Joint Source–Channel Coding: Development of Methods and Utilization in Image Communications

Helge Coward

A dissertation submitted in partial fulfillment
of the requirements for the degree of

DOKTOR INGENIØR

Department of Telecommunications
Faculty of Electrical Engineering and Telecommunications
Norwegian University of Science and Technology

2001

Norwegian University of Science and Technology
Department of Telecommunications
N-7491 Trondheim
Norway

Report no. 420106

ISBN 82-471-5341-6
ISSN 0809-103X

Abstract

In a traditional communication system, the coding process is divided into source coding and channel coding. Source coding is the process of compressing the source signal, and channel coding is the process of error protection. It can be shown that with no delay or complexity constraints and with exact knowledge of the source and channel properties, optimal performance can be obtained with separate source and channel coding. However, joint source–channel coding can lead to performance gains under complexity or delay constraints and offer robustness against unknown system parameters.

Multiple description coding is a system for generating two (or more) descriptions of a source, where decoding is possible from either description, but decoding of higher quality is possible if both descriptions are available. This system has been proposed as a means for joint source–channel coding. In this dissertation, the multiple description coding is used to protect against loss of data in an error correcting code caused by a number of channel errors exceeding the correcting ability of the channel code. This is tried on three channel models: a packet erasure channel, a binary symmetric channel, and a block fading channel, and the results obtained with multiple description coding is compared against traditional single description coding. The results show that if a long-term average mean square error distortion measure is used, multiple description coding is not as good as single description coding, except when the delay or block error rate of the channel code is heavily constrained.

A direct source–channel mapping is a mapping from amplitude continuous source symbols to amplitude continuous channel symbols, often involving a dimension change. A hybrid scalar quantizer–linear coder (HSQLC) is a direct source–channel mapping where the memoryless source signal is quantized using a scalar quantizer. The quantized value is transmitted on an analog channel using one symbol which can take as many levels as the quantizer, and the quantization error is transmitted on the same channel by means of a simple linear coder. Thus, there is a bandwidth expansion, two channel symbols are produced per source symbol. The channel is assumed to have additive white

Gaussian noise and a power constraint. The quantizer levels and the distribution of power between the two symbols are optimized for different source distributions. A uniform quantizer with an appropriate step size gives a performance close to the optimized quantizer both for a Gaussian, a Laplacian, and a uniform memoryless source. The coder performs well compared to other joint source–channel coders, and it is relatively robust against variations in the channel noise level.

A previous image coder using direct source–channel mappings is improved. This coder is a subband coder where a classification following the decorrelating filter bank assigns mappings of different rates to different subband samples according to their importance. Improvements are made to practically all the parts of the coder, but the most important one is that the mappings are changed, and particularly, the bandwidth expanding HSQLC is introduced. The coder shows large improvements compared to the previous version, especially at channel qualities near the design quality. For poor channels or high rates, the HSQLC provides a large portion of the improvement. The coder is compared against a combination of a JPEG 2000 coder and a good channel code, and the performance is competitive with the reference, while the robustness against an unknown channel quality is largely improved. This kind of robustness is very important in broadcasting and mobile communications.

Preface

This dissertation is submitted in partial fulfillment of the requirements for the degree of *doktor ingeniør* at the Faculty of Electrical Engineering and Telecommunications, Norwegian University of Science and Technology (NTNU). The advisor has been Professor Tor A. Ramstad at the Department of Telecommunications, NTNU.

The work has taken place in the period from January 1997 to September 2001. In addition to the research work reported in this dissertation, the degree includes compulsory courses corresponding to full-time studies in two semesters, and the work also includes the equivalent of thirteen months as a teaching assistant. In the period from June to September 2000, I visited the Swiss Federal Institute of Technology (EPFL) in Lausanne, where I did most of the work on multiple description coding (Chapter 2). The rest of the period I have spent at the Department of Telecommunications, NTNU.

For a period of three years, I have been funded by a scholarship from the Research Council of Norway (NFR), under project number 116389/410. For the visit to EPFL, I received another scholarship from NFR, under project number 138642/410. The teaching assistant work was funded by the Department of Telecommunications, NTNU. After my ordinary funding ended in January 2001, I have been hired part-time by Nera Research while being allowed to finish the dissertation. After the dissertation is finished, I will be full-time employed at Nera Research.

Acknowledgments

The work required to obtain a doctoral degree can hardly be done without the help of others. First, I would like to thank my advisor, Professor Tor A. Ramstad. His advice has helped me identifying the problems and looking beyond the results for interpretation and investigation of new topics. He always welcomes questions, and takes the time for careful discussions (sometimes leading to topics far from those of signal processing).

My thanks also go to my other colleagues, past and present, at the Department of Telecommunications. Especially, I would like to thank my former office mate, Dr. Are Hjørungnes. He has been a great inspiration for going into problems that at first sight can look very hard, and discussions with him, both on scientific and other matters, have been very interesting, usually loud, and often fruitful. He has also contributed to proofreading of this dissertation.

In connection with my stay in Lausanne, I would like to thank Dr. Sergio Servetto, first assistant at the Audiovisual Communications Lab., EPFL (now: assistant professor at Cornell University, USA) for introducing me to the field of multiple description coding and for collaboration and help on the work on this topic throughout my stay. I would also like to thank Dr. Raymond Knopp, first assistant at the Mobile Communications Lab., EPFL (now: assistant professor at Institut Eurécom, France) for help with topics on channel coding, especially on the block fading channel. My thanks also go to Professor Martin Vetterli for inviting me and to the rest of the staff at the Audiovisual and Mobile Communications Labs.

Finally, I am very grateful to Nera Research for allowing me to finish the dissertation while being their employee.

Trondheim, September 2001
Helge Coward

Contents

List of Figures

List of Tables

Notation and Symbols

$'$	Differentiation operator
$\lvert\cdot\rvert$	1) Absolute value 2) Cardinality of a set
$\lceil\cdot\rceil$	Rounding towards infinity
$\binom{n}{k}$	Binomial coefficient
α	Exponentially distributed fading state variable
α_l	Exponentially distributed fading state variable in channel number l
$\alpha(\cdot)$	Encoder mapping
$\beta(\cdot)$	Decoder mapping
$\beta_i(\cdot,\cdot)$	HSQLC receiver function on channel branch i
$\tilde{\beta}_i(\cdot)$	HSQLC one-dimensional receiver function on channel branch i
δ	Step size of uniform quantizer
δ_a	Step size of a_i in uniform quantizer
δ_c	Step size of c_i in uniform quantizer
δ_d	Step size of d_i in uniform quantizer
δ_e	Step size of e_i in uniform quantizer
$\delta(\cdot)$	Difference function
Δ	Width of uniform source probability distribution
Δ_P	Minimum distance of channel symbols in PCCOVQ
ϵ	Small positive number
ζ_a	First order constant for a_i in quadratic quantizer
ζ_c	First order constant for c_i in quadratic quantizer

ζ_d	First order constant for d_i in quadratic quantizer
ζ_e	First order constant for e_i in quadratic quantizer
κ	HSQLC quantization error symbol receiver constant
λ	Lagrange multiplier
$\tilde{\lambda}$	Lagrange multiplier
λ_P	Lagrange multiplier in PCCOVQ
Λ	Lattice
ν	Number of bits (or other prime symbol) in one symbol of a block code
ξ_a	Second order constant for a_i in quadratic quantizer
ξ_c	Second order constant for c_i in quadratic quantizer
ξ_d	Second order constant for d_i in quadratic quantizer
ξ_e	Second order constant for e_i in quadratic quantizer
σ_B^2	Variance of HSQLC quantization error symbol
σ_i^2	Variance of subsource i
$\bar{\sigma}_j^2$	Variance of all source symbols mapped to class j
σ_N^2	Noise variance
σ_X^2	Source variance
σ_Y^2	Channel power constraint
$\sigma_{Y_i}^2$	Variance of channel symbol in branch i
a	Multiple description high rate quantization parameter
$\hat{a}(k)$	Observed decoded quantized HSQLC symbol at time instant k
$A(k),\ A$	Random quantized HSQLC symbol (at time instant k)
$\hat{A}(k),\ \hat{A}$	Random decoded quantized HSQLC symbol (at time instant k)
a_i	HSQLC representation value no. i
$\mathcal{A}_j^{(l)}$	Quantization region no. j of description no. l
$\boldsymbol{A}$	BPAM encoder matrix
$\underset{i}{\operatorname{argmax}}\, a_i$	The index i that maximizes a_i
b	Observed HSQLC quantization error symbol

$\hat{b}(k)$	Observed decoded HSQLC quantization error symbol at time instant k
$B(k),\ B$	Random HSQLC quantization error symbol (at time instant k)
$\hat{B}(k),\ \hat{B}$	Random decoded HSQLC quantization error symbol (at time instant k)
$\boldsymbol{B}$	BPAM decoder matrix
C	Channel capacity
$C(S)$	Channel capacity function
$C_a(\cdot)$	Function for finding the channel value from the quantized value in an HSQLC system
$C_x(\cdot)$	HSQLC channel quantizer function
c_i	HSQLC discrete channel value no. i
D	Distortion (mean square error)
$\tilde{D}(\tilde{R})$	Distortion-rate function
d	Distortion with single description coder without block losses
d_0	Central distortion
$d_1,\ d_2$	Side distortions
$d_{1,\max},\ d_{2,\max}$	Side distortion constraints
D_{MD}	Distortion with multiple description coding system
D_{SD}	Distortion with single description coding system
d_q	Distortion obtained with quantizer
$D_j(P_j)$	Distortion on a unit variance source with mapping j using channel power P_j
$\hat{D}_i(\hat{v}_i)$	Distortion by quantizing subsource i with quantizer $\hat{v}_i$ in bit allocation based system
$\tilde{D}_j(\bar{\sigma}_j^2, P_j)$	Distortion on a source of variance $\bar{\sigma}_j^2$ with mapping j using channel power P_j
D_{tot}	Total average distortion
$\hat{D}_{\mathrm{tot}}$	Total average distortion of bit allocation based system
d_i	HSQLC quantizer decision level no. i

$\mathcal{D}_i$	HSQLC quantization interval no. i
$d_{\min}$	Minimum distance
$\hat{d}_{\min}$	Lower bound on minimum distance
$E[\cdot]$	Expectation operator
e_i	HSQLC receiver detection level no. i
$\mathcal{E}_i$	HSQLC receiver detection interval no. i
erf	Error function
$f_\alpha(\alpha)$	Fading state probability density function
$f_B(b)$	Probability density function of quantization error function
$f_N(n)$	Noise probability density function
$f_X(x)$	Source probability density function
$f_{X,Y,\ldots}(x, y, \ldots)$	Generally: Joint probability density function of random variables $X, Y, \ldots$
$f_{X,\ldots\|V,\ldots}(x, \ldots \mid v, \ldots)$	Generally: Joint probability density function of random variable(s) $X, \ldots$ conditional on $V, \ldots$
$F_i(z)$	Transfer function of synthesis filter number i
$G(\cdot)$	Normalized second moment
$h(\cdot)$	Differential entropy
H_q	Entropy obtained with scalar quantizer
$H_q^{(l)}$	Entropy obtained with scalar quantizer for description no. l
$H_i(z)$	Transfer function of analysis filter number i
i	Index
i_{MAP}	Decoded index with MAP receiver
$I_i^{(0)}$	Integral defined by Equation (B.21)
$I_i^{(0-)}$	$I_i^{(0)}$ when integration takes place over negative values
$I_i^{(0+)}$	$I_i^{(0)}$ when integration takes place over positive values
$I_i^{(0\mp)}$	$I_i^{(0)}$ when integration range contains both negative and positive values
$I_i^{(1)}$	Integral defined by Equation (B.24)
$I_i^{(1-)}$	$I_i^{(1)}$ when integration takes place over negative values

$I_i^{(1+)}$	$I_i^{(1)}$ when integration takes place over positive values
$I_i^{(1\mp)}$	$I_i^{(1)}$ when integration range contains both negative and positive values
I_A	Capacity of fading channel given state
$\boldsymbol{I}_{K\times L}$	$K \times L$ matrix with ones on the main diagonal and zeros elsewhere
$\mathcal{I}_j$	Set of subsources assigned to class j
j	Index
J	Number of direct source-channel mappings in image coder
k	1) Time index 2) Horizontal pixel index 3) Number of data symbols in a block code
K	Channel dimension
K_a	HSQLC proportionality constant between a_i and c_i
K_b	HSQLC quantizer error scaling constant
l	1) Index 2) Vertical pixel index
L	1) Source dimension 2) Number of source symbols per block
L_{MD}	Number of source symbols per block with multiple description coding
L_{SD}	Number of source symbols per block with single description coding
$\mathcal{L}$	Lagrangian function
$\tilde{\mathcal{L}}$	Lagrangian function
m	Side size of multiple description index assignment matrix
M	1) Number of bits per block produced by source coder 2) Size of quantizer codebook 3) Number of subbands in parallel filter bank

M_{MD}	Number of bits per block produced by multiple description source coder
M_{SD}	Number of bits per block produced by single description source coder
n	1) Number of symbols per block in a block code 2) Observed scalar noise symbol
n_{MD}	Number of channel symbols per block with multiple description coding
n_{SD}	Number of channel symbols per block with single description coding
$n_{\max}$	Maximum block length
N	Number of subsources
$N(k)$, N	Random noise symbol (at time instant k)
$\boldsymbol{N}(k)$	Random noise vector at time instant k
$n_i(k)$, n_i	Observed noise symbol on channel branch no. i (at time instant k)
$N_i(k)$, N_i	Random noise symbol on channel branch no. i (at time instant k)
$\boldsymbol{N}_i(k)$	Random noise vector in channel branch i at time instant k
n_{diag}	Number of diagonals in index assignment matrix
N_q	Number of cells at each side of the origin in scalar quantizer
p	1) Packet erasure probability 2) Prime as basis for Reed-Solomon code
p_b	Bit error probability
p_e	Block error probability
$p_{e,\mathrm{MD}}$	Block error probability with multiple description coding system
$p_{e,\mathrm{SD}}$	Block error probability with single description coding system
$p_{i,j}$	Channel transition probability
$\tilde{p}_i$	Probability of quantizer cell no. i

$\tilde{p}_j^{(l)}$	Probability of quantizer region no. j of description no. l
p_{out}	Channel outage probability
$p_A(a)$	Probability mass function of quantized HSQLC symbol
$p_{A\|X,\dots}(a \mid x, \dots)$	Generally: Probability mass function of A conditional on $X, \dots$
P	Recieved channel power in block fading channel
P_j	Channel power of mapping j
P_{PAM}	Channel power of direct PAM symbol
P_{obs}	Observed average channel power in power allocation algorithm
P_{tot}	Average available channel power
$\Pr(\cdot)$	Probability of an event
q	Alphabet size of block code
$Q(\cdot)$	Quantizer function
R	Total rate (number of channel symbols per source symbol)
$\tilde{R}$	Information theoretic rate
R_j	Rate of mapping j
$\hat{R}_i(\hat{v}_i)$	Rate when quantizing subsource i with quantizer $\hat{v}_i$ in bit allocation based system
R_c	Channel bit rate
R_s	Source bit rate
$R_s^{(l)}$	Bit rate of description no. l
R_{sum}	Cumulative rate in rate allocation algorithm
R_{tot}	Overall available rate for image coder
$\hat{R}_{\text{tot}}$	Overall available bit rate for bit allocation based system
$\mathbb{R}$	The set of real numbers
s	Number of bits per PAM symbol in side information transmission
S	CSNR
S_L	Sphere in L dimensions

S_x	Horizontal image size
S_y	Vertical image size
t	1) Correcting ability of block code 2) General variable
^{T}	Transpose
$u(\cdot)$	Unit step function
$\hat{u}(k)$	General decoded symbol at time instant k
$U(k)$	General random symbol at time instant k
$v(k)$	General observed received symbol at time instant k
v_i	Class to which subsource i has been mapped
$\hat{v}_i$	Quantizer used to quantize subsource i in bit allocation based system
x	Observed scalar source symbol
$X(k),\ X$	Random scalar source symbol (at time instant k)
$\boldsymbol{X}(k)$	Random source vector at time instant k
$\hat{\boldsymbol{X}}(k)$	Random decoded vector at time instant k
$\hat{X}(k)$	Random decoded scalar source symbol at time instant k
$X_i(k)$	ith element of $\boldsymbol{X}(k)$
$\tilde{X}(k)$	Decoded random signal with single description source decoder
$\tilde{X}_0(k)$	Decoded random signal with central decoder
$\tilde{X}_1(k),\ \tilde{X}_2(k)$	Decoded random signal with side decoder
$\bar{X}$	Expected value of random signal $X(k)$
$\hat{x}_i$	Centroid of quantizer interval no. i
$\hat{x}_j^{(l)}$	Centroid of quantizer region no. j of description no. l
$x(k,l)$	Original pixel value at position (k,l)
$\hat{x}(k,l)$	Decoded pixel value at position (k,l)
$x_{\max}$	Maximum possible pixel value
$x_{\min}$	Minimum possible pixel value
$Y(k)$	Random scalar channel symbol at time instant k
$\boldsymbol{Y}(k)$	Random channel vector at time instant k

$\hat{Y}(k)$	Random received scalar symbol at time instant k
$\hat{\boldsymbol{Y}}(k)$	Random received vector at time instant k
$\hat{y}_i(k)$, $\hat{y}_i$	Observed received symbol of channel branch no. i (at time instant k)
$Y_i(k)$, y_i	Random channel symbol of channel branch no. i (at time instant k)
$\boldsymbol{Y}_i(k)$	Random channel vector of channel branch no. i at time instant k
$\hat{Y}_i(k)$, $\hat{Y}_i$	Random received symbol of channel branch no. i (at time instant k)
$\hat{\boldsymbol{Y}}_i(k)$	Random received vector of channel branch no. i at time instant k
z	The variable in the Z transform
$\mathbb{Z}$	The set of integers

Abbreviations

AWGN	Additive white Gaussian noise
BPAM	Block pulse amplitude modulation
BPSK	Binary phase shift keying
BSC	Binary symmetric channel
CELP	Code-excited linear predictive
CSNR	Channel signal-to-noise ratio
dB	Decibel
DPCM	Differential pulse code modulation
HDA-FEAD	Hybrid digital–analog—fixed encoder, adaptive decoder
HSQLC	Hybrid scalar quantizer–linear coder
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet protocol
ISO	International Standards Organization
JPEG	Joint Photographic Experts Group
LDPC	Low density parity check
MAP	Maximum a posteriori
MD	Multiple description
MDS	Maximum distance separable
MSE	Mean square error
MUX	Multiplexer
OPTA	Optimal performance theoretically attainable
PAM	Pulse amplitude modulation
PCCOVQ	Power-constrained channel-optimized vector quantization
pdf	Probability density function
PR	Perfect reconstruction
PSK	Phase shift keying
PSNR	Peak signal-to-noise ratio
QAM	Quadrature amplitude modulation
RMS	Root mean square

SD	Single description
SNR	Signal-to-noise ratio
TCM	Trellis coded modulation
UEP	Unequal error protection

Chapter 1

Introduction

In the recent years, the area of telecommunication has gone through a large development, both scientifically, technically, and financially. This area has also caught a lot of public attention, perhaps more than any other technological field. The use of wireless services, such as mobile phones, has gone through a tremendous growth, and new services are being developed rapidly. In these new services, there is a high demand for multimedia services, including the transmission of images, video, and high quality audio. Such sources have a high information rate, while the capacity of the wireless channels is limited.

Both compression of the source signals (source coding) and efficient utilization of the channels (channel coding) are topics that are subject to extensive research. Still, it is not always possible to obtain the desired quality of the received signals given the channel constraints that exist. In order to get the most out of the limited resources, new methods should be sought in order to come as close to the performance bounds as possible. Despite good results obtained with separate source and channel coding, unification of these operations can give improvements in some cases. Thus, joint source–channel coding is a concept that deserves further attention. This topic has received increasing attention, although far from that of separate source or channel coding.

In this dissertation, two fundamentally different joint source–channel coding methods are investigated. One is based on multiple description coding, the other is a hybrid digital–analog scheme. The most successful of these, the hybrid digital–analog system, is utilized in an image coder.

1.1 Source and Channel Coding

Figure 1.1 shows a typical block diagram of a digital communication system for transmitting a source signal. There are several ways of sub-dividing the

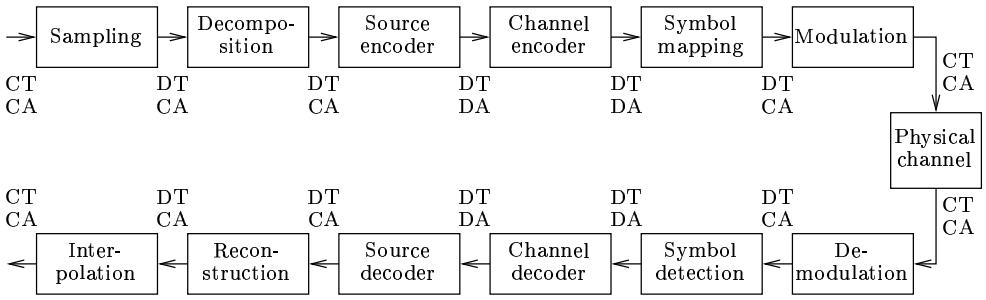


Figure 1.1 Typical blocks of a digital transmission system. The letters at each conjunction indicates the nature of the signal transferred between the blocks. D: Discrete. C: Continuous. T: Time. A: Amplitude

system into blocks in such a schematic overview. Not all systems contain all the blocks, and there might be operations performed in some systems that do not fit well into the framework of Figure 1.1.

The blocks right above each other perform converse operation. The sampling block performs the conversion from a time continuous to a time discrete signal. It is well known from the sampling theorem (Shannon, 1949) that the sampling/interpolation operations can be performed losslessly for a band-limited signal, provided that the sampling frequency is at least twice the bandwidth. The sampling operation creates a signal that is time discrete. In practice, sampling is usually combined with analog-to-digital conversion, creating an amplitude discrete signal. The unavoidable quantization noise introduced in this process is often small compared to the distortion introduced by the rest of the system, and thus, the signal can be viewed as an amplitude continuous signal. The decomposition block performs some sort of frequency analysis, for instance by linear prediction, or transform or subband filtering, in order to reduce or remove the correlation that is present in most natural signals. The source encoder performs such operations as quantization, bit allocation and entropy coding in order to compress the signal, and produces a bit stream (or stream of other finite alphabet symbols). Source coding is often used as a term for the operation of the two blocks last mentioned. In some systems, such as with vector quantization designed for a source with memory or in closed-loop DPCM (differential pulse code modulation) (Gersho and Gray, 1992), these two blocks are joined into one, and in other systems, they are jointly designed, such as if half-whitening predictive coding is performed followed by a coder designed for the signal with some correlation left.

The channel encoder block performs operations to protect the bit stream generated by the source coder against transmission errors by increasing the number of bits. The symbol mapping block is used to represent the operation of generating a stream of real or complex-valued symbols from the bit stream output by the channel encoder block. Typical operations of this block are pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM) and phase shift keying (PSK). As indicated by these names, this operation is often interpreted as a part of the modulation process. Here, the modulation will be used to denote only the process of going from the time discrete signal to a time continuous one by performing pulse shaping and modulating the signal onto a carrier. The symbol mapping process is said to create an amplitude continuous signal, although almost all practical schemes will only use a limited number of amplitudes. However, the use of other amplitudes is assumed to be feasible, and due to channel noise, the symbol detection block can observe a continuum of amplitudes.

For system analysis, the modulation and demodulation blocks are often included in the channel model, creating a time discrete channel. The famous result by Nyquist (1928) states that the discrete symbol rate can be up to twice the channel bandwidth, the Nyquist rate; this can also be shown from the sampling theorem. It is also common to include the symbol mapping and detection blocks in the channel model to get a binary channel.

Soft decision decoding is very common, and this is a joint design of the symbol detection and channel decoder blocks. Even on the encoder side, some form of unification of the channel encoder and symbol mapping blocks is common. A reason for this is that a system performing close to the channel capacity of an analog channel, must have a symbol mapping giving a quite large symbol error rate (Blahut, 1987). Then, the channel coder must know the probabilities of different error events to be well designed.

A successful scheme for jointly designing channel coding and symbol mapping is trellis coded modulation (TCM) (Ungerboeck, 1982), which has been followed by a vast amount of work in the same field. TCM-like methods can be used alone or in combination with a traditional channel code.

1.1.1 The Source–Channel Separation Theorem

The coding of a continuous amplitude source can be described by the rate-distortion function, which gives the minimum rate needed to encode the source with a certain distortion, and a channel can be described by its capacity, which is the largest rate at which the channel can transfer information reliably. Both these measures were defined by Shannon (1948). It can be shown that codes exist that perform arbitrarily close to these bounds (Blahut, 1987). A conse-

quence of this is that a source signal can be transmitted on a channel with a certain distortion if the channel capacity is arbitrarily close to the rate-distortion function of the source at the considered distortion. Furthermore, this can be obtained with a source code that encodes the source at a rate sufficiently close to the rate-distortion function followed by a channel code that gives a sufficiently small probability of error at a rate sufficiently close to the channel capacity. The latter result is often called the *source–channel separation theorem*.

Some issues are worth noticing concerning the source–channel separation theorem. There exist channels for which it does not hold because the capacity in some way depends on the coding method. An example of such a channel is a multiple access channel, where the capacity depends on the correlation between the input signals (Cover and Thomas, 1991), another example is the block fading channel discussed in Section 2.7, where the channel is described in terms of the length of the channel codewords. Another problem with the separation theorem is that the codes must have unconstrained lengths in order to come arbitrarily close to the bounds. This means that for a certain coding delay, the separation of source and channel coding might not be optimal. Perhaps the most important problem is that although separate source and channel coding can be optimal for a certain channel, the separation is highly non-robust. If the channel parameters change so that the capacity is lower than expected, the channel code cannot give a low error probability, and the source code then cannot give a low distortion if the codes are designed to be very close to the bounds. Furthermore, the separation means that if the capacity is higher than expected, the distortion cannot go down.

1.1.2 Joint Source–Channel Coding

The problems mentioned above are reasons for looking at joint source–channel coding, despite the separation theorem. The term joint (or combined) source–channel coding will be used about all methods where at least the source coding and channel coding blocks of Figure 1.1 either are unified into a common block or are jointly designed in some way.

In terms of Figure 1.1, a joint source–channel code can include more blocks than source coding and channel coding. If the code is designed for a source with memory, it means that the decomposition and reconstruction blocks are included. If it is designed to utilize the continuous amplitude nature of the source, for instance by using soft information in the decoder, the symbol mapping and detection blocks are part of the joint source–channel encoder and decoder, respectively. This situation is discussed in Chapter 3.

From the mentioned possible shortcomings of coder separation, it can be

deducted that the use of a joint source–channel code can be justified

- if no known separate source and channel codes can give the performance of the joint source–channel code;
- if the joint source–channel code can get the desired performance with a lower complexity than separate codes;
- if under delay constraints, the joint source–channel code performs better than separate codes;
- if the joint source–channel code gives good results for a larger range of channel qualities, thus it is more robust against unknown channel qualities.

The last point is very important. Although modern communication systems may well include advanced systems for estimation of channel parameters and two-way communication for transmitting such information back from the receiver to the transmitter, there are cases when this is impossible. One example is when the transmission delay is too large for a dynamic adjustment of the coding parameters, such as in space communications. Another example is broadcasting, where the same signal is received by many independent users, each of which sees a different channel.

The use of joint source–channel coding has some disadvantages. Modern communication systems are often quite complex, where data can be transmitted through a chain of communication channels having highly diverse properties. The separation of source and channel coding means that the end users only need to focus on the source coding and do not have to know much about the nature of the channel segments, while the nodes of the communication system only need to focus on appropriate channel coding for the channel segments that they serve, not the source that is being transmitted. If joint source–channel coding is to be applied to such a communication system, the source may have to be re-coded at the nodes to fit to different kinds of channels. This means that the nodes must know about source coding for different types of sources. Thus, joint source–channel coding is best suited for situations where the channel can be known by the end users.

1.2 Previous Work

One approach to joint source–channel coding is that of robust quantization. The idea of robust quantization is that the quantizer is designed so that it not only gives a low distortion provided correctly received quantizer index,

but also so that the distortion associated with typical error events is limited. Fine (1964) gave a framework for optimizing a coding system with a discrete amplitude, possibly noisy channel. Kurtenbach and Wintz (1969) optimized a scalar quantizer for a noisy channel. The algorithm was improved by Farvardin and Vaishampayan (1987), who later extended the work to vector quantization (Farvardin, 1990; Farvardin and Vaishampayan, 1991). Hagen and Hedelin (1999) used a linear block code to design the reconstruction codebook.

Massey (1978) considered linear block codes both for channel coding and for source coding, and demonstrated that the source and channel code could be concatenated, thereby simplifying the realization.

Sayood and Borkenhagen (1991) and Phamdo and Farvardin (1994) used the residual redundancy after source coding with DPCM or vector quantization, respectively, to perform error correction decoding. This way, a source coder becomes a joint source–channel coder by a modification of the decoder.

Most joint source–channel coders are based on fixed length codes, since in variable length codes, errors can lead to loss of synchronization, with a catastrophic result. Buttigieg and Farrell (2000) used a Viterbi algorithm to decode variable length codes designed for joint source–channel coding, and Sayood, Otu and Demir (2000) used DPCM and a combination of a convolutional code and a Huffman code at the encoder and Viterbi decoding.

A much used form of joint source–channel coding is that of unequal error protection (UEP). This means that the channel code is designed such that the most important parts of the coded source signal is protected more heavily than the less important parts. In this case, there is separate source and channel coding, but the two codes are connected in an intelligent manner, which means that the two coders must know each other's operation, making this a form of joint source–channel coding. An early discussion of UEP was done by Masnick and Wolf (1967), later work include (Modestino and Daut, 1979; Hagenauer and Stockhammer, 1999), to mention a few. Goldsmith and Effros (1998) used rate-compatible punctured convolutional coding, which is a form of UEP, in combination with a channel-optimized vector quantization, where the codes were jointly designed and the rate optimally distributed between them.

This section contains only a few of the many works on joint source–channel coding. Section 2.2 discusses previous work on multiple description coding, some of which falls under the category of joint source–channel coding. In Section 3.3, results for some codes which utilize the continuous amplitude nature of the channel are given, while Section 4.1 gives some examples of joint source–channel coding for images. Still, the overview of previous work is not meant to be exhaustive.

1.3 Outline of the Dissertation

The dissertation contains five chapters. The main ones, Chapters 2–4, are mostly self-contained.

Chapter 2: A class of joint source–channel codes based on multiple description coding in combination with traditional channel codes is considered and compared to a traditional single description coding based system. This is investigated for three channel models: a packet erasure channel, a binary symmetric channel, and a block fading channel.

Chapter 3: Methods for mapping analog source symbols directly to analog channel symbols are described, and a bandwidth expanding mapping denoted hybrid scalar quantizer–linear coder (HSQLC) is developed. Its performance for different memoryless source distributions and an additive white Gaussian channel is studied.

Chapter 4: The mappings described in Chapter 3 are applied to an existing image coder. The coder operation is described, and several of the coder blocks are modified. The coding results are investigated, and the coder is analyzed to find the significance of some of the operations performed. It is also compared to a system consisting of state-of-the-art source and channel coders.

Chapter 5: Conclusions are drawn from the results of the previous chapters.

In addition there are three appendices:

Appendix A: The existence of maximum distance separable codes over large enough symbol alphabets is proven. This result is applied in Chapter 2.

Appendix B: Formulas are found for the mean square error and the optimal receivers given a transmitter of the HSQLC from Chapter 3.

Appendix C: The original images used for performance evaluation in Chapter 4 are shown.

1.4 Contributions of the Dissertation

Chapter 2:

- Joint source–channel codes based on multiple description coding and error correcting block codes are proposed for three different channel models.

- It is shown that existing high-rate approximations for multiple description quantization are not uniformly convergent, and that this make them unsuitable when the parameter is allowed to be optimized.
- Comparisons are made between single and multiple description coding under similar conditions, to find when multiple description coding is useful.

Chapter 3 and Appendix B:

- The hybrid scalar quantizer–linear coder system is introduced.
- Formulas for an optimal receiver given an HSQLC transmitter are found for different memoryless sources.
- A simplified receiver is proposed.
- Formulas are found for the mean square error of the proposed system with the simplified receiver.
- The coder parameters are optimized.
- The performance of the HSQLC is investigated and compared to other systems.

Chapter 4:

- New mappings are introduced to an existing joint source–channel image coder, including the HSQLC, which introduces bandwidth expansion to the coder.
- The power and bandwidth allocation algorithm of the coder is improved.
- A new coding strategy of the side information is proposed.
- The coder performance is investigated, including the influence of the HSQLC mapping and the power allocation.
- The coder performance is compared against other systems.

Chapter 2

Joint Source–Channel Codes Based on Multiple Descriptions

Multiple description (MD) coding is a multiuser source coding method, where several (usually two) descriptions of the source signal are given, and decoding should be possible from either of them, while decoding of higher quality should be possible from more than one description. It was originally formulated as an abstraction of a packet network communication problem. Its formulation makes it natural to design a joint source–channel code based on multiple description coding, where robustness against lost blocks of data can be obtained. In this chapter, this robustness is used to protect against loss of data occurring when a traditional channel code fails to correct enough information on a noisy communication channel, and the performance of such a coder is compared to a more traditional system.

The chapter is organized as follows. In Section 2.1, the traditional multiple description coding problem is formulated, and Section 2.2 describes previous work on MD coding. In Section 2.3, an overview of the system and channels considered here is given. Section 2.4 describes and discusses the source coding part of the system. In Sections 2.5, 2.6, and 2.7, a description is given of the erasure channel, the binary symmetric channel, and the block fading channel, respectively, with a discussion on the applied channel coder and results for the total system in each case. Finally, the results are discussed in Section 2.8.

This chapter is based on (Coward, Knopp and Servetto, 2001*a*) and (Coward, Knopp and Servetto, 2001*b*).

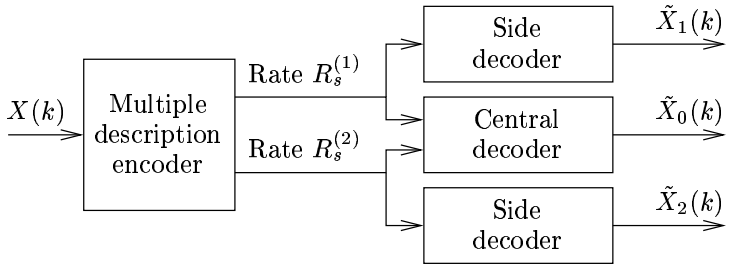


Figure 2.1 Multiple description coder with two descriptions

2.1 Multiple Description Coding Formulation

Multiple description coding is a problem of an encoder producing several outputs, or *descriptions*, and several decoders taking different number of descriptions as input. In general, any number of descriptions can be produced, but in this chapter, as in most of the literature, only the case of two descriptions is considered. This situation is shown in Figure 2.1. The encoder produces two descriptions, and one of the decoders, the *central decoder* takes both descriptions as input, while the other two, the *side decoders*, only take one description each.

The *central distortion* d_0 is defined as the expected distortion between the input $X(k)$ and the output of the central decoder, $\tilde{X}_0(k)$, while the *side distortions* d_1 and d_2, are defined as the expected distortion between the input $X(k)$ and the output of the side decoders, $\tilde{X}_1(k)$ and $\tilde{X}_2(k)$, respectively. For a quadratic distortion measure and real-valued signals, the distortions are given by

$$d_i = E\left[\left(X(k) - \tilde{X}_i(k)\right)^2\right], \; i \in \{0, 1, 2\}. \tag{2.1}$$

The problem is formulated as follows. Given the rates of the two descriptions, $R_s^{(1)}$ and $R_s^{(2)}$, minimize d_0 subject to the constraints

$$d_1 \leq d_{1,\max} \tag{2.2}$$
$$d_2 \leq d_{2,\max}, \tag{2.3}$$

where $d_{1,\max}$ and $d_{2,\max}$ are given constants. If $d_1 = d_2$ and $R_s^{(1)} = R_s^{(2)}$, the descriptions are said to be *balanced*.

2.2 Previous Work

The problem of multiple descriptions was originally formulated at the 1979 IEEE Information Theory Workshop, resulting in works by Witsenhausen (1980), Wolf, Wyner and Ziv (1980), Ozarow (1980), and El Gamal and Cover (1982). All these works seek theoretical bounds on the performance of multiple description coding, not practical coder designs. Later works on performance bounds include (Ahlswede, 1985) and (Zhang and Berger, 1995). Venkataramani, Kramer and Goyal (2001) have found bounds on the achievable performance region for MD coding with more than two descriptions.

One of the first practical coder designs for multiple descriptions was done by Jayant (1981), who used subsampling with a scheme for reducing the perceptual effect of aliasing to perform MD coding of DPCM speech. This approach was later improved by Ingle and Vaishampayan (1995).

Perhaps the most important design principle for MD coders is based on quantization. Already Jayant (1981) mentioned the possibility of using a mid-tread and a mid-rise quantizer to obtain each description, a method that would work even for memoryless sources, but rejected the idea due to the high rate needed on the two descriptions compared to the rate of a single description coder giving the same distortion as the central decoder. However, Vaishampayan (1993) designed multiple description scalar quantizers where the rate of the descriptions can be traded off against the side distortions. This quantizer is obtained by a standard scalar quantizer followed by an index assignment that splits the signal into two descriptions, see Section 2.4.3.1. Vaishampayan and Domaszewicz (1994) considered entropy constrained quantizers, Batllo and Vaishampayan (1997) used transform coding to apply the quantizers to sources with memory, and Vaishampayan and Batllo (1998) found high-rate approximations of the scalar quantizer performance. Jafarkhani and Tarokh (1999) constructed MD trellis coded quantizers, and MD vector quantizers have been studied by Vaishampayan, Sloane and Servetto (2001). Berger-Wolf and Reingold (1999; 2000) found an index assignment and a performance bound for MD scalar quantization for more than two descriptions. Non-balanced MD vector quantization was studied by Fleming and Effros (1999), including more than two descriptions, and by Diggavi, Sloane and Vaishampayan (2000). MD quantizers have been applied by Servetto, Ramchandran, Vaishampayan and Nahrstedt (2000) to design an image coder. Streaming video based on the same principle was developed by Servetto and Nahrstedt (2001).

Another approach to multiple description coding is frequency domain methods, where the descriptions are generated using some sort of frequency analysis, prior to independent quantization of the descriptions. Wang, Orchard,

Vaishampayan and Reibman (2001) used transform coding as the frequency method, and applied the method to image coding. Goyal and Kovačević (2001) also studied transform coding, including the case of more than two descriptions. This was applied to audio coding by Arean, Kovačević and Goyal (2000). Chung and Wang (1999) used lapped orthogonal transforms to design MD coding, with an image coder application. Quantized frame expansions were used by Goyal, Kovačević and Vetterli (1999), again allowing for more than two descriptions. Yang and Ramchandran (2000) designed an MD coder based on filter banks, and Balan, Daubechies and Vaishampayan (2000) used a windowed Fourier method.

A special case of multiple description coding is obtained by relaxing the side distortion constraint on one of the descriptions, thus getting two (or more) descriptions of different importance. The most important description is then needed for good decoding. This problem is known as *successive refinement* or *multiresolution coding*, and has been studied by Equitz and Cover (1991), Rimoldi (1994), and several others.

In this chapter, multiple description coding will be used for joint source–channel coding. Some previous works exist on this problem. Goyal et al. (1999) considered a frequency domain method on an erasure channel. The performance of an MD quantizer on a Rayleigh fading channel was studied by Yang and Vaishampayan (1995), and Alasti, Sayrafian-Pour, Ephremides and Farvardin (2001) investigated the case of an MD quantizer with a congestion network model.

A good historical survey of multiple description coding, giving an alternative presentation from this section, can be found in (Goyal and Kovačević, 2001).

2.3 Problem Considered

The systems considered in this chapter are communication systems for transmitting samples of an analog source over a communication channel. The source is always memoryless, real and Gaussian, and the distortion measure is the mean square error. A system using multiple description coding in combination with channel coding is used as a joint source–channel coder, and the performance is compared against a more traditional system.

It is assumed that there somehow exist two parallel independent channels, and two systems transmitting on these two channels will be considered. One system uses a multiple description (MD) coder which produces two descriptions. These descriptions are channel coded separately and transmitted on the two channels. The other system uses a single description (SD) coder, which

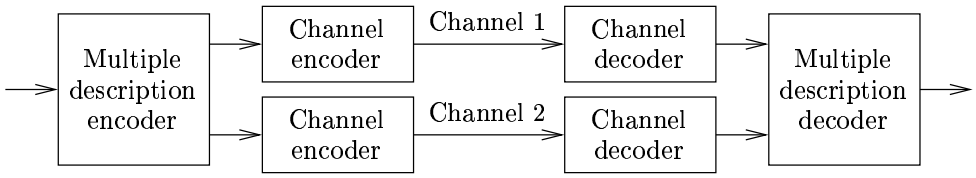


Figure 2.2 Coding system based on multiple description coding

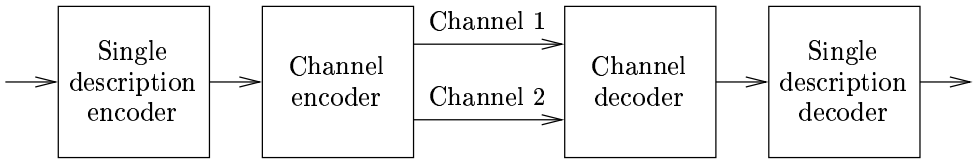


Figure 2.3 Coding system based on single description coding

produces only one output. This is channel encoded in a way suitable for transmission on two channels of the given characteristics and transmitted on the channel pair. The two coding systems are shown in Figures 2.2 and 2.3.

The transmission is block based, and one block from the source coder corresponds to one block from the channel coder. If the number of channel errors exceeds the channel code's correcting ability, the block is lost, and the source coder has to deal with this situation. The MD coder will have an advantage in this case, since decoding of a somewhat reduced quality is possible from the other description if one description is lost. This robustness is deliberately introduced in the source coder so that the channel coder can be made less powerful, and this brings the system into the class of joint source–channel coders.

Three different channel models are considered. One is an erasure channel, which is used as a model of a packet switched network. This channel takes packets as the input, and some of the packets are lost in the transmission. Section 2.5 provides a treatment of this channel.

The second model is a binary symmetric channel, which takes a binary input, and produces bit errors as random events. This channel can be seen as a model of an analog channel with additive white Gaussian noise on which bits are transmitted with a given memoryless modulation. The binary symmetric channel will be discussed in more detail in Section 2.6.

The last channel model is a block fading channel. It has a fading state that can be assumed to be constant for a certain period of time, and two channels with independent fading states. This channel is discussed in more detail in Section 2.7.

The source–channel separation theorem was discussed in Section 1.1.1, and

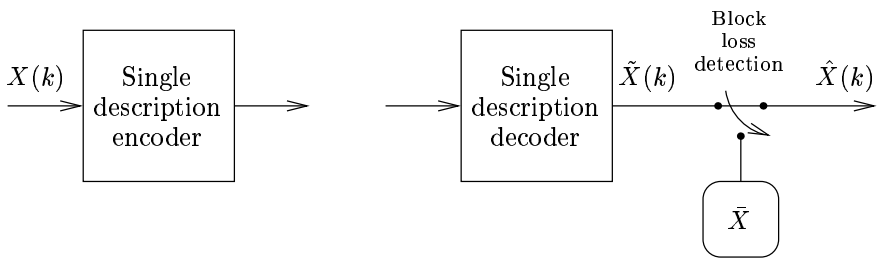


Figure 2.4 Single description source encoder and decoder

it states that for a source with a given rate-distortion function and a channel with a given capacity, an optimal coder for transmitting the source over the channel can be achieved with a source code that reaches the rate-distortion bound followed by a channel code that reaches the channel capacity. The block fading channel is an example of a channel where the theorem does not hold (Ozarow, Shamai and Wyner, 1994). For other channels, in order to get a source code that reaches the rate-distortion bound and a channel code that reaches the capacity, the codes normally have to be infinitely long, thus having infinite delay.

Multiple description coding is a source coding method that is designed to work reasonably well even in the presence of channel errors. This means that it cannot approach the traditional rate-distortion bound unless the constraint on the side distortion is sufficiently relaxed. Thus, in a situation where the source–channel separation theorem holds, multiple description coding will not be useful. Therefore, for the erasure channel and the binary symmetric channel, only cases where the delay is constrained are considered. For the block fading channel, even a case of unconstrained block length is interesting, as the separation theorem does not hold for that channel.

2.4 Source Coding

There are two kinds of source coders, a multiple description coder and a single description coder. Three models for the performance of the coders are considered: The rate-distortion function, which is treated in Section 2.4.1, a high rate quantization model, which is discussed in Section 2.4.2, and a uniform scalar quantizer model, which is described in Section 2.4.3. For all cases, the encoders and decoders can be described as in Figures 2.4 and 2.5. The decoders use information from the channel coder on block losses to decide which output to use. For the single description coder, in case of a block loss, the expected value $\bar{X}$ of the source is output, otherwise the decoded value $\tilde{X}(k)$ is

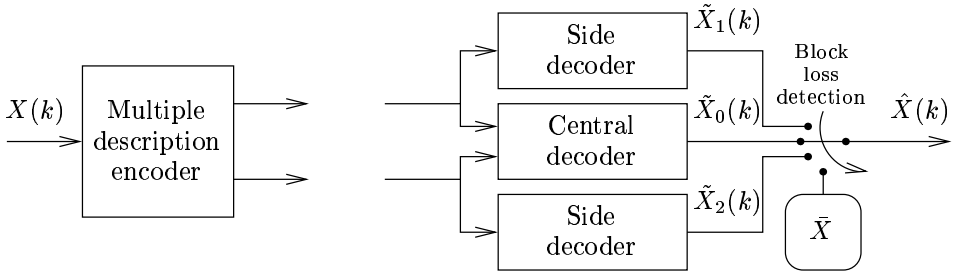


Figure 2.5 Multiple description source encoder and decoder

output. For the multiple description coder, the output of the central decoder, $\tilde{X}_0(k)$, is output if no blocks are lost. If a block in one of the two descriptions is lost, the output of the other side decoder, $\tilde{X}_1(k)$ or $\tilde{X}_2(k)$, is output. If the block in both descriptions is lost, $\bar{X}$ is output.

The source bit rate of the coders is defined as the number of bits produced per input source sample, and is denoted R_s. In the multiple description coder, it is the total number of bits in the two descriptions that is counted.

2.4.1 Rate-Distortion Performance of the Source Coders

The rate-distortion function gives a bound on the distortion of a source coder given a rate, or vice versa. Further, the converse theorem guarantees the existence of codes that perform arbitrarily close to the rate-distortion bound. Thus, the rate-distortion function can be used to find a bound on the performance of the coding systems. In order to come arbitrarily close to the rate-distortion function, no constraint can be put on the delay of the coder, meaning that the block length usually must approach infinity.

For the single description source coder, the rate-distortion function of the source was found by Shannon (1948). The distortion d is defined as the mean square error when no blocks are lost,

$$d = E\left[\left(X(k) - \tilde{X}(k)\right)^2\right]. \tag{2.4}$$

For a memoryless Gaussian unit variance source, the bound is:

$$d \geq 2^{-2R_s} \tag{2.5}$$

For the multiple description coder, the rate-distortion function for a Gaussian input was found by Ozarow (1980). If a quadratic distortion measure

is used, so the distortions are given by (2.1), and the coder is balanced, so $d_1 = d_2$ and the rate of each description is $R_s/2$, the bound is given by:

$$d_1 \geq 2^{-R_s}, \tag{2.6}$$

$$d_0 \geq 2^{-2R_s} \frac{1}{1 - \left(1 - d_1 - \sqrt{d_1^2 - 2^{-2R_s}}\right)^2}. \tag{2.7}$$

2.4.2 High Rate Quantizer Model

The coders are now assumed to be quantizers. A common way of analyzing systems with quantization is to use a high rate model. A high rate model for vector quantization in multiple description coding has been found by Vaishampayan et al. (2001). For a balanced multiple description coder using entropy coding and producing $R_s/2$ bits per sample in each of the two descriptions, the central average distortion $d_0(R_s)$ of an optimal quantizer satisfies

$$\lim_{R_s \to \infty} d_0(R_s) 2^{R_s(1+a)} = \tfrac{1}{4} G(\Lambda) 2^{2h(f_X)}, \tag{2.8}$$

and the average side distortion $d_1(R_s) = d_2(R_s)$ satisfies

$$\lim_{R_s \to \infty} d_1(R_s) 2^{R_s(1-a)} = G(S_L) 2^{2h(f_X)}. \tag{2.9}$$

Here, $G(\Lambda)$ is the normalized second moment of a Voronoi cell of the L-dimensional lattice Λ, L is the vector dimension, $G(S_L)$ is the normalized second moment of a sphere in L dimensions, $h(f_X)$ is the differential entropy of the source, and $a \in (0, 1)$ is a parameter. For the single description coder producing R_s bits per sample, the average distortion $d(R_s)$ of the optimal quantizer satisfies

$$\lim_{R_s \to \infty} d(R_s) 2^{2R_s} = G(\Lambda) 2^{2h(f_X)}. \tag{2.10}$$

The limits above are pointwise convergent for any $a \in (0, 1)$. However, it will be shown that (2.8) is not uniformly convergent. Single description quantization of rate R_s is equivalent to multiple description quantization of a source into two descriptions of rate $R_s/2$ each when the constraints on the side distortions are completely relaxed. This means that the central distortion cannot be smaller than the distortion of a single description coder of the same rate,

$$d(R_s) \leq d_0(R_s). \tag{2.11}$$

Note that Equation (2.10) is independent of a. Thus, if (2.8) is to be uniformly convergent, $d(R_s)2^{2R_s} - d_0(R_s)2^{R_s(1+a)}$ must converge uniformly to $G(\Lambda)2^{2h(f_X)} - \frac{1}{4}G(\Lambda)2^{2h(f_X)} = \frac{3}{4}G(\Lambda)2^{2h(f_X)}$. This means that convergence can be analyzed by studying

$$\delta(a) = \left| d(R_s)2^{2R_s} - d_0(R_s)2^{R_s(1+a)} - \tfrac{3}{4}G(\Lambda)2^{2h(f_X)} \right|. \tag{2.12}$$

If $d(R_s) < d_0(R_s)$, it is possible to find an $a \in (0,1)$ such that $d(R_s)2^{2R_s} - d_0(R_s)2^{R_s(1+a)} < 0$ by choosing an a that satisfies

$$a \in \left(\max\left\{ 1 - \tfrac{1}{R_s}\log_2 \tfrac{d_0(R_s)}{d(R_s)}, 0 \right\}, 1 \right). \tag{2.13}$$

If $d(R_s) = d_0(R_s)$, it is still possible to find an $a \in (0,1)$ such that $d(R_s)2^{2R_s} - d_0(R_s)2^{R_s(1+a)} < \epsilon$ for any $\epsilon > 0$. This is obtained for

$$a \in \begin{cases} \left(\max\left\{ 1 - \tfrac{1}{R_s}\log_2 \tfrac{2^{2R_s}}{2^{2R_s} - \epsilon/d(R_s)}, 0 \right\}, 1 \right) & , \epsilon < 2^{2R_s} d(R_s) \\ (0,1) & , \epsilon \geq 2^{2R_s} d(R_s) \end{cases} \tag{2.14}$$

This means that for any $\epsilon > 0$, there is an $a \in (0,1)$ that satisfies

$$\delta(a) > \tfrac{3}{4}G(\Lambda)2^{2h(f_X)} - \epsilon \tag{2.15}$$

Since $\frac{3}{4}G(\Lambda)2^{2h(f_X)} > 0$ and independent of R_s and a, this means that $\sup_{a\in(0,1)} \delta(a)$ is bounded away from zero, and (2.8) cannot be uniformly convergent.

The fact that the convergence is not uniform means that if the limit performance is used as a approximation for a given finite rate, there will be values of a for which the approximation is poor.

As another way of illustrating this problem, look at the approximation of the distortions obtained by removing the limit operation in Equations (2.8)–(2.10), and denote the approximations $\tilde{d}_0(R_s)$, $\tilde{d}_1(R_s)$, and $\tilde{d}(R_s)$. If the constant a in the multiple description code model is increased, the central distortion becomes smaller and the side distortion larger. If $a \to 1$, the side distortion becomes independent of the rate. This should correspond to the single description case. However,

$$\lim_{a\to 1} \tilde{d}_0(R_s) = \tfrac{1}{4}G(\Lambda)2^{2h(f_X)}2^{-2R_s} = \tfrac{1}{4}\tilde{d}(R_s). \tag{2.16}$$

This means that the estimate of the central distortion of the multiple description coder is one fourth of the estimate of the single description distortion with the same total source rate. This cannot be true for the actual distortions, as

stated in Equation (2.11). When the multiple description coder is to be compared against the single description one, different values of the parameter a must be considered. When the approximation used for the multiple description coder is more optimistic than the one for the single description coder for some values of a, it is difficult to decide which coder is really the best performing one. Thus, the high rate model is unsuitable for the comparisons in this chapter.

2.4.3 Uniform Scalar Quantizer

Since a high rate model is not suitable, a more accurate model must be found. Due to the assumption of a memoryless source, the problem is quite simple, and the performance can be found by some integrals over the probability density function of the source. For a vector quantizer, such integrals will be difficult to calculate due to the complicated integration limits that are obtained. However, with a scalar quantizer, this is not a problem. For a Gaussian source, the integrals can then be expressed by error functions, which can be calculated with any mathematics program. For a scalar quantizer followed by an entropy coder, a uniform quantizer has a performance very close to the optimum (Farvardin and Modestino, 1984), so a uniform scalar quantizer will be assumed.

The single description encoder in Figure 2.4 will now consist of a scalar quantizer followed by an entropy coder, while the single description decoder consists of an entropy decoder followed by a dequantizer. The multiple description coder is given by Figure 2.6, showing that there is an index assignment which distributes the information of the quantizer output between the two descriptions, and separate entropy coding of each index assignment output.

In order to avoid error propagation between blocks and limit the delay introduced by the coder to the number of source samples in one block, the entropy decoding should be carried out separately on each block. It is assumed that the rate out of the entropy coder can be approximated by the actual entropy of the quantized signal, and if this is to be approximately correct, the blocks must contain a relatively large number of source samples. Then, it will also be possible to have a constant number of source samples per block while keeping the blocks equally large. This gives a constant delay.

In the single description coder, if a uniform quantizer of step size δ is used, the distortion of the quantizer is

$$d = d_q = \sum_{i=-\infty}^{\infty} \int_{(i-\frac{1}{2})\delta}^{(i+\frac{1}{2})\delta} (x - \hat{x}_i)^2 f_X(x)\, dx, \tag{2.17}$$

where $f_X(x)$ is the probability density function of the source, and $\hat{x}_i$ is the representation value of interval $i \in \mathbb{Z}$. The optimal representation values are

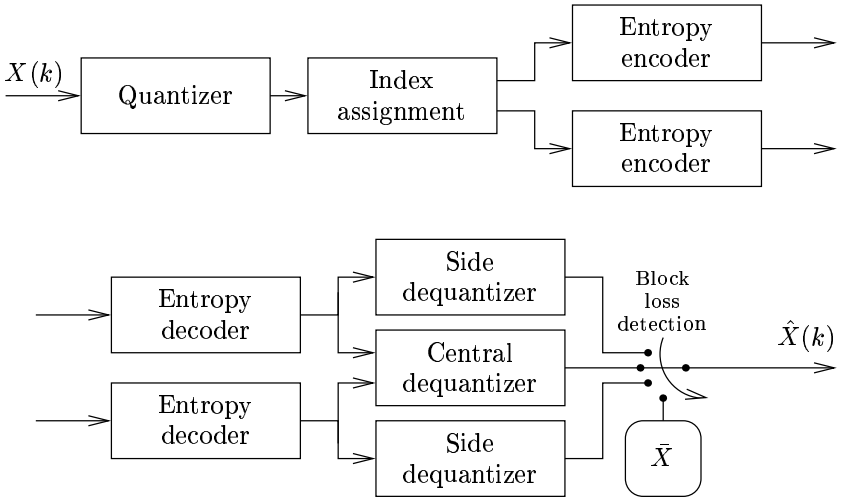


Figure 2.6 Multiple description source coder based on scalar quantization

the centroids of the intervals, given by

$$\hat{x}_i = \frac{\int_{(i-\frac{1}{2})\delta}^{(i+\frac{1}{2})\delta} x f_X(x)\,dx}{\int_{(i-\frac{1}{2})\delta}^{(i+\frac{1}{2})\delta} f_X(x)\,dx}. \tag{2.18}$$

For a Gaussian source X, closed-form expressions for (2.17) and (2.18) suitable for numeric calculation with a mathematics program can be found with the exception of the infinite sum in (2.17)[1]. To be able to calculate the distortion, the sum is taken with i going from $-N_q$ to N_q, where N_q is chosen as

$$N_q = \left\lceil \frac{5\sigma_X}{\delta} \right\rceil, \tag{2.19}$$

where σ_X is the standard deviation of the source and $\lceil t \rceil$ is the smallest integer greater than or equal to t. In addition, the outer integration limits, $(-N_q - 1/2)\delta$ and $(N_q + 1/2)\delta$, are replaced with $-\infty$ and ∞, respectively.

The entropy is

$$R_s = H_q = \sum_{i=-\infty}^{\infty} -\tilde{p}_i \log_2 \tilde{p}_i, \tag{2.20}$$

[1]Such expressions are calculated in Appendix B, Section B.3.1.

where $\tilde{p}_i$ is the probability of interval i,

$$\tilde{p}_i = \int_{(i-\frac{1}{2})\delta}^{(i+\frac{1}{2})\delta} f_X(x)\,dx. \tag{2.21}$$

In order to be able to calculate the entropy, the sum is limited in the same manner as for the distortion.

In the multiple description case, there is an index assignment after the quantizer that splits the quantized symbols into two symbol streams, which are entropy coded separately. The index assignment will be investigated in more detail in Section 2.4.3.1.

The central distortion $d_0 = d_q$, where d_q is as in Equation (2.17). If the m quantization regions of the two descriptions are given by $\mathcal{A}_j^{(l)}$, $j \in \{0, \ldots, m-1\}$, $l \in \{1, 2\}$, then the side distortions are given by

$$d_l = \sum_{j=0}^{m-1} \int_{\mathcal{A}_j^{(l)}} \left(x - \hat{x}_j^{(l)}\right)^2 f_X(x)\,dx, \tag{2.22}$$

where $\hat{x}_j^{(l)}$ is the representation value of each region, given by the centroid of the set $\mathcal{A}_j^{(l)}$,

$$\hat{x}_j^{(l)} = \frac{\int_{\mathcal{A}_j^{(l)}} x f_X(x)\,dx}{\int_{\mathcal{A}_j^{(l)}} f_X(x)\,dx}. \tag{2.23}$$

The entropies of the two signals in the multiple description coder are

$$R_s^{(l)} = H_q^{(l)} = \sum_{j=0}^{m-1} -\tilde{p}_j^{(l)} \log_2 \tilde{p}_j^{(l)}, \tag{2.24}$$

for $l \in \{1, 2\}$, where

$$\tilde{p}_j^{(l)} = \int_{\mathcal{A}_j^{(l)}} f_X(x)\,dx. \tag{2.25}$$

The index assignment is balanced, so $R_s^{(1)} = R_s^{(2)} = R_s/2$ and $d_1 = d_2$.

2.4.3.1 Index Assignment

The index assignment used is proposed by Vaishampayan (1993). This assignment method has two parameters: the side size m of the index assignment

	0	1	2	3	4	5	6	7
0	−10	−9						
1	−8	−7	−5					
2		−6	−4	−3				
3			−2	−1	0			
4					1	3		
5					2	4	5	
6						6	7	9
7							8	10

Figure 2.7 Index assignment matrix, $N_q = 10$, $m = 8$, and $n_{\text{diag}} = 3$.

matrix and the number of diagonals n_{diag}, which must be an odd number or 2. These two parameters are mostly redundant; the parameter that specifies the relationship between the side and the central distortion is the number of diagonals. m should be chosen as small as possible given a number of diagonals.

An example of an index assignment matrix is given in Figure 2.7. A value is first quantized using the uniform quantizer, and the corresponding single description index is found. The placement of the index in the matrix decides the two multiple description indices. In the example, 21 single description indices numbered from -10 to 10 are used. The matrix side size is $m = 8$, giving 8 multiple description indices for each description. The number of diagonals in the matrix is $n_{\text{diag}} = 3$. With three diagonals, there is room for 22 indices, so one cell in the lower one of the three central diagonals is empty.

An example of how the quantization is performed is shown in Figure 2.8. It is assumed that a unit variance Gaussian source is being quantized with a uniform quantizer of step size $\delta = 1/2$. According to Equation (2.19), this gives $N_q = 10$ which corresponds to 21 levels, as in Figure 2.7. Assume that the sample to be quantized lies in the cell with single description index 3. From the matrix, it is seen that the two multiple description indices are 4 and 5. The corresponding side quantization regions can be found by taking the union of the quantization intervals with single description indices given by row 4 and column 5, respectively, which gives

$$\mathcal{A}_4^{(1)} = \left[\tfrac{1}{4}, \tfrac{3}{4}\right) \cup \left[\tfrac{5}{4}, \tfrac{7}{4}\right) \text{ and } \mathcal{A}_5^{(2)} = \left[\tfrac{5}{4}, \tfrac{9}{4}\right) \cup \left[\tfrac{11}{4}, \tfrac{13}{4}\right).$$

The representation value of a side dequantizer is given by Equation (2.23).

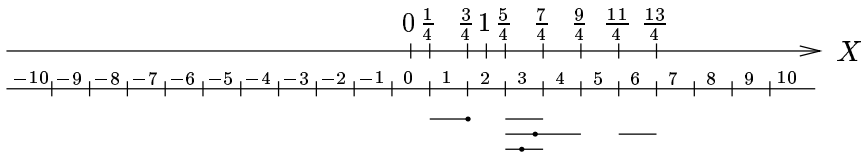


Figure 2.8 An example of the MD quantization of a zero mean Gaussian source of variance $\sigma_X^2 = 1$ using a uniform quantizer of step size $\delta = 1/2$ and the index assignment matrix of Figure 2.7. The upper line shows the variable X that is to be quantized. The next line shows the quantizer cells and the corresponding single description indices. The next three lines show the two side quantizer cells and the central quantizer cell, along with the centroid of the cell, for a value which lies in the cell of single description index 3.

	$\delta = 1/2$				$\delta = 1/10$			
n_{diag}	R_s	d_0	d_1	m	R_s	d_0	d_1	m
2	4.21	0.020	0.077	12	8.74	0.00083	0.0033	52
3	3.37	0.020	0.18	8	7.68	0.00083	0.015	36
5	3.49	0.020	0.65	6	6.36	0.00083	0.11	22
7	3.62	0.020	0.66	6	5.89	0.00083	0.27	18
9	4.10	0.020	0.84	6	5.76	0.00083	0.51	14
11	4.10	0.020	0.84	6	5.87	0.00083	0.62	14

Table 2.1 The source rate (sum of the rates in each description) R_s, the central and side distortions d_0 and d_1, and the matrix side size m for multiple description quantization with a quantizer of step size δ using n_{diag} diagonals

The central quantization region is

$$\mathcal{A}_4^{(1)} \cap \mathcal{A}_5^{(2)} = \left[\tfrac{5}{4}, \tfrac{7}{4}\right).$$

As mentioned in the example, the number of levels used by the multiple description quantizer is given by Equation (2.19), since an infinite number of levels cannot be used with the index assignment algorithm. Thus, the index assignment will depend on how N_q is found. However, it turns out that as long as N_q is chosen large enough, so that the probability of the outer intervals is small, the choice does not influence on the performance of the coder when the number of diagonals is held constant.

The index assignment does not work properly if the number of diagonals is too large. This is demonstrated in Table 2.1. There, the performance of

the multiple description quantizer with a fixed step size is shown for a varying number of diagonals. The central distortion is constant, while the rate and side distortions vary. When the number of diagonals increases, the rate is reduced while the side distortion is increased. However, when the number of diagonals goes from 5 to 7 for $\delta = 1/2$ and from 9 to 11 for $\delta = 1/10$, both the rate and the side distortion increases. By looking at the side size of the index assignment matrix, it is seen that the size is not reduced any more when the number of diagonals reach the saturation levels. The reason is that the index assignment does not allow smaller matrices, although it would be possible in theory; when $\delta = 1/2$, the number of levels is 21, and in 7 diagonals of a matrix of side size $m = 5$, there is room for 23 values. However, it might not be possible to find a balanced index assignment.

This shows that the index assignment algorithm is not perfect. It should also be possible to fill the matrix completely, and thus get a source rate equal to the single description case. This problem is treated by Berger-Wolf and Reingold (2000).

The performance of the MD quantizer is shown in Figure 2.9, where the signal-to-noise ratio (SNR) of the central and side decoder is plotted against the source rate for different number of diagonals. Results where the number of diagonals is too large, as demonstrated in Table 2.1, are omitted. For high rates, the distance between the central and side SNR is constant, and the lowest possible difference is 6 dB for high rates, obtained with two diagonals. Note that for low rates, the central SNR is only marginally higher than with a lower number of diagonals, while the side SNR is significantly lower compared to the case of fewer diagonals. As the rate is somewhat increased, the high-rate performance is obtained, where the increase in central SNR and the decrease in side SNR compared to fewer diagonals is about the same.

2.4.4 Total Distortion

The total average distortion of the system is given by $E[(\hat{X}(k) - X(k))^2]$, see Figures 2.4 and 2.5. In order to calculate this, the probability of block errors p_e must be known. This will be calculated in Sections 2.5, 2.6, and 2.7. Since the two parallel channels are assumed to be independent, block errors in the two channels of the multiple description coding system are independent events. The decoder is assumed always to know when a block is lost, so no corrupted blocks are output from the channel decoder.

For the single description coder, the average distortion is then given by

$$D_{\mathrm{SD}} = (1 - p_{e,\mathrm{SD}})d + p_{e,\mathrm{SD}}\sigma_X^2, \tag{2.26}$$

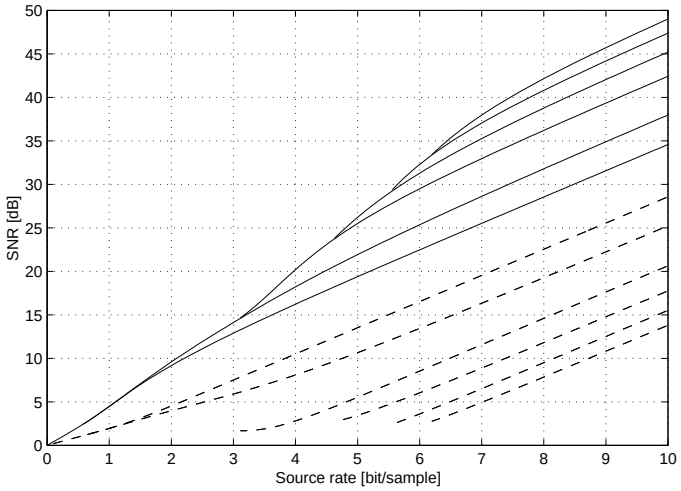


Figure 2.9 The central (solid) and side (dashed) SNR as a function of the source rate (sum of the rates in each description) for the MD quantizer on a Gaussian memoryless source. The number of diagonals is (from below for central SNR, from above for side SNR) 2, 3, 5, 7, 9, and 11.

where σ_X^2 is the signal variance and $p_{e,\mathrm{SD}}$ is the block error probability in this case, while in the multiple description case, the average distortion is

$$D_{\mathrm{MD}} = (1-p_{e,\mathrm{MD}})^2 d_0 + p_{e,\mathrm{MD}}(1-p_{e,\mathrm{MD}})(d_1+d_2) + p_{e,\mathrm{MD}}^2 \sigma_X^2, \qquad (2.27)$$

where $p_{e,\mathrm{MD}}$ is the block error probability in the MD system. The formulas are valid for d, d_0, d_1, and d_2 found by any of the models described in this section.

2.5 Erasure Channels with Large Input Alphabets

When transmitting in packet switched networks, packets might get lost in the transmission. The packet erasure channel can be used as a model for this kind of channel. The channel is assumed to transmit symbols which are erased with a probability p, and erasures are independent events. If a symbol is erased, the receiver knows this, and if it is not erased, it is received correctly with probability one. Since the erasures are independent, the two independent channels depicted in Figures 2.2 and 2.3 are equivalent to one channel with twice the capacity in this case.

For error protection, linear block codes will be considered. In packet switched networks, the symbols will be packets, typically consisting of at least

hundreds of bits. Thus, it is assumed that one symbol (packet) is enough to transmit an isolated block of source data with rate and distortion performance as described in Section 2.4, so the block size of a channel code is allowed to be as low as one. The large symbols also mean that a delay constraint will force the blocks to be quite short.

In the multiple description coder, channel coding is performed separately on each description, while there is only one channel coding operation in the single description coder. In order to provide the same delay, this means that the block length of the channel code can be twice as large for the SD coder compared to the MD coder. If L_{SD} is the number of source symbols per block with single description coding, L_{MD} is the number of source symbols per block with multiple description coding, R_s is the source bit rate, R_c is the channel bit rate (the number of channel bits per data bit, or equivalently, the ratio of data symbols in a block), ν is the number of channel bits per channel symbol, and n_{SD} and n_{MD} are the number of channel symbols per block with SD and MD coding, respectively, then the total rate R measured in channel bits per source symbol is

$$R = \frac{\nu n_{\mathrm{SD}}}{L_{\mathrm{SD}}} = \frac{R_s}{R_c} \tag{2.28}$$

for the single description coder and

$$R = \frac{2\nu n_{\mathrm{MD}}}{L_{\mathrm{MD}}} = \frac{R_s}{R_c} \tag{2.29}$$

for the multiple description coder. This means that for the same delay, $L_{\mathrm{SD}} = L_{\mathrm{MD}}$, $n_{\mathrm{SD}} = 2n_{\mathrm{MD}}$. Note that R_c and R_s need not be the same for the two coding systems, but for a fair comparison, R must be the same.

2.5.1 Channel Coding

A linear block code is used to protect the source bits against channel errors. The block code is described by the block length n, the number of data symbols per block k and the minimum distance $d_{\min}$, and is then called an $(n, k, d_{\min})$ code. The minimum distance decides the number of erasures that the code can correct; it can correct at least $d_{\min} - 1$ erasures in a block.

Maximum distance separable (MDS) codes are codes that have the largest minimum distance possible for a given block length n and number of data symbols k, and their minimum distance is given by (Blahut, 1983)

$$d_{\min} = n - k + 1 \tag{2.30}$$

In Appendix A, it is shown that for any n, MDS codes exist provided that the alphabet size q is large enough.

For a channel of erasure probability p, an $(n, k, n-k+1)$ MDS code gives a channel bit rate $R_c = k/n$ and a correcting ability of $n-k$ erasures in a block, so the block error probability is

$$p_e = 1 - \sum_{i=0}^{n-k} \binom{n}{i} p^i (1-p)^{n-i}. \tag{2.31}$$

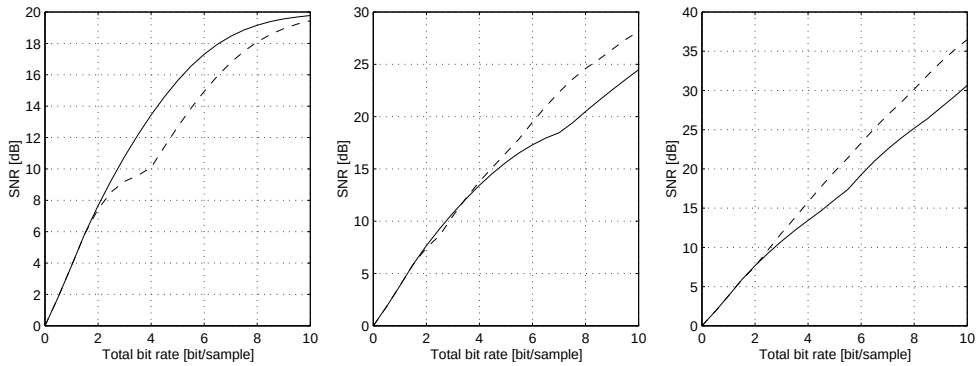
2.5.2 Optimizations

For the erasure channel, both the uniform scalar quantizer and the rate-distortion limit will be considered as source coding models. Since no specific assumption is made for the alphabet size, the source block length L can approach infinity even for a finite block length n, which defends the use of the rate-distortion bound. A maximum block length $n_{\max}$ will be provided, as will an erasure probability p. The block lengths $n_{\mathrm{SD}} \leq 2n_{\max}$ and $n_{\mathrm{MD}} \leq n_{\max}$ are assumed.

The parameters that can be varied are then

- The block length n_{MD} or n_{SD}
- The number of data symbols k
- The quantizer step size δ with the uniform quantizer, or the source bit rate R_s with the rate-distortion bound
- For the multiple description coder, the number of diagonals in the index assignment matrix with the uniform quantizer, or the side distortion d_1 with the rate-distortion bound

Given the parameters above, it is possible to calculate the total distortion and the total rate R. With the uniform scalar quantizer, all the parameters above except δ are discrete. This means that for a desired total rate R, the best scheme can be found using discrete optimization (trying all possibilities) of the discrete parameters, numerically solving for δ in each case. With the rate-distortion bound, both R_s and d_1 are discrete, and in this case, R_s is found from the other parameters, d_1 is optimized numerically using Matlab's Optimization Toolbox (Coleman, Branch and Grace, 1999) for each combination of the discrete parameters, which can subsequently be optimized discretely.



(a) $n_{\max} = 1$ (b) $n_{\max} = 5$ (c) $n_{\max} = 15$

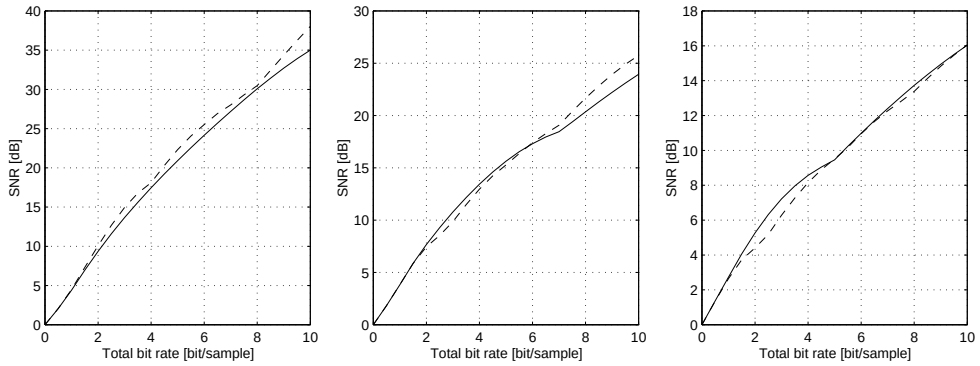
Figure 2.10 Performance of multiple (solid) and single (dashed) description coding on an erasure channel with erasure probability $p = 0.1$ assuming source coding by uniform scalar quantization. The maximum block size is $n_{\max}$.

2.5.3 Comparisons

Figure 2.10 shows the signal-to-noise ratio (SNR) as a function of rate for the coding systems with varying maximum block length $n_{\max}$ on the erasure channel. The uniform scalar quantizer with ideal entropy coding is used as the source coding model. With very short blocks, MD coding will be as least as good as SD coding, but longer blocks give SD coding an advantage. For $n_{\max} = 1$, MD coding is always as least as good as SD coding for the displayed channel quality range. In this case, no channel coding (i.e., a $(1, 1, 1)$ identity code) can be performed for MD coding. This is the situation that most closely resembles the problem formulation of multiple description coding, since that does not assume the existence of any channel code. For SD coding, either no code or a $(2, 1, 2)$ repetition code can be used. This explains the elbow in the SD performance plot for $n_{\max} = 1$ around 4 bit/sample, this point is where the performance curves of no coding and a repetition code cross. The repetition code produces two packets per input source block, and corresponds to MD coding where the index assignment has only one diagonal (although this scheme is not considered in the MD coder), cf. Section 2.4.3.1.

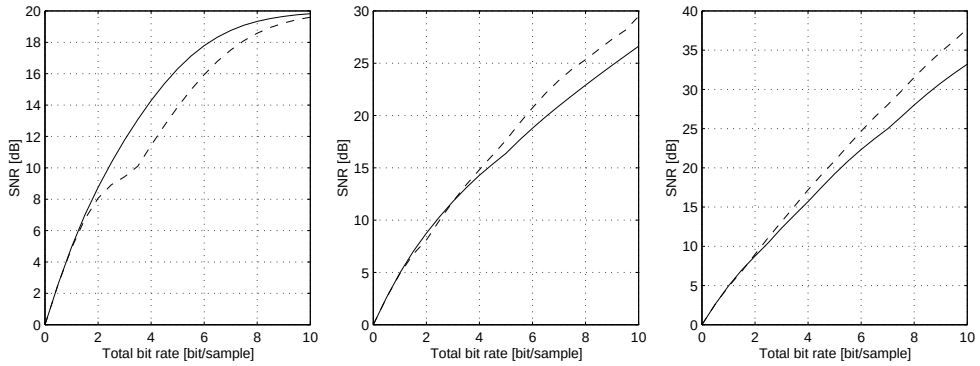
Figure 2.11 shows how the performance varies if the erasure probability is changed. As one would expect, multiple description coding performs relatively better for poor channels, since the shortcomings of a short channel code are more evident there.

In Figure 2.12, the experiments from Figure 2.10 are repeated under the as-



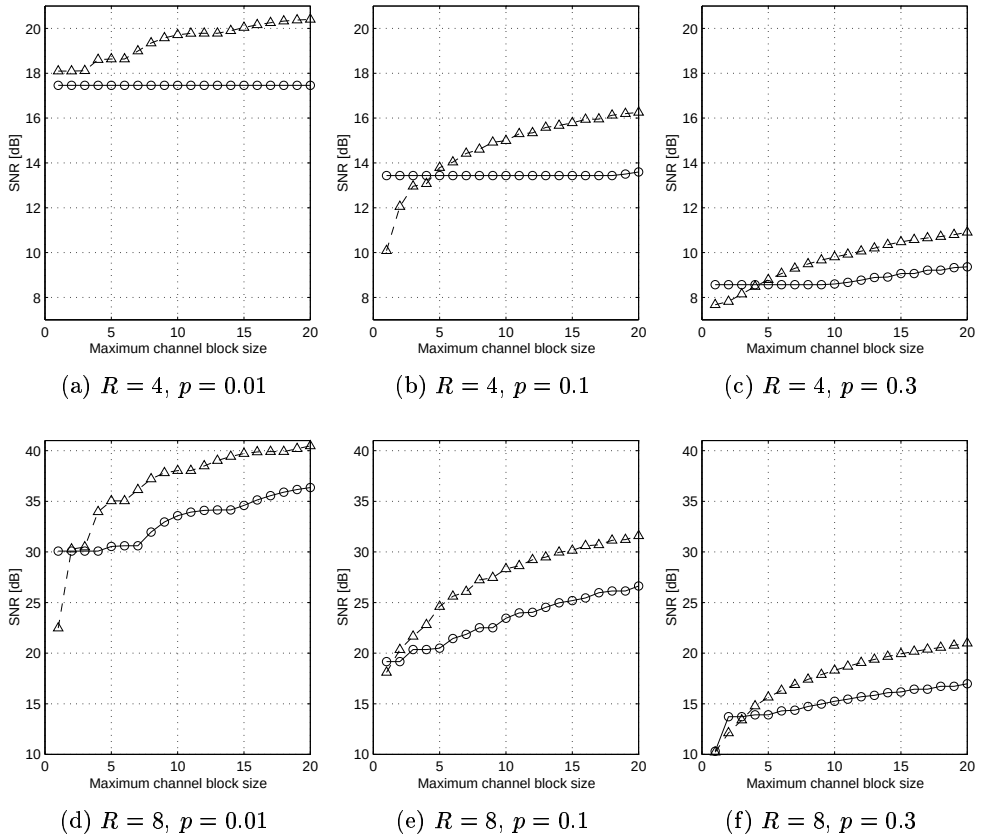
(a) $p = 0.01$ (b) $p = 0.1$ (c) $p = 0.3$

Figure 2.11 Performance of multiple (solid) and single (dashed) description coding on an erasure channel assuming source coding by uniform scalar quantization. The maximum block size is $n_{\max} = 3$, and the erasure probability is p.



(a) $n_{\max} = 1$ (b) $n_{\max} = 5$ (c) $n_{\max} = 15$

Figure 2.12 Performance of multiple (solid) and single (dashed) description coding on an erasure channel with erasure probability $p = 0.1$ assuming source coding performance as the rate-distortion bound. The maximum block size is $n_{\max} = 1$.



(a) $R = 4$, $p = 0.01$ (b) $R = 4$, $p = 0.1$ (c) $R = 4$, $p = 0.3$

(d) $R = 8$, $p = 0.01$ (e) $R = 8$, $p = 0.1$ (f) $R = 8$, $p = 0.3$

Figure 2.13 Performance of multiple (solid with circles) and single (dashed with triangles) description coding on an erasure channel as a function of maximum block size $n_{\max}$ assuming source coding by uniform scalar quantization, for different total rates R and erasure probabilities p.

sumption that the source coder now performs according to the rate-distortion bound. The differences from Figure 2.10 are moderate, but the distance has been reduced when SD coding is better than MD. This is due to the imperfections of the index assignment algorithm described in Section 2.4.3.1, which do not apply to the rate-distortion performance.

The results so far have shown, not unexpectedly, that multiple description coding performs relatively better when the block lengths are constrained to be short. In Figure 2.13, this is investigated by considering the performance as a function of the block size constraint. Note that the MD code often has

a constant performance up to quite high values of $n_{\max}$. This is because no channel code is the preferred scheme in this case even as $n_{\max}$ increases. The figure reveals that SD coding outperforms MD coding already at quite low $n_{\max}$ values, and for the best channel, SD coding is always better at the rate $R = 4$. Still, if the delay constraints prevent the use of any channel code at all, multiple description coding clearly gives an improvement in most cases.

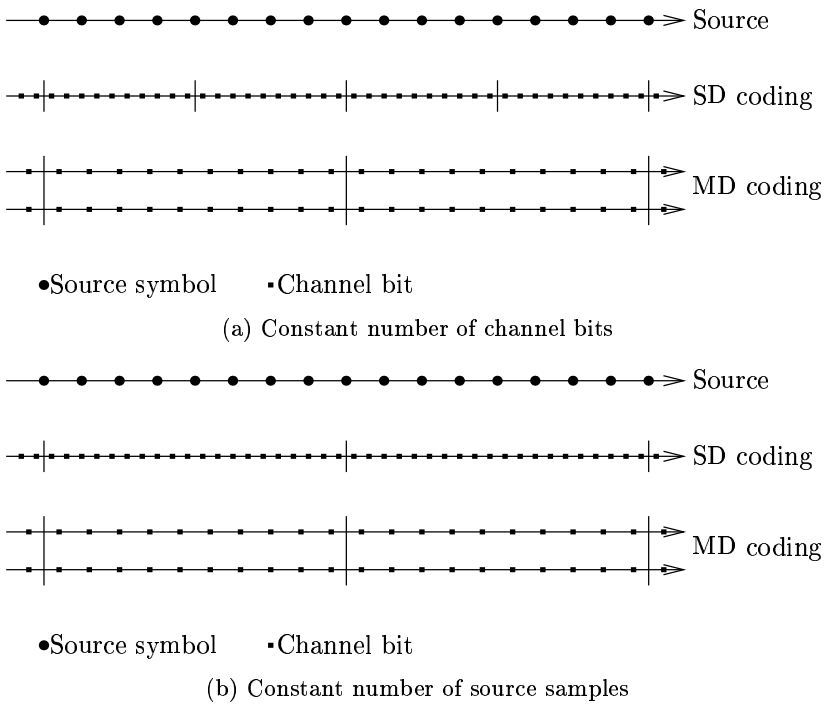
The results reveal that if a certain delay can be accepted, MDS block codes over the alphabet defined by packets can give better protection with single description than with multiple description coding. However, with very tight delays, MD coding does better than SD coding in many cases. The erasure channel model used in this section was introduced as a model of a packet network, but it can also be used as a model of a communication system where block codes are used to protect against bit errors in the transmission. Then, an erasure corresponds to the situation when there are more errors in a block than the code can correct. If the channel coding scheme is fixed and delay constraints prevent the use of a block code on the packet level, MD coding is often better than SD coding. However, if the error protection within the blocks can be varied, it can be jointly optimized with the source coder, changing the results. This situation is investigated in the next two sections for two different channel models.

2.6 Binary Symmetric Channel

The binary symmetric channel (BSC) takes a binary input, and bit errors are independent events and occur at a probability p_b. The independence of the errors means that the two parallel channels in Figures 2.2 and 2.3 are equivalent to one channel with twice the capacity, just as for the erasure channel.

As mentioned, the transmission is block based. Two different cases are studied, illustrated in Figure 2.14. The first case has the following properties:

- The number of channel bits per block is the same in the multiple and the single description coder.
- The delay introduced by the coder in terms of source samples is twice as large in the multiple description coder as compared to the single description coder, assuming that the source decoder operates on blocks no larger than the channel blocks.
- The model is good if a limiting factor is the complexity of the channel coder, or if the family of available channel codes limits the number of available block sizes.



(a) Constant number of channel bits

(b) Constant number of source samples

Figure 2.14 Block sizes of the single and multiple description coder

The properties of the second case are:

- The number of channel bits per block is twice as large for the single description coder as for the multiple description coder.
- The delay introduced by the coder in terms of source samples is the same in both systems.
- The model is good if a limiting factor is the admissible source delay.

The number of source samples per block is L_{SD} and L_{MD} with the single and multiple description coders, respectively, and from this, the source coder produces $M_{\mathrm{SD}} = R_s L_{\mathrm{SD}}$ or $M_{\mathrm{MD}} = R_s L_{\mathrm{MD}}$ bits. In the single description coder, these M_{SD} bits are processed by a channel coder which produces a block of $n_{\mathrm{SD}} = M_{\mathrm{SD}}/R_c$ channel bits. Here, R_c is the channel bit rate measured in data bits per channel bit. The total rate of the system is defined as the number of channel bits per source symbol,

$$R = \frac{n_{\mathrm{SD}}}{L_{\mathrm{SD}}} = \frac{R_s}{R_c} \tag{2.32}$$

In the multiple description coder, there are $M_{\mathrm{MD}}/2$ bits for each description. The channel coder produces a block of $n_{\mathrm{MD}} = \frac{M_{\mathrm{MD}}/2}{R_c}$ channel bits from these $M_{\mathrm{MD}}/2$ source bits. The total rate is now the total number of channel bits produced per source symbol:

$$R = \frac{2n_{\mathrm{MD}}}{L_{\mathrm{MD}}} = \frac{R_s}{R_c}. \tag{2.33}$$

As with the erasure channel, R_c and R_s may be different for MD and SD coding, but R must be the same in a fair comparison.

The binary symmetric channel satisfies the source–channel separation theorem if the block length is infinite. The only way of achieving a source coder whose performance reaches the rate-distortion functions of Section 2.4.1 is to have infinitely long blocks. Then the separation theorem is satisfied, and an optimal coding system can be obtained with a source coder that does not accept errors on the input of the decoder. Thus, the single description coder will be optimal in that case. Because of that, only finite block lengths are considered. The channel coder that is used is described below.

2.6.1 Channel Coding

The error protection is assumed to be performed by an (n, k, d_{min}) linear block code, just as for the erasure channel, but this time the symbol alphabet is binary, since errors occur independently on bits. For the single description coder, $n = n_{\mathrm{SD}}$ and $k = M_{\mathrm{SD}}$, and for the multiple description coder, $n = n_{\mathrm{MD}}$ and $k = M_{\mathrm{MD}}/2$.

The minimum distance decides the number of bit errors that the code can correct; a code can correct at least t bit errors in a block if its minimum distance satisfies

$$d_{\mathrm{min}} \geq 2t + 1. \tag{2.34}$$

Binary MDS codes are not attainable (Blahut, 1983), so the MDS performance assumed for the erasure channel cannot be used here. No specific code is assumed, instead only codes that satisfies the Gilbert bound (Blahut, 1997) are considered. The Gilbert bound for binary linear block codes states that for any integers n and $\hat{d}_{\mathrm{min}}$, $2 \leq \hat{d}_{\mathrm{min}} \leq n$, there exists a binary (n, k, d_{min}) linear code with $d_{\mathrm{min}} \geq \hat{d}_{\mathrm{min}}$ whose dimension k satisfies

$$\sum_{i=0}^{\hat{d}_{\mathrm{min}}-1} \binom{n}{i} \geq 2^{n-k}. \tag{2.35}$$

This means that one can assume the existence of a code of block length n and minimum distance $d_{\min}$ where k is given by

$$k = \left\lceil n - \log_2 \sum_{i=0}^{d_{\min}-1} \binom{n}{i} \right\rceil . \tag{2.36}$$

If such a code does not exist, the Gilbert bound guarantees the existence of a better code of the same block length.

The assumption of codes found from the Gilbert bound needs some justification, since it can be argued that it gives unrealistic performance estimates. For short codes, the Gilbert bound gives a conservative approximation of the attainable code performance, so better codes than predicted are attainable. For long codes, the Gilbert bound will have a different problem: it is not known how to construct codes of the predicted quality. For short codes, an example with practical codes will be given in comparison to Gilbert bound assumption. For long blocks, the Gilbert bound will be used exclusively. The results may be over-optimistic, but it is likely that the demonstrated performance can be found for somewhat longer blocks, and that the Gilbert bound provides insight in how variations of the channel code rate influence the overall system performance.

The channel bit rate is given by $R_c = k/n$. The block error rate p_e can be described in terms of the block length n, the number of bit errors t that can be corrected, and the bit error probability p_b:

$$p_e = 1 - \sum_{i=0}^{t} \binom{n}{i} p_b^i (1-p_b)^{n-i}. \tag{2.37}$$

2.6.2 Optimizations

The performance of the coder will be calculated assuming the uniform scalar quantizer. The performance of the coders will be optimized, assuming that the source block length L or the channel block length n is given, along with the bit error rate p_b. The parameters that can be varied are the following:

- The quantizer step size δ.
- The minimum distance $d_{\min}$ of the channel code.
- For the multiple description coder, the number of diagonals in the index assignment matrix.

This situation is very similar to the erasure channel, and for a desired total rate R, the best scheme can be found by a discrete optimization of the discrete parameters, numerically solving for δ in each case.

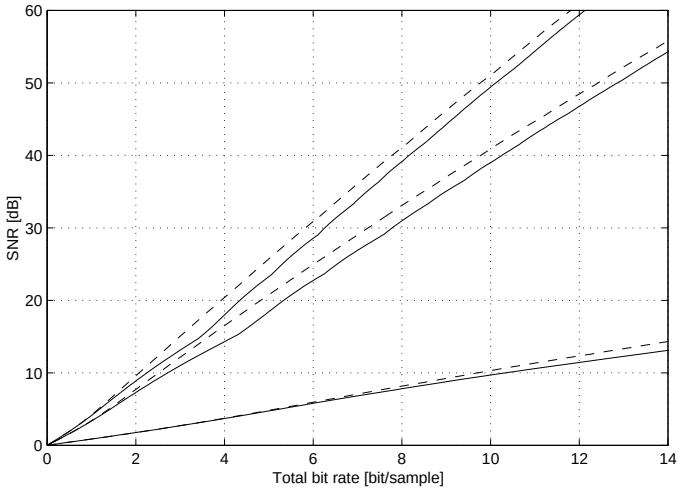


Figure 2.15 Rate-distortion performance for multiple (solid) and single (dashed) description coding with a block size of $n = 1000$ channel bits and bit error probabilities of (from below) 10^{-1}, 10^{-2} and 10^{-3}.

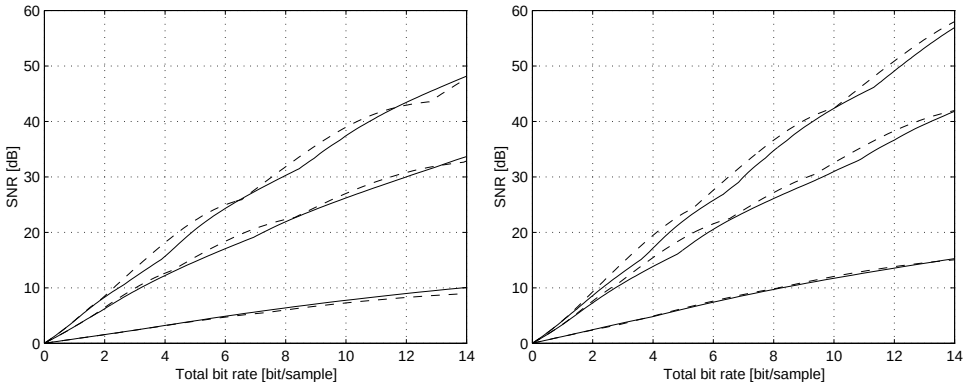
2.6.3 Comparisons

The performance of the multiple and single description coder are to be compared, with the distortion given by Equations (2.26) and (2.27), at specified total rates R. In Section 2.6.3.1, the case of a fixed channel block length is considered, while in Section 2.6.3.3, the source block length is kept constant.

2.6.3.1 Fixed Channel Block Length

In Figure 2.15, the rate-distortion performances of the coders are shown for fixed bit error rates. The graphs show the performance using a channel block code of length $n = 1000$ for both cases, and the performance has been optimized as described in Section 2.6.2. For the multiple description case, also the number of diagonals has been chosen as the best performing one. The rate shown is the total rate R on the channel.

It is seen that for low rates, the performance of the single and multiple description system is about the same, while for high rates, the single description system outperforms the multiple description one. For low bit rates, few bits can be spent on error correction, and this means that the probability of block errors is greater than for high bit rates. Then, the superior performance of multiple description coding in the case of a block error can be exploited. Still, this is



(a) Performance predicted by the Gilbert bound

(b) Performance of known codes

Figure 2.16 Rate-distortion performance for multiple (solid) and single (dashed) description coding with a block size of $n = 63$ channel bits and bit error probabilities of (from below) 10^{-1}, 10^{-2} and 10^{-3}

not enough to perform better than single description coding. This explanation is supported by the fact that the difference between the two systems increases more slowly when the bit error probability is higher, as can be seen from the figure, especially when comparing the case of $p_b = 10^{-1}$ to the other two cases.

In Figure 2.16(a), the results of Figure 2.15 are repeated for shorter blocks of $n = 63$ channel bits. When using the Gilbert bound in this case, multiple description coding is preferred for a poor channel, $p_b = 10^{-1}$. However, the performance obtained with either system in this case is hardly ever acceptable. Note also that the curves are non-smooth. This is because the number of different channel codes available is smaller when the codes are short. This also means that there are areas for $p_b = 10^{-3}$ and $p_b = 10^{-2}$ where the multiple description coder performs better, as there is no available channel code that is well enough suited for the single description coder in that case. With these short blocks, note that the assumptions that the rate can be approximated by the entropy and that the blocks have a constant number of both source and channel symbols, are not well justified anymore.

In Figure 2.16(b), the codes predicted by the Gilbert bound have been replaced with actual codes of length $n = 63$ found in (Peterson and Weldon, 1972, Appendix D). Only the codes with the largest k for each minimum distance are considered. The graphs demonstrate the fact that the Gilbert bound is loose for short blocks, so better codes than predicted can be obtained.

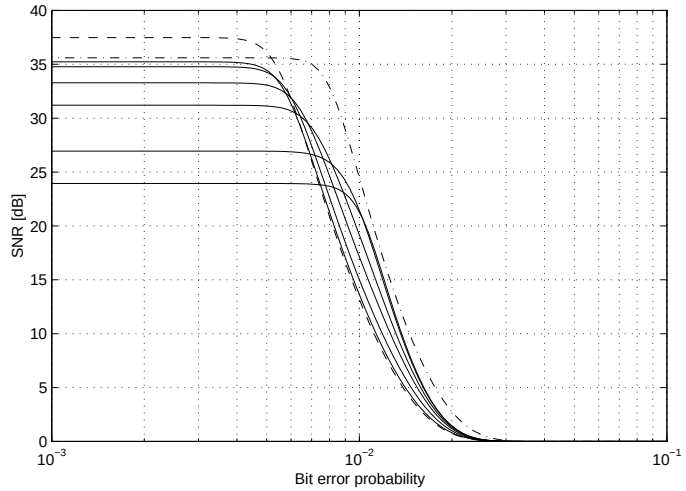


Figure 2.17 SNR as a function of bit error probabilities for a total rate of $R = 8$ bit/sample for the MD and SD case. Solid lines: MD coding with a $(1000, 810, 31)$ channel code and (from below at $p_b = 10^{-3}$) 2, 3, 5, 7, 9, and 11 diagonals. Dashed line: SD coding with a $(1000, 810, 31)$ channel code. Dash-dotted line: SD coding with a $(1000, 771, 39)$ channel code.

In Figure 2.17, the SNR is shown as a function of the bit error probability for multiple and single description coding. This shows the robustness of the coders against variations in the bit error probabilities. For the multiple description case, the number of diagonals changes, while the error correcting code is kept constant. It is seen that as the number of diagonals increases, the performance of the coder improves for good channels, at the cost of poorer performance in the transition region around $p_b = 10^{-2}$. The single description coder with the same error correcting code performs better than all the multiple description coders for good channels, and worse in the transition area, as expected. However, if the error correcting code for the single description case protects some more bits, the performance becomes better than all the multiple description coders for all bit error rates. This means that the increase in robustness that can be achieved with multiple description coding is less than what can be achieved with extra protection in the single description case.

Similar results as in Figure 2.17 can be obtained for other bit rates, packet sizes, and channel code parameters. If the bit rate is lower, the results are less evident, as the difference between different schemes is smaller. If the packets are shorter, it might become impossible for single description coding to

outperform multiple description coding for every bit error probability, because the number of possible channel code parameters is smaller.

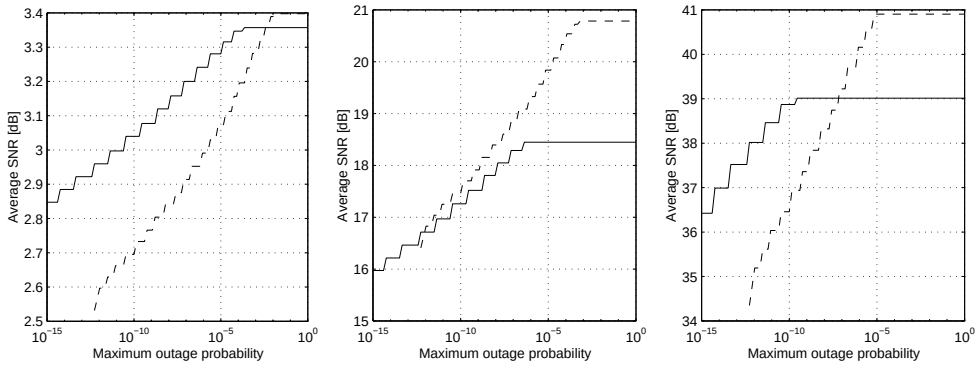
2.6.3.2 Constraints on the Outage Probability

The formulas for the distortion in Equations (2.26) and (2.27) are average distortions, taken over infinite time. In a short period of time, the distortion can be very different from this. For the coders considered in this chapter, there are only two possibilities for the single description case, correctly received blocks with a distortion given by (2.17) and incorrectly received blocks with a distortion equal to the signal variance σ_X^2, that is, a total loss. For the multiple description case, there are three possibilities, two correctly received blocks with a distortion given by (2.17), one correctly received block giving a distortion as in (2.22), and no correctly received blocks, giving a distortion of σ_X^2. The cases of distortion σ_X^2 will be referred to as an *outage*. This case is particularly bad, as the receiver will have no information for the duration of the outage. In many applications, the subjective distortion that this causes will be far greater than the outage's contribution to the total mean square error.

In order to deal with the problem of outages, a constraint will be imposed on the probability of an outage. This means that for each total bit rate, only the channel codes that makes the outage probabilities smaller than a certain level are considered. For the single description coder, the outage probability is equal to the block error probability p_e, while for the multiple description coder, the outage probability is p_e^2.

In the multiple description case, the loss of one description is not considered an outage. However, in some cases, the side distortion may be very large, which means that the loss of one description in practice will be almost as bad as an outage. In order to avoid this situation, it is possible to put a constraint on the side distortions as well, meaning that only numbers of diagonals that give low side distortions are considered.

Figure 2.18 shows the performance of the coders as a function of the maximum outage probability for different rates. Very low outage probabilities cannot be obtained for the single description case, because the channel code with length $n = 1000$ is unable to produce such a low error probability for these channels. As the maximum outage probability decreases, the performance is constant before it starts dropping. The reason for the constant value is that the overall best channel code of course has an outage probability less than one, and the performance will start dropping when the maximum outage probability goes below that level. The constant value is held much longer for the multiple description coder, and this means that it can outperform the single description coder when the maximum outage probability is small enough. This



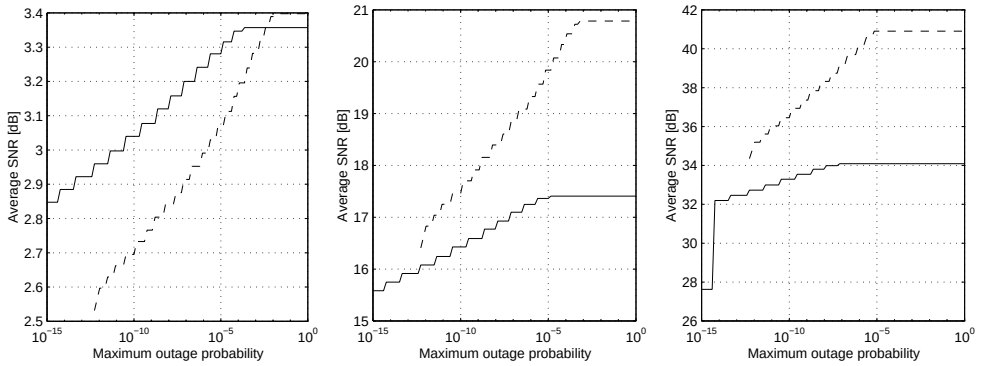
(a) $R = 1$ (b) $R = 5$ (c) $R = 10$

Figure 2.18 Performance of the multiple (solid) and single (dashed) description coders when the maximum outage probability is constrained. The bit error probability $p_b = 10^{-2}$, the channel block length $n = 1000$, and the total rate is R.

is expected, since the outage probability is smaller with the multiple description coder. The reason for the staircase shape of the curves is that the number of possible parameter combinations is finite, and thus, the same system must be used for slightly different maximum outage probabilities, meaning that the obtained maximum outage probability is sometimes smaller than the specified one.

When the total rate increases from 1 bit per sample to 5 bits per sample, the performance of the single description coder increases more than of the multiple description coder, and thus, the range of maximum outage probabilities for which the multiple description coder is the best, is reduced. This can easily be explained, as more error protection can be introduced with a higher total bit rate, and thus, the advantage of multiple description coding becomes smaller. More surprisingly, as the rate increases from 5 to 10 bits per sample, the multiple description coder's performance increases more than the performance of the single description coder. An explanation for this can be that as the source bit rate increases, the possible number of diagonals in the multiple description coding goes up, cf. Section 2.4.3.1. Specifically, it becomes possible to design a multiple description coder that performs close to a single description coder in the central dequantizer, with the cost that the side distortions increase.

This explanation is supported by Figure 2.19. Here, the side SNR measured in dB is constrained to be at least one third of the central SNR. With this constraint, the performance increases more for the single description coder



(a) $R = 1$ (b) $R = 5$ (c) $R = 10$

Figure 2.19 Performance of the multiple (solid) and single (dashed) description coders when the maximum outage probability and the side distortion are constrained. The bit error probability $p_b = 10^{-2}$, the channel block length $n = 1000$, and the total rate is R.

than for the multiple description coder as the rate increases. Note that for 1 bit per sample, there is no difference when imposing the extra constraint.

The examples shown here were all for a bit error probability of 10^{-2}. For higher bit error probabilities, the block error probability will also have to be higher, and thus, the multiple description coder will become more favorable, just as when the total rate is decreased.

2.6.3.3 Fixed Source Block Length

If the number of source samples per block is to be constant, the number of channel bits per block is twice as high for the single description coder as compared to the multiple description coder. Longer block codes will normally give better performance, so this case should be more favorable to the single description coder.

In Figure 2.20, the rate-distortion performances of the coders are shown for fixed bit error rates when the number of source samples per block is constant. The graphs show the performance using blocks of length $L = 500$ source samples for both cases. As expected, this makes the difference between the single and multiple description coder performance larger that when the channel block length is constant, as in Figure 2.15.

For a shorter source block length of $L = 50$, which is shown in Figure 2.21, the single description coder still performs better than the multiple description

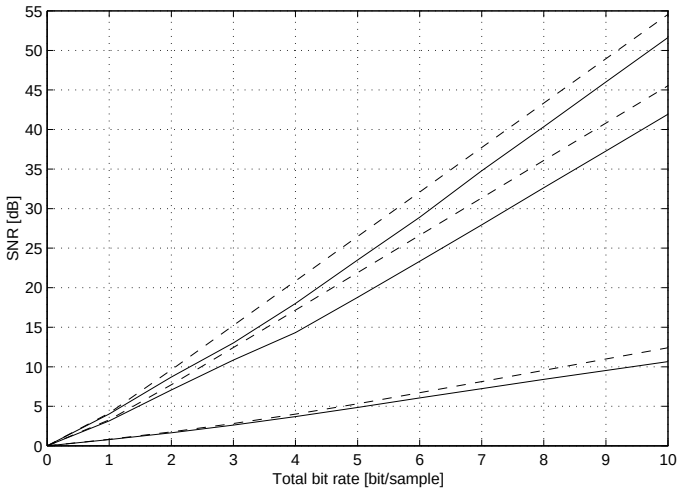


Figure 2.20 Rate-distortion performance for multiple (solid) and single (dashed) description coding with a block size of $L = 500$ source samples and bit error probabilities of (from below) 10^{-1}, 10^{-2} and 10^{-3}.

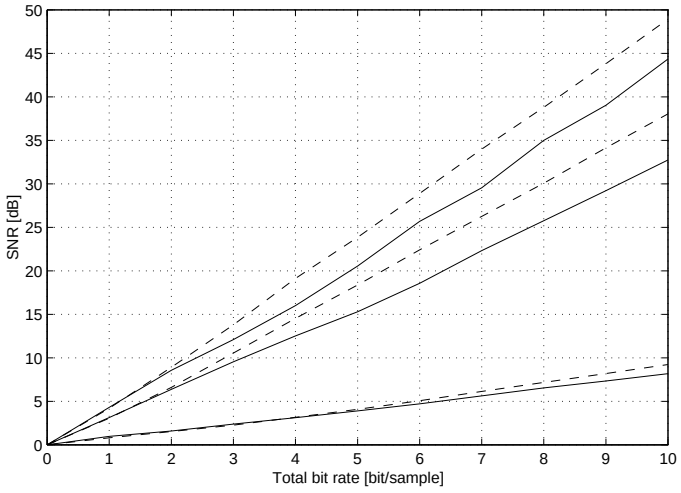
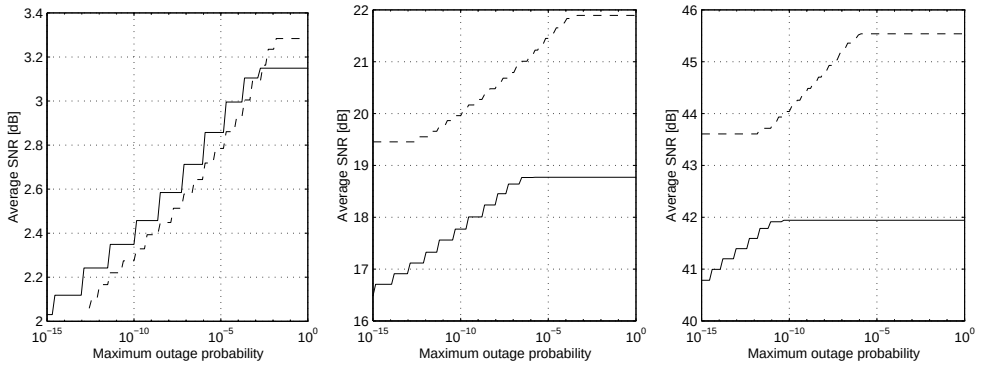


Figure 2.21 Rate-distortion performance for multiple (solid) and single (dashed) description coding with a block size of $L = 50$ source samples and bit error probabilities of (from below) 10^{-1}, 10^{-2} and 10^{-3}.



(a) $R = 1$ (b) $R = 5$ (c) $R = 10$

Figure 2.22 Performance of the multiple (solid) and single (dashed) description coders when the maximum outage probability is constrained. The bit error probability $p_b = 10^{-2}$, the block size is $L = 500$ source samples, and the total rate is R.

coder in all the displayed cases, except for $p_b = 10^{-1}$ and low rates. Here, the channel block length increases by rate, which explains why single description coding is much better at high rates, as opposed to short blocks of fixed channel block length in Figure 2.16(a), where the rate does not influence much on the performance difference for rates above a certain level.

When the block length of the error correcting code is doubled, the block error probability is normally reduced if channel bit rate R_c is the same. For very high bit error probabilities, this might not be the case. However, the delay limit gives only a maximum block size. Thus, shorter blocks can be chosen if they give better performance. In the graphs above, only the maximum block size was considered.

A comparison of the robustness of single and multiple description coding to varying channel quality, as was done in Figure 2.17, will come out more in favor of the single description coder when the source block size is constant, since the only thing that is changed, is that the maximum channel block length of the single description coder is twice that of the multiple description coder.

Also when the outage probability is constrained, the single description coder should come out more favorable when the source block length is the one that is kept constant. This is confirmed by Figures 2.22 and 2.23, which can be compared to Figures 2.18 and 2.19, respectively. In Figure 2.23, the side SNR measured in dB is again restricted to be at most one third of the central distortion.

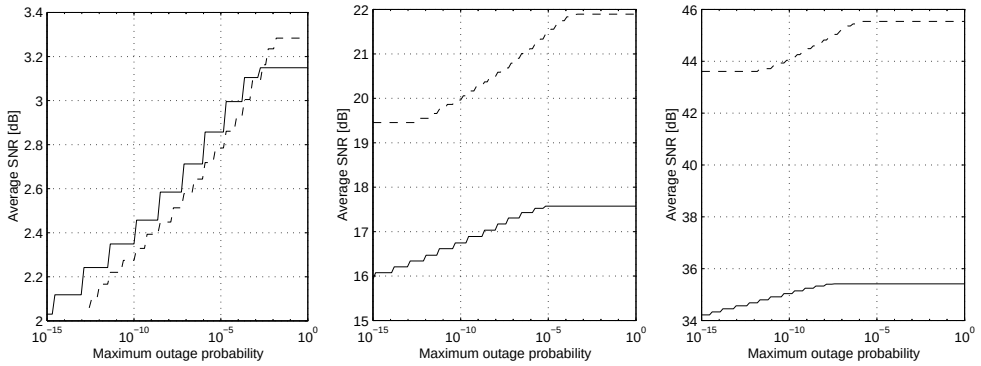


Figure 2.23 Performance of the multiple (solid) and single (dashed) description coders when the maximum outage probability and the side distortion are constrained. The bit error probability $p_b = 10^{-2}$, the block size is $L = 500$ source samples, and the total rate is R.

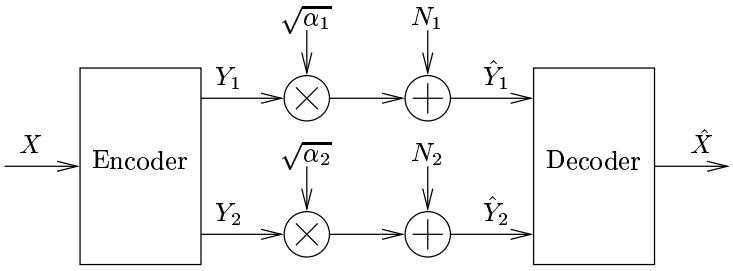


Figure 2.24 Channel model for block fading channel

2.7 Block Fading Channel

A block fading channel (Ozarow et al., 1994) with two independent channel realizations is considered. This channel is shown in Figure 2.24. The encoder takes the input X and creates two real-valued symbols Y_1 and Y_2, which are transmitted on the two channels. The fading is modeled through the multipliers, and the two factors $\sqrt{\alpha_1}$ and $\sqrt{\alpha_2}$ are independent and Rayleigh distributed and vary so slowly that their value can be assumed to be constant throughout the transmission of a block. The signal is corrupted by the independent additive white Gaussian noise signals N_1 and N_2.

The blocks consist of L source samples which are represented by n channel symbols on each channel. The system is characterized by the total rate R

which is defined as

$$R = \frac{2n}{L}, \tag{2.38}$$

that is, the number of channel symbols produced per source symbol. The long-term average power of the received symbols $\hat{Y}_i$, $i \in \{1, 2\}$, is given by

$$P = E\left[\hat{Y}_i^2\right], \tag{2.39}$$

and is assumed to be equal for both channels. It is assumed that the noise power $E[N_i^2] = 1$ for $i \in \{1, 2\}$. The receiver signal-to-noise ratio is defined as the ratio between the received signal power and the noise power, so in linear scale, this ratio will be equal to P.

The channel code is assumed to be based on coded modulation. In the multiple description coder, the channel encoder and decoder blocks of Figure 2.2 are coded modulation and demodulation. The source coder produces $M = R_s L$ bits, with $M/2$ in each description, from a block of L channel bits, and the channel coders produce n channel symbols from the $M/2$ bits. The channel code rate is defined as $R_c = \frac{M}{2n}$, so the total rate $R = R_s/R_c$. In the single description coder, the channel encoder and decoder blocks of Figure 2.3 are a coded modulation block followed by an interleaver, and a de-interleaver followed by coded demodulation, respectively. The coded modulation produces $2n$ channel symbols from the M bits of the source coder, and these channel symbols are distributed between the two channels by the interleaver. The channel code rate is still $R_c = \frac{M}{2n}$, and the total rate is thus $R = R_s/R_c$ as for the multiple description coder.

For this channel, it is impossible to find codes whose block error probability approaches zero as the block length increases. This means that the channel has no capacity in the traditional sense, and that the source–channel separation theorem does not hold. Thus, it is interesting to analyze the performance of the coders when the block length approaches infinity. Assume that the overall rate R, the channel rate R_c, and the source rate R_s are fixed, while the number of source symbols per vector, L, approaches infinity. Then, the number of channel symbols per block, $2n$ in the single description case and n in the multiple description case, will also approach infinity. The assumption that the fading state is constant throughout a block is still assumed to hold. In this case, expressions exist for bounds on the performance, and it is possible to get a performance arbitrarily close to the bounds. For the source coder, the performance bound is the rate-distortion function described in Section 2.4.1. For the channel coder, the performance is described below.

2.7.1 Channel Coder

The difference between the channel coders in the single and multiple description coding systems is that the SD coder uses both the channels with an interleaver, while the MD coder uses the channels separately. The performance bound for the block fading channels has been found by Ozarow et al. (1994). It cannot be described as a traditional channel capacity, as there is no code for which the block error probability approaches zero as the block length approaches infinity on this kind of channel. Instead, the bound is given as a block error probability as a function of the channel code rate R_c.

The block error probability can be stated as

$$\begin{aligned} p_e = \Pr[I_A < R_c] \Pr[\text{Block error} \mid I_A < R_c] \\ + \Pr[I_A \geq R_c] \Pr[\text{Block error} \mid I_A \geq R_c], \end{aligned} \quad (2.40)$$

where I_A is the capacity of the channel at a given fading state. This can be written as a bound:

$$p_e \geq \Pr[I_A < R_c] \Pr[\text{Block error} \mid I_A < R_c] \approx \Pr[I_A < R_c] \quad (2.41)$$

because the probability of block errors is close to 1 if the code rate is greater than the channel capacity. If the block length of the code approaches infinity, it is possible to get the second term in (2.40) to approach zero, Pr[Block error | $I_A < R_c$] will approach 1, and the block error rate will approach the last term in (2.41). This term will be denoted the *channel outage* probability p_{out}, and is used as an approximation of the performance bound. When the block length n approaches infinity, this is the actual bound.

The bound is found by assuming the signaling alphabet of the coded modulation increases, so that the probability density function of the channel output can approach Gaussian. The bound is given by (Ozarow et al., 1994)

$$p_e \approx p_{\text{out}} \geq \Pr\{\log_2(1 + \alpha P) < 2R_c\} \quad (2.42)$$

when one channel realization is used, that is, the multiple description case, and

$$p_e \approx p_{\text{out}} \geq \Pr\{\log_2(1 + \alpha_1 P) + \log_2(1 + \alpha_2 P) < 4R_c\} \quad (2.43)$$

when two channel realizations are used, that is, the single description case. Here, α, α_1, and α_2 are exponentially distributed random variables with a probability density function of

$$f_\alpha(\alpha) = \begin{cases} e^{-\alpha} & , \alpha \geq 0 \\ 0 & , \alpha < 0 \end{cases}, \quad (2.44)$$

and P is the received symbol power as defined in Equation (2.39). For the probability in Equation (2.42), a closed-form expression can be found:

$$p_{\text{out}} \geq \Pr\left\{\alpha < \tfrac{1}{P}\left(2^{2R_c}-1\right)\right\} = \int_0^{\frac{1}{P}\left(2^{2R_c}-1\right)} e^{-\alpha}\, d\alpha = 1 - e^{-\frac{1}{P}\left(2^{2R_c}-1\right)}. \tag{2.45}$$

The expression in Equation (2.43) can be written as:

$$\begin{aligned} p_{\text{out}} &\geq \Pr\left\{\log_2(1+\alpha_1 P) < 4R_c - \log_2(1+\alpha_2 P)\right\} \\ &= \Pr\left\{\alpha_1 > \tfrac{1}{P}\left(2^{4R_c - \log_2(1+\alpha_2 P)} - 1\right)\right\} \\ &= \int_0^{\frac{1}{P}\left(2^{4R_c}-1\right)} \int_0^{\frac{1}{P}\left(2^{4R_c-\log_2(1+\alpha_2 P)}-1\right)} e^{-\alpha_1} e^{-\alpha_2}\, d\alpha_1\, d\alpha_2 \\ &= \int_0^{\frac{1}{P}\left(2^{4R_c}-1\right)} \left(1 - e^{-\frac{1}{P}\left(2^{4R_c-\log_2(1+\alpha_2 P)}-1\right)}\right) e^{-\alpha_2}\, d\alpha_2. \end{aligned} \tag{2.46}$$

This integral will be solved numerically.

From Equation (2.40), it is seen that the block error probability can be dominated by the channel outage probability if it is much greater than the probability of a block error given that the channel is not in an outage state. This means that it is possible to get a block error probability that comes reasonably close to the channel outage probability even without using very long blocks. This problem is investigated by Knopp and Humblet (2000), and it turns out that it is possible to get the outage probability to be close to the bound with a small signaling alphabet.

2.7.2 Coder Models and Optimizations

Assuming that the block length approaches infinity and the source and channel coders reach the performance bounds, Equations (2.5), (2.6), and (2.7) can be used for the performance of the source coder, Equations (2.42) and (2.43) give the performance of the channel coder, and Equations (2.26) and (2.27) give the total distortion. The distortion can be found given the total rate R and the received signal power P. There are parameters to be decided. For both coders, the channel rate R_c can be varied. The source rate R_s will then be given by $R_s = R/R_c$. In the multiple description coder, the side distortion can also vary. If d_1 is chosen to a value that satisfies (2.6), then d_0 will be the smallest value satisfying (2.7). The minimum distortion can be found by numerical optimization of the distortion with R_c and d_1 as variables (only R_c in the single description case).

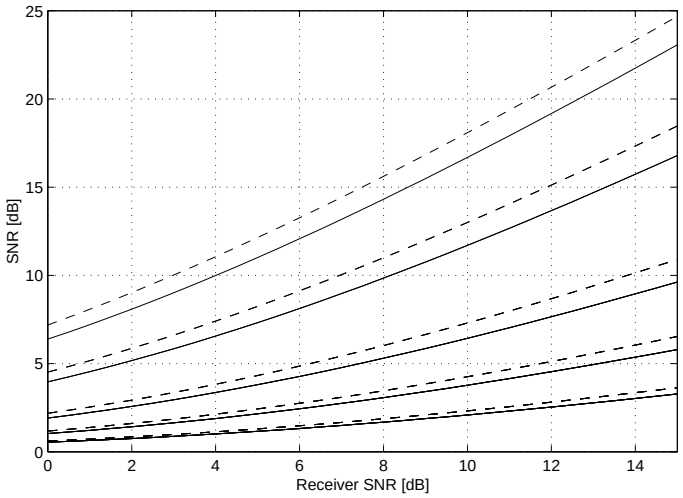


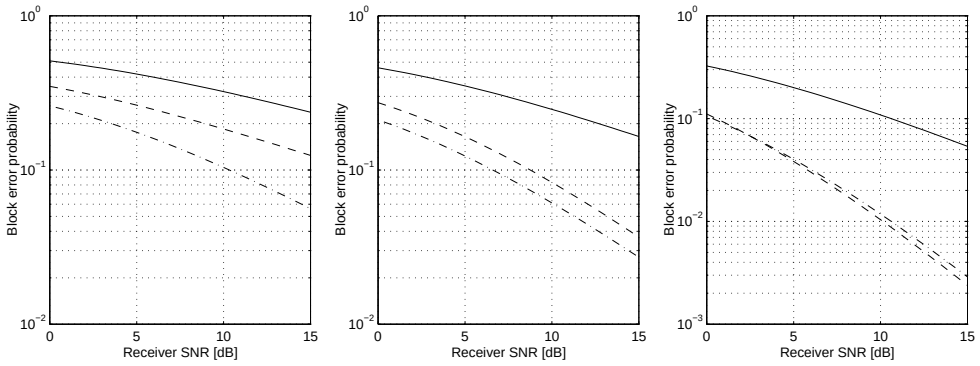
Figure 2.25 Overall SNR as a function of receiver SNR for multiple (solid) and single (dashed) description coding on a block fading channel for total rates R (from below) 0.5, 1, 2, 5 and 10.

It is more difficult to treat a more practical system based on finite block lengths. For the source coder, the uniform scalar quantizer as in Section 2.4.3 can be used. For the channel coder, to the author's knowledge there exists no bound that resembles the Gilbert bound for the binary symmetric channel, which would be an upper bound on the block error probability of the best practical coder of a given channel rate, block length, and signaling alphabet. It is possible to design a family of channel codes of different rates and evaluate their probability. However, the design of codes for coded modulation is outside the scope of this dissertation. In order to get an idea of the performance of a practical coder, results will be calculated using the uniform scalar quantizer performance for the source coder and the channel outage probability bounds for the channel coder. An example will be given in order to show the difference between a practical channel coder and the bound.

2.7.3 Results

In this section, the distortions obtained as described above are calculated for some values of the parameters. The parameters R_c and d_1 are found by minimizing the overall distortion numerically using Matlab's Optimization Toolbox. The integral in Equation (2.46) is also calculated numerically in Matlab.

In Figure 2.25, the performance of the two coders is shown for varying



(a) $R = 0.5$ (b) $R = 2$ (c) $R = 10$

Figure 2.26 Block error probability of the multiple (solid) and single (dashed) description coder along with the MD decoder outage probability (dash-dotted) as a function of receiver SNR for total rate R. The MD decoder outage probability is the square of the MD block error probability.

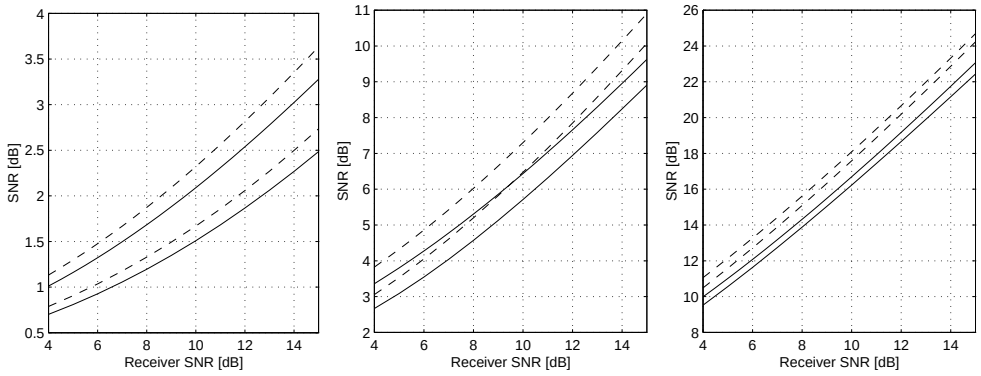
channel qualities and rates. The single description coder performs significantly better than the multiple description coder in all cases.

Figure 2.26 shows the block error probability for the optimized value of R_c in the single and multiple description case. In the single description case, a block error means that the receiver must output the source mean. This situation will be denoted a *decoder outage*, to distinguish it from the channel outage defined in Section 2.7.1. In the multiple description case, the decoder outage probability is p_e^2, and this probability is also shown in the graphs.

For a high rate, the decoder outage probability is about the same for both systems, actually a little higher for the multiple description coder. For lower rates, the multiple description coder has a smaller outage probability than the single description coder. Note that except for high rates and good channels, the outage probability is fairly high, in the order of 10^{-1}.

2.7.3.1 Practical Coder

In order to get an idea of how coders with finite block lengths would work, the rate-distortion function for the source coder is replaced by the performance functions of the uniform scalar quantizer of Section 2.4.3. The results are shown in Figure 2.27, and it is seen that the practical coder performs about 0.5–1 dB worse than the theoretical one. The difference between the single and the multiple description coder is approximately equal for the two source



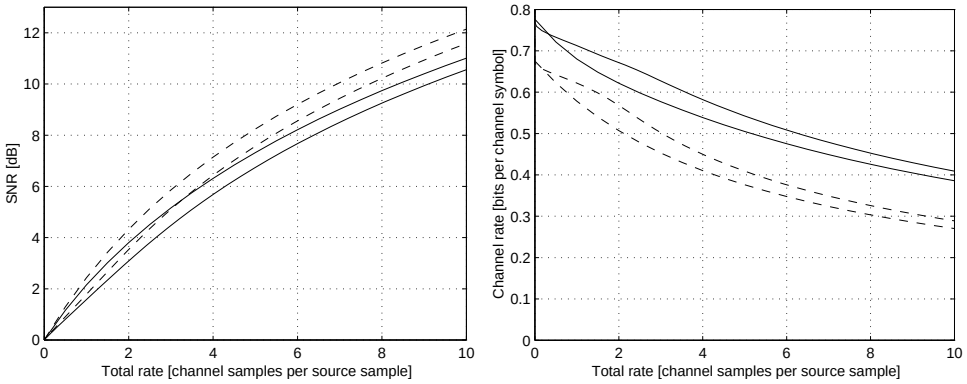
(a) $R = 0.5$ (b) $R = 2$ (c) $R = 10$

Figure 2.27 Performance of the coder with source coder evaluated as rate-distortion function (upper curve) or with uniform scalar quantizer (lower curve) for multiple (solid) and single (dashed) description coding with total rate R

models. By looking at receiver SNR values that give the same performance for the multiple and single description coder, it can be seen that the multiple description coder must have a channel that is 1–3 dB better.

Figure 2.28 shows the optimal channel code rate as a function of total rate for a channel with a receiver SNR of 5 dB, along with the SNR performance of the coding systems. The difference between the practical and theoretical source coder systems is small. The multiple description coder has a higher channel code rate, meaning that it has less error protection. This is expected, as the source code in this case is robust against errors. The best performing practical multiple description coder for the parameters of Figure 2.28 always has two diagonals in the index assignment, that is, high error robustness.

In Figure 2.29, the performance of practical channel coders of $R_c = 0.5$ bits per channel symbol is shown along with the bounds of Equations (2.42) and (2.43). The coders are half-rate convolutional codes with 16 or 64 states, with BPSK modulation, designed for one or two parallel channels. The block error probability has been estimated by simulating 10^5 blocks of length 128 bits. The generators for the convolutional codes can be found in (Knopp and Humblet, 2000) for two parallel channels, and e.g. in (Proakis, 2001, Table 8.2-1) for one channel. It is possible to design codes for other rates as well, as is done in (Knopp and Humblet, 2000) for more than two parallel channels. The results in Figure 2.29 show that the performance of a practical coder can be within approximately 0.5 dB of the bound for one channel (MD coding)



(a) SNR as a function of rate. Upper curve: theoretical. Lower curve: practical.

(b) Channel code rate as a function of total rate. Upper curve: practical. Lower curve: theoretical.

Figure 2.28 Performance of the coder with source coder evaluated as rate-distortion function (theoretical) or with uniform scalar quantizer (practical) for multiple (solid) or single (dashed) description coding. The receiver SNR is 5 dB.

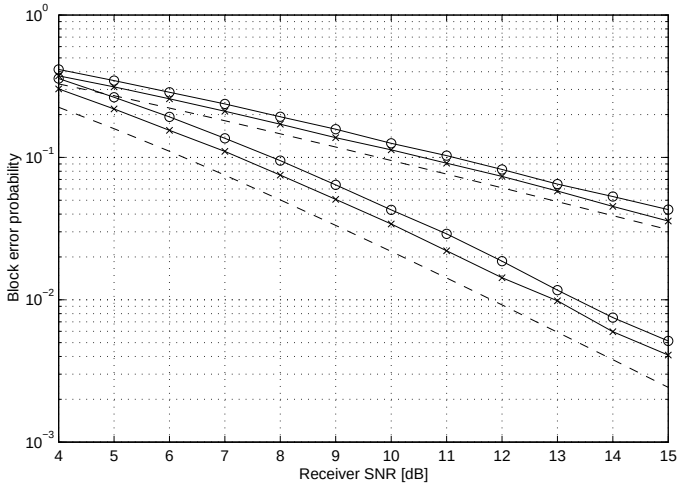


Figure 2.29 Performance of a practical channel coder of rate $R_c = 0.5$ bits per channel sample for one channel (upper curves) and two channels (lower curves). Solid with circles: 16 states. Solid with crosses: 64 states. Dashed: lower bound.

and 1 dB of the bound for two channels (SD coding). Comparing this to the results in Figure 2.27, shows that the difference between the performance of single and multiple description coding may be reduced, but single description coding will still perform better. This conclusion is drawn by assuming the difference between the performance of a practical coder and the bound will be approximately the same as in Figure 2.29 also for other channel code rates.

2.8 Discussion

The results in this chapter show that in most of the examples, multiple description coding gives a performance inferior to that of a traditional single description coding system. This is of course a disappointing result, but nevertheless an interesting one, considering the amount of work that exists on MD coding. It is essential to discuss the results in order to decide to which extent they discourage the use of MD coding for joint source–channel coding.

There are examples where MD coding does outperform SD coding. The common property of these examples is that the channel code is restricted: With the erasure channel and the BSC, it happens when the channel code is short, and thus not very powerful. With the BSC, MD coding better than SD coding was only documented for a codeword length of 63 bits on a poor channel and Gilbert bound performance; with actual codes, SD coding is better even there. The delay constraint imposed by a channel code of less than 63 bits is hardly a problem in multimedia transmission, so such poor codes is only an issue if the complexity of the channel coding units is a limiting factor. Communications hardware becomes more and more powerful, so even for low-cost applications, such a limit is not likely to apply.

For the erasure channel, each codeword is assumed to be long, a typical IP (internet protocol) packet size is in the order of 10–20 Kbits. Then, the delay constraints making MD coding better than SD coding are possible. For example, the transmission of an image of 512×512 pixels at a rate of 0.25 bit per pixel produces 64 Kbits, and this can then correspond to about 3–6 IP packets. In such a situation, MD coding can make sense on the erasure channel, and MD coding of images can be done with fairly good results (Servetto et al., 2000). On a binary symmetric channel or a block fading channel, the small symbol alphabet means that the block sizes will be much larger than 3–6, and the results showing that SD coding is better than MD apply.

Not very long channel codes are needed before SD coding becomes the better scheme on the erasure channel. In some applications, the erasures may not be independent, but there could still be two independent channels. In this case, the use of block codes for error protection might be much less powerful,

and MD coding could be an alternative.

If the outage probability is constrained for the BSC, MD coding is preferred in many cases. This is also a constraint on the channel code, since a code with low error probability is enforced, and the constraint is stricter for the SD code. Note that the distortion when one description is lost with the MD coder can be quite large: even if it is constrained to one third of the SNR in dB of the central SNR, for instance, there is a huge difference between a SNR of 30 dB and 10 dB. For low rates with the MD quantizer considered, it is not possible to find a solution with a much stricter constraint such as one half of the central SNR, however, it is possible for higher rates, where the difference between the side and central SNR is approximately 6 dB when the index assignment matrix has two diagonals. Then, a stricter constraint could be imposed, but results such as in Figure 2.19 would then come out more in favor of SD coding.

With the block fading channel, the MD coder is worse than the SD one for all the cases considered. For this channel, the source–channel separation theorem does not hold, so it was hoped that the multiple description coder could utilize the diversity obtained by two independent channels in an efficient manner. However, it turns out that this diversity is better utilized by the channel coder, because the variation of the channel quality is smaller when the two channels are turned into one by interleaving, and thus, the channel outage probability drops.

Besides the possibility for improvement of the performance for a specific channel quality, joint source–channel coding can be used for obtaining robustness against an unknown channel quality. Figure 2.17 shows that this does not happen with MD coding. This example applies to the BSC for a specific rate, but there is no reason why this could not be generalized. The performance curves as a function of channel quality are just as steep for MD coding as for SD coding, and this means that MD coding does not improve the robustness of a channel code. If another channel model or channel code is used, the steepness of the performance curves will change, but that applies both to MD and SD coding.

A very important question is whether the poor performance of MD coding demonstrated in this chapter can be generalized to other ways of exploiting multiple descriptions in joint source–channel coding. In all the examples except for the tightest delay constraints on the erasure channel, traditional channel coding is utilized. There are two reasons for this. First, to exploit multiple descriptions, the receiver must know whether to use a side or central decoder, and then, there must be a way of telling if errors means that the received data are usable or not. For the BSC, some sort of an error detecting code is the only way of telling this, assuming that the underlying physical channel is not

available. With a long error correcting code, it is usually a low probability of receiving a valid codeword that is not the correct one, and this means that the code is error detecting as well as error correcting. With the block fading channel, it can be possible to avoid the use of a channel code, since the fading state can be estimated, giving information on the probability of error. For the erasure channel, the errors are always detected, so a traditional channel code is not necessary. The second reason for using channel coding is that it provides good protection. In many cases, MD coding cannot provide enough robustness alone and has to be combined with another method. This applies to all the channel models.

There are some other publications on the use of MD coding in joint source–channel coding, with a different conclusion, stating that MD coding does give an improvement. Yang and Vaishampayan (1995) studied the use of MD coding on a Rayleigh fading channel. The differences between their work and the work in this chapter is that error correcting coding is not considered with the MD coder, the appropriate decoder is chosen from an estimate of the fading state in the receiver. In addition, fixed rate MD quantization is used, and this means that the delay is very short. A reference system using error correcting coding is considered, but with the same delay as the MD coder, the performance is lower. These results are certainly valid, but the delay introduced by longer channel codes and entropy coding is often acceptable, especially considering the improvement such coding can give.

For the case of a packet network model, MD coding in a joint source–channel coding setup has been considered by Alasti et al. (2001). They use a more sophisticated channel model than the extremely simple erasure channel model used here, although the channel model is still said to be largely simplified. The model is of a congestion network, and the paper concludes that MD coding is much better than SD coding for high network loading. This corresponds to a large erasure probability in the packet erasure channel model. Besides the different channel model, an important difference is that channel coding on the packets is not considered in (Alasti et al., 2001). Thus, that result is consistent with the result of Figure 2.10(a). The channel model considers two independent channels ("queues"), but packet losses on those two channels are not necessarily independent. That makes the use of channel codes on packets less efficient, and MD coding more suitable. The assumption of two independent channels seems to be motivated by the nature of MD coding with two descriptions, rather than by typical characteristics of actual networks, though.

It is definitely possible to find examples where the introduction of multiple description coding gives an improvement. The problem is that this happens in situations where there is a constraint on the channel coding, and this limits

the attainable performance. Of course, a delay constraint or other constraint on the channel code may very well occur, but they should not be made stricter than necessary. Also, the results obtained in this chapter indicate that quite harsh constraints are needed before MD coding is preferred.

It is clear that the variation of the delivered MSE is higher with SD than with MD coding. This is related to the fact that MD coding is preferred if the outage probability is constrained, at least for the BSC. If the variation of the MSE must be low, MD codes may be a solution. In image coding, for example, MD coders can be designed that give fairly good results both with the central and side decoders (Servetto et al., 2000). Even though better expected performance can probably be obtained with a single description coder, except when very few packets are used on an erasure channel, the performance could be quite unacceptable if important coefficients for the image decoding is lost completely, while a reduction of quality could be acceptable in some cases. This means that MD coding might offer a good trade-off between systems which almost always work, but with a limited quality, and systems that have a low average distortion, but a high probability of unacceptable decoding quality. This is a topic that must be investigated further in order to get clear conclusions, with comparisons between MD coding and other ways of achieving a low probability of unacceptably high distortion.

Chapter 3

Direct Source–Channel Mappings

This chapter considers joints source–channel coding methods that take continuous amplitude source symbols and produce potentially continuous amplitude channel symbols for transmission on an additive white Gaussian noise channel. Any combination of a source coder, channel coder and symbol mapping can be said to perform this operation. Direct source–channel mapping will be used to denote the process when the operation is performed without any intermediate steps, or where intermediate steps are designed in a way that takes the whole process into account. The mappings are assumed to be memoryless vector operators, the only memory allowed is given by the dimensionality of the input and output vectors of the mapping.

The aim of using direct source–channel mappings is twofold. It is hoped that the performance can be better than separate source coding, channel coding and modulation, and that robustness against mismatch between the expected and actual quality of the channel can be better.

The chapter is organized as follows. In Section 3.1, a mathematical definition of the problem is given. Section 3.2 presents a performance bound, and Section 3.3 presents previous work on the topic. In Section 3.4, a system for bandwidth doubling is proposed and expressions for its performance are derived. Section 3.5 presents coding results, and Section 3.6 provides a discussion of the results. Parts of the work in this chapter have been presented in (Coward and Ramstad, 1999; Coward and Ramstad, 2000*a*; Coward and Ramstad, 2000*b*; Coward and Ramstad, 2000*c*).

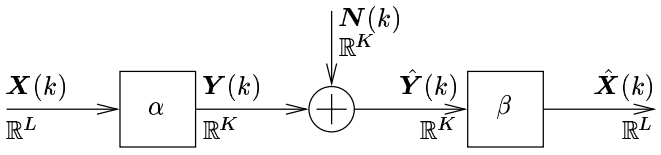


Figure 3.1 Overall system description

3.1 Problem Formulation

The system that is considered is shown in Figure 3.1. A source produces real-valued random vectors $\boldsymbol{X}(k)$ of length L. The elements of the vectors are denoted *source symbols*, and the L-dimensional space in which the vectors reside is called the *source space*. The vectors are transformed by an encoder mapping α,

$$\alpha : \mathbb{R}^L \to \mathbb{R}^K \tag{3.1}$$

into real-valued vectors $\boldsymbol{Y}(k)$ of dimension K, which are to be transmitted on a channel. The elements of these vectors are called *channel symbols*, and they are said to lie in the *channel space*. There is a constraint on the power consumption per channel symbol, given by

$$E\left[\boldsymbol{Y}^T(k)\boldsymbol{Y}(k)\right] = E\left[\alpha(\boldsymbol{X}(k))^T\alpha(\boldsymbol{X}(k))\right] \leq K\sigma_Y^2. \tag{3.2}$$

On the channel, the vectors $\boldsymbol{Y}(k)$ are contaminated by additive noise $\boldsymbol{N}(k)$, forming received vectors $\hat{\boldsymbol{Y}}(k)$. These are transformed by a decoder mapping β,

$$\beta : \mathbb{R}^K \to \mathbb{R}^L \tag{3.3}$$

into vectors $\hat{\boldsymbol{X}}(k)$. The aim of the system is to minimize the mean square error (MSE) D, defined by

$$D = \frac{1}{L}E\left[(\hat{\boldsymbol{X}}(k) - \boldsymbol{X}(k))^T(\hat{\boldsymbol{X}}(k) - \boldsymbol{X}(k))\right] \tag{3.4}$$

Throughout this chapter, the source will be assumed to be stationary and generate memoryless source symbols, i.e., $X_i(k)$ is statistically independent of $X_j(l)$ for $(i,k) \neq (j,l)$, where $X_i(k)$ is the ith component of $\boldsymbol{X}(k)$, and the statistical properties are independent of k and i. The same assumption is made for the noise, which in addition is assumed to have a Gaussian probability

distribution, giving an additive white Gaussian noise (AWGN) channel. For the source, Gaussian as well as Laplacian and uniform probability distributions will be considered.

The noise variance is defined as

$$\sigma_N^2 = \frac{1}{K} E\left[\boldsymbol{N}^T(k)\boldsymbol{N}(k)\right], \tag{3.5}$$

and the channel signal-to-noise ratio (CSNR, measured in dB) is defined as

$$\text{CSNR} = 10\log_{10}\frac{\sigma_Y^2}{\sigma_N^2}. \tag{3.6}$$

The source variance is

$$\sigma_X^2 = \frac{1}{L} E\left[\boldsymbol{X}^T(k)\boldsymbol{X}(k)\right], \tag{3.7}$$

while the signal-to-noise ratio (SNR) in dB of the system is defined as

$$\text{SNR} = 10\log_{10}\frac{\sigma_X^2}{D}. \tag{3.8}$$

The rate of the system is defined as

$$R = \frac{K}{L}, \tag{3.9}$$

i.e., the number of channel symbols produced per source symbol. When $K < L$, there are more source samples than channel samples per time unit, and this situation will be referred to as *bandwidth reduction*. Similarly, the situation when $K > L$ is called *bandwidth expansion*.

3.2 Optimal Performance Theoretically Attainable

The optimal performance theoretically attainable (OPTA) is the theoretical lower bound on the distortion with a coder as described in Section 3.1 given a rate $R = K/L$ and a CSNR (Berger and Tufts, 1967). Note that only the rate, not K and L, is specified, so very large values of K and L may be needed to approach the OPTA. The bound is found from Shannon's channel capacity and rate-distortion bounds (Shannon, 1948). If the channel capacity of the source is given by $C(S)$, where C is the capacity and S is the CSNR measured in linear scale, and the distortion-rate function is given by $\tilde{D}(\tilde{R})$, where $\tilde{D}$ is the distortion and $\tilde{R}$ is the rate, then the OPTA is given by $\tilde{D}(RC(S))$.

For an AWGN channel, the capacity measured in bits per channel symbol is given by (Shannon, 1948)

$$C\left(\frac{\sigma_Y^2}{\sigma_N^2}\right) = \frac{1}{2}\log_2\left(1+\frac{\sigma_Y^2}{\sigma_N^2}\right), \tag{3.10}$$

and for a memoryless Gaussian source, the mean square error distortion at the distortion-rate bound is given by (Shannon, 1948)

$$\tilde{D}(\tilde{R}) = \sigma_X^2 2^{-2\tilde{R}}, \tag{3.11}$$

so in this case, the OPTA is given by

$$D = \sigma_X^2 2^{-R\log_2(1+\sigma_Y^2/\sigma_N^2)} = \sigma_X^2\left(1+\frac{\sigma_Y^2}{\sigma_N^2}\right)^{-R}. \tag{3.12}$$

In this work, the AWGN is the only channel model considered. For source models different from Gaussian, the distortion-rate function can be found using a numerical algorithm (Blahut, 1972).

The distortion-rate function $\tilde{D}(\tilde{R})$ is convex and non-increasing (Berger, 1971), and from (3.10), it can be seen that the capacity of an AWGN channel $C(S)$ is a concave, non-decreasing function. From this, it can be shown that the OPTA distortion D is a convex, non-increasing function of S.

3.2.1 An Optimal System

It is well known that for a memoryless Gaussian source transmitted on an AWGN channel at a rate $R = 1$, a system exists that achieves the OPTA, see for instance (Goblick, 1965). Such a system is obtained by $K = L = 1$, simply by using the transmitter

$$Y(k) = \alpha(X(k)) = \frac{\sigma_Y}{\sigma_X}X(k) \tag{3.13}$$

and the receiver

$$\hat{X}(k) = \beta\left(\hat{Y}(k)\right) = \frac{\sigma_X\sigma_Y}{\sigma_Y^2+\sigma_N^2}\hat{Y}(k). \tag{3.14}$$

This gives the MSE

$$\begin{aligned} D &= E\left[\left(\hat{X}(k)-X(k)\right)^2\right] = E\left[\left(\frac{\sigma_X\sigma_Y}{\sigma_Y^2+\sigma_N^2}\left(\frac{\sigma_Y}{\sigma_X}X(k)+N(k)\right)-X(k)\right)^2\right] \\ &= \frac{\sigma_X^2\sigma_N^2}{\sigma_Y^2+\sigma_N^2} = \sigma_X^2\left(1+\frac{\sigma_Y^2}{\sigma_N^2}\right)^{-1}, \end{aligned} \tag{3.15}$$

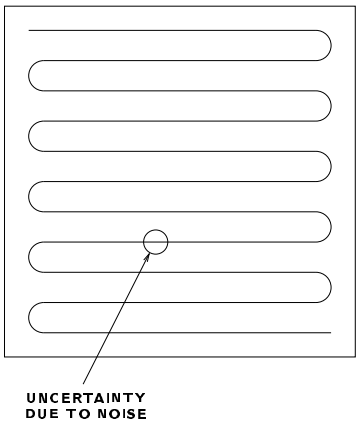


Figure 3.2 Mapping proposed by Shannon (1949)

which is the same as the OPTA for $R = 1$, proving the optimality of the system. This simple system will be denoted a direct PAM (pulse amplitude modulation) system, since a PAM channel sample is created directly from a source sample just by multiplication.

3.3 Previous Work

The concept of direct source–channel mappings was conceived already by Shannon (1949). Figure 3.2 is a reproduction of Figure 4 from Shannon's paper, and is a proposal for a bandwidth expanding mapping. The square represents the two-dimensional channel space in which signaling takes place, and the line winding through it represents the one-dimensional source space. The circle shows how noise in the channel space gives a limited distortion in the source space. Shannon points out that if the noise increases, the uncertainty can lead to an error in which portion of the line that is received, giving a large error. Shannon also mentions the bandwidth reduction case, suggesting a situation like in Figure 3.2, but where the square represents the two-dimensional source space and the line represents the one-dimensional channel space. Here, quantization must be performed by mapping a point in the source space onto a point on the line.

Later works have proposed methods for finding good source–channel mappings for different sources and channels. Berger and Tufts (1967) studied the use of linear PAM methods for various channels, and pointed out that for AWGN channels, one can reach the OPTA when $R = 1$, but not otherwise.

They also proposed a non-linear "Shannon-Cantor" method for other rates. Lee and Petersen (1976) developed the optimal linear block mapping given the statistical properties of the source and channel. This system is discussed further for a memoryless source and channel in Section 3.3.1.

A method that has been used in several varieties is the joint design of quantization and modulation sets. This is a natural combination, since quantization and modulation in a way represent inverse operations. Quantization is the mapping from a source space to an index, and inverse quantization is the mapping from an index to a set of points, the codebook, in the source space. Modulation, on the other hand, is a mapping from an index to a set of points in a channel space, while inverse modulation, or detection, is a mapping from the channel space to an index. Also, the nature of these mappings are usually quite similar, since both in quantization and detection, the set of points being mapped to one index is a simply connected set, and inverse quantization or modulation of this index gives a point contained in that set.

Ayanoğlu and Gray (1987) designed a joint source–channel code based on a trellis code, and Fischer and Marcellin (1991) joined the design of trellis coded quantization and trellis coded modulation, exploiting the similarity between the operations. The trellis introduces memory to the system, which strictly speaking brings it outside the definition of a source–channel mapping used here.

Vaishampayan (1989) performed a joint optimization of the quantizer and modulation sets. In order to make this possible, two simplifications were considered. One system, which was also treated in (Vaishampayan and Farvardin, 1992), considered the use of a linear receiver, while the other used detection demodulation. Liu, Ho and Cuperman (1993) performed optimizations of a joint system with a soft decoder, Skoglund and Hedelin (1999) used a Hadamard transform to design a vector quantizer for use with binary phase shift keying (BPSK) modulation, and Skinnemoen (1994) used Kohonen learning to optimize a vector quantizer for transmission by quadrature amplitude modulation (QAM). Fuldseth and Ramstad (1997*a*) followed up the case with detection demodulation from (Vaishampayan, 1989) with an improvement of the optimization algorithm. In (Fuldseth, 1997), this is elaborated with several more results, including both bandwidth reduction and expansion. The system is described in Section 3.3.2.

Mittal (1999) introduced a set of hybrid digital–analog coding systems with emphasis on robustness against varying channel qualities. By studying performance limits for the digital part and linear direct PAM for the analog part, achievable performance regions for a pair of channel qualities using the same encoder were found. A practical speech coder was obtained by using

a CELP (code-excited linear predictive) coder as the digital part. Furthermore, the duality between an encoder-decoder pair (α, β) and the coder (β, α) obtained by interchanging the encoder and decoder was studied. These topics are also treated in several papers: (Phamdo and Mittal, 2000; Mittal and Phamdo, 2000*a*; Mittal and Phamdo, 2000*b*; Mittal and Phamdo, 2000*c*), and other papers by the same authors. A practical implementation of a bandwidth expanding hybrid digital–analog coder for a general source was found by Skoglund, Phamdo and Alajaji (2000; 2001*a*), where a vector quantizer followed by BPSK modulation was used as the digital part. Skoglund, Phamdo and Alajaji (2001*b*) have also found a hybrid digital–analog system for bandwidth compression, where the analog and digital parts are added and a bandwidth reducing linear on non-linear mapping is used in the analog part. Coward and Ramstad (1999; 2000*a*; 2000*b*) designed a bandwidth expanding system with a multi-level scalar quantizer as the digital system, introducing some analog properties there as well. This system is elaborated in Section 3.4.

3.3.1 Block Pulse Amplitude Modulation

The block pulse amplitude modulation (BPAM) system was introduced by Lee and Petersen (1976)[1]. It is a linear system for transmitting a vector source on a vector channel with additive noise, which corresponds to Figure 3.1 where the operators α and β are multiplication by constant matrices. Then, the transmitter and receiver are given by

$$\boldsymbol{Y}(k) = \alpha(\boldsymbol{X}(k)) = \boldsymbol{A}\boldsymbol{X}(k) \tag{3.16}$$

$$\hat{\boldsymbol{X}}(k) = \beta(\hat{\boldsymbol{Y}}(k)) = \boldsymbol{B}\hat{\boldsymbol{Y}}(k), \tag{3.17}$$

respectively.

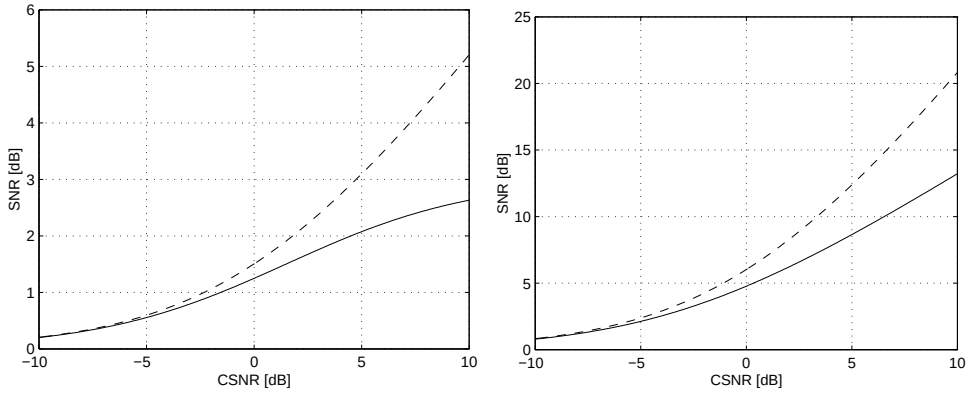
For the case of a vector source derived from a memoryless stationary scalar source and a white noise channel, a set of very simple matrices gives optimum performance. There is only one non-zero element in each row and each column of these matrices, and the transmitter and receiver matrices are given by

$$\boldsymbol{A} = \frac{\sigma_Y}{\sigma_X}\boldsymbol{I}_{K\times L} \qquad \boldsymbol{B} = \frac{\sigma_X\sigma_Y}{\sigma_Y^2+\sigma_N^2}\boldsymbol{I}_{L\times K} \tag{3.18}$$

when $K \leq L$, and

$$\boldsymbol{A} = \sqrt{\frac{K}{L}}\frac{\sigma_Y}{\sigma_X}\boldsymbol{I}_{K\times L} \qquad \boldsymbol{B} = \sqrt{\frac{L}{K}}\frac{\sigma_X\sigma_Y}{\sigma_Y^2+\sigma_N^2}\boldsymbol{I}_{L\times K} \tag{3.19}$$

[1] The term BPAM was not used by Lee and Petersen, but has been used by Vaishampayan (1989) and others.



(a) $R = \frac{1}{2}$ (b) $R = 2$

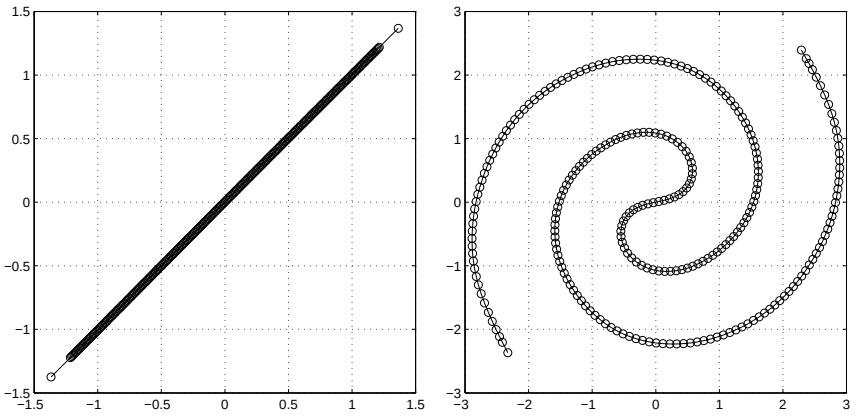
Figure 3.3 Performance of the BPAM system (solid) for different rates R compared to the OPTA (dashed) for a Gaussian memoryless source on an AWGN channel

when $K > L$, where $\boldsymbol{I}_{K\times L}$ is a $K \times L$ matrix with ones on the main diagonal and zeros elsewhere. This means that the BPAM coder simply removes samples to perform bandwidth reduction and inserts zero samples to perform bandwidth expansion. When $K = L$, the BPAM system is similar to a direct PAM system.

The performance of the BPAM system is shown in Figure 3.3 for a Gaussian memoryless source. The BPAM system has performance close to the OPTA for very poor channels (below 0 dB), but the gap between the performance and the OPTA increases with the channel quality. If $K = L$, the BPAM system achieves the OPTA, just as the direct PAM system.

3.3.2 Power-Constrained Channel-Optimized Vector Quantization

Power-constrained channel-optimized vector quantization (PCCOVQ) was proposed in (Fuldseth and Ramstad, 1997*a*; Fuldseth, 1997). The problem considered follows the framework of Section 3.1, and the mapping from source to channel space is performed by a vector quantizer followed by a modulation mapping, giving a mapping from source space to a finite number of points in channel space. The size of the vector quantization codebook is M. The channel symbols are uniformly spaced in each channel dimension, and the receiver uses nearest neighbor detection to pick a value from the reconstruction codebook. The encoding is done by choosing the symbol that minimizes the expected distortion under a power constraint. The distance between the channel symbols



(a) CSNR 0.26 dB (b) CSNR 20.0 dB

Figure 3.4 The reconstruction codebook of a PCCOVQ system with $L = 2$, $K = 1$, and $M = 256$, designed for a memoryless unit variance Gaussian source for transmission on an AWGN channel with a specified CSNR. The circles show the positions of the reconstruction codewords in the 2-dimensional source space, and the lines are drawn between source symbols which correspond to neighboring channel symbols.

and the reconstruction codebook are found by optimization using simulated annealing for given source statistics.

An interesting difference between PCCOVQ and a traditional system is that a high symbol error rate is a desired property. The reason is that when errors occur, the most probable decoded symbols are close in channel space to the transmitted symbol, and the optimization of the reconstruction codebook makes sure that such channel symbols correspond to source symbols of small Euclidean distance. This property is illustrated in Figure 3.4, where it can be seen that the codebook vectors lie on a line through the source space. If the error probability is made smaller while retaining the optimal shape, the number of points must be reduced, which means that the points on the line get further apart. Then, the distortion due to quantization increases without lowering the distortion due to channel noise, and the optimized PCCOVQ system will start to resemble a traditional unconstrained vector quantizer. The performance as a function of the channel quality then reaches saturation. The codebook size decides for what channel quality it happens.

Figure 3.4 also shows the difference in the shape of the codebook according to the CSNR. For a poor channel, the codewords form a straight line. This is equivalent to the BPAM system, except for the finite length. In BPAM, only

one of the two symbols are transmitted, which corresponds to a line along one of the axes of the source space, but the line in Figure 3.4(a) is a 45 degree rotation of that. Such an orthonormal transformation does not alter the performance of the BPAM system. Note that this BPAM equivalent occurs at a CSNR where BPAM has a performance close to the OPTA, cf. Figure 3.3(a). Figure 3.4(b) has the shape of a double spiral, which is more efficient than the shape in Figure 3.2 because the most probable symbols near the origin of the source space get a small amplitude, keeping the power consumption down.

In (Fuldseth, 1997) both bandwidth reduction, as in Figure 3.4, and bandwidth expansion, is considered. It turns out that while bandwidth reduction can give a performance only about 1 dB from the OPTA for memoryless Gaussian sources, bandwidth expansion performs much worse, especially for high CSNR values. This has several causes. One reason is that higher dimensional channel spaces need more points in the constellation to get the same minimum distance between the points in channel space, and the codebook size has to be kept down due to the complexity of the optimization algorithm. Another reason is that the fixed grid in the channel space prevents the algorithm from assuring that the probability of any error giving a high distortion is low enough. There is also an attempt to optimize the constellation in (Fuldseth, 1997), but this algorithm becomes very complex, preventing the use of a large codebook size.

3.4 Hybrid Scalar Quantizer–Linear Coder

In previous works on direct source–channel mappings, few, if any, good results are found for bandwidth expansion on channels with high CSNR. In this section, a coder that follows the hybrid digital–analog scheme of Mittal (1999)[2] is developed for memoryless sources.

In (Mittal, 1999), different hybrid digital–analog systems are presented. The system called System 2 is reproduced in Figure 3.5, using the notation of this chapter. Mittal assumes infinite memory in the digital coder, meaning that the digital source coder can reach the rate-distortion bound and the channel coder can reach the channel capacity. The linear encoder is introduced to bring robustness into the system so that it will work well for different channel qualities, and the system is analyzed to find achievable performance regions for a coder that can be used for two different noise levels.

In this work, all the signals in Figure 3.5 are scalar, and the tandem source and channel coder is replaced by a scalar quantizer. Robustness is achieved by limiting the number of quantizer levels and by using one channel symbol per

[2]The coder presented here was developed independently of Mittal's work.

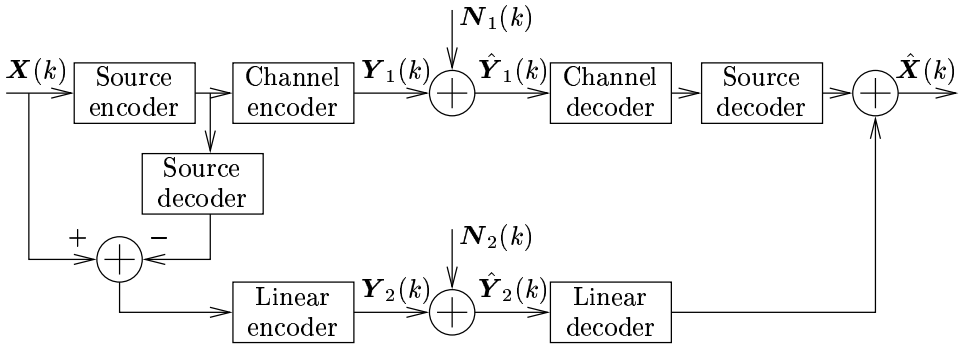


Figure 3.5 Block diagram of System 2 from (Mittal, 1999)

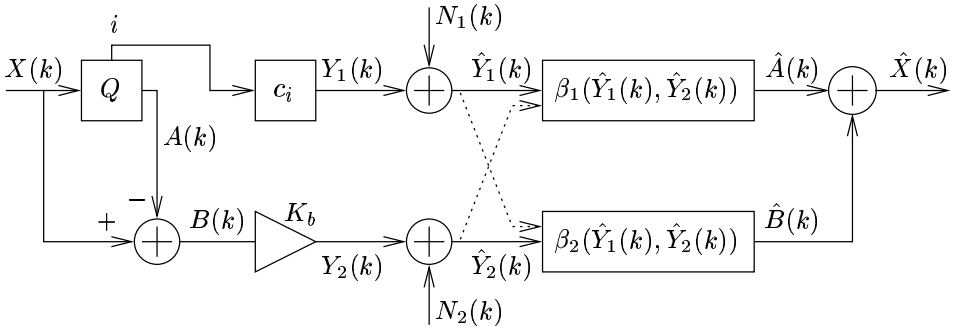


Figure 3.6 The HSQLC system

quantized symbol, where the quantizer levels and the corresponding channel symbols are ordered in the same way. This kind of system will be denoted a *hybrid scalar quantizer–linear coder* (HSQLC) system. Figure 3.6 shows the block diagram of the system. As opposed to the more general system of Figure 3.5, all the blocks are memoryless, and the two branches have to have the same rate. Thus, the source space has dimension $L = 1$ and the channel space has dimension $K = 2$, giving a total rate of $R = 2$.

3.4.1 Transmitter

The transmitter side of the proposed system works as follows. The symbols $X(k)$ are quantized using an M-level quantizer Q. The quantizer uses a set of decision intervals $\mathcal{D}_i = (d_i, d_{i+1}]$, $i \in \{0, \ldots, M-1\}$, where $d_0 = -\infty$ and $d_M = \infty$. It returns the index i and the corresponding representation

value $A(k) = a_i$. For transmission on the channel, the quantization index is represented by the symbol $Y_1(k) = c_i$. The quantization error $B(k) = X(k) - A(k)$ is multiplied by a constant K_b giving the symbol $Y_2(k)$, which is transmitted together with $Y_1(k)$ on the channel.

In order to meet the channel power constraint of Equation (3.2), the values of c_i and K_b must be chosen so that

$$\tfrac{1}{2}E\left[\boldsymbol{Y}^T(k)\boldsymbol{Y}(k)\right] = \tfrac{1}{2}\left(E[Y_1^2(k)] + K_b^2 E[B^2(k)]\right) \leq \sigma_Y^2. \tag{3.20}$$

When the MSE is to be minimized, (3.20) will be satisfied with equality. Note that it is not demanded that $Y_1(k)$ and $Y_2(k)$ have the same power as long as (3.20) is satisfied.

3.4.2 Receiver

Two different receivers will be used: One that is optimal given the transmitter, and a simplified one based on hard decision on the received quantizer symbols, and linear decoding of the quantizer error symbols.

3.4.2.1 Optimal Receiver

In general, the optimal estimate in the MSE sense of a symbol $U(k)$ given received symbols $v(j)$, $j \in \{k, k-1, \dots\}$, is (Gersho and Gray, 1992)

$$\hat{u}(k) = E[U(k) \mid v(k), v(k-1), \dots]. \tag{3.21}$$

Here, both the source, encoder and channel are memoryless, so the optimal receiver is given by

$$\begin{aligned} \hat{x}(k) &= E[X(k) \mid \hat{y}_1(k), \hat{y}_2(k)] \\ &= E[A(k) \mid \hat{y}_1(k), \hat{y}_2(k)] + E[B(k) \mid \hat{y}_1(k), \hat{y}_2(k)]. \end{aligned} \tag{3.22}$$

Thus, estimates $\hat{a}(k)$ and $\hat{b}(k)$ for $A(k)$ and $B(k)$ will be found, and they will be added in order to get an estimate $\hat{x}(k)$ of $X(k)$.

The receiver functions are denoted $\beta_1(\hat{y}_1(k), \hat{y}_2(k))$ and $\beta_2(\hat{y}_1(k), \hat{y}_2(k))$, respectively. When the receivers are known, the MSE can be found as

$$\begin{aligned} D &= E\left[\left(X(k) - \hat{X}(k)\right)^2\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [x - \beta_1(C_x(x) + n_1, K_b(x - Q(x)) + n_2) \\ &\qquad - \beta_2(C(x) + n_1, K_b(x - Q(x)) + n_2)]^2 \\ &\qquad\qquad f_N(n_1) f_N(n_2) f_X(x)\, dn_1\, dn_2\, dx, \end{aligned} \tag{3.23}$$

where $Q(\cdot)$ gives the quantized value drawn from $\{a_i\}$ and $C_x(\cdot)$ gives the corresponding value from $\{c_i\}$, and $f_X(x)$ and $f_N(n)$ are the probability density functions (pdf) of the input signal and the noise, respectively.

In Section B.1 of Appendix B, it is shown that the optimal receivers are:

$$\beta_1(\hat{y}_1, \hat{y}_2) = \frac{\sum_{i=0}^{M-1} a_i f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1} - a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db}{\sum_{i=0}^{M-1} f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1} - a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db} \tag{3.24}$$

$$\beta_2(\hat{y}_1, \hat{y}_2) = \frac{\sum_{i=0}^{M-1} f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1} - a_i} b f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db}{\sum_{i=0}^{M-1} f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1} - a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db} \tag{3.25}$$

For Gaussian, Laplacian, and uniform sources, it is possible to find closed-form expressions for the optimal receivers, see Sections B.1.1–B.1.3. However, to the author's knowledge, it is impossible to find a closed-form expression for the MSE obtained with these receivers. That makes them unsuitable for use with optimization.

3.4.2.2 Simplified Receiver

For the quantized symbol, a simplification of the receiver is the use of hard decision decoding, meaning that the receiver can only return the representation values $a_i,\ i \in \{0, \ldots, M-1\}$. The receiver returns a_i if the received value $\hat{y}_1(k) \in \mathcal{E}_i$, where $\mathcal{E}_i,\ i \in \{0, \ldots, M-1\}$ are disjoint sets such that $\bigcup_{i=0}^{M-1} \mathcal{E}_i = \mathbb{R}$. It is intuitively clear that $\mathcal{E}_i$ have to be intervals, and that one can write $\mathcal{E}_i = (e_i, e_{i+1}]$ provided that the c_i are sorted in ascending order, $c_{i+1} \geq c_i, i \in \{0, \ldots, M-2\}$.

This means that the receiver is given by:

$$\beta_1(\hat{y}_1(k), \hat{y}_2(k)) = \tilde{\beta}_1(\hat{y}_1(k)) = a_i \text{ when } \hat{y}_1(k) \in (e_i, e_{i+1}], \tag{3.26}$$

where $i \in \{0, 1, \ldots, M-1\}$ and $-\infty = e_0 \leq e_1 \leq \cdots \leq e_{M-1} \leq e_M = \infty$. Since this receiver function depends only on $\hat{y}_1(k)$, the notation $\tilde{\beta}_1(\hat{y}_1(k))$ is introduced.

For the quantization error symbol, a simplified receiver is achieved by having a receiver depending only on $\hat{y}_2(k)$, using the optimal linear receiver, that is the linear function

$$\tilde{\beta}_2(\hat{y}_2(k)) = \tfrac{\kappa}{K_b} \hat{y}_2(k) \tag{3.27}$$

that minimizes $E[(B(k) - \hat{B}(k))^2]$. Since both the source, encoder and noise are memoryless, the optimal linear receiver will be the same as the Wiener filter (Therrien, 1992) for $B(k)$ when $\hat{y}_2(k)$ is observed. It is easy to show that the optimal value of κ is

$$\kappa = \frac{K_b^2 \sigma_B^2}{K_b^2 \sigma_B^2 + \sigma_N^2}, \tag{3.28}$$

where σ_B^2 is the variance of $B(k)$. Thus, the optimal linear receiver for $\hat{y}_2(k)$ is given by

$$\hat{b}(k) = \beta_2(\hat{y}_1(k), \hat{y}_2(k)) = \tilde{\beta}_2(\hat{y}_2(k)) = \frac{1}{K_b} \frac{K_b^2 \sigma_B^2}{K_b^2 \sigma_B^2 + \sigma_N^2} \hat{y}_2(k). \tag{3.29}$$

Using these receivers, a closed-form expression for the MSE can be found for Gaussian, Laplacian and uniform sources, assuming that the decision intervals $\mathcal{E}_i$ of the hard decision decoder are known.

3.4.3 Optimizations

The performance of the HSQLC system will be optimized, minimizing the MSE under the power constraint given by Equation (3.20). In order to have a treatable expression for the MSE, the simplified receivers are used. In Section B.3, it is shown that the MSE in this case is given by

$$\begin{aligned} D = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} p_{i,j} \Bigg[& (1-\kappa)^2 \int_{d_i}^{d_{i+1}} x^2 f_X(x)\, dx \\ & + 2(1-\kappa)(\kappa a_i - a_j) \int_{d_i}^{d_{i+1}} x f_X(x)\, dx \\ & + \left((\kappa a_i - a_j)^2 + \frac{\kappa^2 \sigma_N^2}{K_b^2} \right) \int_{d_i}^{d_{i+1}} f_X(x)\, dx \Bigg], \end{aligned} \tag{3.30}$$

where $p_{i,j}$ is the probability of receiving a_j given that a_i was transmitted,

$$p_{i,j} = \int_{e_j - c_i}^{e_{j+1} - c_i} f_N(n_1)\, dn_1. \tag{3.31}$$

Expressions for the value of σ_B^2 which is needed to find κ and for the integrals over different source distributions are also given in Section B.3.

The mean square error given by (3.30) can be minimized numerically with respect to the quantizer parameters $a_i,\ i \in \{0, \ldots, M-1\}$, $c_i,\ i \in \{0, \ldots, M-$

$1\}$, $d_i,\ i \in \{1, \ldots, M-1\}$, $e_i,\ i \in \{1, \ldots, M-1\}$ and K_b. There is a constraint on the channel power as given by Equation (3.20). In order to impose this constraint, K_b is held outside the optimization and is instead found from the other parameters and the power constraint by solving (3.20) with equality. The value of $E[Y_1^2(k)]$ is then needed, and it is found in Section B.3. There are additional constraints that there exists a real-valued solution for K_b, i.e., that $E[Y_1^2(k)] \leq \sigma_Y^2$, and that $d_i \leq d_{i+1}$ and $e_i \leq e_{i+1}$ for all $i \in \{0, \ldots, M-1\}$. These constraints are imposed by the use of penalty functions (Walsh, 1975). The quantizer is assumed to be symmetric around zero, meaning that the number of parameters that need to be optimized is reduced approximately by a factor of 2. The symmetry assumption is justified by the symmetry of both the source and noise probability distribution. The numerical method used is the Nelder-Mead simplex (direct search) method as implemented in the Optimization Toolbox of Matlab (Coleman et al., 1999).

The optimization is performed for increasing noise levels. As initial conditions for the optimization, the result for the previous noise level is used. The initial condition for the lowest noise will be explained in Section 3.5.

3.5 Results

Coding results for the HSQLC system will be given for memoryless sources with three different source distributions: Gaussian, Laplacian, and uniform. The pdf of the distributions are given in Sections B.1.1–B.1.3. In addition, the performance of HSQLC and PCCOVQ is compared to a traditional system with separate source and channel coding.

3.5.1 HSQLC for Gaussian Source

The performance obtained with the HSQLC system with $M = 25$ quantization levels on a Gaussian memoryless source is shown in Figure 3.7. The results are shown both for the simplified receivers and with the use of the optimal receivers and the encoder found by the optimization. For the simplified receivers, the MSE is found from Equation (3.30), while with the optimal receivers, the MSE is found by Monte Carlo simulations, where 10^7 random source samples and a double number of random noise samples are generated for each noise level. The HSQLC system is compared to the OPTA, a PCCOVQ system (Fuldseth, 1997) with $L = 1$, $K = 2$, and $M = 256$ points in the constellation, BPAM (Lee and Petersen, 1976), and hybrid digital–analog—fixed encoder, adaptive decoder (HDA-FEAD) (Skoglund, Phamdo and Alajaji, 2001*a*). The latter coder consists of a vector quantizer of dimension 8 and codebook size

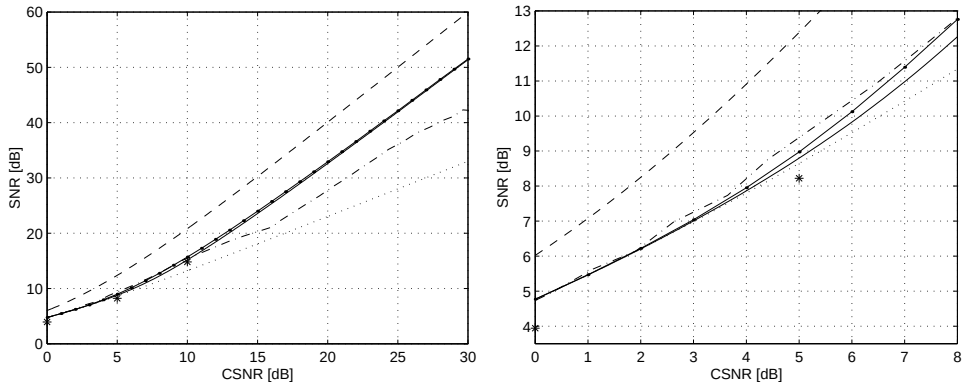
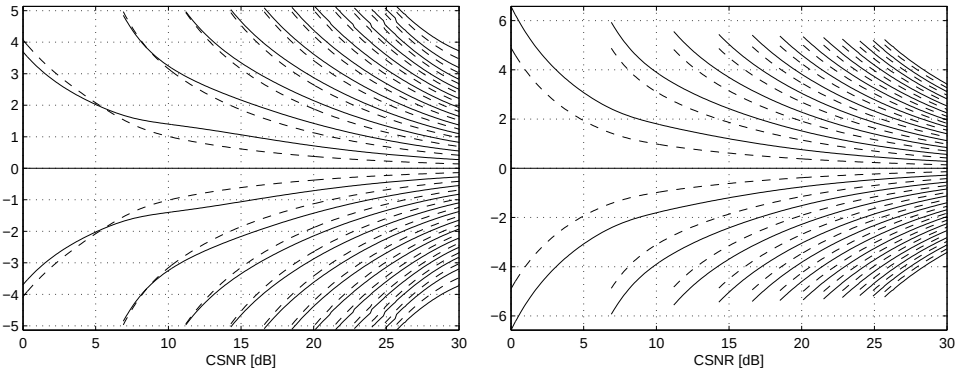


Figure 3.7 Performance of the HSQLC with simplified (solid) and optimal (solid with dots) receiver, compared to the OPTA (dashed), a PCCOVQ system (dash-dot), BPAM (dotted), and HDA-FEAD (stars), for a Gaussian memoryless source. All the systems have rate $R = 2$. The two graphs show the same results over different CSNR ranges.

256 being transmitted on a BPSK channel, and one quantization error symbol per source symbol transmitted with a linear coder. Both symbols have the same power, and the encoding vector quantizer is optimized given a certain channel noise, while the decoder can be re-optimized for the actual CSNR. In Figure 3.7, the design and actual CSNR are equal.

The results reveal that the proposed system is competitive with all the reference systems. For high CSNR values, it performs several dB better. The PCCOVQ system is slightly better than the HSQLC below 9 dB. The use of optimized receivers gives a slight improvement compared to the simplified ones (up to approximately 0.5 dB). However, it is not necessarily worth the increased complexity.

Figure 3.8 shows the quantizer levels obtained with the optimization. The spacing between the levels increases as the channel quality goes down, in order to keep the probability of decoding error of the quantized value low. This also means that the number of effective quantization levels goes down, since the probability of the outer quantization intervals becomes negligible. In the figure, decision levels d_i larger than 5 in magnitude have been removed, since the probability of a Gaussian source having that high a magnitude is only $5.7 \cdot 10^{-7}$. Numerical inaccuracy in the optimization process also means that decision levels larger than 5 and their corresponding representation values get a random-like behavior.



(a) Source representation values a_i (solid) and decision levels d_i (dashed)

(b) Channel representation values c_i (solid) and decision levels e_i (dashed)

Figure 3.8 Optimized representation values and decision limits for a zero mean unit variance Gaussian source on a unit power constrained AWGN channel as a function of CSNR. The number of decision levels is $M = 25$. Source decision levels d_i larger than 5 in magnitude have been removed together with the corresponding representation values.

Figure 3.7 shows that the HSQLC system performs similarly to the BPAM system for low CSNR values. This can be explained from Figure 3.8, since the decision levels for poor channels are quite large, meaning that the probability that the quantized symbol is different from zero is very low. Thus, the HSQLC approaches BPAM when the channel noise increases. This is also demonstrated by Figure 3.9, which shows the percentage of the total power occupied by the two symbols in the HSQLC. At intermediate to good channels, the two symbols get about the same amount of power, but for poor channels, the linearly encoded symbol gets almost all the power, and this explains the similarity to BPAM, where all the power is spent on a linearly encoded symbol.

Since the number of quantization levels $M = 25$ is odd and symmetry is assumed, the quantizer becomes a mid-tread quantizer with a representation level of zero. The optimization process can reduce the number of quantization levels, but not turn the quantizer into a mid-riser quantizer. To check the performance of a mid-riser quantizer, also the case of $M = 24$ has been tried, and the result is shown in Figure 3.10. The use of a mid-tread quantizer is slightly better from 5–12 dB, and there is no difference in performance at higher CSNR values. At CSNRs above 27 dB, $M = 25$ is also better, but that is probably due to the increased number of levels, since 24 or 25 levels is too little to obtain the optimal performance. Below 5 dB, the systems perform

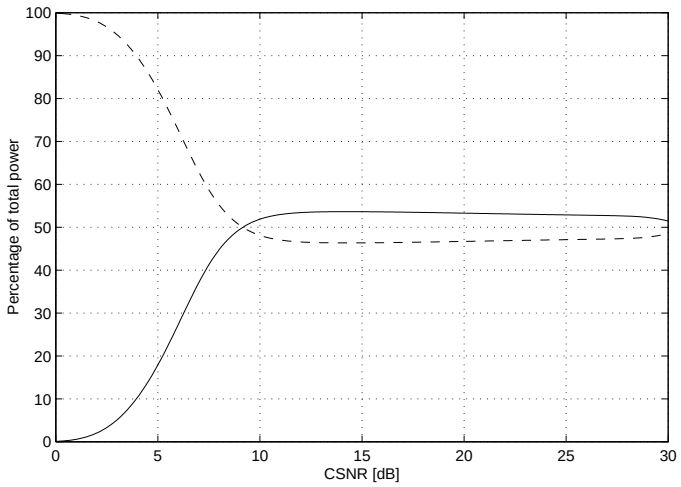
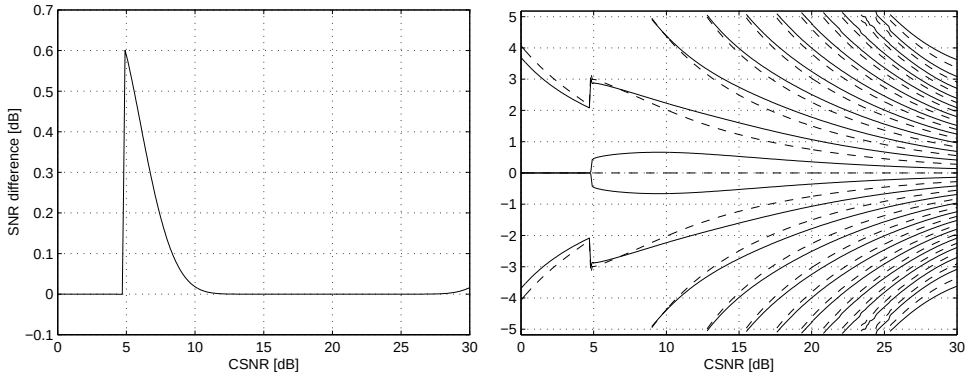


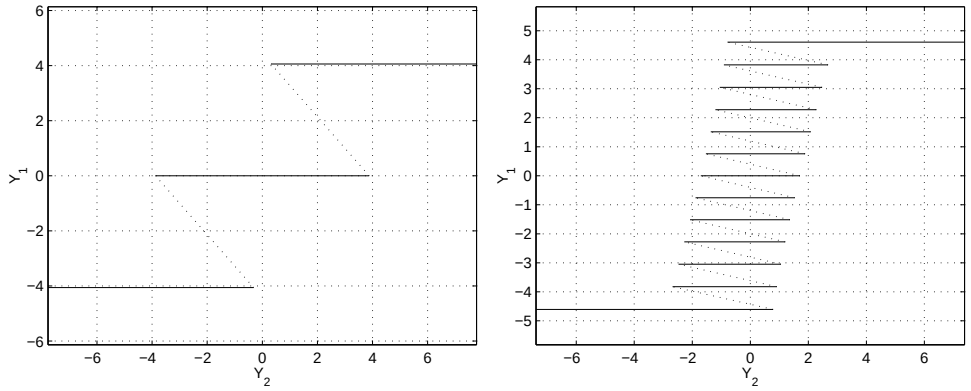
Figure 3.9 Percentage of the total channel power occupied by the quantized symbol $Y_1(k)$ (solid) and the quantization error symbol $Y_2(k)$ (dashed) as a function of CSNR with HSQLC on a Gaussian memoryless source



(a) Performance difference between $M = 25$ and $M = 24$ levels

(b) Source representation values a_i (solid) and decision levels d_i (dashed) using $M = 24$ levels

Figure 3.10 A mid-riser based HSQLC system with simplified receivers and $M = 24$ levels for a unit variance Gaussian source



(a) CSNR 3 dB (b) CSNR 20 dB

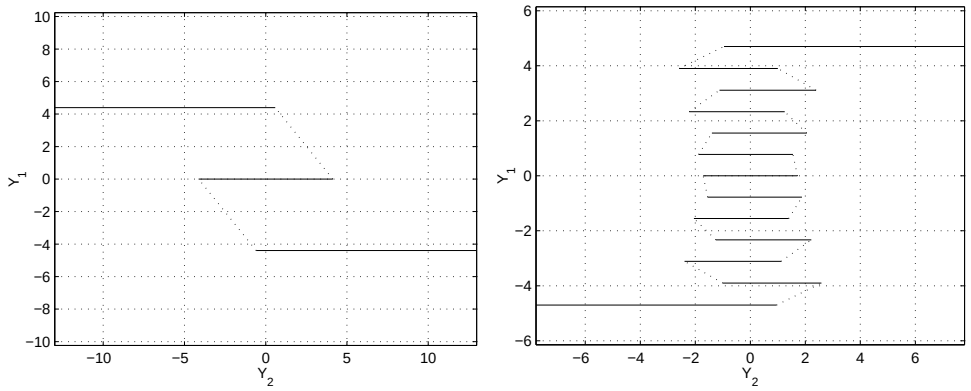
Figure 3.11 HSQLC mappings plotted in channel space for a Gaussian memoryless source and unit channel power constraint

equally again. The reason for this can be seen from Figure 3.10(b), which shows that two representation levels go to zero, creating a de facto mid-tread quantizer. The results show that a mid-tread quantizer can always be used without losing optimality.

Figure 3.8 is one way of visualizing the coder operation. Another way of doing it is shown in Figure 3.11. There, the two-dimensional channel space gives the axes, and the solid line represents the points in the channel space that are used for transmission. Neighboring points on a line segment represents neighbors in the one-dimensional source space, and the dotted lines show which end points of the line segments that are neighbors in the source space. Each solid line is shifted slightly to the left with respect to the line above it; this reduces the consequence of an erroneously decoded quantized symbol. Figure 3.11(a) shows why the representation values are outside the decision intervals for low CSNRs in Figure 3.8(a), it simply means that the outer line segments are shifted so far that they only take up one side of the second axis.

3.5.1.1 Modification of the Shape of the Mapping

If Figure 3.11 is compared to Figure 3.2, two differences come to one's attention. The line in the two-dimensional channel space is discontinuous, which is due to the simplification that a scalar quantizer represents. The other difference is that in Figure 3.2, the line segments go back and forth, while in Figure 3.11, all the line segments go in the same direction. This can be changed by changing



(a) CSNR 3 dB (b) CSNR 20 dB

Figure 3.12 HSQLC mappings for a Gaussian source when the sign of $B(k)$ is changed for every second quantizer interval, plotted in channel space for unit channel power constraint

the sign of $B(k)$ if the quantization index i is odd (or even) at the encoder side, cf. Figure 3.6, and changing it back in an appropriate manner at the decoder. The mapping resulting from doing so, performing new optimizations of the parameters, is shown in Figure 3.12. This mapping has a closer resemblance to Figure 3.2. Its performance compared to the original HSQLC system is shown in Figure 3.13. The original system has better performance. This can be explained by looking at the squared error if the quantized symbol is erroneously decoded and the neighboring symbol is received. In the original system, the error is approximately constant regardless of the value of the quantization error symbol. In the modified system, the error is small if the transmitted channel symbol lies close to a conjunction of line segments and large if it is far from the conjunction. On average, this gives a larger squared error, since the square operator amplifies large values.

3.5.1.2 Uniform Quantizer

By studying Figure 3.8, it can be observed that the spacing between the quantizer levels is close to uniform for any CSNR, except for the highest values, where the outer levels have a larger spacing. This applies especially for the decision limits d_i and e_i, while the representation values a_i and c_i are not quite uniformly spaced. This observation suggests a simplification of the design procedure. If the quantizer levels are forced to be uniform, only the spacing

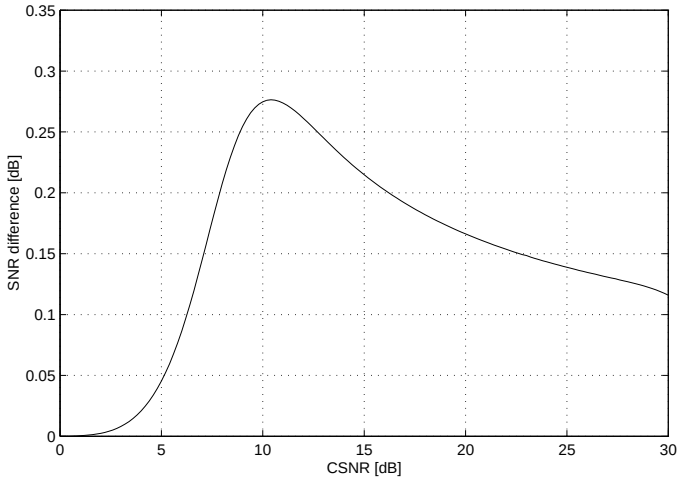


Figure 3.13 Performance difference between the original HSQLC system and the system where the sign of $B(k)$ is changed for every second quantizer interval. Both systems use simplified receivers and $M = 25$ quantizer intervals.

between them needs to be specified, which reduces the number of optimization parameters drastically. If M levels are used, the parameters can be found from

$$a_i = \left(i - \tfrac{M-1}{2}\right)\delta_a,\ i \in \{0, \ldots, M-1\}, \tag{3.32}$$

$$c_i = \left(i - \tfrac{M-1}{2}\right)\delta_c,\ i \in \{0, \ldots, M-1\}, \tag{3.33}$$

$$d_i = \left(i - \tfrac{M}{2}\right)\delta_d,\ i \in \{1, \ldots, M-1\}, \tag{3.34}$$

$$e_i = \left(i - \tfrac{M}{2}\right)\delta_e,\ i \in \{1, \ldots, M-1\}, \tag{3.35}$$

so the optimization can be performed with respect to only the four parameters δ_a, δ_c, δ_d and δ_e. The value of M is not an important parameter, provided it is large enough so that the probability of the outer levels is small. In theory, it could be set to infinity, but in order to calculate the MSE from (3.30), a finite value is needed. M is chosen so that $d_{M-1} \geq 6$, making sure the probability of $X(k)$ being in the outer intervals of the quantizer is very small (less than $2 \cdot 10^{-9}$). M is always odd, giving a mid-tread quantizer.

The performance difference between the original system and the system with uniform quantizer levels is shown in Figure 3.14. The result shows that a quantizer where all levels are optimized is only up to 0.01 dB better than a quantizer with uniform distance. For CSNRs above 27 dB, the uniform quantizer levels perform slightly better. This is because the number of quantizer levels, $M = 25$, used in the original optimization, is too small, while with the

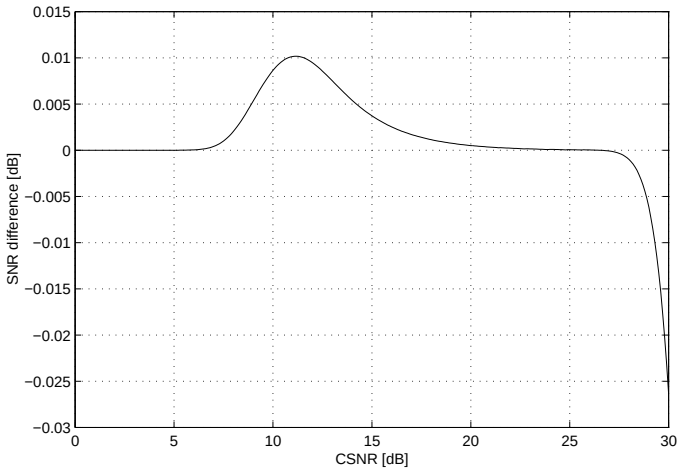


Figure 3.14 Difference between the original HSQLC and the HSQLC with uniformly spaced quantization levels for a Gaussian source. The simplified receiver is used in both cases.

uniform levels, a large number of quantization intervals is easily feasible. This also explains the non-uniform spacing of the decision levels shown in Figure 3.8 for good channels. This non-uniformity is simply due to the limited number of quantizer levels, and will not occur if M is increased. The results plotted in Figure 3.14 are for simplified receivers, but the optimal receivers can be used with the uniform quantizers just as well as with the original ones.

As mentioned earlier, the initial values for optimization of the quantizer is the optimized quantizer for a slightly higher CSNR. This applies to all the varieties of the system. The initial value for the highest CSNR, 30 dB, was originally taken as a Lloyd-Max quantizer (Gersho and Gray, 1992) with M levels for a Gaussian source. This was done for a smaller value of M than 25. When it turned out that the optimal quantizer is approximately uniform, the initial value for a high number of quantizer levels was taken by using an optimized quantizer for a lower number of levels and interpolating the extra levels. For the system with alternating sign of $B(k)$, the quantizer of the ordinary system for 30 dB was used as an initial value, and with the uniform quantizer system, the spacing of the innermost levels of the ordinary system was used as initial values.

3.5.1.3 Channel Mismatch

In many applications, the encoder does not know the actual channel, and in an application like broadcasting, each receiver has a different channel, meaning

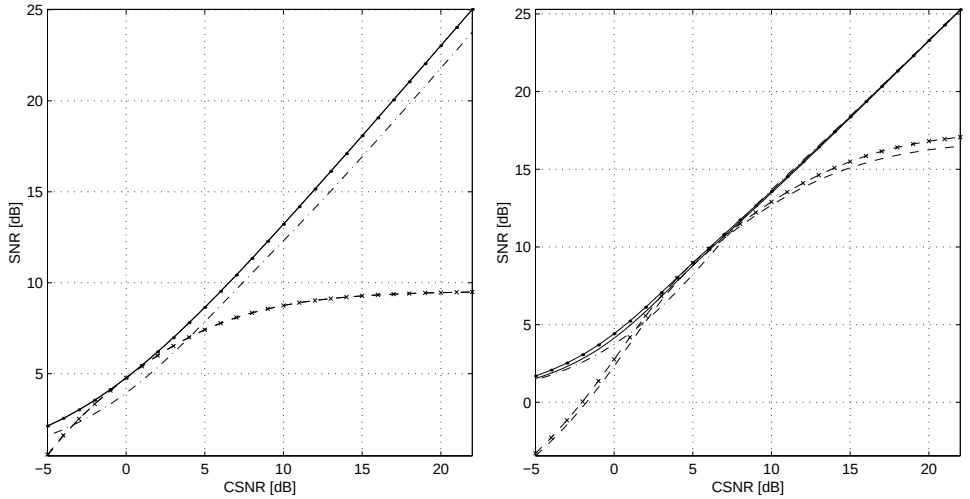
that it is impossible to design an encoder that is optimal for all parties. When studying the performance for such a situation, a design CSNR is assumed for the encoder. For the decoder, there are two possibilities, either it can be designed for the same CSNR as the encoder, or it can be re-designed from the actual CSNR, provided the receiver has a means of estimating the channel quality. If optimal quantizers are used, Equations (3.24)–(3.25) can be used with the actual noise level. If simplified receivers are used, κ must be re-calculated from (3.28), and new values of e_i must be found. The latter could be done by new numerical optimizations, but a simpler way is the use of a maximum a posteriori (MAP) receiver, explained in Section B.2.

The result of channel mismatch for different systems is shown in Figure 3.15. For low design CSNR values, a lot can be gained by re-designing the decoder for the actual CSNR, while this gain is much smaller for high design CSNRs. The use of optimal receivers gives a substantial gain if the design CSNR is high, the actual CSNR is lower, and the decoder is re-designed. The gain is small (or even negative) in most other cases. The HSQLC system is compared to the HDA-FEAD system from (Skoglund et al., 2001*a*). That system has a decoder which is optimized for the actual CSNR. The proposed system with a re-designed decoder performs better than the HDA-FEAD for all cases except when the design CSNR is 5 dB and the actual CSNR is 9 dB or more, where the HDA-FEAD is up to 0.1 dB better. The fact that HDA-FEAD performs best compared to the HSQLC for a design CSNR of 5 dB can be explained from the use of BPSK signaling for the digital part. For 0 dB, it would be better to use a BPAM-like structure, (cf. Figure 3.7, where HDA-FEAD is inferior to BPAM at 0 dB), and at 10 dB, the number of levels in the quantized symbol should be larger than two. At 5 dB design CSNR, the HDA-FEAD is also inferior to BPAM, but when the actual CSNR is higher, the BPSK is useful.

3.5.2 HSQLC for Laplacian Source

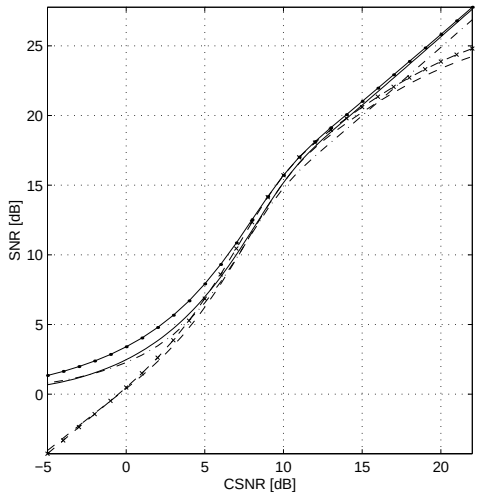
Figure 3.16 shows the performance of the HSQLC, both with optimal and simplified receivers, for a Laplacian source. This time, the number of quantizer levels is $M = 31$. A greater number of quantizer levels is needed than for the Gaussian source because the probability of large amplitudes is greater for the Laplacian distribution. The performance is compared to the OPTA, a PCCOVQ system with $L = 1$, $K = 2$, and $M = 256$ constellation points[3], and to the BPAM. The OPTA gives a higher SNR than for a Gaussian source, while

[3]This system was not reported in (Fuldseth, 1997), but it is designed using the methodology from that work.

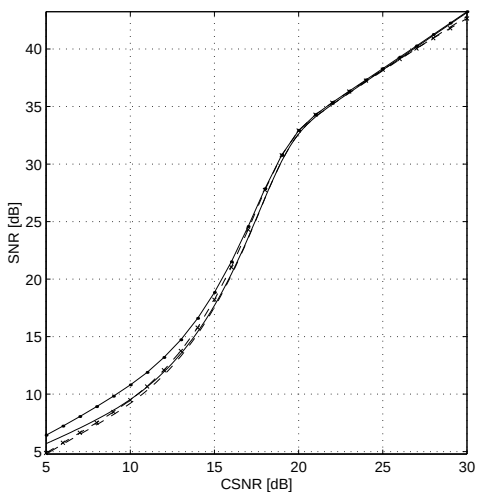


(a) Design CSNR 0 dB

(b) Design CSNR 5 dB



(c) Design CSNR 10 dB



(d) Design CSNR 20 dB

Figure 3.15 Performance at CSNR mismatch for the HSQLC and the HDA-FEAD with a Gaussian source. The solid lines represent the HSQLC with optimal (with dots) or simplified (without dots) receiver designed for the actual CSNR, the dashed lines represent the HSQLC with optimal (with crosses) or simplified (without crosses) receiver designed for the design CSNR, and the dash-dotted line represents the HDA-FEAD system ((a)–(c) only).

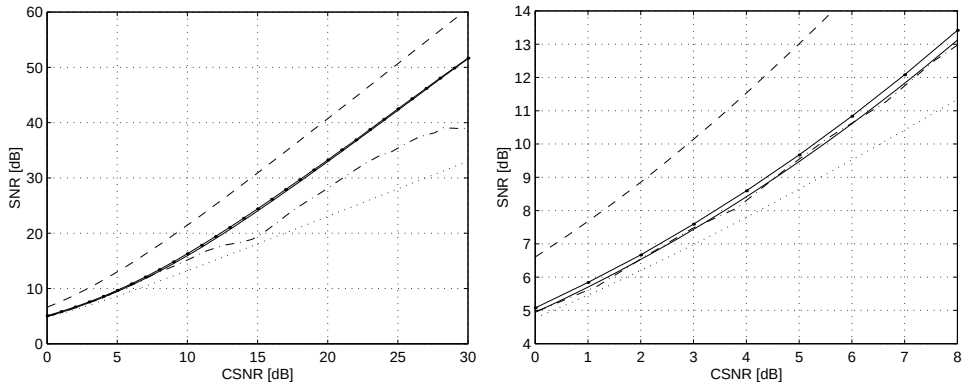


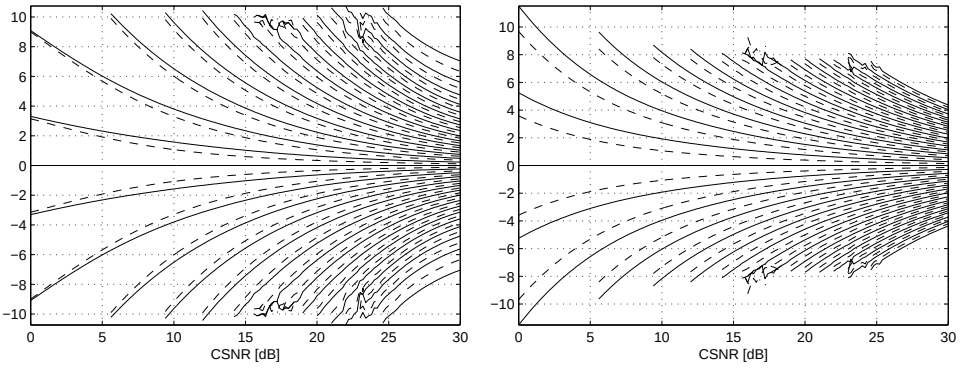
Figure 3.16 Performance of the HSQLC with simplified (solid) and optimal (solid with dots) receiver, compared to the OPTA (dashed), a PCCOVQ system (dash-dot), and BPAM (dotted), for a Laplacian memoryless source. All the systems have rate $R = 2$. The two graphs show the same results over different channel ranges.

the BPAM is unchanged since it is linear. Again, the use of optimal receivers gives a small improvement compared to the simplified ones. The PCCOVQ system performs about equally to the HSQLC with simplified receivers up to approximately 8 dB, but worse than HSQLC with optimal receivers. Above that, HSQLC is clearly better with either receivers.

The quantizer optimized for the Gaussian source was used as initial values for the optimization at 30 dB. The extra levels were interpolated.

The optimized quantizer levels for the Laplacian source are shown in Figure 3.17. Here, source decision levels d_i larger that 10 in magnitude have been removed, the limit has been set higher than for the Gaussian distribution because of the increased probability of large amplitudes. There is some noise on the quantization levels of high amplitudes due to numerical inaccuracy in the optimization process. The significance of this on the performance is negligible.

With the Gaussian source, the probability of the quantized value being anything but zero was very low for CSNR values close to 0 dB. Here, that probability is somewhat higher. This is also illustrated by Figure 3.18 which shows that even at 0 dB, 2 % of the total power is spent on the quantized symbol. This is not much, but enough to distinguish the proposed system from BPAM, something which can also be seen by the performance of those two systems in Figure 3.16.



(a) Source representation values a_i (solid) and decision levels d_i (dashed)

(b) Channel representation values c_i (solid) and decision levels e_i (dashed)

Figure 3.17 Optimized representation values and decision limits for a zero mean unit variance Laplacian source on a unit power constrained AWGN channel as a function of CSNR. The number of decision levels is $M = 31$. Source decision levels d_i larger than 10 in magnitude have been removed together with the corresponding representation values.

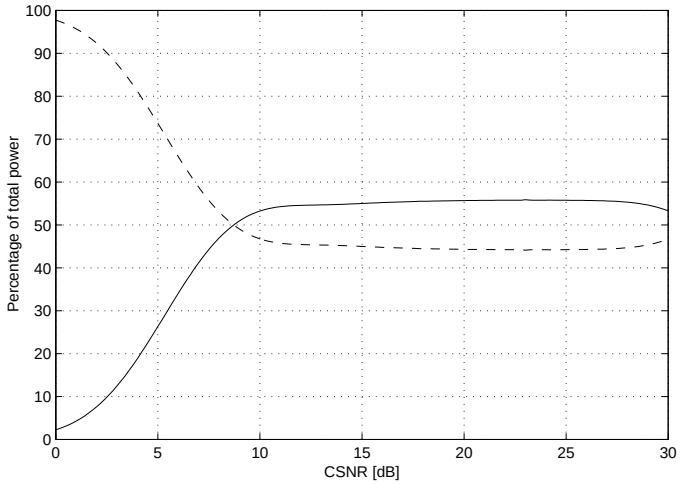


Figure 3.18 Percentage of the total channel power occupied by the quantized symbol $Y_1(k)$ (solid) and the quantization error symbol $Y_2(k)$ (dashed) as a function of CSNR with HSQLC on a Laplacian memoryless source

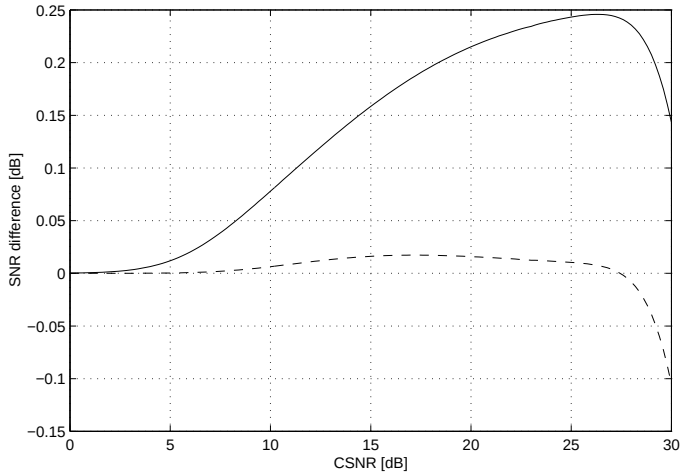


Figure 3.19 Difference between the original HSQLC and the HSQLC with uniform (solid) and quadratic (dashed) quantization levels for a Laplacian source. The simplified receiver is used in all cases.

3.5.2.1 Systematic Quantizer

The use of a uniform quantizer gave a very good result for the Gaussian distribution. The optimized quantizer for the Laplacian distribution is somewhat further from uniformity. Still, the use of a quantizer with levels given by few parameters hugely simplifies the optimization procedure. Thus, two systematic quantizers will be tried: once again a uniform quantizer as given by Equations (3.32)–(3.35), and a quantizer given by the quadratic functions:

$$a_i = \left(i - \tfrac{M-1}{2}\right)\zeta_a + \left(i - \tfrac{M-1}{2}\right)\left|i - \tfrac{M-1}{2}\right|\xi_a,\ i \in \{0, \ldots, M-1\} \quad (3.36)$$

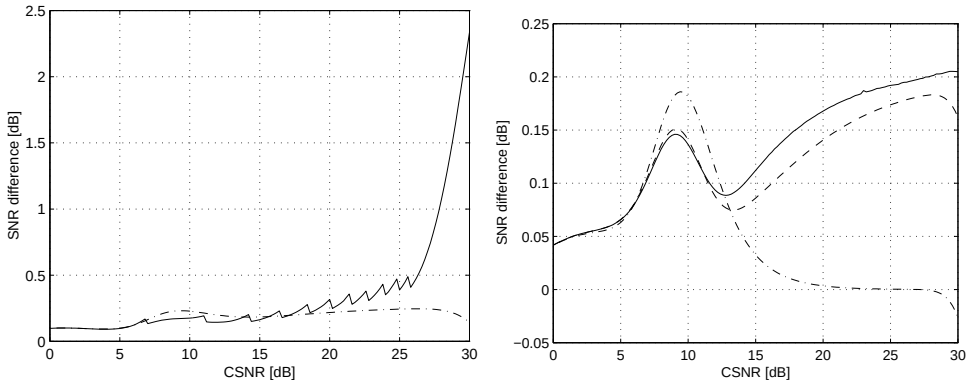
$$c_i = \left(i - \tfrac{M-1}{2}\right)\zeta_c + \left(i - \tfrac{M-1}{2}\right)\left|i - \tfrac{M-1}{2}\right|\xi_c,\ i \in \{0, \ldots, M-1\} \quad (3.37)$$

$$d_i = \left(i - \tfrac{M}{2}\right)\zeta_d + \left(i - \tfrac{M}{2}\right)\left|i - \tfrac{M}{2}\right|\xi_d,\ i \in \{1, \ldots, M-1\} \quad (3.38)$$

$$e_i = \left(i - \tfrac{M}{2}\right)\zeta_e + \left(i - \tfrac{M}{2}\right)\left|i - \tfrac{M}{2}\right|\xi_e,\ i \in \{1, \ldots, M-1\} \quad (3.39)$$

The latter quantizer will be referred to as a quadratic quantizer. It is characterized by eight parameters, and although this is more than the four parameters characterizing the uniform quantizer, it still represents a large reduction of the number of parameters needed when the CSNR is high. In the calculations of the MSE from (3.30), M is chosen to an odd value so that $d_{M-1} \geq 12$ for both the uniform and the quadratic quantizer.

The results of the systematic quantizers are shown in Figure 3.19. The uniform quantizer gives a small deterioration compared to the original one.



(a) Laplacian source, quantizer designed for Gaussian source

(b) Gaussian source, quantizer designed for Laplacian source

Figure 3.20 Difference between the performance of a coder optimized for the actual pdf and the performance of coders optimized for a different shape of the pdf than the actual one. The solid lines show the original quantizer where all the levels are optimized, the dash-dotted lines show the uniform quantizer, and the dashed line shows the quadratic quantizer (Laplacian design pdf only). Simplified receivers are used in all cases. Note the different scale on the two plots.

With the quadratic quantizer, the loss is negligible. At CSNRs above 27.5 dB, the quadratic quantizer performs better than the original, which is caused by the limited number of quantizer levels in the latter.

3.5.2.2 Source Mismatch

A stationary memoryless source producing random symbols is characterized by the pdf of the symbols. Determining that function from the observed source symbols can be a difficult task. However, some parameters like mean and variance are easy to estimate. A pre-determined model can then be used for the shape of the pdf. This model can sometimes be uncertain, and a desired feature is then that the performance of a coder has a low sensitivity to the shape of the pdf. The performance when using a coder designed for a Gaussian source applied to a Laplacian source of the same mean and variance as the design pdf, or vice versa, is shown in Figure 3.20. The graphs show that the sensitivity to pdf mismatch is low in all cases, except when a quantizer where all the levels are optimized for a Gaussian source is applied to a Laplacian source. The reason why this gives a large deterioration is that there are no quantizer levels of large enough amplitude to cover the tails of the Laplacian distribution.

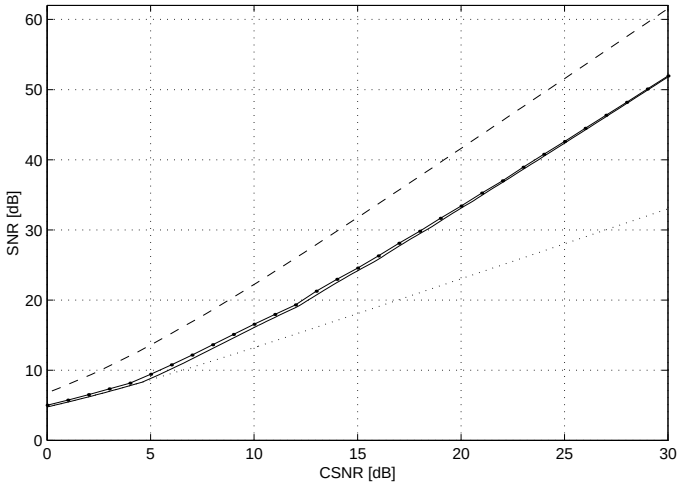
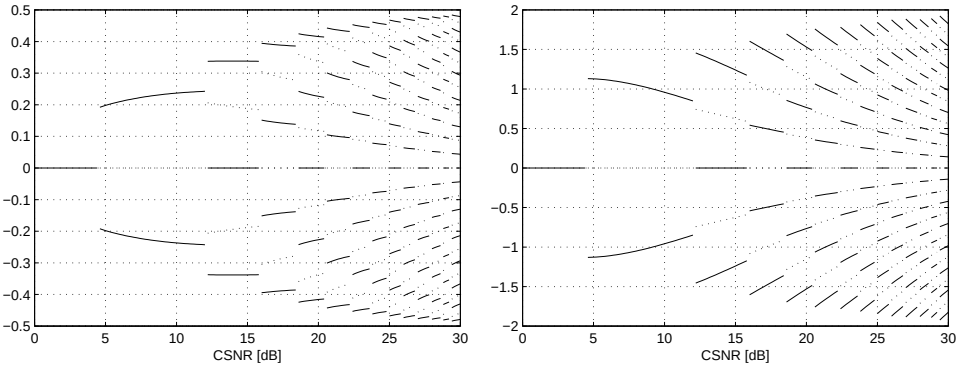


Figure 3.21 Performance of the HSQLC with simplified (solid) and optimal (solid with dots) receiver, compared to the OPTA (dashed) and BPAM (dotted), for a uniform memoryless source. All the systems have rate $R = 2$.

With a uniform quantizer, any amplitudes can be covered. In the performance calculations for systematic quantizers, a value of M suitable for the actual pdf has been used. This is acceptable because the limitation of M is only needed to perform numeric calculations of the MSE with Equation (3.30). In order to perform quantization in practice, M needs not be limited except to what the dynamic range of the processing unit demands. This shows a clear advantage of using systematic quantizers: Larger tails of the actual pdf than the design pdf are handled well.

3.5.3 HSQLC for Uniform Source

Figure 3.21 shows results for a uniformly distributed source. The optimization is performed in a slightly different manner this time. Due to the abrupt change in probability density that occurs in a uniform distribution, the optimization is unable to move quantization levels away from the mean in order to remove them, as happened for the Gaussian and Laplacian distributions. Thus, at each CSNR point, the same number of quantization levels as in the previous (higher CSNR) point is considered, using the optimal quantizer of the previous point as the initial value, along with a number of quantizer levels reduced by one compared to the previous point. If the reduction gives an improvement, another reduction of the number of levels is tried, until the reduction increases



(a) Source representation values a_i (solid) and decision levels d_i (dotted)

(b) Channel representation values c_i (solid) and decision levels e_i (dotted)

Figure 3.22 Optimized representation values and decision limits for a $\left[-\frac{1}{2}, \frac{1}{2}\right]$ uniform source on a unit power constrained AWGN channel as a function of CSNR

the distortion. A uniform quantizer is used as the initial value for the reduced quantizers. The best performing of all the quantizers that have been tried is chosen. For a CSNR of 30 dB, up to 15 levels are considered, but 14 levels give the best performance at 30 dB.

Figure 3.22 shows the quantization levels for the uniform source. Note that the quantizer this time alternates between a mid-tread and a mid-riser quantizer as a function of CSNR. At high CSNR values, the innermost levels seem quite uniformly spaced, while the outer are tighter. For CSNRs below 4.5 dB, the number of levels is one, meaning that the system with simplified receivers becomes exactly equal to the BPAM, something which is also seen in Figure 3.21.

3.5.3.1 Uniform Quantizer

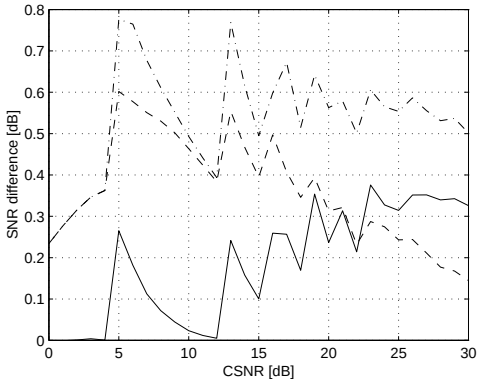
As for the Gaussian and Laplacian sources, a uniform quantizer will be tried for the uniform source. In Section B.3.4, it is shown that if

$$a_i = \left(i - \tfrac{M-1}{2}\right)\tfrac{\Delta}{M},\ i \in \{0, \ldots, M-1\} \tag{3.40}$$

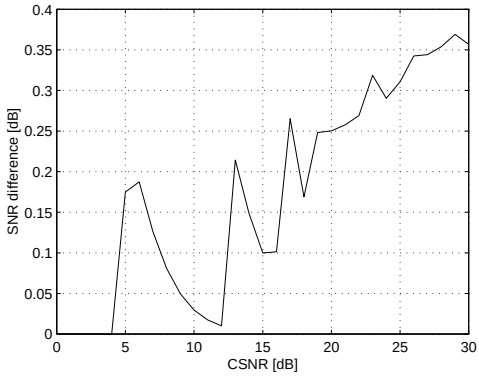
and

$$d_i = \left(i - \tfrac{M}{2}\right)\tfrac{\Delta}{M},\ i \in \{1, \ldots, M-1\}, \tag{3.41}$$

where Δ is the width of the uniform distribution (cf. Equation (B.34)), the expression for the MSE using optimal receivers can be simplified. It is then



(a) Difference between original quantizer with optimal receiver and uniform quantizer with optimal receiver (solid), original quantizer with simplified receiver (dashed), and uniform quantizer with simplified receiver (dash-dotted)



(b) Difference between original quantizer with simplified receiver and uniform quantizer with simplified receiver

Figure 3.23 Performance of HSQLC system with uniform quantizer for a uniform memoryless source

given by a sum of one and two-dimensional integrals. These integrals can be found numerically, and if c_i is chosen to be proportional to a_i,

$$c_i = K_a a_i, \ i \in \{0, \ldots, M-1\}, \tag{3.42}$$

only the constants K_a and K_b need to be found for a given M. One of these constants can be found from the other given a channel power constraint, so the problem has only one variable, and is reduced to a problem of distributing channel power between the symbols. In this case, numerical optimization of the MSE given by numerically solved integrals is feasible, so the optimal receivers can be used in the optimization process. The MSE can be calculated for different M, and the best one will be chosen.

The performance of a uniform quantizer for a uniform source is demonstrated in Figure 3.23. As seen, the loss is relatively small, and more is gained by using optimal receivers than by optimizing the different quantizer levels separately. The difference in performance fluctuates because the quantizer step size with the uniform quantizer only takes values that divide Δ, while the original quantizer can have any spacing between the levels.

The uniform quantizer for the uniform source gives the opportunity to use the optimal receivers in the optimization process, instead of optimizing using

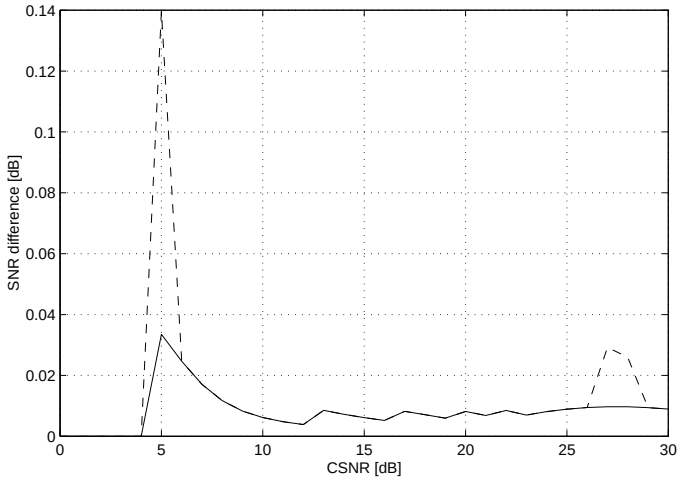
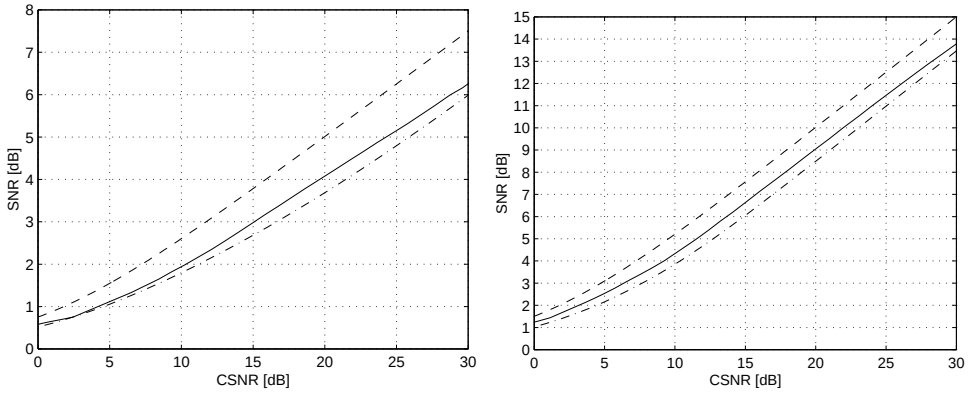


Figure 3.24 Performance difference between the HSQLC system with uniform quantizer optimized assuming optimal receivers and a system with uniform quantizer optimized assuming simplified receivers, but actually using optimal receivers. The solid line shows the case when the number of levels M is optimized using the performance with optimal receivers, while for the dashed line, the performance using simplified receivers has been used for finding the best M.

simplified receivers and then applying the optimal receivers, as must be done in other cases. Figure 3.24 shows how much is gained by assuming optimal receivers in the optimization, which is not much. Note that in some cases, the optimal value of M is different depending on the type of receiver that is assumed when the MSE is calculated, and then, more is gained from optimizing M using the correct receivers than from optimizing the power distribution with the optimal receivers. This indicates that the gain from assuming optimal receivers in the optimization process could be somewhat higher when also the quantizer levels are optimized numerically. Still, no large improvement can be expected, as the gain from introducing optimal receivers in the first place is small.

3.5.4 Comparison to Entropy Coding

Traditional coding systems for compression are often based on entropy coding, since quantization in combination with entropy coding is more efficient than pdf optimized quantization (Ramstad, Aase and Husøy, 1995). Entropy coding lacks the robustness of the direct source–channel mappings, though, since an



(a) $K = 1,\ L = 4$; Total rate $R = \frac{1}{4}$ (b) $K = 1,\ L = 2$; Total rate $R = \frac{1}{2}$

Figure 3.25 Performance of a PCCOVQ system with codebook size $M = 256$ (solid) compared to a uniform quantizer followed by an ideal entropy coder and an ideal channel coder (dash-dotted) and to the OPTA (dashed). The source is memoryless Gaussian.

entropy coder breaks down due to synchronization problems whenever an error occurs.

An entropy coding based system is a pure source coder. In this section, a channel coder that performs as the channel capacity will be assumed, thus, the results for entropy coding are lower bounds on the distortion. A uniform scalar quantizer with centroid representation values is used, since uniform quantization is close to the optimal scalar quantizer in combination with entropy coding (Farvardin and Modestino, 1984). The entropy coder is assumed to give a rate equal to the actual entropy. The entropy and distortion of the quantizer are calculated by integration over the source pdf as was done in Chapter 2, so no high-rate approximation or simulations are used.

In Figure 3.25, the PCCOVQ system of (Fuldseth, 1997) with rates $\frac{1}{4}$ and $\frac{1}{2}$ is compared to the entropy coding based system having the same total rate. As seen, PCCOVQ is better than the entropy coding based system, even though an optimal channel coder is assumed. Of course, the limiting factor of this coder is the scalar quantizer. The PCCOVQ system combines two or four source symbols for the encoding, and if the entropy coder based system could do the same, that is, use a two or four-dimensional vector quantizer, its performance would improve. When a high-rate assumption applies, the gain by using a two-dimensional vector quantizer is approximately 0.17 dB, and for a four-dimensional quantizer, it is approximately 0.39 dB (Lookabaugh and Gray,

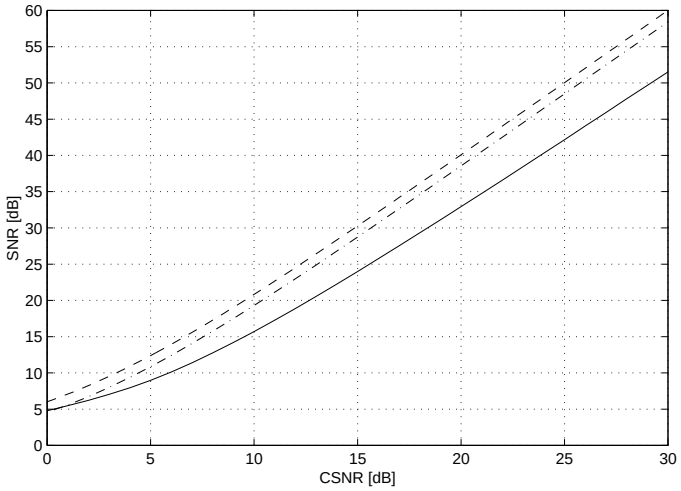
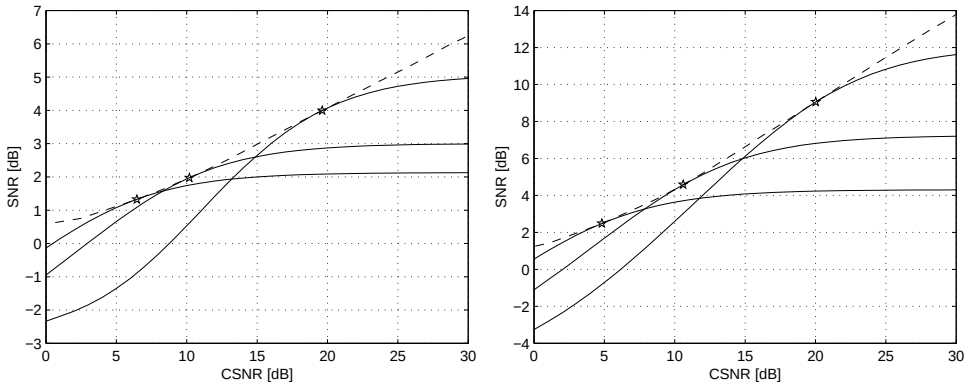


Figure 3.26 Performance of an HSQLC system with optimal receivers (solid) compared to a uniform quantizer followed by an ideal entropy coder and an ideal channel coder (dash-dotted) and to the OPTA (dashed). The source is memoryless Gaussian.

1989). Thus, on good channels where the high-rate assumption holds, the rate $\frac{1}{2}$ entropy coding system will still be inferior to PCCOVQ, while the rate $\frac{1}{4}$ system gets a performance very close to PCCOVQ, if vector quantization of the same dimension as in PCCOVQ is performed.

Figure 3.26 compares the HSQLC system to the entropy coding based system that also has a total rate of two. In this case, the latter system is better, so bandwidth expansion with HSQLC is not as efficient as bandwidth reduction with PCCOVQ. Still, getting a channel coder performance equal to the capacity is impossible, so in practice, the advantage of the entropy coder based system would be reduced or vanish entirely. For the channel qualities shown in Figure 3.26, the HSQLC demands a CSNR up to about 3.5 dB better than the entropy coding based system. A channel code with a performance 3.5 dB below the capacity for high CSNRs is a good code.

The direct source–channel mappings are much more robust than the entropy coder system. Since an optimal channel coder is assumed, no meaningful decoding can be expected below the design CSNR, and the performance will not improve above the design CSNR since the quantization is the only source of distortion. The robustness of the HSQLC coder was demonstrated in Figure 3.15, and the performance of PCCOVQ coders at channel mismatch is shown in Figure 3.27. As seen, the performance decreases relatively slowly



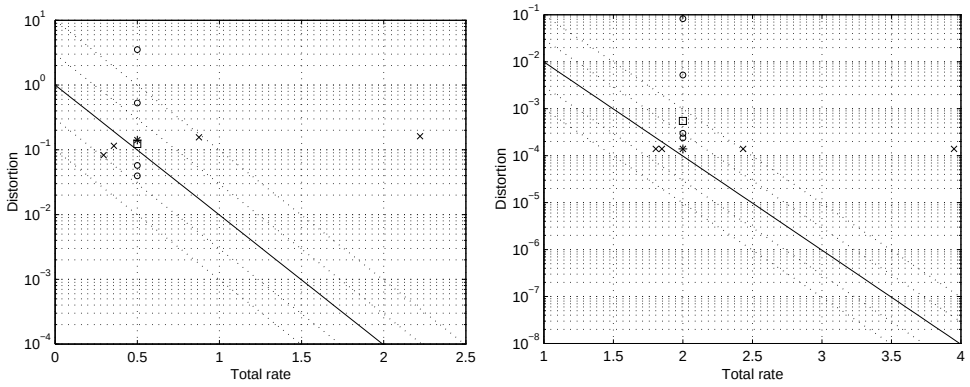
(a) $K = 1,\ L = 4,\ M = 256$ (b) $K = 1,\ L = 2,\ M = 256$

Figure 3.27 The performance of PCCOVQ mappings for a memoryless Gaussian source at CSNR different from the design CSNR (solid). The design CSNR is shown with a star, and the dashed line shows the performance of PCCOVQ evaluated at the design CSNR.

below the design CSNR, and increases somewhat above it before saturation is reached.

So far, only robustness to varying channel quality and pdf shape has been discussed. Another issue is robustness against source variance mismatch. This will be handled differently by a direct source–channel mapping, which has a fixed rate, and an entropy coding based system, where the rate varies according to the source variance. This kind of robustness is investigated in Figure 3.28, where the rate-distortion performance of direct source–channel mappings and the entropy coding based system described above are compared. For the direct source–channel mappings, the channel symbols are assumed to be scaled so that the power constraint is satisfied exactly, and the scaling factor is assumed to be known by the receiver. With the HSQLC, a uniform quantizer is used so that the number of quantizer levels is not limited. In the entropy coder based system, the ideal entropy coder is designed for a Gaussian unit variance source. The results for entropy coding and HSQLC are found by integration over the source pdf as before, while the results for PCCOVQ are found by Monte Carlo simulations over 10^6 channel samples.

From Figure 3.28, just as from Figure 3.26, it can be seen that PCCOVQ performs better than the entropy coding system on the design source, while HSQLC performs worse than it. If the variance of the actual source is higher than expected, the entropy coder system gives a performance far away from the



(a) PCCOVQ, $K = 1$, $L = 2$, $M = 256$

(b) HSQLC, uniform quantizer, simplified receivers

Figure 3.28 Rate-distortion performance of the direct source–channel mappings compared to an entropy coder based system for source variance mismatch. The solid line shows the OPTA at a CSNR of 20 dB for a unit variance Gaussian source, and the dotted lines show the OPTA for Gaussian sources of variance (from below) 0.1, 0.3, 3.0 and 10.0. The square shows the performance of the direct source–channel mapping designed for and applied to a Gaussian unit variance source, and the circles show the performance of the same mapping applied to sources of the other variances mentioned above. The asterisk shows the performance of an entropy coding system designed for a Gaussian unit variance source, and the crosses shows the performance of the same coder applied to Gaussian sources of the variances stated above. A CSNR of 20 dB is assumed in all cases.

OPTA, worse than the direct source–channel mappings. With a source variance below expected, the entropy coder is slightly better. Still, the direct source channel mappings must be said to be more robust against changing source variance, which comes in addition to their increased channel robustness. There is an inherent difference in how the systems react to source variance changes. With the direct source channel mappings, the rate is constant, while the distortion varies. If the channel samples are not scaled, the channel power will vary, while the distortion changes less. With the entropy coder system, on the other hand, the rate varies, and the distortion changes less. If the underlying bit rate is high enough, such as for the rate 2 system of Figure 3.28(b), the distortion is virtually constant. The channel power will always be constant.

3.6 Discussion

3.6.1 HSQLC Performance

For low CSNR values, the performance of the HSQLC is comparable to other direct source channel mapping methods, but for good channels, it is clearly better. The reason is the use of a discrete channel symbol where the number of levels is allowed to increase. The difference between the OPTA and the HSQLC performance increases very slowly with the CSNR. This stands in a clear contrast to the BPAM system, since the SNR as a function of the CSNR (both measured in dB) for the BPAM asymptotically has a slope of one, while the OPTA's asymptotic slope is two.

The performance is better than the HDA-FEAD, which is another hybrid digital–analog system. At moderate design CSNR values (10 dB), the quantizer part of the HDA-FEAD has too few levels, while at low CSNR values (0 and 5 dB) it seems that the HDA-FEAD spends too much power on the quantization error symbol, as it performs worse than BPAM. It might be a problem with BPAM (or HSQLC at low CSNR) that every second symbol carries all the power, but this can be taken care of by an orthonormal linear transformation which would not alter the performance. It would however change the spectrum of the channel signal, and if that is a problem, pseudo-random scrambling of the sign of the output symbols (with correction in the receiver) can be used to recover a white channel signal.

The robustness of the HSQLC is limited, but no worse than the HDA-FEAD. Knowledge of the actual channel quality by the decoder is essential if the encoder is designed for a poor channel. For a high design CSNR, knowledge of the actual quality gives some improvement when the actual channel quality is lower, especially if optimal receivers are used. What happens then is that the quantized symbol is detected using soft decision. Since the channel noise will be larger than the quantization noise in this case, the performance will approach direct PAM transmission with the channel power assigned to the quantized symbol, while the quantization error symbol is not very useful, and will contribute little to the decoded symbol.

Robustness was not addressed in the optimization process, so the robustness that is obtained is a result of the fundamental nature of the HSQLC. With a low dimension of the code, the ability to adapt exactly to the channel quality is limited, and both the use of a naturally ordered quantized symbol and a linearly coded symbol give a certain robustness. It might be possible to re-optimize the coder with robustness as an additional feature, for instance by imposing constraints on the performance at some channel qualities while minimizing the distortion at another channel quality.

Optimal receivers give a very limited performance gain except when the encoder assumes a good channel and the actual channel is poor. This situation does not give a good performance in any case, so often, the encoder will be designed for a channel not much better than the worst channel for which decoding should be possible at all. The optimal receivers require higher complexity than the simplified ones since they are based on non-linear functions requiring a large number of multiplications to be implemented. Still, they are memoryless, and if the HSQLC contributes only negligibly to the total computational complexity of the system, the use of optimal receivers can be defended. Otherwise, when channel qualities far below the design CSNR need not be supported, simplified receivers should be used.

The use of uniform quantizers seems to be a good choice. They work pretty well for both Gaussian, Laplacian and uniform sources, and the optimization becomes much simpler when they are used. Furthermore, the use of a uniform quantizer gives robustness concerning the shape of the source pdf. Since the number of quantizer levels can be unlimited, distributions with long tails are handled well even if the design pdf has shorter tails, and also outliers (a few symbols from another stochastic process) with larger variance can be supported.

3.6.2 Channel Model

In Section 3.1, the channel model was given without any argumentation. An AWGN channel is a very common model, but some issues should be addressed. The time discrete nature of the channel is merely a model for signaling on a time continuous channel. If the physical channel is to be utilized efficiently, signaling must take place at the Nyquist rate (Nyquist, 1928; Blahut, 1987). For a passband system, this means that the discrete system will be derived from either a single sideband PAM system, or a QAM system where the in-phase and quadrature components have been interleaved. Assuming additive white Gaussian noise as the only distortion, these systems are equivalent. Since the model operates with real-valued symbols, it is denoted a PAM model, but it can also be derived from a QAM system, at least in theory.

Another issue is the synchronization. This is always a problem in communication systems, but when the signaling alphabet is limited, clear transitions between symbols will exist, allowing resynchronization in the receiver. With continuous amplitude channel symbols, which are often produced by direct source–channel mappings, finding transitions between symbols can be a difficult task. The problem can be solved by putting pilot symbols in the stream to allow resynchronization. If QAM is used, synchronization must also be assured for the carrier in order to determine the phase of the received symbols,

something which can also be obtained with pilots.

A third issue is fading channels. A slowly fading channel can be turned into an approximate AWGN channel by equalization, but the resulting AWGN channel will have varying amplification and CSNR. The amplification must be known by the receiver when information is transmitted in the amplitude of the symbols, and again pilot symbols may be the solution. The robustness of the mappings would be a useful property for a channel of varying CSNR. The performance is improved if the receiver can estimate the actual CSNR, and even more if this information can also be sent to the transmitter. Finally, to perform the equalization, pilot symbols are also needed if continuous amplitude is used so decision feedback is impossible.

The amount of pilot symbols needed will definitely depend on the channel and the kind of mapping used. In hybrid digital–analog systems, for instance, the digital part may be used for synchronization or decision feedback equalization, reducing the need for pilot symbols.

3.6.3 Future Work

The HSQLC can provide only one rate, namely $R = 2$. It is desirable to develop coders for other rates as well. For rates below one, the PCCOVQ system gives very good performance, and for a rate of one, an optimal system exists for memoryless Gaussian sources. For rates in the range $(1, 2)$, the quantized part of several source symbols could be combined into one symbol, using a PCCOVQ-like method. The codebook vectors would then have a larger minimum distance than shown in Figure 3.4, since a low symbol error rate is desirable when used in conjunction with a quantization error symbol.

For rates larger than two, finding good mappings of the HSQLC type would be more difficult. In order to create coders of integer rates, a natural idea would be to split the symbols using several levels of quantizers and finishing with a linear encoder. This would resemble successive refinement (Equitz and Cover, 1991), but not based on bits, so the number of quantizer levels in each split can be different from two and the least significant bit is replaced by a continuous symbol. The problem with such a method is that it creates some symbols of a very large importance. These need to be protected using few levels and enough power. Such symbols will reduce the robustness of the coder, since the remaining symbols contain very little useful information if the most significant symbol is erroneously decoded. It is possible to include traditional error protection for the most significant symbols, and this might improve the performance at the design CSNR, but the robustness would be further reduced. It would also be possible to use a vector quantizer for the analog part, as in (Skoglund et al., 2001*a*), maybe with other modulation than

BPSK, but such a vector quantizer would probably suffer some of the same problems as with the use of scalar quantizers of several splits, because it is unavoidable that two points of short distance in the channel space will have a large distance in the source space unless large portions of the channel space is unused (as in BPAM). Still, by allowing the modulation set to be optimized, a relatively good performance might be possible. The experience with bandwidth expanding PCCOVQ suggests that it might be very computationally complex to perform the optimizations.

Considering sources with memory might be necessary for some applications. In order to handle that, memory must be introduced in the coder. This can be done with a vector quantizer as Skoglund et al. (2001*a*) did, perhaps with larger modulation sets than BPSK if high CSNR values are to be considered. Also, the power distribution between the symbols should be optimized, and maybe the signal set, too.

Considering other channel models, such as fading channels, can be an important issue for many applications, as can the synchronization problems. Experiments should be carried out to find the amount of pilot symbols needed for a reliable synchronization. Furthermore, methods for handling fading channels by estimating the fading state can be investigated.

Chapter 4

Image Coder Using Joint Source–Channel Coding

As an application of direct source–channel mappings, an image coder is considered in this chapter. Image compression is a topic that has caught a lot of attention, and most of the research is performed for error-free channels. This gives methods that can be used for applications such as storage and transmission, provided that sufficient error protection is used. As mentioned earlier, joint source–channel coding can give advantages such as robustness against unknown channel qualities and lower computational complexity to obtain the same distortion.

Most recent image coders are subband coders (including the special case of wavelet based systems), such as the new standard, JPEG 2000, of which part 1, the basic coder, has been issued as an international standard (ISO/IEC, 2001). The coder proposed here is a subband coder based on a coder proposed by Lervik (1996).

The chapter is organized as follows. In Section 4.1, some previous work on joint source–channel coding of images is presented. Section 4.2 gives a description of the proposed coder explaining the design of each part of the coder. In Section 4.3, coding results with the proposed coder are presented and compared to other coders, and analysis is performed to investigate different aspects of the coder. Section 4.4 provides a discussion of the results. An earlier version of the coder was presented in (Coward and Ramstad, 2000*c*).

4.1 Previous Work

One of the first works in joint source–channel coding of images was done by Modestino and Daut (1979), who traded off the source and channel coder

rates and used unequal error protection for the bits of DPCM coding. This was later extended to transform coding (Modestino, Daut and Vickers, 1981). Later works on unequal error protection in image coding have been done by Fazel and Lhullier (1990) and Tanabe and Farvardin (1992), who considered variable length source coding, and Chande and Farvardin (2000), who combined unequal error protection with progressive transmission. Ruf and Modestino (1999) used operational rate-distortion techniques to trade off source and channel coding rates in an unequal error protection scheme and found information-theoretic performance bounds, and Cai and Chen (2000) also included an all-pass filter to improve the performance. Kozintsev and Ramchandran (1998) and Zheng and Liu (1999) combined unequal error protection with multiresolution modulation, where the modulation set is non-uniform, allowing for different error probability for different bits. The combination of source coding and multiresolution modulation in these papers gives a direct source–channel mapping.

Another approach to joint source–channel image coding is the use of robust quantization, as was done by Vaishampayan and Farvardin (1990), who used bit allocation and quantizers optimized for a channel with errors. Skoglund (1995) used a Hadamard based framework with soft decoding. Chen and Fischer (1998) used robust quantization together with a frequency scrambling to reduce the perceptual impact of the noise. The coder proposed by Lervik and Ramstad (1996) has a quantizer that is jointly designed with a modulation set, as the PCCOVQ system described in Section 3.3.2, but not optimized, and a direct PAM mapping. This system was improved in (Lervik, 1996), where some PCCOVQ-like properties where included. The work in this chapter is based on the framework of that system.

4.2 Coder Structure

The coder is supposed to operate on an AWGN channel as described in Section 3.1, with a fixed noise power and a constraint on the average signal power. The number of channel samples generated for one image is also constrained, and this number is specified by a rate measured as the number of channel samples generated per pixel (picture element).

The structure of the proposed coder is shown in Figure 4.1. The image is first filtered by an analysis filter bank. The subbands are then divided into blocks of subband samples. On these samples, classification is performed, assigning each block to one of $J+1$ classes. Samples in class 0 are not transmitted, while the other classes are transmitted using one of J direct source–channel mappings. The classification table is necessary to perform decoding of the

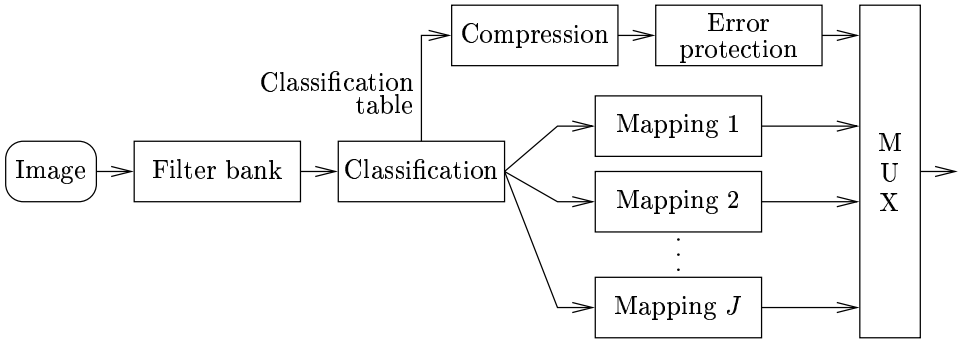


Figure 4.1 Proposed image encoder

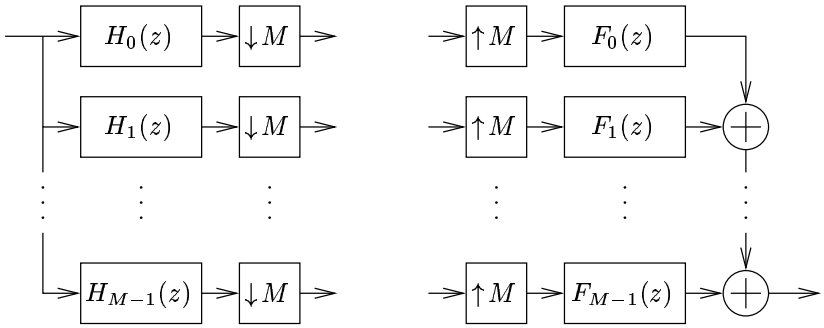


Figure 4.2 M-channel uniform maximally decimated filter bank. Analysis filter bank to the left, synthesis filter bank to the right.

image, and it must be received free of error. Thus, it is compressed and error protected using traditional techniques before it is transmitted.

The decoder receives the classification table and uses that to select the appropriate decoding mapping for each block. The blocks are then placed correctly into a reconstructed subband image, where zero values have been inserted for the blocks of class 0. Then, the synthesis filter bank is applied to generate an approximation of the transmitted image.

Below, the different blocks of the coder are explained in more detail.

4.2.1 Filter Bank

Figure 4.2 shows an *analysis* and a *synthesis filter bank* (Vaidyanathan, 1993). The analysis filter bank is used for decorrelation of the signal, by taking out

different frequency components and placing them in *subbands*. The analysis filter $H_0(z)$ is typically a lowpass filter, $H_{M-1}(z)$ is typically a highpass filter, and the other filters are bandpass filters with different passbands. The synthesis filter bank is used for reconstruction of the signal, and the synthesis filters $F_0(z), \ldots, F_{M-1}(z)$ normally have approximately the same passband frequency ranges as their analysis counterparts. If the output of the synthesis filter bank is equal to a delayed version of the input of the analysis filter bank when the outputs of the analysis filter bank are transferred undistorted to the synthesis filter bank, the filter bank has the *perfect reconstruction* (PR) property.

In the filter bank in Figure 4.2, all the decimators use the same decimation factor. Such a filter bank is called a *uniform* filter bank. Furthermore, the decimation factor is equal to the number of subbands, which means that the total number of symbols per time unit at the output of the analysis filter bank is the same as the number of symbols per time unit at the input. This decimation factor is the highest for which PR can be obtained, and thus, the filter bank is said to be *maximally decimated*.

An image is a two-dimensional signal, and decorrelation is needed in both dimensions. This calls for a two-dimensional filter bank. The most common way to implement this is by applying a one-dimensional filter bank to each dimension. The resulting two-dimensional filter bank is then called *separable*.

Lervik (1996) used the filter bank denoted "32I" proposed by Aase (1993). This is a separable uniform filter bank with eight bands in each direction and 32 filter taps in all the filters. The filter bank has the property of almost perfect reconstruction.

A disadvantage of a uniform filter bank is that there is still much correlation left in the lowpass-lowpass band, i.e., the band that has been lowpass filtered in both dimensions. One way of handling this is to apply another filter bank to the lowpass-lowpass band, as in the system described below.

In this work, the filter bank denoted "System K" proposed by Balasingham (1998) is used. This system consists of an eight-band uniform filter bank and three different two-band filter banks that are applied only to the lowpass band of the previous stage. This results in an overall non-uniform filter bank with a resulting lowpass-lowpass band that is decimated 64 times in each dimension. This kind of splitting is called *dyadic splitting*, and the total filter bank is called a *tree-structured* filter bank. The uniform filter bank used in the first stage has 32 taps and almost perfect reconstruction, just as "32I", and is designed using the same algorithm, but with different optimization parameters specifying a performance closer to PR, giving other filter taps. The two-channel filter banks are different, but they all have perfect reconstruction and have nine filter taps

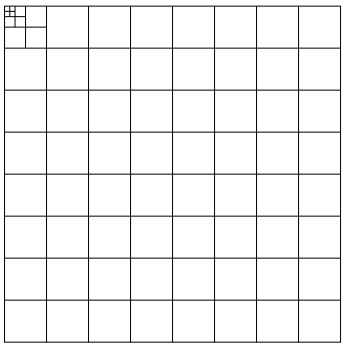


Figure 4.3 Organization of the subbands

in the analysis lowpass and the synthesis highpass filter and seven filter taps in the other two filters. For small images, the number of dyadic splits is reduced in order to avoid creating too small bands. In the coding examples, the image size is 512×512 pixels, which is enough for three splits to be performed.

After filtering with a non-uniform filter bank, several subbands with different decimation factors are obtained. In order to keep a rectangular structure, as for the original image, the subbands are organized as shown in Figure 4.3, where the low frequency subbands are placed towards the left and the top. The lowest frequency subbands are decimated several times and are thus smaller. The figure also gives a good illustration of the structure of the filtering, where the low frequencies are filtered the most.

Natural images do not contain discrete frequency components at other frequencies than zero, so the only subband that will have an expected value different from zero is the lowpass-lowpass band. The mean of that band, however, can give a large contribution to the power of the filtered image. Thus, the mean of the lowpass-lowpass band is subtracted before further coding and is transmitted separately.

In order to analyze the performance of the coding of the subband samples, the connection between noise before and after the synthesis filter bank must be known. If the distortion of the subband signals can be modeled as additive stationary white noise, the distortion on the output of the synthesis filter bank is additive noise of the same variance, provided that the impulse response energy of all the synthesis filters is one (Vaidyanathan, 1993). If the passbands are relatively flat, the noise will be approximately white. Note that for multistage filter banks, it is the overall impulse response through all the filters that must have unit energy. This property is satisfied for the filter bank used here.

4.2.2 Rate and Power Allocation

In (Lervik, 1996; Lervik and Fischer, 1997), an algorithm for allocating rate and power is presented. Here, an improvement of the algorithm is proposed.

The coder operates under rate and power constraints. From the classification point of view, the source can be said to consist of a number of subsources, each having a different source variance. Here, the subsources are blocks of 8×8 subband samples (which are normally smaller than the subbands). Each subsource is assigned to one out of $J + 1$ classes, and to each class, there is a corresponding rate and distortion.

4.2.2.1 Introduction to Rate Allocation

Rate allocation is performed to ensure that the subsources are encoded in a manner that gives the lowest overall distortion while maintaining the overall rate constraint. In practice, that means allocating more rate to the subsources having the largest variances. Rate allocation is closely related to bit allocation, which is a common operation in binary output subband and transform coders (Gersho and Gray, 1992). The rate is here measured in channel samples per pixel, and the distortion is partly due to the channel noise, but otherwise the situation is the same.

The problem of rate allocation can be viewed as an operational rate-distortion problem (Ortega and Ramchandran, 1998). This means that points of distortion as a function of rate that are obtainable with the considered coder framework will be given. These points can be plotted in a diagram, and only the points lying on the lower convex hull of the set of points are used as candidates, because otherwise, a point of the same rate giving a lower distortion can be found by a linear combination of two points on the convex hull. For the case of only one subsource, this is illustrated in Figure 4.4. This graph shows the actual rate-distortion points that are obtained with the mappings used in this coder, which will be described in Section 4.2.3, for a CSNR of 20 dB. In this case, all the points lie on the convex hull. If a mapping is introduced that has an operational point above the dashed line, this mapping should never be used.

In practice, the number of subsources is large. As long as the overall rate constraint is fulfilled, any combination of assignments of subsources to one of the classes can be considered, giving an enormous number of possible rate-distortion points, most of which are not on the convex hull. An algorithm for performing rate allocation while remaining on the convex hull was proposed by Westerink, Biemond and Boekee (1988). An adaptation of this algorithm to the present problem is given in Section 4.2.2.5.

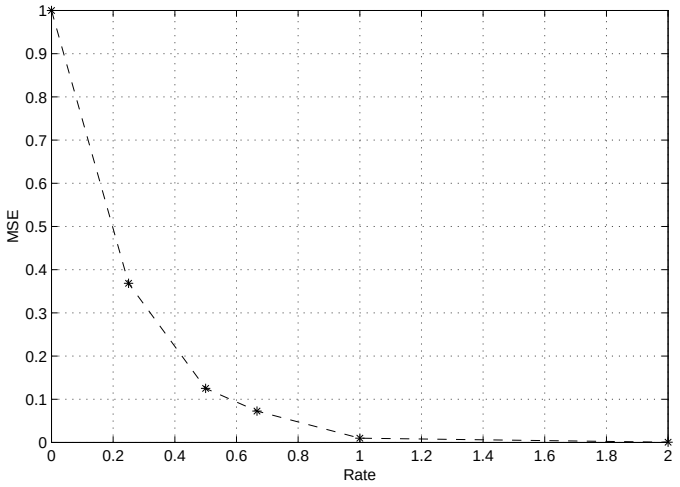


Figure 4.4 Operational rate-distortion performance for one unit variance subsource. The stars show the operational points, and the dashed line shows the convex hull.

4.2.2.2 Introduction to Power Allocation

To the different classes, there is a corresponding rate, which is given by the nature of the mapping used for this class. There is a limitation on the total power of the encoded image signal, but there is no demand that all the classes have the same channel power. Thus, power can be allocated between the classes in a manner so that the total distortion is minimized.

If mappings with every possible rate and power are available and the mappings perform according to the OPTA of a Gaussian source, it has been shown by Lervik (1996) that the channel power of each class should be the same when the distortion is minimized. However, this result does not apply to the proposed system, since the number of possible rates is limited and not all of the mappings perform according to the OPTA. Still, if the mappings provide a set of rates which are not too far from the optimal rates obtained assuming any rate is available, and if their performance is good, the optimal channel power values of each class should not be too different.

4.2.2.3 Finding the Distortions

The rate is fixed for each class, but the distortion is a function of the distribution of source samples assigned to the class and the channel power that the class has been allocated. Finding this fidelity function can be a highly complex

task for an arbitrary source pdf. If the assignment of subsources to classes is assumed to be free, clearly the distribution of source samples will vary. However, a simplification of the problem can be obtained by assuming that the distortion is independent of the shape of the source sample distribution, and depends only on the variance of these samples. As explained in Section 4.2.1, the expected value is zero.

Under the assumption that the distortion is independent of the source pdf shape, the distortion of each class, which is taken as the mean square error between the source samples and the decoded values for class j, can be written as $\tilde{D}_j(\bar{\sigma}_j^2, P_j)$. Here, $\bar{\sigma_j}^2$ is the variance of the entire set of source symbols mapped to class j, and P_j is the channel power for class j. In order to have mappings working for different source variances, all source samples in class j will be divided by $\bar{\sigma}_j$ before the mapping is applied, and multiplied by the same value in the decoder. This means that the mapping will always see a source of unit variance, and the distortion is

$$\tilde{D}_j(\bar{\sigma}_j^2, P_j) = \bar{\sigma}_j^2 D_j(P_j), \tag{4.1}$$

where $D_j(P_j) = \tilde{D}_j(1, P_j)$.

Assume that there are N subsources identified by a unique index $i \in \{0, \ldots, N-1\}$. Define v_i as the class to which subsource i has been assigned and $\mathcal{I}_j$ as the set of subsources that have been assigned to class j. This means that

$$\mathcal{I}_j = \{i \mid v_i = j\}. \tag{4.2}$$

If the subsources have samples of variances σ_i^2, $i \in \{0, \ldots, N-1\}$, then

$$\bar{\sigma}_j^2 = \frac{1}{|\mathcal{I}_j|} \sum_{i \in \mathcal{I}_j} \sigma_i^2, \tag{4.3}$$

where $|\mathcal{I}_j|$ denotes the cardinality of $\mathcal{I}_j$.

The overall distortion (disregarding any distortion due to a non-PR filter bank) of the system given a classification $\{v_i\}_{i \in \{0,\ldots,N-1\}}$ and a power assignment $\{P_j\}_{j \in \{0,\ldots,J\}}$ is then

$$D_{\text{tot}} = \frac{1}{N} \sum_{j=0}^{J} |\mathcal{I}_j| \bar{\sigma}_j^2 D_j(P_j). \tag{4.4}$$

If the rate of class j measured in channel samples per source sample is R_j, the total permitted rate measured in channel samples per pixel is R_{tot}, and

the average channel power constraint is P_{tot}, then the optimization problem can be formulated as follows: Select the classification $\{\mathcal{I}_i\}_{i\in\{0,\dots,N-1\}}$ and the power assignment $\{P_j\}_{j\in\{0,\dots,J\}}$ that minimize D_{tot} under the constraints

$$\frac{1}{N}\sum_{j=0}^{J}|\mathcal{I}_j|R_j \le R_{\text{tot}} \tag{4.5}$$

and

$$\frac{1}{NR_{\text{tot}}}\sum_{j=0}^{J}|\mathcal{I}_j|R_jP_j \le P_{\text{tot}}. \tag{4.6}$$

Finding a solution of this problem is a highly complex task. However, there exist methods for solving the problem of rate allocation for a fixed power assignment, and for solving the problem of power allocation for a fixed classification. These problems will be discussed below.

4.2.2.4 Power Allocation for a Fixed Classification

For a fixed classification, all the v_i and $\mathcal{I}_j$ are given, so the problem is reduced to minimizing (4.4) with respect to P_j, $j \in \{0,\dots,J\}$ under the constraint given by (4.6). Using the Lagrange multiplier method, this problem can be stated as minimizing the unconstrained problem

$$\tilde{\mathcal{L}} = \frac{1}{N}\sum_{j=0}^{J}|\mathcal{I}_j|\bar{\sigma}_j^2 D_j(P_j) + \tilde{\lambda}\frac{1}{NR_{\text{tot}}}\sum_{j=0}^{J}|\mathcal{I}_j|R_jP_j, \tag{4.7}$$

where $\tilde{\lambda}$ is a Lagrange multiplier. Simplifying the problem by multiplying by N and setting $\tilde{\lambda} = R_{\text{tot}}\lambda$, the Lagrangian

$$\mathcal{L} = \sum_{j=0}^{J}|\mathcal{I}_j|\bar{\sigma}_j^2 D_j(P_j) + \lambda\sum_{j=0}^{J}|\mathcal{I}_j|R_jP_j \tag{4.8}$$

is obtained. The minimum value is found by differentiating the Lagrangian with respect to $P_j \ge 0$, $j \in \{0,\dots,J\}$, and setting the derivatives to zero:

$$\frac{\partial\mathcal{L}}{\partial P_j} = |\mathcal{I}_j|\bar{\sigma}_j^2 D_j'(P_j) + \lambda|\mathcal{I}_j|R_j = 0 \tag{4.9}$$

$$\Updownarrow$$

$$-D_j'(P_j) = \lambda\frac{R_j}{\bar{\sigma}_j^2}. \tag{4.10}$$

Here, D_j' is the derivative of D_j. From the formulation of the Lagrangian, it is clear that λ has to be positive, since otherwise, all the class channel powers will go towards infinity.

The function D_j is only defined on $[0, \infty)$. For any reasonable choice of mapping, the function will be non-increasing, so the derivative is non-positive. If the distortion function of a mapping is proportional to the OPTA of the source and rate considered (or of any other source and rate), it will also be convex, so the derivative is increasing, and there is at most one solution. Also, the distortion will approach zero as the power approaches infinity, and a constant greater than or equal to one as the power approaches zero. This means that the derivative also approaches zero as the power approaches infinity.

The non-negativity constraint on the class channel powers can be handled by Kuhn–Tucker conditions (Walsh, 1975). However, the problem is simple, so for each $j \in \{0, \ldots, J\}$, if (4.10) has a solution for $P_j \in [0, \infty)$, that solution is chosen, and if $-D_j'(P_j) < \lambda \frac{R_j}{\bar{\sigma}_j^2}$ for all $P_j \in [0, \infty)$, the optimal value is $P_j = 0$.

A problem that has not been discussed so far is that of finding $D_j'(P_j)$ and its inverse. For most practical mappings, the distortion has to be found experimentally. The method that is used is to produce a table of distortions as a function of the class channel power. Between the points, interpolation can be used, giving a quadratic spline estimate of the distortion. The estimated derivative then consists of straight line segments, and the inverse of the derivative can be found by a search. There is no guarantee that the convexity or even the non-increasing property of the distortion function is retained when the distortion of a practical mapping is used. This means that all possible solutions must be tried, so the one giving the smallest Lagrangian can be chosen. Due to the simplicity of the problem, this can be done separately for each class $j \in \{0, \ldots, J\}$.

Another problem is finding the Lagrange multiplier. Solving directly for it dramatically increases the complexity of the problem, since it creates an equation that depends on all the unknowns. However, it can be found iteratively, as will be explained in Section 4.2.2.6.

4.2.2.5 Rate Allocation for a Fixed Power Assignment

The rate allocation part of the algorithm from (Lervik, 1996; Lervik and Fischer, 1997) is based on a bit allocation algorithm proposed by Westerink et al. (1988). The problem considered there is the problem of bit allocation for N subsources that are to be quantized, where the distortion to be minimized is

given by

$$\hat{D}_{\text{tot}} = \sum_{i=0}^{N-1} \hat{D}_i(\hat{v}_i) \tag{4.11}$$

and the rate constraint is given by

$$\hat{R}_{\text{tot}} = \sum_{i=0}^{N-1} \hat{R}_i(\hat{v}_i), \tag{4.12}$$

where $\hat{v}_i$ is the quantizer that is used for subsource i, $\hat{D}_i(\hat{v}_i)$ is the distortion of quantizing subsource i using quantizer $\hat{v}_i$, and $\hat{R}_i(\hat{v}_i)$ is the bit rate obtained for subsource i with $\hat{v}_i$.

In this work, the quantizers are replaced with mappings, each having a distortion that depends on the signal and noise power. Furthermore, if power allocation has been performed so that the channel power of the mappings are different, the power consumption must be taken into account, so that the Lagrangian given by Equation (4.8) is the objective function to be minimized. The Lagrangian can be written as

$$\begin{aligned}
\mathcal{L} &= \sum_{j=0}^{J} |\mathcal{I}_j| \bar{\sigma}_j^2 D_j(P_j) + \lambda \sum_{j=0}^{J} |\mathcal{I}_j| R_j P_j \\
&= \sum_{j=0}^{J} |\mathcal{I}_j| \frac{1}{|\mathcal{I}_j|} \sum_{i \in \mathcal{I}_j} \sigma_i^2 D_j(P_j) + \lambda \sum_{j=0}^{J} \sum_{i \in \mathcal{I}_j} R_j P_j \\
&= \sum_{i=0}^{N-1} \left(\sigma_i^2 D_{v_i}(P_{v_i}) + \lambda R_{v_i} P_{v_i} \right).
\end{aligned} \tag{4.13}$$

If the signal (and noise) power is fixed, this is principally the same situation as the bit allocation problem. The rate in this work is the number of channel symbols per source symbol, as opposed to the bit rate. If (4.13) is compared to (4.11) and (4.5) to (4.12), it is seen that if $\hat{v}_i = v_i$, $\hat{D}_i(\hat{v}_i) = \sigma_i^2 D_{v_i}(P_{v_i}) + \lambda R_{v_i} P_{v_i}$, and $\hat{R}_i(\hat{v}_i) = R_{v_i}$, then $\mathcal{L} = \hat{D}_{\text{tot}}$ and $N R_{\text{tot}} = \hat{R}_{\text{tot}}$, making the problems equivalent when λ and P_j are fixed for all $j \in \{0, \ldots, J\}$[1]. If power allocation has not been performed and the class source power is the same for all the classes, this corresponds to setting $\lambda = 0$.

The algorithm is proven to be optimal in (Westerink et al., 1988), so for a fixed power allocation, if the correct Lagrangian multiplier does not change

[1] The inequality sign in Equation (4.5) does not change the principle.

from the classification, and under the assumption that changing the shape of the source pdf in each class does not change the distortion of the class, an optimal rate allocation algorithm exists.

4.2.2.6 A Rate and Power Allocation Algorithm

It will now be assumed that the classes are sorted so that $R_k > R_j$ for $k > j$ and that symbols in class 0 are not transmitted, so $R_0 = 0$ and $D_0(P_0) = 1$ for all P_0. This means that P_0 can be chosen arbitrarily, and that the term corresponding to $j = 0$ in (4.8) is constant.

Combining the rate and power allocation methods proposed in the two previous sections, the following algorithm can be obtained.

1. Initialize $P_j = P_{\text{tot}}$ for $j \in \{1, \ldots, J\}$, $v_i = 0$ for $i \in \{0, \ldots, N-1\}$, $R_{\text{sum}} = 0$, and $\lambda = 0$.

2. Find the subsource i and the class $j > v_i$ that maximizes the fraction[2]

$$\frac{R_j - R_{v_i}}{\sigma_i^2 \left(D_j(P_j) - D_{v_i}(P_{v_i})\right) + \lambda \left(R_j P_j - R_{v_i} P_{v_i}\right)}$$

   considering only i and j that satisfy $R_{\text{sum}} + R_j - R_{v_i} \leq N R_{\text{tot}}$.

3. Increase R_{sum} by $R_j - R_{v_i}$ and set v_i to j.

4. If $N R_{\text{tot}} - R_{\text{sum}} \geq \min\limits_{i;j>v_i} (R_j - R_{v_i})$, go to 2.

5. If $\lambda = 0$, set λ to λ_0.

6. Find the power distribution $\{P_j\}_{j \in \{1,\ldots,J\}}$ that minimizes the Lagrangian

$$\mathcal{L} = \sum_{j=1}^{J} \left(|\mathcal{I}_j| \bar{\sigma}_j^2 D_j(P_j) + \lambda |\mathcal{I}_j| R_j P_j \right).$$

7. Calculate the consumed power,

$$P_{\text{obs}} = \frac{1}{N R_{\text{tot}}} \sum_{j=1}^{J} |\mathcal{I}_j| R_j P_j.$$

8. If P_{obs} differs more from P_{tot} than a predefined value, guess a new λ, and go to 6.

[2] This fraction is obtained by modifying a similar fraction from the algorithm in (Westerink et al., 1988) as described in Section 4.2.2.5.

9. Calculate the average distortion

$$D_{\text{tot}} = \frac{1}{N} \sum_{j=0}^{J} |\mathcal{I}_j| \bar{\sigma}_j^2 D_j(P_j)$$

10. If the relative change of D_{tot} is larger than a predefined value, go to 2. Otherwise quit.

The guess of λ in step 8 is done by linear interpolation if values of P_{obs} both above and below P_{tot} have been observed during the the iterations, otherwise λ is multiplied by a constant (larger than one if $P_{\text{obs}} > P_{\text{tot}}$, and smaller than one if $P_{\text{obs}} < P_{\text{tot}}$). Note that in step 2, only the subsource of largest variance for each class needs to be considered,.

In this algorithm, steps 1–4 correspond to the bit allocation algorithm from (Westerink et al., 1988), with a modified distortion function. There are two major differences between the proposed algorithm and the algorithm from (Lervik, 1996; Lervik and Fischer, 1997). Firstly, that algorithm performed power allocation in a similar manner as the rate allocation, using only a finite set of power values, instead of using a Lagrange multiplier method, and secondly, it only performed one rate allocation followed by a power allocation, and no iteration between these, as in the proposed algorithm. This means that it was not necessary to take the power into account in the rate allocation algorithm, as is done here by using the Lagrangian as the distortion function.

The optimality of the rate allocation algorithm given a fixed power allocation is shown from (Westerink et al., 1988) in Section 4.2.2.5, and the power allocation algorithm is optimal for a fixed rate allocation since every critical point of the derivative of the Lagrangian is considered. In both cases, the optimality is of course conditioned on the correctness of the estimated distortion functions. However, the optimality of each part of the algorithm does not guarantee that the total algorithm is optimal. The alternation between the rate and power allocation might bring the distortion into a local minimum as a function of all the parameters, even though the two parts give global optima with respect to subsets of the parameters. The starting condition of equal power for all classes is well justified from the fact that this is optimal for a related problem, so hopefully, the algorithm gives the global optimum in most cases.

4.2.3 Direct Source–Channel Mappings

The source symbols belonging to one class are transmitted using the same mapping. In the classification, it is assumed that the performance of a mapping

is independent of the shape of the source symbol distribution. However, very few actual mappings have that property. Thus, in order to analyze the performance of the mappings, a source distribution must be assumed. The samples within one subsource are well modeled as Gaussian, but when the number of classes is low, meaning that many subsources of different variance are assigned to the same class, a Laplacian distribution is a better model, cf. (Joshi and Fischer, 1995; Lervik, 1996). Thus, a Laplacian source distribution is assumed in the design and performance evaluation of the mappings.

In this work $J = 5$ mappings are used to produce channel symbols. The mappings have rates $\frac{1}{4}$, $\frac{1}{2}$, $\frac{2}{3}$, 1, and 2. For the mappings with rate below 1, power-constrained channel-optimized vector quantization (PCCOVQ) coders (Fuldseth, 1997) are used. The mapping with rate 1 is a direct PAM mapping, while the mapping with rate 2 is a HSQLC mapping (cf. Chapter 3).

4.2.3.1 PCCOVQ Mappings

Power-constrained channel-optimized vector quantization (Fuldseth, 1997) is briefly described in Section 3.3.2. The mappings have source dimension L, channel dimension K, and codebook size M. They are characterized by the reconstruction codebook, the minimum distance of the channel symbols Δ_P, and a Lagrangian multiplier λ_P handling the power constraint.

Figure 4.5 shows the codebook for a Laplacian memoryless source. Compared to Figure 3.4(b) which shows the codebook for a Gaussian source, notice the changed shape, due to the shape of a two-dimensional Laplacian pdf.

The PCCOVQ mappings used in this work are all designed for a memoryless Laplacian source. Three sets of parameters are employed: $L = 4$, $K = 1$, $M = 256$, giving a rate of $\frac{1}{4}$, $L = 2$, $K = 1$, $M = 256$, giving a rate of $\frac{1}{2}$, and $L = 3$, $K = 2$, $M = 1024$, giving a rate of $\frac{2}{3}$. In all cases, the mappings are optimized for different channel CSNR values with approximately 1 dB spacing. Although the codebook is larger for the third parameter set, this mapping reaches saturation for a lower channel quality than the other two. The reason is the channel dimension of 2 in this case, giving only $\sqrt{1024} = 64$ channel symbols in each dimension. A codebook size of $M = 256^2 = 65536$ would give too high complexity in the optimization process, and possibly also in the encoding process of the signal. Of these three mappings, the one of rate $\frac{1}{2}$ is the only one reported in (Fuldseth, 1997). The other two were designed using the same methodology.

The performance of the three mappings is shown in Figure 4.6. It can be seen that the performance compared to the OPTA is fairly good, especially for the mappings of rates $\frac{1}{2}$ and $\frac{1}{4}$. Note that the two mappings with 256 channel symbols per channel dimension reach saturation at CSNR around 45–50 dB

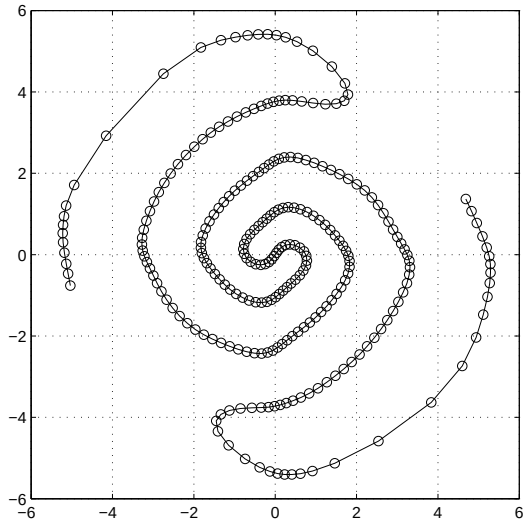


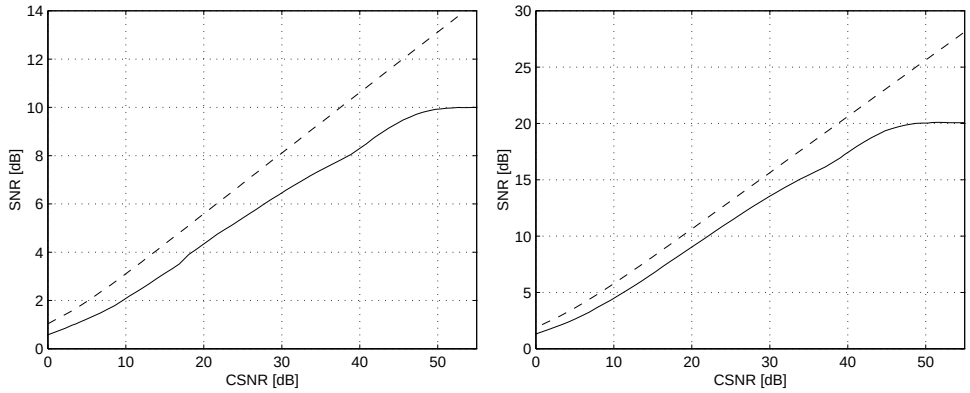
Figure 4.5 The reconstruction codebook of a PCCOVQ system with $L = 2$, $K = 1$, and $M = 256$, designed for a memoryless unit variance Laplacian source for transmission on an AWGN channel with a CSNR of 23.1 dB. The circles show the positions of the reconstruction codewords in the 2-dimensional source space, and the lines are drawn between source symbols which correspond to neighboring channel symbols.

whereas the mapping with 64 channel symbols per channel dimension reaches saturation around 30 dB.

As mentioned in Section 4.2.2.4, CSNR values between the points for which optimization has been performed can be chosen by the power allocation algorithm. For such points, the reconstruction codebook for the nearest point is chosen, while the parameters Δ_P and λ_P are interpolated linearly (using linear CSNR values) between the two neighboring points of the desired CSNR value. Of course, this does not assure the exact performance as estimated with the quadratic spline interpolation of the distortion, but due to the relatively small spacing between the points, the approximation is quite good. The reason why the codebook is not interpolated is that at some points, an abrupt change from one codebook shape to another occurs. At these points, an interpolation of the codebook vectors might lead to a constellation of poor performance.

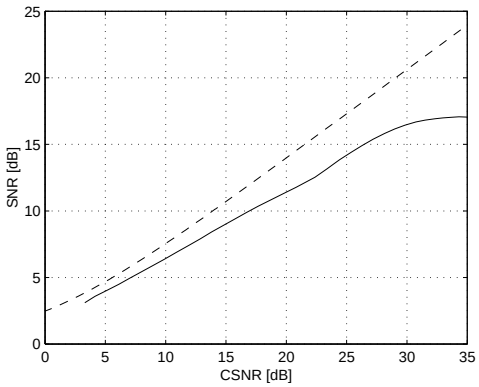
4.2.3.2 Direct PAM Mapping

For a memoryless Gaussian source and a rate of one, an optimal mapping is found simply by using multiplication by a constant as the encoder and decoder,



(a) $L = 4, K = 1, M = 256$

(b) $L = 2, K = 2, M = 256$



(c) $L = 3, K = 2, M = 1024$

Figure 4.6 SNR performance of the three PCCOVQ mappings optimized for different CSNR values for a Laplacian memoryless source (solid), compared to the OPTA for the same source (dashed)

cf. Section 3.2.1. This system is denoted a direct PAM system. If the source has unit variance, the desired channel power is P_{PAM}, and the noise power is σ_N^2, the encoder should multiply by $\sqrt{P_{\text{PAM}}}$, and the decoder should multiply by $\sqrt{P_{\text{PAM}}}/(P_{\text{PAM}} + \sigma_N^2)$, cf. Equations (3.13) and (3.14).

For source probability distributions other than Gaussian, this mapping is no longer optimal, but the performance is the same as for the Gaussian case, which is still very good for many distributions. The performance for a memoryless Laplacian source is shown in Figure 4.7.

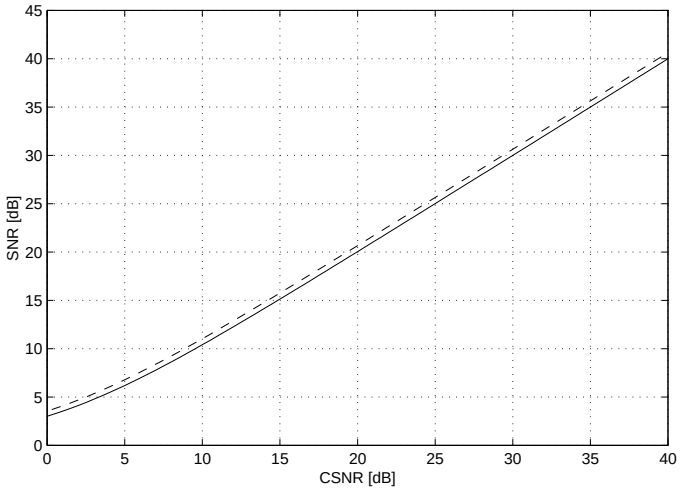


Figure 4.7 SNR performance of the direct PAM mapping for different CSNR values for a Laplacian memoryless source (solid), compared to the OPTA for the same source (dashed)

4.2.3.3 HSQLC Mapping

The hybrid scalar quantizer–linear coder (HSQLC) system is described in Chapter 3. It was shown that the use of uniform quantizers gives very little deterioration compared to optimized quantizers, and therefore, uniform quantizers will be used here. They have the advantage that they have an unlimited number of quantization intervals, which gives a better handling of probability distributions that have longer tails than the distribution for which they are optimized, cf. the performance of a quantizer optimized for a Gaussian source applied to a Laplacian source in Figure 3.20(a).

The coder is given by five parameters that vary as a function of channel quality: the four distances δ_a, δ_c, δ_d, and δ_e, and the quantization error amplification K_b. The performance of the coder on a uniform Laplacian source is shown in Figure 4.8.

As for the PCCOVQ mappings, the coder performance and parameters are tabulated for the rate and power allocation. Due to the simple optimization, the distance between the CSNR points is only 0.1 dB, and for values between the points, linear interpolation of all the parameters is used.

4.2.3.4 Normalization of the Channel Power

All the mappings are designed to give an output signal of the pre-defined desired channel power when the input is zero mean, unit variance Laplacian.

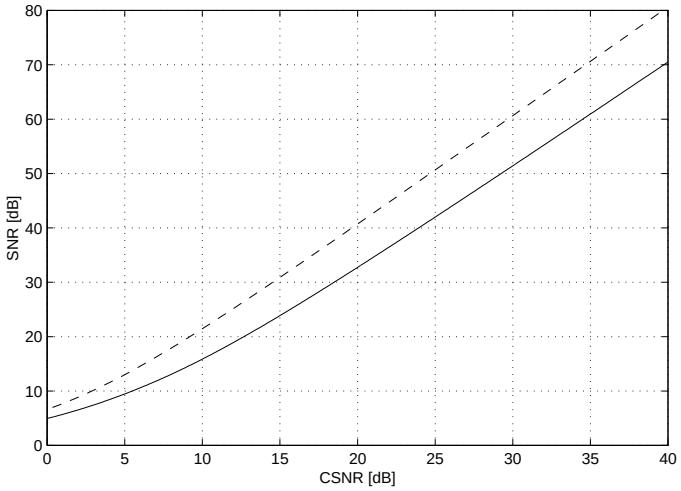


Figure 4.8 SNR performance of the HSQLC mapping optimized for different CSNR values for a Laplacian memoryless source (solid), compared to the OPTA for the same source (dashed)

The zero mean is assured by the filter bank and the subtraction of the mean of the lowpass-lowpass band, and the unit variance is obtained by dividing the source samples in each class by the observed root mean square (RMS) value. However, the probability distribution is not always exactly Laplacian. This, combined with statistical uncertainty due to a limited number of samples, means that the actual mean square output value can be different from the desired output power. To make sure that the power constraint of the output signal is satisfied and that no class consumes more power than it has been assigned at the expense of the other classes, the output values of each mapping is divided by the ratio between the RMS value of the output values of that class and the square root of the desired class power. At the receiver, the samples are multiplied by this fraction to get back the original value. For the direct PAM mapping, the fraction will always be exactly one.

4.2.4 Coding of the Side Information

The largest amount of side information from the coder is the classification table. This table has to be transmitted error free, since it is used to determine how the the received symbols are interpreted by the decoder. This means that it has to be error protected using traditional error correcting coding. The table is highly correlated and non-uniformly distributed, and can thus be

compressed. Since error protection is used and any uncorrected error is fatal anyway, traditional entropy coding can be used for compression.

In addition to the classification table, the following information needs to be transmitted error free:

- The size of the image
- The mean of the lowpass-lowpass band
- The standard deviation of the source symbols in the classes 1–5
- The Lagrange multiplier λ of the power allocation
- The ratio between the output RMS value and the square root of the desired class power for each class

The class channel power values obtained by the power allocation can be recalculated in the receiver from the Lagrange multiplier in combination with the classification table and the class source standard deviations.

The size of the image is transmitted as two unsigned 16 bit values. The mean of the lowpass-lowpass band, the class source standard deviations, and λ are all transmitted as 32 bit IEEE 754 standard floating point numbers (IEEE, 1985). The ratio between the output RMS value and the square root of the desired class power is a number close to one, and is encoded with a 12 bit uniform quantizer in the interval $(0, 2)$. The value for the direct PAM mapping is always one, so this is not transmitted. This gives a total of 320 bits. The additional side information is not compressed. The accuracy of the 32 bit floating point numbers is higher than actually needed, so the amount of data could be reduced.

4.2.4.1 Compression of the Classification Table

The classification table is correlated because neighboring blocks in one subband tend to have about the same source variance due to the image correlation, and because blocks on similar location within a subband tend to have correlated source variance values as details produce high variances in all subbands. In addition, low frequency subbands tend to have higher class source variance than high frequency subbands.

The compression of the classification table is carried out through first order adaptive arithmetic coding (Nelson and Gailly, 1996). In order to exploit the correlation, the classification table is scanned as shown in Figure 4.9. First, the blocks from the lowpass-lowpass band of the parallel filter bank (which have been filtered again with a tree-structured filter bank) are scanned, then the

1	8	9	16	23	130	143	250	24	129	144	249	31	122	151	242
2	7	10	15	40	113	160	233	39	114	159	234	32	121	152	241
3	6	11	14	53	100	173	220	54	99	174	219	61	92	181	212
4	5	12	13	70	83	190	203	69	84	189	204	62	91	182	211
17	136	137	256	22	131	142	251	25	128	145	248	30	123	150	243
46	107	166	227	41	112	161	232	38	115	158	235	33	120	153	240
47	106	167	26	52	101	172	221	55	98	175	218	60	93	180	213
76	77	196	197	71	82	191	202	68	85	188	205	63	90	183	210
18	135	138	255	21	132	141	252	26	127	146	247	29	124	149	244
45	108	165	228	42	111	162	231	37	116	157	236	34	119	154	239
48	105	168	225	51	102	171	222	56	97	176	217	59	94	179	214
75	78	195	198	72	81	192	201	67	86	187	206	64	89	184	209
19	134	139	254	20	133	140	253	27	126	147	246	28	125	148	245
44	109	164	229	43	110	163	230	36	117	156	237	35	118	155	238
49	104	169	224	50	103	170	223	57	96	177	216	58	95	178	215
74	79	194	199	73	80	193	200	66	87	186	207	65	88	185	208

Figure 4.9 Scanning order for the classification table, demonstrated on 4×4 parallel subbands with two dyadic splits and a total of 16×16 blocks. The numbers indicate the order in which the block class numbers are organized. The solid lines denote subband limits, and the dotted lines denote block limits. The subband image is organized with lower frequencies to the left and to the top.

other bands are scanned so that blocks with the same position in neighboring bands are placed after each other. The figure shows the situation for a smaller filter bank than in this work in order to demonstrate the principle.

4.2.4.2 Error Protection of the Side Information

For protection of the side information, traditional error correcting coding is used, since no error is allowed in this part of the signal. Because the amount of side information is quite small and it must be possible to transmit one image individually, the block length of the error correcting code must be small. In addition to the error correcting code, a signaling constellation must be selected since the channel has continuous amplitude.

Reed–Solomon coding (Blahut, 1983) is a frequently used block code. It operates on a coding alphabet p^ν, where p is a prime and ν is a positive integer. The codeword length is $n = p^\nu - 1$, and the number of data symbols per code word is k. The code can then correct up to $t = (n - k)/2$ symbol errors in a

block. Usually, $p = 2$, and the codeword length is then νn bits.

Here, binary symbols are used to encode the side information, so a Reed–Solomon code of $p = 2$ is used. The signaling represents s bits per channel symbol, so the signaling is 2^s-PAM. The bit error probability p_b can be found from the CSNR and s (Proakis, 2001), and from n and k, the block error rate of the Reed–Solomon code can be found as in Equation (2.37), except that p_b is replaced with the symbol error probability $1 - (1 - p_b)^s$. For a given block size n and a maximum error probability, the combination of s and k that maximizes the number of data bits per channel symbol, $(ks)/(\nu n)$, is used.

In this work, CSNR values down to 10 dB are considered. The signaling is 2-PAM, the code symbols consist of 5 bits, and the codewords have 31 symbols with 25 data symbols and 6 parity symbols. This gives a block error rate of $6.74 \cdot 10^{-6}$ at 10 dB.

4.2.4.3 Finding the Size of the Side Information

The size of the side information of an image is not known until the classification and entropy coding have been performed. If a certain number of channel samples is permitted for an image, this means that the number of channel samples available for transmitting the subband image samples is not known when the classification is performed. Thus, an iterative method is used to find the side information size.

Since block coding is used to protect the side information from channel errors, the side information size is quantized, which makes it easier to find. First, an initial guess is made for the side information size, based on the coding rate and the image size. Then, the corresponding number of channel samples for transmitting the subband samples is calculated, and classification is performed based on that number of channel samples. If entropy and channel coding of the side information gives the same size as the estimate, the classification is finished, otherwise, the side information size is assumed not to change, and classification is performed again. This is normally done until the estimate of the side information size is equal to the actual size. Because the size of the coded side information changes little as the rate of the remaining channel signal is changed, convergence is normally reached quickly, within 3–4 iterations. Sometimes, however, the estimate will never be correct. In that case, the smallest of the side information estimates considered in the iterations that have given a total number of channel symbols below the desired number is chosen.

If the coding rate of the image does not have to be specified exactly, a coding rate that does not include the side information can be given. This will speed up the encoding, since classification is performed only once.

4.2.5 A Summary of Changes

As mentioned, the framework of the proposed coder is the same as for the coder from (Lervik, 1996). Most of the coding blocks have been changed, though. In order to provide an overview of the differences, the significant changes are listed below:

- The filter bank has been changed.
- The power allocation is performed using a Lagrange multiplier method.
- Iteration between rate and power allocation has been introduced.
- The power and bandwidth allocation is adjusted to an expected channel quality, while in (Lervik, 1996), the channel assumption was fixed.
- All the mappings have been changed, except the direct PAM mapping. Particularly:
  - Lervik had a mapping combining two symbols of different significance, which were given nominal rates of $\frac{3}{4}$ and $\frac{1}{4}$.[3] This has been replaced with a mapping of rate $\frac{2}{3}$, which takes three symbols of the same significance.
  - A mapping of rate 2 has been introduced.
  - The mappings are designed according to the assumed channel quality.
- The coding of the side information has been changed, introducing arithmetic coding of the classification table. This calls for finding the side information size iteratively.

4.3 Coding Results

The proposed coder is evaluated by encoding some images at different rates and assumed channel qualities, adding white Gaussian noise to the signal, and decoding. The results are mainly evaluated using a mean square error distortion measure, given as the peak signal-to-noise ratio (PSNR) measured

[3] A problem with this mapping was that the symbol of nominal rate $\frac{1}{4}$ got a significantly higher distortion than the symbols of rate $\frac{1}{4}$ encoded with another mapping.

in dB, which is defined as

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{(x_{\max}-x_{\min})^2}{\frac{1}{S_xS_y}\sum_{k=0}^{S_x-1}\sum_{l=0}^{S_y-1}(x(k,l)-\hat{x}(k,l))^2}\right), \tag{4.14}$$

where S_x and S_y are the horizontal and vertical image sizes, $x(k,l)$ is the pixel value at position (k,l), $\hat{x}(k,l)$ is the decoded pixel value, and $x_{\max}$ and $x_{\min}$ are the maximum and minimum *possible* pixel values of the original image, so for an image represented with one 8 bit integer per pixel, $x_{\max} - x_{\min} = 255$.

It is well known that distortion measures based on square error give limited insight in the perceptive distortion of an image. Still, they are frequently used due to their simple calculation and the lack of good alternatives. Also, when comparing images which have the same type of errors, square error measures correspond quite well to the perceptive quality. In order to show what kind of artifacts the coder introduces, a few image examples will be given.

The images that are used to evaluate the coder performance are shown in Appendix C. All the pictures have size 512×512 and are represented with 8 bit grayscale values.

4.3.1 Performance of the Proposed Coder

Figure 4.10 shows the result when coding these images with the proposed coder for different rates and channel qualities. The coder has been optimized for the actual channel quality in each case. The PSNR values are calculated simulating five transmissions of each image[4] and averaging the MSE; this applies to all reported results for the proposed coder. As expected, the performance increases with the CSNR and the rate. The PSNR values are quite different for the four images at the same rate and CSNR. This has to do with the level of detail, and a detailed image naturally needs more resources to be transmitted satisfactorily. However, the subjective difference is not as large as the PSNR values could indicate, because in general, less detailed images do not need as high PSNR values as less detailed ones to get the same subjective quality.

Figure 4.11 shows the percentage of the channel signal that must be spent on side information. This percentage naturally decreases as the rate goes up, however, the absolute size of the side information increases with the rate. The

[4] A number as low as five is chosen because the size of the images (512×512 pixels) means that a lot of channel symbols are generated for one image, and the variation in decoded quality from image to image is therefore relatively small.

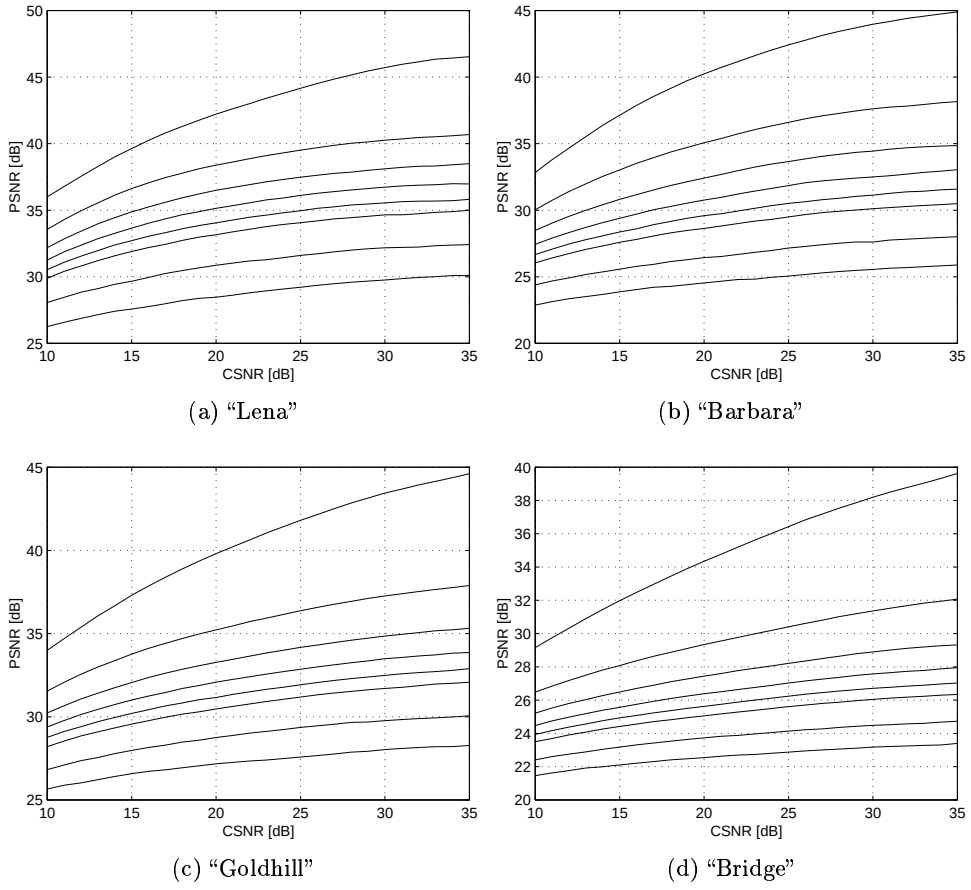
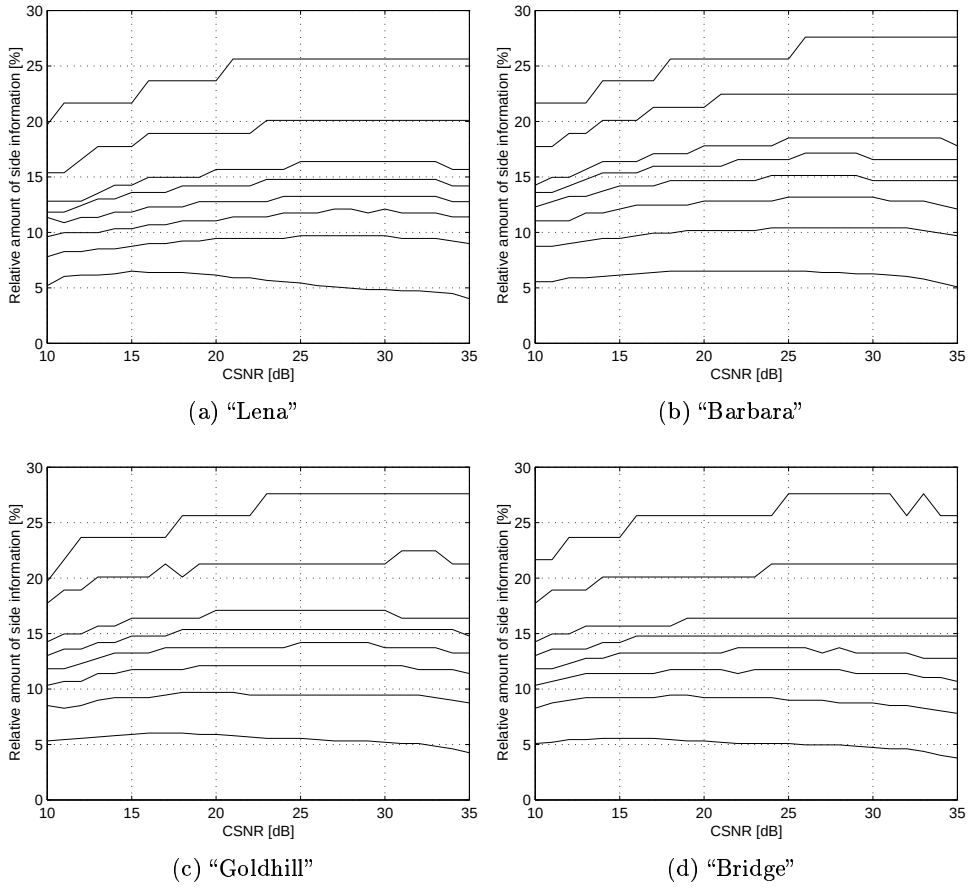


Figure 4.10 Coding results for the for test images as a function of CSNR for a coder optimized for the actual channel quality. The rates are (from below) 0.03, 0.05, 0.083, 0.1, 0.125, 0.166, 0.25 and 0.5 channel symbols per pixel.

side information percentage takes only a finite number of values due to the block coding of the side information.

The probability of side information error in an image depends on the number of coding blocks used. For the images, rate, and channels considered in Figure 4.10, the number of Reed–Solomon blocks needed to transmit the side information varies from 10 to 55, with higher numbers for higher coding rates. With 55 blocks at 10 dB, the image error probability is $3.71 \cdot 10^{-4}$, which corresponds to one error in 2698 transmitted images. Larger images can of course get higher error probabilities. For channels below 10 dB, the error protection



(a) "Lena" (b) "Barbara"

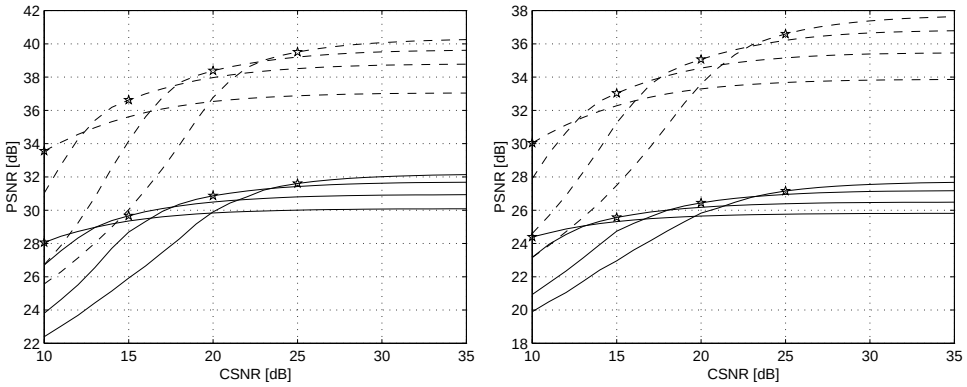
(c) "Goldhill" (d) "Bridge"

Figure 4.11 Percentage of side information in the encoded image streams as a function of CSNR for different images and rates. The rates are (from above) 0.03, 0.05, 0.083, 0.1, 0.125, 0.166, 0.25 and 0.5 channel symbols per pixel.

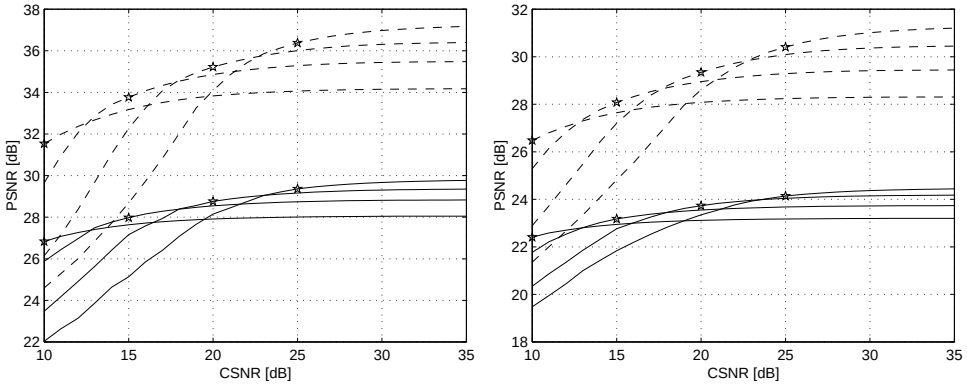
is not sufficient, and if such channels can occur, the protection must be improved. Above 10 dB, the image error probability is quickly reduced, at 13 dB, the block error rate is $4.89 \cdot 10^{-15}$. If the channel quality is never allowed to go down to 10 dB, the error protection could be weakened.

4.3.2 Robustness

An important motivation for using joint source–channel coding is robustness against changing channel qualities. The results in Figure 4.10 assume that the coder knows the actual channel quality. This is not always the case. Fig-



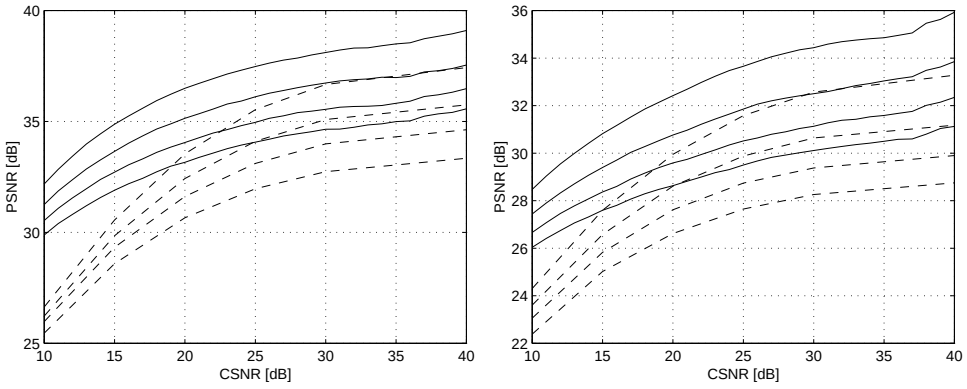
(a) "Lena" (b) "Barbara"



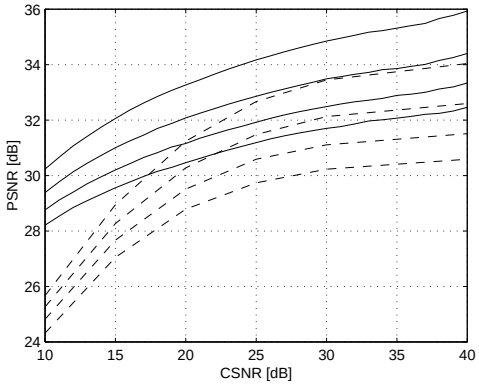
(c) "Goldhill" (d) "Bridge"

Figure 4.12 Performance at channel qualities different from the design CSNR for design CSNR values of 10, 15, 20, and 25 dB. The rates are 0.05 (solid) and 0.25 (dashed), and the design CSNR of each curve is marked with a star.

ure 4.12 shows the results for different design CSNR values when the channel quality changes. For channels worse than the design CSNR, the performance drops quite rapidly, but some dB below the design CSNR, the results are still acceptable. Above the design CSNR, some improvement in obtained, but not as much as if the design CSNR were higher.



(a) "Lena" (b) "Barbara"



(c) "Goldhill"

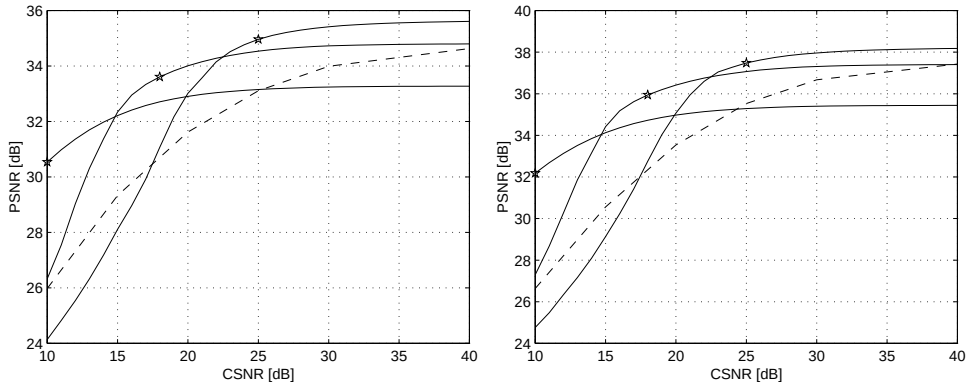
Figure 4.13 Performance comparison of the proposed coder (solid) to the coder from (Lervik, 1996) (dashed). The rates are (from below) 0.083, 0.1, 0.125, and 0.166 channel samples per pixel.

4.3.3 Comparison to Earlier Coder

Figure 4.13 shows the coder performance[5] compared to the coding results reported by Lervik (1996). The results are substantially better for all the rates, channel qualities, and images. The reference coder has been designed with side information protection that works down to 10 dB, just as the proposed coder.

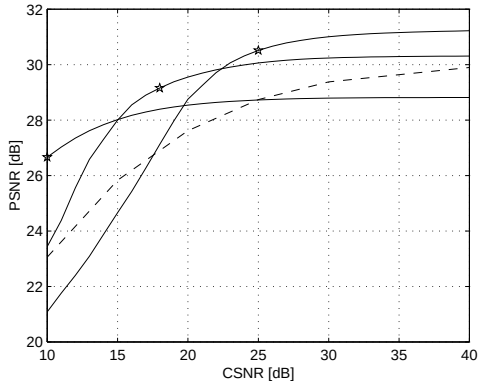
The coder in (Lervik, 1996) did not perform any optimization for different

[5]The increased slope between CSNR values of 35 and 40 dB can be explained by the similar behavior for the PCCOVQ mappings of rates $\frac{1}{4}$ and $\frac{1}{2}$ at these channel qualities, cf. Figure 4.6(a)–(b).

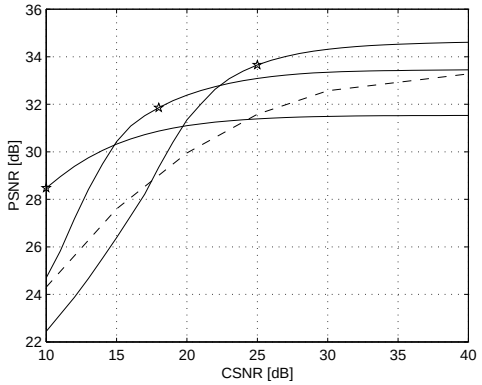


(a) "Lena", rate 0.1

(b) "Lena", rate 0.166



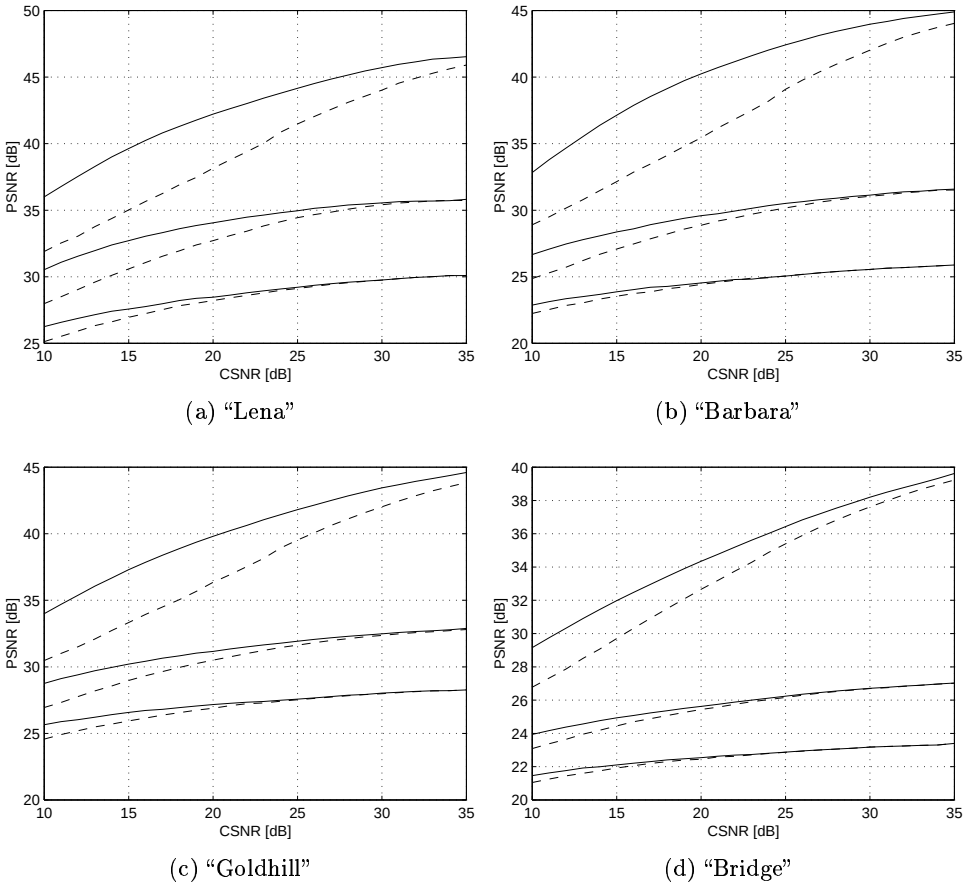
(c) "Barbara", rate 0.1



(d) "Barbara", rate 0.166

Figure 4.14 Coding results for various images and rates compared to the coder from (Lervik, 1996) at CSNR values different from the design CSNR. The solid line is the proposed coder optimized for CSNR values of 10, 18, and 25 dB (marked with a star), and the dashed line is the coder from (Lervik, 1996).

channel qualities. Thus, the results reported there are valid also when the encoder does not know the channel quality. In order to obtain that property also for the proposed coder, it must be designed for one particular CSNR. The result when this is done is illustrated in Figure 4.14. As seen, the design CSNR will decide where the proposed coder outperforms the reference. For a design CSNR of 18 dB, the proposed coder outperforms the reference at all the displayed channel qualities for the images and rates shown in Figure 4.14. For CSNR values around the design CSNR, the improvement is large.



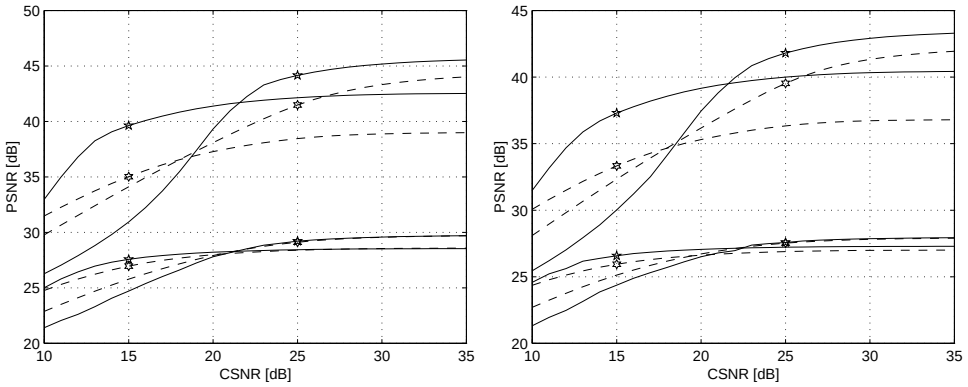
(a) "Lena" (b) "Barbara"

(c) "Goldhill" (d) "Bridge"

Figure 4.15 Coding results with (solid) and without (dashed) the use of the rate 2 HSQLC mapping. The rates are (from below) 0.03, 0.1, and 0.5 channel symbols per pixel.

4.3.4 Influence of the Bandwidth Expanding Mapping

One of the changes from the coder from (Lervik, 1996) is the inclusion of a bandwidth expanding mapping, the HSQLC mapping of rate 2. The improvement obtained by including that mapping is shown in Figure 4.15. This mapping has a large influence on the result for high rates and poor channels, while for low rates and good channels, there is no difference at all, simply because the HSQLC mapping is never allocated in these situations. For high rates, the rate 2 mapping is naturally used more frequently than for low rates. For poor channels, it is more desirable than for good channels, because the important



(a) "Lena" (b) "Goldhill"

Figure 4.16 Performance at channel qualities different from the design CSNR with (solid) and without (dashed) the use of the rate 2 HSQLC mapping. The rates are (from below) 0.03 and 0.5, and the design CSNR values are 15 and 25 dB, marked with a star.

symbols need protection on the poor channels at the expense of representation of the details, while for good channel, sufficient protection can be obtained with a mapping of lower rate, allowing also less important subband samples to be transmitted. Also, the improvement by using the HSQLC mapping is smaller for a detailed image like "Bridge" than for a less detailed image like "Lena". This is because the relative difference between subband variances are larger in an image of low detail level, so the high rate mapping can be used more frequently.

A problem with the HSQLC mapping is that it is less robust than the direct PAM mapping. The consequences of this is illustrated in Figure 4.16, which shows that when the design CSNR is high, the performance of the coder is worse at low CSNR values when the HSQLC mapping is included. This applies even when the gain at the design CSNR is small, cf. the rate 0.03 at a design CSNR of 25 dB.

The visual effect of the bandwidth expanding mapping is shown in Figure 4.17. For a design CSNR and actual CSNR of 25 dB, the difference is barely noticeable, but when the CSNR drops to 15 dB, there is a difference. However, it is questionable whether the image with the HSQLC mapping is the worst, even though it has the lowest PSNR. The low PSNR value can be partially explained from the darkened area around the left side of the mouth, which is probably due to a decoding error of the quantized symbol in the

(a) 35.0 dB/34.5 dB (b) 27.7 dB/26.9 dB (c) 32.7 dB/31.9 dB

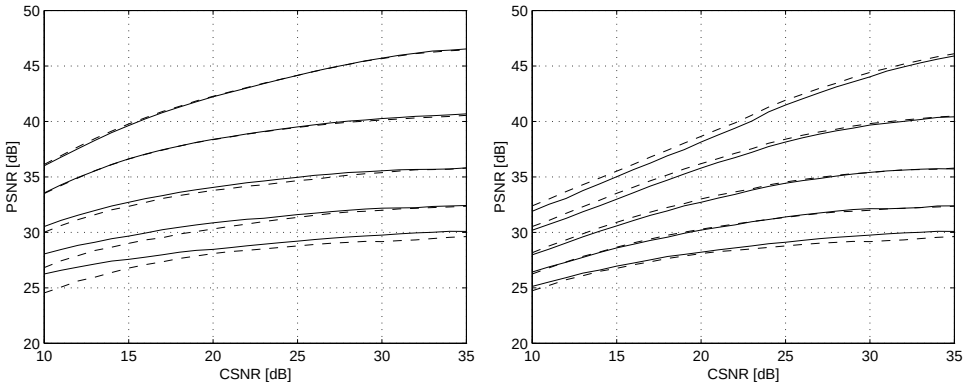
(d) 34.4 dB/33.7 dB (e) 29.5 dB/28.4 dB (f) 30.6 dB/29.6 dB

Figure 4.17 Extract from "Lena" at a rate 0.1 with (upper row) and without (lower row) the bandwidth expanding mapping. In the left column, the design CSNR and the actual CSNR are both 25 dB, in the middle column, the design CSNR is 25 dB and the actual CSNR is 15 dB, and in the right column, the design and actual CSNR are both 15 dB. Under each extract, the PSNR of the whole image (with the applied noise sequence) and the PSNR of the shown segment are stated.

HSQLC mapping. Such spots caused by fairly large errors in low frequency symbols can be seen all over the image, and can be somewhat more annoying in the background, which is not shown here. For the design CSNR of 15 dB with the same actual CSNR, the coder with HSQLC is clearly preferable, both visually and from the PSNR value.

4.3.5 Influence of the Tree-Structured Filter Bank

The filter bank consists of an 8 band uniform filter bank followed by three dyadic splits of the lowpass-lowpass band of the uniform bank, cf. Section 4.2.1.



(a) With HSQLC mapping (b) Without HSQLC mapping

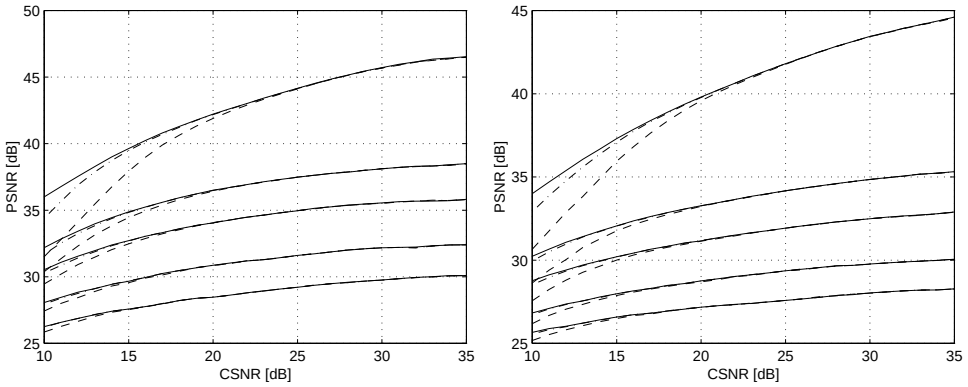
Figure 4.18 Results for "Lena" with (solid) and without (dashed) dyadic splitting of the lowpass-lowpass band of the parallel filter bank at rates (from below) 0.03, 0.05, 0.1, 0.25, and 0.5.

Such a tree-structured filter bank, which is not used in (Lervik, 1996), has been introduced to remove more of the correlation of the image before encoding. The significance of this is shown in Figure 4.18, which shows the coding results with the proposed filter bank compared to results with the uniform filter bank part of this, where the dyadic splits are not performed. The figure shows that more gain is obtained by the dyadic splits at low rates; for a rate of 0.5, the performance actually drops. If the bandwidth expanding mapping is not used, the performance is better without the dyadic splits for most of the rates.

An explanation of this behavior is that the dyadic splits not only remove correlation, they also increase the variance of the subband samples in the lowest frequency bands. Thus, with the use of a tree-structured filter bank, some symbol of a very large variance need to be transmitted. This is best done with a mapping of high rate relative to the total rate. Such a mapping is available for low rates, but not high, if the rate of the mapping is taken relative to the total rate. Removing the mapping of the highest rate is then worse when the dyadic splits are used.

4.3.6 Influence of the Power Allocation

The significance of the power allocation is shown in Figure 4.19. The figure compares the results of the proposed coder to the results obtained when power allocation is disabled, i.e., when all the class channel powers are equal to the



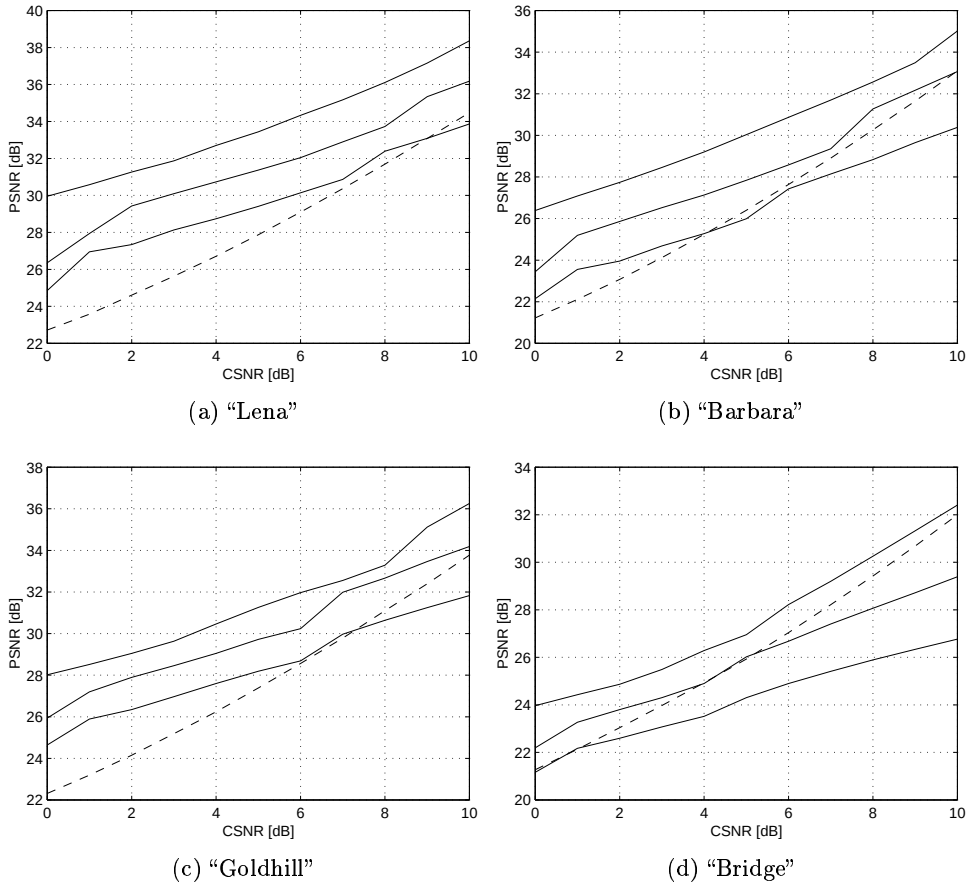
(a) "Lena" (b) "Goldhill"

Figure 4.19 Coding result with power allocation (solid), without power allocation (dashed), and with power allocation after the rate allocation, but no iterations to redo the rate allocation (dash-dotted). The rates are (from below) 0.03, 0.05, 0.1, 0.166, and 0.5, and the design CSNR is equal to the actual CSNR.

power constraint of the channel. In addition, results when the iterations between rate and power allocation described in Section 4.2.2.6 are disabled, i.e., when the algorithm stops the first time point 10 is reached. This corresponds to the algorithm from (Lervik, 1996; Lervik and Fischer, 1997), except that the power allocation is performed by a different method.

The results show that for poor channels, in particular for high rates, a large improvement is obtained by the power allocation, while in other cases, the improvement is insignificant. Most of the improvement is obtained without the iterations between rate and power allocation, but in some cases, a significant gain can be observed when introducing it, for instance, a 1.5 dB improvement is obtained for "Lena" at a CSNR of 10 dB and a rate of 0.5. For good channels, where the improvement is small, the difference in allocated channel power between the classes is relatively small, which explains the little change in performance.

It is no coincidence that the cases where the bandwidth expanding mapping gives the largest improvement are also the cases where power allocation is most useful. For poor channels, protection of the most important symbols to avoid the full amount of channel noise to deteriorate this portion of the signal is essential. This can be obtained either by higher rate or by increased power for these symbols. For poor channels, it can be observed that the classes of the highest rates also have the largest power allocated to them. This indicates



(a) "Lena" (b) "Barbara"

(c) "Goldhill" (d) "Bridge"

Figure 4.20 Performance of the proposed coder on poor channels. The rates not including side information are 0.25, 0.5, 1.0 (solid, from below), and 1.5 (dashed).

that mappings of higher rates would be desirable.

4.3.7 Performance on Poor Channels

So far, only CSNR values down to 10 dB have been considered, and the channels in the lower quality segment of the considered channels have been denoted "poor channels". In many applications, these are actually good channels, and thus, also channels of lower quality should be considered. The side information protection that has been used only works down to around 10 dB, and with the method used, it is only possible to protect well for channels down to approximately 7 dB. In order to code the side information for lower channel qualities,

a different method, such as a convolutional code (Blahut, 1983) in combination with a Reed–Solomon code, must be considered. This is not investigated further. Instead, results given for channels below 10 dB will be stated for rates not including the side information. The actual rates will thus be somewhat higher.

Figure 4.20 shows coding results for CSNR values between 0 and 10 dB. The rates considered are higher than before, since a high rate is needed to get a reasonable quality on such a channel. The most striking result is that the rate of 1.5 performs much worse than lower rates. This is due to a shortcoming of the rate and power allocation algorithm. The rate allocation part will always allocate as much of the specified rate as possible, even if allocation steps lead to deterioration in quality. The reason why allocating more rate can lead to a quality loss, is that when too many blocks are allocated rate 2, the power of that class cannot be as high as when the class has fewer blocks allocated. For instance, for "Lena" with a CSNR of 0 dB and a rate of 1.0, the symbols that have been allocated a rate of two has a power giving those symbols a channel signal-to-noise ratio of 11.8 dB (the low rate symbols have very little power allocated), while for a rate of 1.5, the channel signal-to-noise ratio of the rate 2 symbols is only 1.76 dB. Thus, the protection of the most important symbols drops, but less important symbols are better protected. This result is a convincing proof that mappings of higher rates are essential.

Figure 4.21 shows the visual quality of the proposed coder. At 0 dB, the attainable quality is limited, although it will be sufficient for some applications, given enough rate. For 5 dB, more acceptable results can be obtained.

4.3.8 Comparison to a Traditional System

Although some joint source–channel image coders exist, the far most common way to deal with the problem of image transmission is to perform source and channel coding separately. Recent developments in channel coding have resulted in codes with a performance very close to the channel capacity for AWGN channels (Nickl, Hagenauer and Burkert, 1997; MacKay and Neal, 1997; Chung, Forney, Richardson and Urbanke, 2001), and this fact supports the use of separate source and channel coding. However, similar robustness as obtained with the proposed coder has not been demonstrated.

It should be noted that the distance from the Shannon limit of 0.27 and 0.0045 dB stated in the titles of the abovementioned papers refer to the limit for a *binary input* AWGN channel, not the capacity of a general input AWGN channel. The papers consider low CSNR values where these two limits are close, but for the CSNR values considered for the proposed coder, the bounds are far apart. Since the channel model accepts any input symbol, and this is utilized

(a) 24.6 dB/23.9 dB (b) 26.1 dB/25.6 dB (c) 30.0 dB/29.1 dB

(d) 29.3 dB/28.5 dB (e) 31.4 dB/30.5 dB (f) 33.4 dB/32.9 dB

Figure 4.21 Extract from "Lena" at a CSNR of 0 dB (upper row) and 5 dB (lower row) and rates not including side information of (from left) 0.25, 0.5, and 1.0. Under each extract, the PSNR of the whole image (with the applied noise sequence) and the PSNR of the shown segment are stated.

in the proposed coder, a coding method that uses more general modulation schemes than BPSK is needed in order to come close to the channel capacity for channels with CSNR from 10 dB and up. Myhre, Markhus and Øien (2001) use QAM modulation with Gray coding and low density parity check (LDPC) codes[6] (Gallager, 1962; MacKay, 1999) with soft decision. For CSNR values in the range 10–25 dB, performance 3.0–3.2 dB away from the channel capacity of a general input AWGN channel is obtained, and the error probability curves are very steep. Based on this result, the channel code will be assumed to give zero error probability for CSNR values more than 3.0 dB to the right of the the channel capacity curve as a function of CSNR, and a high probability of

[6]Also known as Gallager codes

Total rate	Design CSNR		
	15 dB	20 dB	25 dB
0.1	0.2037	0.2838	0.3659
0.25	0.5093	0.7095	0.9147

Table 4.1 Bit rates in bits per pixel for the JPEG 2000 coder corresponding to different total rates and channel qualities

error for lower CSNRs. The steepness of the error probability curves justifies this assumption. The channel capacity is given by (Shannon, 1948)

$$C = \frac{1}{2}\log_2\left(1 + \frac{P_{\text{tot}}}{\sigma_N^2}\right), \tag{4.15}$$

where P_{tot} is the power constraint and σ_N^2 is the noise power of the channel.

A new image coding standard, JPEG 2000, has recently become an international standard. This coder can certainly be called state-of-the-art, and thus, as a reference image source coder, the JPEG 2000 part 1 baseline coder (ISO/IEC, 2001), implemented in verification model 8.0, is used. The standard supports insertion of resynchronization symbols in order to limit the consequences of channel bit errors, but due to the steepness of the error curves of the channel code, the bit error rate is almost always either very high or very low, so such symbols are not useful.

The proposed coder is designed for a certain CSNR. When setting up the reference coder, the channel code will be assumed to have a channel rate R_c, measured in bits per channel symbol, equal to the capacity of a channel 3 dB below that CSNR,

$$R_c = \frac{1}{2}\log_2\left(1 + 10^{\frac{\text{CSNR}-3}{10}}\right). \tag{4.16}$$

The rate of the proposed coder, R_{tot} is given as the number of channel samples per pixel, and if the reference coder shall give the same rate, the bit rate R_s of the source coder must be given by

$$R_s = R_{\text{tot}} R_c. \tag{4.17}$$

Figure 4.22 shows coding results with the proposed coder compared to the results with the reference coder. The proposed coder is clearly more robust since it can give fairly good results also below the design CSNR, while the reference coder cannot offer decoding below the design CSNR. For better

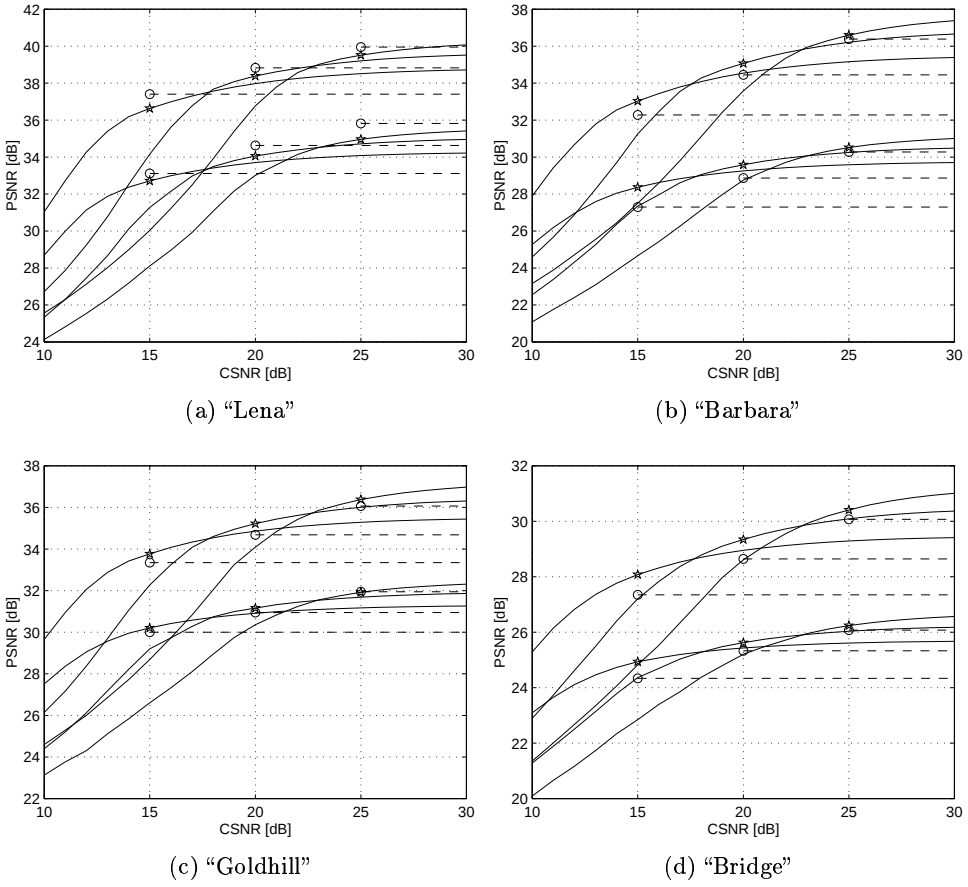


Figure 4.22 Coding result with the proposed coder (solid) and the JPEG 2000 coder followed by a channel coder (dashed). The total rate is (from below) 0.1 and 0.25 channel samples per pixel, and the design CSNR values are 15, 20, and 25 dB (marked with a star or a circle). The PSNR of the JPEG 2000 coder is not given below the design CSNR, since correct reception of the image cannot be expected there. The bit rates of the JPEG 2000 coder are given in Table 4.1.

(a) 34.6 dB/34.3 dB (b) 34.0 dB/33.4 dB (c) 32.9 dB/31.9 dB

(d) 30.9 dB/29.5 dB (e) 31.2 dB/29.7 dB (f) 30.3 dB/28.5 dB

Figure 4.23 Extracts from "Lena" and "Goldhill" with a design CSNR of 20 dB and a total rate of 0.1 channel symbols per pixel. The left column shows the JPEG 2000 coding result, the middle column shows the result with the proposed coder at the design CSNR, and the right column shows the proposed coder at a CSNR of 17 dB. Under each extract, the PSNR of the whole image (with the applied noise sequence) and the PSNR of the shown segment are stated.

channels than expected, the proposed coder gives a certain gain, which is not the case for the reference coder. The performance at the design CSNR varies from image to image. For "Lena", the reference coder is clearly better, while for the other images, the proposed coder is usually better even at the design CSNR. There is a tendency that the proposed coder is better for detailed images, while the reference coder is better for less detailed images, although the set of images is too small to verify such a claim.

Some image examples are shown in Figure 4.23, which demonstrates the difference in the artifacts obtained with the two coders. For "Lena", the JPEG 2000 coder gives a smoother look of the face, which is visually pleasant,

although the original image has a fine texture that is lost in the coding process. For the proposed coder, this texture seems to be exaggerated. This is because the highest frequency components of the texture have been removed and channel noise has appeared in the image. The last statement is shown clearly for a worse channel, where the speckled pattern is stronger. Note that with the JPEG 2000 coder and the assumed channel coder, decoding would be impossible for a channel of 17 dB, while the quality with the proposed coder is still satisfactory. For "Goldhill", the proposed coder reproduces the image with better detail, for instance at the roof tiles. It seems that the proposed coder is visually better than JPEG 2000 for a detailed image, and vice versa for a non-detailed image. The difference is larger that the PSNR difference would indicate.

4.4 Discussion

The proposed coder has shown good results compared to both the coder from (Lervik, 1996) and the JPEG 2000 coder with a channel coder. Especially the second result is interesting, since the proposed coder has a performance similar to a combination of state-of-the-art source and channel coders at the design channel quality, but with an added feature of robustness against changing channel qualities.

The JPEG 2000 coder uses entropy coding, while the proposed coder allocates a fixed-rate coder from a limited set of mappings to each block. These schemes were compared in a simplified manner in Section 3.5.4, where it was shown that PCCOVQ mappings perform better than a combination of scalar quantization, ideal entropy coding and ideal channel coding. The same would apply for the direct PAM mapping, since its performance is equal to the OPTA for a Gaussian source. It was also shown that the mappings provide a greater robustness against varying source variances. However, the entropy coding scheme gave a better performance than the HSQLC, although that was with a channel coder that obtains the capacity; with a channel coder 3 dB below it, as assumed here, they would have about the same performance. Entropy coding based systems also have some other advantages over the direct source–channel mappings. Any rate is achievable, while the mappings can be designed only for a few rates. Furthermore, with adaptive entropy coding, it is possible to adjust to the actual distribution of the samples that are to be coded with the same entropy coder, to a large extent overcoming the robustness problems demonstrated in Section 3.5.4. All in all, these points give an indication of why the two coders give quite similar results. It should be noted that the JPEG 2000 coder is a result of a long process involving hundreds of scientists

optimizing all parameters and processes, and a similar process for the proposed coder would probably give room for a substantial improvement. On the other hand, JPEG 2000 is not merely optimized for MSE performance, but also for complexity, progressive decoding abilities, to some extent visual quality, and other factors.

In the design of the coder, robustness has not been addressed directly, both the mappings and the classification algorithm are designed with performance at the design channel quality as the only criterion, robustness is a side effect. The robustness is a property of the mappings; since they have no or very fine quantization, the performance will change with the channel quality, and since the dimensionality of the mappings is low, the adaptation cannot be done to one exact channel quality. The comparisons to (Lervik, 1996) reveal, however, that while the performance near the design CSNR has improved a lot, the performance far away from it does not improve much. If robustness is taken to mean the relative change in performance when the channel quality is different from expected, the robustness has been reduced, but if it is taken to mean the performance over a large range of channel qualities, it is improved, since it is better for some qualities and equally good for others. The results when the HSQLC mapping is disabled, reveals that this mapping to a large extent causes the effect that may be called reduced robustness.

A way of increasing the robustness could be to design the mappings with an extra criterion of robustness. This could for instance be done by imposing a constraint on the performance of one or more CSNR values different from the value for which the mapping is optimized. When considering the HSQLC mapping, the problem concerning the robustness below the design CSNR is that the probability of error of the quantized symbol becomes quite large when the CSNR drops below the design CSNR. Thus, the quantization step size should probably be increased somewhat compared to the optimal value at the design CSNR, in order to get higher robustness below that value.

If the decoder is able to estimate the actual channel quality, the performance of the HSQLC can be improved, as was demonstrated in Section 3.5.1. With a high design CSNR, the improvement is moderate, though. It should also be possible to perform a similar adaptation of the decoders for the other mappings. If a coder is to be designed for operation on channels of lower qualities than what is considered in this work, adaptation of the decoders to the actual CSNR will be more important.

Several of the results indicate that mappings of rates higher than two are needed in many situations. The problem is how such mappings could be designed, which is discussed in Section 3.6.3. One possibility is a combination of a traditional bit based quantizer (scalar or vector) followed by a channel coder,

and transmission of the quantization error, as shown in Figure 3.5. If such a scheme is used, the quantized symbols could be channel coded together with the side information. This would make it possible to use more efficient error correcting coding with larger block lengths, and the classification algorithm could be designed such that the blocks are filled up. A problem with such a scheme is that the robustness is reduced, since the error correcting code will not work below a certain level, and only the analog symbol gives performance improvement above that level. For very poor channels (near 0 dB), the need for higher rate mappings is the largest. In this case, simple BPAM mappings give a good performance compared to the OPTA, cf. Section 3.3.1, and can be used.

For very poor channels, the side information must be better protected than what has been done in the proposed coder. This means that the relative amount of data spent on side information increases. If the error protection of the side information is done with a fixed code regardless of the channel quality, this applies in all cases. Coders designed for very good channels will not give acceptable performance on very poor channels even if the side information is decoded correctly. Thus, it should be possible to adapt the lowest channel quality for which the side information can be decoded to the rest of the coder. This can be done by transmitting information on how the side information is coded first. This very small amount of data must be well protected.

The power allocation gives a large improvement of the coder performance in some cases, but the results indicate that this is because of the lack of mappings with high enough rates. If such mappings are provided, the influence of the power allocation would probably be small, and the coder could then be simplified by removing it.

The rate and power allocation algorithm is likely to find the global minimum of the estimated distortion function given a rate and power constraint, cf. the discussion in Section 4.2.2.6. However, it does not take the size of the side information into account. It might be possible to make a modification of the algorithm so that allocating blocks to an improbable class gets an extra penalty due to the increased side information size it causes. Still, the exact contribution is difficult to calculate due to the complexity of the side information coding. There is also no guarantee that the estimated distortion is correct, since this is calculated for an ideal Laplacian source, while the actual pdf of a subsource, and thus the distortion, depends on the actual classification performed.

The rates considered for the mappings are chosen from what mappings are easy to design. It is clear that mappings with rates larger than two should be considered, but the number of mappings of lower rates has not been optimized.

It would probably be possible to design PCCOVQ mappings that have rates of other simple fractions, such as $\frac{1}{6}$ or $\frac{2}{5}$. As long as the performance of a mapping lies on the convex hull of the operational rate-distortion function, as in Figure 4.4, it will always lower the distortion of the coder for a given rate. This is however a rate not including side information, and the side information will normally increase when new mappings are added. Thus, the number of mappings is a tradeoff between the total rate including side information and the distortion. In this work, no experiments have been performed to find which mappings should be used and which should be left out.

In Section 3.6.2, the applicability of the channel model is discussed, and that discussion also applies here. In order to get the synchronization that is needed, pilot symbols may have to be inserted in the channel signal, which will reduce the performance of the coder somewhat. Pilot symbols might also be needed for traditional coding systems, but not necessarily as frequently.

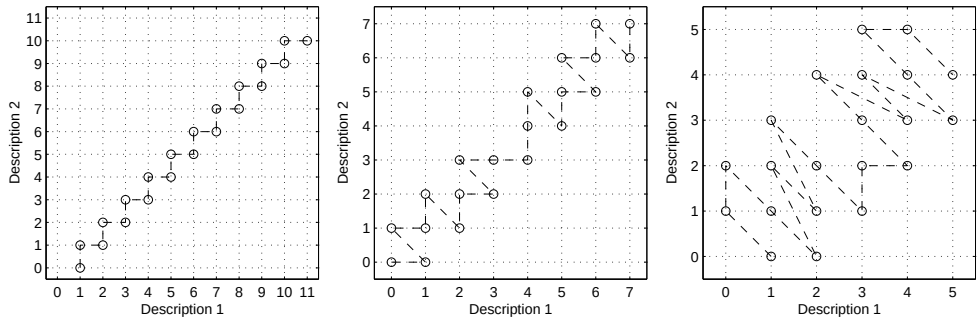
Chapter 5

Conclusions

In this dissertation, two fundamentally different ways of performing joint source–channel coding have been explored, and the conclusions drawn in the two cases are also quite different. For the multiple description coding based system, it was concluded that it did not provide any improvement in most cases. The hybrid scalar quantizer–linear coder system, on the other hand, gives an improvement. It could be argued that Figure 3.26 shows that the performance of the HSQLC is inferior to the combination of a scalar quantizer and an ideal channel coder, but the channel coder must have a very large delay to perform close to the channel capacity, while the HSQLC has practically no delay. The necessary length of a channel code needed to get similar performance as with a separate channel code has not been explored, but since the HSQLC is memoryless, it is clear that the delay will be larger.

The source–channel separation theorem states that separate source and channel coding can be arbitrarily close to the optimum for most channels, which is why the advantages that joint source–channel coding gives in each case should be compared to a realistic alternative with separate source and channel coding. For the HSQLC, advantages are obtained in terms of robustness and delay. For the MD coder, it turned out that in many cases, no advantages were obtained. Note that although joint source–channel coding is not always profitable, a joint optimization of the source and channel code can be important. There is a trade-off between rate spent on source coding, giving lower distortion in the case of error-free transmission, and rate spent on channel coding, reducing the probability of error. On the block fading channel, for example, the source and channel coder were adjusted so that the block error probability was relatively high both with MD and SD coding, cf. Figure 2.26.

The channel code is really the key difference between the two systems. The MD coder is of a nature that makes it suitable for use in combination



(a) 2 diagonals (b) 3 diagonals (c) 5 diagonals

Figure 5.1 Graphic view of the index assignment matrix for 21 levels, as in Table 2.1 for $\delta = \frac{1}{2}$. The circles show the MD index combinations used, and the dashed line shows which points have neighboring SD index values.

with channel coding. However, the results show that when the channel code is introduced, it often protects better alone than in combination with MD coding, except when the channel code is heavily constrained. In the HSQLC (or other direct source–channel mappings such as PCCOVQ), there is no traditional channel code, and instead, the analog nature of the channel is utilized by the coder. In a hybrid digital–analog framework as the HSQLC, it is possible to use a channel code in the digital branch, but the analog branch should be of a nature such that the channel noise is added directly to the signal, thus taking advantage of the unimodal, zero mean nature of a typical noise process (such as Gaussian noise). In the HSQLC, even the quantized symbol utilizes the channel noise properties directly, by making sure that errors with the smallest quadratic value are the most probable.

The utilization of the analog nature of the channel makes the HSQLC fairly robust compared to systems based on good channel codes. With the MD code, Figure 2.17 indicates that the robustness is not increased compared to the underlying channel code. This means that one of the potential advantages of joint source–channel coding is not achieved in this case.

The HSQLC and the multiple description quantizer have in common that they operate by splitting one symbol into two separate ones. Both the cost measure and the channel distortion are different. The cost measure is in terms of power for the HSQLC and entropy for the MD quantizer, and the channel distortion is AWGN for the HSQLC and symbol erasure for the MD quantizer. Still, a resemblance can be seen by comparing the graphic view of the MD index assignment, shown in Figure 5.1, to the plots of the HSQLC in channel

space, shown in Figure 3.11. One difference is that the channel space points go along the diagonal for the MD quantizer, while they go along the second axis, or a line rotated somewhat compared to that, for the HSQLC. This is related to the fact that the MD quantizer is balanced, so the two descriptions contain the same amount of information about the source. This is not the case for the HSQLC, where the quantized symbol is of greater importance than the quantization error symbol.[1] This means that the HSQLC is more related to the special case of MD coding known as multiresolution coding or successive refinement (Equitz and Cover, 1991), where one of the side distortion constraints has been relaxed. The quantized symbol corresponds to the most significant bits in successive refinement, while the quantization error symbol corresponds to the least significant bits. The difference in the shape (other than the angle) between the two systems has to do with the distortion properties. MD coding protects against an erasure of any symbol, while HSQLC protects against noise on both the symbols.

Joint source–channel coding is interesting only for a source on which some distortion can be allowed, and results are perhaps most interesting when evaluated on a real source. The HSQLC has been demonstrated to give good results in an image coder where it was used in combination with other mappings. The case of still image coding may not be as interesting as video coding, since the need for compression and limited delay is more evident there, but video coding usually uses a still image coder to encode difference frames and intra-frame coded frames. The difference frames do usually not require as high rate as still images, so the need for a bandwidth expanding mapping could be smaller. A video coder based on PCCOVQ was proposed in (Fuldseth, 1997; Fuldseth and Ramstad, 1997*b*). This coder does use a rate 2 mapping, and it is likely that replacing this mapping with HSQLC will give an improvement. The need for mappings of higher rates observed in Chapter 4 might be less of a problem for a video coder.

The MD coding system was not applied to any real source, since the results were not good compared to SD coding. The experiments on MD coding were performed with the intention of using an MD video coder (Servetto and Nahrstedt, 2001) for transmission on a radio channel. However, the results with a model source on the binary symmetric and block fading channels, which can be used to model the radio channel under different circumstances, did not encourage this application. If subjective criteria were taken into account, the situation could change, since the variation of the received quality is smaller

[1] The two symbols could be given equal importance by rotating all the points in the channel space prior to transmission, with an inverse rotation at the receiver. A rotation is an orthonormal operation, which does not alter the properties of the white Gaussian noise.

with MD coding, although the average mean square error is usually larger.

Directions for further work have been proposed in the previous chapters. The results suggest further research on direct source–channel mappings to a larger extent than on MD coding used as a joint source–channel coding method. For the mappings, in addition to what has been mentioned in Chapter 3, comparisons to systems using separate source and channel coding with actual codes should be made, in order to find how much longer delay is needed to get the same performance and how this will influence on the robustness. The use of HSQLC for other sources than image coding, such as video or audio coding, should also be considered.

Appendix A

Existence of MDS Codes over Huge Symbol Alphabets

In this appendix, it will be proven for a block code that if the size of the symbol alphabet is large enough compared to the block size, maximum distance separable codes always exist.

The Gilbert bound (Blahut, 1997) for q-ary linear block codes states that for any integers n and $\hat{d}_{\min}$, $2 \leq \hat{d}_{\min} \leq n$, there exists a q-ary $(n, k, d_{\min})$ linear code with $d_{\min} \geq \hat{d}_{\min}$ whose dimension k satisfies

$$\sum_{i=0}^{\hat{d}_{\min}-1} \binom{n}{i} (q-1)^i \geq q^{n-k}. \tag{A.1}$$

If

$$\log_q \left(\sum_{i=0}^{d_{\min}-1} \binom{n}{i} (q-1)^i \right) < d_{\min}, \tag{A.2}$$

then there exists a k such that $d_{\min} > n - k$, and since the $d_{\min} \leq n - k + 1$, there must be a k such that $d_{\min} = n - k + 1$, giving an MDS code.

If the alphabet size is large, $q - 1 > n$, the largest element of the sum in (A.2) will be for $i = d_{\min} - 1$, so

$$\begin{aligned} \log_q \left(\sum_{i=0}^{d_{\min}-1} \binom{n}{i} (q-1)^i \right) &< \log_q \left(d_{\min} \binom{n}{d_{\min}-1} (q-1)^{d_{\min}-1} \right) \\ &< \log_q \left(d_{\min} n^{d_{\min}-1} (q-1)^{d_{\min}-1} \right) \\ &= \log_q d_{\min} + (d_{\min}-1) \log_q (n(q-1)) \\ &< \log_q d_{\min} + (d_{\min}-1)(1 + \log_q n). \end{aligned} \tag{A.3}$$

This means that the existence of an MDS code is assured when

$$\begin{aligned} \log_q d_{\min} + (d_{\min} - 1)(1 + \log_q n) &< d_{\min} \\ 1 + \log_q n &< \frac{d_{\min} - \log_q d_{\min}}{d_{\min} - 1} \\ \log_q n &< \frac{1 - \log_q d_{\min}}{d_{\min} - 1} \\ n &< q^{\frac{1-\log_q d_{\min}}{d_{\min}-1}}. \end{aligned} \tag{A.4}$$

Since $d_{\min} \leq n$ for any code,

$$q^{\frac{1-\log_q d_{\min}}{d_{\min}-1}} \geq q^{\frac{1-\log_q n}{n-1}} = \left(\frac{q}{n}\right)^{\frac{1}{n-1}}, \tag{A.5}$$

so

$$n < \left(\frac{q}{n}\right)^{\frac{1}{n-1}} \quad \Leftrightarrow \quad q > n^n \tag{A.6}$$

is a sufficient condition for MDS codes to exist. The requirement that $q-1 > n$ is always fulfilled if (A.6) holds. The Gilbert bound is valid for $d_{\min} \geq 2$, but $d_{\min} = 1$ corresponds to a code where $n = k$ and no erasures can be corrected, which is obtained by an identity code, so an MDS code can be assumed even then. This result tells us that for any block size, any alphabet size over a certain value ensures the existence of an MDS code. The alphabet size has to be huge compared to the codeword length, but $q = 2^\nu$ where ν is the number of bits per packet, and typical IP packets consist of hundreds of bits or more, while delay constraints prohibit the use of long blocks. In addition, the bound obtained in (A.6) is quite conservative. Thus, when working with erasure channels, MDS codes will be assumed.

In Section 2.5.3, block lengths up to 20 for MD coding and 40 for SD coding are considered. These block lengths do not violate (A.6) for typical packet networks, since $\log_2(40^{40}) < 213$, meaning that if the symbols are packets of 213 bits or more, a maximum distance code exists. This is less than the number of bits in a typical internet packet.

Appendix B

Derivations Concerning HSQLC

In this appendix, formulas for the performance of the HSQLC needed in Chapter 3 are derived. Section B.1 treats optimal receivers, in Section B.2, expressions for a MAP receiver for the quantized symbol are derived, and in Section B.3, expressions for the MSE of the system assuming simplified receivers are found.

B.1 Derivation of Optimal Receivers

Here, the general formula for the optimal receivers of the HSQLC system will be derived.

Equation (3.22) shows that $E[A(k) \mid \hat{y}_1(k), \hat{y}_2(k)]$ and $E[B(k) \mid \hat{y}_1(k), \hat{y}_2(k)]$ must be found. Since there is no memory in the system, and thus no dependence on the time index k, it will be dropped for the derivation here. From the definition of conditional expectation,

$$E[A \mid \hat{y}_1, \hat{y}_2] = \sum_{i=0}^{M-1} a_i p_{A|\hat{Y}_1,\hat{Y}_2}(a_i \mid \hat{y}_1, \hat{y}_2). \tag{B.1}$$

Here, $p_{A|\hat{Y}_1,\hat{Y}_2}(a_i \mid \hat{y}_1, \hat{y}_2)$ is the conditional probability that $A = a_i$ given $\hat{y}_1$ and $\hat{y}_2$. Now, define $f_{A,\hat{Y}_1,\hat{Y}_2}(a, \hat{y}_1, \hat{y}_2)$ as the joint probability distribution of A, $\hat{Y}_1$, and $\hat{Y}_2$. Note that this function is discrete in a and continuous in $\hat{y}_1$ and $\hat{y}_2$, so that integrating over $\hat{y}_1$ and $\hat{y}_2$ gives the discrete probability mass function of A, while summing over all a_i, $i \in \{0, \ldots, M-1\}$ gives the joint probability *density* function of $\hat{Y}_1$ and $\hat{Y}_2$. From the definition of conditional

probability (Walpole and Myers, 1993),

$$p_{A|\hat{Y}_1,\hat{Y}_2}(a_i \mid \hat{y}_1, \hat{y}_2) = \frac{f_{A,\hat{Y}_1,\hat{Y}_2}(a_i, \hat{y}_1, \hat{y}_2)}{\sum_{i=0}^{M-1} f_{A,\hat{Y}_1,\hat{Y}_2}(a_i, \hat{y}_1, \hat{y}_2)}. \tag{B.2}$$

In order to find $f_{A,\hat{Y}_1,\hat{Y}_2}$, look at the joint probability of both branches at both the transmitter and receiver end, $f_{A,B,\hat{Y}_1,\hat{Y}_2}$. From this,

$$f_{A,\hat{Y}_1,\hat{Y}_2}(a_i, \hat{y}_1, \hat{y}_2) = \int_{-\infty}^{\infty} f_{A,B,\hat{Y}_1,\hat{Y}_2}(a_i, b, \hat{y}_1, \hat{y}_2)\, db. \tag{B.3}$$

In order to find the joint probability distribution $f_{A,B,\hat{Y}_1,\hat{Y}_2}$, consider the following.

$$f_{A,B,\hat{Y}_1,\hat{Y}_2}(a_i, b, \hat{y}_1, \hat{y}_2) = f_{A,B}(a_i, b) f_{\hat{Y}_1,\hat{Y}_2|A,B}(\hat{y}_1, \hat{y}_2 \mid a_i, b) \tag{B.4}$$

where $f_{A,B}$ is the joint probability density function of A and B, and $f_{\hat{Y}_1,\hat{Y}_2|A,B}$ is the conditional probability density on $\hat{Y}_1$ and $\hat{Y}_2$ given A and B. Since N_1 and N_2 are statistically independent of the input and of each other, one can write

$$f_{N_1,N_2|A,B}(n_1, n_2 \mid a_i, b) = f_N(n_1) f_N(n_2), \tag{B.5}$$

where f_N is the probability density function (pdf) of the identically distributed noise symbols N_1 and N_2. By noting that $\hat{Y}_1$ and $\hat{Y}_2$ are connected to the noise by

$$\hat{Y}_1 = c_i + N_1 \tag{B.6}$$

$$\hat{Y}_2 = K_b B + N_2, \tag{B.7}$$

one can change variables from $\hat{Y}_1$ and $\hat{Y}_2$ to N_1 and N_2 to get

$$\begin{aligned} f_{\hat{Y}_1,\hat{Y}_2|A,B}(\hat{y}_1, \hat{y}_2 \mid a_i, b) &= f_{N_1,N_2|A,B}(\hat{y}_1 - c_i, \hat{y}_2 - K_b b \mid a_i, b) \\ &= f_N(\hat{y}_1 - c_i) f_N(\hat{y}_2 - K_b b). \end{aligned} \tag{B.8}$$

Thus,

$$f_{A,B,\hat{Y}_1,\hat{Y}_2}(a_i, b, \hat{y}_1, \hat{y}_2) = f_{A,B}(a_i, b) f_N(\hat{y}_1 - c_i) f_N(\hat{y}_2 - K_b b). \tag{B.9}$$

Then, to find $f_{A,B}$, assume that the input symbol X is in the interval $\mathcal{D}_i = (d_i, d_{i+1}]$. Then $A = a_i$, $B = X - a_i$, and the probability distribution of B is

$$f_{B|X\in\mathcal{D}_i}(b \mid x \in \mathcal{D}_i) = \begin{cases} \dfrac{f_X(b + a_i)}{\Pr(X \in \mathcal{D}_i)} & \text{when } b \in (d_i - a_i, d_{i+1} - a_i], \\ 0 & \text{otherwise,} \end{cases} \tag{B.10}$$

where f_X is the pdf of the source symbol X. Since $A = a_i$ when $X \in \mathcal{D}_i$,

$$\begin{aligned} f_{A,B}(a_i, b) &= f_{B|A}(b \mid a_i) \Pr(A = a_i) \\ &= f_{B|X\in\mathcal{D}_i}(b \mid x \in \mathcal{D}_i) \Pr(X \in \mathcal{D}_i) \\ &= \begin{cases} f_X(b + a_i) & \text{when } b \in (d_i - a_i, d_{i+1} - a_i], \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{B.11}$$

Inserting (B.9) and (B.11) into (B.3) gives

$$\begin{aligned} f_{A,\hat{Y}_1,\hat{Y}_2}(a_i, \hat{y}_1, \hat{y}_2) &= \int_{d_i - a_i}^{d_{i+1}-a_i} f_X(b + a_i) f_N(\hat{y}_1 - c_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1}-a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db, \end{aligned} \tag{B.12}$$

and from (B.1), (B.2), and (B.12),

$$E[A \mid \hat{y}_1, \hat{y}_2] = \frac{\displaystyle\sum_{i=0}^{M-1} a_i f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1}-a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db}{\displaystyle\sum_{i=0}^{M-1} f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1}-a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db} \tag{B.13}$$

as stated in (3.24).

To find the other conditional expectation, $E[B(k) \mid \hat{y}_1(k), \hat{y}_2(k)]$, again state the definitions

$$E[B \mid \hat{y}_1, \hat{y}_2] = \int_{-\infty}^{\infty} b f_{B|\hat{Y}_1,\hat{Y}_2}(b \mid \hat{y}_1, \hat{y}_2)\, db \tag{B.14}$$

and

$$f_{B|\hat{Y}_1,\hat{Y}_2}(b \mid \hat{y}_1, \hat{y}_2) = \frac{f_{B,\hat{Y}_1,\hat{Y}_2}(b, \hat{y}_1, \hat{y}_2)}{\displaystyle\int_{-\infty}^{\infty} f_{B,\hat{Y}_1,\hat{Y}_2}(b, \hat{y}_1, \hat{y}_2)\, db}, \tag{B.15}$$

where $f_{B|\hat{Y}_1,\hat{Y}_2}(b \mid \hat{y}_1, \hat{y}_2)$ is the conditional pdf that $B = b$ given $\hat{y}_1$ and $\hat{y}_2$, and $f_{B,\hat{Y}_1,\hat{Y}_2}$ is the joint probability density of these variables. This joint pdf can be found by summing $f_{A,B,\hat{Y}_1,\hat{Y}_2}(a_i, b, \hat{y}_1, \hat{y}_2)$ over i, so from (B.9) and (B.11),

$$\begin{aligned} f_{B,\hat{Y}_1,\hat{Y}_2}(b, \hat{y}_1, \hat{y}_2) &= \sum_{i=0}^{M-1} f_{A,B,\hat{Y}_1,\hat{Y}_2}(a_i, b, \hat{y}_1, \hat{y}_2) \\ &= \sum_{i=0}^{M-1} f_X(b + a_i) f_N(\hat{y}_1 - c_i) f_N(\hat{y}_2 - K_b b) u(b - (d_i - a_i)) u((d_{i+1} - a_i) - b), \end{aligned} \tag{B.16}$$

where $u(\cdot)$ is the unit step function,

$$u(t) = \begin{cases} 0 & \text{when } t < 0 \\ 1 & \text{when } t > 0. \end{cases} \tag{B.17}$$

This gives

$$E[B \mid \hat{y}_1, \hat{y}_2] = \frac{\sum_{i=0}^{M-1} f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1} - a_i} b f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db}{\sum_{i=0}^{M-1} f_N(\hat{y}_1 - c_i) \int_{d_i - a_i}^{d_{i+1} - a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db}, \tag{B.18}$$

as stated in (3.25).

B.1.1 Optimal Receivers for a Gaussian Source

In order to find analytic expressions for (B.13) and (B.18), the probability density functions $f_N(n)$ and $f_X(x)$ must be known. The noise is Gaussian with zero mean and variance σ_N^2, so

$$f_N(n) = \frac{1}{\sqrt{2\pi}\sigma_N} e^{-\frac{n^2}{2\sigma_N^2}}, \tag{B.19}$$

and for a Gaussian zero mean source of variance σ_X^2,

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{x^2}{2\sigma_X^2}}. \tag{B.20}$$

Two integrals need to be solved. First, look at

$$\begin{aligned} I_i^{(0)} &\stackrel{\text{def}}{=} \int_{d_i - a_i}^{d_{i+1} - a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= \frac{1}{2\pi\sigma_X\sigma_N} \int_{d_i - a_i}^{d_{i+1} - a_i} e^{-\frac{(b+a_i)^2}{2\sigma_X^2} - \frac{(\hat{y}_2 - K_b b)^2}{2\sigma_N^2}}\, db \\ &= \frac{1}{2\pi\sigma_X\sigma_N} \int_{d_i - a_i}^{d_{i+1} - a_i} e^{-\frac{1}{2}\left(\frac{1}{\sigma_X^2} + \frac{K_b^2}{\sigma_N^2}\right) b^2 - \left(\frac{1}{\sigma_X^2} a_i - \frac{K_b}{\sigma_N^2}\hat{y}_2\right) b - \frac{1}{2}\left(\frac{1}{\sigma_X^2} a_i^2 + \frac{1}{\sigma_N^2}\hat{y}_2^2\right)}\, db. \end{aligned} \tag{B.21}$$

This integral can be solved by completing the square in the exponent and applying the definition of the error function,

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt, \tag{B.22}$$

or by using a symbolic mathematics program, such as Maple. The solution is

$$I_i^{(0)} = \frac{1}{2\sqrt{2\pi}\sqrt{\sigma_N^2 + K_b^2\sigma_X^2}}\, e^{-\frac{1}{2}\frac{(\hat{y}_2 + a_i K_b)^2}{\sigma_N^2 + K_b^2\sigma_X^2}}$$
$$\cdot \left[\operatorname{erf}\left(\frac{1}{\sqrt{2}} \left(\frac{\sqrt{\sigma_N^2 + K_b^2\sigma_X^2}}{\sigma_X\sigma_N}(d_{i+1} - a_i) + \frac{\sigma_N^2 a_i - K_b\sigma_X^2\hat{y}_2}{\sigma_X\sigma_N\sqrt{\sigma_N^2 + K_b^2\sigma_X^2}} \right) \right) \right.$$
$$\left. - \operatorname{erf}\left(\frac{1}{\sqrt{2}} \left(\frac{\sqrt{\sigma_N^2 + K_b^2\sigma_X^2}}{\sigma_X\sigma_N}(d_i - a_i) + \frac{\sigma_N^2 a_i - K_b\sigma_X^2\hat{y}_2}{\sigma_X\sigma_N\sqrt{\sigma_N^2 + K_b^2\sigma_X^2}} \right) \right) \right]. \quad \text{(B.23)}$$

Then look at,

$$I_i^{(1)} \stackrel{\text{def}}{=} \int_{d_i - a_i}^{d_{i+1} - a_i} b f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db$$
$$= \frac{1}{2\pi\sigma_X\sigma_N} \int_{d_i - a_i}^{d_{i+1} - a_i} b e^{-\frac{(b+a_i)^2}{2\sigma_X^2} - \frac{(\hat{y}_2 - K_b b)^2}{2\sigma_N^2}}\, db$$
$$= \frac{1}{2\pi\sigma_X\sigma_N} \int_{d_i - a_i}^{d_{i+1} - a_i} b e^{-\frac{1}{2}\left(\frac{1}{\sigma_X^2} + \frac{K_b^2}{\sigma_N^2}\right)b^2 - \left(\frac{1}{\sigma_X^2}a_i - \frac{K_b}{\sigma_N^2}\hat{y}_2\right)b - \frac{1}{2}\left(\frac{1}{\sigma_X^2}a_i^2 + \frac{1}{\sigma_N^2}\hat{y}_2^2\right)}\, db. \quad \text{(B.24)}$$

This integral can also be solved by completing the square in the exponent and changing the variable of integration, or by using a symbolic mathematics program. This gives

$$I_i^{(1)} = \frac{1}{2\pi}\frac{\sigma_N\sigma_X}{\sigma_N^2 + K_b^2\sigma_X^2}$$
$$\cdot \left[-e^{-\frac{1}{2\sigma_N^2}(K_b(d_{i+1} - a_i) - \hat{y}_2)^2 - \frac{1}{2\sigma_X^2}d_{i+1}^2} + e^{-\frac{1}{2\sigma_N^2}(K_b(d_i - a_i) - \hat{y}_2)^2 - \frac{1}{2\sigma_X^2}d_i^2} \right]$$
$$- \frac{a_i\sigma_N^2 - K_b\sigma_X^2\hat{y}_2}{\sigma_N^2 + K_b^2\sigma_X^2} I_i^{(0)} \quad \text{(B.25)}$$

Now, analytical expressions for (B.13) and (B.18) are found by inserting Equations (B.19), (B.23) and (B.25).

B.1.2 Optimal Receivers for a Laplacian Source

A zero mean Laplacian source has the source pdf

$$f_X(x) = \frac{1}{\sqrt{2}\sigma_X} e^{-\frac{\sqrt{2}|x|}{\sigma_X}}, \quad \text{(B.26)}$$

where σ_X^2 is the variance. The noise is still Gaussian, so Equation (B.19) applies. Due to the absolute value in the source pdf, three different cases have to be treated. First, consider the case $d_i < d_{i+1} \leq 0$. Then,

$$\begin{aligned} I_i^{(0-)} &= \int_{d_i-a_i}^{d_{i+1}-a_i} f_X(b+a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= \frac{1}{2\sqrt{\pi}\sigma_X\sigma_N} \int_{d_i-a_i}^{d_{i+1}-a_i} e^{-\frac{(\hat{y}_2 - K_b b)^2}{2\sigma_N^2} + \frac{\sqrt{2}(b+a_i)}{\sigma_X}}\, db \\ &= \frac{1}{2\sqrt{\pi}\sigma_X\sigma_N} \int_{d_i-a_i}^{d_{i+1}-a_i} e^{-\frac{K_b^2}{2\sigma_N^2}b^2 + \left(\frac{\sqrt{2}}{\sigma_X} + \frac{K_b}{\sigma_N^2}\hat{y}_2\right)b + \left(\frac{\sqrt{2}}{\sigma_X}a_i - \frac{1}{2\sigma_N^2}\hat{y}_2^2\right)}\, db. \end{aligned} \tag{B.27}$$

As in the Gaussian case, the integral can be solved by completing the square in the exponent, or by a symbolic mathematics program. This gives

$$\begin{aligned} I_i^{(0-)} = \frac{1}{2\sqrt{2}K_b\sigma_X} e^{\frac{\sqrt{2}K_b^2\sigma_X a_i + \sqrt{2}K_b\sigma_X\hat{y}_2 + \sigma_N^2}{K_b^2\sigma_X^2}} \\ \cdot \left[\operatorname{erf}\left(\frac{K_b}{\sqrt{2}\sigma_N}(d_{i+a} - a_i) - \frac{\hat{y}_2}{\sqrt{2}\sigma_N} - \frac{\sigma_N}{K_b\sigma_X}\right)\right. \\ \left. - \operatorname{erf}\left(\frac{K_b}{\sqrt{2}\sigma_N}(d_i - a_i) - \frac{\hat{y}_2}{\sqrt{2}\sigma_N} - \frac{\sigma_N}{K_b\sigma_X}\right)\right]. \end{aligned} \tag{B.28}$$

Similarly, for $0 \leq d_i < d_{i+1}$,

$$\begin{aligned} I_i^{(0+)} &= \int_{d_i-a_i}^{d_{i+1}-a_i} f_X(b+a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= \frac{1}{2\sqrt{2}K_b\sigma_X} e^{\frac{-\sqrt{2}K_b^2\sigma_X a_i - \sqrt{2}K_b\sigma_X\hat{y}_2 + \sigma_N^2}{K_b^2\sigma_X^2}} \\ &\quad \cdot \left[\operatorname{erf}\left(\frac{K_b}{\sqrt{2}\sigma_N}(d_{i+a} - a_i) - \frac{\hat{y}_2}{\sqrt{2}\sigma_N} + \frac{\sigma_N}{K_b\sigma_X}\right)\right. \\ &\quad \left. - \operatorname{erf}\left(\frac{K_b}{\sqrt{2}\sigma_N}(d_i - a_i) - \frac{\hat{y}_2}{\sqrt{2}\sigma_N} + \frac{2\sigma_N}{K_b\sigma_X}\right)\right]. \end{aligned} \tag{B.29}$$

For $d_i < 0 < d_{i+1}$, the integral can be divided, giving

$$\begin{aligned} I_i^{(0\mp)} &= \int_{d_i-a_i}^{d_{i+1}-a_i} f_X(b+a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= \int_{d_i-a_i}^{-a_i} f_X(b+a_i) f_N(\hat{y}_2 - K_b b)\, db + \int_{-a_i}^{d_{i+1}-a_i} f_X(b+a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= I_i^{(0-)}\Big|_{d_{i+1}=0} + I_i^{(0+)}\Big|_{d_i=0}. \end{aligned} \tag{B.30}$$

The other integral also has to be treated for three different cases. For $d_i < d_{i+1} \leq 0$,

$$\begin{aligned} I_i^{(1-)} &= \int_{d_i - a_i}^{d_{i+1}-a_i} b f_X(b+a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= \frac{\sigma_N}{2\sqrt{\pi} K_b^2 \sigma_X} \\ &\cdot \left[-e^{-\frac{1}{2\sigma_N^2}(K_b(d_{i+1}-a_i)-\hat{y}_2)^2 + \frac{\sqrt{2}}{\sigma_X^2} d_{i+1}} + e^{-\frac{1}{2\sigma_N^2}(K_b(d_i-a_i)-\hat{y}_2)^2 + \frac{\sqrt{2}}{\sigma_X^2} d_i} \right] \\ &- \frac{\hat{y}_2 K_b \sigma_X + \sqrt{2}\sigma_N^2}{K_b^2 \sigma_X} I_i^{(0-)}, \end{aligned} \tag{B.31}$$

using the same method for solving the integral as for the Gaussian source. For $0 \leq d_i < d_{i+1}$,

$$\begin{aligned} I_i^{(1+)} &= \int_{d_i - a_i}^{d_{i+1}-a_i} b f_X(b+a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= \frac{\sigma_N}{2\sqrt{\pi} K_b^2 \sigma_X} \\ &\cdot \left[-e^{-\frac{1}{2\sigma_N^2}(K_b(d_{i+1}-a_i)-\hat{y}_2)^2 - \frac{\sqrt{2}}{\sigma_X^2} d_{i+1}} + e^{-\frac{1}{2\sigma_N^2}(K_b(d_i-a_i)-\hat{y}_2)^2 - \frac{\sqrt{2}}{\sigma_X^2} d_i} \right] \\ &- \frac{\hat{y}_2 K_b \sigma_X - \sqrt{2}\sigma_N^2}{K_b^2 \sigma_X} I_i^{(0+)}, \end{aligned} \tag{B.32}$$

For $d_i < 0 < d_{i+1}$, again the integral can be divided,

$$\begin{aligned} I_i^{(1\mp)} &= \int_{d_i - a_i}^{d_{i+1}-a_i} b f_X(b+a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= I_i^{(1-)}\Big|_{d_{i+1}=0} + I_i^{(1+)}\Big|_{d_i=0}. \end{aligned} \tag{B.33}$$

B.1.3 Optimal Receivers for a Uniform Source

For a uniform source, the source pdf is given by

$$f_X(x) = \begin{cases} \frac{1}{\Delta} & \text{when } -\frac{\Delta}{2} \leq x \leq \frac{\Delta}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{B.34}$$

The mean of the source symbols is then zero, and the variance is $\sigma_X^2 = \frac{\Delta^2}{12}$. For this distribution, it will be assumed that the endpoints of the quantizer are

given by $d_0 = -\frac{\Delta}{2}$ and $d_M = \frac{\Delta}{2}$. The noise pdf is still given by Equation (B.19). Now,

$$\begin{aligned} I_i^{(0)} &= \int_{d_i - a_i}^{d_{i+1} - a_i} f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= \frac{1}{\Delta\sqrt{2\pi}\sigma_N} \int_{d_i - a_i}^{d_{i+1} - a_i} e^{-\frac{(\hat{y}_2 - K_b b)^2}{2\sigma_N^2}}\, db \\ &= \frac{1}{2\Delta K_b} \left[\operatorname{erf}\left(\frac{K_b(d_{i+1} - a_i) - \hat{y}_2}{\sqrt{2}\sigma_N} \right) - \operatorname{erf}\left(\frac{K_b(d_i - a_i) - \hat{y}_2}{\sqrt{2}\sigma_N} \right) \right], \end{aligned} \tag{B.35}$$

and

$$\begin{aligned} I_i^{(1)} &= \int_{d_i - a_i}^{d_{i+1} - a_i} b f_X(b + a_i) f_N(\hat{y}_2 - K_b b)\, db \\ &= \frac{1}{\Delta\sqrt{2\pi}\sigma_N} \int_{d_i - a_i}^{d_{i+1} - a_i} b e^{-\frac{(\hat{y}_2 - K_b b)^2}{2\sigma_N^2}}\, db \\ &= \frac{\sigma_N}{\Delta\sqrt{2\pi}K_b^2} \left[-e^{-\frac{(K_b(d_{i+1} - a_i) - \hat{y}_2)^2}{2\sigma_N^2}} + e^{-\frac{(K_b(d_i - a_i) - \hat{y}_2)^2}{2\sigma_N^2}} \right] + \frac{\hat{y}_2}{K_b} I_i^{(0)}. \end{aligned} \tag{B.36}$$

For the special case of a uniform quantizer with midpoint representation values and a quantizer step size that divides the width Δ of the uniform source distribution, the expressions of Equations (B.13) and (B.18) can be simplified. Assume that the quantizer has M quantization intervals of width $\delta = \Delta/M$. This means that $d_i = (i - M/2)\delta$, and that $a_i = (i - (M-1)/2)\delta$, as midpoint representation values are used. Now, $d_{i+1} - a_i = \delta/2$, and $d_i - a_i = -\delta/2$, so

$$I_i^{(0)} = \frac{1}{2\Delta K_b} \left[\operatorname{erf}\left(\frac{\frac{K_b\delta}{2} - \hat{y}_2}{\sqrt{2}\sigma_N} \right) - \operatorname{erf}\left(\frac{-\frac{K_b\delta}{2} - \hat{y}_2}{\sqrt{2}\sigma_N} \right) \right] \tag{B.37}$$

and

$$I_i^{(1)} = \frac{\sigma_N}{\Delta\sqrt{2\pi}K_b^2} \left[-e^{-\frac{\left(\frac{K_b\delta}{2} - \hat{y}_2\right)^2}{2\sigma_N^2}} + e^{-\frac{\left(-\frac{K_b\delta}{2} - \hat{y}_2\right)^2}{2\sigma_N^2}} \right] + \frac{\hat{y}_2}{K_b} I_i^{(0)}. \tag{B.38}$$

Note that in this case, $I_i^{(0)}$ and $I_i^{(1)}$ are independent of i, and can thus be

written $I^{(0)}$ and $I^{(1)}$. Inserting this in Equations (B.13) and (B.18) gives:

$$\beta_1(\hat{y}_1,\hat{y}_2) = E[A \mid \hat{y}_1,\hat{y}_2] = \frac{\sum_{i=0}^{M-1} a_i \frac{1}{\sqrt{2\pi}\sigma_N} e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N^2}} I^{(0)}}{\sum_{i=0}^{M-1} \frac{1}{\sqrt{2\pi}\sigma_N} e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N^2}} I^{(0)}} = \frac{\sum_{i=0}^{M-1} a_i e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N^2}}}{\sum_{i=0}^{M-1} e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N^2}}} \tag{B.39}$$

and

$$\begin{aligned}\beta_2(\hat{y}_1,\hat{y}_2) = E[B \mid \hat{y}_1,\hat{y}_2] &= \frac{\sum_{i=0}^{M-1} \frac{1}{\sqrt{2\pi}\sigma_N} e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N^2}} I^{(1)}}{\sum_{i=0}^{M-1} \frac{1}{\sqrt{2\pi}\sigma_N} e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N^2}} I^{(0)}} = \frac{I^{(1)}}{I^{(0)}} \\ &= \sqrt{\frac{2}{\pi}}\frac{\sigma_N}{K_b}\frac{-e^{-\frac{\left(\frac{K_b\delta}{2}-\hat{y}_2\right)^2}{2\sigma_N^2}} + e^{-\frac{\left(-\frac{K_b\delta}{2}-\hat{y}_2\right)^2}{2\sigma_N^2}}}{\operatorname{erf}\left(\frac{\frac{K_b\delta}{2}-\hat{y}_2}{\sqrt{2}\sigma_N}\right) - \operatorname{erf}\left(\frac{-\frac{K_b\delta}{2}-\hat{y}_2}{\sqrt{2}\sigma_N}\right)} + \frac{\hat{y}_2}{K_b}.\end{aligned} \tag{B.40}$$

Note that in this case, β_1 is independent of $\hat{y}_2$ and β_2 is independent of $\hat{y}_1$, just as for the simplified receivers.

B.2 Derivation of MAP Receiver

A way of finding the the receiver decision intervals $\mathcal{E}_i$ of the simplified receiver is the use of a maximum a posteriori (MAP) receiver for the quantized symbols. A MAP receiver is designed so that it finds the most probable symbol to have been transmitted given the received symbol. Strictly speaking, this is not the same as minimizing the MSE. However, if the quantized symbol is correctly received, this gives no contribution to the squared error, as opposed to an erroneous reception, so maximizing the probability of correct reception also gives a low MSE, and is unlikely to give any significant difference from the optimal detection.

The MAP receiver outputs the i that maximizes the conditional probability

$$p_{A|\hat{Y}}(a_i \mid \hat{y}_1) = \frac{f_{\hat{Y}_1|A}(\hat{y}_1 \mid a_i)p_A(a_i)}{f_{\hat{Y}_1}(\hat{y}_1)}. \tag{B.41}$$

Since $\hat{Y}_1$ is a result of adding noise to the transmitted signal Y_1, it is clear that

$$f_{\hat{Y}_1|A}(\hat{y}_1 \mid a_i) = f_N(\hat{y}_1 - c_i). \tag{B.42}$$

The probability of $A = a_i$ is

$$p_A(a_i) = \int_{d_i}^{d_{i+1}} f_X(x)\, dx. \tag{B.43}$$

Since $f_{\hat{Y}_1}(\hat{y}_1)$ is constant with respect to i, it can be kept out of the optimization. The MAP index is then found by

$$i_{\text{MAP}} = \underset{i\in\{0,\dots,M-1\}}{\text{argmax}} f_{\hat{Y}_1|A}(\hat{y}_1 \mid a_i)p_A(a_i) = \underset{i\in\{0,\dots,M-1\}}{\text{argmax}} f_N(\hat{y}_1 - c_i)p_A(a_i). \tag{B.44}$$

The noise is zero mean Gaussian, so by inserting (B.19),

$$i_{\text{MAP}} = \underset{i\in\{0,\dots,M-1\}}{\text{argmax}} \frac{1}{\sqrt{2\pi}\sigma_N^2} e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N}} p_A(a_i) = \underset{i\in\{0,\dots,M-1\}}{\text{argmax}} e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N^2}} p_A(a_i). \tag{B.45}$$

The decision limits can now be found from the points where the objective functions given neighboring i values are equal. Thus, e_i is given by the $\hat{y}_1$ that solves the equation

$$e^{-\frac{(\hat{y}_1-c_{i-1})^2}{2\sigma_N^2}} p_A(a_{i-1}) = e^{-\frac{(\hat{y}_1-c_i)^2}{2\sigma_N^2}} p_A(a_i), \tag{B.46}$$

which gives

$$e_i = \tfrac{1}{2}(c_i + c_{i-1}) - \frac{\sigma_N^2}{c_i - c_{i-1}} \ln \frac{p_A(a_i)}{p_A(a_{i-1})} \tag{B.47}$$

for $i \in \{1, \dots, M-1\}$, $e_0 = -\infty$, and $e_M = \infty$. This requires that all the transmitted values have an interval for which it maximizes the conditional probability. If not, there will be an i where $e_{i-1} > e_i$ when calculated by (B.47). Then, the c_i for the value of i that gives the problem must be removed and the decision levels recalculated. In practice, this is not a problem.

The integral in (B.43) is found for different probability distributions in Section B.3.

In the beginning of the section, it was claimed that a MAP receiver is not equivalent to the decision limits that minimize the MSE, but that the difference should be small. This can be verified from Figure 3.15, where the HSQLC with simplified receiver designed for the actual CSNR uses the MAP receiver, while the receiver designed for the design CSNR used the decision limits that minimize the MSE. At the design CSNR, there is no significant difference between the results.

B.3 Derivation of Mean Square Error

The MSE of the HSQLC system is given by Equation (3.23). When the simplified receivers are used, $\beta_1(\hat{y}_1, \hat{y}_2)$ depends only on $\hat{y}_1$ and $\beta_2(\hat{y}_1, \hat{y}_2)$ depends only on $\hat{y}_2$, and they can be written $\beta_1(\hat{y}_1, \hat{y}_2) = \tilde{\beta}_1(\hat{y}_1)$ and $\beta_2(\hat{y}_1, \hat{y}_2) = \tilde{\beta}_2(\hat{y}_2)$. This gives

$$\begin{aligned} D &= E[(X - \hat{X})^2] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[x - \tilde{\beta}_1(C_x(x) + n_1) - \tilde{\beta}_2(K_b(x - Q(x)) + n_2)]^2 \\ &\qquad f_N(n_1)f_N(n_2)f_X(x)\,dn_1\,dn_2\,dx. \end{aligned} \tag{B.48}$$

The functions C_x and Q return $C_x(x) = c_i$ and $Q(x) = a_i$ when $x \in (d_i, d_{i+1}]$. By inserting this in (B.48),

$$\begin{aligned} D = \sum_{i=0}^{M-1}\int_{d_i}^{d_{i+1}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}&[x - \tilde{\beta}_1(c_i + n_1) - \tilde{\beta}_2(K_b(x - a_i) + n_2)]^2 \\ &f_N(n_1)f_N(n_2)f_X(x)\,dn_1\,dn_2\,dx. \end{aligned} \tag{B.49}$$

Furthermore, $\tilde{\beta}_1(c_i + n_1) = a_j$ when $c_i + n_1 \in (e_j, e_{j+1}]$. This means that

$$\begin{aligned} D = \sum_{i=0}^{M-1}\int_{d_i}^{d_{i+1}}\int_{-\infty}^{\infty}\sum_{j=0}^{M-1}\int_{e_j - c_i}^{e_{j+1}-c_i}&[x - a_j - \tilde{\beta}_2(K_b(x - a_i) + n_2)]^2 \\ &f_N(n_1)f_N(n_2)f_X(x)\,dn_1\,dn_2\,dx. \end{aligned} \tag{B.50}$$

Then, apply $\tilde{\beta}_2(\hat{y}_2) = \frac{\kappa}{K_b}\hat{y}_2$ and rearrange to get

$$\begin{aligned} D = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1}\int_{d_i}^{d_{i+1}}\int_{-\infty}^{\infty}&\left[(1-\kappa)x + (\kappa a_i - a_j) - \frac{\kappa n_2}{K_b}\right]^2 \\ &f_N(n_2)f_X(x)\,dn_2\,dx\int_{e_j - c_i}^{e_{j+1}-c_i} f_N(n_1)\,dn_1. \end{aligned} \tag{B.51}$$

Now, introduce the probability $p_{i,j}$ of receiving a_j given that a_i was transmitted. This happens if the noise N_1 is in the interval $(e_i - c_j, e_{i+1} - c_j]$, so

$$p_{i,j} = \int_{e_j - c_i}^{e_{j+1}-c_i} f_N(n_1)\,dn_1. \tag{B.52}$$

This gives

$$
\begin{aligned}
D = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} p_{i,j} \int_{d_i}^{d_{i+1}} \int_{-\infty}^{\infty} \Bigg[& (1-\kappa)^2 x^2 + 2(1-\kappa)(\kappa a_i - a_j)x + (\kappa a_i - a_j)^2 \\
& - \frac{2\kappa}{K_b}(1-\kappa) n_2 x - \frac{2\kappa}{K_b}(\kappa a_i - a_j) n_2 + \frac{\kappa^2 n_2^2}{K_b^2} \Bigg] f_N(n_2) f_X(x)\, dn_2\, dx \\
= \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} p_{i,j} \Bigg[& \int_{d_i}^{d_{i+1}} \int_{-\infty}^{\infty} f_N(n_2)\, dn_2 \\
& \cdot \left((1-\kappa)^2 x^2 + 2(1-\kappa)(\kappa a_i - a_j)x + (\kappa a_i - a_j)^2 \right) f_X(x)\, dx \\
& - \frac{2\kappa}{K_b} \int_{d_i}^{d_{i+1}} \int_{-\infty}^{\infty} n_2 f_N(n_2)\, dn_2 ((1-\kappa)x - \kappa a_i + a_j) f_X(x)\, dx \\
& + \frac{\kappa^2}{K_b^2} \int_{d_i}^{d_{i+1}} \int_{-\infty}^{\infty} n_2^2 f_N(n_2)\, dn_2\, f_X(x)\, dx \Bigg] \\
= \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} p_{i,j} \Bigg[& \int_{d_i}^{d_{i+1}} \left((1-\kappa)^2 x^2 + 2(1-\kappa)(\kappa a_i - a_j)x \right. \\
& \left. + (\kappa a_i - a_j)^2 \right) f_X(x)\, dx + \frac{\kappa^2 \sigma_N^2}{K_b^2} \int_{d_i}^{d_{i+1}} f_X(x)\, dx \Bigg],
\end{aligned}
\tag{B.53}
$$

where it has been used that f_N is a pdf of a variable with zero mean and variance σ_N^2. A rearrangement gives

$$
\begin{aligned}
D = \sum_{i=0}^{M-1}\sum_{j=0}^{M-1} p_{i,j} \Bigg[& (1-\kappa)^2 \int_{d_i}^{d_{i+1}} x^2 f_X(x)\, dx \\
& + 2(1-\kappa)(\kappa a_i - a_j) \int_{d_i}^{d_{i+1}} x f_X(x)\, dx \\
& + \left((\kappa a_i - a_j)^2 + \frac{\kappa^2 \sigma_N^2}{K_b^2} \right) \int_{d_i}^{d_{i+1}} f_X(x)\, dx \Bigg],
\end{aligned}
\tag{B.54}
$$

which is the same as Equation (3.30).

The constant κ is given by Equation (3.28), which shows that the variance of B, σ_B^2 must be found. The distribution $f_B(b)$ can be found as

$$
f_B(b) = \sum_{i=0}^{M-1} f_{A,B}(a_i, b), \tag{B.55}
$$

and $f_{A,B}(a_i, b)$ was found in Equation (B.11). Using that,

$$\begin{aligned}
\sigma_B^2 &= \int_{-\infty}^{\infty} b^2 f_B(b)\, db = \sum_{i=0}^{M-1} \int_{-\infty}^{\infty} b^2 f_{A,B}(a_i, b)\, db \\
&= \sum_{i=0}^{M-1} \int_{d_i - a_i}^{d_{i+1} - a_i} b^2 f_X(b + a_i)\, db = \sum_{i=0}^{M-1} \int_{d_i}^{d_{i+1}} (x - a_i)^2 f_X(x)\, dx \\
&= \sum_{i=0}^{M-1} \left[\int_{d_i}^{d_{i+1}} x^2 f_X(x)\, dx - 2a_i \int_{d_i}^{d_{i+1}} x f_X(x)\, dx + a_i^2 \int_{d_i}^{d_{i+1}} f_X(x)\, dx \right] \\
&= \sigma_X^2 + \sum_{i=0}^{M-1} \left[-2a_i \int_{d_i}^{d_{i+1}} x f_X(x)\, dx + a_i^2 \int_{d_i}^{d_{i+1}} f_X(x)\, dx \right].
\end{aligned} \tag{B.56}$$

When performing optimizations of the MSE, the channel power constraint of Equation (3.20) must be satisfied. That means that $\sigma_{Y_1}^2 = E[Y_1^2]$ and $\sigma_{Y_2}^2 = E[Y_2^2]$ must be found. $\sigma_{Y_2}^2 = K_b^2 \sigma_B^2$, so what remains is finding $\sigma_{Y_1}^2$, which is given by

$$\sigma_{Y_1}^2 = \sum_{i=0}^{M-1} c_i^2 \Pr(Y_1 = c_i) = \sum_{i=0}^{M-1} c_i^2 \Pr(X \in \mathcal{D}_i) = \sum_{i=0}^{M-1} c_i^2 \int_{d_i}^{d_{i+1}} f_X(x)\, dx. \tag{B.57}$$

To find D from (B.54), a closed-form expression for $p_{i,j}$ is needed. The noise is always assumed to be Gaussian, so f_N is given by (B.19). Using the definition of the error function,

$$p_{i,j} = \frac{1}{\sqrt{2\pi}\sigma_N} \int_{e_i - c_i}^{e_{i+1} - c_i} e^{-\frac{n_1^2}{2\sigma_N^2}}\, dn_1 = \frac{1}{2} \left[\operatorname{erf}\left(\frac{e_{j+1} - c_i}{\sqrt{2}\sigma_X} \right) - \operatorname{erf}\left(\frac{e_j - c_i}{\sqrt{2}\sigma_X} \right) \right]. \tag{B.58}$$

Now, closed-form expressions for D, κ, $\sigma_{Y_1}^2$, and $\sigma_{Y_2}^2$ can be found by solving the three integrals $\int_{d_i}^{d_{i+1}} f_X(x)\, dx$, which is the probability of $X \in (d_i, d_{i+1}] = \mathcal{D}_i$, $\int_{d_i}^{d_{i+1}} x f_X(x)\, dx$, and $\int_{d_i}^{d_{i+1}} x^2 f_X(x)\, dx$. In Sections B.3.1, B.3.2, and B.3.3, these integrals are calculated for Gaussian, Laplacian, and uniform source distributions, respectively.

B.3.1 Gaussian Source

The pdf of a Gaussian source is given by (B.20), and

$$\begin{aligned}\int_{d_i}^{d_{i+1}} f_X(x)\,dx &= \frac{1}{\sqrt{2\pi}\sigma_X}\int_{d_i}^{d_{i+1}} e^{-\frac{x^2}{2\sigma_X^2}}\,dx \\ &= \frac{1}{2}\left[\operatorname{erf}\left(\frac{d_{i+1}}{\sqrt{2}\sigma_X}\right) - \operatorname{erf}\left(\frac{d_i}{\sqrt{2}\sigma_X}\right)\right], \qquad \text{(B.59)}\end{aligned}$$

$$\begin{aligned}\int_{d_i}^{d_{i+1}} x f_X(x)\,dx &= \frac{1}{\sqrt{2\pi}\sigma_X}\int_{d_i}^{d_{i+1}} x e^{-\frac{x^2}{2\sigma_X^2}}\,dx \overset{t=\frac{x^2}{2\sigma_X^2}}{=} \frac{1}{\sqrt{2\pi}\sigma_X}\int_{\frac{d_i^2}{2\sigma_X^2}}^{\frac{d_{i+1}^2}{2\sigma_X^2}} e^{-t}\sigma_X^2\,dt \\ &= \frac{\sigma_X}{\sqrt{2\pi}}\left[-e^{-\frac{d_{i+1}^2}{2\sigma_X^2}} + e^{-\frac{d_i^2}{2\sigma_X^2}}\right], \qquad \text{(B.60)}\end{aligned}$$

and, using integration by parts,

$$\begin{aligned}\int_{d_i}^{d_{i+1}} x^2 f_X(x)\,dx &= \frac{1}{\sqrt{2\pi}\sigma_X}\int_{d_i}^{d_{i+1}} x^2 e^{-\frac{x^2}{2\sigma_X^2}}\,dx \\ &= \frac{1}{\sqrt{2\pi}\sigma_X}\int_{d_i}^{d_{i+1}} x\left(x e^{-\frac{x^2}{2\sigma_X^2}}\right)dx \\ &= \frac{\sigma_X}{\sqrt{2\pi}}\left[-d_{i+1}e^{-\frac{d_{i+1}^2}{2\sigma_X^2}} + d_i e^{-\frac{d_i^2}{2\sigma_X^2}} + \int_{d_i}^{d_{i+1}} e^{-\frac{x^2}{2\sigma_X^2}}\,dx\right] \\ &= \frac{\sigma_X}{\sqrt{2\pi}}\left[-d_{i+1}e^{-\frac{d_{i+1}^2}{2\sigma_X^2}} + d_i e^{-\frac{d_i^2}{2\sigma_X^2}}\right] \\ &\qquad + \frac{\sigma_X}{2}\left[\operatorname{erf}\left(\frac{d_{i+1}}{\sqrt{2}\sigma_X}\right) - \operatorname{erf}\left(\frac{d_i}{\sqrt{2}\sigma_X}\right)\right]. \qquad \text{(B.61)}\end{aligned}$$

B.3.2 Laplacian Source

First consider the case $d_i < d_{i+1} \leq 0$. The probability density function is given by (B.26), and

$$\int_{d_i}^{d_{i+1}} f_X(x)\,dx = \frac{1}{\sqrt{2}\sigma_X}\int_{d_i}^{d_{i+1}} e^{\frac{\sqrt{2}x}{\sigma_X}}\,dx = \frac{1}{2}\left[e^{\frac{\sqrt{2}d_{i+1}}{\sigma_X}} - e^{\frac{\sqrt{2}d_i}{\sigma_X}}\right], \tag{B.62}$$

$$\begin{aligned}\int_{d_i}^{d_{i+1}} x f_X(x)\,dx &= \frac{1}{\sqrt{2}\sigma_X}\int_{d_i}^{d_{i+1}} x e^{\frac{\sqrt{2}x}{\sigma_X}}\,dx \\ &= \frac{1}{2}\left[d_{i+1}e^{\frac{\sqrt{2}d_{i+1}}{\sigma_X}} - d_i e^{\frac{\sqrt{2}d_i}{\sigma_X}} - \int_{d_i}^{d_{i+1}} e^{\frac{\sqrt{2}x}{\sigma_X}}\,dx\right] \\ &= \frac{1}{2}\left[\left(d_{i+1} - \frac{\sigma_X}{\sqrt{2}}\right)e^{\frac{\sqrt{2}d_{i+1}}{\sigma_X}} - \left(d_i - \frac{\sigma_X}{\sqrt{2}}\right)e^{\frac{\sqrt{2}d_i}{\sigma_X}}\right],\end{aligned} \tag{B.63}$$

and

$$\begin{aligned}\int_{d_i}^{d_{i+1}} x^2 f_X(x)\,dx &= \frac{1}{\sqrt{2}\sigma_X}\int_{d_i}^{d_{i+1}} x^2 e^{\frac{\sqrt{2}x}{\sigma_X}}\,dx \\ &= \frac{1}{2}\left[d_{i+1}^2 e^{\frac{\sqrt{2}d_{i+1}}{\sigma_X}} - d_i^2 e^{\frac{\sqrt{2}d_i}{\sigma_X}} - 2\int_{d_i}^{d_{i+1}} x e^{\frac{\sqrt{2}x}{\sigma_X}}\,dx\right] \\ &= \frac{1}{2}\left[\left(d_{i+1}^2 - \sqrt{2}\sigma_X d_{i+1} + \sigma_X^2\right)e^{\frac{\sqrt{2}d_{i+1}}{\sigma_X}} - \left(d_i^2 - \sqrt{2}\sigma_X d_i + \sigma_X^2\right)e^{\frac{\sqrt{2}d_i}{\sigma_X}}\right],\end{aligned} \tag{B.64}$$

where integration by parts was used in the last two equations. Similarly, for $0 \leq d_i < d_{i+1}$,

$$\int_{d_i}^{d_{i+1}} f_X(x)\,dx = \frac{1}{2}\left[-e^{-\frac{\sqrt{2}d_{i+1}}{\sigma_X}} + e^{-\frac{\sqrt{2}d_i}{\sigma_X}}\right], \tag{B.65}$$

$$\int_{d_i}^{d_{i+1}} x f_X(x)\,dx = \frac{1}{2}\left[-\left(d_{i+1} + \frac{\sigma_X}{\sqrt{2}}\right)e^{-\frac{\sqrt{2}d_{i+1}}{\sigma_X}} + \left(d_i + \frac{\sigma_X}{\sqrt{2}}\right)e^{-\frac{\sqrt{2}d_i}{\sigma_X}}\right], \tag{B.66}$$

and

$$\begin{aligned}&\int_{d_i}^{d_{i+1}} x^2 f_X(x)\,dx \\ &= \frac{1}{2}\left[-\left(d_{i+1}^2 + \sqrt{2}\sigma_X d_{i+1} + \sigma_X^2\right)e^{-\frac{\sqrt{2}d_{i+1}}{\sigma_X}} + \left(d_i^2 + \sqrt{2}\sigma_X d_i + \sigma_X^2\right)e^{-\frac{\sqrt{2}d_i}{\sigma_X}}\right].\end{aligned} \tag{B.67}$$

When $d_i < 0 < d_{i+1}$, the integrals must be written as a sum of two integrals, one going from d_i to 0, and another going from 0 to d_{i+1}. The previous expressions can be used for each of these integrals. This gives:

$$\int_{d_i}^{d_{i+1}} f_X(x)\,dx = \frac{1}{2}\left[2 - e^{-\frac{\sqrt{2}d_{i+1}}{\sigma_X}} - e^{\frac{\sqrt{2}d_i}{\sigma_X}}\right], \tag{B.68}$$

$$\int_{d_i}^{d_{i+1}} x f_X(x)\,dx = \frac{1}{2}\left[-\left(d_{i+1} + \frac{\sigma_X}{\sqrt{2}}\right)e^{-\frac{\sqrt{2}d_{i+1}}{\sigma_X}} - \left(d_i - \frac{\sigma_X}{\sqrt{2}}\right)e^{\frac{\sqrt{2}d_i}{\sigma_X}}\right], \tag{B.69}$$

and

$$\int_{d_i}^{d_{i+1}} x^2 f_X(x)\,dx = \frac{1}{2}\left[2\sigma_X^2 - \left(d_{i+1}^2 + \sqrt{2}\sigma_X d_{i+1} + \sigma_X^2\right)e^{-\frac{\sqrt{2}d_{i+1}}{\sigma_X}} - \left(d_i^2 - \sqrt{2}\sigma_X d_i + \sigma_X^2\right)e^{\frac{\sqrt{2}d_i}{\sigma_X}}\right]. \tag{B.70}$$

B.3.3 Uniform Source

In this case, the pdf is given by (B.34). Assuming $d_0 = -\frac{\Delta}{2}$ and $d_M = \frac{\Delta}{2}$,

$$\int_{d_i}^{d_{i+1}} f_X(x)\,dx = \frac{1}{\Delta}\int_{d_i}^{d_{i+1}} dx = \frac{d_{i+1} - d_i}{\Delta}, \tag{B.71}$$

$$\int_{d_i}^{d_{i+1}} x f_X(x)\,dx = \frac{1}{\Delta}\int_{d_i}^{d_{i+1}} x dx = \frac{d_{i+1}^2 - d_i^2}{2\Delta}, \tag{B.72}$$

and

$$\int_{d_i}^{d_{i+1}} x^2 f_X(x)\,dx = \frac{1}{\Delta}\int_{d_i}^{d_{i+1}} x^2\,dx = \frac{d_{i+1}^3 - d_i^3}{3\Delta}. \tag{B.73}$$

B.3.4 MSE Using Optimal Receivers for Uniform Source and Uniform Quantizers

When the source is uniform, and the quantizer is also uniform with midpoint representation values and quantization interval $\delta = \frac{\Delta}{M}$, the quantized symbol A and the quantization error B are independent of each other. This can be seen by realizing that the conditional distribution $f_{B|A}$ is given by

$$f_{B|A}(b \mid a_i) = \begin{cases} \frac{1}{\delta} & \text{when } -\frac{\delta}{2} < b \leq \frac{\delta}{2}, \\ 0 & \text{otherwise,} \end{cases} \tag{B.74}$$

regardless of i. Thus, $f_B(b) = f_{B|A}(b \mid a_i)$, and A and B are statistically independent. In Section B.1.3, it was shown that the optimal receivers in this case depend only on the received version of the variable they are estimating.[1] Since the noise sequences are independent of both the signal and of each other, the received signal $\hat{Y}_1$ will be independent of B and of $\hat{Y}_2$, and the received signal $\hat{Y}_2$ will be independent of A and of $\hat{Y}_1$, and as the receiver functions depend only on one variable, $\hat{A}$ will be independent of $\hat{B}$ and B, and $\hat{B}$ will be independent of $\hat{A}$ and A. This means that the expression for the MSE can be rewritten as

$$\begin{aligned} D &= E[(X - \hat{X})^2] = E[(A - \hat{A} + B - \hat{B})^2] \\ &= E[(A - \hat{A})^2] + 2E[(A - \hat{A})(B - \hat{B})] + E[(B - \hat{B})^2] \\ &= E[(A - \hat{A})^2] + E[(B - \hat{B})^2] \end{aligned} \tag{B.75}$$

since both A, B, $\hat{A}$ and $\hat{B}$ have zero mean due to symmetry. Since the optimal receivers depend only on one variable in this case, they will be written $\tilde{\beta}_1(\hat{y}_1)$ and $\tilde{\beta}_2(\hat{y}_2)$ as for the simplified receivers. The MSE is now given by

$$\begin{aligned} D &= E[(A - \tilde{\beta}_1(C_a(A) + N_1))^2] + E[(B - \tilde{\beta}_2(K_b B + N_2))^2] \\ &= \sum_{i=0}^{M-1} \int_{-\infty}^{\infty} (a_i - \tilde{\beta}_1(c_i + n_1))^2 f_N(n_1) \Pr(A = a_i)\, dn_1 \\ &\quad + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (b - \tilde{\beta}_2(K_b b + n_2))^2 f_B(b) f_N(n_2)\, db\, dn_2, \end{aligned} \tag{B.76}$$

where $C_a(A)$ returns the channel sample value corresponding to the quantized value A, so $C_a(a_i) = c_i$. Due to the uniform distribution and quantizer, $\Pr(A = a_i) = \frac{\delta}{\Delta} = \frac{1}{M}$ for all i, so,

$$\begin{aligned} D = \frac{1}{M}\frac{1}{\sqrt{2\pi}\sigma_N} \sum_{i=0}^{M-1} \int_{-\infty}^{\infty} (a_i - \tilde{\beta}_1(c_i + n_1))^2 e^{-\frac{n_1^2}{2\sigma_N^2}}\, dn_1 \\ + \frac{1}{\delta}\frac{1}{\sqrt{2\pi}\sigma_N} \int_{-\infty}^{\infty}\int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} (b - \tilde{\beta}_2(K_b b + n_2))^2 e^{-\frac{n_2^2}{2\sigma_N^2}}\, dn_2\, db. \end{aligned} \tag{B.77}$$

The expressions for the receivers can now be inserted, and the remaining problem is the integrals in (B.77). There is no closed-form expression for these integrals with the receivers of Equations (B.39) and (B.40), but numerical one and two-dimensional integration is feasible in combination with the one-dimensional numerical optimization needed in this problem.

[1] This can also be shown from the statistical independence.

Appendix C

Original Images

Figures C.1–C.4 show the four original images used in the coding examples in Chapter 4.

Figure C.1 Original image "Lena"

Figure C.2 Original image "Barbara"

Figure C.3 Original image "Goldhill"

Figure C.4 Original image "Bridge"

References

Aase, S. O. (1993), Image Subband Coding Artifacts: Analysis and Remedies, Dr. ing. dissertation, Norwegian Institute of Technology.

Ahlswede, R. (1985), "The rate-distortion region for multiple descriptions without excess rate", *IEEE Trans. Inform. Theory*, vol. IT-31, no. 6, pp. 721–726, November 1985.

Alasti, M., Sayrafian-Pour, K., Ephremides, A. and Farvardin, N. (2001), "Multiple description coding in networks with congestion problem", *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 891–902, March 2001.

Arean, R., Kovačević, J. and Goyal, V. K. (2000), "Multiple description perceptual audio coding with correlating transforms", *IEEE Trans. Speech Audio Processing*, vol. 8, no. 2, pp. 140–145, March 2000.

Ayanoğlu, E. and Gray, R. M. (1987), "The design of joint source and channel trellis waveform coders", *IEEE Trans. Inform. Theory*, vol. IT-33, no. 6, pp. 855–865, November 1987.

Balan, R., Daubechies, I. and Vaishampayan, V. (2000), "The analysis and design of windowed Fourier frame based multiple description source coding schemes", *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2491–2536, November 2000.

Balasingham, I. (1998), On Optimal Perfect Reconstruction Filter Banks for Image Compression, Dr. ing. dissertation, Norwegian University of Science and Technology.

Batllo, J.-C. and Vaishampayan, V. A. (1997), "Asymptotic performance of multiple description transform codes", *IEEE Trans. Inform. Theory*, vol. 43, no. 2, pp. 703–707, March 1997.

Berger, T. (1971), *Rate Distortion Theory*, Prentice-Hall, Inc., Englewood Cliffs, NJ, USA.

Berger, T. and Tufts, D. W. (1967), "Optimum pulse amplitude modulation—Part I: Transmitter-receiver design and bounds from information theory", *IEEE Trans. Inform. Theory*, vol. IT-13, no. 2, pp. 196–208, April 1967.

Berger-Wolf, T. Y. and Reingold, E. M. (1999), Optimal multichannel communication under failure, *in* "Proc. Tenth Annual ACM-SIAM Symp. on Discrete Algorithms (SODA '99)", Baltimore, MD, USA, January 1999.

Berger-Wolf, T. Y. and Reingold, E. M. (2000), "Index assignment for multichannel communication under failure", Submitted to *IEEE Trans. Inform. Theory*, November 2000. Available from `http://feline.cs.uiuc.edu:1024/~tanyabw/research/`.

Blahut, R. E. (1972), "Computation of channel capacity and rate-distortion functions", *IEEE Trans. Inform. Theory*, vol. IT-18, no. 4, pp. 460–473, July 1972.

Blahut, R. E. (1983), *Theory and Practice of Error Control Codes*, Addison-Wesley, Reading, MA, USA.

Blahut, R. E. (1987), *Principles and Practice of Information Theory*, Addison-Wesley, Reading, MA, USA.

Blahut, R. E. (1997), *Theory and Practice of Data Transmission Codes*, 2nd ed., Dept. ECE, Univ. of Illinois, Urbana-Champaign. Preliminary draft.

Buttigieg, V. and Farrell, P. G. (2000), "Variable-length error-correcting codes", *IEE Proc.-Commun.*, vol. 147, no. 4, pp. 211–215, August 2000.

Cai, J. and Chen, C. W. (2000), "Robust joint source-channel coding for image transmission over wireless channels", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, September 2000.

Chande, V. and Farvardin, N. (2000), "Progressive transmission of images over memoryless noisy channels", *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 850–860, June 2000.

Chen, Q. and Fischer, T. R. (1998), "Image coding using robust quantization for noisy digital transmission", *IEEE Trans. Image Processing*, vol. 7, no. 4, pp. 496–505, April 1998.

Chung, D.-M. and Wang, Y. (1999), "Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 6, pp. 895–908, September 1999.

Chung, S.-Y., Forney, Jr., G. D., Richardson, T. J. and Urbanke, R. (2001), "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit", *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58–60, February 2001.

Coleman, T., Branch, M. A. and Grace, A. (1999), *Optimization Toolbox: For Use with MATLAB*, The Math Works Inc. Version 2.

Cover, T. M. and Thomas, J. A. (1991), *Elements of Information Theory*, John Wiley & Sons, Inc., New York, NY, USA.

Coward, H., Knopp, R. and Servetto, S. D. (2001*a*), "On the performance of a natural class of joint source–channel codes based on multiple descriptions", Submitted to *IEEE Trans. Inform. Theory*, August 2001.

Coward, H., Knopp, R. and Servetto, S. D. (2001*b*), On the performance of multiple description codes over bit error channels, *in* "Proc. IEEE Int. Symp. Inform. Theory (ISIT)", Washington, D.C., USA, June 2001, p. 240.

Coward, H. and Ramstad, T. A. (1999), Hybrid digital-analog transmission of analog source signals, *in* "Proc. Norwegian Signal Processing Symp. (NORSIG)", Asker, Norway, September 1999, pp. 17–21.

Coward, H. and Ramstad, T. A. (2000*a*), Bandwidth doubling in combined source-channel coding of memoryless Gaussian sources, *in* "Proc. IEEE Int. Symp. Intell. Signal Processing Commun. Syst. (ISPACS)", vol. 1, Honolulu, HI, USA, November 2000, pp. 571–576.

Coward, H. and Ramstad, T. A. (2000*b*), Quantizer optimization in hybrid digital-analog transmission of analog source signals, *in* "Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)", vol. 5, Istanbul, Turkey, June 2000, pp. 2637–2640.

Coward, H. and Ramstad, T. A. (2000*c*), Robust image communication using bandwidth reducing and expanding mappings, *in* "Conf. Record 34th Asilomar Conf. Signals, Syst., Comput.", vol. 2, Pacific Grove, CA, USA, October 2000, pp. 1384–1388.

Diggavi, S. N., Sloane, N. J. A. and Vaishampayan, V. A. (2000), Design of asymmetric multiple description lattice vector quantizers, *in* "Proc. IEEE Data Compression Conf. (DCC)", Snowbird, UT, USA, March 2000, pp. 490–499.

El Gamal, A. A. and Cover, T. M. (1982), "Achievable rates for multiple descriptions", *IEEE Trans. Inform. Theory*, vol. IT-28, no. 6, pp. 851–857, November 1982.

Equitz, W. H. R. and Cover, T. M. (1991), "Successive refinement of information", *IEEE Trans. Inform. Theory*, vol. 37, no. 2, pp. 269–275, March 1991.

Farvardin, N. (1990), "A study of vector quantization for noisy channels", *IEEE Trans. Inform. Theory*, vol. 36, no. 4, pp. 799–809, July 1990.

Farvardin, N. and Modestino, J. W. (1984), "Optimum quantizer performance for a class of non-Gaussian memoryless sources", *IEEE Trans. Inform. Theory*, vol. IT-30, no. 3, pp. 485–497, May 1984.

Farvardin, N. and Vaishampayan, V. (1987), "Optimal quantizer design for noisy channels: An approach to combined source-channel coding", *IEEE Trans. Inform. Theory*, vol. 33, no. 6, pp. 827–838, November 1987.

Farvardin, N. and Vaishampayan, V. (1991), "On the performance and complexity of channel-optimized vector quantizers", *IEEE Trans. Inform. Theory*, vol. 37, no. 1, pp. 155–160, January 1991.

Fazel, K. and Lhullier, J. J. (1990), Application of unequal error protection codes on combined source-channel coding of images, *in* "Conf. record IEEE Int. Conf. Commun. (ICC)", vol. 3, April 1990, pp. 898–903.

Fine, T. (1964), "Properties of an optimum digital system and applications", *IEEE Trans. Inform. Theory*, vol. IT-10, no. 4, pp. 287–296, October 1964.

Fischer, T. R. and Marcellin, M. W. (1991), "Joint trellis coded quantization/modulation", *IEEE Trans. Commun.*, vol. 39, no. 2, pp. 172–176, February 1991.

Fleming, M. and Effros, M. (1999), Generalized multiple description vector quantization, *in* "Proc. IEEE Data Compression Conf. (DCC)", Snowbird, UT, USA, March 1999, pp. 3–12.

Fuldseth, A. (1997), Robust Subband Video Compression for Noisy Channels with Multilevel Signaling, Dr. ing. dissertation, Norwegian University of Science and Technology.

Fuldseth, A. and Ramstad, T. A. (1997*a*), Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space, *in* "Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)", vol. 4, Munich, Germany, April 1997, pp. 3093–3096.

Fuldseth, A. and Ramstad, T. A. (1997*b*), Channel-optimized subband video coding for channels with a power constraint, *in* "Proc. IEEE Int. Conf. Image Processing (ICIP)", vol. 3, Santa Barbara, CA, USA, October 1997, pp. 428–431.

Gallager, R. G. (1962), "Low density parity check codes", *IRE Trans. Inform. Theory*, vol. IT-8, pp. 21–28, January 1962.

Gersho, A. and Gray, R. M. (1992), *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, MA, USA.

Goblick, Jr., T. J. (1965), "Theoretical limitations on the transmission of data from analog sources", *IEEE Trans. Inform. Theory*, vol. IT-11, no. 4, pp. 558–567, October 1965.

Goldsmith, A. J. and Effros, M. (1998), "Joint design of fixed-rate source codes and multiresolution channel codes", *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1301–1312, October 1998.

Goyal, V. K. and Kovačević, J. (2001), "Generalized multiple description coding with correlating transforms", *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2199–2224, September 2001.

Goyal, V. K., Kovačević, J. and Vetterli, M. (1999), Quantized frame expansions as source–channel codes for erasure channels, *in* "Proc. IEEE Data Compression Conf. (DCC)", Snowbird, UT, USA, March 1999, pp. 326–335.

Hagen, R. and Hedelin, P. (1999), "Robust vector quantization by a linear mapping of a block code", *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 200–218, January 1999.

Hagenauer, J. and Stockhammer, T. (1999), "Channel coding and transmission aspects for wireless multimedia", *Proc. IEEE*, vol. 87, no. 10, pp. 1764–1777, October 1999.

IEEE (1985), *IEEE Standard for Binary Floating-Point Arithmetic.* (ANSI/IEEE Std 754-1985).

Ingle, A. and Vaishampayan, V. A. (1995), "DPCM system design for diversity systems with applications to packetized speech", *IEEE Trans. Speech Audio Processing*, vol. 3, no. 1, pp. 48–58, January 1995.

ISO/IEC (2001), *Information Technology—JPEG 2000 Image Coding System*. International Standard 15444-1.

Jafarkhani, H. and Tarokh, V. (1999), "Multiple description trellis-coded quantization", *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 799–803, June 1999.

Jayant, N. S. (1981), "Subsampling of a DPCM speech channel to provide two "self-contained" half-rate channels", *Bell Syst. Tech. J.*, vol. 60, no. 4, pp. 501–509, April 1981.

Joshi, R. L. and Fischer, T. R. (1995), "Comparison of generalized Gaussian and Laplacian modeling in DCT image coding", *IEEE Signal Processing Lett.*, vol. 2, no. 5, pp. 81–82, May 1995.

Knopp, R. and Humblet, P. A. (2000), "On coding for block fading channels", *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 189–205, January 2000.

Kozintsev, I. and Ramchandran, K. (1998), "Robust image transmission over energy-constrained time-varying channels using multiresolution joint source–channel coding", *IEEE Trans. Signal Processing*, vol. 46, no. 4, pp. 1012–1026, April 1998.

Kurtenbach, A. J. and Wintz, P. A. (1969), "Quantizing for noisy channels", *IEEE Trans. Commun. Technol.*, vol. COM-17, no. 2, pp. 291–302, April 1969.

Lee, K.-H. and Petersen, D. P. (1976), "Optimal linear coding for vector channels", *IEEE Trans. Commun.*, vol. COM-24, no. 12, pp. 1283–1290, December 1976.

Lervik, J. M. (1996), Subband Image Communication over Digital Transparent and Analog Waveform Channels, Dr. ing. dissertation, Norwegian University of Science and Technology.

Lervik, J. M. and Fischer, T. R. (1997), Robust subband image coding for waveform channels with optimum power- and bandwidth allocation, *in* "Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)", vol. 4, Munich, Germany, April 1997, pp. 3089–3092.

Lervik, J. M. and Ramstad, T. A. (1996), Robust image communication using subband coding and multilevel modulation, *in* "Proc. 1996 Symp. Visual Commun. Image Processing (VCIP)", vol. SPIE 2727, part 2, Orlando, FL, USA, March 1996, pp. 524–535.

Liu, F.-H., Ho, P. and Cuperman, V. (1993), Joint source and channel coding using a non-linear receiver, *in* "Conference Record, IEEE Int. Conf. on Communications", Geneva, Switzerland, June 1993, pp. 1502–1507.

Lookabaugh, T. D. and Gray, R. M. (1989), "High-resolution quantization theory and the vector quantizer advantage", *IEEE Trans. Inform. Theory*, vol. 35, no. 5, pp. 1020–1033, September 1989.

MacKay, D. J. C. (1999), "Good error-correcting codes based on very sparse matrices", *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 399–431, March 1999.

MacKay, D. J. C. and Neal, R. M. (1997), "Near Shannon limit performance of low density parity check codes", *Electronics Letters*, vol. 33, no. 6, pp. 457–458, March 1997.

Masnick, B. and Wolf, J. (1967), "On linear unequal error protection codes", *IEEE Trans. Inform. Theory*, vol. IT-13, no. 4, pp. 600–607, October 1967.

Massey, J. L. (1978), Joint source and channel coding, *in* J. K. Skwirzynski, ed., "Communication Systems and Random Process Theory", Nato Advanced Study Institute Series, Sijthoff & Noordhoff, Alphen aan den Rijn, the Netherlands, pp. 279–293.

Mittal, U. (1999), Broadcasting, Robustness and Duality in a Joint Source-Channel Coding System, Ph.D. dissertation, State University of New York at Stony Brook.

Mittal, U. and Phamdo, N. (2000*a*), Achievable distortion regions of Gaussian broadcast systems, *in* "Proc. IEEE Int. Symp. Inform. Theory (ISIT)", Sorrento, Italy, June 2000, p. 24.

Mittal, U. and Phamdo, N. (2000*b*), "Duality theorems for joint source–channel coding", *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1263–1275, July 2000.

Mittal, U. and Phamdo, N. (2000*c*), "Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications", Submitted to *IEEE Trans. Inform. Theory*, November 1998. Revised, February 2000.

Modestino, J. W. and Daut, D. G. (1979), "Combined source-channel coding of images", *IEEE Trans. Commun.*, vol. COM-27, no. 11, pp. 1644–1659, November 1979.

Modestino, J. W., Daut, D. G. and Vickers, A. L. (1981), "Combined source-channel coding of images using the block cosine transform", *IEEE Trans. Commun.*, vol. COM-29, no. 9, pp. 1261–1274, September 1981.

Myhre, B., Markhus, V. and Øien, G. E. (2001), LDPC coded adaptive multilevel modulation for slowly varying Rayleigh-fading channels, *in* "Proc. Norwegian Signal Processing Symp. (NORSIG)", Trondheim, Norway, October 2001.

Nelson, M. and Gailly, J.-L. (1996), *The Data Compression Book*, 2nd ed., M & T Books.

Nickl, H., Hagenauer, J. and Burkert, F. (1997), "Approaching Shannon's capacity limit by 0.27 dB using simple Hamming codes", *IEEE Commun. Lett.*, vol. 1, no. 5, pp. 130–132, September 1997.

Nyquist, H. (1928), "Certain topics in telegraph transmission theory", *Trans. A. I. E. E.*, vol. 47, pp. 617–644, February 1928.

Ortega, A. and Ramchandran, K. (1998), "Rate-distortion methods for image and video compression", *IEEE Signal Processing Mag.*, vol. 15, no. 6, pp. 23–50, November 1998.

Ozarow, L. H. (1980), "On a source-coding problem with two channels and three receivers", *Bell Syst. Tech. J.*, vol. 59, no. 10, pp. 1909–1921, December 1980.

Ozarow, L. H., Shamai, S. and Wyner, A. D. (1994), "Information theoretic considerations for cellular mobile radio", *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

Peterson, W. W. and Weldon, Jr., E. J. (1972), *Error-Correcting Codes*, 2nd ed., The MIT Press, Cambridge, MA, USA.

Phamdo, N. and Farvardin, N. (1994), "Optimal detection of discrete Markov sources over discrete memoryless channels—applications to combined source-channel coding", *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 186–193, January 1994.

Phamdo, N. and Mittal, U. (2000), A CELP-based hybrid digital-analog (HDA) joint source-channel speech coder, *in* "Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)", vol. 3, Istanbul, Turkey, June 2000, IEEE, pp. 1487–1490.

Proakis, J. G. (2001), *Digital Communications*, 4th ed., McGraw-Hill.

Ramstad, T. A., Aase, S. O. and Husøy, J. H. (1995), *Subband Compression of Images: Principles and Examples*, Elsevier Science B.V., Amsterdam, the Netherlands.

Rimoldi, B. (1994), "Successive refinement of information: Characterization of the achievable rates", *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 253–259, January 1994.

Ruf, M. J. and Modestino, J. W. (1999), "Operational rate-distortion performance for joint source and channel coding of images", *IEEE Trans. Image Processing*, vol. 8, no. 3, pp. 305–320, March 1999.

Sayood, K. and Borkenhagen, J. C. (1991), "Use of residual redundancy in the design of joint source/channel coders", *IEEE Trans. Commun.*, vol. 39, pp. 838–846, June 1991.

Sayood, K., Otu, H. H. and Demir, N. (2000), "Joint source/channel coding for variable length codes", *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 787–794, May 2000.

Servetto, S. D. and Nahrstedt, K. (2001), "Broadcast quality video over IP", *IEEE Trans. Multimedia*, vol. 3, no. 1, pp. 162–173, March 2001.

Servetto, S. D., Ramchandran, K., Vaishampayan, V. A. and Nahrstedt, K. (2000), "Multiple description wavelet based image coding", *IEEE Trans. Image Processing*, vol. 9, no. 5, pp. 813–826, May 2000.

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

Shannon, C. E. (1949), "Communication in the presence of noise", *Proc. IRE*, vol. 37, pp. 10–21, January 1949.

Skinnemoen, P. H. (1994), Robust Communication with Modulation Organized Vector Quantization, Dr. ing. dissertation, Norwegian Institute of Technology.

Skoglund, M. (1995), A soft decoder vector quantizer for a Rayleigh fading channel – Application to image transmission, *in* "Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)", vol. 4, Detroit, MI, USA, May 1995, pp. 2507–2510.

Skoglund, M. and Hedelin, P. (1999), "Hadamard-based soft decoding for vector quantization over noisy channels", *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 515–532, March 1999.

Skoglund, M., Phamdo, N. and Alajaji, F. (2000), VQ-based hybrid digital–analog joint source–channel coding, *in* "Proc. IEEE Int. Symp. Inform. Theory (ISIT)", Sorrento, Italy, June 2000, p. 403.

Skoglund, M., Phamdo, N. and Alajaji, F. (2001*a*), "Design and performance of VQ-based hybrid digital–analog joint source–channel codes", Submitted to *IEEE Trans. Inform. Theory*, March 2000. Revised, June 2001.

Skoglund, M., Phamdo, N. and Alajaji, F. (2001*b*), Hybrid digital–analog coding for bandwidth compression/expansion using VQ and turbo codes, *in* "Proc. IEEE Int. Symp. Inform. Theory (ISIT)", Washington, D.C., USA, June 2001, p. 260.

Tanabe, N. and Farvardin, N. (1992), "Subband image coding using entropy-coded quantization over noisy channels", *IEEE J. Select. Areas Commun.*, vol. 10, no. 5, pp. 926–943, June 1992.

Therrien, C. W. (1992), *Discrete Random Signals and Statistical Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, USA.

Ungerboeck, G. (1982), "Channel coding with multilevel/phase signals", *IEEE Trans. Inform. Theory*, vol. IT-28, no. 1, pp. 55–67, January 1982.

Vaidyanathan, P. P. (1993), *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs.

Vaishampayan, V. A. (1989), Combined Source-Channel Coding for Bandlimited Waveform Channels, Ph.D. dissertation, University of Maryland.

Vaishampayan, V. A. (1993), "Design of multiple description scalar quantizers", *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 821–834, May 1993.

Vaishampayan, V. A. and Batllo, J.-C. (1998), "Asymptotic analysis of multiple description quantizers", *IEEE Trans. Inform. Theory*, vol. 44, no. 1, pp. 278–284, January 1998.

Vaishampayan, V. A. and Domaszewicz, J. (1994), "Design of entropy-constrained multiple-description scalar quantizers", *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 245–250, January 1994.

Vaishampayan, V. A. and Farvardin, N. (1990), "Optimal block cosine transform image coding for noisy channels", *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 327–336, March 1990.

Vaishampayan, V. A. and Farvardin, N. (1992), "Joint design of block source codes and modulation signal sets", *IEEE Trans. Inform. Theory*, vol. 38, no. 4, pp. 1230–1248, July 1992.

Vaishampayan, V. A., Sloane, N. J. A. and Servetto, S. D. (2001), "Multiple-description vector quantization with lattice codebooks: Design and analysis", *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 1718–1734, July 2001.

Venkataramani, R., Kramer, G. and Goyal, V. K. (2001), Bounds on the achievable region for certain multiple description coding problems, *in* "Proc. IEEE Int. Symp. Inform. Theory (ISIT)", Washington, D.C., USA, June 2001, p. 148.

Walpole, R. E. and Myers, R. H. (1993), *Probability and Statistics for Engineers and Scientists*, 5th ed., Prentice Hall, Englewood Cliffs, NJ, USA.

Walsh, G. R. (1975), *Methods of Optimization*, John Wiley & Sons, London, England.

Wang, Y., Orchard, M. T., Vaishampayan, V. and Reibman, A. R. (2001), "Multiple description coding using pairwise correlating transforms", *IEEE Trans. Image Processing*, vol. 10, no. 3, pp. 351–366, March 2001.

Westerink, P. H., Biemond, J. and Boekee, D. E. (1988), An optimal bit allocation algorithm for sub-band coding, *in* "Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)", New York, NY, USA, April 1988, pp. 757–760.

Witsenhausen, H. S. (1980), "On source networks with minimal breakdown degradation", *Bell Syst. Tech. J.*, vol. 59, no. 6, pp. 1083–1087, July–August 1980.

Wolf, J. K., Wyner, A. D. and Ziv, J. (1980), "Source coding for multiple descriptions", *Bell Syst. Tech. J.*, vol. 59, no. 8, pp. 1417–1426, October 1980.

Yang, S.-M. and Vaishampayan, V. A. (1995), "Low-delay communication for Rayleigh fading channels: An application of the multiple description quantizer", *IEEE Trans. Commun.*, vol. 43, no. 11, pp. 2771–2783, November 1995.

Yang, X. and Ramchandran, K. (2000), "Optimal subband filter banks for multiple description coding", *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2477–2490, November 2000.

Zhang, Z. and Berger, T. (1995), "Multiple description source coding with no excess marginal rate", *IEEE Trans. Inform. Theory*, vol. 41, no. 2, pp. 349–357, March 1995.

Zheng, H. and Liu, K. J. R. (1999), "The subband modulation: A joint power and rate allocation framework for subband image and video transmission", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 5, pp. 823–838, August 1999.