_1 > R(P,\Delta_1)$, and $\frac{1}{n}\log \parallel f_1^{(n)} \parallel\parallel f_2^{(n)} \parallel\leq R_2$, where $0 < R_2 < R(P,\Delta_1,\Delta_2)$. Then, it holds that

$$\lim_n \left\{-\frac{1}{n}\log\left[1-e^{(n)}(P,R_1,R_2,\Delta_1,\Delta_2)\right]\right\}=G(R_1,R_2,\Delta_1,\Delta_2),$$

where

$$G(R_1,R_2,\Delta_1,\Delta_2)=\inf_Q\left(D(Q\|P)+|R(Q,R_1\Delta_1,\Delta_2)-R_2|^+\right),$$

with $|x|^+$ denoting $\max\{x,0\}$. The proof is identical to that for one-level coding (*c.f.*, *e.g.*, Csiszár-Körner [13, Problem 2.4.6, pp.158-159]) and relies on the strong converse (Rimoldi, [46, Theorem 1, converse part]).

Chapter 4

Upper Bounds for the Redundancies of One- and Two-Level Block Coding

4.1 Introduction

Let $\{X_t\}_{t=1}^{\infty}$ be a $\mathcal{X}$-valued DMS with pmf P, where $\mathcal{X}$ is a finite set. Let $\mathcal{Y}$ be the (finite) reproduction alphabet. Let $R \geq 0$ be a nonnegative number and $d : \mathcal{X} \times \mathcal{Y} \mapsto \mathbf{R}^+$ be a nonnegative-valued mapping with $\min_y d(x, y) = 0,\ x \in \mathcal{X}$.

From rate-distortion theory, we know that we can find a n-length block code of rate R, such that the expected distortion under P incurred by this code is, for a sufficiently large value of the block-length n, arbitrarily close to the distortion-rate function defined by

$$D(P, R) = \min_{W: I(X \wedge Y) \leq R} \mathbf{E} d(X, Y),$$

where the $\mathcal{X} \times \mathcal{Y}$-valued random variable (X, Y) has joint pmf PW. However, most available rate-distortion results (*e.g.*, the rate-distortion theorem) for DMS's characterize asymptotic behaviour, in that they can be approached only when arbitrarily complex operations (such as exhaustive search) on very long source sequences are allowed. It is, therefore, of interest to characterize the rate at which these limits are attained as the block length n increases.

We shall show that the rate of convergence of the expected distortion of a good code to the distortion-rate function is bounded above by a term which decays with the block length n as $O(\frac{\log n}{n})$. The method used is that of random coding, namely we shall show that if we pick codes at random, the expected value (with respect to the probability law governing the selection of the codes) of the distortion redundancy (the difference between the expected distortion and the distortion-rate function) decays with the block length n as $O(\frac{\log n}{n})$.

The proof for single-level coding is based on a paper by R. J. Pilc [45], where it has been shown that there exists a code for which the distortion redundancy is bounded above by a function which decays as $\alpha(\ln n)/n$. We provide a more rigorous presentation of Pilc's proof, which is based on the notion of types. This makes it easier to extend the proof to the multiresolution case (*c.f.*, Section 4.3).

Next, we shall provide an upper bound for the distortion redundancy in the case of two-level coding, and show that it too decays as $O(\frac{\log n}{n})$.

4.2 An Upper Bound for One-Level Coding

Consider a $\mathcal{X}$-valued DMS with pmf P, and a nonnegative number R, as described in the previous section. A n-length block code (f_n, ϕ_n) of rate R comprises the pair (f_n, ϕ_n), where

$$f_n : \mathcal{X}^n \mapsto \{1, \dots, M\},$$

$$\phi_n : \{1, \dots, M\} \mapsto \mathcal{Y}^n$$

and

$$\frac{\log M}{n} \leq R.$$

We focus attention on those codes (f_n, ϕ_n) where for a given ϕ_n the encoder f_n is chosen so as to obey the nearest-neighbour rule, i.e.,

$$f_n(\mathbf{x}) = \arg \min_{1 \leq i \leq M} d(\mathbf{x}, \phi_n(i)), \tag{4.1}$$

where $d : \mathcal{X}^n \times \mathcal{Y}^n \mapsto [0, \infty)$ is the distortion measure induced by d, i.e.,

$$d(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{t=1}^{n} d(x_t, y_t), \quad \mathbf{x} \in \mathcal{X}^n, \mathbf{y} \in \mathcal{Y}^n.$$

It is obvious that, for a fixed decoder, this choice of f_n yields minimal expected distortion is fixed. Let G be the pmf on $\mathcal{Y}$ given by

$$G(y) = \sum_{x \in \mathcal{X}} P(x) W_P^*(y|x), \quad y \in \mathcal{Y},$$

where W_P^* is the distortion-rate achieving channel, *i.e.*,

$$W_P^* = \arg \min_{W : I(X \wedge Y) \leq R} \mathbf{E} d(X, Y), \tag{4.2}$$

where the $\mathcal{X} \times \mathcal{Y}$-valued rv (X, Y) has joint pmf PW. We shall assume that W_P^* in (4.2) is unique.

Let $\mathbf{Z}_i, i = 1, \ldots, M$, be i.i.d. $\mathcal{Y}^n$-valued rv's each with i.i.d. components and distributed according to a pmf G on $\mathcal{Y}$. Then,

$$\begin{aligned} \Pr\{\mathbf{Z}_1 = \mathbf{z}_1, \ldots, \mathbf{Z}_M = \mathbf{z}_M\} &= \prod_{i=1}^{M} \Pr\{\mathbf{Z}_i = \mathbf{z}_i\} \\ &= \prod_{i=1}^{M} \prod_{t=1}^{n} G(z_{it}). \end{aligned}$$

Observe that any realization of the $(\mathcal{Y}^n)^M$-valued rv $(\mathbf{Z}_1, \ldots, \mathbf{Z}_M)$ defines a set of M $\mathcal{Y}^n$-valued codewords, and, hence, a decoding rule ϕ_n which, in turn, determines an encoder f_n according to (4.1). If we denote by (F_n, Φ_n) the random code defined by the rv's $(\mathbf{Z}_1, \ldots, \mathbf{Z}_M)$, then

$$\Pr\{(F_n, \Phi_n) = (f_n, \phi_n)\} = \prod_{i=1}^{M} \prod_{t=1}^{n} G(z_{it}),$$

where $\phi_n(i) = \mathbf{z}_i,\ i = 1, \ldots, M$.

The expected distortion corresponding to the random code (F_n, Φ_n) with respect to the input pmf, P, is

$$\begin{aligned} d(F_n, \Phi_n) &= \sum_{\mathbf{x} \in \mathcal{X}^n} P^n(\mathbf{x}) d\left(\mathbf{x}, \Phi_n(F_n(\mathbf{x}))\right) \\ &= \sum_{\mathbf{x} \in \mathcal{X}^n} P^n(\mathbf{x}) \left\{ \min_{1 \le i \le M} d(\mathbf{x}, \Phi_n(i)) \right\}. \end{aligned}$$

Note that $d(F_n, \Phi_n)$, as defined above, is a random variable. The expected average distortion over the ensemble of codes $\{(F_n, \Phi_n)\}$ is, therefore,

$$\begin{aligned} d_{av} &= \sum_{(f_n, \phi_n)} \Pr\{(f_n, \phi_n)\} d(f_n, \phi_n) \\ &= \sum_{\mathbf{x} \in \mathcal{X}^n} P^n(\mathbf{x})\ \mathbf{E} \left\{ \min_{1 \le i \le M} d(\mathbf{x}, \mathbf{Z}_i) \right\} \\ &= \sum_{\mathbf{x} \in \mathcal{X}^n} P^n(\mathbf{x}) \ \int_0^\infty \Pr\{ \min_{1 \le i \le M} d(\mathbf{x}, \mathbf{Z}_i) > \delta \} d\delta. \end{aligned} \tag{4.3}$$

Let $G_{i,\mathbf{x}}$ denote the cdf of the rv $d(\mathbf{x}, \mathbf{Z}_i)$.

Claim 4.1 ([45, p. 879]) *For each* $\mathbf{x} \in \mathcal{X}^n$*, it holds that*

$$G_{i,\mathbf{x}} = G_{\mathbf{x}},\ i = 1, \ldots, M.$$

Furthermore, $G_{\mathbf{x}} = G_{T_{\mathbf{x}}}$*, where* $T_{\mathbf{x}}$ *is the type of* $\mathbf{x}$.

Proof: Observe for $i = 1, \ldots, M$, and $\mathbf{x} \in \mathcal{X}^n$, that

$$\begin{aligned} G_{i,\mathbf{x}}(\delta) &= \Pr\{d(\mathbf{x}, \mathbf{Z}_i) \le \delta\} \\ &= \Pr\{\frac{1}{n} \sum_{t=1}^{n} d(x_t, Z_{it}) \le \delta\}. \end{aligned}$$

The distribution of the rv $\sum_{t=1}^{n} d(x_t, Z_{it})$ clearly depends on $\mathbf{x}$ only through its type as a consequence of the fact that $Z_{it},\ t = 1, \ldots, n$ are i.i.d. Thus,

$$G_{i,\mathbf{x}} = G_{i,T_{\mathbf{x}}}.$$

Next, the i.i.d. property of $\mathbf{Z}_i,\ i = 1, \ldots, M$, yields that

$$G_{i,\mathbf{x}} = G_{T_{\mathbf{x}}}, \quad i = 1, \ldots, M,\ \mathbf{x} \in \mathcal{X}^n.$$

■

Consequently, in (4.3),

$$\begin{aligned}
\Pr\left\{\min_{1\le i\le M} d(\mathbf{x}, \mathbf{Z}_i) > \delta\right\} &= \Pr\left\{d(\mathbf{x}, \mathbf{Z}_i) > \delta,\ i = 1, \ldots, M\right\} \\
&= \prod_{i=1}^{M} \Pr\left\{d(\mathbf{x}, \mathbf{Z}_i) > \delta\right\} \\
&= \prod_{i=1}^{M} \left(1 - \Pr\left\{d(\mathbf{x}, \mathbf{Z}_i) \le \delta\right\}\right) \\
&= \left(1 - G_{T_{\mathbf{x}}}(\delta)\right)^M,
\end{aligned}$$

by Claim 4.1.

Next, we obtain exponential bounds in n on $G_{T_{\mathbf{x}}}(\delta)$. To this end, let $q = T_{\mathbf{x}}$ and define for each $\delta > 0$,

$$\begin{aligned}
R^*(q, \delta) &\triangleq \min_{V: d(q,V)\le\delta} \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} q(x)V(y|x) \log \frac{V(y|x)}{G(y)} \\
&= \min_{V: d(q,V)\le\delta} D(qV \| q \times G).
\end{aligned} \tag{4.4}$$

Observe that the minimizer in (4.4) does exist.

Claim 4.2 ([51, 19, 45, eq. (92)]) *It holds that*

$$h_n \exp\left[-n\left(R^*(q, \delta) - \eta_n\right)\right] \le G_q(\delta) \le H_n \exp\left[-nR^*(q, \delta)\right],$$

where h_n and H_n are polynomial in n, and η_n is $O(\frac{1}{n})$.

Remarks: According to the notation adopted in Chapter 2,

$$R^*(q,\delta) = \tilde{R}(q,G,\delta).$$

Also, as observed by Linder *et al* ([38]), if $d_R \triangleq D(P,R)$, then

$$R^*(P,d_R) = R(P,d_R) = R.$$

Note that G is the output pmf induced by the minimizing distortion-rate channel for input pmf P and rate R, or, equivalently, the minimizing rate-distortion channel for pmf P and distortion d_R.

Furthermore, for every $\delta > 0$, $R^*(q,\delta)$ has continuous second order derivatives with respect to q. ■

Proof: Observe that

$$\begin{aligned} G_q(\delta) &= \Pr\{d(\mathbf{x},\mathbf{Z}) \leq \delta\} \\ &= \sum_{V:d(q,V)\leq\delta} \sum_{\mathbf{y}\in\mathcal{T}_V^n(\mathbf{x})} G^n(\mathbf{y}), \end{aligned}$$

where the first summation is over all conditional types V given $\mathbf{x}$ (V-*shells* of $\mathbf{x}$). Then, by [13, §1.2],

$$G_q(\delta) = \sum_{V:d(q,V)\leq\delta} |\mathcal{T}_V^n(\mathbf{x})| \exp\left[-n(H(g_V) + D(g_V\|G))\right],$$

where g_V is a pmf on $\mathcal{Y}$ given by

$$g_V(y) = \sum_{x\in\mathcal{X}} q(x)V(y|x), \quad \mathbf{y}\in\mathcal{Y}.$$

Continuing the bounding above,

$$\begin{aligned} G_q(\delta) &\leq \sum_{V:d(q,V)\leq\delta} \exp\left[-n\left(I(q,V) + D(g_V\|G)\right)\right] \\ &= \sum_{V:d(q,V)\leq\delta} \exp\left[-nD(qV\|q\times G)\right] \\ &\leq (n+1)^{|\mathcal{X}||\mathcal{Y}|} \exp\left[-n \min_{V:d(q,V)\leq\delta} D(qV\|q\times G)\right] \\ &= H_n \exp\left[-nR^*(q,\delta)\right], \end{aligned}$$

where $H_n \triangleq (n+1)^{|\mathcal{X}||\mathcal{Y}|}$.

Let W^* be the minimizing conditional pmf in (4.4) for distortion $\delta - \eta_n$, where $\eta_n \triangleq |\mathcal{Y}| \frac{d_{max}}{n}$, i.e.,

$$R^*(q, \delta - \eta_n) = D(qW^* \| q \times G).$$

Let V^* be a conditional type, such that

$$|V^*(y|x) - W^*(y|x)| \leq \frac{1}{n}, \quad x \in \mathcal{X}, y \in \mathcal{Y}. \tag{4.5}$$

Then, for $\mathbf{y} \in \mathcal{T}^n_{V^*}(\mathbf{x})$, it holds that

$$\begin{aligned} d(\mathbf{x}, \mathbf{y}) &= \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} q(x) V^*(y|x) d(x, y) \\ &\leq \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} q(x) \left(W^*(y|x) + \frac{1}{n} \right) d(x, y) \\ &\leq d(q, W^*) + \frac{d_{max}}{n} |\mathcal{Y}| \\ &= \delta. \end{aligned}$$

Consequently,

$$\begin{aligned} G_{\mathbf{x}}(\delta) &\geq G^n(\mathcal{T}^n_{V^*}(\mathbf{x})) \\ &\geq (n+1)^{-|\mathcal{X}||\mathcal{Y}|} \exp\left\{ -nD(qV^* \| q \times G) \right\}. \end{aligned} \tag{4.6}$$

Furthermore,

$$\begin{aligned} &D(qV^* \| q \times G) - D(qW^* \| q \times G) \qquad (4.7) \\ &\quad = \sum_{x \in \mathcal{X}} q(x) \left[H(W^*|x) - H(V^*|x) \right] + \sum_{x \in \mathcal{X}} q(x) \sum_{y \in \mathcal{Y}} \left[W^*(y|x) - V^*(y|x) \right] \log G(y). \end{aligned}$$

In view of (4.5), [13, Lemma 1.2.7] implies that the first term in the sum above does not exceed $|\mathcal{Y}| \frac{\log n}{n}$. The second term is easily seen to be bounded above by $\frac{1}{n} A(G)$, where $A(G) \triangleq -\sum_{y \in \mathcal{Y}} \log G(y) < \infty$.

Thus, from (4.7),

$$D(qV^*\|q \times G) \leq R^*(q, \delta - \eta_n) + \left[|\mathcal{Y}|\frac{n}{\log n} + \frac{1}{n}A(G)\right],$$

so that from (4.6),

$$G_q(\delta) \geq h_n \exp\left\{-nR^*(q, \delta - \eta_n)\right\},$$

where $h_n \triangleq (n+1)^{-(|\mathcal{X}|+1)|\mathcal{Y}|} \exp\left(A(G)\right)$. ■

We set the size of the codebook to be $M = \exp\{nR\}$, so that the rate of the code is R. The average distortion will then be

$$\begin{aligned} d_{av} &= \sum_{q \in \mathcal{P}} \sum_{\mathbf{x} \in \mathcal{T}_q^n} P^n(\mathbf{x}) \int_0^{d_{max}} \left[1 - G_q(\delta)\right]^M d\delta \\ &\leq \sum_{q \in \mathcal{P}} P^n(\mathcal{T}_q^n) \int_0^{d_{max}} \left[1 - h_n \exp\{-nR^*(q, \delta - \eta_n)\}\right]^{\exp\{nR\}} d\delta, \quad (4.8) \end{aligned}$$

where $\mathcal{P}_n$ is the set of all types in $\mathcal{X}^n$, and $d_{max} \triangleq \max_{x \in \mathcal{X}, y \in \mathcal{Y}} d(x, y)$. Since $R^*(q, \delta)$ is decreasing in δ (as can be observed from its definition (4.4)), decreasing in δ, $1 - h_n e^{-nR^*(q,\delta)}$ will also decrease in δ and is bounded above by unity. Let $d_R(q)$ be chosen such that

$$R^*(q, d_R(q)) = R. \quad (4.9)$$

In terms of the mismatched distortion-rate function, $d_R(q)$ can be written as

$$d_R(q) = \tilde{D}(q, G, R),$$

and equals $D(q, R)$ only for $q = P$. The integral in (4.8) is bounded above by

$$[d_R(q) + \Delta_n] \cdot 1 + [d_{max} - d_R(q) - \Delta_n] \cdot \left[1 - h_n e^{-nR^*(q, d_R(q) + \Delta_n - \eta_n)}\right]^{e^{nR}},$$

which is further bounded above by

$$d_R(q) + \Delta_n + d_{max} \exp\left\{-h_n \exp\left\{-n\left[R - R^*(q, d_R(q) + \Delta_n - \eta_n)\right]\right\}\right\}, \quad (4.10)$$

using the inequality $1-a \leq \exp(-a),\ a \in [0,1]$. Here, Δ_n is a sequence of positive numbers to be determined later. Using (4.9), we get

$$
\begin{aligned}
& R - R^*(q, d_R(q) + \Delta_n - \eta_n) && (4.11)\\
= & -(\Delta_n - \eta_n)\frac{\partial}{\partial\delta}s_q - \frac{(\Delta_n-\eta_n)^2}{2}\frac{\partial^2}{\partial\delta^2}R^*(q, d_R(q)+\delta)\bigg|_{\delta=\tilde{\delta}}, && (4.12)
\end{aligned}
$$

where

$$
s_q = \frac{\partial}{\partial\delta}\ R^*(q,\delta)|_{\delta=d_R(q)}
$$

and $0 \leq \tilde{\delta} \leq \Delta_n - \eta_n$, for n large enough (so that $\Delta_n - \eta_n \geq 0$).

If we take Δ_n to satisfy

$$
A\frac{\log n}{(-s_q)n} \leq \Delta_n \leq (A+\epsilon)\frac{\log n}{(-s_q)n},
$$

where $\epsilon > 0$ is suitably small and $A = (|\mathcal{X}|+1)|\mathcal{Y}|$, then the second term on the right side of (4.12) is easily seen to decay as $o(\frac{\log n}{n})$ and the second term of the bound in (4.10) decays as $\frac{\log n}{n}$, whereas the third term decays as $\exp\{-n^A\}$, i.e., exponentially fast. Thus the upper bound for the integral in (4.8) becomes

$$
d_R(q) + (A+\epsilon)[1+o(1)]\frac{\log n}{(-s_q)n}.
$$

Hence,

$$
d_{av} \leq \sum_{q\in\mathcal{P}_n} P^n(\mathcal{T}_q^n)\left\{d_R(q) + (A+\epsilon)[1+o(1)]\frac{\log n}{(-s_q)n}\right\}. \qquad (4.13)
$$

The final step is to take the Taylor series expansion of the expression on the right side of (4.13) around $q = P$. Thus,

$$
\begin{aligned}
d_{av} \leq\ & d_R(q) + (A+\epsilon)[1+o(1)]\frac{\log n}{(-s_P)n}\\
& + \left[\nabla_q\left(d_R(q) + (A+\epsilon)[1+o(1)]\frac{\log n}{(-s_q)n}\right)\right]^T_{q=P} \cdot \sum_{q\in\mathcal{P}_n} P^n(\mathcal{T}_q^n)(q-P)\\
& + \sum_{q\in\mathcal{P}_n} P^n(\mathcal{T}_q^n)(q-P)^T\overline{Q}(q-P),
\end{aligned}
$$

where $\overline{Q}$ is a $|\mathcal{X}|$ by $|\mathcal{X}|$ Jacobian matrix given by

$$\overline{Q} \triangleq \left[\nabla_q \nabla_q^T \left(d_R(q) + (A+\epsilon)[1+o(1)] \frac{\log n}{(-s_q)n} \right)\right]_{q=\overline{q}}.$$

Here, P and q are treated as $|\mathcal{X}|$-length vectors.

The following Claim 4.3 appears in Pilc's work [45] without proof. A simple proof is provided below.

Claim 4.3 ([45, p. 882]) *It holds that*

$$\sum_{q\in\mathcal{P}_n} P^n(\mathcal{T}_q^n)(q-P) = 0, \tag{4.14}$$

for all n. *Furthermore,*

$$\sum_{q\in\mathcal{P}_n} P^n(\mathcal{T}_q^n)(q-P)^T\overline{Q}(q-P) = O(\frac{1}{n}). \tag{4.15}$$

Proof: It is easier to compute first the moment-generating function $\Phi(k) = \mathbf{E}_q\left[e^{k^Tq}\right]$, where k is a $|\mathcal{X}|$-length vector and then the first and second moments (equations (4.14) and (4.15)). Observe that

$$\begin{aligned}
\Phi(k) &= \sum_{q\in\mathcal{P}_n} P^n(\mathcal{T}_q^n)e^{k^Tq} \\
&= \sum_{q\in\mathcal{P}_n} \frac{n!}{\prod_{i\in\mathcal{X}}(nq(i))!}\left(\prod_{i\in\mathcal{X}} P(i)^{nq(i)}\right) e^{\sum_{i\in\mathcal{X}} k_i q(i)} \\
&\overset{n_i \triangleq nq(i)}{=} \sum \frac{n!}{\prod_{i\in\mathcal{X}} n_i!}\prod_{i\in\mathcal{X}}\left(P(i)e^{\frac{1}{n}k_i}\right)^{n_i} \qquad (4.16)\\
&= \left(\sum_{i\in\mathcal{X}} P(i)e^{\frac{1}{n}k_i}\right)^n, \qquad (4.17)
\end{aligned}$$

where the sum in (4.16) is taken over all positive integers $n_i,\ i\in\mathcal{X}$ such that $\sum_{i\in\mathcal{X}} n_i = n$, and (4.17) follows from the standard multinomial formula. Taking the derivatives of $\Phi(k)$ with respect to k_i and the second order derivatives with respect to k_i and k_j we get

$$\begin{aligned}
\sum_{q\in\mathcal{P}_n} \frac{n!}{\prod_{i\in\mathcal{X}}(nq(i))!} \prod_{i\in\mathcal{X}} P(i)^{nq(i)} q(j) &= P(j), \quad j \in \mathcal{X} \\
\sum_{q\in\mathcal{P}_n} \frac{n!}{\prod_{i\in\mathcal{X}}(nq(i))!} \prod_{i\in\mathcal{X}} P(i)^{nq(i)} q(j)q(m) &= \frac{1}{n}(P(j)\delta_{jm} - P(j)P(m)) \\
&\quad +P(j)P(m), \quad j,m \in \mathcal{X},
\end{aligned}$$

with $\delta_{jm} = \mathbf{1}(j = m)$ being the Kronecker delta function. We can finally easily verify (4.14) and (4.15). ■

As a result of the previous analysis we state the following :

Theorem 4.1 *The minimum attainable expected distortion for a DMS with pmf P is bounded above by*

$$d_{av} \leq D(P, R) + (A + \epsilon)[1 + o(1)] \frac{\log n}{(-s_P)n},$$

for all n large enough.

4.3 The two-level coding case

As in Chapter 3, let $\mathcal{X}, \mathcal{Y}_1$ and $\mathcal{Y}_2$ be finite alphabets and $d_1 : \mathcal{X} \times \mathcal{Y}_1 \mapsto [0, \infty), d_2 : \mathcal{X} \times \mathcal{Y}_2 \mapsto [0, \infty)$ be (finite) distortion measures. Suppose that we wish to describe the $\mathcal{X}$-valued DMS $\{X_t\}_{t=1}^{\infty}$ with pmf P, at rate R_1 with codewords drawn from $\mathcal{Y}_1$. We have shown in Chapter 3 that there exists a n-length block code of rate R_1 and distortion that converges as $\frac{\log n}{n}$ to the distortion-rate function $D(P, R_1)$.

Suppose now that we wish to provide additional information at rate $R_2 - R_1$ (so that the overall rate is R_2) in order to achieve a better description (with less distortion). We know from Chapter 3, Theorem 3.1, that the minimum distortion which can be achieved using a two-level n-length block code, asymptotically in n, is $D(P, R_1, R_2)$. We are interested in investigating the rate of convergence of the expected value of the distortion to $D(P, R_1, R_2)$, where the two-level distortion-rate function $D(P, R_1, R_2)$ is defined as

$$D(P, R_1, R_2) \triangleq \inf_{\substack{P_{XY_1Y_2}: \quad P_X = P \\ I(X \wedge Y_1) \leq R_1 \\ I(X \wedge Y_1 Y_2) \leq R_2 \\ \mathbf{E} d_1(X;Y_1) \leq D(P, R_1)}} \mathbf{E} d_2(X; Y_2).$$

To this end, recall the definition of two-level n-length block codes. A two-level n-length block code of rates (R_1, R_2) comprises

- a n-length block code $(f_1^{(n)}, \phi_1^{(n)})$, where

  $$f_1^{(n)} : \mathcal{X}^n \mapsto \{1, \ldots, M\},$$

  $$\phi_1^{(n)} : \{1, \ldots, M\} \mapsto \mathcal{Y}_1^n$$

  and

  $$\frac{\log M}{n} \leq R_1.$$

- a n-length refinement block code $(f_2^{(n)}, \phi_2^{(n)})$, where

  $$f_2^{(n)} : \mathcal{X}^n \mapsto \{1, \ldots, N\},$$

  $$\phi_2^{(n)} : \{1, \ldots, M\} \times \{1, \ldots, N\} \mapsto \mathcal{Y}_2^n$$

  and

  $$\frac{\log MN}{n} \leq R_2.$$

We shall concentrate on those two-level block codes for which, for given decoders $\phi_1^{(n)}, \phi_2^{(n)}$, the encoders $f_1^{(n)}, f_2^{(n)}$ are chosen so as to obey the nearest-neighbour rule, i.e.,

$$f_1^{(n)}(\mathbf{x}) = \arg \min_{1 \leq i \leq M} d_1(\mathbf{x}, \phi_1^{(n)}(i)).$$

and

$$f_2^{(n)}(\mathbf{x}) = \arg \min_{1 \leq j \leq N} d_1(\mathbf{x}, \phi_2^{(n)}(f_1^{(n)}(\mathbf{x}), j))). \tag{4.18}$$

Let G be the pmf on $\mathcal{Y}_1 \times \mathcal{Y}_2$ given by

$$G(y_1, y_2) = \sum_{x \in \mathcal{X}} P(x) U^*(y_1, y_2|x), \quad y_1 \in \mathcal{Y}_1, y_2 \in \mathcal{Y}_2, \tag{4.19}$$

where $U^* : \mathcal{X} \mapsto \mathcal{Y}_1 \times \mathcal{Y}_2$ is a channel given by

$$U^* = \arg \min_{\substack{U: \quad I(X \wedge Y_1) \leq R_1 \\ I(X \wedge Y_1 Y_2) \leq R_2 \\ \mathbf{E} d_1(X; Y_1) \leq \Delta_1}} \mathbf{E} d_2(X; Y_2),$$

and the $\mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2$-valued rv (X, Y_1, Y_2) has joint pmf PU.

Further, define the conditional pmf G as

$$G(y_2|y_1) = \frac{G(y_1, y_2)}{\sum_{y_2 \in \mathcal{Y}_2} G(y_1, y_2)}, \quad y_1 \in \mathcal{Y}_1, y_2 \in \mathcal{Y}_2.$$

Let $\{\mathbf{z}_i \in \mathcal{Y}_1, i = 1, \ldots, M\}$, be the codeword set of a good code, in the same sense as in Chapter 2, for the first level and let $\mathbf{W}_{ij},\ i = 1, \ldots, M,\ j = 1, \ldots, N$, be i.i.d. $\mathcal{Y}_2^n$-valued rv's with

$$\Pr\{\mathbf{W}_{ij} = \mathbf{w}_{ij},\ i = 1 \ldots, M,\ j = 1 \ldots, N\} = \prod_{i=1}^{M} \left[\prod_{j=1}^{N} \prod_{t=1}^{n} G(w_{ijt}|z_{it}) \right].$$

Again, observe that any realization of the $(\mathcal{Y}_2^n)^{MN}$-valued random set $\{\mathbf{W}_{ij},\ i = 1, \ldots, M,\ j = 1, \ldots, N\}$ defines a set of $M \times N$ $\mathcal{Y}_2^n$-valued codewords, and, hence, a decoding rule $\phi_2^{(n)}$ which, in turn, determines an encoder $f_2^{(n)}$ according to (4.18).

Following an approach similar to that in the previous section, the expected average distortion over the ensemble of codes $\{(f_2^{(n)}, \phi_2^{(n)})\}$ will be

$$\begin{aligned} d_{av} &= \sum_{\mathbf{x}\in\mathcal{X}^n} P^n(\mathbf{x})\, \mathbf{E}\left\{\min_{1\le j\le N} d_2(\mathbf{x}, \mathbf{W}_{f_1^{(n)}(\mathbf{x}),j})\right\} \\ &= \sum_{\mathbf{x}\in\mathcal{X}^n} P^n(\mathbf{x}) \int_0^\infty \Pr\{\min_{1\le j\le N} d_2(\mathbf{x}, \mathbf{W}_{f_1^{(n)}(\mathbf{x}),j}) > \delta\} d\delta. \end{aligned}$$

It it straightforward to extend the methodology of the previous section to the present case. Thus,

$$\Pr\left\{\min_{1\le j\le N} d(\mathbf{x}, \mathbf{W}_{f_1^{(n)}(\mathbf{x}),j}) > \delta\right\} = \prod_{j=1}^N \Pr\left\{d(\mathbf{x}, \mathbf{W}_{f_1^{(n)}(\mathbf{x}),j}) > \delta\right\}.$$

Claim 4.4 *It holds that :*

$$\begin{aligned} & h_n \exp\left[-n\left(\delta R^*(T_{\mathbf{x}\mathbf{z}_{f_1^{(n)}(\mathbf{x})}}, \delta) - \eta_n\right)\right] \le \\ & \quad \Pr\left\{d(\mathbf{x}, \mathbf{W}_{f_1^{(n)}(\mathbf{x}),j}) > \delta\right\} \le \quad H_n \exp\left[-n\delta R^*(T_{\mathbf{x}\mathbf{z}_{f_1^{(n)}(\mathbf{x})}}, \delta)\right], \end{aligned}$$

where $T_{\mathbf{x}\mathbf{z}_i}$ is the joint type of $\mathbf{x}$ and $\mathbf{z}_i$,

$$\begin{aligned} \delta R^*(r,\delta) &\triangleq \min_{U: d_2(r,U)\le\delta} \sum_{x,y_1,y_2} r(x,y_1)U(y_2|x,y_1)\log\frac{U(y_2|x,y_1)}{G(y_2|y_1)} \\ &= \min_{U: d_2(r,U)\le\delta} D(rU\|q\times vG). \end{aligned} \tag{4.20}$$

for every pmf r on $\mathcal{X}\times\mathcal{Y}_1$ with marginals q on $\mathcal{X}$ and v on $\mathcal{Y}_1$, h_n and H_n are polynomial in n, and $\eta_n = O(\frac{1}{n})$. Here, U ranges over all stochastic matrices $U: \mathcal{X}\times\mathcal{Y}_1 \mapsto \mathcal{Y}_2$ and

$$d_2(r,U) = \sum_{x,y_1,y_2} r(x,y_1)U(y_2|x,y_1)d(x,y_2).$$

The function $\delta R^*(r, \delta)$ corresponds to the excess rate-distortion function, which is defined as

$$\delta R^*(r, \delta) \triangleq \inf_{P_{XY_1Y_2}: P_{XY_1}=r, \mathbf{E}d_2(X;Y_2)\le\delta} I(X \wedge Y_2|Y_1).$$

Also, observe that the minimizer in (4.20) does exist.

Proof: The proof is analogous to that for Claim 4.2. The only difference is that the functions h_n and H_n decay as $n^{-(|\mathcal{X}||\mathcal{Y}_1|+1)|\mathcal{Y}_2|}$ and $n^{-|\mathcal{X}||\mathcal{Y}_1||\mathcal{Y}_2|}$ respectively. ■

We set $M = \exp\{nR_1\}$ and $N = \exp\{n\delta R\}$, where $\delta R \triangleq R_2 - R_1$, so that the overall rate is R_2. The average distortion will then be bounded as

$$d_{av} \le \sum_{\mathbf{x}\in\mathcal{X}} P^n(\mathbf{x}) \int_0^{d_{max}} \left[1 - h_n \exp\{-n\delta R^*(T_{\mathbf{xz}_{f_1^{(n)}(\mathbf{x})}}, \delta - \eta_n)\}\right]^{\exp\{n\delta R\}} d\delta, \tag{4.21}$$

and the integral in (4.21) is bounded above by

$$d_{\delta R}(r) + \Delta_n + d_{max} \exp\left\{-h_n \exp\left\{-n\left[\delta R - \delta R^*(r, d_{\delta R}(r) + \Delta_n - \eta_n)\right]\right\}\right\},$$

where $r \triangleq T_{\mathbf{xz}_{f_1^{(n)}(\mathbf{x})}}$ and $d_{\delta R}(r)$ is chosen such that

$$\delta R^*(r, d_{\delta R}(r)) = \delta R. \tag{4.22}$$

Finally, Δ_n is a positive sequence which is chosen to satisfy

$$A\frac{\log n}{(-s_r)n} \le \Delta_n \le (A+\epsilon)\frac{\log n}{(-s_r)n},$$

with $A = (|\mathcal{X}||\mathcal{Y}_1| + 1)|\mathcal{Y}_2|$ and $\epsilon > 0$ being suitably small. Here, s_r is the slope of the excess rate-distortion function at distortion $d_{\delta R}(r)$, i.e.,

$$s_r \triangleq \frac{\partial}{\partial\delta}\delta R^*(r, \delta)\bigg|_{\delta=d_{\delta R}(r)}.$$

Thus, inequality (4.21) yields

$$d_{av} \le \sum_{\mathbf{x}\in\mathcal{X}} P^n(\mathbf{x}) \left\{d_{\delta R}(r) + (A+\epsilon)[1+o(1)]\frac{\log n}{(-s_r)n}\right\}. \tag{4.23}$$

If r were approximately equal to PW_P^*, where W_P^* is the minimizing channel in (4.4) for pmf P and rate R_1, i.e.,

$$I^*(P, W_P^*) \triangleq \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}_1} P(x) W_P^*(y|x) \log \frac{W_P^*(y|x)}{G(y)} = R_1$$

and

$$d_1(P, W_P^*) \triangleq \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P(x) W_P^*(y|x) d_1(x, y) = D(P, R_1),$$

then the proof of Theorem 4.2 below would be complete. Thus, the assertion of Theorem 4.2 below is contingent on the validity of the following Claim 4.5 whose proof is not available to date.

Claim 4.5 *If $(\hat{f}_1^{(n)}, \hat{\phi}_1^{(n)})$ is a good n-length block code of rate R_1 for the first level, then for n suitably large, $r_{\mathbf{x}}^{(n)}$ is close to PW_P^* for most source sequences $\mathbf{x} \in \mathcal{X}^n$. To be specific, for all n suitably large,*

$$P^n\left(\{\mathbf{x} : \|r_{\mathbf{x}} - PW_P^*\|_\infty > \gamma_n\}\right) \leq \eta_n,$$

where $\gamma_n, \eta_n = O(\frac{\log n}{n})$, and $r_{\mathbf{x}}$ is the joint type of $\mathbf{x}$ and $\phi_1^{(n)}(f_1^{(n)}(\mathbf{x}))$.

Note that the set $\{\mathbf{x} : \|r_{\mathbf{x}} - PW_P^*\|_\infty > \gamma_n\}$ is the set $\mathcal{N}_n(\gamma_n)$, according to the notation adopted in Section 2.6. Thus, Claim 4.5 is simply an extension of Lemma 2.4. According to the latter result, the probability of the set $\mathcal{N}_n(\gamma_n)$ is bounded above by $\exp(-n\delta_n)$, where δ_n depends on γ_n. Clearly, we must take $\gamma_n = O(\frac{\log n}{n})$. Hence, if the dependence of δ_n on γ_n is linear, the statement of the claim is true. Referring to the proof of Lemma 2.2, if we assume that the mismatched rate-distortion function $\tilde{R}(Q, T_{\mathbf{y}}, \Delta)$ is differentiable with respect to its arguments, then the functions c_1, c_2 and c_3 are locally linear for small values of γ. Under this assumption, the dependence of the exponent δ on the "width" γ is linear, and the result of the Claim 4.5 follows.

In view of Claim 4.5, (4.23) yields

$$d_{av} \leq \eta_n d_{max} + \sum_{\mathbf{x}:\|r_{\mathbf{x}}-PW_P^*\|_\infty \leq \epsilon_n} P^n(\mathbf{x}) \left[d_{\delta R}(r_{\mathbf{x}}) + (A+\epsilon)(1+o(1)) \frac{\log n}{(-s_{r_{\mathbf{x}}})n} \right]. \tag{4.24}$$

If we take the Taylor expansion of (4.24) around $r_{\mathbf{x}} = PW_P^*$, and use the fact that the sequences $\mathbf{x}$ on the right side of (4.24) satisfy $\|r_{\mathbf{x}} - PW_P^*\|_\infty \leq \epsilon_n$, we conclude that

$$d_{av} \leq D(P, R_1, R_2) + (A+\epsilon)[1+o(1)] \frac{\log n}{(-s_{PW_P^*})n}.$$

We, therefore, state the following :

Theorem 4.2 *The minimum attainable expected distortion for two-level compression of a DMS with pmf P satisfies*

$$d_{av} \leq D(P, R_1, R_2) + (A+\epsilon)[1+o(1)] \frac{\log n}{(-s_{PW_P^*})n}.$$

Chapter 5

Multicasting over Network Links with Capacity Constraints

5.1 Introduction

Multicasting is a data transmission method in which a source broadcasts information of the same nature to multiple recipients which are connected via a communication network. The development of Broadband Integrated-Services Digital Networks (B-ISDN) will render multicasting techniques more feasible and, perhaps, a standard feature in network services. In this work, we address the problem of *optimal multicasting* and devise efficient algorithms which will achieve near optimal (if not optimal) performance.

The source coding technique under consideration is hierarchical (or multiresolution) data compression. This is a signal representation technique in which the source signal is partitioned into substreams (levels), each of which represents a portion of the signal. The performance of this scheme, as should be expected, is inferior to that of conventional, single-level data-compression schemes in terms of compression rate. On the other hand, this scheme is preferable for applications such as multicasting, owing to its flexibility.

The optimization problem considered below is one of selecting the quantization rates and the transmission routes for each level of hierarchically compressed information, given the number of levels which we wish to transmit and the capacities of the network links. There is a trade-off in this choice. We can use higher rates for each level in order to improve the quality of the signal received by the various users. However, increasing the rates will result in fewer levels reaching the users and, hence, only limited service will be afforded to some users owing to a saturation of link capacities. As an example, consider the case where a small number of users is linked to the source by paths with high capacity values while the rest of the users receive the source information via paths of lower capacities. We can either use a low quantization rate and transmit to all network nodes, or use a higher rate which will result in fewer users receiving the information. In the former situation, all users will receive a low-quality description, whereas in the latter some will receive a high-quality description, while others may receive nothing at all.

In order to systematically decide which of the two situations in the previous example is preferable, we introduce two cost criteria : the maximum and the average. The former seeks to improve the lot of the user with the worst reception quality, thereby ensuring that all users receive the same information, whereas the latter is concerned with the average overall distortion incurred by all users. The maximum problem is the less complex of the two and we present an algorithm which fully solves it. We refer the interested reader to [18] for related work on packing spanning trees in capacitated graphs.

The average cost problem is an extremely complex one. It resembles the

Steiner Tree Problem (*c.f.*, *e.g.*, [26]) in that it involves searching through all possible spanning trees to find the optimal routing scheme. There exist algorithms which solve the Steiner problem in exponential (in the number of nodes) time, *e.g.*, [24, 53]. We would, therefore, expect that complexity of the average cost problem to be at least exponential in both the number of nodes and the number of information levels. In order to address the (NP-complete) Steiner problem in graphs, a number of heuristic algorithms have been proposed (*c.f.*, *e.g.*, [34]), which construct a suboptimal spanning tree using one of several known algorithms (*c.f.*, *e.g.*, [8]). Most of the work to date in multicasting involves optimization with respect to the average cost or delay of the multicast tree, *e.g.*, [2, 4, 14, 31, 64].

Our work is closer to that of N. Shacham [49, 50], where optimality is sought with respect to a criterion characterizing the quality of reception under a capacity constraint. Our formulation differs from that of Shacham [49, 50] in a number of ways. Firstly, we allow different levels to be routed over different paths. Next, our optimality criterion (the distortion-rate function) is a function of all the rate levels, instead of only their sum. Also, the multicast paths in our approach can be set up by a distributed algorithm; however, unlike in the minimum cost (Steiner Tree) problem, where the entire algorithm can be implemented in a distributed manner (*c.f.*, *e.g.*, [5]), our rate selection procedure must be centralized. For the average case, we propose a heuristic algorithm which is a variant of our maximum algorithm, and which can also be implemented in a distributed manner (for the multicast path set-up part); this algorithm yields "near-optimal" performance as is shown by our simulation results.

The remainder of this chapter is organized as follows. In Section 5.2, we

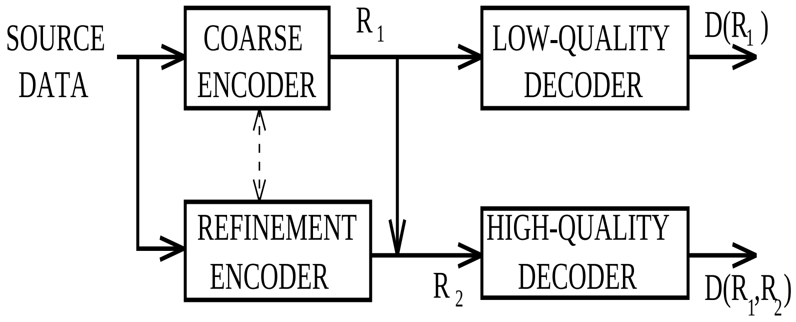

Figure 5.1: Two-stage encoding scheme

present the distortion-rate function for multiple-level quantization and the properties which render it an appropriate choice as the performance criterion in the multicasting problem. In Section 5.3, we formulate the problem of selecting the multicast routes and rates. Sections 5.4 and 5.5 address the maximum and average cost problems, respectively. Simulation results are presented in Section 5.6. Concluding remarks are contained in Section 5.7.

5.2 The distortion-rate function for hierarchical data compression

A schematic for hierarchical data compression is shown in Figure 5.1. The encoder at the first level of compression uses a low quantization rate, *e.g.* a few bits per source sample, to provide a "coarse" description of the source signal. In the next (second) level of compression, "add-on" information is provided at an additional rate (in bits per source sample), yielding a more detailed and enchanced description of the original source signal. The decoder is capable of producing two different reconstructions of the source signal. The first decoder produces a reconstruction of lower quality based only on the first level of compression. The second decoder, on the other hand, combines the information contained in both levels to produce a reconstruction of better quality (which is finer or more detailed). This idea can be extended to K ($K \geq 2$) levels of description, so that decoder k can produce a progressively refined reconstruction of the signal by combining the information contained in levels $1, \ldots, k,\ k = 1, \ldots, K$.

In order to be economical, we shall encode each level at the least possible rate adequate for the distortion specified for that level. In other words, for the first level, we compress the signal optimally so as to attain a prespecified quality or distortion. For the second level of refinement, we shall use as low a rate as possible under the restriction that the first level is compressed optimally, and so on. In order to be specific, we should mention what these "optimal rates" are for each stage.

We begin by outlining the information theoretic underpinnings of our approach to the problem of multicasting. Specifically, we shall briefly describe the distortion-rate function and some of its properties which make it an appropriate choice as a cost criterion for optimum rate selection and routing in a multicasting situation. We refer the reader to [7] for a detailed treatment of the distortion-rate function. A block code consists of an encoder and a decoder which operate on blocks of source symbols (rather than on individual source symbols) so as to exploit the statistical structure inherent in longer source blocks. The rate of such a code is defined as the normalized logarithm of the number of codewords (*e.g.*, corresponding to the quantization levels) used by the encoder, where the normalization is with respect to the number of symbols in a source block. Thus, the rate of the code can be viewed as the number of bits or the precision with which the encoder describes the source signal.

We now summarize two relevant and key information-theoretic results concerning the minimum rate of a code with an expected distortion not exceeding a prespecified level, as well as the minimum expected distortion which can be achieved by a code of given rate. The reader is referred to [7, 13] for details.

We are given an independent and identically distributed (i.i.d.) random process model for an information source which emits sumbols from a (finite) alphabet $\mathcal{X}$. The probability mass function (pmf) P on $\mathcal{X}$, which characterizes this i.i.d. process, is assumed to be known. The distortion between a sequence of source outputs and its representation by a codeword is quantified in terms of a distortion measure $d(\cdot,\cdot)$ defined on $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{Y}$ is the (finite) reproduction alphabet from which sumbols constituting the codeword are selected. Examples of distortion measures are the Euclidean distance $d(\S, \dagger) = |\S - \dagger|$ and the Hamming distance $1(\S = \dagger)$, $\S \in \mathcal{X}$, $\dagger \in \mathcal{Y}$, where $1(\cdot)$ denotes the indicator function. We are also given a number $\Delta > 0$ which denotes the expected value of distortion that can be tolerated in representing a sequence of source symbols by its codeword. It can then be shown (*c.f.*, *e.g.*, [7, 13]), that there exists a block code for this source, with coldwords of block-length n symbols where n is sufficiently large, for which the expected value of distortion (as measured by the "distortion metric" $d : \mathcal{X} \times \mathcal{Y} \mapsto \mathbf{R}^+$) does not exceed Δ, and whose rate is arbitrarily close to the *rate-distortion function* defined by

$$R(\Delta) = \min_{P_{XY}: P_X = P, \mathbf{E}d(X,Y) \le \Delta} I(X \wedge Y).$$

Here, $I(X \wedge Y)$ denotes the mutual information between the $\mathcal{X}$- and $\mathcal{Y}$-valued random variables (rv's) X and Y with joint pmf P_{XY} and $\mathcal{X}$-marginal pmf $P_X =$

P.

Next, given a rate $R > 0$, the *distortion-rate function*, defined as

$$\Delta(R) = \min_{P_{XY}:P_X=P,\ I(X\wedge Y)\leq R} \mathbf{E}d(X,Y),$$

quantifies the minimum expected distortion which can be achieved by a block code of sufficiently large block-length n, with rate not exceeding R.

The notions of rate-distortion function and distortion-rate function can be extended to the case of two-level coding. Specifically, we can characterize the minimum coding rates which are achievable for two-level coding when the expected values of distortion in the first and the second levels do not exceed Δ_1 and Δ_2, respectively, as measured by the "distortion metrics" d_1 and d_2 (*c.f.*, *e.g.*, [46, 29]). Next, assume that the level rates (*e.g.*, code rates for the two levels of compression) R_1 and R_2 are given, where R_1 is the rate of the first-level code, and R_2 is the "add-on" rate of the second-level code. We can then define a *two-level distortion-rate function* as

$$\Delta(R_1, R_1+R_2) = \inf_{\substack{P_X=P \\ \mathbf{E}d_1(X,Y_1)\leq\Delta(R_1) \\ I(X\wedge Y_1)\leq R_1 \\ I(X\wedge Y_1Y_2)\leq R_1+R_2}} \mathbf{E}d_2(X,Y_2), \quad 0 \leq R_1, R_2 \tag{5.1}$$

where (X, Y_1, Y_2) is a $\mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2$-valued rv with pmf $P_{XY_1Y_2}$. The two-level distortion-rate function quantifies the smallest value of expected distortion which the decoder can achieve in reconstruction, when the first level is encoded *optimally* at rate R_1 and the second level contains additional information encoded at rate R_2.

The two-level distortion-rate function, $\Delta(R_1, R_1 + R_2)$, defined above has some interesting properties, which can be easily verified :

1. $\Delta(R_1, R_1 + R_2)$ is jointly convex in $(R_1, R_1 + R_2)$;

2. $\Delta(R_1, R_1 + R_2)$ is jointly continuous in $(R_1, R_1 + R_2)$, except possibly at $R_1 = R_2 = 0$;

3. $\Delta(R_1, R_1 + R_2)$ is nonincreasing in $R_1, R_2 \geq 0$.

If we are allowed a maximum of R bits per source symbol to represent the source signal and choose to encode it at two levels — the first at rate $R_1 \leq R$ bits per source symbol and the second at $\delta R = R - R_1$ bits per source symbol — then, the optimal expected distortion in reconstruction for the two-step procedure, $\Delta(R_1, R)$, is obviously no smaller than that achievable in a single step, $\Delta(R)$, assuming that the same distortion metric $d_1 = d_2 = d$ is used throughout. Thus,

$$\Delta(R_1, R) \geq \Delta(R), \quad 0 \leq R_1 \leq R.$$

The condition under which equality holds above was determined by Koshélev [32] and Equitz-Cover [16], and subsequently given a geometric interpretation by Rimoldi [46]. This condition requires that the source rv X and the two reproduction rv's Y_1 and Y_2 in (5.1) satisfy a Markov condition, (*c.f.*, *e.g.*, [11]).

We conclude this section by noting that the two-level distortion-rate function, $\Delta(R_1, R_1 + R_2)$, can be extended to the case of K $(K > 2)$ levels of coding, so as to quantify the minimum value of expected distortion which can be achieved by the decoder when each level k is optimally compressed with R_k, bits, $k = 1, \ldots, K-1$. It is reasonable to use this function as a measure of the quality of the received signal.

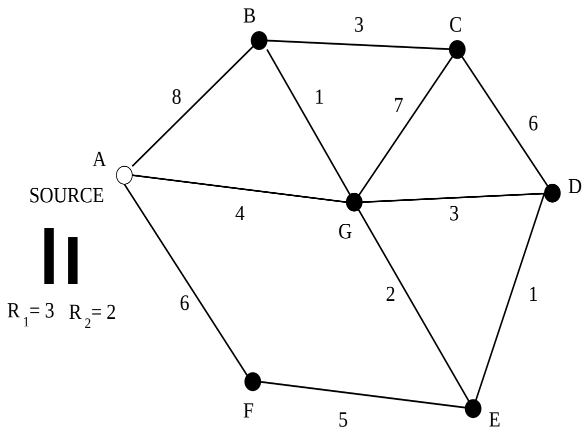

Figure 5.2: A multicasting example with a source which uses only two compression levels at rates $R_1 = 3$ and $R_2 = 2$.

5.3 The Multicasting Problem

In this section, we formulate the problem of multiresolution data transmission in a communication network. Our approach is motivated by the work of Shacham [49], which deals with the optimal selection of the multicast routing tree and the bandwidth assignment to each level assuming that multiple levels of information are sent to a user along the same route. The performance of such a scheme can, of course, be improved by assigning different routes to the various levels. This improvement in performance, which can be significant, is achieved at the cost of additional algorithmic complexity in implementation, as will be seen below.

5.3.1 Preliminaries

The topology of a multicasting problem can be described in terms of an undirected graph whose vertices represent the source and the receivers, and whose edges correspond to links with specified capacities. *We shall assume that every node in the graph (other than the source) is a receiver,* i.e., *requires the source information at some level or levels, and may have to transmit a portion or all of it to some of its neighboring nodes.* Thus, the problem of multicasting hierarchically compressed information becomes one of broadcasting. This situation is not typical in practice where, at any point in time, only a small fraction of the number of nodes in a network constitute intended receivers while the remaining nodes may serve as relays. However, our simplifying assumption above serves to provide useful insight into the workings of the multicasting algorithms proposed below. Also, there is no fundamental difficulty in extending our approach to a *bona fide* multicasting situation. Next, any collection of paths on which information is sent can be represented by a spanning tree with the source node being the tree root. Clearly, any such collection of paths must form a tree as no loops are allowed; the existence of a loop is tantamount to the same information level being simultaneously sent to a node on two different paths. Furthermore, this tree must span all the nodes, since one node is the source and all the remaining nodes are receivers.

Figure 5.2 shows an example of a graph corresponding to a simple network with 7 nodes. In this example, the source node transmits information encoded at two levels with rates $R_1 = 3$ and $R_2 = 2$ bits/source symbol, respectively. The capacity of each link is denoted by the number adjacent to it. We first consider the situation where if a node receives both levels, it is assumed that both these levels were sent to the node along the same path. If an information packet, containing both levels, encounters a link with capacity lower than $R_1 + R_2$, then as many levels are dropped as are necessary for the resulting rate not to exceed

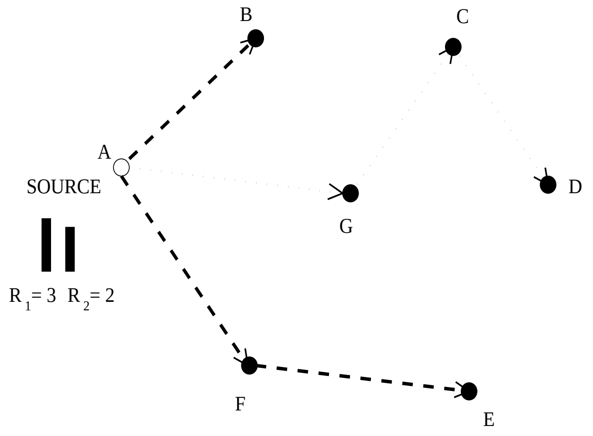

Figure 5.3: A spanning tree which delivers the two levels over the same path.

the capacity of that link. Thus, some nodes receive both levels while others receive only one level. In the present situation, it is easily seen that under the best circumstances, nodes B, E and F can receive both information levels while nodes C, D and G receive only the first level. Figure 5.3 depicts a spanning tree which achieves this purpose. This tree is shown as consisting of directed dashed and dotted links; the dashed links carry both information levels while the dotted links carry only the first level. Unused links are not displayed.

There are clear advantages in routing both levels to a receiver over the same path, such as reduced delay jitter at the receiver and simplification of session management. It is obvious, however, that this scheme is restrictive in that it does not provide the best utilization of network link capacities as not all nodes receive both levels of information. On the other hand, by allowing the two levels of information to be sent on different paths so as to take advantage of the unused capacity, all nodes in Figure 5.2 can receive both levels. To see this, observe that Figure 5.4 shows the network of Figure 5.2 with remaining (available) link

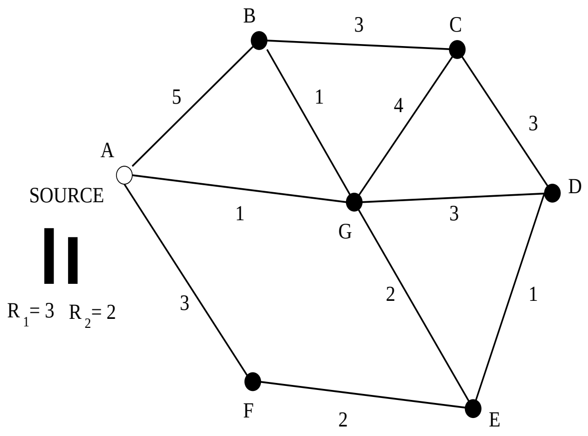

Figure 5.4: The network of Figure 5.2 with remaining (available) link capacities after the first level has been routed according to the scheme proposed in Figure 5.3.

capacities after the first level has been routed according to the scheme proposed in Figure 5.3. Then, Figure 5.5 shows a spanning tree for routing the second information level to all the receiver nodes; this level is routed along the directed dashed links. Unused links are not shown in Figure 5.5.

Our objective is to simultaneously identify the compression rates and routing trees for each level which lead to optimal reception quality — the latter will be characterized by the multi-level distortion-rate function. We formulate below the problem of optimal rate and source selection for distributing hierarchically compressed information to multiple recipients with different requirements (or reception capabilities) under given link capacity constraints.

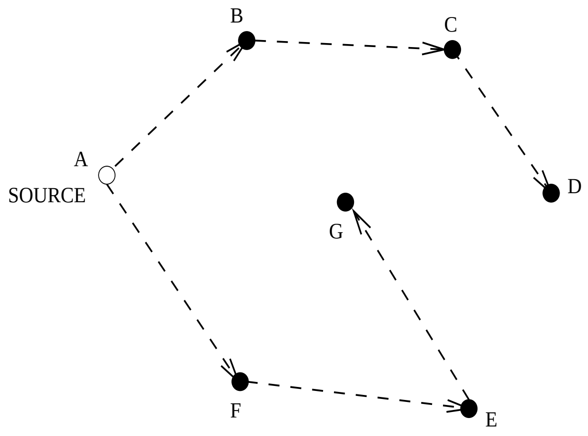

Figure 5.5: A spanning tree for routing the second information level.

5.3.2 Problem formulation

Let $(\mathcal{V}, \mathcal{E})$ be an undirected connected graph with set of nodes $\mathcal{V}$, set of edges $\mathcal{E} = \mathcal{V} \times \mathcal{V}$ and link capacities $c(\mathbf{e}), \mathbf{e} \in \mathcal{E}$, where $c : \mathcal{E} \mapsto \mathbf{R}^{+}$; if link $(i, j),\ i, j \in \mathcal{V}$, does not exist, we simply set $c\left((i, j)\right) = 0$. We assume that $\mathbf{s} \in \mathcal{V}$ is the source node and that all the other nodes $\mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}$ are receiver nodes. The source information can be represented at $K > 1$ levels of rates $R_1, \ldots, R_K$ bits/source symbol, where $R_1, \ldots, R_K$, are to be suitably chosen so as to satisfy certain criteria. The routing for information level $k,\ 1 \leq k \leq K$, is specified by a (directed) routing tree $(\mathcal{V}_k, \mathcal{E}_k)$. Note that $\mathcal{V}_k$ is the set of the nodes which receive and use information level k. There may exist nodes which receive level k only to retransmit it, since they did not receive a lower level. We, therefore, require that $\mathcal{V}_k \subset \mathcal{V}_{k-1}$ in order to ensure continuity of the received information. However, $\mathcal{E}_k$ and $\mathcal{E}_\ell$ $(k \neq \ell)$ need not possess any containment property. Also, $(\mathcal{V}_k, \mathcal{E}_k)$ is a connected subgraph of $(\mathcal{V}, \mathcal{E})$ and is not allowed any loops, so that there exists only one path from $\mathbf{s}$ to every $\mathbf{v} \in \mathcal{V}_k$.

In view of the nature of our optimality criteria, which will be specified below, we shall require without any essential loss of generality that any routing policy of interest must have the following attributes.

(i) Each receiver node receives at least some information; therefore, the first level must reach all receivers. To this end, we impose the condition that $\mathcal{V}_1 = \mathcal{V}$, and

$$R_1 \leq \min_{\mathbf{e} \in \mathcal{E}_1} c(\mathbf{e}),$$

where $\mathcal{E}_1$ is to be suitably chosen (as described in Sections 5.4 and 5.5 below).

(ii) For each of the next levels $k = 2, \ldots, K$, we start with $\mathcal{V}_k = \mathcal{V}_{k-1}$ and a suitable spanning tree $\mathcal{E}_k$ which spans the nodes of $\mathcal{V}_k$. We then choose R_k (in accordance with the criteria described below) and check each link to see if the capacity constraint is satisfied. If the capacity constraint is violated for link $\mathbf{e} \in \mathcal{E}_k$, *i.e.*,

$$\sum_{k : \mathbf{e} \in \mathcal{E}_k} R_k > c(\mathbf{e}),$$

then the tree $\mathcal{E}_k$ is *pruned* at branch $\mathbf{e}$, *i.e.*, branch $\mathbf{e}$ and all the "subsequent" branches are removed from $\mathcal{E}_k$ and the corresponding nodes are removed from $\mathcal{V}_k$.

(iii) The previous step is repeated for the next value of k until $k = K$ is reached.

Observe that for *fixed* routing $(\mathcal{V}_k, \mathcal{E}_k)$, $k = 1, \ldots, K$, the number of information levels received by each node is a function of the level rates $R_1, \ldots, R_K$. Each node $\mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}$ receives $l_{\mathbf{v}}$ levels of information, with the resulting distortion at that node being equal to

$$D_{\mathbf{v}} = \Delta(R_1, \ldots, R_1 + \ldots + R_{l_{\mathbf{v}}}). \tag{5.2}$$

We shall study the following two problems :

(I) The maximum distortion problem : Subject to the requirements and constraints mentioned above, choose $R_1, \ldots, R_K$ so as to *minimize* the maximun distortion (over all receivers)

$$\max_{\mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}} D_{\mathbf{v}}. \tag{5.3}$$

(II) The average distortion problem : Subject to the requirements and constraints mentioned above, choose $R_1, \ldots, R_K$ so as to *minimize* the average distortion (over all receivers)

$$\frac{1}{|\mathcal{V}| - 1} \sum_{\mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}} D_{\mathbf{v}}. \tag{5.4}$$

Remark : The minima in problems **(I)** and **(II)** exist, since the rate and route selection in both problems involves searching among a finite number of candidates for optimality, as can be seen below in Sections 5.4 and 5.5, respectively.

Heretofore, we have assumed that the routing trees are *fixed* at each level. Clearly, it is of interest to relax this assumption, and determine the routing trees which minimize the maximum and average distortions (*i.e.*, distortions) among all routing trees; the resulting problems would now involve joint optimization with respect to the level rates and the routing trees. An obvious solution to the difficult problem of finding the optimal routing trees, namely an exhaustive search, is prohibitively complex. It is, therefore, of interest to devise "good" algorithms based on a set of well-chosen heuristics which yield solutions that are close to the optimal. These issues are addressed below in Section 5.5.

5.4 The maximum cost problem

We now examine the maximum distortion case in which, loosely speaking, fairness is ensured for all the receivers by maximizing the reception quality of the node receiving the least amount of information. In other words, we minimize the distortion for the receiver node which incurs the maximum value of the distortion-rate function.

Observe that the resulting maximun value of distortion is determined by those nodes which receive the fewest levels of information. Sending additional information levels to the other nodes does not offer any improvement in the value of maximun distortion, and only makes inefficient use of the available bandwidth. Hence, the approach results in all nodes receiving the same number of information levels. Bearing this in mind, we can reformulate as follows the maximum distortion problem for the set of nodes $\mathcal{V} = \{1, \ldots, n\}$ (with node 1 being the source) and the ($n \times n-$) capacity matrix $c : \mathcal{V} \times \mathcal{V} \mapsto \mathbf{R}^+$, for a source with K levels of information.

Pick K sets $\mathcal{E}_k \subset \mathcal{V} \times \mathcal{V}$, with $|\mathcal{E}_k| = n - 1$, and rates $R_k,\ k = 1, \ldots, K$, so as to minimize

$$\Delta(R_1, R_1 + R_2, \ldots, R_1 + \ldots + R_K) \tag{5.5}$$

subject to the following constraints :

1. (*Capacity constraint*) : For each link $\mathbf{e}$ of the graph,

$$\sum_{k:\mathbf{e}\in\mathcal{E}_k} R_k \leq c_{\mathbf{e}}, \quad \mathbf{e} \in \mathcal{V} \times \mathcal{V};$$

2. For each level $k \in \{1, \ldots, K\}$:

(a) (*Spanning constraint*) : For each receiver node $\mathbf{v} \in \mathcal{V}$, there exists another receiver node $\mathbf{v}' \in \mathcal{V}$, such that $(\mathbf{v}, \mathbf{v}') \in \mathcal{E}_k$;

(b) (*Connectivity constraint*) : For each pair of distinct nodes $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$, there exist nodes $\mathbf{v}_1, \ldots, \mathbf{v}_p$, such that $(\mathbf{v}, \mathbf{v}_1), (\mathbf{v}_1, \mathbf{v}_2), \ldots, (\mathbf{v}_p, \mathbf{v}') \in \mathcal{E}_k$, for some integer p, $1 \leq p \leq n-1$.

Remark : Constraint 2 ensures that each $\mathcal{E}_k$, $k = 1, \ldots, K$, is a connected spanning tree of the network $\mathcal{V} \times \mathcal{V}$; here, we assume that links $(\mathbf{v}, \mathbf{v}')$ and $(\mathbf{v}', \mathbf{v})$ are equivalent.

We determine below the spanning trees and rates, $\{(\mathcal{E}_k, R_k), k = 1, \ldots, K\}$ which yield the maximum distortion (*cf.*, (5.3)), subject to the constraints 1 and 2 above. It will be useful to introduce the following

Definition 5.1 *The* minimum capacity *of a set of links $\mathcal{E}' \subset \mathcal{E}$ is defined as the smallest capacity among all the links in $\mathcal{E}'$, and is denoted by $c(\mathcal{E}')$;* i.e.,

$$c(\mathcal{E}') \triangleq \min_{\mathbf{e} \in \mathcal{E}'} c(\mathbf{e}).$$

For ease of exposition, we shall address in detail the case $K = 2$; the results so obtained can be easily generalized to $K \geq 2$ levels of compression. We now focus on the case $K = 2$ and determine the optimal $\{(\mathcal{E}_1, R_1), (\mathcal{E}_2, R_2)\}$. Let $\mathcal{E}_{12} = \mathcal{E}_1 \bigcap \mathcal{E}_2$. Then, constraint 1 is given by :

$$\begin{aligned} R_1 &\leq \min\nolimits_{\mathbf{e} \in \mathcal{E}_1 \setminus \mathcal{E}_{12}} c(\mathbf{e}) &= c(\mathcal{E}_1 \setminus \mathcal{E}_{12}) \\ R_2 &\leq \min\nolimits_{\mathbf{e} \in \mathcal{E}_2 \setminus \mathcal{E}_{12}} c(\mathbf{e}) &= c(\mathcal{E}_2 \setminus \mathcal{E}_{12}) \\ R_1 + R_2 &\leq \min\nolimits_{\mathbf{e} \in \mathcal{E}_{12}} c(\mathbf{e}) &= c(\mathcal{E}_{12}), \end{aligned} \tag{5.6}$$

with the convention $c(\emptyset) = \infty$, where $\emptyset$ denotes the empty set.

The maximum solution $\{(\mathcal{E}_1, R_1), (\mathcal{E}_2, R_2)\}$, is characterized below in Proposition 5.2. The earlier Proposition 5.1 paves the way by identifying a useful property of the maximum solution.

Proposition 5.1 *For fixed spanning trees $\mathcal{E}_1, \mathcal{E}_2$, the optimum rates R_1 and R_2 can be determined by the following steps :*

(i) *pick R_1 to equal the minimum capacity of the tree $\mathcal{E}_1$;*

(ii) *subtract R_1 from the capacity of every link in $\mathcal{E}_1$ to obtain a set of reduced (available) capacity values of the links in $\mathcal{E}_1$.*

(iii) *pick R_2 to equal the minimum capacity of $\mathcal{E}_2$ with the understanding that a link in $\mathcal{E}_2$, which is also in $\mathcal{E}_1$, will possess a capacity equal to the reduced value from step (ii).*

Proof: We can write the minimum in (5.3) as :

$$\min_{\substack{R_1 \le c(\mathcal{E}_1 \setminus \mathcal{E}_{12}), R_2 \le c(\mathcal{E}_2 \setminus \mathcal{E}_{12}) \\ R_1 + R_2 \le c(\mathcal{E}_{12})}} \Delta(R_1, R_1 + R_2) \tag{5.7}$$

$$\begin{aligned}
&= \min_{R_1 \le \min\{c(\mathcal{E}_1 \setminus \mathcal{E}_{12}), c(\mathcal{E}_{12})\}} \min_{R_2 \le \min\{c(\mathcal{E}_2 \setminus \mathcal{E}_{12}), c(\mathcal{E}_{12}) - R_1\}} \Delta(R_1, R_1 + R_2) \\
&= \min_{R_1 \le \min\{c(\mathcal{E}_1 \setminus \mathcal{E}_{12} \cup \mathcal{E}_{12})\}} \Delta(R_1, R_1 + \min\{c(\mathcal{E}_2 \setminus \mathcal{E}_{12}), c(\mathcal{E}_{12}) - R_1\}) \\
&= \min_{R_1 \le \min\{c(\mathcal{E}_1\}} \Delta(R_1, \min\{R_1 + c(\mathcal{E}_2 \setminus \mathcal{E}_{12}), c(\mathcal{E}_{12})\}) \\
&= \Delta(\min\{c(\mathcal{E}_1\}, \min\{c(\mathcal{E}_1) + c(\mathcal{E}_2 \setminus \mathcal{E}_{12}), c(\mathcal{E}_{12})\}),
\end{aligned} \tag{5.8}$$

where the second and the last inequalities follow from the fact that $\Delta(\cdot, \cdot)$ is nonincreasing. The form of the arguments of $\Delta(\cdot, \cdot)$ in the expression for the minimum distortion in (5.8), immediately leads to the assertion of the Proposition. ■

Remarks :

(i) For fixed spanning trees $\mathcal{E}_1, \mathcal{E}_2$, it is clear from (5.8) above that the optimum rates R_1 and R_2 are given by

$$R_1 = c(\mathcal{E}_1), \quad R_2 = \min\{c(\mathcal{E}_2 \setminus \mathcal{E}_{12}), c(\mathcal{E}_{12}) - c(\mathcal{E}_1)\}. \tag{5.9}$$

(ii) For the general case of $K \geq 2$ levels of compression, if the spanning trees $\mathcal{E}_1, \ldots, \mathcal{E}_K$ are fixed, it is easy to see that the optimum level rates $R_1, \ldots, R_K$, can be obtained by an obvious extension of the procedure described in Proposition 5.1.

Proposition 5.2 *The optimum spanning trees $\mathcal{E}_1, \mathcal{E}_2$, and the optimum rates R_1, R_2, can be determined by the following steps :*

(i) *pick $\mathcal{E}_1$ to be the spanning tree of the original graph, which possesses the largest minimum capacity,* i.e., *the set of $n-1$ links with the highest capacity values which also form a tree;*

(ii) *set R_1 to equal the minimum capacity of $\mathcal{E}_1$ from step (i);*

(iii) *subtract R_1 from the capacity of every link in $\mathcal{E}_1$ to obtain a set of reduced capacity values of the links in $\mathcal{E}_1$;*

(iv) *pick $\mathcal{E}_2$ to be the spanning tree of the original graph which possesses the largest minimum capacity, with the understanding that a link in $\mathcal{E}_2$, which is also in $\mathcal{E}_1$, will possess a capacity equal to the reduced value from the previous step;*

(v) *set R_2 to equal the corresponding minimum capacity of $\mathcal{E}_2$;*

Proof: Consider the minimum in (5.3) (for the case $K = 2$) subject to constraints 1 and 2. From Proposition 5.1, by (5.7) and (5.8), this equals

$$\min_{\mathcal{E}_1, \mathcal{E}_2} \Delta \left(c(\mathcal{E}_1), c(\mathcal{E}_1) + \min \left\{ c(\mathcal{E}_2 \setminus \mathcal{E}_{12}), c(\mathcal{E}_{12}) - c(\mathcal{E}_1) \right\} \right), \tag{5.10}$$

where the minimum is with respect to all spanning trees which satisfy constraint 2. Consider the expression in (5.10) as a minimum with respect to $\mathcal{E}_2$ first (for a fixed $\mathcal{E}_1$), followed as a minimum with respect to $\mathcal{E}_1$. By the fact that $\Delta(\cdot, \cdot)$ is nonincreasing and from the form of its arguments in (5.10), it is clear that for a fixed $\mathcal{E}_1$, the minimizing $\mathcal{E}_2$ is given by steps (iii) and (iv) of the Proposition; then, the minimizing $\mathcal{E}_1$ is given by step (i) of the Proposition. Furthermore, it is evident from the form of the arguments of $\Delta(\cdot, \cdot)$ in (5.10) that the optimum values of R_1 and R_2 are as given by steps (ii) and (v) of the Proposition. Finally, the assertion of step (vi) is straightforward. This completes the proof of the Proposition. ■

Remark : For the general case of $K \geq 2$ levels of compression, the optimum spanning trees and level rates, $\{(\mathcal{E}_k, R_k) : k = 1, \ldots, K\}$, can be obtained by an obvious extension of the procedure described in Proposition 5.2 is applied to each candidate (spanning tree) until level K is reached; at this point, a comparison of the corresponding values of distortion is used to resolve a tie at the earlier level k.

Several algorithms exist for the systematic determination of a spanning tree with largest minimum capacity. One such algorithm can be found in [49]. We shall refer to it as the Maximum Capacity Tree (MCT) algorithm. It operates by calculating the path from each receiver to the source, which has the largest minimum capacity (hereafter referred to as the *Maximum Capacity Path* (MCP)), using a Dijkstra-like algorithm (*c.f.*, *e.g*, [8]); these MCP's from each receiver node to the source node are subsequently merged to form the MCT. We conclude this section by presenting our MCT Algorithm for determining the spanning tree with the largest minimum capacity. This algorithm enables an algorithmic implementation of the maximum solution in Proposition 5.2. The main component of our MCT Algorithm is the determination of the MCP between the source node and each receiver node by a Bellman-Ford variant (*c.f.*, *e.g.*, [8]) of the algorithm given in [49].

Algorithm 5.1 (The MCT Algorithm)

Let $C_{\mathbf{v}}$ denote the minimum capacity of the current path from the source node $\mathbf{s}$ to the receiver node $\mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}$, i.e., *the smallest capacity among all links on this path.*

1. *Initially, for each $\mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}$, let the path to the source $\mathbf{s}$ be that which consists of a single link $(\mathbf{s}, \mathbf{v})$, and let*

$$C_{\mathbf{v}} = c\left((\mathbf{s}, \mathbf{v})\right), \quad \mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}.$$

2. *Each node $\mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}$ looks at its neighbouring nodes $\mathbf{v}' \in \mathcal{V} \setminus \{\mathbf{s}, \mathbf{v}\}$ (*i.e., *those neighbouring nodes with which it is connected by a single link) to determine if there is an alternative path (from $\mathbf{s}$ to $\mathbf{v}'$ and then to $\mathbf{v}$) with minimum capacity exceeding $C_{\mathbf{v}}$. In case of a tie between the minimum capacities $C_{\mathbf{v}_1}$ and $C_{\mathbf{v}_2}$ of two neighbour nodes $\mathbf{v}_1, \mathbf{v}_2$ of $\mathbf{v}$ respectively, the second highest capacities of the paths leading to $\mathbf{v}_1$ and $\mathbf{v}_2$ are compared and so on, if necessary. Thus, it is necessary to store at each node the (sorted) list of the capacities of all links of the current MCP. The identity of the node $\mathbf{v}'$ which affords the maximum value of the minimum capacity*

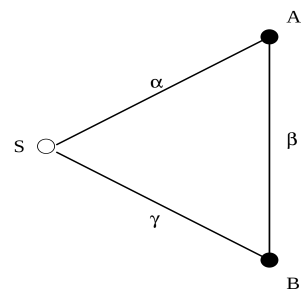

Figure 5.6: An example of simple network with only two receiving nodes.

is stored; this $\mathbf{v}'$ *is referred to as the parent node of* $\mathbf{v}$ *in the MCT. Each* $C_{\mathbf{v}}$ *is updated according to*

$$C_{\mathbf{v}} = \max_{\mathbf{v}' \in \mathcal{V} \setminus \{\mathbf{s}\}} \left\{ C_{\mathbf{v}}, \min \left\{ C_{\mathbf{v}' \in \mathcal{V}}, c\left((\mathbf{v}', \mathbf{v}) \right) \right\} \right\}.$$

3. *Step 2 is repeated until none of the* $C_{\mathbf{v}}$*'s changes.*

4. *Merge the MCP's thus obtained to form the MCT.*

Remark : The previous algorithm converges in at most n iterations. It is simple to implement, and works in a distributed and adaptive manner.

Finally, we observe that the maximum solution in Proposition 5.2 together with its implementation by means of Algorithm 5.1, can be applied to any distortion measure $\Delta(\cdot, \ldots, \cdot)$ as in (5.5), which is nonincreasing with respect to $R_1, R_2, \ldots, R_K$. Furthermore, the approach of this section can be easily extended to a network where not all the nodes are receivers.

5.5 The Average Cost Problem

The problem of determining the routing trees and level rates which minimize the average distortion (5.4) is significantly more difficult than that involving the maximum distortion treated in the previous section. To our knowledge, there is no efficient systematic solution for this problem; the only means of solution seems to entail an exhaustive search, whose complexity grows exponentially with the number of nodes in the network.

The complexity of the average distortion problem is already evident in the simple situation considered below in Example 1.

Example 1

Consider a simple network consisting of a source node **s** and two receiver nodes **A** and **B**, as illustrated in Figure 5.6. The capacities of links $(\mathbf{s}, \mathbf{A})$, $(\mathbf{A}, \mathbf{B})$ and $(\mathbf{s}, \mathbf{B})$ are α, β and γ, respectively, where $\alpha \geq \gamma$. Let us initially assume that $\gamma = 0$, so that the network, in effect, has only two links ($(\mathbf{s}, \mathbf{A})$ and $(\mathbf{A}, \mathbf{B})$) in tandem. Further, assume that the source node **s** produces only one information level. Even for this simplest possible situation, finding a solution for the average distortion problem can entail an exhaustive search as is seen below. If $\alpha < \beta$, then the source node must send information at rate $R_1 \leq \alpha$ in order for nodes **A** and **B** to receive the first level. The optimal choice is

$$R_1 = \alpha,$$

in which case, both nodes receive information at rate α; the resulting average distortion is

$$d_{ave} = \frac{1}{2}\Delta(\alpha) + \frac{1}{2}\Delta(\alpha) = \Delta(\alpha).$$

Next, if $\alpha > \beta$, the situation becomes interesting. Then the source node has two options : either send the first level to both receivers at rate $R_1 \leq \beta$ or send it only to node A at rate $R_1 \in (\beta, \alpha]$. Correspondingly, the optimal selection for the level rate is either $R_1 = \beta$, when the average distortion is

$$d_{ave} = \frac{1}{2}\Delta(\beta) + \frac{1}{2}\Delta(\beta) = \Delta(\beta),$$

or $R_1 = \alpha$, when the average distortion is

$$d_{ave} = \frac{1}{2}\Delta(\alpha) + \frac{1}{2}\Delta(0).$$

Depending on the specific form of the distortion-rate function, and the actual values of α and β, either of the average distortions above can be smaller than the other. (The concavity of $\Delta(\cdot)$ does not suffice to resolve this issue.) Thus, both average distortions above must be explicitly calculated in order to determine the optimal solution. In general, determining the optimal rate will, thus, involve an exhaustive comparison of a prohibitively large number of potential candidates, typically polynomial in the number of nodes. ■

Since even the selection of optimal level rates for a simple network is not a straightforward problem, it is not surprising that the selection of optimal routing trees is an even more complex task. The following example illustrates this fact.

Example 2

Consider the previous example with the difference that we now assume that $\alpha \leq \gamma$. By a simple but tedious procedure, it turns out that in order to determine the optimal routing tree at each level, we need to consider six possible orderings of the three link capacities and the sum of the two smallest ones. This leads to nine possibilities for the (routing) trees $\mathcal{E}_1$ and $\mathcal{E}_2$, each of which has been examined. We obtain the optimum level rates for each possible routing scheme based on the results of Section 5.3, and calculate the corresponding optimum distortion. After eliminating those routing schemes which are obviously not candidates for optimality (on the basis of larger distortions), there still remain routing schemes different from the MCT scheme in five out of the six cases, which for some values of the link capacities are optimal. Obviously, these are pruned versions of possible spanning tree pairs, where the two nodes do not receive both levels; otherwise, if both nodes receive both levels, the problem becomes equivalent to the maximum cost problem, in which case the optimal routing scheme is given in terms of MCT's. These alternative schemes and the corresponding aggregate distortions are presented in Table 5.1, along with the MCT scheme for each case and its associated aggregate distortion.

It is obvious from Table 5.1 that even for this simple example, the optimal routing trees need not be MCT's. Even for the first information level, it may transpire that an MCT is not an optimal choice; consider, for instance, the case $\beta \leq \alpha \leq \gamma \leq \alpha + \beta$ in Table 5.1. ■

We can, however, make the following simple observation concerning the structure of the optimal solution. If $\mathcal{E}_k,\ k = 1, \ldots, K$, is a fixed collection of K spanning trees, then the level rates $R_k,\ k = 1, \ldots, K$, which lead to the minimum average distortion are such that each level rate saturates at least one link in the network. In other words, the rate of each level should be equal to the remaining (residual) capacity of one of the links of the routing tree for that level. In particular, let $\mathcal{C}_1$ be the (ordered) set of distinct capacities of the paths from the source node to the receiver nodes in $\mathcal{E}_1$, *i.e.*,

$$\mathcal{C}_1 = \left\{ C_{\mathbf{v}}^{(1)}, \mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\} \right\},$$

<table>
<tr><th>MCT's $\mathcal{E}_1, \mathcal{E}_2$</th><th>$D_{\mathbf{A}} + D_{\mathbf{B}}$</th><th>alternative $\mathcal{E}_1, \mathcal{E}_2$</th><th>$D_{\mathbf{A}} + D_{\mathbf{B}}$</th></tr>
<tr><td colspan="4">Case : $\alpha \leq \beta \leq \alpha + \beta \leq \gamma$</td></tr>
<tr><td rowspan="2">$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{A}, \mathbf{B})$
$\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$</td><td rowspan="2">$2\Delta(\beta, \alpha{+}\beta)$</td><td>$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$
$\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B})$</td><td>$\Delta(\beta, \gamma)$
$+\Delta(\beta, \beta)$</td></tr>
<tr><td>$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{A}, \mathbf{B})$
$\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B})$</td><td>$\Delta(\alpha, \alpha{+}\gamma)$
$+\Delta(\alpha, \alpha)$</td></tr>
<tr><td colspan="4">Case : $\alpha \leq \beta \leq \gamma \leq \alpha + \beta$</td></tr>
<tr><td rowspan="2">$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{A}, \mathbf{B})$
$\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$</td><td rowspan="2">$2\Delta(\beta, \gamma)$</td><td>$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$
$\mathcal{E}_2$ $(\mathbf{s}, \mathbf{A})$</td><td>$\Delta(\beta, \alpha{+}\beta)$
$+\Delta(\beta, \beta)$</td></tr>
<tr><td>$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{A}, \mathbf{B})$
$\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B})$</td><td>$\Delta(\alpha, \alpha + \gamma)$
$+\Delta(\alpha, \alpha)$</td></tr>
<tr><td colspan="4">Case : $\alpha \leq \gamma \leq \alpha + \gamma \leq \beta$</td></tr>
<tr><td>$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$
$\mathcal{E}_2$ $(\mathbf{s}, \mathbf{A}), (\mathbf{A}, \mathbf{B})$</td><td>$2\Delta(\gamma, \alpha{+}\gamma)$</td><td></td><td></td></tr>
</table>

Table 5.1: MCT's and alternative routing schemes and the corresponding distortions, for various values of the link capacities in Example 2.

where $C_{\mathbf{v}}^{(1)}$ is the capacity of the path in $\mathcal{E}_1$ from the source node $\mathbf{s}$ to the receiver node $\mathbf{v}$, for the original graph $(\mathcal{V}, \mathcal{E})$ with given capacities. By definition, the elements of $\mathcal{C}_1$ satisfy

$$c_{11} < c_{12} < \ldots < c_{1|\mathcal{C}_1|}.$$

The significance of the set $\mathcal{C}_1$ lies in the fact that if the first level rate R_1 is in the interval $(c_{1i-1}, c_{1i}]$ for some $i \in \{1, \ldots, |\mathcal{C}_1|\}$, then the first level will reach all receiver nodes $\mathbf{v}$ which have $C_{\mathbf{v}}^{(1)} \geq c_{1i}$, whereas the nodes with $C_{\mathbf{v}}^{(1)} \leq c_{1i}$ will

MCT's $\mathcal{E}_1, \mathcal{E}_2$	$D_{\mathbf{A}} + D_{\mathbf{B}}$	**alternative** $\mathcal{E}_1, \mathcal{E}_2$	$D_{\mathbf{A}} + D_{\mathbf{B}}$
Case : $\alpha \le \gamma \le \beta \le \alpha + \gamma$			
$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$ $\mathcal{E}_2$ $(\mathbf{s}, \mathbf{A}), (\mathbf{A}, \mathbf{B})$	$2\Delta(\gamma, \beta)$	$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$ $\mathcal{E}_2$ $(\mathbf{s}, \mathbf{A})$	$\Delta(\gamma, \alpha+\gamma)$ $+\Delta(\gamma, \gamma)$
Case : $\beta \le \alpha \le \alpha + \beta \le \gamma$			
$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{s}, \mathbf{B})$ $\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$	$2\Delta(\alpha, \alpha+\beta)$	$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{s}, \mathbf{B})$ $\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B})$	$\Delta(\alpha, \gamma)$ $+\Delta(\alpha, \alpha)$
		$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{A}, \mathbf{B})$ $\mathcal{E}_2$ $(\mathbf{s}, \mathbf{A})$	$\Delta(\beta, \beta+\gamma)$ $+\Delta(\beta, \beta)$
Case : $\beta \le \alpha \le \gamma \le \alpha + \beta$			
$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{s}, \mathbf{B})$ $\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B}), (\mathbf{B}, \mathbf{A})$	$2\Delta(\alpha, \gamma)$	$\mathcal{E}_1$ $(\mathbf{s}, \mathbf{A}), (\mathbf{A}, \mathbf{B})$ $\mathcal{E}_2$ $(\mathbf{s}, \mathbf{B})$	$\Delta(\beta, \beta+\gamma)$ $+\Delta(\beta, \beta)$

Table 5.1: (*continued*) Three more cases for the simple 2-node network problem.

not receive the first information level. Thus, the number of nodes receiving the first level does not depend on the exact value of R_1 but rather on the interval to which it belongs as defined by the elements of $\mathcal{C}_1$. Therefore, if we are given that $R_1 \in (c_{1i-1}, c_{1i}]$, then the monotonicity property of $\Delta(\cdot, \ldots, \cdot)$ requires us to pick $R_1 = c_{1i}$ in order to maximize the quality of the received signal at each receiver node.

Subsequently, we define successively the sets $\mathcal{C}_k, \quad k = 2, \ldots, K$, in the same manner, with the difference that for $k \geq 2$, the elements of $\mathcal{C}_k$ correspond to capacity constraints for the aggregate rate $R_1 + \ldots + R_k$ of the kth level. In other words, the role of R_1 in the previous step is now played by the aggregate

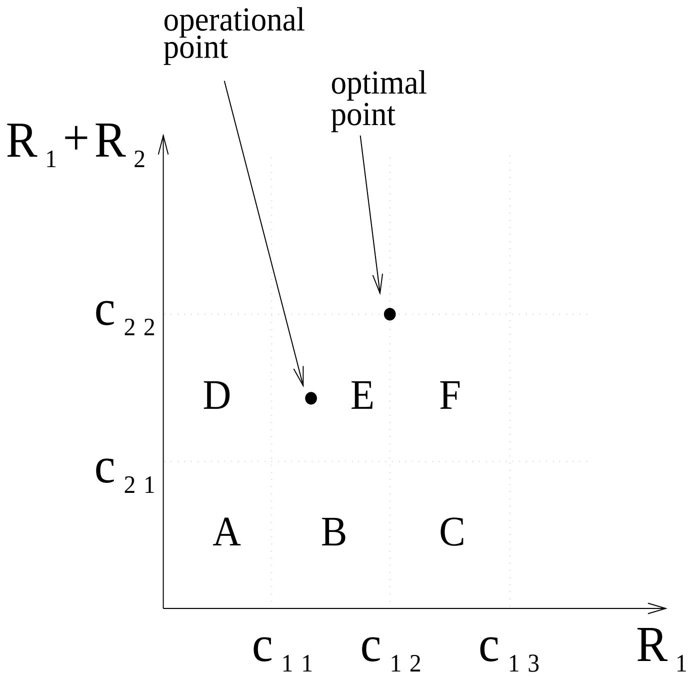

Figure 5.7: Position of the rate vector $(R_1, R_1 + R_2)$ for $K = 2$ information levels.

rate $R_1 + \ldots + R_k, k = 2, \ldots, K$. In particular, for each $c' \in \mathcal{C}_1$, let $C^{(2)}_{\mathbf{v}}(c')$ be the the capacity of the path in $\mathcal{E}_2$ from the source node $\mathbf{s}$ to the receiver node $\mathbf{v}$, where the used capacity c' is subtracted from each link of $\mathcal{E}_2$ which is also an active link of $\mathcal{E}_1$, *i.e.*, carries the first information level. Then, $\mathcal{C}_2$ is defined as

$$\mathcal{C}_2 = \left\{ c' + C^{(2)}_{\mathbf{v}}(c'), \mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}, c' \in \mathcal{C}_1 \right\},$$

and $\mathcal{C}_k, \quad k = 3, \ldots, K$ are defined similarly. This definition of the $\mathcal{C}_k$'s is motivated by the fact that the first level has higher priority over the other levels. Hence, it can be sent on any link of $\mathcal{E}_1$ with capacity no smaller than R_1; similarly, the second level can be sent only on the links of $\mathcal{E}_1 \bigcap \mathcal{E}_2$ with capacity at least equal to $R_1 + R_2$, and those of $\mathcal{E}_2 \setminus \mathcal{E}_1$ with capacity at least equal to R_2, and so on.

The elements of the $\mathcal{C}_k$'s partition the $\mathbf{R}_+^K$ space into $\prod_{k=1}^K (|\mathcal{C}_k| + 1)$ regions, out of which we have to consider only the $\prod_{k=1}^K |\mathcal{C}_k| < N^K$ K-dimensional "rectangles", as shown in Figure 5.7, for the simple case of $K = 2$ levels of information. In the situation illustrated in Figure 5.7, we have assumed that the sets $\mathcal{C}_1$ and $\mathcal{C}_2$ consist of 3 elements and 2 elements, respectively. Thus, there are six cases to be considered, namely those corresponding to the point $(R_1, R_1 + R_2)$ belonging to each of the rectangles marked A, B, C, D, E and F. For each of these rectangles, the upper right corner is the point yielding minimum distortion, by virtue of the monotonicity of the distortion-rate function. The outer regions are not of interest because if $(R_1, R_1 + R_2)$ falls within such a region, at least one of the levels will not reach any node.

Proposition 5.3 *Assume that the K-tuple $(R_1, R_1+R_2, \ldots, R_1+\ldots+R_K)$ lies in one of the rectangles in $\mathcal{C}_1 \times \ldots \times \mathcal{C}_K$,* i.e.,

$$\sum_{k'=1}^{k} R_{k'} \in (c_{ki_k-1}, c_{ki_k}], \quad k = 1, \ldots, K. \tag{5.11}$$

Then, the minimum distortion occurs at the vertex

$$(c_{1,i_1}, \ldots, c_{K,i_K}) \in \mathcal{C}_1 \times \ldots \times \mathcal{C}_K.$$

Proof: The assertion is a simple consequence of the fact that $\Delta(\cdot, \ldots, \cdot)$ is monotonically decreasing.

All the points in each rectangle (see, for example, Figure 5.7) result in the same number of levels being received by each node, so that the number of levels which each node receives is determined by (5.11) and the capacity and continuity restrictions (constraints 1 and 2 in Section 5.4, respectively). Thus, node $\mathbf{v} \in$

$\mathcal{V} \setminus \{\mathbf{s}\}$ receives level k (at rate R_k) if

$$C_{\mathbf{v}}^{(k)}(R_1, \ldots, R_{k-1}) + \sum_{k'=1}^{k-1} R_{k'} \geq c_{ki_k}$$

and it has received levels $1, \ldots, k-1$. Let $l_{\mathbf{v}}$ be the number of levels received by node $\mathbf{v}$. Also, let n_k be the number of nodes that receive (exactly) k levels, *i.e.*,

$$n_k = \left|\{\mathbf{v} \in \mathcal{V} : l_{\mathbf{v}} = k\}\right|, \quad k = 1, \ldots, K.$$

Then, the resulting average distortion will be

$$\begin{aligned} &d_{ave}(R_1, \ldots, R_K) = \\ &\quad \frac{1}{|\mathcal{V}| - 1} \sum_{k=1}^{K} n_k \cdot \Delta(R_1, R_1 + R_2, \ldots, \sum_{k'=1}^{k-1} R_{k'}, \sum_{k'=1}^{k} R_{k'}, \ldots, \sum_{k'=1}^{k} R_{k'}). \end{aligned}$$

It is easy to see that d_{ave} above inherits the monotonicity property of the (K-level) distortion-rate function $\Delta(\cdot, \ldots, \cdot)$; hence, its minimum occurs at the maximum values of the allowable rates, namely, for

$$\sum_{k'=1}^{k} R_{k'} = c_{ki_k}, \quad k = 1, \ldots, K.$$

■

Although the previous claim establishes that we only have to look at the vertices of these "rectangles," instead of all of $\mathbf{R}_+^K$, there is no known economical method for determining which of these vertices is the minimizing one, as has been illustrated by the first example. One sure method, of course, is to search through all the combinations of elements of the $\mathcal{C}_k$'s for each fixed routing scheme. In fact, it is not necessary, to search through *all* combinations. For instance, for each $c' \in \mathcal{C}_1$ we only have to consider the elements of the set $\left\{c' + C_{\mathbf{v}}^{(2)}(c'), \mathbf{v} \in \mathcal{V} \setminus \{\mathbf{s}\}\right\} \subset \mathcal{C}_2$.

Remark : Note that we can restrict ourselves to spanning trees alone, instead of "pruned" trees. Indeed, for any collection of (not necessarily spanning) trees $\mathcal{E}'_k,\ k = 1, \ldots, K$, there exists a collection of spanning trees $\mathcal{E}_k,\ k = 1, \ldots, K$, with average distortion not exceeding that of the former. This is easily seen, by augmenting each $\mathcal{E}'_k$ by as many links as are needed to make it a spanning tree. Then the various information levels reach the same nodes as earlier and may also reach some additional nodes not served by $\mathcal{E}'_k$. This leads to an increase in the aggregate received rate, and, consequently, a smaller average distortion.

Note that in order to find the optimal routing trees, we need to search over all collections of possible spanning-trees, which is a set of prohibitively large size (of the order $O(n^n)$). Thus, although the method of exhaustive search guarantees a solution, it is prohibitively costly in terms of running time.

On the other hand, the algorithm proposed in Section 5.4 (by restricting the search to the class of MCT's) may not always yield an optimal solution. Note, however, that in Table 5.1 most of the alternative routing trees (which are candidates for optimality) are just pruned versions of the MCT. It is, therefore, reasonable to expect that by choosing the $\mathcal{E}_k$'s to be the MCT's, the resulting distortion should not be much larger than the optimal distortion. On this basis, we propose the following heuristic suboptimal algorithm.

Algorithm 5.2 (Heuristic Algorithm)

1. *Find the MCT $\mathcal{E}_1$ for the initial graph $(\mathcal{V}, \mathcal{E})$ and derive the set $\mathcal{C}_1$ from this tree.*

2. *For every element $c_{1i} \in \mathcal{C}_1$, set $R_1 = c_{1i}$ and subtract R_1 from all the links of $\mathcal{E}_1$ which can serve level 1,* i.e., *those links with capacity no less than R_1*

and for which level 1 reaches one of their end nodes.

3. *For the resulting graph with reduced link capacity values, set the routing tree for the next level to be equal to the new MCT and repeat the same procedure recursively until all possible combinations have been examined.*

The complexity of the previous algorithm is polynomial in the number of nodes and can be no worse than $O(n^K)$. Its main advantage is that, by using MCT's, it avoids the most complex part of search viz. that for the best routing trees.

5.6 Simulation Results

In order to compare the performance of our algorithms with that of the existing algorithm in [49], and the optimal scheme (in the average distortion case), we have performed a number of simulations for randomly created graphs.

The first category of graphs examined consists of *loosely connected random graphs*, which are generated as follows : select N random nodes, and for each pair of nodes (i, j), $i, j \in \{1 \dots N\}$, we establish a link between i and j with probability p. The capacity of this link is uniformly distributed over the interval $[0, 0.1)$, in order to avoid saturation of the distortion-rate function (see (5.12) below). It is obvious that as the number of nodes N increases, the number of links attached to each node will also increase. Thus, even for small values of p (*e.g.*, $p = 0.1$) a moderately large value for N renders the graph extremely "dense" and unrealistic. In practical situations, the nodes of a network are only connected to a small number of neighbouring nodes and the number of alternative paths between any two nodes of the network remains relatively low.

In order to choose a suitable value for the probability p (as a function of N) we have generated random graphs for different numbers of nodes and tested them for connectivity. It turns out that for $N < 100$ and

$$p = \exp\left\{-\alpha\sqrt{N}\right\},$$

with $\alpha = 0.25$, about 99% of the generated graphs are connected, whereas for larger values of α the resulting graphs are, with relatively high probability, not connected. In our simulation experiments, we have used $\alpha = 0.2$.

We consider various network sizes (in terms of the number of nodes N) and a number of information levels K. For each random graph, we calculate the average and maximum distortion for two different routing and rate-selection techniques : our proposed algorithms (*i.e.*, the Maximum Algorithm and the Heuristic Algorithm for the maximum and the average distortion cases, respectively) and the algorithm proposed by Shacham [49]. If we denote by D the (average or maximum) distortion yielded by our algorithm, and by D_S that yielded by Shacham's algorithm, then the normalized difference

$$\mathrm{PR} = 100\frac{D_S - D}{D},$$

which will be referred to as *Performance Ratio* (PR), can be interpreted as the percentage improvement (in terms of distortion) which results by using our algorithm instead of Shacham's. The PR's of 10000 random graphs for various values of N and K were calculated, and their average was assumed to be the expected value of the normalized performance difference. The performance (or distortion) measure used in this simulation is the binary-alphabet Hamming-

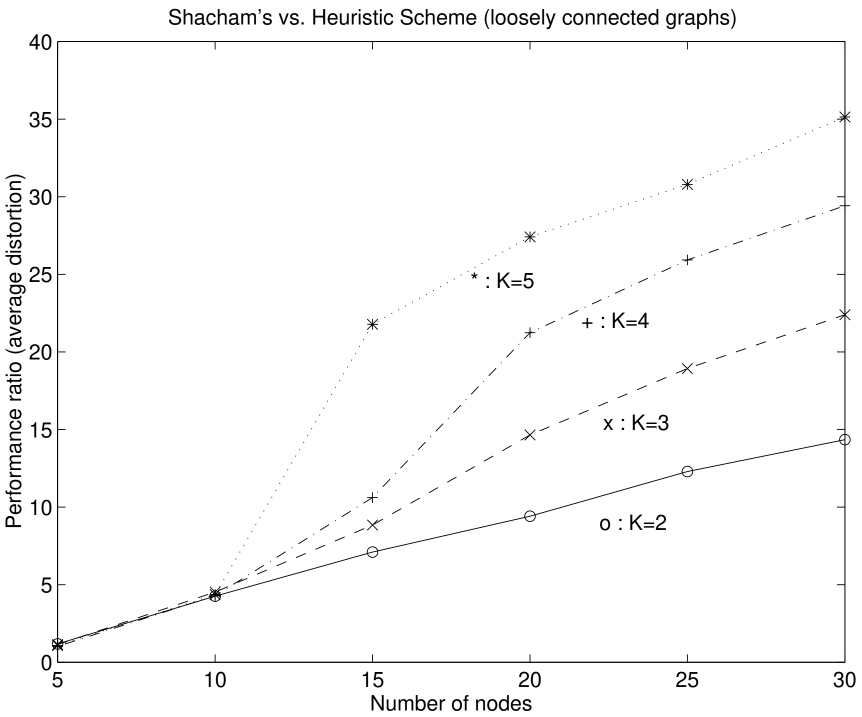

Figure 5.8: Simulation results for *loosely connected graphs* (average cost case).

distance distortion-rate function with equiprobable input symbols, given by

$$\Delta(R) = \begin{cases} 0, \;\; R > \frac{1}{2} \\ h_b^{-1}(1-R), \;\; 0 \leq R \leq \frac{1}{2} \end{cases}, \tag{5.12}$$

with R being the *overall* rate $R_1 + \ldots + R_K$ and $h_b(\cdot)$ the binary entropy function

$$h_b(p) = -p \log p - (1-p) \log(1-p), \quad p \in [0,1] \,.$$

The PR's for $N = 5, \ldots, 30$ and $K = 2, \ldots, 5$ are shown in Figures 5.8 and 5.9. Our algorithms exhibit considerable performance gains over the single-tree scheme (*i.e.*, that proposed by Shacham), both in the average and in the maximum distortion cases.

In the average distortion case, it is obvious that the line PR $= 0$ is the performance difference for $K = 1$, since Shacham's algorithm always yields the optimal

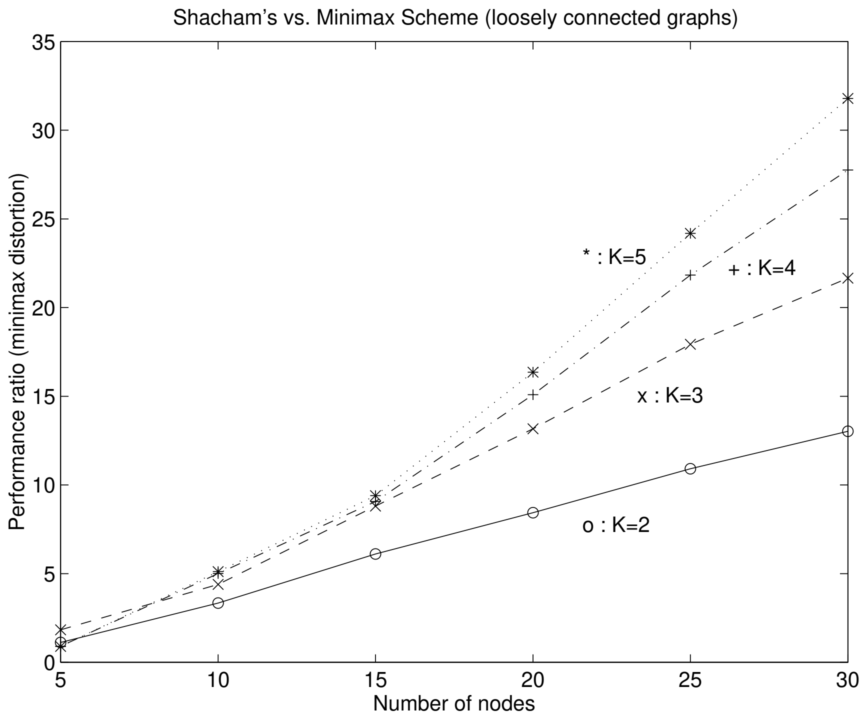

Figure 5.9: Simulation results for *loosely connected graphs* (maximum cost case).

path for one level of information. As should be expected, the performance ratio increases with the number of levels K, but is not proportional to K; rather, it decreases as K gets larger. This occurs as the rate of information reaching each node depends only on the capacity of the paths leading to it, which (on the average for random graphs) is limited; there exists an upper limit for the aggregate capacity of the paths leading to each node. This limit can be attained by using more information levels, thereby taking advantage of unused capacity, but convergence to the maximum overall rate is slow and gains from the introduction of more levels are minimal.

In the maximum distortion case, the performance ratio for $K = 1$ is again zero, because of the restriction imposed that at least one level has to reach each node. This comparison might not be very meaningful, since Shacham's scheme is not designed according to the maximum criterion. However, the significance of the corresponding graph lies in showing the improvement obtained in terms of maximum distortion, every time an additional level of information is used.

The second category of graphs with which we deal has also been presented in [54], and used by other researchers as well (*e.g.*,[4], [14], [64]); we shall refer to them as *locally connected random graphs*. These graphs differ from those described previously in that links between neighbouring nodes are favored over links between distant nodes. In particular, each node was placed at a random location on the square $[0,1] \times [0,1]$, according to the uniform probability distribution. A link between nodes u and v is created with probability

$$P_e(u,v) = \beta \exp\left\{-\frac{d(u,v)}{2\alpha}\right\}, \tag{5.13}$$

where $d(u,v)$ is the Manhattan distance between nodes u and v, *i.e.*,

$$d(u,v) = |x_u - x_v| + |y_u - y_v|,$$

with (x_u, y_u) being the coordinates of node u. If a link is established, it is assigned a random capacity uniformly on $[0, 0.1]$. Graphs generated in this manner resemble several networks encountered in practical situations in that connections between distant nodes are less likely than connections between neighbouring nodes.

In [14], the coordinates of the nodes can only be integers and the Euclidean distance metric is used, but we allow the coordinate values to be real numbers and have selected the Manhattan distance metric, as the resulting theoretical calculations are less complex. For this model, we have calculated the expected number of links emanating from each node (*i.e.*, the degree of the network) as a

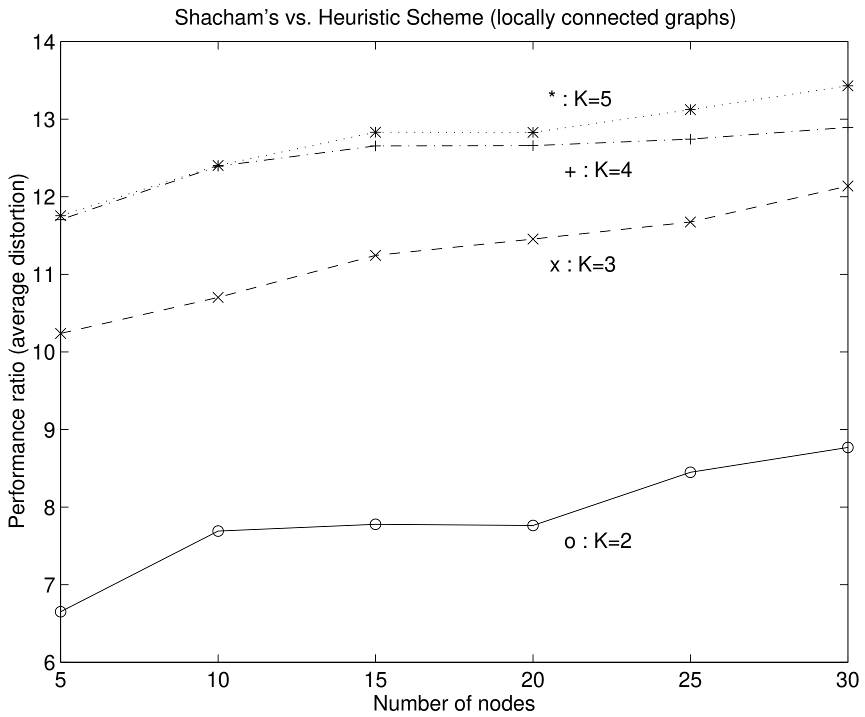

Figure 5.10: Simulation results for *locally connected graphs* (average cost case).

function of the parameters α and β in (5.13). We have used the values $\alpha = 0.5$ and $\beta = \frac{10}{N}$, which result in an average network degree of approximately 5.

The difference in performance between our proposed algorithms and the single-tree routing scheme of Shacham for these degree-constrained graphs is displayed in Figures 5.10 and 5.11. These results are similar to those in Figures 5.8 and 5.9. There are, however, a couple of noteworthy differences.

For the average distortion case, the performance difference again increases with the number of nodes N, but levels off for large values of N. This occurs as the information rate received by a node does not depend on the number of nodes in the network, but rather on the average number of links leading to it (*i.e.*, on the degree of the network), which in this case is fixed. As the number of nodes increases, the performance improvement is very small, because adding

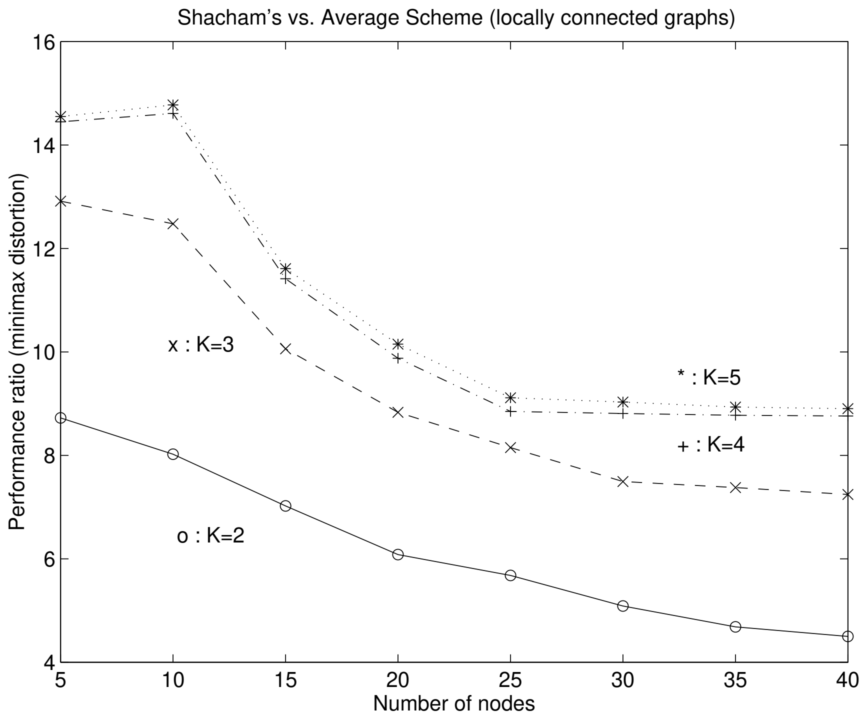

Figure 5.11: Simulation results for *locally connected graphs* (maximum cost case).

more nodes does not add more paths leading to each node, owing to the degree constraint. This suggests that the results obtained for $N > 10$ are valid for degree-constrained networks of a much larger size.

In the maximum distortion plot, in Figure 5.11, there is a considerable drop in the performance ratio as N grows larger. This is due to the fact that the maximum distortion depends on the capacity of the paths leading from the source to each node, namely the minimum capacity of the path links. As the number of network nodes increases, the average length (*i.e.*, number of intermediate nodes) of these paths also increases, yielding smaller path capacities. We also observe that because the network degree is limited to 5 links per node, the improvement in using 4 levels instead of 3, and (especially) 5 instead of 4, is very small, suggesting that no more than $d - 1$ levels of information should be

Loosely connected

N	PR
4	0.126
5	0.211
6	0.583

Locally connected

N	PR
4	0.303
5	0.587
6	1.023

Table 5.2: Heuristic vs. optimal scheme for loosely and locally connected graphs ($K = 2$)

used in a network with degree constrained to d. We should note though, that this comparison shows the improvement in maximum distortion incurred by the use of additional information levels. It suggests that the Maximum Algorithm might yield much better results for large networks. This is not surprising, since this very simple and straightforward algorithm is specifically designed for the maximum distortion criterion.

The reason for the absence of similar behaviour in loosely connected graphs is that such graphs are not degree-constrained. In fact, the network degree in this case varies from 2.5 for $N = 5$ to 11 for $N = 40$, which results in larger performance gains, as explained previously.

It is difficult to compare the performance of our Heuristic Algorithm with the optimal average distortion performance, because the exhaustive search solution is very costly in terms of running time. We can only obtain results for very small network sizes. Only for $N < 7$ does the exhaustive search algorithm terminate in a reasonable amount of time. The results are shown in Table 5.2, and exhibit a very small difference in average distortion. The PR in this case is

$$\mathrm{PR} = \frac{D - D_O}{D_O},$$

where D_O is the optimal average distortion and D is the average distortion using the Heuristic Algorithm. Based on the previous observation that the average performance should not depend much on the network size (at least for degree-constrained networks), we conjecture that the average distortion performance of the Heuristic Algorithm should not differ much from the optimal solution (not more than 5%) for larger values of N.

5.7 Discussion

Multicasting in a network with capacity restrictions requires data compression schemes which are more flexible than the conventional ones. Multiresolution data compression is a suitable choice; however, the problem of appropriately selecting the routes and rates of each coding level may not be easy to solve. When the criterion used is the maximum distortion incurred by any user, the optimal solution is given by the *Maximum Algorithm*. For the average distortion criterion, however, there is no polynomially complex algorithm which yields the optimal solution. Instead, we have to resort to exhaustive search, a method which is computationally prohibitive. In order to alleviate this difficulty, we propose a Heuristic Algorithm which is of polynomial complexity $O(n^K)$, and yields an average distortion close to the optimal (as observed in our simulation experiments). We remark that it may be possible to devise a means of overcoming exhaustive search for the rate selection part or devise a better heuristic approach (*e.g.*, by assigning to each link weights pertaining to the decrease in the average distortion, and thus reducing it to a Steiner-tree problem) of smaller complexity.

At this juncture, we note that in many practical situations, it is more desirable to use the maximum distortion criterion rather than the average distortion criterion. This is so, as the former is fairer than the latter, treating all the receivers equally and ensuring that all receivers obtain the same information. However, one drawback of the maximum cost criterion is that if all the links leading to a remote node saturate, then no additional information can be sent to that node or to any other node in the network, since this would not affect the value of the maximum distortion. This means that even if it were possible to send extra levels to *some* of the nodes, it will not lead to any improvement in maximum distortion, because we cannot send those levels to *all* the nodes. One possible strategy for improving the reception quality of the more capable nodes, while simultaneously guaranteeing fairness, is to exclude from the graph used to determine the optimal routes for the kth level, those nodes which are isolated after routing the previous $k-1$ levels (*i.e.*, those nodes for which all the links leading to them are saturated), and repeate this procedure until $k = K$.

It would also be interesting to investigate the case where not all nodes are receivers. However, it has been pointed out in Section 5.4 that the Maximum Algorithm gives the optimal solution for the maximum distortion problem in this case, too. For the average distortion problem, we can expect the discrepancy between the optimal solution and that given by the Heuristic Algorithm to be even smaller than in the situation where all nodes are receivers.

Finally, we remark that instead of the multi-level distortion-rate function, we can use as our criterion any other function $\hat{D} : \mathbf{R}_+^K \mapsto \mathbf{R}_+$ with similar properties. Also, note that the concavity of $\Delta(\cdot, \ldots, \cdot)$ has not been exploited above; it may be possible to exclude some of the alternatives in the exhaustive search in our Heuristic Algorithm, simply by using this property, thereby, further reducing its complexity.

Chapter 6

Optimal Bit Allocation in Universal Vector Quantization

6.1 Introduction

In the classical rate-distortion coding of an i.i.d. source, we describe all possible source sequences of a certain length by a relatively small number of reproduction vectors. Although the size of the set of possible source sequences grows exponentially with the sequence length n, with exponent equal to the logarithm of the source alphabet size $|\mathcal{X}|$ (*i.e.*, the number of possible source sequences is $\exp\{n \log |\mathcal{X}|\}$), we can find a set of reproduction vectors that incur distortion not more than a prespecified level, whose size is also exponential, but with a smaller exponent, specified by the rate-distortion function. Equivalently, we can find a collection of exponentially many reproduction vectors of a prespecified rate (exponent), which incur expected distortion given by the distortion-rate function.

The codes which are designed for the aforementioned situation are "matched" to the statistics of the source, which are assumed to be known. In general, the statistical behaviour of the source can only be estimated by observing a large number of source sequences, and it is possible that the "true" statistics differ from the estimated or assumed ones. The first question which arises is : what would be the performance of a code designed for a specific source when it is applied to a different source, *i.e.*, with different statistics. In the case of variable-length noiseless codes, we know that this results in a penalty, in terms of expected rate, given by the Kullback-Leibler divergence between the true and the assumed probability mass functions (pmf's) (*c.f.*, *e.g.*, [11]). However, to the best of our knowledge, there exists no similar result for fixed-rate lossy coding. Lapidoth in [35] has studied a different case of lossy coding mismatch, where the decoding rule (or distance measure) used at the encoder for the design of the code differs from the rule that is used in the decoder. He has also examined the performance of random Gaussian codebooks when used for encoding non-Gaussian sources, which, in a way, is a special case of the (previously mentioned) fixed-rate lossy mismatched coding for continuous-alphabet sources. An indication of what the answer to this problem might be is given by Theorem 6.1 in this chapter. This result leads us to the conjecture that the rate penalty for mismatched coding is given by a new quantity, the rate-distortion discrepancy.

The problem of mismatched coding can be addressed by universally encoding the source sequences. The term universal coding refers to the technique used to design codes whose performance approaches the optimal (namely, that achieved by a code matched to the true statistics) for an entire class of sources. The various universal coding techniques ([38], MDL in [47], etc.) usually allocate a small portion of the available bits for the description of the source statistics and the remainder for the description of the data. The number of bits used for the source description grows sublinearly with the block-length of the code, so that the per symbol contribution decays to zero as more complex codes (with larger block-lengths) are used, whereas the data bits are usually codevectors drawn from a codebook matched to the predicted source statistics. However, the schemes proposed in the literature to date do not always allocate the available rate (bits) optimally. For instance, Linder *et al* in [38] use a lossy fixed-rate coding scheme which has excess rate (because of the universality of their code) decaying as $|\mathcal{X}|\frac{\log n}{n}$. The main question here is whether this is the optimal rate at which one can describe the source statistics. Moreover, it is not clear how to determine what the optimal rate is, and provide a precise characterization of the optimality criterion.

In this chapter we shall adopt a criterion similar in spirit to that in [10]. We shall refer to it as the expected excess distortion (EED) criterion. In order to give a precise definition of this criterion, we must first provide a vector quantization (VQ) formulation of the universal lossy fixed-rate coding. In the VQ setup, the encoder comprises two vector quantizers. The first quantizes the source statistics space (which, in general, is nondiscrete) at some rate r and produces an estimate of the source statistics, as one of a finite number of statistics vectors. The second quantizer has access to a finite collection of codebooks, each corresponding to a quantization level of the first quantizer. By looking at the output of the first stage, we choose the appropriate codebook and use it to quantize the source sequence. The codebook picked is, in general, matched to a source close to but different from the true one (in the sense of statistical behaviour); so its performance under the true statistics will be described by the rate-distortion discrepancy. Assuming a prior on the statistics space, we can define the EED criterion to be the expected (over the set of all possible source statistics) rate-distortion discrepancy. This formulation assumes that the quantizer (*i.e.*, the encoder) is "omniscient," namely it has prior knowledge of the underlying source statistics. In reality, the prediction of the source behaviour is based on the source sequence to be quantized itself. The identity of the quantizer chosen on this basis must then be communicated to the receiver at rate r. However, it is shown in [10] that there is always a source-sequence based VQ scheme which outperforms the omniscient scheme; in other words, we can consider only omniscient codes without loss of generality.

Clearly, increasing the number of bits allocated for a description of the source will serve to improve the estimate of the source statistics, thereby decreasing the EED incurred by the quantizer used at the second stage. On the other hand, fewer bits will be available for use in the second stage (since the total number of bits is fixed), which results in higher distortion (worse description of the data). Hence, there is a trade-off between the quality of the descriptions of the source and the data, which corresponds to the choice of the rate r. We shall, therefore, characterize the optimal r as that which minimizes the EED criterion.

This chapter is organized as follows : In Section 6.2, we define the mismatched rate-distortion and distortion rate functions, and the rate-distortion discrepancy, which will be critical in the definition of the EED criterion, and investigate their properties. We also state a coding theorem for S-typical codes, leading to a conjecture on the performance of mismatched lossy codes. In Section 6.3, we state some general results for optimal rate allocation, and compare our results with existing results for noiseless variable-rate coding. Finally, in Section 6.4, we examine the special case of the universal compression of a binary alphabet source with the Hamming distortion measure, for which more specific theoretical results can be obtained; numerical results are compared with analytical predictions.

6.2 The Mismatched Rate-Distortion Function

The notion of a "mismatched" rate-distortion function which also involves an output pmf (in addition to an input pmf) is not new and has appeared in the earlier work of other researchers. For instance, Zhang *et al.* [63] have used the term "lower mutual information" to describe, in essence, the same quantity. We have introduced the mismatched rate-distortion function in Chapter 2.

The term "mismatched rate-distortion function" (mRDf) is motivated by the property that it coincides with the (conventional) rate-distortion function (RDf) only when the output pmf is "matched" to the input pmf, *i.e.*, the output pmf coincides with that induced by the minimizing channel in the definition of the RDf (*c.f.*, *e.g.*, [11]). For all other output pmf's, the mRDf is never less than the RDf.

6.2.1 Definitions

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets representing the source and reproduction alphabets, respectively. Let P be a given pmf on $\mathcal{X}$, and S a given pmf on $\mathcal{Y}$. Let $d : \mathcal{X} \times \mathcal{Y} \mapsto \mathbf{R}^+$ be a nonnegative-valued mapping; we shall assume without loss of generality that $\min_{x,y} d(x,y) = 0$. We define the mRDf as

$$\tilde{R}(P,S,\Delta) \stackrel{\triangle}{=} \inf_{\substack{W: \quad P\cdot W = S \\ d(P,W) \leq \Delta}} I(P,W), \quad \Delta \geq \Delta_{min}, \tag{6.1}$$

where W ranges over all stochastic matrices $W : \mathcal{X} \mapsto \mathcal{Y}$, and $P \cdot W$ is a pmf on $\mathcal{Y}$ given by

$$(P \cdot W)(y) = \sum_x P(x)W(y|x), \quad y \in \mathcal{Y}.$$

Also,

$$d(P, W) \triangleq \sum_{x,y} P(x)W(y|x)d(x, y),$$

and Δ_{min} is the smallest value of $\Delta \geq 0$ for which the constraint set in (6.1) is nonempty.

In a manner analogous to the definition of the distortion-rate function (*c.f.*, *e.g.*, [7, 11]), we define the mismatched distortion-rate function (mDRf) as

$$\tilde{D}(P, S, R) = \inf_{\substack{W: \; P \cdot W = S \\ I(P,W) \leq R}} d(P, W), \quad 0 \leq R_{max}, \tag{6.2}$$

where $R_{max} \triangleq \tilde{R}(P, S, \Delta_{min})$.

The operational significance of $\tilde{R}(P, S, \Delta)$ and $\tilde{D}(P, S, R)$ is provided by Theorem 6.1 below. Let $\{X_t\}_{t=1}^{\infty}$ be a $\mathcal{X}$-valued discrete memoryless source (DMS), *i.e.*, an independent and identically distributed (i.i.d.) process, with pmf P. Then, $\tilde{R}(P, S, \Delta)$ is the normalized (per symbol) logarithm of the minimum number of S-typical (reproduction) vectors needed to represent the DMS within distortion Δ, whereas $\tilde{D}(P, S, R)$ is the minimum possible expected (per symbol) distortion resulting from the representation of the DMS with 2^{nR} S-typical reproduction vectors.

The RDf for the DMS $\{X_t\}_{t=1}^{\infty}$ is given by

$$R(P, \Delta) = \inf_{W: d(P,W) \leq \Delta} I(P, W)$$

and is a lower bound for the mRDf, since the minimization in (6.1) is performed over a smaller set. If the minimizing channel in the definition of $R(P, \Delta)$ is unique, then the mRDf and RDf coincide iff S equals the output pmf corresponding to the minimizing channel. In fact, we can write

$$R(P,\Delta) = \min_S \tilde{R}(P,S,\Delta).$$

6.2.2 Properties

The functions $\tilde{R}(P,S,\Delta)$ and $\tilde{D}(P,S,R)$ share some of the properties of the rate-distortion and distortion-rate functions, namely convexity and monotonicity with respect to the appropriate variables.

The monotonicity of $\tilde{R}(P,S,\Delta)$ in Δ and $\tilde{D}(P,S,R)$ in R is clear from the definitions (6.1) and (6.2). For a fixed pmf P on $\mathcal{X}$, the convexity of $\tilde{R}(P,S,\Delta)$ with respect to (S,Δ), can also be shown easily. Let W_1^* and W_2^* be the minimizing channels in (6.1) for output pmf's S_1, S_2 and distortions Δ_1, Δ_2 respectively. Let $\alpha \in [0,1]$. If $\Delta_1 \geq \Delta_{min}(P,S_1)$ and $\Delta_2 \geq \Delta_{min}(P,S_2)$ then it is easy to verify that $\alpha\Delta_1 + (1-\alpha)\Delta_2 \geq \Delta_{min}(P, \alpha S_1 + (1-\alpha)S_2)$. Also,

$$\begin{aligned}
\tilde{R}(P,\alpha S_1 + (1-\alpha)S_2, \alpha\Delta_1 + (1-\alpha)\Delta_2) &\overset{(A)}{\leq} I(P, \alpha W_1^* + (1-\alpha)W_2^*) \\
&\overset{(B)}{\leq} \alpha I(P,W_1^*) + (1-\alpha)I(P,W_2^*) \\
&= \alpha\tilde{R}(P,S,\Delta_1) + (1-\alpha)\tilde{R}(P,S,\Delta_2),
\end{aligned}$$

where (B) results from the convexity of of $I(P,W)$ in W for a fixed P, and (A) holds because $\alpha W_1^* + (1-\alpha)W_2^*$ satisfies the conditions

$$P \cdot (\alpha W_1^* + (1-\alpha)W_2^*) = \alpha S_1 + (1-\alpha)S_2,$$

and

$$\sum_{x,y} P(x)(\alpha W_1^*(y|x) + (1-\alpha)W_2^*(y|x))d(x,y) \leq \alpha\Delta_1 + (1-\alpha)\Delta_2.$$

The convexity of $\tilde{D}(P,S,R)$ with respect to (S,R) can also be shown in a similar way.

It is not hard to show that the mDRf is the inverse of the mRDf, *i.e.*,

$$\tilde{R}(P,S,\tilde{D}(P,S,R)) = R, \ \ 0 \leq R_{max}.$$

Similarly,

$$\tilde{D}(P,S,\tilde{R}(P,S,\Delta)) = \Delta, \ \ \Delta_{min} \leq \Delta \leq \Delta_{max},$$

where Δ_{max} is the smallest value of Δ for which $\tilde{R}(P,S,\Delta) = 0$ and can be shown to equal

$$\Delta_{max} = \sum_{x,y} P(x)S(y)d(x,y).$$

To this end, given the pmf's P and S on $\mathcal{X}$ and $\mathcal{Y}$, respectively, we define $D(r) \triangleq \tilde{D}(P,S,r)$ and $R(\delta) \triangleq \tilde{R}(P,S,\delta)$, and let W_P^* be the channel which achieves the minimum in (6.1) at distortion δ. Then, it satisfies

$$\begin{aligned} I(P,W_P^*) &= R(\delta), \\ P \cdot W_P^* &= S, \\ d(P,W_P^*) &\leq \delta. \end{aligned}$$

We shall show that $D(R(\delta)) = \delta$. Define

$$\delta' \triangleq D(R(\delta)).$$

Then, by the definition of the mDRf (6.2), $\delta' \leq d\left(P,W_P^*\right)$, since W_P^* satisfies the first two of the conditions above. From the fourth condition, we get $\delta' \leq \delta$.

In a similar manner, we get $R(D(r)) \leq r$; setting $r = R(\delta)$, we have $R(\delta') \leq R(\delta)$. The monotonicity of $R(\cdot)$ results in $\delta' \geq \delta$. We, therefore, conclude that $\delta' = \delta$. Thus, the mRDf and mDRf are inverses of each other.

6.2.3 Coding with S-Typical Codewords

In the forward part of the rate-distortion theorem for a DMS with pmf P, namely in the construction of a code whose performance approaches that predicted by the rate-distortion function, $R(P,\Delta)$, we randomly pick codewords from the $P \cdot W_P^*$-typical set $\mathcal{T}_{[P \cdot W_P^*]} \subset \mathcal{Y}^n$, where $P \cdot W_P^*$ is the output pmf corresponding to the rate-distortion achieving channel W_P^*. If, however, we restrict ourselves to using only codewords of some prespecified type, which differs from $P \cdot W_P^*$, then the minimum rate at which we can describe the source within a distortion level Δ, will be no smaller than $R(P,\Delta)$. This result is formally stated below as

Theorem 6.1 *Consider a $\mathcal{X}$-valued DMS $\{X_t\}_{t=1}^{\infty}$ with pmf P, and a distortion measure d. Let S be a given pmf on the reproduction alphabet $\mathcal{Y}$. For each $\Delta \geq \Delta_{min}$ there exists a n-length block code with codewords from the S-typical set $\mathcal{T}_{[S]}^n \subset \mathcal{Y}^n$ for which the expected distortion does not exceed Δ, and whose rate is arbitrarily close to $\tilde{R}(P,S,\Delta)$, provided n is sufficiently large; conversely, any n-length block code consisting of S-typical codewords, for which the expected distortion does not exceed Δ, must have a rate of at least $\tilde{R}(P,S,\Delta)$, for every n.*

Similarly, for each $0 \leq R \leq R_{max}$, there exists a n-length block code of rate R with codewords from $\mathcal{T}_{[S]}^n \subset \mathcal{Y}^n$, for which the expected distortion converges to $\tilde{D}(P,S,R)$ as the block-length goes to infinity; conversely, the expected distortion of any n-length block code of rate R with S-typical codewords does not exceed $\tilde{D}(P,S,R)$.

The previous theorem, however, does not imply that the performance of any code with S-typical codewords on a DMS with pmf P can be characterized in terms of $\tilde{R}(P,S,\Delta)$ or $\tilde{D}(P,S,R)$. If such a code is additionally well-behaved in that it is optimal for a DMS with pmf P' (for which S constitutes the "best" pmf on $\mathcal{Y}$), it is conceivable that its performance under P may be characterized in terms of $\tilde{R}(P,S,\Delta)$ or $\tilde{D}(P,S,R)$.

Proof: We sketch a proof of the second assertion of the theorem. Namely, we show that there exists a n-length block code of rate R with codewords in $\mathcal{T}^n_{[S]}$ for which the expected distortion with respect to the pmf P converges to $\tilde{D}(P,S,R)$ as n tends to infinity; conversely, the expected distortion (with respect to the pmf P) of any n-length block code of rate R and codewords in $\mathcal{T}^n_{[S]}$ cannot be less than $\tilde{D}(P,S,R)$. Our approach follows the proof of the distortion-rate theorem (*c.f.*, *e.g.*, [7]) and relies on the method of types [13]. The proof of the first part of the theorem is similar in spirit, and is omitted.

Considering first the converse part of the assertion above, let $\mathcal{C}_n \subset \mathcal{T}^n_{[S]}$ be the codebook of a n-length block code of rate R, *i.e.*,

$$\frac{1}{n}\log|\mathcal{C}_n| \leq R,$$

with S-typical codewords; let $g : \mathcal{X}^n \mapsto \mathcal{C}_n$ denote the map which assigns codewords to the source sequences.

Denote by $A_n(\mathbf{y})$ the set of sequences which are mapped to a particular $\mathbf{y} \in \mathcal{C}_n$, *i.e.*,

$$A_n(\mathbf{y}) \triangleq g^{-1}(\{\mathbf{y}\}) = \{\mathbf{x} \in \mathcal{X}^n : g(\mathbf{x}) = \mathbf{y}\}, \quad \mathbf{y} \in \mathcal{C}_n.$$

The average distortion of this code, with respect to the (i.i.d.) pmf P on $\mathcal{X}^n$, is

$$\begin{aligned}\mathbf{d}(g) &= \mathbf{E}d(X^n, g(X^n)) \\ &= \sum_{\mathbf{x}\in\mathcal{X}^n} P^n(\mathbf{x})d(\mathbf{x}, g(\mathbf{x}))\end{aligned}$$

$$= \sum_{Q} \sum_{\mathbf{x}\in\mathcal{T}_Q^n} P^n(\mathbf{x}) d(\mathbf{x}, g(\mathbf{x})),$$

where the first sum is over all types Q of n-length sequences in $\mathcal{X}^n$. Since only those types which are "close" to P will be significant in the last sum above, we can approximate the average distortion as follows

$$\begin{aligned}
\mathbf{d}(g) &\simeq \sum_{\mathbf{x}\in\mathcal{T}_P^n} P^n(\mathbf{x}) d(\mathbf{x}, g(\mathbf{x})) \\
&= \exp\{-nH(P)\} \sum_{\mathbf{x}\in\mathcal{T}_P^n} d(\mathbf{x}, g(\mathbf{x})) \\
&= \exp\{-nH(P)\} \sum_{\mathbf{y}\in\mathcal{C}_n} \sum_{\mathbf{x}\in\mathcal{T}_P^n \bigcap A_n(\mathbf{y})} d(\mathbf{x}, \mathbf{y}) \\
&= \exp\{-nH(P)\} \sum_{\mathbf{y}\in\mathcal{C}_n} \sum_{V:T_{\mathbf{y}}\cdot V=P} \sum_{\mathbf{x}\in\mathcal{T}_V^n(\mathbf{y}) \bigcap A_n(\mathbf{y})} d(T_{\mathbf{y}}, V) \\
&= \exp\{-nH(P)\} \sum_{\mathbf{y}\in\mathcal{C}_n} \sum_{V:T_{\mathbf{y}}\cdot V=P} |\mathcal{T}_V^n(\mathbf{y}) \bigcap A_n(\mathbf{y})| d(T_{\mathbf{y}}, V),
\end{aligned}$$

where the second summation is over all the V-shells of $\mathbf{y}$ which induce the pmf P on $\mathcal{X}$, and $T_{\mathbf{y}}$ is the empirical distribution (type) of the output sequence $\mathbf{y}$.

For each $\mathbf{y} \in \mathcal{C}_n$ define $\alpha_n(\mathbf{y})$ by

$$\frac{1}{n} \log |A_n(\mathbf{y}) \cap \mathcal{T}_P^n| = H(P) - \alpha_n(\mathbf{y}).$$

Next, note that

$$A_n(\mathbf{y}) \cap \mathcal{T}_P^n = \bigcup_{V:T_{\mathbf{y}}\cdot V=P} \mathcal{T}_V^n(\mathbf{y}) \cap A_n(\mathbf{y}),$$

and since there are only polynomially many (in n) V-shells of $\mathbf{y}$, each with exponentially large cardinality, there must exist at least one V with $T_{\mathbf{y}} \cdot V = P$ for which

$$|\mathcal{T}_V^n(\mathbf{y}) \cap A_n(\mathbf{y})| = \exp\left[n\left(H(P) - \alpha_n(\mathbf{y})\right)\right], \tag{6.3}$$

i.e., for which

$$|\mathcal{T}_V^n(\mathbf{y})| \geq |A_n(\mathbf{y}) \cap \mathcal{T}_P^n|,$$

i.e.,

$$H(V|T_{\mathbf{y}}) \geq H(P) - \alpha_n(\mathbf{y}),$$

i.e.,

$$I(T_{\mathbf{y}}, V) \leq \alpha_n(\mathbf{y}). \tag{6.4}$$

Hence, we can further approximate the average distortion by

$$\mathbf{d}(g) \simeq \exp\{-nH(P)\} \sum_{\mathbf{y}\in\mathcal{C}_n} \sum_{V:T_{\mathbf{y}}\cdot V=P, I(T_{\mathbf{y}},V)\leq\alpha_n(\mathbf{y})} |\mathcal{T}_V^n(\mathbf{y}) \cap A_n(\mathbf{y})| d(T_{\mathbf{y}}, V).$$

Clearly, by (6.3),

$$\mathbf{d}(g) \overset{\sim}{\geq} \exp\{-nH(P)\} \sum_{\mathbf{y}\in\mathcal{C}_n} \exp\{n[H(P) - \alpha_n(\mathbf{y})]\} \min_{V:T_{\mathbf{y}}\cdot V=P, I(T_{\mathbf{y}},V)\leq\alpha_n(\mathbf{y})} d(T_{\mathbf{y}}, V),$$

so that

$$\mathbf{d}(g) \geq \sum_{\mathbf{y}\in\mathcal{C}_n} \exp\{-n\alpha_n(\mathbf{y})\} \tilde{D}(P, T_{\mathbf{y}}, \alpha_n(\mathbf{y})).$$

The quantity on the right side is a function of the $\exp\{nR\}$ variables $\alpha_n(\mathbf{y})$, $\mathbf{y} \in \mathcal{C}_n$, which satisfy

$$\sum_{\mathbf{y}\in\mathcal{C}_n} |A_n(\mathbf{y}) \cap \mathcal{T}_P^n| = |\mathcal{T}_P^n|,$$

i.e.,

$$\sum_{\mathbf{y}\in\mathcal{C}_n} \exp\{-n\alpha_n(\mathbf{y})\} = \exp\{-nH(P)\} \cdot |\mathcal{T}_P^n|.$$

If all the codewords are drawn from the set $\mathcal{T}_S^n \subset \mathcal{Y}^n$, then the problem of minimizing the expected distortion is described by the functional

$$\sum_{i=1}^{\exp\{nR\}} \exp\{-nx_i\} f(x_i) + \lambda \sum_{i=1}^{\exp\{nR\}} \exp\{-nx_i\},$$

with $f(\cdot)$ being the (decreasing) distortion-rate function $\tilde{D}(P, S, \cdot)$. Obviously, the minimum value is achieved for

$$\alpha_n(\mathbf{y}) = const \simeq R, \quad \mathbf{y} \in \mathcal{C}_n.$$

Therefore, the average distortion of a code of rate R and codewords of type S is bounded below as

$$\mathbf{d}(g) \tilde{\geq} \tilde{D}(P, S, R).$$

Now, let us consider the forward part. Let $\mathcal{C}_n$ be a (random) codebook of rate R with codewords belonging to $\mathcal{T}_S^n$. The codewords $\{\mathbf{y}_i, 1 \leq i \leq M\}$ are drawn independently according to the uniform distribution on $\mathcal{T}_S^n$, *i.e.*,

$$\Pr\left\{\mathcal{C}_n = \{\mathbf{y}_1, \ldots, \mathbf{y}_M\}\right\} = \prod_{i=1}^{M} \frac{\mathbf{1}_{\mathcal{T}_S^n}(\mathbf{y}_i)}{|\mathcal{T}_S^n|},$$

where $M \triangleq \exp\{nR\}$. The ensemble average value (over all possible realizations of the random code) of the expected distortion can be written as

$$\mathbf{E}_{\mathcal{C}_n} \mathbf{d}(\mathcal{C}_n) = \int_0^{d_{max}} \Pr\left\{\mathbf{d}(\mathcal{C}_n) > \delta\right\} d\delta.$$

Define

$$\Phi_n(\delta) \triangleq \Pr\left\{\mathbf{d}(\mathcal{C}_n) > \delta\right\}.$$

Since Φ_n is nonincreasing, it is easy to see that

$$\mathbf{E}_{\mathcal{C}_n} \mathbf{d}(\mathcal{C}_n) \leq \delta + d_{max} \cdot \Phi_n(\delta), \quad \delta \in [0, d_{max}], \tag{6.5}$$

where $d_{max} \triangleq \max_{x,y} d(x, y)$. In order to proceed further with the bound in (6.5), we need to estimate Φ_n. To this end, first note that we can represent the random code $\mathcal{C}_n$ in terms of the i.i.d. rv's $\mathbf{Y}_1, \ldots, \mathbf{Y}_M$, each uniformly distributed on $\mathcal{T}_S^n$. Then,

$$\begin{aligned}
\Phi_n(\delta) &\simeq \sum_{\mathbf{x} \in \mathcal{T}_P^n} P^n(\mathbf{x}) \Pr\left\{\min_{1 \leq i \leq M} d(\mathbf{x}, \mathbf{Y}_i) > \delta\right\} \\
&= \sum_{\mathbf{x} \in \mathcal{T}_P^n} P^n(\mathbf{x}) \prod_{i=1}^{M} \Pr\{d(\mathbf{x}, \mathbf{Y}_i) > \delta\}.
\end{aligned}$$

Next, observe that for a given $\mathbf{x} \in \mathcal{T}_P^n$, the probability values in the product above are identical, because the $\mathbf{Y}_i$'s are i.i.d. Moreover, each term in the product equals

$$\begin{aligned}
\Pr\{d(\mathbf{x},\mathbf{Y}) > \delta\} &= 1 - \frac{1}{|\mathcal{T}_S^n|} \sum_{\mathbf{y}\in\mathcal{T}_S^n} \mathbf{1}\left(d(\mathbf{x},\mathbf{y}) \le \delta\right) \\
&= 1 - \frac{1}{|\mathcal{T}_S^n|} \sum_{W: P\cdot W = S, d(P,W)\le\delta} |\mathcal{T}_W^n(\mathbf{x})| \\
&\simeq 1 - \frac{1}{|\mathcal{T}_S^n|} \sum_{W: P\cdot W = S, d(P,W)\le\delta} \exp\left\{n\left[H(S) - I(P,W)\right]\right\} \\
&\simeq 1 - \exp\left\{-n\tilde{R}(P,S,\delta)\right\},
\end{aligned}$$

because the sum will be dominated by the exponent which corresponds to the channel W^* which minimizes the mutual information $I(P,W)$ under the constraints $P\cdot W = S$ and $d(P,W) \le \delta$. Using this approximation, we can bound Φ_n above as

$$\begin{aligned}
\Phi_n(\delta) &\simeq \sum_{\mathbf{x}\in\mathcal{T}_P^n} P^n(\mathbf{x}) \left(1 - \exp\left\{-n\tilde{R}(P,S,\delta)\right\}\right)^M \\
&\tilde{\le} \sum_{\mathbf{x}\in\mathcal{T}_P^n} P^n(\mathbf{x}) \exp\left(-M\exp\left\{-n\tilde{R}(P,S,\delta)\right\}\right) \\
&= \sum_{\mathbf{x}\in\mathcal{T}_P^n} P^n(\mathbf{x}) \exp\left(-\exp\left\{n\left[R - \tilde{R}(P,S,\delta)\right]\right\}\right),
\end{aligned}$$

using the well-known inequality $(1-a) \le 2^{-a}$, $0 < a \le 1$. So, $\Phi_n(\delta)$ converges to one if the exponent in the expression above is negative and to zero if the exponent is nonnegative, *i.e.*,

$$\lim_n \Phi_n(\delta) = \begin{cases} 1, & R \ge \tilde{R}(P,S,\delta) \\ 0, & R < \tilde{R}(P,S,\delta) \end{cases} = \begin{cases} 1, & \delta \le \tilde{D}(P,S,R) \\ 0, & \delta > \tilde{D}(P,S,R). \end{cases} \tag{6.6}$$

Substituting this in (6.5), we get

$$\mathbf{E}_{\mathcal{C}_n}\mathbf{d}(\mathcal{C}_n) \le \inf_{\delta:\lim_n \Phi_n(\delta)=0} \delta + o(1)$$

$$\leq \inf_{\delta:\delta>\tilde{D}(P,S,R)} +o(1)$$
$$\tilde{\leq} \quad \tilde{D}(P,S,R)+o(1),$$

where the second inequality above follows from (6.6). We then conclude in the standard manner that there exists a realization of the random code $\mathcal{C}_n$ whose expected distortion does not exceed $\tilde{D}(P,S,R)+o(1)$. ■

The motivation for Theorem 6.1 arises from the fact that most of the codewords of a "good code" which is matched to some pmf P, will be of type close to $P \cdot W_P^*$, as has been shown in Chapter 2. It is, therefore, of interest to investigate the properties of a mismatched code, namely a code whose codewords are drawn according to a pmf different from $P \cdot W_P^*$, and try to establish lower and (if possible) upper bounds on the performance of such a code.

6.2.4 The Rate-Distortion Discrepancy

A relevant quantity in the assessment of mismatched coding performance is the rate-distortion discrepancy. Let $\{X_t\}_{t=1}^{\infty}$ be a $\mathcal{X}$-valued DMS with pmf P. Let $\Delta \geq 0$ be a given value of distortion. Then the rate-distortion discrepancy is defined as the difference between the mDRf when the output pmf is $S^*(Q,\Delta) \triangleq Q \cdot W_Q^*$ (namely the output pmf induced by the minimizing channel in the rate-distortion problem for some input pmf $Q \neq P$) and the rate-distortion function. Its significance will become apparent later in this section. We define the rate-distortion discrepancy between P and Q as

$$\begin{aligned} R_d(P,Q;\Delta) &\triangleq \tilde{R}(P,S^*(Q,\Delta),\Delta) - R(P,\Delta) \\ &= \inf_{\substack{W: \; P\cdot W=S^*(Q,\Delta) \\ d(P,W)\leq\Delta}} I(P,W)-I(P,W_P^*), \;\; \Delta\geq\Delta_{min}(P,S^*(Q,\Delta)). \end{aligned} \quad (6.7)$$

We can also define the distortion-rate discrepancy in a similar manner as

$$
\begin{aligned}
D_d(P,Q;R) &\triangleq \tilde{D}(P,S^*(Q,R),R) - D(P,R) \\
&= \inf_{\substack{W:\ \ P\cdot W = S^*(Q,R) \\ I(P,W)\le R}} d(P,W) - d(P,W_P^*), \quad 0 \le R \le R_{max}(P,S^*(Q,\Delta)).
\end{aligned}
$$

where $S^*(Q,R) \triangleq Q\cdot W_Q^*$ and W_Q^* is the minimizing channel in the distortion-rate problem for input pmf Q and rate R.

Unlike the mRDf and mDRf, it is not clear if the rate-distortion discrepancy given by (6.7) is either monotone or convex in Δ. Observe that, for $\Delta < \Delta_{min}$ the constraint set in (6.7) is empty and the discrepancy has to be defined using the fact that it equals the excess mismatched-coding rate as stated in Theorem 6.1.

An important property of the rate-distortion discrepancy $R_d(P,Q;\Delta)$ is that it coincides with the rate distortion function $R(P,\Delta)$ for $P = Q$. However, it is not true that it is continuous as a function of Q in some neighbourhood of P, as can be seen by considering the situation where $\Delta = 0$.

Let $\mathcal{X} = \mathcal{Y}$, and let the function $d(x,y)$ satisfy $d(x,y) = 0$ iff $x = y$. For the case $\Delta = 0$, it might then be expected that the rate-distortion discrepancy $R_d(P,Q;0)$ equals the (lossless) expected coding redundancy $D(P\|Q)$; the latter would result from using a variable-length code based on the pmf Q when the "true" pmf is P. However, $R_d(P,Q;0)$ cannot be defined from (6.7), since the constraint set $\{W : P\cdot W = Q\cdot W_Q^*, d(P,W) = 0\}$ is empty for $P \neq Q\cdot W_Q^*$, as $d(P,W) = 0$ iff $\Pr\{X = Y\} = 1$, which in turn implies $P_X = P_Y$. Although we could define $R_d(P,Q;0)$ to equal $D(P\|Q)$, Theorem 6.1 implies that it should be equal to $\mathbf{1}(P \neq Q)\cdot H(P)$. Indeed, a block code for a DMS with pmf Q, for $\Delta = 0$, comprises codewords that only cover $\mathcal{T}^n_{[Q]}$; thus, any source with pmf $P \neq Q$ will not be covered at all by the existing codewords. We should therefore use a new set of $2^{nH(P)}$ P-typical codewords to cover $\mathcal{T}^n_{[P]}$ noiselessly. Hence, the rate-distortion discrepancy $R_d(P, Q; 0)$ is equal to $H(P)$ for $P \neq Q$ and 0 for $P = Q$.

This should not be surprising, though. The essential difference is that here we are considering block coding whereas noiseless coding, where the discrepancy is characterized by $D(P\|Q)$, involves variable-length codes.

Based on the observation that the codewords of a good code designed for some particular pmf Q have empirical pmf close to the optimal output pmf corresponding to Q we state the following

Conjecture 6.1 *Let $\{X_t\}_{t=1}^{\infty}$ be a DMS with pmf Q. Let $\Delta > 0$ be an acceptable distortion level. Consider a distortion-rate code of rate R which is matched to Q,* i.e., *it incurs expected distortion $D(Q, R)$. The expected excess distortion resulting from using this code on a DMS with pmf P is given by the distortion-rate discrepancy $D_d(P, Q; R)$, where the expected excess distortion is defined as the difference between the expected distortion incurred by the code and the minimum achievable distortion for the DMS, namely $D(P, R)$.*

This means that the discrepancy is both a lower and an (asymptotic) upper bound of the performance of a good mismatched code. The proof for the lower bound is based on the proof of Theorem 6.1 and the fact that the type of most of the codewords of a good code is close to the output distribution induced by the minimizing rate-distortion channel as has been shown in Chapter 2. The proof for the upper bound remains elusive; loosely speaking, since the codebook is optimal for $\mathcal{T}^n_{[Q]}$ its codewords have to be as "spaced out" as possible (in the

Euclidean topology, they will form a lattice). The optimal S-typical code (with $S = Q \cdot W_Q^*$) will also have "spaced out" codewords, and since its performance converges to the optimal, it is possible that it is matched by that of the optimal code for the pmf Q.

To demand that *every good code* designed for some pmf Q performs nearly optimally, in the distortion-rate discrepancy sense, for any pmf $P \neq Q$ is actually more than we need. In fact, in order to substantiate the results of Section 6.3, it suffices to demonstrate that among the good codes which are matched to the pmf Q, there exists at least one that incurs excess distortion close to $D_d(P, Q; R)$ for *every* pmf P. The following theorem provides an indication of the existence of such a code.

Theorem 6.2 *Let Q and P be given pmf's on $\mathcal{X}$. There exists a n-length block code of rate R and $Q \cdot W_Q^*$-typical codewords, for which the expected distortion under Q and P is arbitrarily close to $D(Q, R)$ and $D(P, Q \cdot W_Q^*; R)$, respectively, for all n suitably large.*

Proof: We know from Theorem 6.1 that the expected distortion under Q (P) of any code of rate R and $Q \cdot W_Q^*$-typical codewords, represented by $g : \mathcal{X}^n \mapsto \mathcal{Y}^n$, is bounded below by $D(Q, R)$ (resp. $D(P, Q \cdot W_Q^*, R)$). Moreover, from the proof of the forward part of Theorem 6.1, if the code g is chosen randomly among all possible $Q \cdot W_Q^*$-typical codes, then

$$
\begin{aligned}
\mathbf{E}_{\mathcal{C}_n} \left\{ \mathbf{E}_Q \left[d(X^n, g(X^n)) | \mathcal{C}_n = g \right] \right\} &< D(Q, R) + \epsilon, \\
\mathbf{E}_{\mathcal{C}_n} \left\{ \mathbf{E}_P \left[d(X^n, g(X^n)) | \mathcal{C}_n = g \right] \right\} &< D(P, Q \cdot W_Q^*, R) + \epsilon,
\end{aligned}
$$

where $\mathbf{E}_{\mathcal{C}_n}$ denotes expectation with respect to the distribution of the random code $\mathcal{C}_n$ and $\mathbf{E}_Q$ and $\mathbf{E}_P$ denote conditional expectations with respect to the input pmf's Q and P, respectively, given a realization g of the random code $\mathcal{C}_n$. Thus, we get

$$\mathbf{E}_{\mathcal{C}_n}\left\{\mathbf{E}_Q\left[d(X^n, g(X^n))|\mathcal{C}_n = g\right] + \mathbf{E}_P\left[d(X^n, g(X^n))|\mathcal{C}_n = g\right]\right\}$$

$$< D(Q, R) + D(P, Q \cdot W_Q^*, R) + 2\epsilon.$$

Hence, there exists a realization, say $\tilde{g}$, of the random code $\mathcal{C}_n$ for which the sum of the expected distortions with respect to Q and P is close to $D(Q, R) + D(P, Q \cdot W_Q^*, R)$. This, together with the lower bounds on these expected distortions mentioned at the beginning of the proof, yield for this realization $\tilde{g}$ of the random code $\mathcal{C}_n$, that each of the expected distortion is close to its optimal value, as claimed. ■

The result of Theorem 6.2 can also be extended for the case of a pmf Q and any finite number of pmf's $P \neq Q$. Thus, we can claim that there exists a code matched to Q, with expected distortion close to $D(P, Q \cdot W_Q^*, R)$ for every type P, which will prove to be useful in the next section.

6.3 Optimal Bit Allocation for Source Description

Let $\{X_t\}_{t=1}^{\infty}$ be a $\mathcal{X}$-valued DMS with unknown pmf. Let Θ be the set of all pmf's on $\mathcal{X}$; hereafter, we shall often refer to Θ as the pmf space. Since $\mathcal{X}$ has been assumed to be a finite set, we can take Θ to be the $(|\mathcal{X}| - 1)$-dimensional space probability simplex.

We shall study a two-stage vector quantization (VQ) approach, namely to first quantize the pmf space Θ and then the data space $\mathcal{X}^n$. We are allowed a total of R bits per source symbol to describe sequences from $\mathcal{X}^n$ of which we shall allot $r = r_n$ bits for the description of the source pmf and the remaining $nR - r$ bits for the description of the n-length data sequence from $\mathcal{X}^n$. This results in an allocation of $R - r/n$ bits per source symbol for the data. Hereafter, we shall refer to r as the pmf rate and to $R - r/n$ as the data rate.

Let $\hat{g}_r = \phi_r \circ f_r$ be the quantizer for the pmf space Θ, with $f_r : \Theta \mapsto \{1, \ldots, M_r\}$ being the encoder, $\phi_r : \{1, \ldots, M_r\} \mapsto \Theta$ the decoder and $M_r \triangleq 2^r$ the number of quantization levels. Then, each $P \in \Theta$ will be approximated by

$$\hat{P}_r \triangleq \hat{g}_r(P),$$

and for each of the quantization levels $\hat{P}_r$, we define the corresponding output pmf

$$\hat{S}_{r,R_0} \triangleq \hat{P}_r \cdot W_{\hat{P}_r,R_0},$$

where W_{Q,R_0} is the minimizing channel in the distortion-rate problem for pmf Q and rate R_0. Thus, the minimum average distortion incurred by this two-stage VQ scheme, when the true pmf of the DMS is P, is

$$\tilde{D}(P, \hat{S}_{r,R-\frac{r}{n}}, R - \frac{r}{n}),$$

whereas, if we used an optimal code of rate R which is "matched" to P without having to describe the pmf, it would be $D(P, R)$. Therefore, we pay a distortion penalty for using the two-stage scheme, which equals

$$\Delta_n(P, r, R) \triangleq \tilde{D}(P, \hat{S}_{r,R-\frac{r}{n}}, R - \frac{r}{n}) - D(P, R).$$

At this juncture, note that, for a finite value of the block-length n, the minimum average distortion incurred by the two-stage VQ scheme should include an additional term which decays to zero as $O(\log n/n)$. However, this will affect the optimal choice of the pmf rate r_n, which should grow as $A \log n$, as is shown later in this section, only through the constant A. This qualification had been omitted in an earlier version of this thesis, and we thank Tamas Linder [36] for a careful reading of this chapter followed by a question which alerted us to this omission.

If we allow the pmf rate r_n to grow with the block-length n, then the approximated (or estimated) source pmf $\hat{P}_r$ will, under the best of circumstances, converge to P as n goes to infinity. Also, if the dependence of r_n on n is such that $\lim r_n/n = 0$, then as n goes to infinity, the rate needed to describe the source within distortion Δ converges to the rate-distortion function $R(P, \Delta)$; hence, the distortion penalty will decay to 0. There are three different kinds of universality (*c.f.*, *e.g.*, [10]). We say that a sequence of two-stage n-length block codes $\{\hat{g}_r^{(n)}\}$ is *weakly minimax* universal if

$$\lim_n \Delta_n(P, r, R) = 0,$$

for every $P \in \Theta$ and *strongly minimax* universal if the convergence is uniform in Θ. Furthermore, if we assume a prior w on Θ, then $\{\hat{g}_r^{(n)}\}$ is called *weighted* universal if

$$\lim_n \mathbf{E}_w \Delta_n(P, r, R) = 0.$$

The three kinds of universal convergence are not equivalent, but can be treated as such. To this end, note that weighted universality implies the existence of a $\Theta_1 \subset \Theta$ with arbitrarily small measure under w, on which the weighted-universal sequence of codes is also weakly minimax universal. Similarly, weakly minimax universality implies that the distortion penalty for $\{\hat{g}_r^{(n)}\}$ converges to zero uniformly for "most" of the $P \in \Theta$. We shall examine the weighted universal case, since it is the easiest.

For a given block-length n, we wish to choose r so as to minimize

$$\bar{\Delta}_n(r, R) \stackrel{\triangle}{=} \mathbf{E}_w \Delta_n(P, r, R).$$

In general, it is difficult to calculate $\tilde{D}(P, S, \Delta)$, and hence, $\bar{\Delta}_n(r, R)$. We shall, therefore, examine the (somewhat equivalent) problem of minimizing the excess rate. If we fix an acceptable distortion level Δ, then the minimum rate at which we can describe a source with pmf P is $R(P, \Delta)$. In the universal coding scheme, though, we allocate r bits for the description of the pmf and $n\tilde{R}(P, \hat{S}_{r,\Delta}, \Delta)$ bits for the data. Thus, the overall rate redundancy per source symbol is

$$\rho_n(P, r, \Delta) \stackrel{\triangle}{=} \frac{r}{n} + \tilde{R}(P, \hat{S}_{r,\Delta}, \Delta) - R(P, \Delta) = \frac{r}{n} + R_d(P, \hat{P}_r; \Delta). \tag{6.8}$$

Here, $\hat{S}_{r,\Delta_0} \stackrel{\triangle}{=} \hat{P}_r \cdot W_{\hat{P}_r,\Delta_0}$, with W_{Q,Δ_0} being the minimizing channel in the distortion-rate problem for pmf Q and distortion Δ_0. Our objective is to choose r so as to minimize

$$\bar{\rho}_n(r, \Delta) \stackrel{\triangle}{=} \mathbf{E}_w \rho_n(P, r, \Delta). \tag{6.9}$$

The existence of a unique value of r which minimizes the expected excess rate given by (6.9) is not always guaranteed. Observe that the first part on the right side of (6.8) is linearly increasing in r, whereas the second part (rate-distortion discrepancy) is decreasing in r. Indeed, as r grows larger the quantization of the pmf space becomes finer and the quantization level $\hat{P}_r \stackrel{\triangle}{=} \hat{g}_r(P)$ corresponding to some $P \in \Theta$ comes closer to P. If the rate-distortion discrepancy is continuous in $\hat{P}$ in some region of P (which may not be the case, as pointed out in Section 6.2.4), then the discrepancy $R_d(P, \hat{P}_r; \Delta)$ will converge to 0.

The minimizing value of r is given by setting the first derivative of $\bar{\rho}_n(r, \Delta)$ with respect to r equal to zero, *i.e.*,

$$\frac{\partial}{\partial r} \tilde{R}(P, \hat{S}_{r,\Delta}, \Delta) = -\frac{1}{n}. \tag{6.10}$$

It is possible that the partial derivative of the mismatched rate-distortion function crosses the line $y = -1/n$ on an XY plot more than once, so the different values of r resulting from (6.10) may correspond only to local minima. A sufficient condition for the existence of a unique r which minimizes (6.8) is that the mismatched rate-distortion function is convex in r. Then, its partial derivative with respect to r will be decreasing and it will assume the value $-1/n$ only once.

Although we know that $\tilde{R}(P, S, \Delta)$ is convex in S, in order to ensure convexity in r we have to assume that $\hat{S}_{r,\Delta}$ is also convex in r. This is true for a binary source with the Hamming distortion measure, but we have been unable to make a similar assertion for the general case, when we do not know the exact form of $\hat{S}_{r,\Delta}$.

If W_P^* is unique, and $\tilde{R}(P, \hat{S}_{r,\Delta}, \Delta)$ is differentiable in $\hat{P}$, then, by taking its Taylor expansion around P, we can show that the rate-distortion discrepancy is locally quadratic. Indeed, since the discrepancy achieves its minimum for $\hat{P} = P$, its first derivative should be zero at P, and what remains is the polynomial terms of order 2 and higher. We can, therefore, write the approximation

$$\rho_n(P, r, \Delta) \simeq \frac{r}{n} + (\hat{P} - P)^T C_1 (\hat{P} - P),$$

where C_1 is a matrix not depending on $\hat{P}$. This is consistent with the assumption in [10] that the rate-distortion discrepancy is locally quadratic in $\hat{P}$. The difference $\|\hat{P} - P\|$ is a function of r and P. If the quantizer $\hat{g}_r$ is based on a uniform scalar quantizer, then the L_∞ norm $\|\hat{P} - P\|_\infty$ will be inversely proportional to the number of quantization levels $M_r = 2^r$ of the pmf quantizer $\hat{g}_r$, which means that

$$\bar{\rho}_n(r, \Delta) \simeq \frac{r}{n} + C_2 2^{-2r}.$$

Minimizing this quantity with respect to r yields an optimal value given by

$$r_{opt} \triangleq \frac{1}{2}\log(2(\log 2)C_2 n),$$

as shown in [47] for noiseless codes and in [10] for more general setups.

It turns out, however, that in some cases (like the universal quantization of a binary source, presented in the next section) the expected discrepancy is not exactly quadratic, which is quite interesting in itself.

6.4 The Binary Case

We shall examine the case of a DMS with input and output alphabets $\mathcal{X} = \mathcal{Y} = \{0, 1\}$, and the Hamming distance distortion measure. This is the simplest case which can be studied, because the minimization problem in the definition of the mismatched distortion function turns trivial: the constraint set in (6.1) is either empty or contains exactly one element, namely the joint pmf which is the solution to the constraint equations. Note further that the source pmf depends only on one parameter, namely the probability of 0 (or 1). Thus, the pmf space to be considered is $\Theta = (0, 1)$, for which we know that partitioning into intervals is the optimal quantization scheme [6, 21, 60].

6.4.1 Mismatched Rate-Distortion Function

In this section, we determine the mismatched rate-distortion function $\tilde{R}(P, S, \Delta)$ for the binary case. Here, P is the input pmf, S the output pmf and $\Delta > 0$ an acceptable distortion level.

Let $p \triangleq \Pr\{X = 0\} = P(0)$, $s \triangleq \Pr\{Y = 0\} = S(0)$. Then the joint pmf

P_{XY} that satisfies

$$\begin{aligned} P_X(0) = P_{XY}(0,0) + P_{XY}(0,1) &= p \\ P_Y(0) = P_{XY}(0,0) + P_{XY}(1,0) &= s \\ d(P_{XY}) = P_{XY}(0,1) + P_{XY}(1,0) &= \Delta \\ \sum_{x,y} P_{XY}(x,y) &= 1 \end{aligned}$$

is fully determined by these conditions, and is given by the following table.

$P_{XY}(x,y)$	$y=0$	$y=1$
$x=0$	$\frac{p+s-\Delta}{2}$	$\frac{p-s+\Delta}{2}$
$x=1$	$\frac{-p+s+\Delta}{2}$	$\frac{2-p-s+\Delta}{2}$

Not all triplets $(p,s,\Delta) \in [0,1]^3$ are admissible: the condition is that the entries of the table be nonnegative. Considering p to be a parameter, s and Δ must satisfy

$$\begin{aligned} p \leq \; & s + \Delta \; \leq 2 - p \\ -p \leq \; & s - \Delta \; \leq p. \end{aligned}$$

Furthermore, since the mDRf is decreasing in R, the maximum distortion level occurs when $I(P_{XY}) = 0$, in which case the random variables X and Y are independent. Thus, $P_{XY} = PS$ and

$$\begin{aligned} \Delta_{max} &= \sum_{x,y} P(x)S(y)d(x,y) \\ &= p(1-s) + s(1-p) \\ &= p + s - 2sp, \end{aligned}$$

so that the required compression rate is 0 for

$$\Delta \geq (1-2p)s + p.$$

Summarizing, the mismatched rate-distortion function is given by

$$\tilde{R}(P,S,\Delta) = \begin{cases} 0, & \Delta \geq (1-2p)s + p \\ h(p) - sh\left(\frac{\Delta+s-p}{2s}\right) - (1-s)h\left(\frac{\Delta-s+p}{2(1-s)}\right), & |s-p| \leq \Delta \leq (1-2p)s + p. \end{cases}$$

The binary mismatched rate-distortion function is an upper bound for the rate-distortion function. Indeed, for $|s-p| \leq \Delta \leq (1-2p)s+p$, since $h(\cdot)$ is a concave function, we have

$$\begin{aligned} \tilde{R}(P,S,\Delta) &= h(p) - sh\left(\frac{\Delta+s-p}{2s}\right) - (1-s)h\left(\frac{\Delta-s+p}{2(1-s)}\right) \\ &\geq h(p) - h\left(\frac{\Delta+s-p}{2} + \frac{\Delta-s+p}{2}\right) \\ &= h(p) - h(\Delta) \\ &= R(P,\Delta), \end{aligned}$$

with equality iff

$$\frac{\Delta+s-p}{2s} = \frac{\Delta-s+p}{2(1-s)},$$

which results in the optimal output probability

$$s^*(p,\Delta) = \frac{p-\Delta}{1-2\Delta}, \quad p \in [\Delta, 1-\Delta]. \tag{6.11}$$

The behaviour of $\tilde{R}(P,S,\Delta)$ as a function of s and Δ is illustrated in Figure 6.1 for $p \leq 1/2$, where the various regions of the mismatched rate-distortion function are shown for the binary source on the (s,Δ) plane. $\tilde{R}(P,S,\Delta)$ is well-defined in the unhighlighted region and zero in the lighter shaded region. The solid curve corresponds to the optimal output pmf for a source governed by $P=(p,1-p)$, given by (6.11). When $p \geq 1/2$, the behaviour is similar, as can be seen in Figure 6.2.

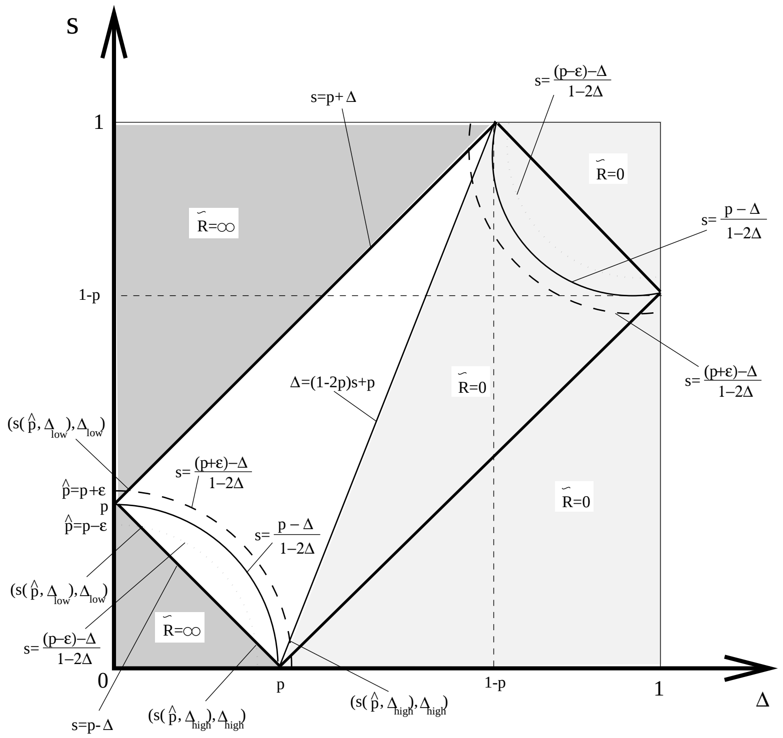

Figure 6.1: mRDf for the binary source as a function of the distortion and the output pmf, for $p < 1/2$.

If a code is designed to be optimal for some different input pmf, say $\hat{P} = (\hat{p}, 1-\hat{p})$, then s (the output probability of 0) is described by a similar curve (dashed for $\hat{p} < p$ and dotted for $\hat{p} > p$ in Figure 6.1). This, however, will not lie entirely in the well-defined region ($\Delta \geq |s-p|$). Hence, the rate-distortion discrepancy between P and $\hat{P}$, when $\hat{p} \in [\Delta, 1-\Delta]$ and $\hat{p}$ is not "too far" from p, is given by

$$R_d(P,\hat{P};\Delta) = \begin{cases} R_0, & \Delta > \Delta_{high} \\ h(\Delta) - s^* h\Big(\Delta - \frac{p-\hat{p}}{2s^*}\Big) & \\ \quad -(1-s^*) h\Big(\Delta + \frac{p-\hat{p}}{2(1-s^*)}\Big), & \Delta \in [\Delta_{low}, \Delta_{high}] \end{cases} \tag{6.12}$$

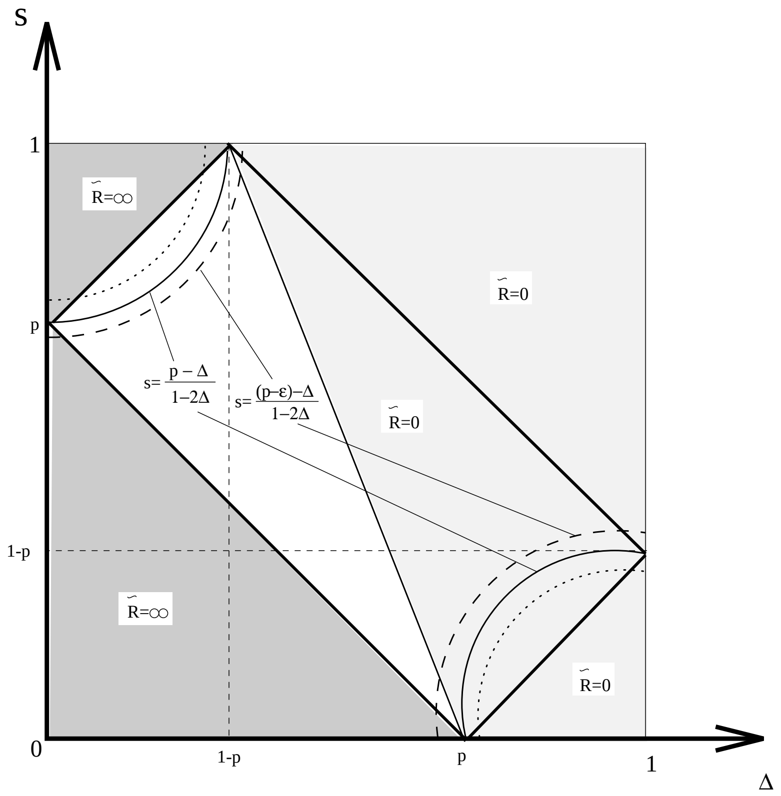

Figure 6.2: mRDf for the binary source as a function of the distortion and the output pmf, for $p > 1/2$.

where

$$\Delta_{low} \triangleq \begin{cases} \frac{1-p-\sqrt{1-2\hat{p}+p^2}}{2}, & \hat{p} \geq p \\ \frac{p-\sqrt{p^2-2p+2\hat{p}}}{2}, & \hat{p} \leq p \end{cases}, \; \Delta_{high} \triangleq \begin{cases} \begin{cases} \frac{1-\sqrt{(1-2p)(1-2\hat{p})}}{2}, & \hat{p} \geq p \\ \frac{p+\sqrt{p^2-2p+2\hat{p}}}{2}, & \hat{p} \leq p \end{cases}, & p \geq \frac{1}{2} \\ \begin{cases} \frac{1-p+\sqrt{1-2\hat{p}+p^2}}{2}, & \hat{p} \geq p \\ \frac{1+\sqrt{(1-2p)(1-2\hat{p})}}{2}, & \hat{p} \leq p \end{cases}, & p \leq \frac{1}{2} \end{cases}$$

are the intersections of the rate-distortion curve $s = (\hat{p} - \Delta)/(1 - 2\Delta)$ with the limiting curves $|s - p| = \Delta$ and $(1-2p)s + p = \Delta$. Also, the minimum achievable rate

$$R_0 \triangleq \tilde{R}(P, S^*(\hat{P}, \Delta_{high}), \Delta_{high}) - R(P, \Delta)$$

is 0 only for $(\hat{p}-p)(1/2-p) \geq 0$, as can be seen from Figures 6.1 and 6.2.

Finally, if $\hat{p} \in [0, \Delta)$, then $s^*(\hat{p}, \Delta) = 0$ and

$$\tilde{R}(P, S^*(\hat{P}, \Delta), \Delta) = 0, \quad p \leq \Delta,$$

whereas, if $\hat{p} \in (1-\Delta, 1]$, then $s^*(\hat{p}, \Delta) = 1$ and

$$\tilde{R}(P, S^*(\hat{P}, \Delta), \Delta) = 0, \quad 1-p \leq \Delta.$$

6.4.2 Universal Quantization of a Binary Source

Let us assume that we have $M \stackrel{\triangle}{=} 2^r$ quantization levels to describe the pmf of a binary source. Here $\Theta = (0, 1)$, the probability simplex in the one-dimensional space. Obviously, given the quantization levels $\hat{p}_i$, $i = 1, \ldots, M$, the Voronoi shell for each $\hat{p}_i$ will be an interval around $\hat{p}_i$, say $[\hat{p}_i - \epsilon_i^{(1)}, \hat{p}_i + \epsilon_i^{(2)}]$.

Assuming a prior pdf $w(p)$ on the pmf space Θ, the average performance (in terms of excess rate) of this scheme will be

$$\begin{aligned} \bar{\rho}_n(r, \Delta) &= \bar{\rho}_n(\log M, \Delta) \\ &= \frac{\log M}{n} + \int_0^1 w(p) R_d(p, \hat{g}_r(p); \Delta) dp \\ &= \frac{\log M}{n} + \sum_{i=1}^{M} \int_{\hat{p}_i - \epsilon_i^{(1)}}^{\hat{p}_i + \epsilon_i^{(2)}} w(p) R_d(p, \hat{p}_i; \Delta) dp. \end{aligned}$$

It is easy to verify for the binary source, that

$$R_d(p, \hat{p}; \Delta) = R_d(1-p, 1-\hat{p}; \Delta).$$

It is, therefore, obvious that, owing to the symmetry of the problem, we have to choose the quantization levels $\hat{p}_i$, $i = 1, \ldots, M$ to be symmetric around $1/2$. Thus,

$$\bar{\rho}_n(r, \Delta) = \frac{\log M}{n} + \sum_{i=1}^{M/2} \int_{\hat{p}_i - \epsilon_i^{(1)}}^{\hat{p}_i + \epsilon_i^{(2)}} [w(p) + w(1-p)] R_d(p, \hat{p}_i; \Delta) dp.$$

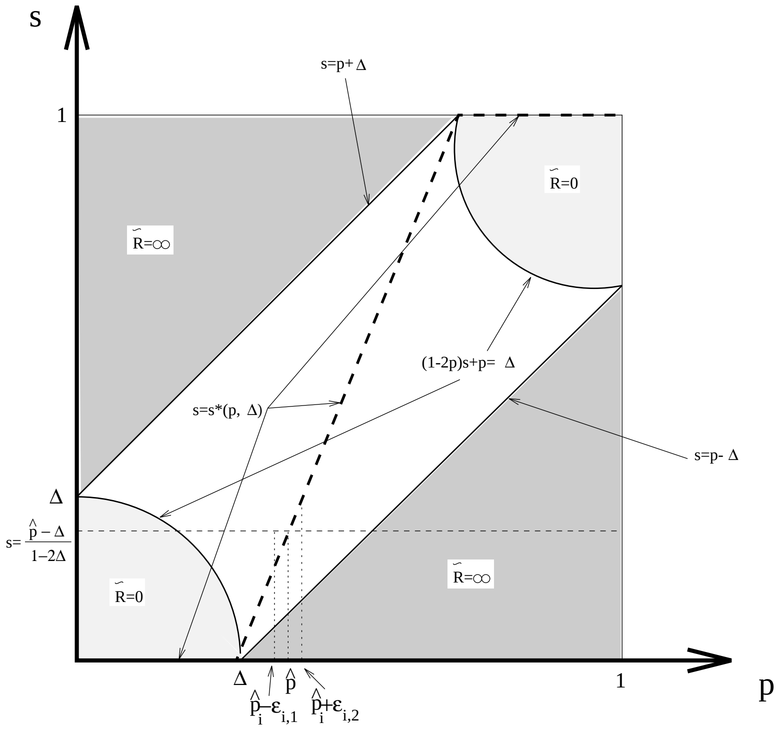

Figure 6.3: mRDf for binary source as a function of the input and output pmf's.

The behaviour of $\tilde{R}(p, s, \Delta)$ as a function of p and s is displayed in Figure 6.3. The constraint set in the definition of $\tilde{R}(P, S, \Delta)$ (6.1) is nonempty only within the strip defined by lines $p = s+\Delta$ and $p = s-\Delta$. The optimal output probability s of 0 for a particular input probability p of 0 is given by the three dashed line segments. When s is "matched" to some $\hat{p}$, the rate-distortion discrepancy is given by the points of the horizontal thin dashed line $s = \frac{\hat{p}-\Delta}{1-2\Delta}$.

Clearly, if $\hat{p}$ is greater than, but close to Δ, then $\tilde{R}(p, s^*(\hat{p}_1, \Delta), \Delta) = 0$ for

$$p \leq p_{min} \triangleq \frac{\Delta(1-2\Delta) - (p-\Delta)}{1-2\hat{p}} \simeq \Delta - (p - \Delta).$$

We shall, therefore, choose $\hat{p}_i, \ i = 1, \ldots, M/2$ uniformly spaced in the interval $[\Delta, 1/2]$, *i.e.*,

$$\hat{p}_i = \Delta + (i - \frac{1}{2})\frac{1-2\Delta}{M}, \quad i = 1, \ldots, M/2,$$

and

$$\epsilon_i^{(1)} = \epsilon_i^{(2)} = \epsilon(M) \triangleq \frac{1/2-\Delta}{M}, \quad i = 1, \ldots, M/2,$$

except for $\epsilon_1^{(1)}$, which is such that $\hat{p}_1 - \epsilon_1^{(1)} = 0$, *i.e.*, $\epsilon_1^{(1)} = \Delta + (1/2-\Delta)/M$, so that the first quantization interval is $[0, \Delta + \epsilon(M)]$.

This choice of $\hat{p}_i,\ i = 1, \ldots, M/2$ is most likely suboptimal, but it will provide a good upper bound for $\bar{\rho}_n(r, \Delta)$. Thus, setting $\bar{w}(p) \triangleq \frac{1}{2}w(p) + \frac{1}{2}w(1-p)$,

$$\begin{aligned}\bar{\rho}_n(r,\Delta) \quad &\leq \quad \frac{\log M}{n} + \sum_{i=1}^{M/2} \int_{\hat{p}_i - \epsilon(M)}^{\hat{p}_i + \epsilon(M)} 2\bar{w}(p) R_d(p, \hat{p}_i; \Delta) dp \\ &\qquad + \int_{p_{min}}^{\Delta} 2\bar{w}(p) \tilde{R}(p, s^*(\hat{p}_i, \Delta), \Delta) dp. \qquad (6.13)\end{aligned}$$

The rate-distortion discrepancy $R_d(p,\hat{p}_i;\Delta)$, for $|p-\hat{p}|$ small, is a locally quadratic function of the difference $p - \hat{p}$, as can be seen from (6.12). Indeed,

$$\begin{aligned}R_d(p, \hat{p}_i; \Delta) \quad &= \quad h(\Delta) - s_i^* h\left(\Delta - \frac{p-\hat{p}_i}{2s_i^*}\right) - (1-s_i^*) h\left(\Delta + \frac{p-\hat{p}_i}{2(1-s_i^*)}\right) \\ &= \quad -h''(\Delta) \frac{1}{s_i^*(1-s_i^*)} \frac{(p-\hat{p}_i)^2}{8} + o((p-\hat{p}_i)^2).\end{aligned}$$

If we assume that the prior $w(p)$ is sufficiently smooth, then

$$\begin{aligned}\bar{\rho}_n(r,\Delta) \quad &\leq \quad \frac{\log M}{n} + \sum_{i=1}^{M/2} 2\bar{w}(\hat{p}_i) \int_{-\epsilon(M)}^{\epsilon(M)} \left[\frac{-h''(\Delta)}{8s_i^*(1-s_i^*)}\beta^2 + o(\beta^2)\right] d\beta \\ &\qquad + 2\bar{w}(\Delta) \int_{p_{min}}^{\Delta} \tilde{R}(p, s^*(\hat{p}_i, \Delta), \Delta) dp \\ &= \quad \frac{\log M}{n} + \frac{\epsilon^3(M)}{6\Delta(1-\Delta)} \left[\sum_{i=1}^{M/2} \bar{w}(\hat{p}_i) \frac{1}{s_i^*(1-s_i^*)}\right] + o(\epsilon^3(M)) \\ &\qquad + \epsilon^2(M) \left[\log\frac{1-\Delta}{\Delta} + \frac{1-2\Delta}{6\Delta(1-\Delta)}\right] \bar{w}(\Delta).\end{aligned}$$

For the uniform prior $w(p) = 1$, we have

$$\begin{aligned}\bar{\rho}_n(r,\Delta) \quad &\leq \quad \frac{\log M}{n} + \frac{(1/2-\Delta)^3}{6\Delta(1-\Delta)} \frac{\log M + 1}{M^2} \\ &\qquad + \frac{(1-2\Delta)^2}{M^2} \left[\log\frac{1-\Delta}{\Delta} + \frac{1-2\Delta}{6\Delta(1-\Delta)}\right] + o(M^{-2}).\end{aligned}$$

Our objective is to determine the value of M (equivalently r) which minimizes $\bar{\rho}_n(r,\Delta)$. For large values of n, since r (and consequently M) grows with n, the optimal value of M will also be large. Thus, for large n,

$$\bar{\rho}_n(\log M,\Delta) \leq \frac{\log(M)}{n} + c(\Delta)\frac{\log(M)}{M^2} + o\left(\frac{\log(M)}{M^2}\right),$$

where

$$c(\Delta) \triangleq \frac{(1/2-\Delta)^3}{6\Delta(1-\Delta)},$$

and the optimal M is determined by solving the equation $\frac{\partial}{\partial M}\bar{\rho}_n(\log M,\Delta) = 0$, *i.e.*,

$$\frac{M^2}{\ln M - 1/2} = \frac{2}{\ln 2}c(\Delta)n,$$

or, in terms of the rate r,

$$\frac{2^{2r}}{r - \frac{1}{2\ln 2}} = 2c(\Delta)n.$$

For moderately large values of r (corresponding to a large value of n), the left side of the above equation behaves almost as 2^{2r}. Hence, for n sufficiently large,

$$\frac{1}{2}\log 2c(\Delta) + \frac{1}{2}\log n \leq r_{opt}^{(n)} \leq \frac{\log 2c(\Delta) + \log n}{2-\eta}, \quad \eta > 0.$$

Note : if $\Delta < 1/4$, then $\hat{p}_i + \epsilon(M)$ is in the undefined region (the darker shaded region in Figure 6.3) for $i < 1/4\Delta$, we must, therefore, adjust $\hat{p}_1,\ldots,\hat{p}_{\frac{1}{4\Delta}}$, but this will only result in a constant factor multiplying $c(\Delta)$, and not exceeding 4. The new $\hat{p}_i$'s will be

$$\hat{p}_i = \Delta + \frac{2-4\Delta}{1-4\Delta}\epsilon(M)(i-1), \quad i = 1,\ldots,\frac{1}{4\Delta}.$$

■

In order to verify the theoretically obtained results, we calculate numerically the optimal rate allocation for a binary source. The results are summarized in

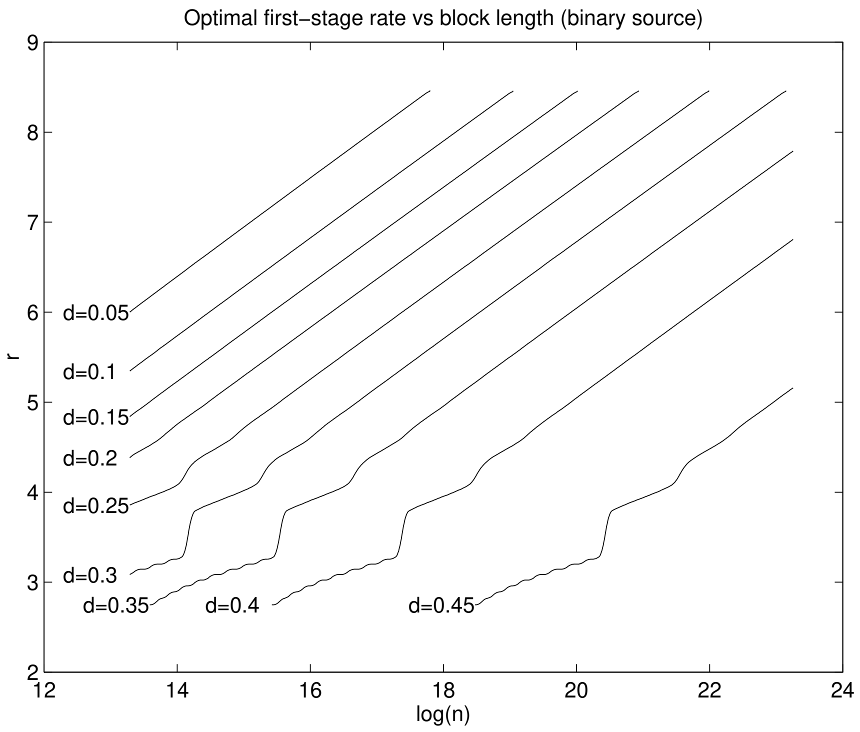

Figure 6.4: Optimal pmf rate *vs* block-length for a binary source.

Figures 6.4 and 6.5, where the optimal pmf rate r is plotted as a function of the logarithm of the block-length n for various values of the distortion level d (Δ in our theoretical calculations) and as a function of d for a few values of n.

These results were obtained by calculating the expected excess rate from (6.13) for different values of n, r and Δ and finding the value of r which corresponds to the minimum $\bar{\rho}_n(r, \Delta)$. The numerical results justify the approximations we made for large n. Thus, the curves of r *vs* $\log n$ in Figure 6.4, asymptotically look like straight lines of slope $1/2$ and the curves of $r - 0.5 \log n$ *vs* d, which show the dependence of the optimal rate on the distortion level Δ, are very close to the theoretically calculated "distortion overhead," represented by the dashed line in Figure 6.5.

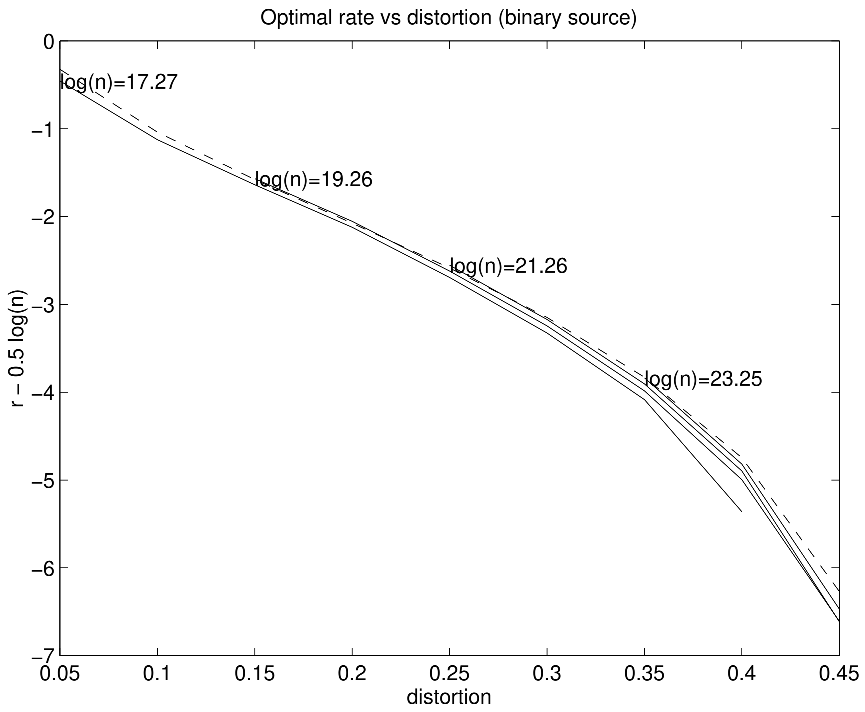

Figure 6.5: Optimal pmf-rate overhead *vs* distortion level for a binary source.

6.5 Discussion

We have addressed the problem of optimally allocating a prespecified number of bits in a two-stage universal vector quantization scheme. The main issue in the fixed-rate universal coding is to assess the performance of a code designed for a specific source pmf under a different pmf. Conjecture 6.1 proposes a reasonable answer to this question, however, we are unable at this point to provide a proof for the upper bound to the expected mismatch performance.

The proof of the upper bound is not essential for the selection of the optimal bit allocation scheme; all we need is the result of Theorem 6.2, which proves the existence of a code which has performance close to that predicted by the conjecture for every source pmf. Turning to the second stage of the universal code, the code which is matched to the quantized source pmf $\hat{P}$ can be selected to be this "universally optimal" code. Thus, the rate performance of the universal code will be given by (6.8).

In general, it is hard to compute the mismatched rate-distortion function in closed form. However, we can use a Blahut-like algorithm [9] to calculate its values for various choices of P, S and Δ. The optimal rate allocation scheme can be obtained by numerically minimizing (6.8).

For most cases, it is true that the pmf rate should grow linearly with the logarithm of the block-length. The general conditions under which the rate-distortion discrepancy is locally quadratic are still to be determined. Even for the simple binary case, it turns out that it is "subquadratic," although this does not greatly affect the asymptotic form of the optimal pmf rate.

Chapter 7

Conclusions

7.1 Brief Summary

In this dissertation, a number of issues have been addressed concerning the multi-level compression of information and its subsequent transmission in a multicasting application.

We begin with a study of the characteristics of good lossy codes and establish certain typicality properties. This constitutes the basis for our subsequent work on characterizing the performance of hierarchical codes. We determine the expected distortion redundancy for such codes by extending Pilc's [45] result for the convergence of the expected distortion of single-level codes to the distortion-rate function. We show that the rate of convergence for multi-level coding is of the same order, with the only difference being that more compression levels result in a larger constant multiplier.

The typicality properties of a good code have another important application in the evaluation of the performance of a mismatched lossy code. This is an important issue which, to the best of our knowledge, has not been addressed in the literature, although similar results do exist for different kinds of mismatch and for lossless coding. We introduce the notion of a mismatched distortion-rate function and establish that it provides a lower bound for the expected distortion of a mismatched code; further, this bound is shown to be achievable.

Next, we use the mismatched distortion-rate function as a performance criterion for optimal rate allocation between the two stages of a universal vector quantizer. This is a more precise and tractable measure than the abstract *lth-order divergence* used in [10]. Assuming suitable smoothness properties for the mismatched distortion-rate function, we can verify existing results on optimal rate-allocation under less restrictive conditions. The mismatched rate-distortion function is easily computable for a binary source with the Hamming-distance distortion measure; closed-form as well as numerical results are presented for this case.

We also address the issue of error exponents for hierarchical coding. It turns out that the error exponents in this case are obtained by minimizing the information divergence under certain constraints, similar to the single-level error exponent. The multi-level error exponent constitutes an interesting performance measure, and is probably more useful than the multi-level rate-distortion function. It is shown, by means of a simple example, that even when the multi-level rate-distortion function coincides with the corresponding single-level function, the overall multi-level error exponent may be worse (smaller) than the single-level exponent.

An interesting application of hierarchical data compression arises in multicasting over a network with link-capacity constraints. We use the multi-level distortion-rate function as a measure of the quality of the received signal, and address the problem of the optimal selection of routes and compression rates so as to minimize the maximum and the average distortions over all users. A simple algorithm is presented which solves the maximum distortion problem. The algorithmic complexity of the optimal solution for the average distortion problem, on the other hand, is prohibitively large. Therefore we consider a suboptimal scheme, based on a set of heuristics, with performance close to the optimal as indicated by simulation results.

7.2 Open Issues

In Chapter 2, we establish the exponential decay of the probability of the set of source sequences which are not jointly typical with their corresponding codewords, as the block-length of the code grows to infinity. However, we have not explicitly determined the dependence of the exponent on the corresponding parameters (*e.g.*, the "width" γ of the typical codeword subset $\mathcal{C}_n^T(\gamma)$). By referring to the proof of Lemma 2.2, it is easy to observe that the only point which needs to be clarified is the behaviour of the mismatched rate-distortion function $\tilde{R}(P, S, \Delta)$ around the minimizing output distribution $S = P \cdot W_P^*$. In general, it would be useful to characterize the sensitivity of $\tilde{R}(P, S, \Delta)$ with respect to its arguments, namely, to determine the precise manner in which variations in P, S and Δ translate to corresponding variations in the mismatched rate-distortion function, *i.e.*, to characterize

$$\tilde{R}(P + \delta P, S + \delta S, \Delta + \delta\Delta) - \tilde{R}(P, S, \Delta).$$

An observation which may prove useful, involves the relationship between the expected distortion induced by a block code and the corresponding error exponent. Let $g : \mathcal{X}^n \mapsto \mathcal{Y}^n$ represent the block code. Since the distortion between a random source sequence X^n and its corresponding codeword $g(X^n)$ is a positive random variable, the expected distortion can be written as

$$\mathbf{E}d(X^n, g(X^n)) = \int_0^{d_{max}} \Pr\{d(X^n, g(X^n)) > \delta\}\, d\delta.$$

The asymptotic behaviour of the error probability $\Pr\{d(X^n, g(X^n)) > \delta\}$ is completely characterized by the error exponent result [40] for $\delta > D(P, R)$, so that asymptotically it holds that

$$\mathbf{E}d(X^n, g(X^n)) \overset{\sim}{\leq} D(P, R) + \int_{D(P,R)}^{d_{max}} \exp\{-nF(P, R, \delta)\} d\delta.$$

If we are able to provide a qualitative description of the dependence of the error exponent function $F(P, R, \Delta)$ on R and Δ, it might be possible to obtain a better estimate of the asymptotic dependence of the distortion redundancy on the block length n. Observe that this approach would yield a lower bound on the expected distortion redundancy as well.

An extension to the multi-level universal code of the rate-allocation result for a single-level code is also of interest. For this situation, the optimal allocation of rates for multi-level codes should be done jointly for all compression levels. In order to extend our analysis to hierarchical codes, we must first characterize the mismatched multi-level rate-distortion function, $\tilde{R}(P, S, \Delta_1, \Delta_2, \ldots)$. Then, the optimization procedure will be along the same lines as in the single-level case.

The problem of performance evaluation of a mismatched code is not yet fully resolved. In particular, we have conjectured that the expected distortion redundancy of any good code, with some additional suitable structure, is close to the mismatched distortion-rate function. In order to show this, we need to study in detail the structure of a good S-typical code. For instance, its more subtle characteristics, such as its Voronoi shells and placement of their centroids, may have to be addressed.

Turning next to the multicasting problem, although the requirement that each network node is a user is valid in several practical applications, *e.g.*, in small local wireless networks, in general it is not a realistic assumption for large communication networks. In the latter situation, it is meaningful to study the case of a few users in a network with a large number of relay nodes. This is a simple extension of our work, but we need to provide appropriate numerical results to support the effectiveness of the proposed algorithms.

As has been pointed out in Chapter 5, the optimal route selection in Algorithms 5.1 and 5.2 is distributed in nature, in the sense that the multicasting paths can be set up at the network nodes using information from neighbouring nodes. The optimal rate selection, however, has to be centralized, since knowledge of the overall state of the network is necessary. This results in a lot of excess overhead traffic in the network from the nodes to the source, in reporting the states of the network links. This overhead can be reduced by devising smart methods of estimating the network state, which would then be used by the source based on minimal side information from the nodes. Furthermore, frequent changes of the link capacities would lead to different selections of optimal rates, causing the source to frequently redesign the hierarchical code. Instead, we could require the source to select from among a small number of rate-allocation schemes (along with the corresponding optimal codes). Although these techniques would result in suboptimal coding and transmission policies, they are obviously more appropriate for real-time applications.

Bibliography

[1] R. Ahlswede. The Rate-Distortion Region for Multiple Descriptions Without Excess Rate. *IEEE Transactions on Information Theory*, 31(6):721–726, November 1985.

[2] M. H. Amar, S. Y. Cheung, and C. M. Scoglio. Routing Multipoint Connections Using Virtual Paths in an ATM Network. In *Proceedings of IEEE Infocom'93*, pages 98–105, 1993.

[3] V. Anantharam. A Large Deviations Approach to Error Exponents in Source Coding and Hypothesis Testing. *IEEE Transactions on Information Theory*, 36(4):938–943, July 1990.

[4] F. Bauer and A. Varma. Degree-Constrained Multicasting in Point-to-Point Networks. In *Proceedings of IEEE Infocom'95*, pages 369–376, 1995.

[5] F. Bauer and A. Varma. Distributed Algorithms for Multicast Setup in Data Networks. *IEEE/ACM Transactions on Networking*, 4(2):181–191, April 1996.

[6] W. R. Bennett. Spectra of Quantized Signals. *Bell System Technical Journal*, 27:446–472, 1948.

[7] T. Berger. *Rate Distortion Theory: A Mathematical Basis for Data Compression.* Prentice-Hall, Englewood Cliffs, NJ, 1971.

[8] D. Bertsekas and R. G. Gallager. *Data Networks.* Prentice-Hall, NJ, 1987.

[9] R.E. Blahut. *Principles and Practice of Information Theory.* Reading, MA : Addison-Wessley, 1978.

[10] P. A. Chou, M. Effros, and R. M. Gray. A Vector Quantization Approach to Universal Noiseless Coding and Quantization. *IEEE Transactions on Information Theory*, 42:1109–1138, July 1996.

[11] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* New York: Wiley, 1991.

[12] I. Csiszár. *I*-Divergence Geometry of Probability Distributions and Minimization Problems. *The Annals of Probability*, 3(1):146–158, 1975.

[13] I. Csiszár and J. Körner. *Information Theory : Coding Theorems for Discrete Memoryless Systems.* Academic Press, NY, 1981.

[14] M. Doar and I. Leslie. How Bad is Naive Multicast Routing? In *Proceedings of IEEE Infocom'93*, pages 82–89, 1993.

[15] A. El-Gamal and T. M. Cover. Achievable Rates for Multiple Descriptions. *IEEE Transactions on Information Theory*, 28(6):851–857, November 1982.

[16] W. H. Equitz and T. M. Cover. Successive Refinement of Information. *IEEE Transactions on Information Theory*, 37(2):269–275, March 1991.

[17] W. H. Equitz and T. M. Cover. Addendum to "Successive Refinement of Information". *IEEE Transactions on Information Theory*, 39(4):1465–1466, July 1993.

[18] H. N. Gabow and K.S. Manu. Packing Algorithms for Arborescences (and Spanning Trees) in Capacitated Graphs. preprint, 1997.

[19] R. G. Gallager. Lower Bounds on the Tails of Probability Distributions. Quart. Progress Report 77, M.I.T. Research Laboratory of Electronics, April 1965.

[20] R. G. Gallager. *Information Theory and Reliable Communications.* Wiley, New York, 1968.

[21] A. Gersho. Asymptotically Optimal Block Quantization. *IEEE Transactions on Information Theory*, 25(4):373–380, July 1979.

[22] R. M. Gray. Conditional Rate-Distortion Theory. Technical Report 6502-2, Stanford Electronics Laboratory, Stanford University, Stanford, CA, October 1972.

[23] R. M. Gray. A New Class of Lower Bounds to Information Rates of Stationary Sources via Conditional Rate-Distortion Functions. *IEEE Transactions on Information Theory*, 19(4):480–489, July 1973.

[24] S. Hakimi. Steiner's Problem in Graphs and its Implications. *Networks*, 1:113–133, 1971.

[25] T. S. Han and S. Verdú. Approximation Theory of Output Statistics. *IEEE Transactions on Information Theory*, 39(3):752–772, May 1993.

[26] F. K. Hwang and D. S. Richards. Steiner Tree Problems. *Networks*, 22:55–89, 1992.

[27] A. Kanlis. Multicasting of hierarchically compressed information in a network. In *IEEE International Symposium on Information Theory 1997*, 1997.

[28] A. Kanlis, S. Khudanpur, and P. Narayan. Typicality of a Good Rate-Distortion Code. *Problems of Information Transmission (Problemy Peredachi Informatsii)*, January 1996. Special issue in honor of M. S. Pinsker.

[29] A. Kanlis and P. Narayan. Error Exponents for Successive Refinement by Partitioning. *IEEE Transactions on Information Theory*, 42(1):275–282, January 1996.

[30] A. H. Kaspi and T. Berger. Rate-Distortion for Correlated Sources with Partially Separated Encoders. *IEEE Transactions on Information Theory*, 28(6):828–840, November 1982.

[31] V. P. Kompella, J. C. Pasquale, and G. C. Polyzos. Multicasting for Multimedia Applications. In *Proceedings of IEEE Infocom'92*, pages 2078–2085, 1992.

[32] V. Koshélev. Hierarchical Coding of Discrete Sources. *Problemy Peredachi Informatsii*, 16(3):31–49, 1980.

[33] V. Koshélev. Estimation of Mean Error for a Discrete Successive-Approximation Scheme. *Problemy Peredachi Informatsii*, 17(3):20–33, July-September 1981.

[34] L. Kou, G. Markowsky, and L. Berman. A Fast Algorithm for Steiner Trees. *Acta Informatica*, 15:141–145, 1981.

[35] A. Lapidoth. On the Role of Mismatch in Rate-Distortion Theory. *IEEE Transactions on Information Theory*, 43(1):38–47, January 1997.

[36] T. Linder. personal communication.

[37] T. Linder, G. Lugosi, and K. Zeger. Rates of Convergence in the Source Coding Theorem, in Empirical Quantizer Design, and in Universal Lossy Source Coding. *IEEE Transactions on Information Theory*, 42(6):1728–1740, November 1994.

[38] T. Linder, G. Lugosi, and K. Zeger. Fixed Rate Universal Lossy Source Coding and Rates of Convergence for Memoryless Sources. *IEEE Transactions on Information Theory*, 41(3):665–676, May 1995.

[39] T. D. Lookabaugh and R. M. Gray. High-Resolution Quantization Theory and the Vector Quantizer Advantage. *IEEE Transactions on Information Theory*, 35(5):1020–1033, September 1989.

[40] K. Marton. Error Exponent for Source Coding with a Fidelity Criterion. *IEEE Transactions on Information Theory*, 20(2):197–199, March 1974.

[41] K. Marton. A Simple Proof of the Blowing-Up Lemma. *IEEE Transactions on Information Theory*, 32(3):445–446, May 1986.

[42] J. K. Omura. Coding Theorem for Discrete-Time Sources. *IEEE Transactions on Information Theory*, 19(4):490–497, July 1973.

[43] J. K. Omura. A Lower Bounding Method for Channel and Source Coding Probabilities. *Information and Control*, 27:148–177, 1975.

[44] L. Ozarow. On a Source-Coding Problem with Two Channels and Three Receivers. *Bell System Technical Journal*, 59(10):1909–1921, December 1980.

[45] R. J. Pilc. The Transmission Distortion of a Source as a Function of the Encoding Block Length. *Bell System Technical Journal*, 47(6):827–885, July-August 1968.

[46] B. Rimoldi. Successive Refinement of Information : Characterization of the Achievable Rates. *IEEE Transactions on Information Theory*, 40(1):253–259, January 1994.

[47] J. Rissanen. Universal Coding, Information, Prediction and Estimation. *IEEE Transactions on Information Theory*, 30(4):629–636, July 1984.

[48] T. R. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

[49] N. Shacham. Multipoint Communication by Hierarchically Encoded Data. In *Proceedings of IEEE Infocom'92*, pages 2107–2114, 1992.

[50] N. Shacham and J. S. Meditch. An Algorithm for Optimal Multicast of Multimedia Streams. In *Proceedings of IEEE Infocom'94*, 1994.

[51] C. E. Shannon. unpublished work.

[52] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948.

[53] M. L. Shore, L. R. Foulds, and P. B. Gibbons. An Algorithm for the Steiner Tree Problem in Graphs. *Networks*, 12:323–333, 1982.

[54] B. Waxman. Routing of Multipoint Connections. *IEEE Journal on Selected Areas in Communications*, 6:1617–1622, December 1988.

[55] H. S. Witsenhausen and A. D. Wyner. Source Coding for Multiple Descriptions II: A Binary Source. *Bell System Technical Journal*, 60(10):2281–2292, December 1981.

[56] J. K. Wolf, A. D. Wyner, and J. Ziv. Source Coding for Multiple Description. *Bell System Technical Journal*, 60(10):2281–2292, December 1981.

[57] A. D. Wyner and J. Ziv. The Rate-Distortion Function for Source Coding with Side Information at the Decoder. *IEEE Transactions on Information Theory*, 22(1):1–10, January 1976.

[58] H. Yamamoto. Source Coding Theory for Cascade and Branching Communication Systems. *IEEE Transactions on Information Theory*, 27(3):299–308, May 1981.

[59] B. Yu and T. P. Speed. A rate of convergence result for a universal d-semifaithful code. *IEEE Transactions on Information Theory*, 39(3):813–820, May 1993.

[60] P. Zador. Asymptotic Quantization Error of Continuous Signals and the Quantization Dimension. *IEEE Transactions on Information Theory*, 28(2):139–149, March 1982.

[61] K. Zeger, A. Bist, and T. Linder. Universal Source Coding with Codebook Transmission. *IEEE Transactions on Communications*, 42(2/3/4):336–346, February/March/April 1994.

[62] Z. Zhang and T. Berger. New Results in Binary Multiple Descriptions. *IEEE Transactions on Information Theory*, 33(4):502–521, July 1987.

[63] Z. Zhang, E. Yang, and V. W. Wei. The Redundancy of Source Coding with a Fidelity Criterion – Part One: Known Statistics. *IEEE Transactions on Information Theory*, 43(1):71–91 , January 1997.

[64] Q. Zhu, M. Parsa, and J.J. Garcia-Luna-Aceves. A Source-Based Algorithm for Delay-Constrained Minimum-Cost Multicasting. In *Proceedings of IEEE Infocom'95*, pages 377–385, 1995.