Joint Source-Channel Coding for Image Transmission over Flat Fading Channels

Greg Harald Håkonsen

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Philosophiae doctor
· 2007 ·

Department of Electronics and Telecommunications
Norwegian University of Science and Technology

NTNU
Innovation and Creativity

Greg Harald Håkonsen

Joint Source-Channel Coding for Image Transmission over Flat Fading Channels

Thesis for the degree philosophiae doctor

Trondheim, June 2007

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics
and Electrical Engineering
Department of Electronics and Telecommunications

NTNU
Innovation and Creativity

NTNU
Norwegian University of Science and Technology

Thesis for the degree philosophiae doctor

Faculty of Information Technology, Mathematics and Electrical Engineering
Department of Electronics and Telecommunications

ISBN 978-82-471-2985-2 (printed version)
ISBN 978-82-471-2999-9 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2007:136

Printed by NTNU-trykk

Abstract

In this thesis, transmission of images over a flat fading channel using *joint source-channel coding* (JSCC) is considered. Through the use of nonlinear dimensional changing mappings, the system becomes robust. The system will not experience a clear breakdown, but have a graceful degradation which is visually pleasant. It is shown how multiple ways of adapting the source to the channel, and using the knowledge about the channel, results in a performance comparable to state-of-the-art systems but with less complexity. By relying on the robustness, reduction in channel information does not mean a large loss in performance.

The proposed system has low computational complexity as there are no separate source and channel coders. Compression and generation of channel symbols are done in one operation. By using nonlinear dimensional changing mappings, the dependency between the channel symbols are low. This leads to a system where the received image can be progressively decoded, and where the received information is still usable if the transmission stops unexpectedly.

By allowing a small variation around target time and transmission power constraints, the variation of the quality of the received image is kept small. This is done through planning and on-the-fly adaptation of the transmission. The planning depends on the distribution of the channel quality, and makes sure the channel is used bandwidth efficiently.

Through the results, the impact and importance of the design of some of the system parameters are analyzed and discussed.

By using theoretical models, it is shown how practical limitations of the system contributes to loss in performance. Similar techniques are also used to analyze where in the system effort should be made to improve the system.

The proposed system has a framework that can be easily extended to other scenarios.

Preface

This thesis is submitted in partial fulfillment of the requirements for the doctoral degree of *philosophiae doctor, Ph.D.* at the Norwegian University of Science and Technology (NTNU).

The research was carried out in the period from July 2002 to March 2007 at the Department of Electronics and Telecommunications, the Norwegian University of Science and Technology. The work was funded by Norwegian Research Council (NFR) through the BEATS project. Professor Tor A. Ramstad, Department of Electronics and Telecommunications, the Norwegian University of Science and Technology, has been supervisor.

Acknowledgments

I would first of all like to thank professor Tor A. Ramstad who has been my supervisor during the Ph.D. studies, and who got me to apply in the first place. Tor is a very nice person to work with, and a supervisor who truly cares about his students. I would also like to thank Geir E. Øien who has been project manager of the BEATS project. Through the BEATS project I have seen and experienced things I will always remember.

My fellow Ph.D students have also contributed to a fun and interesting time. I would especially like to mention my office buddy Fredrik Hekland who I have had many interesting and enlightening discussions with. I would also thank Duc Van Duong, Gang Lin, Øystein Birkenes, Saeeid Tahmasbi Oskuii, Anna Kim, Sébastien de la Kethulle de Ryhove and other current and previous Ph.D. students that contributed in making my time in Trondheim great. I also wish to thank all my colleagues at the Department of Electronics and Telecommunications at NTNU.

Bojana Gajic also deserves a special thanks for lots of support and encouraging words and great fun. Lina Hanssen is another person that has helped me and supported me through the studies.

Further I would like to thank Anders Gjendemsjø for supplying the channel rates for one of the reference channel transmission schemes.

Finally, I would like to thank my parents and sister for having faith in me all this time.

Greg Harald Håkonsen, May 2007.

Contents

List of Figures

Acronyms

3G *third-generation mobile technology*

ACM *adaptive coded modulation*

ASE *average spectral efficiency*

AWGN *additive white Gaussian noise*

BER *bit error rate*

BMF *bandwidth matching functions*

BPAM *block pulse amplitude modulation*

BSC *binary symmetric channel*

CAS *channel adaptation strategy*

COVQ *channel optimized vector quantizer*

CRC *cyclic redundancy check*

CSI *channel state information*

CSNR *channel signal-to-noise ratio*

DAB *digital audio broadcasting*

dB *decibel*

DCT *discrete cosine transform*

DPCM *differential pulse code modulation*

FM *frequency modulation*

GSL *GNU scientific library*

HDA *hybrid digital-analog*

HSQLC *hybrid scalar quantizer-linear coder*

Hz *Hertz*

i.i.d. *independent and identically distributed*

JPEG *joint photographic experts group*

JSCC *joint source-channel coding*

LDPC *low-density parity-check*

MSE *mean square error*

M-QAM *multilevel quadrature amplitude modulation*

OFDM *orthogonal frequency division multiplex*

OPTA *optimal performance theoretically attainable*

PAM *pulse amplitude modulation*

PCCOVQ *power constrained channel optimized vector quantizer*

pdf *probability density function*

PR *perfect reconstruction*

PSNR *peak signal-to-noise ratio*

QoS *quality of service*

RCPC *rate compatible punctured convolution*

RMS *root mean squared*

ROI *region of interest*

RS *Reed-Solomon*

SPIHT *set partitioning in hierarchical trees*

SNR *signal-to-noise ratio*

TuCM *turbo-coded modulation*

UEP *unequal error protection*

VQ *vector quantizer*

WTSOM *wavelet transform self organizing map*

List of symbols

$'$	Differentiation operator
$2D$	Two dimensional
$3D$	Three dimensional
α	Channel gain
α_{A_m}	Channel gain value for the centroid value of the m'th CSNR region
β	Decoder for linear mapping
γ	CSNR value
$\bar{\gamma}$	Average CSNR value
$\bar{\gamma}_e$	Expected CSNR value
$\bar{\gamma}_t$	True CSNR value
γ_{A_m}	Centroid value for the m'th CSNR region
γ_{T_m}	Upper CSNR threshold for the m'th channel region
$\gamma^*_{C_n}$	CSNR value optimized for block n
γ_C	CSNR a mapping is optimized for
γ_{C_m}	Representation level for region m
κ	Number of source samples in each block
λ	Lagrange multiplier
μ	Distortion level
Ψ	CSNR mismatch distortion by choosing γ_{A_m} as representation points
$\sigma^2_{D_n}$	Resulting distortion for the n'th source block
$\sigma^2_{X_n}$	Variance of the n'th source block
σ^2_N	Noise power on channel
σ^2_S	Transmit power on channel
σ^2_{Sm}	Transmit power on channel for channel state m
$\bar{\sigma}^2_S$	Average transmit power
$\sigma^2_{G_m}$	Power a mapping optimized for channel state m gives out
$\sigma^{2*}_{G_m}$	Mapping power optimized for block n

$\bar{\sigma}^2_{G_m}$	Average mapping power for blocks in channel state m
$\varkappa$	Barrier parameter
ν	CSNR penalty for the j'th mapping compared to OPTA
$\nu_{\hat{r}_n}$	CSNR penalty for the mapping with rate $\hat{r}$, source block n has been allocated to
$\mathcal{C}$	Continuous channel gain adaptation
$\mathcal{N}$	No direct channel gain adaptation
$\mathcal{S}$	Single scalar per region used to adapt to channel gain
$\mathfrak{B}$	Mapping-rate mismatch compensation on a per block basis
$\mathfrak{NB}$	Mapping-rate mismatch compensation on a per block basis is not used
ξ	Encoder for linear mapping
$B \times B$	Number of source samples in a block
C	Channel capacity in bits/channel symbol
c	Speed of light
D	Statistical sample size
$E[]$	Expectation operator
f_c	Carrier frequency
f_m	Doppler frequency
$f(\gamma)$	Probability density function of γ
$g(k)$	Signal after mapping
$\hat{g}(k)$	Signal before demapping
I_m	Set of blocks allocated to channel state m
J	Number of available mappings for implemented coder
K	Total number of channel samples
k	Time index
K_j	Channel dimension of a mapping
K_{out}	Number of channel samples in outage
L	Total number of source samples
$\mathcal{L}$	Object function
L_j	Source dimension of a mapping
$M+1$	Total number of channel states
N	Total number of source blocks
$n(k)$	Noise sample k
p_{o}	Probability of channel being in outage
p_{t}	Probability of transmitting information
p	Probability
p_m	Probability for channel state m

$\hat{p}_m$	Probability for preallocated channel state m
R	Rate of source in bits/source sample
$\mathcal{R}_m$	CSNR region m
R_c	Channel rate in bits/channel symbol
R_s	Source rate in bits/channel symbol for reference coder
r_{avg}	Total ratio of number of channel symbols/ number of source symbols
r_{avg_t}	Total ratio of number of channel symbols/ number of source symbols for single simulation
$\hat{r}_n$	Mapping rate block n is allocated to
$\hat{r}_j$	Rate of implemented mapping j
$\hat{r}_n$	Rate of implemented mapping block n is allocated to
Rx	Receiver
$s(k)$	Transmitted signal
S_x	Horizontal image size, in pixels
S_y	Vertical image size, in pixels
T_s	Transmitted symbol duration
Tx	Transmitter
v	Relative speed of mobile
w_m	Channel gain mismatch equalizer factor for channel state m
$x(k)$	Source sample at time instance k
$\hat{x}(k)$	Decoded source sample at time instance k
$x(i,j)$	Original value of image pixel at position (i,j)
$x_{\min}$	Minimum possible pixel value
$x_{\max}$	Maximum possible pixel value
$\hat{x}(i,j)$	Decoded value of image pixel at position (i,j)
$y(k)$	Received signal

Chapter 1

Introduction

Wireless devices are used in an increasingly larger part of our everyday life. In more and more areas habits are shifting from stationary to mobile. Currently there are more portable than stationary computers being sold, but maybe the area that has gone through the largest change is telecommunications. In relatively few years the mobile phone has been transformed from a rather big contraption confined to the car, to a small device.

From being a thing for the few, it has grown to be something you have to be rather unusual not to have. From an steadily younger age, most people get their own mobile phone, they rely on being available at all times. Being unavailable is a thing of the past. The traditional fixed phone is even losing ground at home, being replaced by the mobile phone.

The mobile phone industry is also one of the busiest industries world wide, with rapid changes over very few years. Both the telephony service providers and the mobile phone manufacturers invest huge amounts of money in attracting costumers to their products. As an illustration on how much is at stake, the auctions for the licenses to use the frequency spectrum assigned to the *third-generation mobile technology* (3G) system in Europe ended up at about \$ 100 billion, (10^{11}), American dollars [Goldsmith, 2005]. So naturally the companies getting these licences would like to get the most out of their spectrum. The 3G case has shown how much spectrum is perceived to be worth these days, maybe making the hunt for bandwidth efficient systems even more important than it has ever been.

Even though making phone calls still is the most important function of a mobile phone, other functions are increasing in popularity. At the time of writing, cameras on mobile phones are used actively for taking still pictures, and even videos are catching on. Listening to music and the use of miscellaneous Internet services are also functions that are being used increasingly by mobile phone owners.

One of the largest problems with many of these functions is that they use power. Typically, the more complex a function is, the more power it uses. At the same time, as long as wireless devices are wireless, they rely on batteries. So to avoid having to charge batteries to often, the different functions need to use as little power as possible.

When transmitting data from one place to another it is wise to consider the content of the data. Data can be divided into two groups. On the one hand there is information that is digital in nature, such as documents, text files or computer programs. On the other hand, there is data that is analog in nature, e.g. sound/speech and images. The first group is very sensitive to errors. Even a few errors can make a text document unreadable or prevent a program from running. The correctness of these data is so essential that it might be a good strategy to retransmit the data if errors are found. For the second group the data is more *robust* towards errors, meaning that it can tolerate more errors yet be meaningful for the receiver. Often there are also other *quality of service* (QoS) *constraints* on the transmission. In the case of video there is usually at least a time constraint on the transmission; the viewer would notice the delay exceeding this constraint in the stream even if every frame is received error free. As long as the mobile device is wireless there will also, as mentioned above, be a power constraint. So the principle of practical data transmission is not just minimizing the errors. How the receiver perceives the data should also be taken into consideration. As an example Figure 1.1 shows four examples of the same image transmitted with errors using four different approaches. Since each image is coded and transmitted with a different approach, the errors will appear differently for each image.

1.1 Source and channel coding

From the above discussion it is easy to see that channel bandwidth is expensive, and it is highly desired to use as little as possible. To be able to transmit an information signal over a channel, *compression* is needed, either in terms of bandwidth or equivalently, data-rate. As the general communication channel is noisy the transmitted data has to be *protected* in some sense.

A general communication system can be modeled as in Figure 1.2. The transmitter encodes the information, and the receiver decodes the information. Traditionally these coding operations are split further into two separate operations in a *tandem* structure, as shown in Figure 1.3. The purpose of *source coding* is to remove any redundancies, and describe the information source in an efficient manner. The *channel encoder* on the

(a) JPEG (b) JPEG2000

(c) WTSOM (d) Proposed

Figure 1.1: Example of errors effects that appear in images for four different strategies. Images (a), (b) and (c) reprinted from [Boeglen and Chatellier, 2006] with permission.

other hand, introduces redundancy whose purpose is to provide protection against channel noise. Claude Shannon proved through the *separation principle* that the source coder and channel coder can be optimized separately and still achieve a jointly optimal system [Shannon, 1948a, 1959]. The lowest rate a source can be coded with for a given fidelity, is given by the *rate distortion* function $R(D)$. The channel *capacity* C, gives the maximum amount of information that can be transmitted error free over a given channel. So as long as $R(D) < C$, information can in principle be transmitted reliably from a transmitter to a receiver.

In the following the ideas of separate source coding and channel coding will be further presented. After that the motivation and principle for *joint source-channel coding* (JSCC) will be given.

Figure 1.2: General communication system

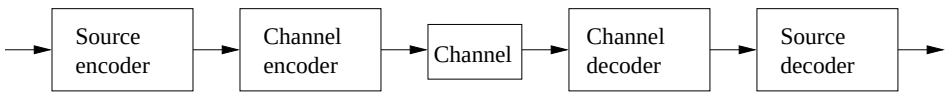


Figure 1.3: Communication system with separately designed source/channel coders

1.1.1 Source coding

The process of representing an information source is denoted source coding. For all practical scenarios this will be to represent the information with as little resources as possible for a given fidelity. For most schemes resources are given in terms of bits, but in general other ideas might be used as well.

All information sources such as images, sound, speech and such, will have samples which are *correlated*, one sample contains some of the same information as others. This redundancy can be removed without affecting the information content. For efficient coding some sort of decorrelation mechanism must thus be present. This can be done through some sort of *whitening* process, either in the time domain by predicting future samples, e.g. *differential pulse code modulation* (DPCM) [Gersho and M. Gray,

1992], or in the frequency domain by using a *filter bank*[1] or transform, such as in *joint photographic experts group* (JPEG) [ISO/IEC, 1991] or JPEG2000 [ISO/IEC, 2000].

Since the human senses are not perfect, some distortion of the signal can be tolerated as it will not be noticed. Such distortion is denoted *irrelevancy*.

The process of removing redundancy and irrelevancy is not enough if the needed compression of the signal is high. As long as the original signal is analog, the coding has to be *lossy*. Some information is lost in the coding process, and the representation is a distorted version of the original. This process is usually done through quantization [Gersho and M. Gray, 1992].

1.1.2 Channel coding

All practical channels are noisy. This means that to transmit information over a channel, the information will have to be protected against transmission errors. The purpose of the channel codes is to protect the source from the noise on the channel by adding redundancy in a clever way that matches a given channel. The receiver is then better equipped to detect the transmitted symbols, after they have been altered by noise.

There are many different channel models used, both for wired and wireless channels. For the case of wired transmission, the channel is usually considered to have a fixed quality for a given connection, so the need for adaptation is relatively low. Wireless channels on the other hand, change conditions relatively often for mobile connections. To be able to deal with such a channel, one possibility is to construct very robust channel codes, capable of dealing with the varying conditions. Another possibility is to adapt to the channel as it changes, using codes optimized for different channel conditions, and in this way being able to transmit more data on average.

1.1.3 Regarding tandem coding

Designing source and channel codes separately has many advantages. Some of them are listed below.

- When designing the source code, it is possible to ignore the channel. The same applies for the design of the channel code, focus can be kept on one thing at the time.

- Separate design makes a portable system, since the source can be coded separately of the channel, the source can be coded once, and

[1] The term filterbank is a very wide term, including more specific designs such as the *wavelet transform [Vaidyanathan, 1993]*

be transmitted over different channels by changing the channel code.

- It is proven that an optimal system can be achieved by separate design of source and channel code [Shannon, 1948b].

There are however a few drawbacks with these points. The optimality of the separate design is for systems allowing infinite *complexity* and infinite *delay*. Higher complexity implies higher power consumption, which further implies higher need for battery. The delay considerations are most important for real-time systems.

Example of good channel codes, are the *low-density parity-check* (LDPC) codes [Gallager, 1963], which in the later years have been gaining popularity since they were rediscovered in [MacKay and Neal, 1996]. It was demonstrated in [Chung et al., 2001] that one can come only 0.04 *decibel* (dB) away from the theoretical capacity limit for a small *bit error rate* (BER) (10^{-6}) on a binary input *additive white Gaussian noise* (AWGN) channel by using LDPC codes. This is however at the cost of a block length of 10^7, where the block length refers to the number of bits transmitted on the channel that are dependent of each other. For such long block lengths the delay can be quite significant. Other popular channel codes, e.g. the turbo codes [Berrou et al., 1993], also rely on long block-lengths and high computational cost.

These concerns about separate source and channel coder design, make it appealing to also consider a joint approach.

1.2 Related work on JSCC

The use of the term *joint source-channel coding* (JSCC) is very wide. In the literature, any system involving cooperative interaction of information considering the source and channel can be called a JSCC scheme. There is a plethora of solutions on how this is done in practice, some of which will be considered in the following sections. Ways of classifying these different schemes are almost as many as there are schemes, but in this thesis, different JSCC schemes are classified into three categories, to be explained further in the following.

Digital systems

The first category is fully digital systems, where the source and channel code are designed using information about each other. Some of the earliest work on JSCC belongs to this category. In [Fine, 1964], a communication system with a discrete noisy channel was considered, and a framework for optimizing encoders and decoders was presented. Later a scalar quantizer was optimized for noisy channels in [Kurtenbach and Wintz, 1969],

where the quantizer design includes channel noise. This work was improved in [Farvardin and Vaishampayan, 1987], and extended to vector quantization in [Farvardin, 1990; Farvardin and Vaishampayan, 1991], denoted *channel optimized vector quantizer* (COVQ). This work was then extended to include power constraints in [Fuldseth and Ramstad, 1997; Fuldseth, 1997].

Most conventional source coders generate a bit stream where some bits have greater impact on the decoded source than others. An intuitive idea is to apply stronger protection on the important bits. This technique of varying the channel code depending on the importance of the bit stream is commonly referred to as *unequal error protection* (UEP). The authors of [Modestino and Daut, 1979] had an early treatment of UEP, but came to the conclusion that more flexible code rates were needed for JSCC design. Later work picked up this thread, and *rate compatible punctured convolution* (RCPC) codes [Hagenauer, 1988] gave such a possibility and became a popular channel code for UEP schemes. RCPC channel codes provide an easy way of changing the protection through puncturing, and has been used in [Tanabe and Farvardin, 1992], where a RCPC code was matched to a subband source coder. The optimal bit allocation and channel code rate was found through an exhaustive search. In [Goldsmith and Effros, 1998] an iterative approach was used to find the best choice of COVQ and RCPC code design for an AWGN channel by minimizing the average end-to-end distortion. This work was extended to a fading channel in [Tie et al., 1998]. Later work has included RCPC codes in connection with *cyclic redundancy check* (CRC) [Sun and Xiong, 2006].

Another related approach for JSCC was presented in [Kozintsev and Ramchandran, 1998] and [Zheng and Liu, 1999], where the concept of multi-resolution modulation was used. By using sub-constellations within a constellation, it is possible to place important information on the main constellation, and less important information on the sub-constellations.

Hybrid Digital Analog

Traditional tandem systems use a channel code that is designed for a worst case *channel signal-to-noise ratio* (CSNR) where it can guarantee an error rate below a certain level. A source coder is then matched to the bit rate the channel code is designed for, resulting in a certain distortion for the transmitted source. If the true CSNR level on the physical channel is exactly the same as the CSNR the system is designed for, everything is fine. However, if there is a mismatch, there are two scenarios. If the true CSNR is larger than the CSNR the system is designed for, the system suffers from what is called the *leveling-off effect*. As the true CSNR rises,

the rate the channel supports increases, but the performance of the system remains constant after a certain threshold. This is due to the lossy part of the source coder. The design of the quantizer sets a certain distortion-level which yields the target rate of the channel code. It is not possible to track the increase of the rate on the channel without redesigning the source coder. Should the true CSNR be lower than the design-CSNR, the system would instead experience what is called the *threshold effect*. This is due to error correcting capabilities of the channel code breaking down very fast if the channel noise is larger than expected. For traditional source coders, this would then lead to a huge error, due to the nonlinearities in the quantizers.

The leveling-off effect is targeted in *hybrid digital-analog* (HDA) systems. By keeping an analog component the system can track the improved channel condition. Whereas digital systems have discrete jumps for different designs, the analog part will ensure *graceful improvement*. Examples of such systems is given in [Mittal and Phamdo, 2002; Coward and Ramstad, 2000c; Skoglund et al., 2002; Wang et al., 2005; Skoglund et al., 2006].

Fully analog systems

HDA systems target the leveling-off effect, but there is still the problem of the threshold effect. One of the characteristics of *frequency modulation* (FM)-radio, is that it has graceful degradation; as the channel quality is lowered, the reception becomes poorer and poorer, but it is still possible to get information through. A fully digital system, does not behave in the same manner. In a digital audio broadcasting system, e.g. *digital audio broadcasting* (DAB) [DAB, 1997], the audio quality will be very good as long as the reception is good enough. If the channel quality drops below a certain level, the sound drops out, it is all or nothing. The *robustness* is greater for an analog system than for a digital one. The problem with FM, is that it is much less bandwidth efficient than a digital system e.g. DAB.

In his paper [Shannon, 1949], Shannon suggested the use of *nonlinear source-channel mappings* by looking at the problem of transmitting source samples on a channel in a geometrical perspective. Each point in the source space correspond to a point in the channel space. The mapping is a translation between these two spaces. These mappings can be used for both bandwidth compression and expansion, depending on the target fidelity of the source. Independently of Shannon, Kotel'nikov developed theory for a similar kind of system, with focus on bandwidth expansion [Kotel'nikov, 1959]. Due to the analog nature of a system using such mappings, they will be robust towards channel variations, both with regard to the threshold effect, and the leveling-off effect. Such mappings were discussed further for bandwidth expansion in [McRae, 1971; Thomas et al., 1975].

After Shannon and Kotel'nikov many similar mappings have been developed. The mappings developed in [Fuldseth and Ramstad, 1997; Fuldseth, 1997; Coward, 2001] still contained a discrete part, but lead to later development of fully analog continuous mappings for both bandwidth expansion and compression [Floor and Ramstad, 2006b,a,c], and a thorough analysis of one particular bandwidth compression mapping in [Hekland et al., 2005]. Analysis on how such a mapping could connect with the backbone network is done in [Hekland and Ramstad, 2006b,a].

1.3 Outline of the thesis

The remainder of the thesis is organized as follows

Chapter 2 The theoretical models needed in the proposed system are presented. Parameters needed in later chapters are also defined here.

Chapter 3 The proposed image transmission system is presented. The structure and different parts of the system are described.

Chapter 4 A model to estimate the theoretical limit of the proposed system is presented. A framework to analyze the theoretical improvement possible for the proposed system by developing extra nonlinear dimensional changing mappings is also presented.

Chapter 5 Simulations and discussion of the results are done for the proposed system for a Rayleigh flat fading channel. Comparison with reference systems is also done.

Chapter 6 Conclusions are drawn from the results and discussions of the previous chapters. The main contributions of the thesis are presented. Ideas for future research topics are given.

In addition to the main chapters, there are five appendixes. Appendix A gives a small proof on the optimal allocation of source-blocks to channel states. In Appendix B the channel gain mismatch receiver filters are found. Appendix C presents issues with numerical optimization regarding the channel regions and representation points. Appendix D have plots of the used regions and representation points when the CSNR range is split in $M = 1$ and $M = 2$ regions. In Appendix E the original images used in the simulations in the thesis are given as a reference.

Chapter 2

Theoretical aspects of Joint Source-Channel coding

In this chapter, most of the theoretical background needed in the thesis is presented. In Section 2.1 the framework concerning the source is presented, and some theoretical limits are given. Section 2.2 presents theory concerning the transmission medium, the channel, for two different models. The AWGN channel is presented in Section 2.2.1, and properties and capacity of a fading channel are presented in Section 2.2.2. In Section 2.3, the theoretical limits for channel and source are combined.

2.1 Describing the source

Rate-distortion theory is the theory covering the minimum resources needed to describe an amplitude-continuous source X with known distribution, for a given distortion, σ_D^2. For most distributions this is unfortunately not known, but for Gaussian sources this is known. The variance of a source X is found by

$$\sigma_X^2 = E(X^2), \tag{2.1}$$

when the mean value is zero, and where E is the expectation operator. The rate distortion function for a white Gaussian source with variance σ_X^2, is given by [Berger, 1971],

$$R = \begin{cases} \frac{1}{2}\log_2\left(\frac{\sigma_X^2}{\sigma_D^2}\right), & \text{if } 0 \leq \sigma_D^2 \leq \sigma_X^2 \\ 0, & \text{if } \sigma_D^2 > \sigma_X^2, \end{cases} \tag{2.2}$$

measured in bits/source sample, where the block distortion σ_D^2 is found by the *mean square error* (MSE)

$$\sigma_D^2 = E[(X - \hat{X})^2], \tag{2.3}$$

where $\hat{X}$ is the decoded estimation of the source X. In practice, the source variance is estimated by

$$\sigma_X^2 = \frac{1}{L}\sum_{l=1}^{L} x^2(l), \tag{2.4}$$

and the distortion is estimated by

$$\sigma_D^2 = \frac{1}{L}\sum_{l=1}^{L} (x(l) - \hat{x}(l))^2, \tag{2.5}$$

where L is the length of the source sequence, x, and $\hat{x}$ is the decoded source sequence. This is correct for ergodic sources when $L \to \infty$.

In the case of multiple *independent and identically distributed* (i.i.d.) white Gaussian sources with variances $\sigma_{X_1}^2, \ldots, \sigma_{X_N}^2$, the average rate is given by [Berger, 1971]

$$R = \frac{1}{N}\sum_{n=0}^{N-1} \frac{1}{2}\log_2\left(\frac{\sigma_{X_n}^2}{\sigma_{D_n}^2}\right) \text{bits/source sample}, \tag{2.6}$$

where

$$\sigma_{D_n}^2 = \min(\mu, \sigma_{X_n}^2), \tag{2.7}$$

and μ is the distortion-level for sources with $\sigma_X^2 \geq \mu$. This implies that only the sources with variance larger than μ is represented. Sources with $\sigma_X^2 \geq \mu$, all get an equal distortion $\sigma_D^2 = \mu$, leading to what is called *reverse water filling* [Cover and Thomas, 1991].

The *signal-to-noise ratio* (SNR) for one source is given in dB by

$$\text{SNR} = 10\log_{10}\left(\frac{\sigma_X^2}{\sigma_D^2}\right). \tag{2.8}$$

A composite source can be decomposed into multiple sub-sources. If this is done by a filterbank, and if the overall impulse response of the synthesis filters is one, the additive white noise will appear as additive noise of the same variance on the output. If the passbands are relatively flat, the noise

will be approximately white. Assuming i.i.d. sub-sources, a total SNR in dB can be defined by

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\sum_{n=0}^{N-1}\sigma_{X_n}^2}{\sum_{n=0}^{N-1}\sigma_{D_n}^2}\right). \tag{2.9}$$

So for a given SNR, it is possible to find the corresponding distortion-level μ. So in a setting with multiple i.i.d. sub-sources it is possible to set a given SNR value, and find the corresponding distortion-level μ.

2.2 Channel description

A medium used to transmit information from a transmitter to a receiver is called the channel. Depending on the physical channel, the transmitted signal will be perturbed differently from channel to channel, so different models are used to be able to analyze a given system. Two models for a wireless channel are the AWGN channel and the fading channel[1]. The capacity of a channel is defined as the *mutual information* maximized over all input distributions.

$$C = \max_{f(s)} I(S;Y), \tag{2.10}$$

where the mutual information for amplitude-continuous variables is defined as

$$I(S;Y) = \int\int f(s,y)\log_2\frac{f(s,y)}{f(s),f(y)}dsdy, \tag{2.11}$$

where S is the transmitted random variable, Y is the received random variable, $f(s)$ is the *probability density function* (pdf) of s, and $f(s,y)$ is the joint pdf [Cover and Thomas, 1991].

Shannon's theory [Shannon, 1948a, 1949] have shown that a code exists, that can achieve a data rate R_c, close to the capacity C with error probability going towards zero. On the other hand, any code with error probability going towards zero must have data rate smaller than the channel capacity. In other words, this means that there exists a code that can transmit C bits/channel symbol, with arbitrary small probability of error over a given channel.

2.2.1 AWGN channel model

In an AWGN channel model, the received signal $y(k)$ can be written as

$$y(k) = s(k) + n(k), \tag{2.12}$$

[1]There are many models for fading channels, but for now a general fading model is assumed.

where $s(k)$ is the transmitted signal and $n(k)$ is a white (uncorrelated) random noise process with a Gaussian distribution. The capacity is given by the well-known formula given by Shannon in [Shannon, 1948a]

$$C = \frac{1}{2}\log_2(1+\gamma) \text{ bits/channel symbol,} \tag{2.13}$$

and γ is the CSNR given by

$$\gamma = \frac{\sigma_S^2}{\sigma_N^2}, \tag{2.14}$$

where σ_S^2 is the power of the received signal, and σ_N^2 is the noise power within the transmitted bandwidth.

2.2.2 Fading channels

When transmitting signals over a wireless channel, an AWGN model is often not sufficient. Electromagnetic waves will experience reflection, scattering and diffraction due to obstacles they may encounter such as buildings, terrain or other objects. These objects may then create additional copies of the signal known as *multipath components* which can be attenuated, delayed and shifted in phase with respect to the original transmitted signal. In this thesis the focus will be on fluctuations in the received power due to constructive and destructive multipath components. The result of this in the received signal is known as *fading*[2]. Determining these effects by calculating the path of each component would be to complex in practice. Hence, statistical modelling is used instead [Goldsmith, 2005].

In addition to the natural time variation of a channel on its own, the channel will vary depending on the relative motion between the mobile and the base station, or the motion of objects in the channel. When moving, the received signal at either the base station or mobile will experience a broadening of the frequency spectrum called *Doppler spread*. The time dual to this phenomenon is called *coherence time*. If the coherence time is greater than the channel symbol period, it means that the channel variations are slower than the baseband signal variations, and *slow fading* is occurring [Rappaport, 1996]. In other words, the channel will not change during the transmission of a channel symbol. Throughout this thesis a slow fading channel will be assumed. The impact of the speed of the mobile is expressed through the maximum Doppler shift, or maximum Doppler frequency f_m. The Doppler shift can be found through the well known relation

$$f_m = \frac{v f_c}{c} \text{ Hz,} \tag{2.15}$$

[2]More correct small scale fading for the scenario in this thesis.

where v is the relative speed of the mobile and base station, f_c is the carrier frequency, and c is the speed of light ($3 \cdot 10^8$ m/s).

If a system has signal bandwidth much smaller than the inverse delay spread, i.e. the coherence bandwidth, it is called narrowband. Being narrowband means that all the frequency components are attenuated equally within the signals bandwidth, giving *flat fading*.

Consider a frequency-flat slow fading channel where it is assumed that the received signal $y(k)$ can be written as

$$y(k) = \sqrt{\alpha(k)}s(k) + n(k), \tag{2.16}$$

where $\sqrt{\alpha(k)}$ is the ergodic and stationary channel gain, $s(k)$ is the sent signal and $n(k)$ is AWGN. The channel power gain, $\alpha(k)$, will be assumed to be independent of the channel input and have an expected value of unity.

Let $\bar{\sigma}^2_{S_e}$ denote the long term average transmission power and σ^2_N be the power of the in-band noise on the channel. Assuming no power allocation, the instantaneous *pre-adapted* CSNR, $\gamma(k)$, is defined as

$$\gamma(k) = \frac{\bar{\sigma}^2_{S_e}\alpha(k)}{\sigma^2_N}. \tag{2.17}$$

The expected value of the CSNR, $\bar{\gamma}_e$, is then

$$\bar{\gamma}_e = \frac{\bar{\sigma}^2_{S_e}}{\sigma^2_N}. \tag{2.18}$$

Looking at equation (2.17) it can be seen that the gain $\alpha(k)$ and CSNR $\gamma(k)$ are connected through a multiplicative constant. Hence, $\gamma(k)$ is distributed with the same distribution as $\alpha(k)$. In the case that $\sqrt{\alpha(k)}$ follows a Rayleigh distribution, $\alpha(k)$ will follow an exponential distribution and $\gamma(k)$ will also be distributed with an exponential distribution [Stüber, 2001]

$$f(\gamma) = \frac{1}{\bar{\gamma}_e}\exp^{-\frac{\gamma}{\bar{\gamma}_e}}. \tag{2.19}$$

Assuming a specific time instance, k, the *instantaneous* channel capacity is given by

$$C(k) = \frac{1}{2}\log_2\left(1 + \gamma(k)\right), \tag{2.20}$$

which for simplicity will be written without the time reference k for the remainder of the thesis

$$C(\gamma) = \frac{1}{2}\log_2\left(1 + \gamma\right). \tag{2.21}$$

By rewriting (2.17), an expression for α for a given pre-adaptation CSNR value is given by

$$\alpha = \frac{\sigma_N^2 \gamma}{\bar{\sigma}_S^2}. \tag{2.22}$$

2.2.3 Channel model

Traditionally a complex baseband model is assumed. This indicate that a complex *multilevel quadrature amplitude modulation* (M-QAM) signal is transmitted. In this thesis real *pulse amplitude modulation* (PAM) signals will be used, indicating that the phase carries no information. In the model for Rayleigh fading, the in-phase and quadrature component can, by the use of the central limit theorem, be approximated by two independent white Gaussian processes. The envelope of the sum of the in-phase and quadrature component will then be Rayleigh distributed, e.g [Rappaport, 1996; Goldsmith, 2005]. In this thesis it will be assumed that the transmitted signal still experiences Rayleigh fading, even though the transmitted signal is real. For a PAM signal there is no information in the quadrature component, but the carrier will still have a phase, contributing to Rayleigh distribution of the envelope. Assuming that the receiver has full channel information, the phase-error will be compensated for, denoted *coherent* detection, avoiding a shift of sign in the decoded signal that might happen if the phase-error is large enough. The noise can be complex with independent in-phase and quadrature component, but since the receiver has full channel knowledge, only the in-phase component distorts the information signal.

2.2.3.1 Capacity

The capacity of a fading channel is dependent on the amount of information at the receiver and transmitter. In this thesis two different scenarios will be looked at. For both cases it is assumed that both the receiver and transmitter know the distribution of the CSNR, $f(\gamma)$. But both do not necessarily know the instantaneous value of the CSNR of the channel, γ, at all times. This information is called the *channel state information* (CSI). For a practical system, the CSI has to be estimated by the receiver and can be transmitted back to the transmitter through a return-channel. For simplicity the return-channel is assumed to be delay- and error free.

All the systems in this thesis are assumed to only have one transmitting antenna and one receiving antenna, so no diversity techniques are used.

Assuming that the transmitter does not have any CSI means that the transmitter can not adapt to the channel as it varies. The data rate that

can be transmitted over the channel is constant with value depending on the power available. For a traditional channel coding system, this means that the channel code has to be long enough to be able to cope with the deep fades encountered. Due to the capacity being found through averaging over the pdf of γ, it is also denoted the *ergodic capacity*, and is given in [Caire and Shamai, 1999], from results in [McEliece and Stark, 1984] as

$$C = \frac{1}{2}\int_0^\infty \log_2(1+\gamma)f(\gamma)d\gamma. \tag{2.23}$$

As can be seen from the integration limits, this system will transmit all the time, and hence has no *outage*, i.e. periods where no data are received. It will be up to the receiver to combat the problems occurring from a deep fade.

By letting the transmitter have information about the channel conditions, it can adapt to the channel by using power and rate adaptation, and choose not to transmit if the channel gets too bad, thus saving power for better channel conditions. Under a power constraint given by

$$\int_0^\infty \sigma_S^2(\gamma)f(\gamma)d\gamma \leq \bar{\sigma}_{S_e}^2, \tag{2.24}$$

the capacity can be found by maximizing

$$C = \frac{1}{2}\int_0^\infty \log_2(1+\frac{\sigma_S^2(\gamma)}{\bar{\sigma}_{S_e}^2}\gamma)f(\gamma)d\gamma, \tag{2.25}$$

with respect to $\sigma_S^2(\gamma)$ [Goldsmith and Varaiya, 1997]. The optimal power adaptation is given by

$$\frac{\sigma_S^2(\gamma)}{\bar{\sigma}_{S_e}^2} = \begin{cases} \frac{1}{\gamma_0} - \frac{1}{\gamma}, & \text{if } \gamma \geq \gamma_0 \\ 0, & \text{if } \gamma < \gamma_0. \end{cases} \tag{2.26}$$

The solution leads to a *water-filling* in time as the γ level varies. When the channel conditions are good, it is beneficial to use more power than when the channel conditions are poor. In equation (2.26) there is however still one unknown. The optimal outage level γ_0 has to be found numerically, which can be done by changing the inequality in equation (2.24) to an equality and inserting equation (2.26) into it, giving that the optimal γ_0 has to satisfy

$$\int_{\gamma_0}^\infty \left(\frac{1}{\gamma_0} - \frac{1}{\gamma}\right) f(\gamma)d\gamma = 1. \tag{2.27}$$

The resulting expression for the capacity can then be found by substituting equation (2.26) into equation (2.25),

$$C = \frac{1}{2}\int_{\gamma_0}^\infty \log_2\left(\frac{\gamma}{\gamma_0}\right) f(\gamma)d\gamma. \tag{2.28}$$

When not adapting the channel power, it was shown in [Goldsmith and Varaiya, 1997] that the capacity for a fading channel with CSI at both transmitter and receiver is equal to the capacity for a fading channel with only CSI at the receiver, (given by (2.23)). The difference between the capacity of a channel with CSI at the receiver only, and the capacity of a channel having CSI at both receiver and transmitter, given in Figure 2.1, is only due to power adaptation. For most practical considerations, this gap is insignificant, but to the best of the authors knowledge, there are currently no good code designs that achieve rates close to the capacity given by (2.23) [Alouini and Goldsmith, 2000].

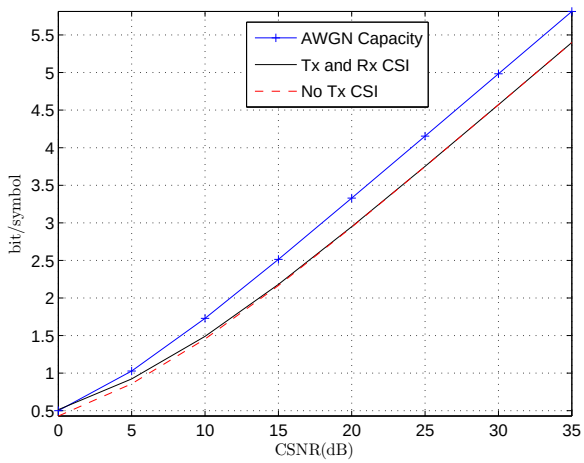


Figure 2.1: Comparison between capacity of an AWGN channel, and different channel capacities for a Rayleigh fading channel.

2.2.3.2 Channel adaptation techniques

Design of practical systems that approach the capacity of fading channels will make use of techniques indicated above. No-outage systems will have to use robustness, often including long codewords. For correlated channels, an interleaver is often used to spread the impact of deep fades over multiple symbols.

For systems with CSI, there are several cases possible. The transmission rate can be adapted, the power can be adapted, or both.

A common technique to increase the *average spectral efficiency* (ASE) for a fading channel, is to split the CSNR range into multiple regions, and to use different settings of the transmitter within each region. An example

of this is given in Figure 2.2. Each region $\{\mathcal{R}_m\}_{m=0}^{M}$, is defined to be the pre-adapted CSNR range of $[\gamma_{T_{m-1}}, \gamma_{T_m})$, see equation (2.17). $\{\gamma_{T_m}\}_{m=0}^{M}$ will be denoted *threshold values* throughout the thesis. $\mathcal{R}_0$ is the region where the system is in *outage*, $\gamma = [0, \gamma_{T_0})$. For convenience $\gamma_{T_M} = \infty$. For traditional channel transmission systems, it is possible to use *adaptive coded modulation* (ACM), where the channel codes, modulation, or both, is changed depending on the region the channel is in. Either for a given BER in the case of a set of channel codes, modulation constellations and transmitted power [Goldsmith and Chua, 1997; Hole et al., 2000; Vishwanath and Goldsmith, 2000, 2003], when using a fixed number of capacity achieving channel codes with fixed power [Holm et al., 2003], or variable transmit power [Gjendemsjø et al., 2005; Lin et al., 2006]. Common for these systems is that they use codes that require a certain CSNR level γ_{C_m}, to transmit reliably. If the received CSNR is lower than γ_{C_m}, the channel codes will break down. When using power adaptation, γ_{C_m} can be increased within each region through

$$\gamma_{C_m} = \frac{\sigma_S^2(\gamma)}{\bar{\gamma}_e}\gamma. \tag{2.29}$$

γ_{C_m} will be denoted as the *representation point* of the m'th channel region.

In practical systems, it is not possible to continuously adjust the transmission power, but in [EIA/TIA-95, 1989] the transmission power can be adjusted in 1 dB intervals over a dynamic range of 60 dB, which makes the flexibility of the power adjustment high.

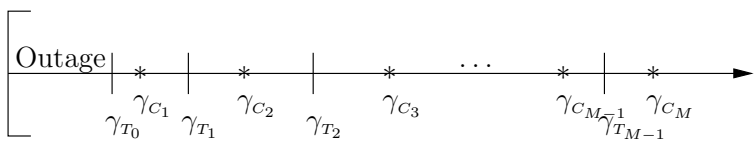


Figure 2.2: CSNR range divided into regions with thresholds $\gamma_{T_m}, m = 0, \ldots, M-1$ $(\gamma_{T_M} = \infty)$, with corresponding coded CSNR,$\gamma_{C_m}, m = 1, \ldots, M$.

2.3 Optimal systems

When designing communication systems, the optimal performance is very interesting, if it is known. By comparing with the optimal system, the strengths and weaknesses of different practical systems can be seen, or even what parts of a given system that contribute the most towards achieving a limit-approaching system.

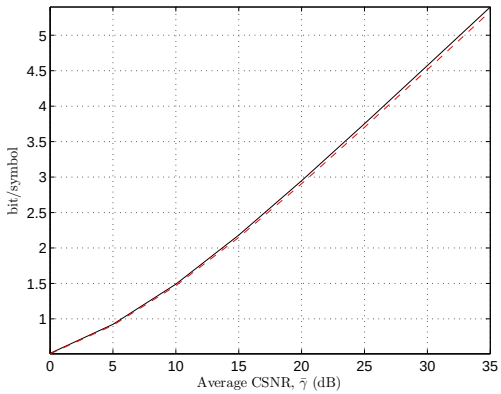


Figure 2.3: Channel transmission rate, R_c, of system in [Gjendemsjø et al., 2005] with four channel regions and infinite power level adaptation compared with channel capacity for Rayleigh fading channel with Tx and Rx CSI.

2.3.1 Optimal performance theoretical attainable

The *optimal performance theoretically attainable* (OPTA) is the theoretical limit for any practical communication system transmitting a given source over a given channel with respect to given fidelity criterion [Berger and Tufts, 1967]. In essence OPTA is a combination of the rate distortion function of a source, and the channel capacity of a channel. It can be given for an arbitrary bandwidth change, r_{avg}, between source and channel, or the ratio between the number of source samples and channel samples.

OPTA can be generally expressed by

$$r_{\text{avg}} = \frac{R(\text{SNR})}{C(\text{CSNR})} \text{channel samples/source samples}, \tag{2.30}$$

where $C(\text{CSNR})$ is the channel capacity, and $R(\text{SNR})$ is the rate distortion function for a given source.

OPTA can be written in many ways. One is as the upper bound of the SNR resulting from transmitting a source over a channel, as a function of the CSNR of a given channel under a bandwidth change r_{avg}. A second case looks at the CSNR as a function of the SNR for a given bandwidth change r_{avg}, and gives the lower bound for the CSNR needed to be able to transmit a given source with a certain SNR through the channel. A third variant is the bandwidth change needed for a given SNR and CSNR.

For the case of a white Gaussian source transmitted over an AWGN

channel, the three different versions of OPTA are given by combining equation (2.30) with (2.13) and (2.2),

$$r_{\text{avg}}(\gamma, \sigma_X^2/\sigma_D^2) = \frac{\log_2(\frac{\sigma_X^2}{\sigma_D^2})}{\log_2(1+\gamma)} \tag{2.31}$$

$$\frac{\sigma_X^2}{\sigma_D^2}(\gamma, r_{\text{avg}}) = (1+\gamma)^{r_{\text{avg}}}, \tag{2.32}$$

or

$$\gamma(\sigma_X^2/\sigma_D^2, r_{\text{avg}}) = (\frac{\sigma_X^2}{\sigma_D^2})^{1/r_{\text{avg}}} - 1. \tag{2.33}$$

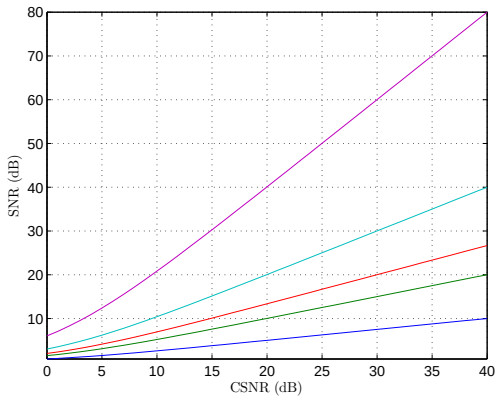


Figure 2.4: OPTA for a white Gaussian source transmitted over an AWGN channel, for different values of r_{avg}. From below: $r_{\text{avg}} = \{1/4, 1/2, 2/3, 1, 2\}$.

2.3.2 Linear vs nonlinear systems

An optimal transmission system with very low complexity and delay was suggested in [Goblick, 1965]. This system transmits a white Gaussian source over an AWGN channel by scaling the source signal by a constant ξ, and scaling with a different factor β in the receiver using continuous amplitude PAM symbols on the channel. If ξ and β are defined by

$$\xi = \frac{\sigma_S}{\sigma_N} \tag{2.34}$$

and

$$\beta = \frac{\sigma_X \sigma_S}{\sigma_N^2 + \sigma_Y^2}, \tag{2.35}$$

where $s(k)$ is the transmitted signal and $x(k)$ is the original signal, the transmitted signal is given by

$$s(k) = \xi x(k) = \frac{\sigma_S}{\sigma_N} x(k). \tag{2.36}$$

The received signal will be

$$y(k) = s(k) + n(k). \tag{2.37}$$

After the decoding, the reconstructed signal is given by

$$\hat{x}(k) = \beta y(k) = \beta(s(k) + n(k)) = \beta \left(\frac{\sigma_S}{\sigma_N} x(k) + n(k) \right), \tag{2.38}$$

and the MSE (distortion,σ_D^2) is then given by

$$\sigma_D^2 = E[(x(k) - \hat{x}(k))^2] = \frac{\sigma_X^2 \sigma_N^2}{\sigma_S^2 + \sigma_N^2} = \sigma_X^2 \left(1 + \frac{\sigma_S^2}{\sigma_N^2} \right)^{-1}. \tag{2.39}$$

Which is the same as equation (2.32) when $\gamma = \sigma_S^2/\sigma_N^2$ and $r_{\text{avg}} = 1$.

This system will be revisited in Section 3.2.2.3. The system structure is seen in Figure 2.5.

Lee and Petersen [Lee and Petersen, 1976], derived an optimum linear transform to transmit a vector source over a vector channel with AWGN. Later this system was named *block pulse amplitude modulation* (BPAM) [Vaishampayan, 1989]. The transmitter is simply a constant matrix, and the receiver is another constant matrix. When bandwidth reduction is wanted, the transmitter removes some samples altogether, and when bandwidth expansion is wanted, zeros are added. For the case of no bandwidth change, this system is the same as in [Goblick, 1965].

In [Berger and Tufts, 1967] it was pointed out that a linear PAM system can be made optimal for the case when there is no bandwidth change between source and channel when transmitting on AWGN channels. When there is bandwidth change between the source and channel, a linear system cannot be made optimal. An non-linear code was also suggested to combat this.

BPAM is also implemented for an image transmission system, i.e. in [Kafedziski, 1998] the design technique from [Lee and Petersen, 1976] is used for image transmission over fading channels when using a finite state channel model. This system is only considered for low CSNR values as the BPAM becomes increasingly suboptimal as the CSNR value increases.

Shannon pointed out in [Shannon, 1949], that to expand or compress bandwidth, there has to be a nonlinear connection between the source and channel space. This was confirmed in [Ramstad, 2006], where it was shown how a colored Gaussian source can be transmitted over a colored additive Gaussian channel by using *bandwidth matching functions* (BMF). These functions map a source-signal frequency range from the source spectrum to a frequency range in the channel spectrum. It turns out that in general these functions need to be nonlinear to achieve an optimal system. In this system the exception is again the case where the bandwidth of the source is equal to the bandwidth of the channel.

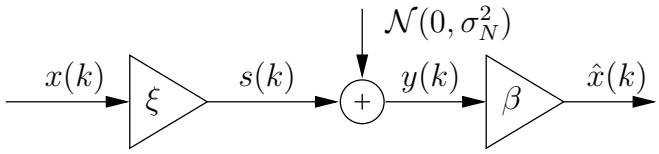


Figure 2.5: Optimal linear system for AWGN channel

Chapter 3

Image coder

Most of the image coders that are designed, are designed for error free channels. For a fading channel, a low complexity, robust and spectral efficient system is generally hard to design. This chapter presents a complete JSCC image transmission system for fading channels. Through the use of rate and power adaptation and by the means of nonlinear mappings, the system is very robust even with low complexity. By tracking the channel variations the system can be bandwidth efficient, and can be progressively decoded with little dependency between the channel symbols. By relying on the robustness of nonlinear mappings, the system gives a "as good as it gets" performance. The coder presented here is based on work presented in [Lervik, 1996], [Fuldseth, 1997] and [Coward, 2001].

The chapter is organized as follows. In Section 3.1 some previous image transmission systems using JSCC are presented. Section 3.2 presents the structure and properties of the different parts of the image coder system. Section 3.3 presents how a traditional tandem source and channel system can be used as a reference system without going through the notion of bits.

This chapter is partly based on [Håkonsen and Ramstad, 2005, 2006a; Håkonsen et al., 2006]

3.1 Previous work

One of the earliest image transmission schemes using JSCC was developed in [Modestino and Daut, 1979], for the case of DPCM they traded off rate for source and channel coding and used UEP for resulting symbols. Later they extended this for *discrete cosine transform* (DCT) in [Modestino et al., 1981]. [Kozintsev and Ramchandran, 1998; Zheng and Liu, 1999] proposed systems using multi-resolution modulation resulting in UEP through allocating important information to coarse constellation clouds, and less

important information to points within each cloud.

Another field that has received much attention is *progressive* image coding. The ability to start the decoding of an image before all the information is received is attractable in wireless communication. If the transmission stops unexpectedly, the previous received information can still be used. It is a "go with what we have" approach, which for most cases is much better than nothing. Based on the *set partitioning in hierarchical trees* (SPIHT) coder [Shapiro, 1993; Said and Pearlman, 1996], a progressive image JSCC transmission system using an inner RCPC code in concatenation with an outer CRC code was presented in [Sherwood and Zeger, 1997a,b] By changing the number of bits for the channel and source coder, the system can operate on different channel qualities. Later a general method to achieve the best trade-off between source and channel coding of video and images by using universal distortion rate characteristics was proposed [Bystrom and Modestino, 1998]. They showed the optimal rate allocation scheme as a function of channel statistics. One of the key feature of this progressive bit-stream is that it ceases to be useful past the first unrecoverable error [Nosratinia et al., 2003]. So the schemes based on the SPIHT coder will stop decoding after the first error is found. Different channels and channel codes are tried in [Nosratinia et al., 2003] (RCPC and CRC over *binary symmetric channel* (BSC)), [Thomos et al., 2005] (turbo coded and *Reed-Solomon* (RS) over fading channels) and [Pan et al., 2006] (LDPC over fading channels).

Other schemes are for example [Zhang et al., 2004a,b] where image transmission over Rayleigh fading channels is done by using adaptive coded modulation in connection with LDPC. The modulation is changed according to the state of the channel, and choosing channel regions is done by using performance curves of schemes used in each channel region for a given BER. [Boeglen and Chatellier, 2006] presented a JSCC scheme for image transmission over a flat Rayleigh fading channel. They use a wavelet transform together with a self organizing map *wavelet transform self organizing map* (WTSOM). To get added protection and correction they can choose to use an RS code, and an interleaver.

In the case of a robust quantization there has been work on index assignment in [Farvardin, 1990] and when using a COVQ [Farvardin and Vaishampayan, 1991]. This was later picked up to include power constraints in [Lervik and Fischer, 1997a; Lervik, 1996]. In [Coward and Ramstad, 2000b; Coward, 2001] this was extended to include a HDA part [Coward and Ramstad, 2000a]. The systems in [Lervik, 1996; Coward, 2001] were optimized for an AWGN channel. In this chapter a similar system is developed for a fading channel based on the same framework.

3.2 Coder structure

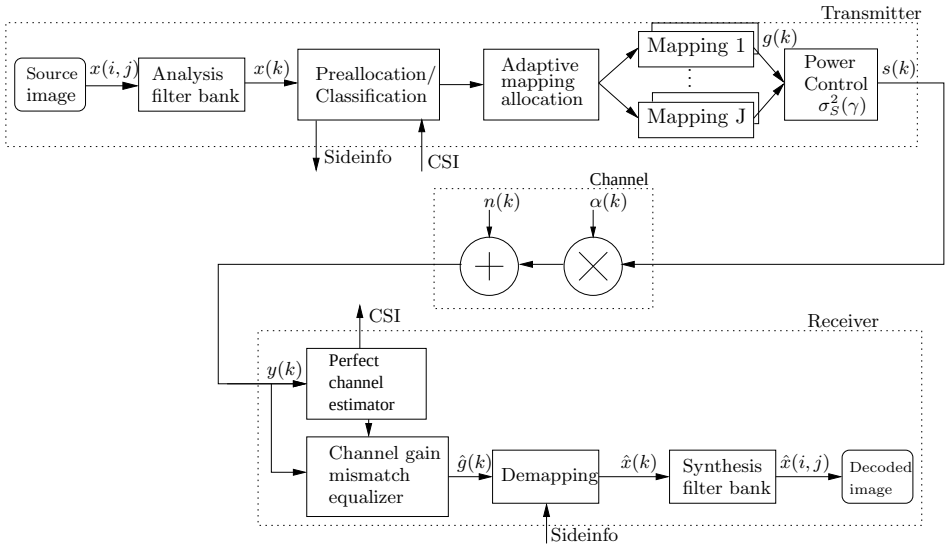


Figure 3.1: Proposed JSCC system

The structure of the proposed system is given in Figure 3.1. Through the use of CSI, the system can adapt to a time varying channel. The level of adaptation depends on the level of complexity and CSI. Two constraints are used. The first is the long term average transmission power $\bar{\sigma}^2_{S_e}$. The second is the ratio r_{avg}, which is given by

$$r_{\text{avg}} = \frac{K}{L}, \tag{3.1}$$

where K is the total number of transmitted samples on the channel, and L is the number of source samples(pixels).

First the image is decorrelated through an analysis filter bank, before the filtered image is split into a set of N source-blocks. To be able to decode the received signal, the variance of the source-blocks needs to be sent to the receiver as side-information. Since the transmission order is based on the side-information, it has to be received fully before transmission of the main image information can start. An error in the side-information will then lead to a complete breakdown of the system. It is assumed that the side-information is decoded correctly throughout the thesis. The source-blocks are *preallocated* to one out of J nonlinear mappings, and to one out of M transmission states the channel can be in. This preallocation is done based on the long term statistics of the channel, and is done to be

able to plan power consumption and the number of channel samples. The level of CSI will play a factor in how the power consumption is planned. During transmission, the source-blocks will be mapped with a mapping and transmitted according to the preallocation-plan, but since the channel will not behave according to the long term statistics for a finite number of channel symbols, the transmitter will have to adapt blocks to new mappings and channel states on the go as well. This will lead to a slight variation from the preallocated parameters, but will on average be correct for multiple transmissions. At transmission, power can be used to adapt better to the channel. The channel will distort the signal by a time varying channel gain, and an AWGN component. From the side-information, the receiver can find the transmission order, by assuming that the receiver has the same information about the long term statistics as the transmitter. Since it is the receiver that estimates the CSI, it will know at all times know the information the transmitter has. As it is assumed that the receiver has full CSI at all times, it can try and compensate for the lack of CSI at the receiver. It is assumed that there are no synchronization problems at the receiver. After the receive-filter, the symbols are demapped according to the mapping used to map them in the transmitter. The demapped symbols are then filtered through a synthesis filter bank, and the image is reconstructed.

The following sections will describe each part of the system in more detail.

3.2.1 Decorrelation

Natural images are all highly correlated, which is the reason why most image coders use some sort of decorrelation of the image before further coding. In the proposed scheme, a pair of *filter banks* designed by Balasingham in [Balasingham, 1998] denoted "system K" is used to decorrelate the image.

Two different filter banks are used, an analysis filter bank, and a synthesis filter bank, where the analysis filter bank decorrelates the signal, and the synthesis filter bank reconstructs the signal [Vaidyanathan, 1993].

The filter banks consists of an eight band uniform filter bank, and three different two-band filter banks, which are used in a tree structure so each filter is applied to the lowpass-lowpass band of the previous stage, called *dyadic splitting*. The overall filter bank makes what is called a *tree-structured* filter bank. Since an image is a two-dimensional signal it needs to be decorrelated in both dimensions. A common way of dealing with this is by using a *separable* filter bank, which means that the image is filtered by the same filter twice, once for each direction.

The resulting sub-bands are organized as shown in Figure 3.2, where

the low frequency bands are placed top left. The reason why the lowpass-lowpass band is further filtered can be seen in Figure 3.3, where the sub-bands after the uniform filtering is shown. Natural images have more power at low frequencies than at high frequencies, so if only the uniform filter-bank where to be used, the lowpass-lowpass band would still be highly correlated [Taubman and Marcellin, 2001]. After each filtering stage the signal is decimated. The filter banks used here are *maximally decimated* which keeps the number of pixels equal before and after filtering.

The lowpass-lowpass band is the only band with expected mean different from zero, so to reduce the power of the image signal, the mean is subtracted from this band and sent as side-information.

In the case where an analysis and synthesis filter bank are connected directly together, and only generates a delay of the original signal, the overall filter-bank is said to have *perfect reconstruction* (PR). The 8×8 uniform filter bank used here is almost PR, and the effect of this can only be measured for very good image qualities, and is not a problem for any practical scenarios.

The analysis and synthesis filter banks used here are the same as used in [Coward, 2001].

To cope with the local statistical differences within each band, the sub-band filtered image is split into N blocks of $B \times B$ sub-band pixels. To be able to decode the transmitted signal the receiver needs knowledge about the variances of all the source blocks. This information is sent as side information. By using small blocks, the local statistics are captured better, but then the number of blocks will be larger, resulting in larger side information. Taking practical considerations into account, 8×8 sub-band pixels has been found to be a suitable block-size. The variance $\sigma^2_{X_n}$ of each block is estimated using the *root mean squared* (RMS) value.

The variance of each block will then reflect the activity within a block. A block with large variance has much activity and is hence important for the image quality. This can also be seen from rate distortion theory. In Section 2.1 it was said that for a given distortion, a source with large variance needs to be described with more bits than a source with small variance. In a more general sense, it can be said that more resources are needed to get the same distortion for a block of large variance compared to a block of smaller variance.

3.2.2 Dimensional changing mappings

Traditional communication systems are based on the idea that discrete time, discrete amplitude source symbols are transmitted over a channel under the protection of a channel code. The source encoder has to reduce

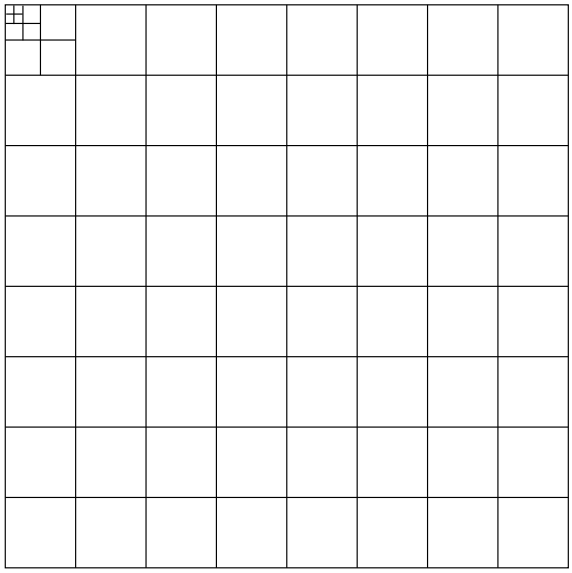


Figure 3.2: Organization of sub-bands

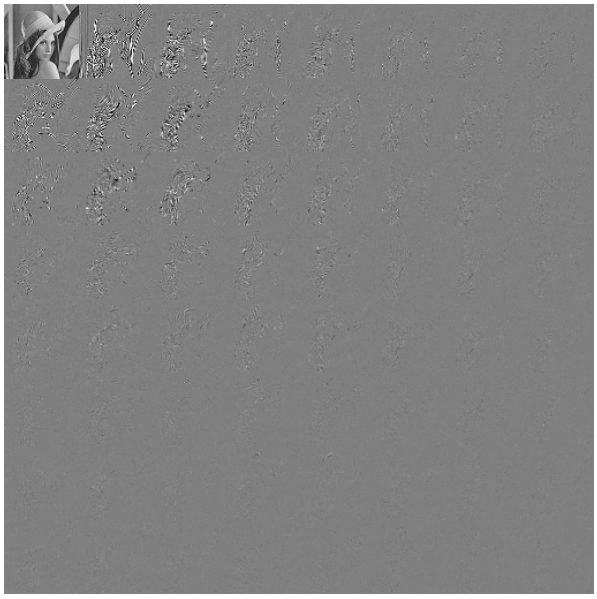


Figure 3.3: Example of sub-band decomposed image using eight band uniform filterbank.

the quality of the source to match the amount of information the channel can support with an acceptable number of errors.

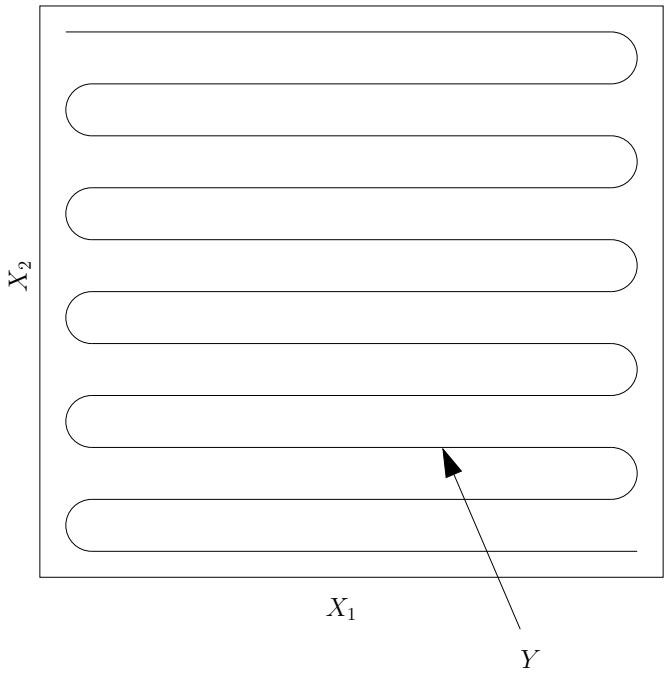


Figure 3.4: Mapping similar to Shannon's original suggested mapping.

Using Shannon-Kotel'nikov mappings will, however, allow some distortion being added by the channel. The channel is not assumed to be error-free any more. Discrete time, but amplitude continuous source symbols are being mapped into the channel space by *nonlinear mappings*, and transmitted as discrete time, continuous amplitude PAM symbols. The mappings work by taking L_j *source samples* and represent these samples by K_j *channel samples*, where L_j and K_j are integers such that $\hat{r} = K_j/L_j$. The resulting ratio $\hat{r}$, will from now on be denoted as the *rate* of the mapping. An example can be that a mapping of $\hat{r} = 2$, represents each source sample as 2 channel samples, thus adding redundancy to protect each sample. Different mapping-rates can be split into three different categories. The first is when a mapping produces fewer channel samples than the number of source samples (compression). Secondly there is the case when the number of channel symbols after the mapping process is the same as the number of source symbols (amplification), and the third case is when a mapping produces more channel symbols than the number of source symbols going in (expansion).

Explaining how compression and expansion can be achieved, is prob-

ably easiest done by looking at Figure 3.5 first for the case of $\hat{r} = 1/2$. Compression is achieved by representing two source samples by one channel sample. A $2D$ vector composed of two source samples is represented by (*). This point is then mapped to the closest point on the spiral, represented by (o). The mapped point will then be transmitted as a continuous amplitude PAM symbol. Points on the dotted line can be represented by negative channel symbols, while points on the solid line can be represented by positive channel symbols. Channel noise will move this point so that the received point will have a different value, represented by ($\Diamond$). The total distortion in the reconstructed sample will be due to the approximation noise and channel noise. In this specific example, the distortion will be distributed over 2 source samples, thus reducing the impact of noise for each sample. Further, it can be seen that designing good compression mappings, is a trade-off between approximation noise (quantization) and channel noise. By letting the spiral arms be denser, the approximation noise will go down, but to achieve the same transmission power, the transmitted signal must be downscaled. Scaling the signal up again in the receiver will then also amplify the channel noise. A thorough study of this for the $\hat{r} = 1/2$ case has been done by in [Hekland et al., 2005].

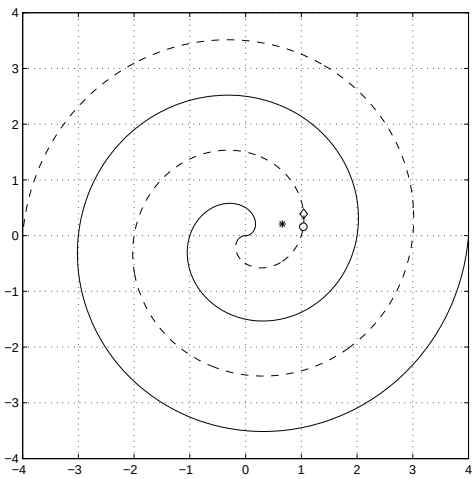


Figure 3.5: Example of mapping with $r_j = \frac{1}{2}$. $2D$ input vector marked by (*) is mapped to the closest point (o) in the channel space (the spiral). The channel noise will move the point along the spiral ($\Diamond$).

Figure 3.5 can also be used to explain a mapping of $\hat{r} = 2$. Letting the spiral represent the source space, a source sample, represented by ($\Diamond$), can take any value along the spiral. By representing this sample as a $2D$ vector, the value of each coordinate can be transmitted as a continuous

amplitude PAM symbol. After noise is added, the resulting $2D$ vector has been moved in the noise-space, represented by (*). The receiver knows that the original point has to be on the spiral, and maps it into the closest point along the spiral, represented by (o). From this it can be pointed out a problem with expanding mappings. If the noise power is so strong that the received point (*) is not closest to the dashed line any more, but will be decoded to the solid line instead, it would result approximately in a shift of sign for the source symbol. The resulting distortion will then be very large, compared to the regular noise moving the source point in an area which is close to the original source point. This property of expanding mappings, leads to a relatively fast breakdown of the performance for the mapping when the channel quality is lower than expected.

From these examples it is seen how optimizing mappings is a trade-off between approximation noise and channel noise. To deal with this, the mappings are optimized for a given CSNR value, γ_C. The configuration of a mapping is optimized for a channel noise power σ_N^2, such that the signal power of the mapped symbols is σ_G^2. Throughout this thesis, such a CSNR value will be denoted as a *representation point*, and give the CSNR value the source symbols are coded for

$$\gamma_C = \frac{\sigma_G^2}{\sigma_N^2}. \tag{3.2}$$

The *robustness* of a mapping is the ability to perform outside the optimized CSNR level. If the noise power is larger than assumed, the mapping should give graceful degradation. If the noise power is lower than assumed, the mapping should not suffer from the leveling-off effect.

Making a good mapping for a general dimension change is in general quite hard, and there is still work going on with these mappings, so for simplicity, the mappings already used in [Coward, 2001], which were based on mappings designed in [Lervik, 1996] and [Fuldseth, 1997], will be used. The scope of this thesis is not to optimize each mapping, but to show how these mappings can be used in a larger communication system.

The distribution of samples within a block can be modeled very closely with a Gaussian distribution. However, when grouping several blocks with different variances, the resulting distribution is closer to a Laplacian distribution [Joshi and Fischer, 1995; Lervik and Ramstad, 1996]. Since each of the mappings used here will be used for many blocks, each mapping is optimized for a Laplacian source.

To make the mappings work for several block-variances, each block is downscaled by its variance before being mapped. Thus the mappings can be optimized for a unit variance source. In the receiver the samples in a

block are scaled up again with the block variance. The distortion of a block can then be given by

$$\sigma_D^2 = \sigma_X^2 D(\sigma_{G_j}^2, r_j), \tag{3.3}$$

where D is the resulting optimized distortion for a mapping of rate r_j, with resulting output power $\sigma_{G_j}^2$.

3.2.2.1 Choosing a set of mappings

A good performing mapping might need high dimensional vectors as input, and output, but designing good mappings for large dimension changes is increasingly more difficult as the dimensionality goes up, due to the high number of parameters that need to be optimized, similar to *vector quantizer* (VQ) [Gersho and M. Gray, 1992]. Finding a mapping of arbitrary rate is generally hard, so due to this, the mappings that are implemented in the proposed coder all have rates from a ratio of integers of relatively low value.

$$\hat{r}_j \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{2}{3}, 1, 2\}, j = 0, \ldots, 5. \tag{3.4}$$

Throughout the thesis, $\hat{r}_j$ means that the rate is picked from one in the set in equation 3.4, while r can be any real number.

3.2.2.2 PCCOVQ mappings

Fuldseth [Fuldseth and Ramstad, 1997; Fuldseth, 1997] proposed a way of designing bandwidth compression mappings by using a VQ followed by a modulation mapping, optimized under a power constraint. The resulting scheme was called *power constrained channel optimized vector quantizer* (PCCOVQ). The optimization of the codebook is done by minimizing the total distortion of the decoded symbols for a given source distribution. The spacing of the representation points in the codebook is nonuniform, but are transmitted as an equal space multiple level PAM signal. An example of a reconstruction codebook for a Laplacian source over an AWGN channel is given in Figure 3.6. Neighbouring points along the line in the figure are transmitted as neighbouring points on the channel. This will, as the example given in Figure 3.5, give a robust mapping. In the proposed system three mappings designed by this method is used, $\hat{r} = \{1/4, 1/2, 2/3\}$. The robustness of the different PCCOVQ mappings can be seen in Figure 3.7, together with the best performance of the mappings and the OPTA curves for the different mapping rates. The rate distortion function for Laplacian sources can not been found analytically, so it has to be estimated using the Blahut algorithm, [Blahut, 1972]. The performance of these mappings will be very similar to that of an analog mapping, e.g. the mapping given

in Figure 3.5. Using a finite number of PAM symbols will not influence the performance until a relatively high CSNR where the mappings will saturate. Figure 3.7(a), 3.7(b) and 3.7(c) show how the different PCCOVQ mappings saturate as the CSNR increases. The codebook size is 256 for the mappings of rate $\hat{r}_1 = 1/4$ and $\hat{r}_2 = 1/2$, which gives a saturation at about 50 dB CSNR. For the mapping of rate $\hat{r}_3 = 2/3$ the performance saturates at about 30 dB, even though the codebook size is 1024. The explanation can be found when considering the fact that this mapping has a channel dimension of two, which means that there are only $\sqrt{1024} = 32$ code points in each dimension. To get the same performance, this mapping would then need a codebook of size $256^2 = 65536$ which would give to high complexity for the optimization of the mappings.

Each mapping is characterized by the reconstruction codebook, partitions, the minimum distance between channel symbols, Δ, and a Lagrange multiplier λ to deal with the power constraint. The PCCOVQ mappings are optimized for CSNR values with a spacing of about 1 dB. To be able to choose transmit power outside these points, the codebook corresponding to the closest optimized CSNR is chosen and the Δ and λ parameters are found through linear estimation from the neighboring optimized parameters. Due to the nature of the optimization process, the shape of the codebook might be too different for an interpolation. It is not possible to guarantee that an interpolated codebook would give good performance.

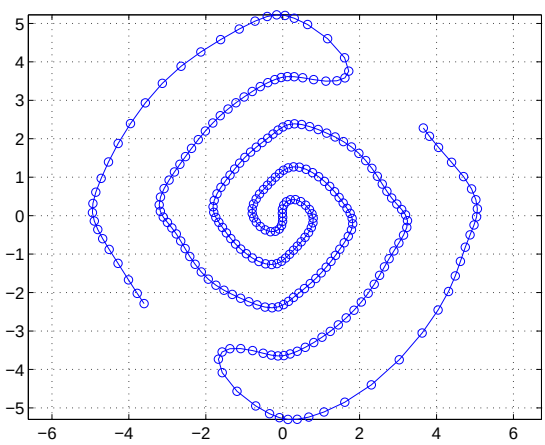
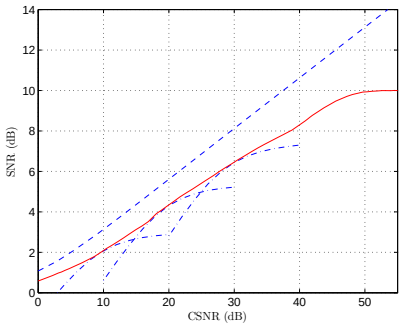
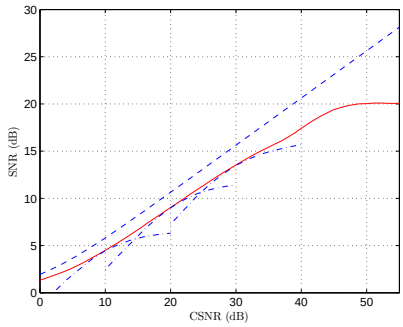


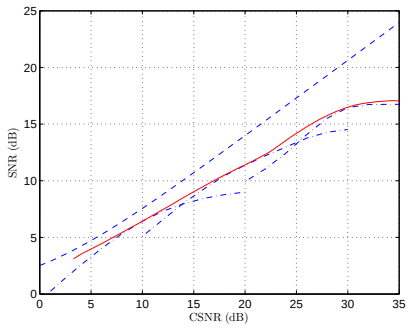
Figure 3.6: Reconstruction codebook for PCCOVQ mapping of rate 2 : 1 optimized for a memoryless unit variance Laplacian source at CSNR = 25.1 dB.



(a) $\hat{r}_1 = 1/4,\ L = 1024$



(b) $\hat{r}_2 = 1/2,\ L = 256$



(c) $\hat{r}_3 = 2/3,\ L = 256$

Figure 3.7: Performance (solid) and robustness (dash-dotted) compared to OPTA (dashed) for the PCCOVQ mappings for a Laplacian source. Robustness shown for mappings optimized for $\gamma_C = \{10, 20, 30\}$ dB.

3.2.2.3 Direct PAM mapping

For a rate $\hat{r}_4 = 1$, the optimal system is found for a Gaussian memoryless source, transmitted over an AWGN channel in section 2.3.2. It is a continuous amplitude PAM system. For a Laplacian source, the mapping will not be optimal any more, as OPTA is slightly higher for a Laplacian source compared to a Gaussian source, but will perform equal to the Gaussian source case.

This linear mapping is very robust due to the linear connection to the noise. So the performance of the direct PAM mapping will always follow the performance curve in Figure 3.8 as the noise power varies.

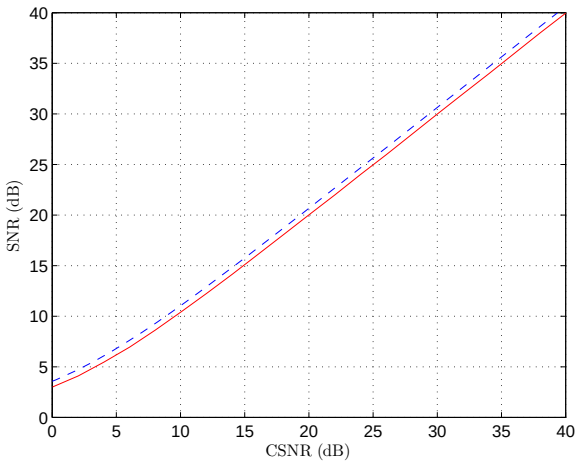


Figure 3.8: $\hat{r}_4 = 1$. Performance (solid) compared to OPTA (dashed) for the direct PAM mapping optimized for a Laplacian source. CSNR channel mismatch will follow the same performance curve.

3.2.2.4 HSQLC mapping

For bandwidth expansion the mapping called *hybrid scalar quantizer-linear coder* (HSQLC) designed by Coward and presented in [Coward and Ramstad, 1999, 2000a,c; Coward, 2001] is used. It is possible to design expanding mappings using the same methods as done for the PCCOVQ mappings, but the complexity will make it impractical due to the high number of points needed in the codebook [Fuldseth, 1997].

The HSQLC works by transmitting an uncoded scalar quantized symbol together with the quantization error. Hence increasing the number of channel symbols by a factor of two, giving the HSQLC system a rate of

$\hat{r}_5 = 2$. By keeping an analog part in the system, the mapping will avoid the leveling off effect, but the quantized part will lead to a faster breakdown when the channel is worse than expected. The performance and robustness of the HSQLC mapping can be seen in Figure 3.9 together with the OPTA for $\hat{r}_5 = 2$. Good dimensional expanding mappings are generally harder to find than dimensional reducing mappings, and it is seen from Figure 3.9 that the distance to OPTA is much larger than for the PCCOVQ mappings. As mentioned previously in Section 3.2.2, it is possible to design a bandwidth expanding mapping by using fully analog amplitude system. An example of this is presented in [Floor and Ramstad, 2006b,a] where it can be seen that such a system has about the same performance, and the same properties as the HSQLC system.

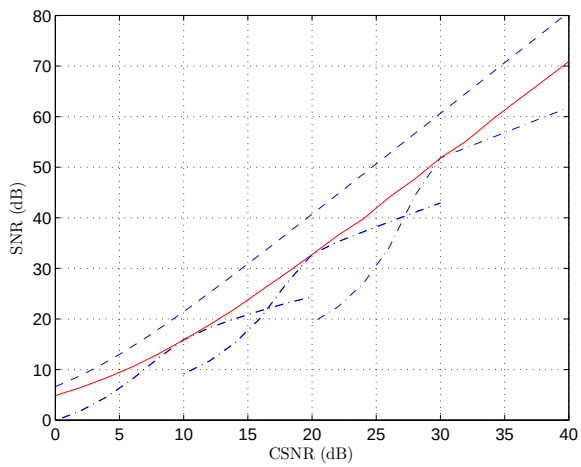


Figure 3.9: $\hat{r}_1 = 2$. Performance (solid) and robustness (dash-dotted) compared to OPTA (dashed) for the HSQLC mapping for a Laplacian source. Robustness shown when mappings optimized for CSNR $= \{10, 20, 30\}$ dB.

3.2.2.5 Finding mappings to use

In Section 2.1 it was said that a general source can be described by its rate distortion function, unfortunately this is generally not known, but in the case of a white Gaussian source, this is known, and given by equation (2.2), measured in bits/source sample.

For many channels the capacity is known and can be given in bits/channel sample. In Section 2.2 a few of them were mentioned.

By looking at the ratio between the rate distortion function of a source, and the channel capacity, the ratio measured in *channel samples/source*

samples, is obtained. In the case of a white Gaussian source with signal variance σ_X^2, distortion σ_D^2 where $\sigma_D^2 = \min(\mu, \sigma_X^2)$, and an AWGN channel with CSNR γ, this ratio is given by

$$r = \frac{\log_2\left(\frac{\sigma_X^2}{\sigma_D^2}\right)}{\log_2(1+\gamma)} \text{channel/source samples.} \tag{3.5}$$

This ratio is the bandwidth change needed, to be able to transmit a Gaussian source signal X, over an AWGN channel with CSNR= γ, resulting in distortion σ_D^2. The dimensions-changing properties of the mappings mentioned earlier are also given in channel samples/source samples. Equation (3.5) then gives the rate of the mapping needed.

Looking at equation (3.5), it can be seen that r can in theory take any value, but using a continuous set of dimensional changing mappings is, however, highly impractical, so in practice r has to be approximated to the closest match amongst the available set of mappings given by (3.4). Since the mappings are not ideal, the actual performance of each mapping has to be considered. This can be done by using the distortion as parameter to choose mapping from. Since the different mappings are optimized for different CSNR values, the distortion for a given CSNR value is known. By using this value to find the closest match of μ, the actual performance of each mapping can be included in the choice of mapping. The mapping with distortion closest to μ for a given CSNR value is chosen. μ is set on the basis of a target image quality.

The CSNR value for fading channels is proportional to the channel gain, α. Ideally this would mean that the transmitter should continuously choose mapping parameter setting as the channel changes. This is, however, complex and impractical, so a simpler strategy has to be used. For fading channels a common used technique to combat fading is to split the CSNR range into a set of *regions*, as mentioned in section 2.2.3.2. For channel-region m, a CSNR value γ_{C_m} is chosen to be the value the information in that region is coded for. γ_{C_m} is usually chosen conservatively to avoid that the CSNR value on the channel drops below this value, and the channel coder breaks down. Assuming that a channel is AWGN, the *instantaneous* channel capacity for channel-region m is then given by

$$C_m = \frac{1}{2}\log_2(1+\gamma_{C_m}). \tag{3.6}$$

A source n with variance $\sigma_{X_n}^2$, transmitted over channel state m, will need a mapping of rate r given by

$$r_{n,m} = \frac{\log_2\left(\frac{\sigma_{X_n}^2}{\sigma_D^2}\right)}{\log_2(1+\gamma_{C_m})} \text{ channel/source samples,} \tag{3.7}$$

to be able to achieve distortion μ. This framework will help to find the best suited mapping for any given source transmitted over a channel with a given CSNR.

When a block is allocated to a specific mapping, the rate of block n is denoted $\hat{r}_n$, where $\hat{r}$ is chosen from equation 3.4.

3.2.3 Finding channel regions and representation points

So far the mappings that are best fitted for a certain source-block variance and channel representation CSNR have been discussed. How these regions and representation points are chosen, have, however, not yet been discussed. In section 2.2.3.2 a few examples were briefly mentioned for the case of BER for codes and constellations. Using waterfall curves, the CSNR for a given BER were found directly, giving the regions in where different constellations should be used. A similar approach would not work in the case of mappings, as all the mappings can be used for any CSNR value. Ideally an object function should be found that would give the optimal regions and γ_{C_m} values, capturing the characteristics of the system. The best setting would be the one that minimizes the distortion σ_D^2, for an image given a

- power constraint $\bar{\sigma}_{S_e}^2$,
- rate constraint r_{avg},
- fixed set of mappings J.

For a set of M regions where symbols are being transmitted, assuming that the mappings used have optimal performance, the distortion can be written as

$$D = \sum_{n=0}^{N-1} \sigma_{D_n}^2 = \sum_{m=1}^{M} \sum_{n \in I_m} \sigma_{X_n}^2 \left(1 + \gamma_{C_m}\right)^{-\hat{r}_n} . \tag{3.8}$$

The number of blocks in a region is given by the probability for that region, but also through the mapping-rate of the block, since each mapping yields different number of samples. The mapping-rate depends on the representation points γ_{C_m} and block variance $\sigma_{X_n}^2$. In addition, the mapping-rate needed might not be available, leading to an extra distortion. An attempt to capture this in an optimization problem is done in the system presented in [Håkonsen et al., 2006]. Due to non-continuities introduced by the source-blocks, and the general complexity, the overall optimal system is, however, hard to find.

Instead of optimizing the system with source and channel information together, an approach of finding the channel related parameters first, then

the source related parameters, will be tried. To understand the problem better, two approaches will be tried when finding the channel regions $\{\gamma_{T_m}\}_{m=0}^{M-1}$, and representation points $\{\gamma_{C_m}\}_{m=1}^{M}$. Both will maximize the channel rate, but with different complexity.

Before going closer into the different optimization approaches, the possible levels of adaptation to the channel will be presented.

In this thesis, the different levels of adaptation can be summarized as follows

- The level of protection of the source signal can be adapted by setting the mapping-rate differently depending on the channel states.
- The level of protection provided by a mapping might not be exactly the needed. This can be compensated for by changing the transmission power for each source block, coding each block with mappings optimized for different CSNR values.
- The transmit power can also be used to compensate for the channel gain on the channel by scaling the transmitted symbols.

Traditional transmission schemes using digital channel codes, need to make sure the CSNR is above a certain threshold to be able to guarantee a certain BER. As seen previously, the mappings described in Section 3.2.2 are robust in nature. So a scheme using these mappings, does not necessarily have to be limited to a fixed received CSNR, but can achieve a good performance by relying on the robustness of the mappings.

The different mappings are optimized for a CSNR level, γ_C. In practice this means that a given setting for a mapping will give the best performance for a certain noise power, σ_N^2, through

$$\gamma_C = \frac{\sigma_G^2}{\sigma_N^2}, \tag{3.9}$$

where σ_G^2 is the power of the signal after being mapped. The optimization of a mapping is done with respect to minimization of the sum distortion of the approximation noise and additive channel noise. For a fading channel there will be a gain α distorting the information signal, which means that the mappings will not work optimally if the channel gain is not considered. Even if the noise power is equal to the one the mapping was designed for, the channel gain will move the received CSNR away from the CSNR the mapping was designed for. Due to this it would be beneficial to try and compensate for the gain both in the transmitter and receiver. A given way of dealing with this will be denoted as a *channel adaptation strategy* (CAS).

As explained in section 2.2.2, a fading channel is often characterized by its pdf. From equation (2.17) the instantaneous CSNR, written γ for short, is given as a function of the long term average transmit power, $\bar{\sigma}^2_{S_e}$, noise power, σ^2_N, and channel gain, α. By allowing the transmit power, $\sigma^2_S(\gamma)$, to vary as a function of the channel condition, an expression for the *post-adapted* channel capacity C, in bits/channel symbol, was found in[Goldsmith and Varaiya, 1997] as

$$C = \max_{P:\int_\gamma \sigma^2_S(\gamma)f(\gamma)d\gamma=\bar{\sigma}^2_{S_e}} \frac{1}{2}\int_\gamma \log_2\left(1+\frac{\sigma^2_S(\gamma)}{\bar{\sigma}^2_{S_e}}\gamma\right) f(\gamma)d\gamma, \tag{3.10}$$

under power constraint

$$\int_\gamma \sigma^2_S(\gamma)f(\gamma)d\gamma \leq \bar{\sigma}^2_{S_e}, \tag{3.11}$$

where the optimum is found when using water-filling in time, see section 2.2.3.1. For $M+1$ channel states with outage for $m=0, 0 \leq \gamma < \gamma_{T_0}$, the maximum number of bits/channel symbol R_c, is given by

$$R_c = \max_{\sigma^2_{S_m}:\sum_m \int_m \sigma^2_{S_m}(\gamma)f(\gamma)d\gamma=\bar{\sigma}^2_{S_e}} \frac{1}{2}\sum_{m=1}^{M}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \log_2\left(1+\frac{\sigma^2_{S_m}(\gamma)}{\bar{\sigma}^2_{S_e}}\gamma\right) f(\gamma)d\gamma, \tag{3.12}$$

under power constraint

$$\sum_{m=1}^{M}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \sigma^2_{S_m}(\gamma)f(\gamma)d\gamma \leq \bar{\sigma}^2_{S_e}. \tag{3.13}$$

The power constraint can also be written in a normalized version

$$\sum_{m=1}^{M}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{\sigma^2_{S_m}(\gamma)}{\bar{\sigma}^2_{S_e}} f(\gamma)d\gamma \leq 1. \tag{3.14}$$

Adapting the information signal to the channel to make sure a specific CSNR is received, can be done by scaling down the signal by the channel gain. The transmitted power will then depend on the gain the transmitter pre-scales with. Scaling the transmitted signal down with channel gain α before transmission, will give a transmitted power given by

$$\sigma^2_{S_m}(\gamma) = \frac{\sigma^2_{G_m}}{\alpha} = \frac{\sigma^2_{G_m}\bar{\sigma}^2_{S_e}}{\gamma\sigma^2_N} = \gamma_{C_m}\frac{\bar{\sigma}^2_{S_e}}{\gamma}. \tag{3.15}$$

Where equation (2.22) has been used in the second step.

In the sub-sections 3.2.3.2 to 3.2.3.4 three different strategies for adapting to the channel gain based on the CSI will be presented. The two first strategies require knowledge about which CSNR region the channel is in, whereas the last require full channel knowledge at all times. The latter can be seen as an upper limit as it is not practical due to the need for an infinite amount of CSI. These three strategies will be used throughout the thesis. Expressions for the planned power consumption and theoretical channel rate will be found for these three strategies. For each strategy, the expressions are used to find a setting for the channel regions and representation points.

3.2.3.1 Practical considerations

Before going further into the details of the different CAS ideas, a few practical considerations will be presented. To maximize equation (3.12) under the constraint given by equation (3.13), numerical methods will be used. The maximization is done with respect to the channel region thresholds $\{\gamma_{T_m}\}_{m=0}^{M-1}$ and representation points $\{\gamma_{C_m}\}_{m=1}^{M}$. Two approaches will be tried in this thesis. The first is a computationally expensive algorithm, maximizing equation (3.12) under a power constraint given by equation (3.13). This procedure will be denoted *complex* in the following. For practical systems, this operation might be too computational expensive, so a simpler approach might be desirable. In this thesis an approach described in further detail in Appendix C is used. In short the procedure is as follows:

- Fix $\gamma_{T_0} = 2$[1], set channel regions $\{\mathcal{R}_m\}_{m=1}^{M}$ such that each region has equal probability.
- In each region, set the representation points $\{\gamma_{C_m}\}_{m=1}^{M}$ equal to the centroid of the region.
- Scale the representation points by an equal factor for all regions such that the power equals the constraint.
- Run maximizing of equation (3.12) with the method described in Appendix C to a local maximum.

Below $\bar{\gamma} = 14$ dB, this approach has to be initialized by preset values to converge, as it is not possible to fulfill the power constraint with equal

[1]This is done to simplify computation. The level is set from the basis that for a theoretical system, the outage level will never be larger than one for a Rayleigh fading channel, [Alouini and Goldsmith, 1999]. For practical systems this will be larger. A very low CSNR value would also mean that a mapping with higher dimensionality than the largest available should be used.

probability. These points have been found through trial and error for low $\bar{\gamma}$ values and interpolated to $\bar{\gamma} = 14$ dB. It is not claimed that this approach is optimal in any way, but is used as an example of a simple approach. In the following, this approach will be denoted *simple*. By comparing this approach with the approach when using the regions and representation points using *complex* for optimization, it will give an indication on how sensitive the system is towards choice of regions and representation points. The results should also give an indication of the different factors regarding regions and representation points that give good system performance.

Both algorithms are constrained to keep the m'th representation point within region m, and to keep the thresholds in the right order.

3.2.3.2 No pre-scaling of channel symbols

What if the amount of CSI the transmitter has is low, so it can not track the channel changes continuously, but only the state the channel is in? A straight forward approach is to optimize the mappings for M different states, transmit with constant power, and let the receiver try to compensate for the gain mismatch on the channel. For the remainder of the thesis this *channel adaptation strategy* (CAS) will be denoted $\mathcal{N}$ for *none*, reflecting that no pre-scaling of the transmitted channel symbols is done to compensate for the channel gain. The transmit power is then given by

$$\sigma^2_{S_m}(\gamma) = \sigma^2_{G_m}, \tag{3.16}$$

resulting in channel rate R_c in bits/channel samples, given by

$$R_c = \max_{\sigma^2_{S_m}:\sum_m \int_m \sigma^2_{S_m}(\gamma) f(\gamma) d\gamma = \bar{\sigma}^2_{S_e}} \frac{1}{2} \sum_{m=1}^{M} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \log_2 \left(1 + \frac{\sigma^2_{G_m}}{\bar{\sigma}^2_{S_e}} \gamma \right) f(\gamma) d\gamma, \tag{3.17}$$

with power constraint given by

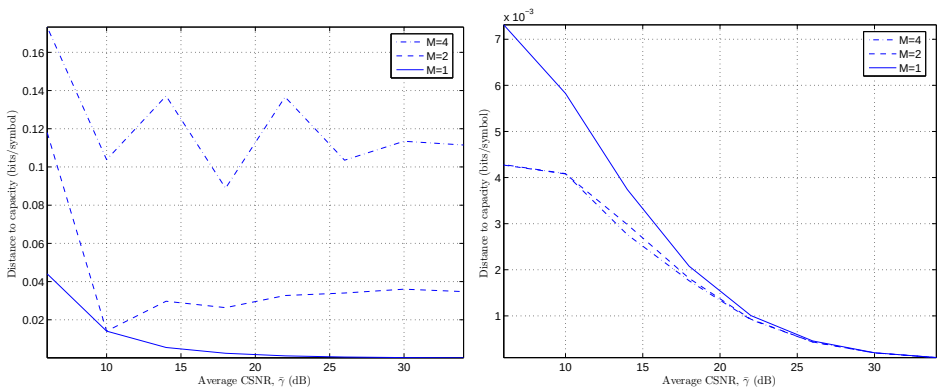
$$\sum_{m=1}^{M} \sigma^2_{G_m} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma \leq \bar{\sigma}^2_{S_e}, \tag{3.18}$$

which can be written as

$$\sum_{m=1}^{M} \frac{\sigma^2_{G_m}}{\bar{\sigma}^2_{S_e}} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma \leq 1. \tag{3.19}$$

From equation (3.17) it can be seen that for each region m, the output-power will be a fixed value, making the post-adapted CSNR change with

the channel. A changing CSNR will mean that the transmission rate is a continuous variable. For $M = 1$ this was shown in [Goldsmith and Varaiya, 1997] to come very close to the capacity. The effect of this, can be seen in Figure 3.10 where the distance in bits/sample, to the channel capacity for a Rayleigh fading channel is plotted for different values of M, both for the *simple* algorithm and the *complex* algorithm. The optimal channel rate can only increase as the number of channel regions increases, since more degrees of freedom are introduced. The possible gain is, however, small when looking at Figure 3.10. For the optimization, finding the optimal parameters becomes harder as the number of channel regions increases. This is clearly reflected in values found for $\{\gamma_{C_m}\}_{m=1}^{4}$, $\{\gamma_{T_m}\}_{m=0}^{3}$, and region probability $\{p_m\}_{m=1}^{4}$, plotted in Figure 3.11(a),3.11(c) and 3.11(e) for the *simple* algorithm, and Figure 3.11(b), 3.11(d) and 3.11(f) for the *complex* algorithm. Even though the values of $\{\gamma_{C_m}\}_{m=1}^{4}$ and $\{\gamma_{T_m}\}_{m=0}^{3}$ for the *simple* algorithm seem to behave better, the simpler nature of the algorithm makes the channel-rate, R_c decrease as M increases. For the *complex* algorithm the channel-rate, R_c increases as M increases, but some regions seems to collapse. The region probabilities are accumulated, so the upper curve will give the probability of transmission. For the case of $M = 2$ and $M = 1$, the representation points, thresholds and region probability, are plotted in Figures D.1 and D.4.



(a) *simple* (b) *complex*

Figure 3.10: Distance to the channel capacity for a Rayleigh fading channel for $M = \{1, 2, 4\}$ number of channel regions, CAS set to $\mathcal{N}$.

Since the transmission power is constant within each region, the CSNR will vary as the channel gain varies. So when assuming that the receiver has full channel information, it will be up to the receiver to try to compensate for the channel gain mismatch. This can be done due to the robustness

of the mappings. An initial thought could be to invert the channel gain. The problem with this is that the noise could actually be enhanced. To minimize the overall perturbation by the channel, a strategy can be to keep the difference between the mapped signal $g(k)$ and the signal being demapped $\hat{g}(k)$ as small as possible. Taking both the channel gain and additive noise into consideration, a simple *channel gain mismatch receiver filter* is given by

$$w_m(k) = \frac{\sqrt{\alpha(k)}}{\alpha(k) + \frac{\sigma_N^2}{\sigma_{Sm}^2(\gamma)}}. \tag{3.20}$$

Since there is no scaling of the signal before transmission in this case, $\sigma_{Sm}^2(\gamma) = \sigma_{Gm}^2$. The full calculation can be found in Appendix B.1.

3.2.3.3 Fixed channel gain within a region

For $\mathcal{N}$ the transmitter knew the channel state, so it could adapt the mapping rates accordingly, but the transmitter did not do any pre-scaling of the channel symbols before transmisson to compensate for the channel gain. If it is assumed that the amount of CSI is the same, another option could the be to pre-scale the channel symbols using one gain value as a representation for the whole region. For the remainder of the thesis this *channel adaptation strategy* (CAS) will be denoted $\mathcal{S}$ for *single*, reflecting a single scalar used for each channel region. The transmitted power for each region is then expressed as

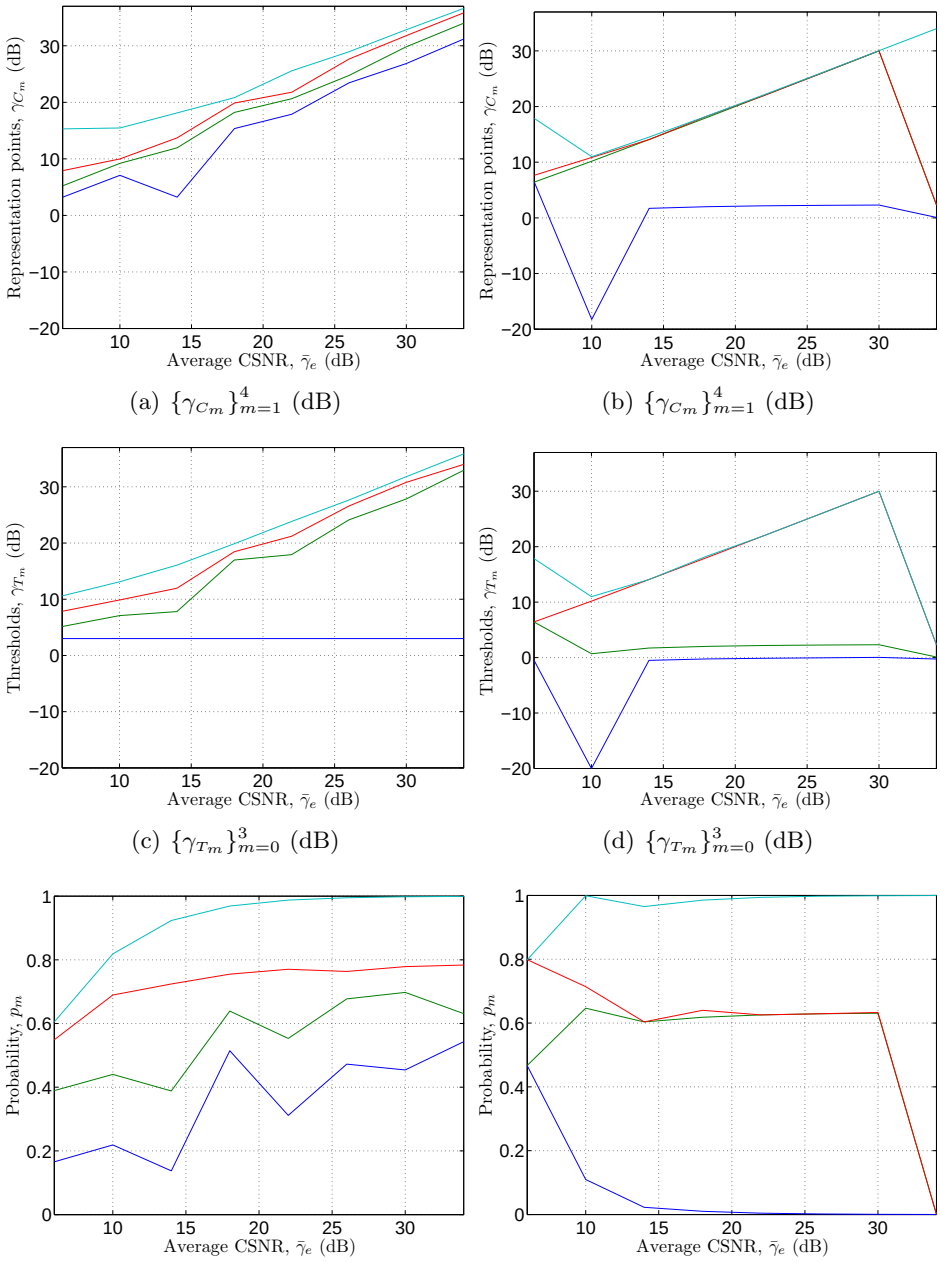
$$\sigma_{Sm}^2(\gamma) = \frac{\sigma_{Gm}^2}{\alpha_{A_m}} = \frac{\sigma_{Gm}^2 \bar{\sigma}_{S_e}^2}{\gamma_{A_m} \sigma_N^2} = \gamma_{C_m} \frac{\bar{\sigma}_{S_e}^2}{\gamma_{A_m}} \tag{3.21}$$

where γ_{A_m} is the assumed *pre-adapted* scaling-CSNR. To make the error between the assumed gain and the exact as small as possible, the distribution of the gain has to be taken into consideration, or the distribution of γ, $f(\gamma)$. The average error introduced by choosing a fixed gain value can be expressed, when using a squared error metric, as

$$\Psi = \sum_{m=1}^{M} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} (\gamma - \gamma_{A_m})^2 f(\gamma) d\gamma. \tag{3.22}$$

From the optimization of a scalar quantizer in [Gersho and M. Gray, 1992], it is known that the γ_{A_m} values that minimize Ψ, are the centroids in each region, given by

$$\gamma_{A_m} = \frac{\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \gamma f(\gamma) d\gamma}{\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma}. \tag{3.23}$$



(a) $\{\gamma_{C_m}\}_{m=1}^{4}$ (dB)

(b) $\{\gamma_{C_m}\}_{m=1}^{4}$ (dB)

(c) $\{\gamma_{T_m}\}_{m=0}^{3}$ (dB)

(d) $\{\gamma_{T_m}\}_{m=0}^{3}$ (dB)

(e) $\{p_m\}_{m=1}^{4}$

(f) $\{p_m\}_{m=1}^{4}$

Figure 3.11: Representation points, γ_{C_m}, thresholds, γ_{T_m} for CAS set to $\mathcal{N}$, and accumulated region probability p_m. Found by *simple* algorithm (left column), *complex* algorithm (right column), $M = 4$.

Inserting equation (3.21) into equation (3.12), yields

$$R_c = \max_{\sigma^2_{Sm}:\sum_m \int_m \sigma^2_{Sm}(\gamma)f(\gamma)d\gamma=\bar{\sigma}^2_{Se}} \frac{1}{2}\sum_{m=1}^{M}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \log_2\left(1+\frac{\gamma_{C_m}}{\gamma_{A_m}}\gamma\right) f(\gamma)d\gamma, \tag{3.24}$$

with power constraint

$$\sum_{m=1}^{M}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \sigma^2_{Sm}(\gamma)f(\gamma)d\gamma = \sum_{m=1}^{M}\frac{\gamma_{C_m}\bar{\sigma}^2_{S_e}}{\gamma_{A_m}}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma)d\gamma \leq \bar{\sigma}^2_{S_e}, \tag{3.25}$$

which can be simplified to

$$\sum_{m=1}^{M}\frac{\gamma_{C_m}}{\gamma_{A_m}}\int_{\gamma_{T_m}}^{\gamma_{T_m}} f(\gamma)d\gamma \leq 1. \tag{3.26}$$

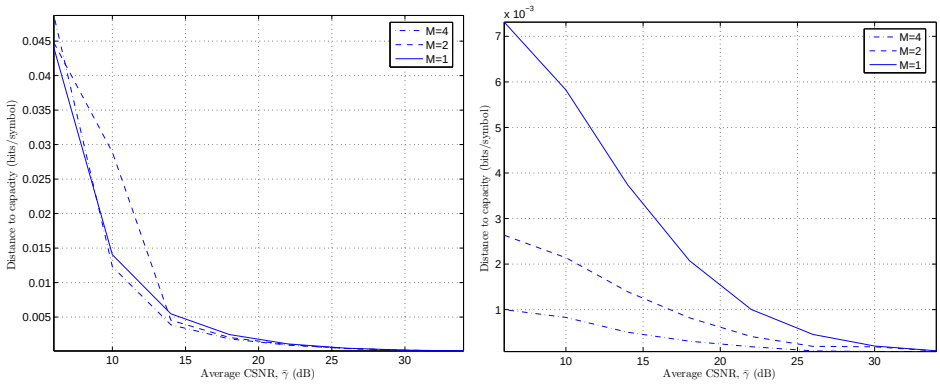
It is assumed that the *transmitter* only knows the channel state, and not the exact channel gain at all times. However, assuming that the *receiver* knows the exact channel gain at all times, the receiver can use a channel gain mismatch equalizer. The full calculation to find this expression is shown in Appendix B.2. When using the centroid γ_{A_m} as a scaling factor, the optimal channel gain mismatch received filter will be given by

$$w_m(k) = \frac{\sqrt{\alpha(k)}\sqrt{\alpha_{A_m}}}{\alpha(k)+\frac{\sigma^2_N}{\sigma^2_{Sm}(\gamma)}}. \tag{3.27}$$

As in the case of not adapting to the channel gain, using one value to represent a channel region leads to near optimal channel rate. Optimizing Equation (3.24) will give very similar results whether the optimization is done with the *complex* algorithm or the *simple* algorithm. For comparison the distance from the channel capacity for $M = \{1, 2, 4\}$, is given in Figure 3.12(a) for *simple*, and in Figure 3.12(b) for *complex*. Looking at the representation points $\{\gamma_{C_m}\}_{m=1}^{4}$, region thresholds $\{\gamma_{T_m}\}_{m=0}^{3}$, and region probabilities $\{p_m\}_{m=1}^{4}$ in Figure 3.13, is is seen that the optimized parameters are smoother for the *complex* algorithm compared to the case of CAS being set to $\mathcal{N}$. For the case of $M = 2$ and $M = 1$, the representation points, thresholds and region probability, are plotted in Figures D.2 and D.4.

3.2.3.4 Perfect channel information

By getting perfect CSI with no delay, the transmitter can adapt continuously to the channel. This means that it is possible to scale down with



(a) *simple* (b) *complex*

Figure 3.12: Theoretical channel rate for a Rayleigh fading channel for $M = \{1, 2, 4\}$ number of channel regions, CAS set to $\mathcal{S}$, compared to the capacity.

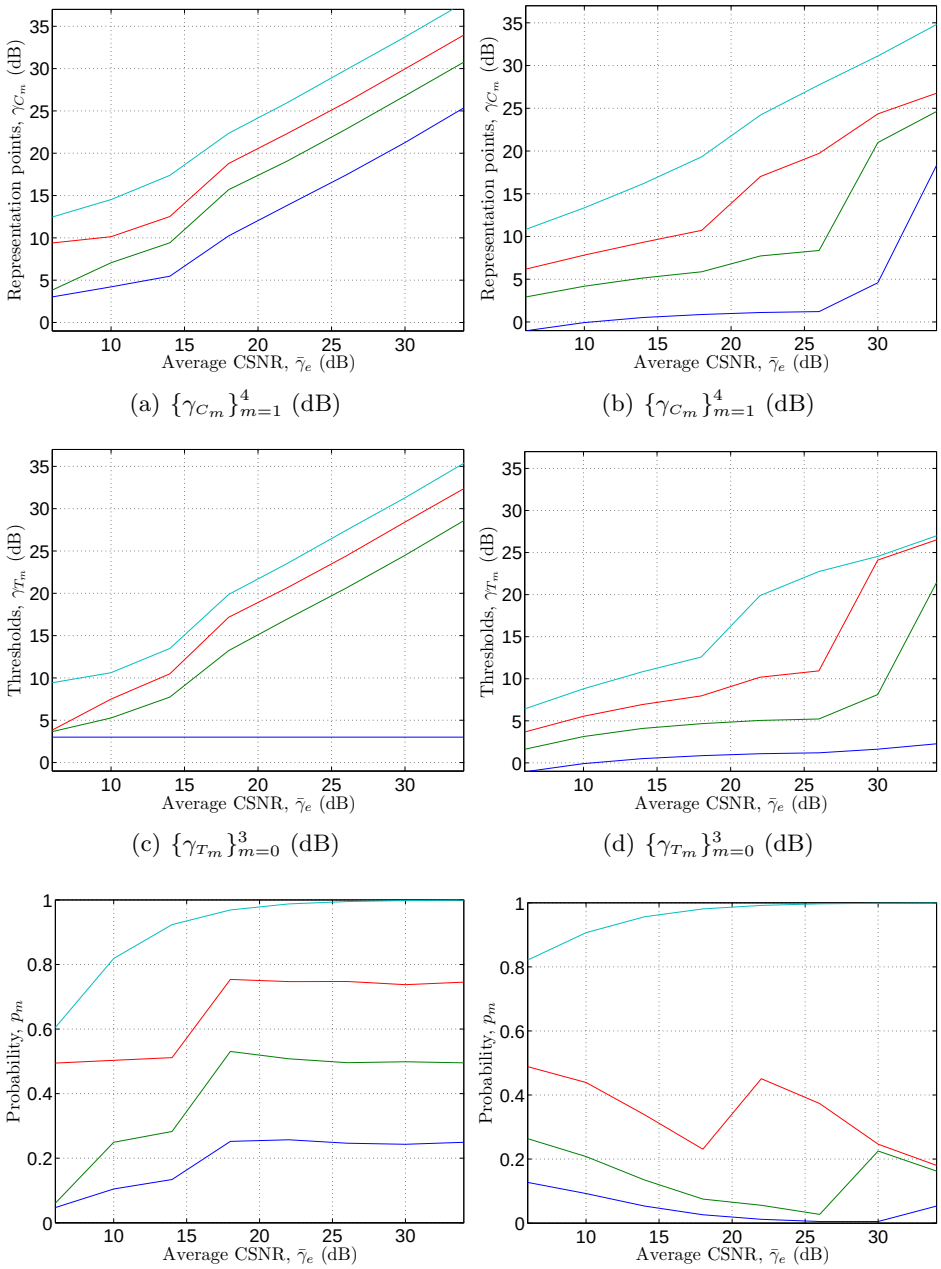
the channel gain for each sample, and the post-adapted CSNR is fixed to $\{\gamma_{C_m}\}_{m=1}^{M}$ for the M different channel states. Scaling down with the channel gain for each channel sample leads to a channel inversion for each channel state. Extra power is used when the instantaneous CSNR is lower than γ_{C_m}, and power is saved if the instantaneous CSNR is larger than γ_{C_m}. For the remainder of the thesis this *channel adaptation strategy* (CAS) will be denoted $\mathcal{C}$ for *continuous*. The transmitted power will then be given by equation (3.15), and the maximum number of bits/channel symbol can be found by maximizing the expression found by inserting equation (3.15) into equation (3.12), resulting in

$$R_c = \max_{\sigma^2_{S\,m}:\sum_m \int_m \sigma^2_{S\,m}(\gamma)f(\gamma)d\gamma=\bar{\sigma}^2_{S_e}} \frac{1}{2}\sum_{m=1}^{M}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \log_2\left(1+\gamma_{C_m}\right) f(\gamma)d\gamma \quad (3.28)$$

$$= \max_{\sigma^2_{S\,m}:\sum_m \int_m \sigma^2_{S\,m}(\gamma)f(\gamma)d\gamma=\bar{\sigma}^2_{S_e}} \frac{1}{2}\sum_{m=1}^{M}\log_2\left(1+\gamma_{C_m}\right)\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma)d\gamma, \quad (3.29)$$

with power constraint given by inserting equation (3.15) into equation (3.13) resulting in

$$\bar{\sigma}^2_{S_e} \geq \sum_{m=1}^{M}\gamma_{C_m}\bar{\sigma}^2_{S_e}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{1}{\gamma}f(\gamma)d\gamma, \quad (3.30)$$



(a) $\{\gamma_{C_m}\}_{m=1}^{4}$ (dB)

(b) $\{\gamma_{C_m}\}_{m=1}^{4}$ (dB)

(c) $\{\gamma_{T_m}\}_{m=0}^{3}$ (dB)

(d) $\{\gamma_{T_m}\}_{m=0}^{3}$ (dB)

(e) $\{p_m\}_{m=1}^{4}$

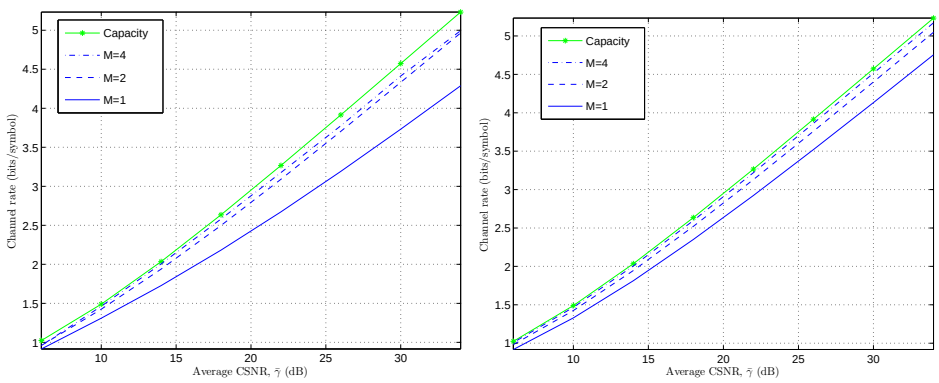
(f) $\{p_m\}_{m=1}^{4}$

Figure 3.13: Representation points, γ_{C_m}, and thresholds, γ_{T_m} for CAS set to $\mathcal{S}$. Found by *simple* algorithm (left column), *simple* algorithm (right column), $M = 4$.

which can be simplified to

$$1 \geq \sum_{m=1}^{M} \gamma_{C_m} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{1}{\gamma} f(\gamma) d\gamma. \tag{3.31}$$

Equation (3.29) is maximized with respect to $\{\gamma_{C_m}\}_{m=1}^{M}$, and $\{\gamma_{T_m}\}_{m=1}^{M-1}$ for *simple*, and $\{\gamma_{C_m}\}_{m=1}^{M}$, and $\{\gamma_{T_m}\}_{m=0}^{M-1}$ for *complex*. The channel rates, R_c, resulting from solving equation (3.29) numerically for a Rayleigh fading channel, is plotted in Figure 3.14(a) for *simple* and Figure 3.14(b) for *complex* , for different number of channel regions M. The channel capacity is plotted as a reference. It should be mentioned that for a Rayleigh fading channel, a system with no outage using this approach would lead to $R_c = 0$. For both cases the channel rate R_c approaches the capacity as the number of regions increases. The biggest difference can be seen for the case of $M = 4$ and $M = 1$. For $M = 1$ the only difference is due to the outage level γ_{T_0}. The resulting $\{\gamma_{C_m}\}_{m=1}^{4}$, $\{\gamma_{T_m}\}_{m=1}^{4}$ and $\{p_m\}_{m=1}^{4}$ values are plotted in Figure 3.15. It can clearly be seen how the *simple* algorithm converges to a local optimum compared to the *complex* algorithm, but it is interesting how the curves for the channel rate, R_c, still is very smooth as seen in Figure 3.14(a). For the case of $M = 2$ and $M = 1$, the representation points, thresholds and region probability, are plotted in Figures D.3 and D.4.



(a) *simple* (b) *complex*

Figure 3.14: Theoretical channel rate for a Rayleigh fading channel for $M = \{1, 2, 4\}$ number of channel regions, CAS set to $\mathcal{C}$, compared to the capacity.

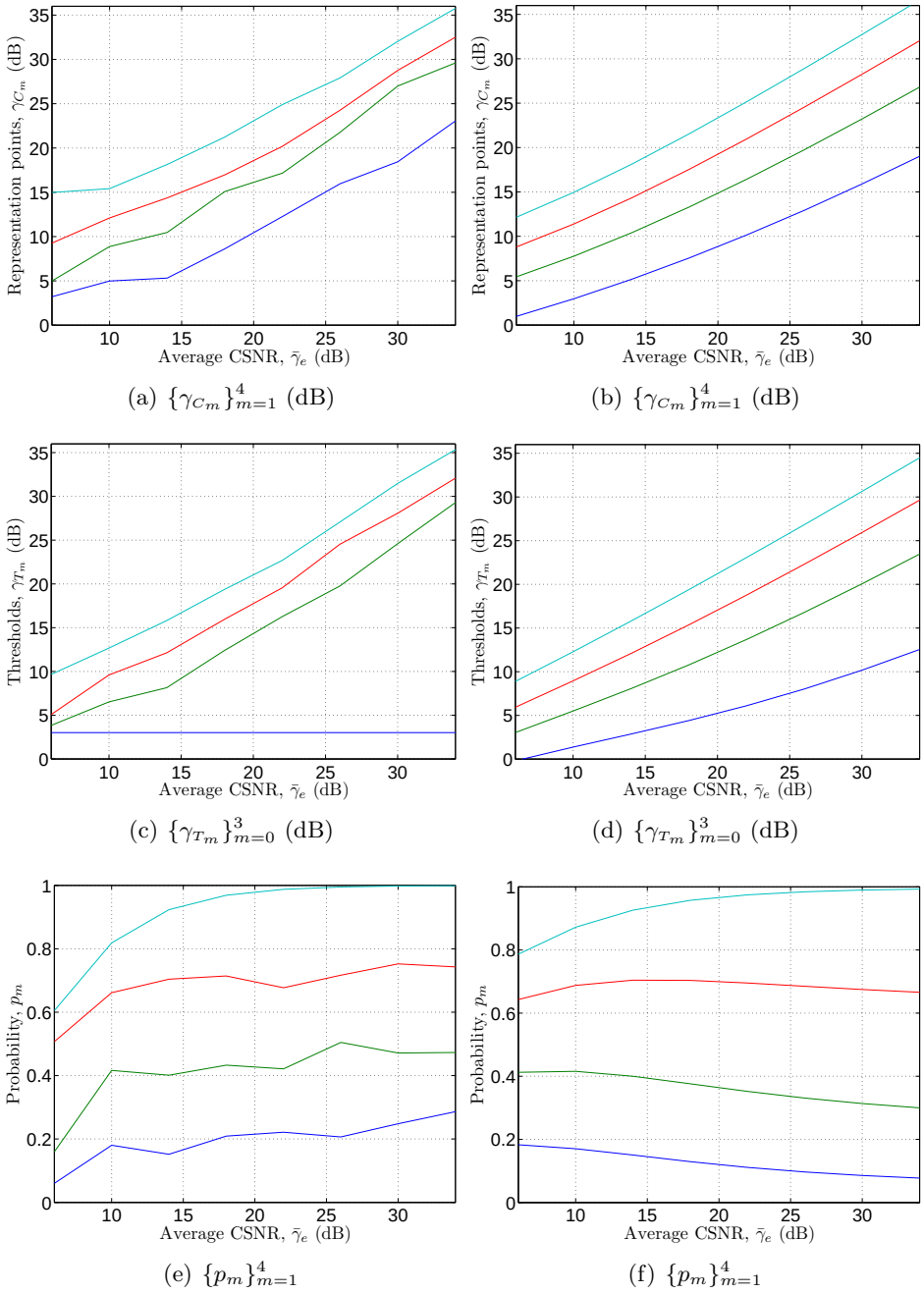


Figure 3.15: Representation points, γ_{C_m}, and thresholds, γ_{T_m} for CAS set to $\mathcal{C}$. Found by *simple* algorithm (left column), *complex* algorithm (right column), $M = 4$.

3.2.4 Preallocation

When transmitting information efficiently and reliably on a fading channel is it essential to have a plan for combating the varying conditions. To be able to plan the transmitted power, or the time it takes to transmit a certain amount of information ahead of transmission, the decisions must be based on the statistics of the channel. For this thesis, it has been decided to *preallocate* the source-blocks to different channel-states according to certain wanted properties.

After the filterbank decomposition, the image is decorrelated and arranged according to the frequency content. The source-block variance, $\sigma^2_{X_n}$, will indicate the level of activity within a block, and hence show the importance of a given block. The CSNR range is split into different channel-regions with a given channel quality γ_{C_m} representing each region. A problem arising then, is to decide what source-block to transmit on what channel-region. There are however a few desired properties that can help us:

- There is only a limited set of mappings, and the largest mapping-rate is 2.
- The overall number of channel samples should be as small as possible.

To avoid the need of high dimensional expanding mappings, it can be seen from equation (3.7) that to minimize r for a given source block variance $\sigma^2_{X_n}$ and distortion, μ, the CSNR value to code for, γ_{C_m}, should be as large as possible. So the largest block variances should be sent on the channel-region with the highest CSNR. To minimize the overall number of channel samples, the average mapping-rate needed for the different blocks, given by

$$r_{\text{avg}} = \frac{1}{N}\sum_{m=1}^{M}\sum_{n\in I_m} r_{n,m} = \frac{1}{N}\sum_{m=1}^{M}\sum_{n\in I_m} \frac{\log_2\left(\frac{\sigma^2_{X_n}}{\sigma^2_D}\right)}{\log_2\left(1+\gamma_{C_m}\right)}, \tag{3.32}$$

where I_m is the set of source blocks preallocated to channel state m, can be minimized.

To minimize equation (3.32), the source blocks with largest variance should be sent on the channel-state with largest γ_C. The source blocks should then be preallocated in a descending order to the channel states until the source block with the smallest variance is preallocated to the poorest channel-state. Proof is given in Appendix A.

The number of blocks that can be allocated to a given channel-state, will depend on the probability of the channel-state, and the mapping a block

uses. Since a block may need different mapping-rates on different channel-states, the preallocation needs to be done in a way so that the probability of a given channel-state, matches the theoretical state-probability as closely as possible. An exact match will not always be possible as it is desirable to avoid splitting a block up on several states, and the number of channel samples a block produces, depends greatly on the mapping used. A block coded with the mapping of rate 2 will for example produce eight times as many samples as a block coded with a mapping of rate 1/4.

For a given target r_{avg}, the total number of channel samples K can be expressed by

$$K = Lr_{\text{avg}}, \tag{3.33}$$

where L is the total number of pixels in an image. The theoretical state probability p_m, has to be translated into a number of channel samples when allocating blocks to different regions. For the m'th region, the probability given from the number of blocks allocated to it, is given by

$$\hat{p}_m = \frac{B^2 \sum_{n \in I_m} \hat{r}_n}{K}, \tag{3.34}$$

where B^2 is the number of pixels in each block. Normalizing with respect to the block-size, the expression for the probability becomes

$$\hat{p}_m = \frac{\sum_{n \in I_m} \hat{r}_n}{N r_{\text{avg}}}, \tag{3.35}$$

where N is the total number of source blocks.

γ_{C_4}	$\sigma^2_{X_1}$			
γ_{C_3}	$\sigma^2_{X_2}$	$\sigma^2_{X_3}$	$\sigma^2_{X_4}$	
γ_{C_2}	$\sigma^2_{X_5}$	$\sigma^2_{X_6}$	$\sigma^2_{X_7}$	$\sigma^2_{X_8}$
γ_{C_1}	$\sigma^2_{X_9}$	$\sigma^2_{X_{10}}$	$\sigma^2_{X_{11}}$	$\sigma^2_{X_{12}}$

Figure 3.16: Example of preallocation with twelve source blocks and four channel states. $\gamma_{C_1} \geq \ldots \geq \gamma_{C_4}$, $\sigma^2_{X_1} \geq \ldots \geq \sigma^2_{X_{12}}$.

The number of source-blocks that will be transmitted, is decided by the distortion-level μ. Large values of μ imply that some of the source blocks are discarded according to rate distortion theory [Berger, 1971]. This is expressed through equations (2.2) and (2.7).

By using equation (2.9), it is possible to find the value of μ that yields a target SNR. This can be the scenario if the receiver needs a certain quality in the received image. Another scenario is when there is a real time requirement where the transmission has to be finished within a certain time. This will put a rate constraint on (3.32), and the corresponding μ for this rate needs to be found.

To make sure the preallocated rate is as close to the target rate r_{avg} as desired, iteration is done over the distortion value μ, with linear approximation between each iteration. The number of blocks that fulfill $\sigma^2_{X_n} \leq \mu$ will be preallocated for transmission. This will also mean that the allocated probabilities, $\{\hat{p}_m\}_{m=0}^{M}$, are iterated over to make sure they are as close to the theoretical probabilities, $\{p_m\}_{m=0}^{M}$ as wanted. For the experimental results in this thesis, an accumulated error of the preallocated probability of 0.01 has been used. As long as the region probability is independent of the image, an adapted version of the rate allocation algorithm in [Westerink et al., 1988a] could have been used instead.

3.2.5 Mapping rate mismatch compensation

The mapping-rate, $\hat{r}_n$, for each block, does in general not match the needed rate $r_{n,m}$ exactly. Since the mappings are found by the closest match in distortion, it means that the protection a mapping gives is sometimes better than needed, and sometimes worse than needed. This leads to a variation in the resulting distortion that is not necessarily optimal. By coding each block for different CSNR values, i.e., each block is coded with mappings optimized for separate CSNR values γ_{C_n}, it is partly possible to compensate for this. The optimal power distribution after the source blocks have been preallocated to a channel state and mapping-rate, can be found by minimizing the total distortion given by

$$D = \sum_{n=0}^{N-1} \sigma^2_{X_n} D_n(\gamma_{C_n}) \tag{3.36}$$

where, for practical mappings, $D_n(\gamma_{C_n})$ is the tabulated distortion of the mapping the n'th block is given, designed for CSNR γ_{C_n}. The normalized power constraint will be used, and is given by equation (3.14). This results in a problem that can be solved using Lagrange multipliers, with an object

function

$$\mathcal{L} = \sum_{n=0}^{N-1} \sigma_{X_n}^2 D_n(\gamma_{C_n}) + \lambda \sum_{m=1}^{M} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{\sigma_{S_m}^2(\gamma)}{\bar{\sigma}_S^2} f(\gamma) d\gamma, \tag{3.37}$$

where λ is the Lagrange multiplier.

To find the optimal power distribution under an overall power constraint, (3.37) is differentiated with respect to γ_{C_n} and set to zero. As seen in Sections 3.2.3.2 to 3.2.3.4, the expression for power $\sigma_{S_m}^2(\gamma)$, will depend on the choice of CAS, and will be found on a per block basis in the following sections.

For theoretical mappings following OPTA, the distortion as a function of the power is a convex function. The derivative for the distortion with respect to the power of a theoretical mapping will then be negative and increasing as the power increases. For practical mappings, the distortion is not necessarily convex as a function of the power, but for any reasonable choice of mapping, the derivative will go towards zero as the power increases.

Since $D(\gamma_C)$ is tabulated, the derivative does not exist, and it has to be estimated. This is done by differentiating the quadratic spline estimates between the optimized CSNR points, resulting in a derivative consisting of straight line segments. For practical mappings, there is no guarantee that the distortion results in a convex estimate, or even a non-increasing estimate. To compensate for this, a full search has to be made to find the optimal parameter. The parameter resulting in the smallest Lagrangian is chosen. This has to be done for all blocks that are going to be transmitted.

The power for each block has the additional physical constraint that it has to be positive, which can be solved by the Karush-Kuhn-Tucker conditions, [Nocedal and Wright, 1999], but due to the simplicity of the problem, the power for these blocks will simply be set to zero.

Finding the optimal λ value is a task that has to be done through a search. For the optimal λ, the optimal $\gamma_{C_n}^*$ values satisfy the total power in

$$\sum_{m=1}^{M} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{\sigma_{S_m}^{2*}(\gamma)}{\bar{\sigma}_{S_e}^2} f(\gamma) d\gamma = 1. \tag{3.38}$$

To find the total average transmission power when each block can be coded for a special CSNR level, the average transmission power $\bar{\sigma}_{G_m}^2$ for each channel region has to be found since the symbols in each region might be treated differently. Due to the blocks being allocated to different rates $\hat{r}_n$, each source block will result in different number of channel samples, resulting in different relative energy consumption for each block. This

means that the rate for each blocks has to be included in the calculation. The average mapping output power $\bar{\sigma}^2_{G_m}$ for region m, can be expressed as

$$\bar{\sigma}^2_{G_m} = \frac{\sum_{n \in I_m} \hat{r}_n \sigma^2_{G_n}}{\sum_{n \in I_m} \hat{r}_n}, \tag{3.39}$$

where $\sigma^2_{G_n}$ is the output power of the mapping $\hat{r}_n$, source block n is allocated to.

For the remainder of the thesis the scenarios where mapping-rate mismatch compensation is enabled will be denoted as $\mathfrak{B}$, for *block*-basis. $\mathfrak{NB}$ will be used to represent the cases when mapping-rate mismatch is not used.

As mentioned above, the derivative of the distortion for practical mappings does not exist. In the following sub-sections, expressions that the estimated derivative has to match is found for the different CAS variants.

3.2.5.1 No pre-scaling of channel symbols

When choosing not to adapt to the varying channel gain at the transmitter, the transmitter can still choose to compensate for inaccurate mapping-rate for each block. The transmit power for a given channel state for $\mathcal{N}$ is generally given by equation (3.16). Setting equation (3.39) in for $\sigma^2_{G_m}$ in equation (3.16) results in

$$\sigma^2_{S_m}(\sigma^2_G) = \frac{\sum_{n \in I_m} \hat{r}_n \sigma^2_{G_n}}{\sum_{n \in I_m} \hat{r}_n}. \tag{3.40}$$

The total normalized power can then be found as a function of γ by replacing $\sigma^2_G = \gamma_C \sigma^2_N$ resulting in

$$\frac{\sigma^2_{S_m}(\gamma)}{\bar{\sigma}^2_{S_e}} = \frac{\sigma^2_N \sum_{n \in I_m} \hat{r}_n \gamma_{C_n}}{\bar{\sigma}^2_{S_e} \sum_{n \in I_m} \hat{r}_n}. \tag{3.41}$$

This expression can be inserted into equation (3.37) resulting in an object function

$$\mathcal{L} = \sum_{n=0}^{N-1} \sigma^2_{X_n} D_n(\gamma_{C_n}) + \lambda \sum_{m=1}^{M} \frac{\sigma^2_N \sum_{n \in I_m} \hat{r}_n \gamma_{C_n}}{\bar{\sigma}^2_{S_e} \sum_{n \in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \tag{3.42}$$

Differentiating equation (3.42) with respect to γ_{C_η} results in

$$\frac{\partial \mathcal{L}}{\partial \gamma_{C_\eta}} = \sigma^2_{X_\eta} D'_\eta(\gamma_{C_\eta}) + \lambda \frac{\sigma^2_N \hat{r}_\eta}{\bar{\sigma}^2_{S_e} \sum_{n \in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \tag{3.43}$$

Setting the derivative equal to zero yields

$$-D'_\eta(\gamma_{C_\eta}) = \lambda \frac{\sigma^2_N \hat{r}_\eta}{\sigma^2_{X_\eta} \bar{\sigma}^2_{S_e} \sum_{n \in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \tag{3.44}$$

3.2.5.2 Fixed channel gain within a region

Using a single factor $\{\gamma_{A_m}\}_{m=1}^{M}$, to compensate in the transmitter for the channel gain per region, it was found in section 3.2.3.3 that the optimal value for γ_{A_m} is the centroid in each region. Inserting $\sigma_{G_m}^2$ in equation (3.39) into equation (3.21) results in the transmitted power in region m

$$\sigma_{S_m}^2(\gamma) = \frac{\bar{\sigma}_{S_e}^2 \bar{\sigma}_{G_m}^2}{\gamma_{A_m}\sigma_N^2} = \frac{\bar{\sigma}_{S_e}^2 \sum_{n\in I_m} \hat{r}_n \gamma_{C_n}}{\gamma_{A_m} \sum_{n\in I_m} \hat{r}_n}. \tag{3.45}$$

Normalizing the power, results in

$$\frac{\sigma_{S_m}^2(\gamma)}{\bar{\sigma}_{S_e}^2} = \frac{\sum_{n\in I_m} \hat{r}_n \gamma_{C_n}}{\gamma_{A_m} \sum_{n\in I_m} \hat{r}_n}. \tag{3.46}$$

Inserting equation (3.46) into equation (3.37), the object function is obtained

$$\mathcal{L} = \sum_{n=0}^{N-1} \sigma_{X_n}^2 D_n(\gamma_{C_n}) + \lambda \sum_{m=1}^{M} \frac{\sum_{n\in I_m} \hat{r}_n \gamma_{C_n}}{\gamma_{A_m} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \tag{3.47}$$

Differentiating (3.47) with respect to γ_{C_η}, leads to

$$\frac{\partial \mathcal{L}}{\partial \gamma_{C_\eta}} = \sigma_{X_\eta}^2 D'_\eta(\gamma_{C_\eta}) + \lambda \frac{\hat{r}_\eta}{\gamma_{A_m} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma \tag{3.48}$$

setting equal to zero

$$-D'_\eta(\gamma_{C_\eta}) = \lambda \frac{\hat{r}_\eta}{\sigma_{X_\eta}^2 \gamma_{A_m} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \tag{3.49}$$

3.2.5.3 Perfect channel information

With perfect channel knowledge, the scaling factor is continuously tracked, so the transmission power in channel region m is given by equation (3.15). Setting equation (3.39) in for $\sigma_{G_m}^2$ gives the transmitted power for region m as

$$\sigma_{S_m}^2(\gamma) = \frac{\bar{\sigma}_{G_m}^2 \bar{\sigma}_{S_e}^2}{\gamma \sigma_N^2} = \frac{\bar{\sigma}_{S_e}^2 \sum_{n\in I_m} \hat{r}_n \sigma_{G_n}^2}{\sigma_N^2 \gamma \sum_{n\in I_m} \hat{r}_n} = \frac{\bar{\sigma}_{S_e}^2 \sum_{n\in I_m} \hat{r}_n \gamma_{C_n}}{\gamma \sum_{n\in I_m} \hat{r}_n}. \tag{3.50}$$

Normalizing the power results in

$$\frac{\sigma_{S_m}^2(\gamma)}{\bar{\sigma}_{S_e}^2} = \frac{\sum_{n\in I_m} \hat{r}_n \gamma_{C_n}}{\gamma \sum_{n\in I_m} \hat{r}_n}. \tag{3.51}$$

Inserting equation (3.51) into equation (3.37) results in

$$\mathcal{L} = \sum_{n=0}^{N-1} \sigma_{X_n}^2 D_n(\gamma_{C_n}) + \lambda \sum_{m=1}^{M} \frac{\sum_{n \in I_m} \hat{r}_n \gamma_{C_n}}{\gamma \sum_{n \in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \tag{3.52}$$

Differentiating equation (3.52) with respect to γ_{C_η} leads to

$$\frac{\partial \mathcal{L}}{\partial \gamma_{C_\eta}} = \sigma_{X_\eta}^2 D'_\eta(\gamma_{C_\eta}) + \lambda \frac{\hat{r}_\eta}{\sum_{n \in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{1}{\gamma} f(\gamma) d\gamma \tag{3.53}$$

setting equal to zero

$$-D'_\eta(\gamma_{C_\eta}) = \lambda \frac{\hat{r}_\eta}{\sigma_{X_\eta}^2 \sum_{n \in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{1}{\gamma} f(\gamma) d\gamma. \tag{3.54}$$

3.2.6 Transmission

Above it has been shown how the source-blocks are preallocated both with respect to channel state m, and corresponding mapping. This is based upon the statistics of an ergodic channel, which in practice may not be fully ergodic for a limited number of channel samples. Especially if the image is compressed hard, when r_{avg} is low, there is a high probability that there will be a mismatch between the long term channel statistics, and the channel the transmitter will see when transmitting a single image. So the actual channel state probability will not necessarily be the same as assumed. To deal with this, the transmitter will need to have a strategy for channel statistics mismatch.

Since the source blocks are normalized, it is possible to combine pixels from different blocks for the same mapping as long as they are preallocated to the same channel state and mapping. In case there are no more blocks preallocated to the same mapping, or no more blocks are preallocated to a given state, the transmitter pads the input vector to the mappings with source dimension larger than one with zeros. Since it is assumed that the transmitter has instantaneous information about the channel, the transmitter will stop the coding of a given block in case the channel state changes. Since there is a maximum number of two channel symbols that are strictly dependent, the transmitter can change block to pick samples from in the middle of a block.

When the transmitter runs out of preallocated blocks for a given channel state, blocks preallocated to channel states with larger γ_{C_m} are chosen. When transmitting a source-block on a channel state that is worse than the one it is preallocated for, it means that the block might need to be transmitted with a mapping of higher rate to be able to get the same distortion.

Which will lead to an r_{avg} that is larger than assumed. The reason for still doing it this way, is that in the case of a tight time constraint, the receiver might have to stop receiving after a while, and then it is important to have the most important source-blocks received first. A similar argument can be made for a progressive image transmission-system, where the image is decoded as the information is received. It is in the interest of the receiver to have a good image quality as early as possible in the transmission process.

Transmission algorithm

1. Choose regions and representation points from pdf and average transmission power.
2. Preallocate source blocks to different channel regions such that the probabilities for each region match the theoretical as closely as wanted.
3. If rate-mismatch compensation on a block level is wanted, calculate optimal λ that minimizes distortion.
4. Start transmitting
   (a) Check channel state.
   (b) Are there more preallocated blocks for given channel state? If not, check if there are more blocks preallocated for channel states with higher CSNR. If not, check states for lower CSNR.
   (c) Pick samples from blocks preallocated to channel state. If the block was preallocated for a different channel state than the current, find new mapping suitable for new channel state. Encode samples with chosen mapping.
   (d) Scale output according to chosen CAS.
   (e) As long as there are more blocks left. Goto 4a. Else stop.

3.3 Reference systems

There are quite a few joint source channel coding systems for fading channels published, [Liu and Daut, 2005; Srinivasan and Chellappa, 1997; Thomos et al., 2005; Pan et al., 2006; Zhang et al., 2004b], to mention a few. The far most common schemes are however tandem schemes. As it is easier to compare the proposed system with tandem systems, two slightly different reference tandem systems will be used as comparison in this thesis.

As a source coder, the baseline JPEG2000 [Taubman and Marcellin, 2001; ISO/IEC, 2000] has been chosen. This is a well defined standard,

and can certainly be called state-of-the-art. For channel transmission, two different schemes are used, the first scheme employs an ACM consisting of M transmission rates (bits/channel symbol), each assumed to achieve AWGN capacity for a given CSNR. To further maximize the ASE, continuous power adaptation is used within each CSNR region [Gjendemsjø et al., 2005]. This scheme performs very close to the capacity for a Rayleigh fading channel, and the parameter settings are comparable to the proposed system when using CAS set to $\mathcal{C}$ for $M = 4$. The second scheme uses adaptive *turbo-coded modulation* (TuCM) to combat a flat fading channel. By varying the transmit power, constellation size and the turbo code, the scheme comes within 3 dB CSNR of the Rayleigh fading capacity when using perfect CSI. The target BER is set to be 10^{-6} [Vishwanath and Goldsmith, 2000, 2003]. Even though it was only shown that the TuCM scheme comes within 3 dB $\bar{\gamma}$ of the Rayleigh fading capacity for a average CSNR range of $0-10$ dB, it is assumed in the comparisons that the performance is valid for any CSNR value. The TuCM scheme uses $M = 3$, three different settings of the turbo encoder and constellation, plus an outage region.

The examples of the proposed coder are given for a certain average CSNR, ($\bar{\gamma}$), and an overall target compression-rate, r_{avg} (channel samples/pixel). To be able to compare with the reference systems, a bit-rate in bits/channel sample, R_c, is found for a given $\bar{\gamma}$ for the channel transmission reference scheme, and the resulting source bit-rate, R_s, in bits/pixel is found and given as a parameter to the reference image coder through

$$R_s = r_{\text{avg}} R_c. \tag{3.55}$$

Through equation (3.55) the source rate R_s is found for a given average rate change r_{avg}. This rate (bits/pixels) is then fed to the JPEG2000 encoder, and the *peak signal-to-noise ratio* (PSNR) value is found. In the case of the TuCM scheme, the average CSNR value is shifted 3 dB compared to the channel capacity.

The PSNR is defined as the ratio between the squared maximum pixel value, and the MSE on a pixel basis for the whole image, in dB

$$\text{PSNR} = 10 \log 10 \left(\frac{(x_{\max} - x_{\min})^2}{\frac{1}{S_x S_y} \sum_{i=0}^{S_x-1} \sum_{j=0}^{S_y-1} (x(i,j) - \hat{x}(i,j))^2} \right) \tag{3.56}$$

where S_x and S_y are the horizontal and vertical image sizes. $x(i,j)$ is the pixel value at position (i,j) in the original image, and $\hat{x}(i,j)$ is the pixel value at position (i,j) in the decoded image. $x_{\max}$ and $x_{\min}$ is the maximum and minimum possible pixel value. So for an image with 8 bits per pixel,$(x_{\max} - x_{\min})^2 = 255$.

Chapter 4

Theoretical JSCC systems

When considering practical implementation of communication systems, there will always be trade-offs:

- Where should the resources(bits) be used? Under a bit constraint, could the overall performance of the system be improved by using more bits in the channel coding, or in the source coding?
- During construction of a system, where in the system would an extra effort with respect to optimization contribute to the largest performance gain?
- Is extra complexity when implementing the system worth the effort?

Common for most source coders is the use of some sort of decorrelation as one of the first parts of the encoder.

There are many different schemes for compressing an decomposed image signal. Some schemes do not assume that the sub-bands are fully decomposed, and use relatively complex techniques such as vector quantization [Ramamurthi and Gersho, 1986; Westerink et al., 1988b] and trellis coded quantization [Jafarkhani et al., 1994; Kleider and Abousleman, 2000]. Most of the signal correlation should however be removed by the signal decomposition, and the decomposed signal is often modeled as memoryless, leaving room for simpler scalar quantization. The most successfull image coding standard, JPEG [ISO/IEC, 1991], uses a zig-zag scan of the cosine transform coefficients before quantization to achieve runs of similar coded symbols.

In subband coding, a solution can be to quantize each subband independently with a scalar quantizer [Tanabe and Farvardin, 1992; Taubman and Marcellin, 2001]. Other schemes might also exploit the inter sub-band energy dependencies by zero-tree coding [Shapiro, 1993; Said and Pearlman, 1996].

This chapter presents a method to approximate an upper bound for the performance of an image coder system, based on the use of source-splitting [Kang et al., 1994]. There it is assumed that an image can be decomposed into independent sub-sources $\{\sigma^2_{X_n}\}_{n=0}^{N-1}$. Other systems have tried similar ideas, e.g. in [Vaishampayan and Farvardin, 1990] a Gauss-Markov image model was used to estimate the rate distortion function of an image. In [Ruf and Modestino, 1999] they found the operational rate distortion performance of an image coder using a wavelet transform with separate coding of each subband.

The chapter is organized as follows: In Section 4.1 a scheme to estimate the theoretical limit for the system in Chapter 3 is presented. In Section 4.2 a system allowing arbitrary number of mappings with any given performance is presented.

This chapter is partly based on [Håkonsen and Ramstad, 2006b].

4.1 Estimating ideal image transmission system

For the scenario in this thesis, an image is decomposed by using a filter bank, and the image will be represented by a set of sub-bands, where each subband contains the information in a certain frequency band of the image. Due to the non-whiteness of the image, the statistics in different subbands will be different. In addition there will be local statistics *within* each subband. To capture these local variations, blocks of $B \times B$ subband samples are used to represent a sub-source of the image. In natural images, there are no discrete frequency components other than at zero. The mean of the lowpass-lowpass band is removed and coded seperately. The variance of each block can be estimated by using the mean squared value of the samples within a block given by equation (2.1). It has been indicated that the distribution of the samples within each such block can be modeled as Gaussian [Lervik and Ramstad, 1996]. As long as the blocks are smaller than a given subband, and a subband is sufficiently narrow, they can also be assumed to be white. Further, for a filterbank with ideally separated bands, the sub-band samples can also be assumed to be independent. By using these assumptions the rate distortion function for an image can be estimated by equation (2.6). The mismatch introduced by these assumptions is not considered, as it will be the same for both the proposed practical image transmission scheme, and for the theoretical scheme presented in this chapter. The same filter-bank, and blocks are used in the simulation of both cases.

For a given channel with known capacity, it is then possible to find the

average compression[1], r_{avg}, in channel samples/source samples, by combining equation (2.6) with the channel capacity, C, in a way given by equation (2.30), resulting in

$$r_{\text{avg}} = \frac{\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{2}\log_2\left(\frac{\sigma^2_{X_n}}{\sigma^2_{D_n}}\right)}{C(\text{CSNR})}\text{channel samples/source samples.} \tag{4.1}$$

From equation (4.1) it is seen that the channel capacity will have an impact on the performance of the theoretical system. An example is given in Figure 4.1 for the comparison of an AWGN channel, and a Rayleigh fading channel with and without CSI at the transmitter.

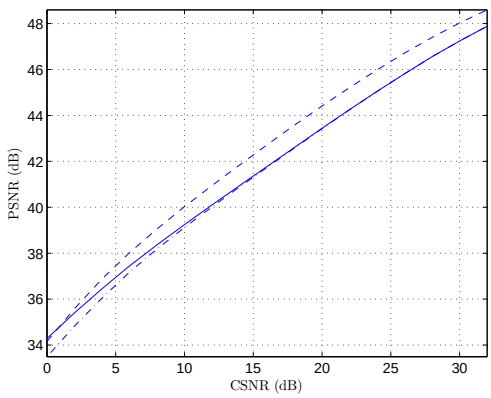


Figure 4.1: Estimated optimal performance for the "Lena" image with block-size 8×8. Transmitted over an AWGN channel (dashed) (2.13), and Rayleigh fading channel with Tx and Rx CSI (solid) equation (2.25), with only Rx CSI (dash-dotted) equation (2.23). Compression is at ratio 1 : 2, meaning that $r_{\text{avg}} = 0.5$.

Measuring image quality

For a white Gaussian source σ^2_X, the distortion distribution that achieves the rate distortion function is the Gaussian [Cover and Thomas, 1991]. To simulate the performance of a theoretical system for an actual image, AWGN with variance σ^2_D, is added to the source-blocks where $\mu < \sigma^2_{X_n}$. This way the image can be reconstructed through the synthesis filter-bank

[1]Instead of compression, bandwidth change could also be used.

after distortion is added. The reasoning behind this is that it gives the opportunity of calculating the PSNR value of the received image.

Block-size

The block-size $B \times B$ sets the level of how well it is possible to capture the local statistics in an image. In a practical system, the gain from reducing the block-size must be compared to the increased side-information. If the decoder is going to be able to decode such a system, the block variances and the mean of the lowpass-lowpass band would have to be known, thus giving a trade-off between the amount of side-information and the accuracy in capturing the local statistics. To avoid comparing on different premises, the block-size should then be kept equal for the theoretical estimator and the practical system.

Examples of different images simulated transmitted over an AWGN channel for three different block-sizes are given in Figure 4.2. Due to the different content in an image, the performance will be different for each image. For the "Lena" and the "Bridge" image there is about 8 dB difference in system performance. This shows that to be able to compare a practical image transmission system, the comparisons should be done on a per picture basis, to avoid the differences in performance due to the image statistics. It can also be seen that the block-sizes also play a vital role in the performance. As said above, the smaller the blocks are, the more accurate the statistics can be tracked, which can also be seen in Figure 4.2. One thing that has to considered is the number of blocks. For an image of size 512×512 this results in $(512/2)^2 = 65536$ number of blocks for the 2×2 block-size case, and $(512/8)^2 = 4096$ number of blocks for the 8×8 block-size case.

The difference in PSNR value for the "Lena", "Bridge" and "Goldhill" images seen in Figure 4.2 is due to the different frequency content of the images. By calculating an estimation of the spectral flatness measure based on the block variances, the difference can be found. The estimated spectral flatness Γ_X^2, is found from the ratio between the geometrical mean and arithmetical mean of the source-block variances

$$\Gamma_X^2 = \frac{\sqrt[N]{\prod_{n=0}^{N-1} \sigma_{X_n}^2}}{\frac{1}{N}\sum_{n=0}^{N-1} \sigma_{X_n}^2}. \tag{4.2}$$

Since the spectral flatness is a number less than one, the dB value of the inverse of Γ_X^2 is given in Table 4.1 for the different images. The difference in value between the different images corresponds well with the differences for the PSNR values seen in Figure 4.2.

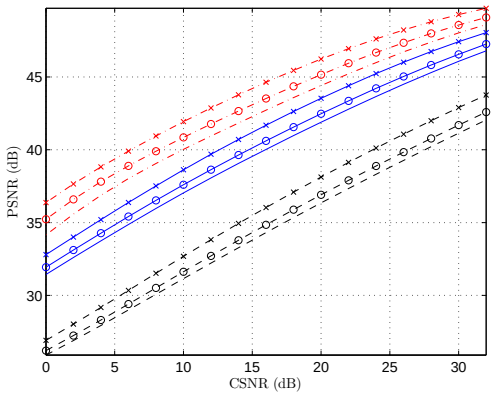


Figure 4.2: Estimated optimal performance for the "Lena" (dash-dotted), "Goldhill" (solid) and "Bridge" (dashed) images transmitted over an AWGN channel. Block-sizes are 2×2 ($\times$), 4×4 (o) and 8×8 (none). $r_{\text{avg}} = 0.5$

	Lena	Goldhill	Bridge
$1/\Gamma_X^2$ (dB)	20.0	17.6	12.3

Table 4.1: The inverse of spectral flatness for the different images in dB using 8×8 blocks.

The dB values of the source block variances sorted in decreasing order are given for comparison in Figure 4.3 for the different images.

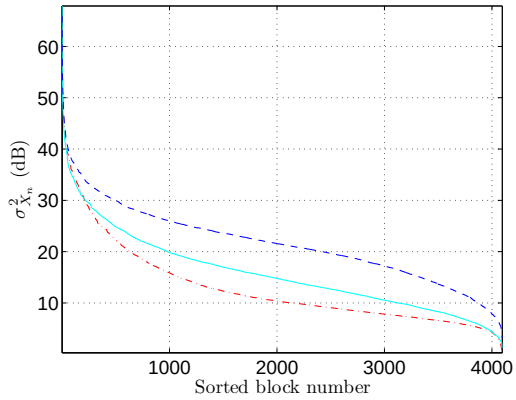


Figure 4.3: Sorted image block-variances in dB for 8×8 blocks for the "Lena" (dash-dotted), "Goldhill" (solid) and "Bridge" (dashed) images.

4.2 Nonlinear mappings

Using bandwidth changing mappings introduces the problem of finding the actual need for a given mapping-rate. Earlier systems using such mappings, e.g. [Lervik and Fischer, 1997b; Fuldseth, 1997; Coward and Ramstad, 2000b], first designed a mapping, then implemented it in a system to see the improvement in performance. Designing good mappings can be both hard and tedious. It would therefore be beneficial to be able to get an indication on the increased performance of a system if a mapping of a certain rate and performance is included in the available set, before actually designing the mapping.

The method presented in section 4.1 estimated an ideal system without considering implementation issues. In the case of dimensional changing mappings, this could be interpreted as a system having an infinite set of perfect mappings. For every arbitrary rate, there is a perfect mapping. In practice this would of course be impossible, not only because they are hard to design, but also because the system would be infinitely complex. So in practical systems, the design will lead to a suboptimal system. But by how much? To answer this it is possible to look at OPTA again. In section 2.3.1 the theoretical bounds for one given source transmitted over a given channel was presented. OPTA curves were also presented in Figure 2.4 for some di-

mensional changing rates. Inspired by the source-splitting aforementioned, it is then possible to look at each sub-source as a source on its own, and use the OPTA expressions to find the performance of a given sub-source being transmitted over a given channel. Combining expressions for all the different sub-sources would then give the performance of such a system when removing the sub-optimality factor of a given mapping. However, it would not be able to find the actual mapping-rate a block should be given, as this is a very complex task. The help such a system would give, is to able to examine the performance of a given block-to-mapping allocation. Finding the best block-mapping allocation was discussed in chapter 3.2.2.5. For now, it is assumed that each block already has been allocated to a mapping with rate $\hat{r}_n$.

The actual simulation of the transmission of each block would then work the same way as the simulation of the whole system. As each block is assumed to be i.i.d. white Gaussian, the optimal distortion is white Gaussian, with variance found by rewriting equation (2.32) slightly to

$$\sigma_{D_n}^2 = \sigma_{X_n}^2 (1+\gamma)^{-\hat{r}_n}, \tag{4.3}$$

where $\hat{r}_n$ is the mapping-rate block n is allocated. By adding this distortion to sub-source n, optimal transmission on a block basis is simulated. An example of this is given in Figure 4.4, for the image "Lena" transmitted over an AWGN channel. The blocks where the different mappings are used, can be clearly seen in a staircase pattern, where the SNR value of each step corresponds to the SNR for each mapping at $\bar{\gamma} = 24$ dB in Figure 2.4.

To see how good a mapping is, it is either possible to look at the difference in SNR for a given CSNR value, or it is possible to look at the difference in CSNR for a given SNR value. When simulating imperfect mappings, either a distance in SNR or CSNR compared to OPTA can be used. Looking at the performance of real mappings, the distance to OPTA in SNR is usually lower for low γ values than for high γ values. By using a fixed CSNR distance from OPTA when simulating mapping performance, this property is captured, as the slope of OPTA decreases for decreasing CSNR values. An example of this is given in Figure 4.5.

To find the distortion of a mapping with a penalty in CSNR equation (4.3) can be rewritten as

$$\sigma_{D_n}^2 = \sigma_{X_n}^2 \left(1 + 10^{\left(\frac{\gamma_{C_n\,\mathrm{dB}} - \nu_{\hat{r}_n\,\mathrm{dB}}}{10}\right)}\right)^{-\hat{r}_n}, \tag{4.4}$$

where $\gamma_{C_n\,\mathrm{dB}}$ is the representation value of the n'th block in dB, and $\nu_{\hat{r}_n\,\mathrm{dB}}$ is the CSNR penalty in dB for the mapping the n'th block is allocated to.

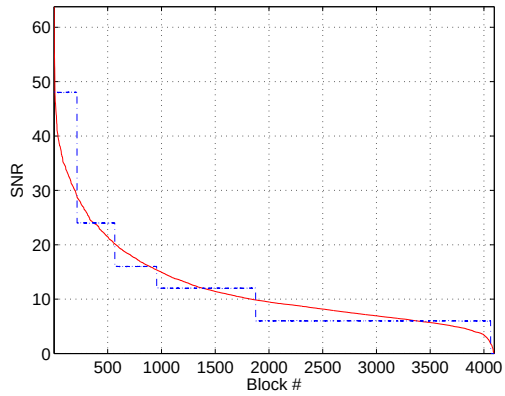


Figure 4.4: SNR of sorted blocks from the image "Lena", using ideal mappings. Optimal performance (solid), constant transmission power (dash dotted). $\hat{r}_{\text{avg}} = 0.5$, AWGN channel, $\bar{\gamma} = 24$ dB. Mapping rates used $\hat{r}_j \in \{0, 1/4, 1/2, 2/3, 1, 2\}$

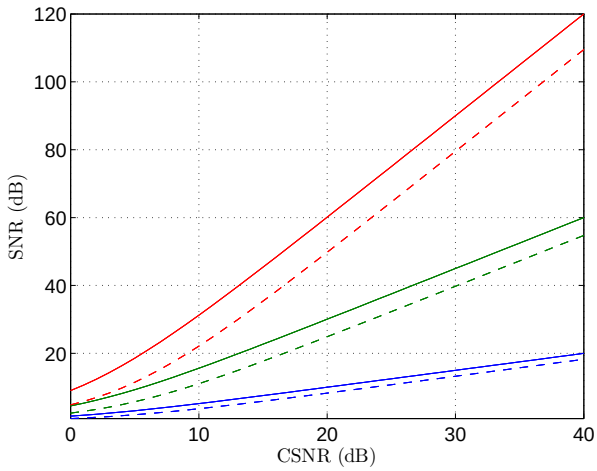


Figure 4.5: Performance of mappings of rate $\hat{r} = \{1/2, 3/2, 3\}$ from below. OPTA(solid), penalty of $\gamma = 3.5$ dB(dashed).

Equation (4.4) can be rewritten as

$$\sigma_{D_n}^2 = \sigma_{X_n}^2 \left(1 + \frac{\gamma_{C_n}}{\nu_{\hat{r}_n}}\right)^{-\hat{r}_n}. \tag{4.5}$$

So by using equation (4.5) to find the distortion, an arbitrary CSNR penalty can be added to the OPTA function for any value of $\hat{r}_n$. By using the presented ideas, a system with arbitrary many mappings with any given performance can be simulated, thus helping towards analyzing how much performance gain can be obtained by adding a certain mapping.

Preallocation when using theoretical mappings

To be able to compare on equal terms for theoretical mappings and implemented mappings, the same region thresholds and representation points will be used for theoretical mappings as for the implemented mappings. The difference in this chapter compared to Chapter 3, is that when finding the mapping a block should be coded with, the theoretical distortion is considered and not the distortion of the implemented mapping. This way a block might be coded with a mapping of lower rate due to increased performance. The optimization of the compensation for mapping rate-mismatch was previously done by estimating the derivative of the different mappings. In the following section, the same expressions will be found, but since theoretical mappings are used, expressions exists, and closed form solutions can be found.

4.2.1 Mapping rate-mismatch for simulated mappings

Through equation (4.5) it is now possible to set up an expression for the total distortion for a set of blocks by rewriting equation (3.36),

$$D = \sum_{n=0}^{N-1} \sigma_{D_n}^2 = \sum_{n=0}^{N-1} \sigma_{X_n}^2 \left(1 + \frac{\gamma_{C_n}}{\nu_{\hat{r}_n}}\right)^{-\hat{r}_n}. \tag{4.6}$$

Following the same lines as equation (3.37), a general Lagrangian can be set up,

$$\mathcal{L} = \sum_{n=0}^{N-1} \sigma_{X_n}^2 \left(1 + \frac{\gamma_{C_n}}{\nu_{\hat{r}_n}}\right)^{-\hat{r}_n} + \lambda \sum_{m=1}^{M} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{\sigma_{S_m}^2(\gamma)}{\bar{\sigma}_S^2} f(\gamma) d\gamma, \tag{4.7}$$

where $\sigma_{S_m}^2(\gamma)/\bar{\sigma}_S^2$ depends on the CAS.

If the transmitter has perfect channel information the CSNR value each block will be coded for, will also be the CSNR value the receiver sees. For

any other CAS this will not be the case. An extra distortion will be added due to mismatch on each block. The amount of distortion for each CAS is hard to estimate, so when simulating the performance of a theoretical estimate of the system performance, each block will get the distortion as if the CSNR level at the receiver is correct. For this reason a CAS that work poorer in the practical system, might perform better for the theoretical case. The exception is when the transmitter has full channel information, and $\mathcal{C}$ is used as CAS.

4.2.1.1 No pre-scaling of channel symbols

When not scaling the output power to adapt to the channel state, the expression for the normalized transmission power per channel state $\sigma^2_{S_m}(\gamma)/\bar{\sigma}^2_S$, is given equation (3.41). Inserting this into equation (4.7) results in

$$\mathcal{L} = \sum_{n=0}^{N-1} \sigma^2_{X_n} \left(1 + \frac{\gamma_{C_n}}{\nu_{\hat{r}_n}}\right)^{-\hat{r}_n} + \lambda \sum_{m=1}^{M} \frac{\sigma^2_N \sum_{n\in I_m} \hat{r}_n \gamma_{C_n}}{\bar{\sigma}^2_S \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \tag{4.8}$$

Differentiation with respect to γ_{C_η} yields

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \gamma_{C_\eta}} = & \frac{\sigma^2_{X_\eta}}{\nu_{\hat{r}_\eta}} (-\hat{r}_\eta) \left(1 + \frac{\gamma_{C_\eta}}{\nu_{\hat{r}_\eta}}\right)^{(-\hat{r}_\eta - 1)} \\ & + \lambda \frac{\sigma^2_N \hat{r}_\eta}{\bar{\sigma}^2_S \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \end{aligned} \tag{4.9}$$

Setting the derivative equal to zero results in

$$\frac{\sigma^2_{X_\eta}}{\nu_{\hat{r}_\eta}} (1 + \frac{\gamma_{C_\eta}}{\nu_{\hat{r}_\eta}})^{(-\hat{r}_\eta - 1)} = \lambda \frac{\sigma^2_N}{\bar{\sigma}^2_S \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \tag{4.10}$$

Isolating γ_{C_η} yields the expression for the optimal CSNR $\gamma^*_{C_\eta}$. This is the CSNR which the mapping source-block η is allocated to should be designed for. The result is

$$\gamma^*_{C_\eta} = \left(\exp \left(\frac{\log \left(\lambda \frac{\nu_{\hat{r}_\eta} \sigma^2_N}{\bar{\sigma}^2_S \sigma^2_{X_\eta} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma \right)}{(-\hat{r}_\eta - 1)} \right) - 1 \right) \nu_{\hat{r}_\eta}, \tag{4.11}$$

where λ is such that $\gamma^*_{C_\eta}$ fulfills the power constraint given by

$$\frac{\sigma^2_N}{\bar{\sigma}^2_S} \sum_{m=1}^{M} \frac{\sum_{n\in I_m} \hat{r}_n \gamma^*_{C_n}}{\sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma \leq 1. \tag{4.12}$$

4.2.1.2 Fixed channel gain within a region

When adapting to the channel on a state basis alone, the normalized transmission power per channel state, $\sigma^2_{Sm}(\gamma)/\bar{\sigma}^2_S$, is given equation (3.46). Inserting this into equation (4.7) yields

$$\mathcal{L} = \sum_{n=0}^{N-1} \sigma^2_{X_n} \left(1 + \frac{\gamma_{C_n}}{\nu_{\hat{r}_n}}\right)^{-\hat{r}_n} + \lambda \sum_{m=1}^{M} \frac{\sum_{n\in I_m} \hat{r}_n \gamma_{C_n}}{\gamma_{A_m} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \quad (4.13)$$

Differentiation with respect to γ_{C_η} yields

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \gamma_{C_\eta}} &= \frac{\sigma^2_{X_\eta}}{\nu_{\hat{r}_\eta}} (-\hat{r}_\eta) \left(1 + \frac{\gamma_{C_\eta}}{\nu_{\hat{r}_\eta}}\right)^{(-\hat{r}_\eta - 1)} \\ &\quad + \lambda \frac{\hat{r}_\eta}{\gamma_{A_m} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma. \end{aligned} \quad (4.14)$$

Setting the derivative equal to zero,

$$\frac{\sigma^2_{X_\eta}}{\nu_{\hat{r}_\eta}} \left(1 + \frac{\gamma_{C_\eta}}{\nu_{\hat{r}_\eta}}\right)^{(-\hat{r}_\eta - 1)} = \lambda \frac{1}{\gamma_{A_m} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma, \quad (4.15)$$

and isolating γ_{C_η} to find the expression for the optimal value yields

$$\gamma^*_{C_\eta} = \left(\exp \left(\frac{\log \left(\lambda \frac{\nu_{\hat{r}_\eta}}{\sigma^2_{X_\eta} \gamma_{A_m} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma \right)}{(-\hat{r}_\eta - 1)} \right) - 1 \right) \nu_{\hat{r}_\eta}, \quad (4.16)$$

where λ is such that $\gamma^*_{C_\eta}$ fulfills the normalized power constraint given by

$$\sum_{m=1}^{M} \frac{\sum_{n\in I_m} \hat{r}_n \gamma^*_{C_n}}{\gamma_{A_m} \sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} f(\gamma) d\gamma \leq 1. \quad (4.17)$$

4.2.1.3 Perfect channel information

For the case of perfect channel information, the object function is found by inserting the equation for the normalized transmission power per channel region $\sigma^2_{Sm}(\gamma)/\bar{\sigma}^2_S$, given by equation (3.51), into equation (4.7) giving

$$\mathcal{L} = \sum_{n=0}^{N-1} \sigma^2_{X_n} \left(1 + \frac{\gamma_{C_n}}{\nu_{\hat{r}_n}}\right)^{-\hat{r}_n} + \lambda \sum_{m=1}^{M} \frac{\sum_{n\in I_m} \hat{r}_n \gamma_{C_n}}{\sum_{n\in I_m} \hat{r}_n} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{1}{\gamma} f(\gamma) d\gamma. \quad (4.18)$$

Differentiating with respect to γ_{C_η} yields

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \gamma_\eta} &= \frac{\sigma^2_{X_\eta}}{\nu_{\hat{r}_\eta}}(-\hat{r}_\eta)\left(1+\frac{\gamma_{C_\eta}}{\nu_{\hat{r}_\eta}}\right)^{(-\hat{r}_\eta-1)} \\ &\quad + \lambda \frac{\hat{r}_\eta}{\sum_{n\in I_m}\hat{r}_n}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{1}{\gamma} f(\gamma)d\gamma.\end{aligned} \tag{4.19}$$

Setting the derivative equal to zero results in

$$\frac{\sigma^2_{X_\eta}}{\nu_{\hat{r}_\eta}}\left(1+\frac{\gamma_{C_\eta}}{\nu_{\hat{r}_\eta}}\right)^{(-\hat{r}_\eta-1)} = \lambda \frac{1}{\sum_{n\in I_m}\hat{r}_n}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{1}{\gamma} f(\gamma)d\gamma. \tag{4.20}$$

Isolating γ_{C_η} to find the optimal value

$$\gamma^*_{C_\eta} = \left(\exp\left(\frac{\log\left(\lambda \frac{\nu_{\hat{r}_\eta}}{\sigma^2_{X_\eta}\sum_{n\in I_m}\hat{r}_n}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}}\frac{1}{\gamma}f(\gamma)d\gamma\right)}{(-\hat{r}_\eta-1)}\right)-1\right)\nu_{\hat{r}_\eta}, \tag{4.21}$$

is found. To find the optimal value of λ, this expression for $\gamma^*_{C_\eta}$ has to satisfy the expression for the normalized power constraint, given by

$$\sum_{m=1}^{M}\frac{\sum_{n\in I_m}\hat{r}_n\gamma^*_{C_n}}{\sum_{n\in I_m}\hat{r}_n}\int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}}\frac{1}{\gamma}f(\gamma)d\gamma \leq 1. \tag{4.22}$$

4.2.2 Effect of mapping-rate mismatch for ideal system

An example of the effect when adding mapping-rate mismatch compensation for ideal mappings for the case of transmitting "Lena" over an AWGN channel is given in Figure 4.6. The figure shows how the mapping-rate mismatch compensation tries to approximate the optimal performance by shifting power for the different mappings.

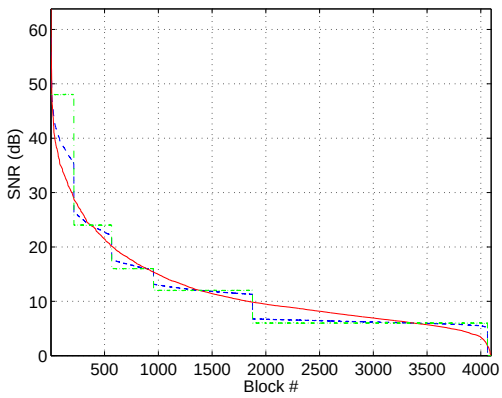


Figure 4.6: SNR of sorted blocks from the image "Lena", using ideal mappings. Optimal performance(solid), constant transmission power, $\mathfrak{NB}$ (dash dotted), mapping-rate mismatch compensation(dashed). $\hat{r}_{\text{avg}} = 0.5$, AWGN channel, $\bar{\gamma} = 24$ dB. Mapping rates used $\hat{r}_j \in \{0, 1/4, 1/2, 2/3, 1, 2\}$

Chapter 5

Simulations

In this chapter, simulation results for the proposed system is presented for three different images. The original images are printed in Appendix E. The channel is simulated by multiplying the transmitted symbols with a correlated Rayleigh distributed random sequence, before adding white Gaussian noise. This means that it is assumed that the fading is constant within one transmitted channel symbol. The channel is simulated according to the Jakes correlation model [Jakes, 1974]. As a reference, comparisons with the ideal system in Chapter 4 are also included.

In Section 2.2.2, the expected CSNR, $\bar{\gamma}_e$, was given by equation 2.18. This was based on the long term transmission power, $\bar{\sigma}^2_{S_e}$. During all calculations done for the preallocation of source-blocks with respect to power and rate, $\bar{\sigma}^2_{S_e}$ was used. For the transmission of a single image, there will be a finite number of transmitted channel symbols. The channel will in such a case not be seen as *ergodic*, i.e. it will not follow its pdf exactly. A consequence of this, is that when the transmitter has CSI, the *true* transmitted power, $\bar{\sigma}^2_{S_t}$, will not be exactly equal to the long term transmitted power $\bar{\sigma}^2_{S_e}$. Transmitting the same image for several channel realisations will result in a different value of $\bar{\sigma}^2_{S_t}$ for each transmission. To show the spread of transmission power the transmitter experiences, the true CSNR for a transmission, $\bar{\gamma}_t$, will be defined as

$$\bar{\gamma}_t = \frac{\bar{\sigma}^2_{S_t}}{\sigma^2_N}. \tag{5.1}$$

As a reference, the average of $\bar{\gamma}_t$ over D transmissions will be given by

$$\bar{\gamma} = \frac{1}{D}\sum_{d=0}^{D-1} \bar{\gamma}_{t_d}. \tag{5.2}$$

Since the expected value of the channel gain is unity, $\bar{\gamma}$ will converge to $\bar{\gamma}_e$ for increasing sample size D. In addition to changing the true transmission power, the PSNR value of the received image will change depending on the channel. The PSNR value from the transmission of one image will be given by PSNR_t. For most of the results presented, the average value of PSNR_t over D simulations will be used. The last parameter which will change depending on the channel is the ratio of the number of channel symbols and source symbols r_{avg}. For one transmission of an image, the resulting ratio will be given by r_{avg_t}. Most of the results will however be given for a target ratio r_{avg}. In this thesis, the *performance* of the system for a given image for a set of parameters, are found by plotting the average PSNR values found from Monte Carlo simulation. The average PSNR values are plotted as a function of $\bar{\gamma}$.

In this chapter, results of simulations done with the framework presented in Chapter 3 and Chapter 4 will be presented. The main focus will be on performance and analysis of the system, and comparison with the given reference systems. Special emphasis will be given to the different ways of adapting to the channel gain. Either through a given *channel adaptation strategy* (CAS), which can be not to adapt the transmitted signal at the output, $\mathcal{N}$, adapting with a single factor for each channel state, $\mathcal{S}$, or continuously scaling of the output with full *channel state information* (CSI) at the transmitter, $\mathcal{C}$. The differences when compensating for rate mismatch due to a finite set of mappings through power allocation on a block basis, $\mathfrak{B}$, and not, $\mathfrak{NB}$, will also be analyzed.

The chapter is organized as follows. In Section 5.1 the results and importance of the different levels of adaptation are given. Section 5.2 presents the impact of the Doppler shift on the system. In Section 5.3 the proposed system is compared with different ideal scenarios, and two reference systems. A discussion of the results is provided on Section 5.4.

Transmission parameters

Unless said otherwise, the parameter values used in the simulations are given by Table 5.1.

5.1 Effect of using CSNR regions

5.1.1 One CSNR region

Assume that the transmitter only knows the pdf of the channel. Then the adaptation will depend on the receiver. If it is assumed that the receiver has

Parameter	Symbol	Value
Carrier frequency	f_c	2.0 GHz
Symbol duration	T_s	4 μs
Doppler shift	f_m	100 Hz
Statistical sample size	D	2000
Mobile velocity	v	15 m/s

Table 5.1: Simulation parameters

perfect CSI, the receiver can partly compensate for the channel variations. At the other extreme case, the receiver does not have any CSI either. For a practical scenario the situation is likely to be somewhere in between. As the CSI has to be estimated, it will not be possible to do this perfectly, whether this is done through the use of pilots or some other method[1].

5.1.1.1 Blind broadcast: No Tx or Rx CSI

Although the assumed model using PAM symbols over a fading channel presented in Section 2.2.3 does not hold when the receiver does not have full channel information, an example is included to demonstrate the inherent robustness of the system.

In a broadcast situation where neither the transmitter nor the receiver have knowledge about a time varying channel, a transmission system will have to rely on robustness to deliver the information. For the image coder used here, it means that the CSNR will be relatively far away from the value the mappings were optimized for. Due to the robustness of the mappings it is, however, still possible to see the content of the image. One consequence of the lack of adaptation to the channel, is that the channel will be different during the transmission of an image from time to time. So for each transmitted image, the source samples will be distorted differently. The effect of this can be seen in Figure 5.1, where two examples are given for the same transmission parameters. Even though the PSNR values are similar, the two images are quite different. The most striking difference is that Figure 5.1(a) has more contrast than Figure 5.1(b), indicating that the intensity of the pixel values have been reduced in Figure 5.1(b). To compensate for this, one might think of a scheme where the contrast of an image is captured by some parameter and sent as side-information. It might also be an idea to allocate the source-blocks to a higher mapping-

[1]It is possible to consider a scheme where the transmit power is constant, and the receiver compensates for channel gain by estimating received power over some time window.

(a) PSNR = 17.1 dB (b) PSNR = 17.5 dB

Figure 5.1: Two examples of the received "Lena" image for the case of $\bar{\gamma} = 14.8$ dB, $r_{\text{avg}}= 0.5$. No CSI knowledge at transmitter and receiver.

rate than usually to increase the robustness of the system. This way rate is traded off for robustness.

5.1.1.2 No Tx CSI(no outage), perfect Rx CSI

If the receiver has perfect CSI, it is possible for the receiver to try and compensate for the perturbation done by the channel. One option is simply to invert the channel gain $\sqrt{\alpha}$, but by doing so the additive noise will also be enhanced as mentioned in Section 3.2.3.2. A better solution is to find a trade-off between the channel gain and additive noise by using a filter given by equation (3.20). By inverting the channel gain, the noise is enhanced for symbols resulting in distinct distorted areas in the images as seen in an example in Figure 5.2(a). Above the brim of the hat, a checkered pattern is clearly visible. For the case when the filter is used, these areas are smoother, giving a more pleasant image to look at, even though the PSNR value is smaller, as seen in Figure 5.2(b).

In the case of no outage, the transmission power is equal for each simulation, as the transmitter does not adapt to the channel during transmission. Since the statistics for the source-blocks are slightly different for different images, $\bar{\gamma}_t$ will not be exactly equal to the target value $\bar{\gamma}_e$. The statistics of the source-blocks will not be exactly equal to the ones used in the design of the mappings. For the "Lena" image this will, for instance, result in $\bar{\gamma}_t = 13.8$ dB, when aiming for $\bar{\gamma}_e = 14$ dB. For other images this might be slightly above, e.g. "Bridge".

(a) PSNR = 29.3 dB (b) PSNR = 28.5 dB

Figure 5.2: Example of received image with inversion of channel gain (left) and when using channel gain mismatch filter (right), for the "Lena" image. $\bar{\gamma} = 13.8$ dB, $r_{\text{avg}}= 0.5$. No CSI at Tx.

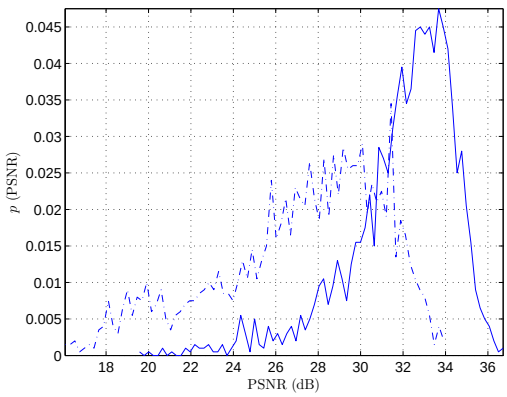
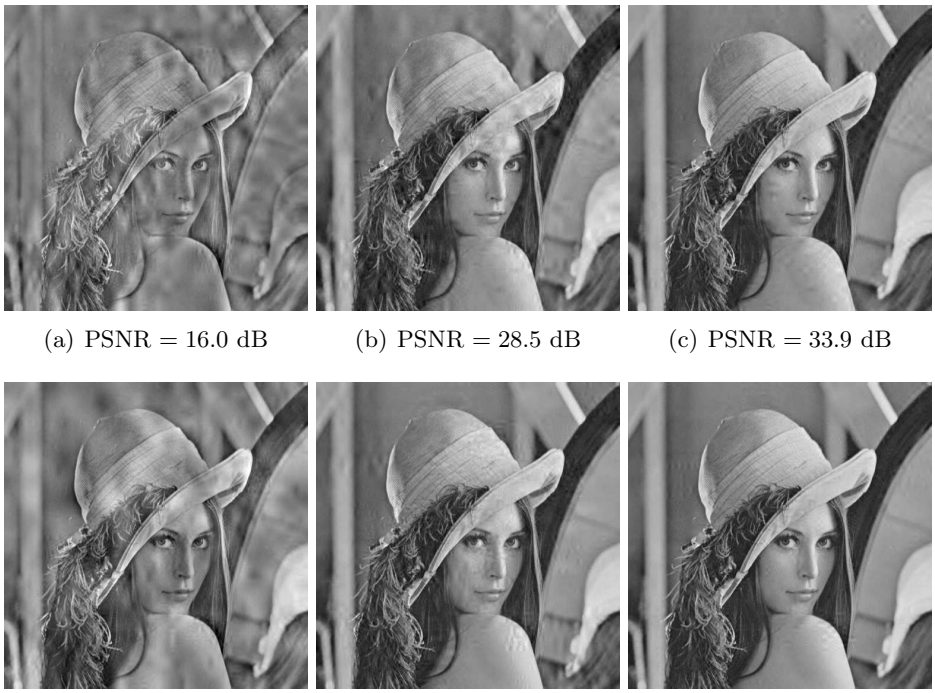


Figure 5.3: Estimated pdf of received image quality. "Lena" image. $\bar{\gamma} = 13.8$ dB, $r_{\text{avg}} = 0.5$, $\mathfrak{NB}$ (dash-dotted), $\mathfrak{B}$ (solid). No CSI at Tx, full CSI at Rx.



(a) PSNR = 16.0 dB (b) PSNR = 28.5 dB (c) PSNR = 33.9 dB

(d) PSNR = 20.5 dB (e) PSNR = 32.5 dB (f) PSNR = 36.9 dB

Figure 5.4: Image comparison for the received "Lena" image. $\bar{\gamma} = 13.8$ dB, $r_{\text{avg}} = 0.5$. Lower part of estimated pdf (left), mean (middle) and high (right). $\mathfrak{NB}$ (upper row), $\mathfrak{B}$ (lower row) No CSI at Tx. Full CSI at Rx.

The spread in the quality of the received image can be shown as an estimated pdf. In this thesis, such a pdf is estimated by a normalized histogram[2].

So far the results have been given for the case when all source-blocks are coded for a common CSNR level γ_{C_1}. In Figure 5.3 an estimated pdf is plotted when using mapping-rate mismatch compensation $\mathfrak{B}$, and without ($\mathfrak{NB}$). Compensating for the mapping-rate mismatch is done through shifting power from blocks that are over-protected to the blocks that are under-protected. So that each block is coded with a mapping optimized for different CSNR values. In Figure 5.4 examples of images representing the extreme and mean values in Figure 5.3 are shown. It can be seen how compensating for mapping-rate mismatch can give an overall better average quality.

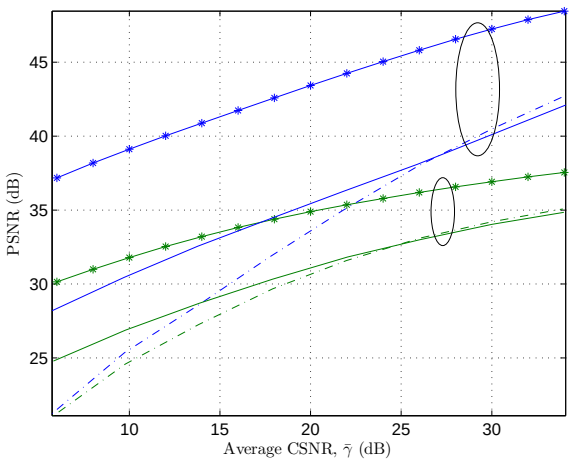


Figure 5.5: Performance with no Tx CSI, full Rx CSI for the "Lena" image, with block rate compensation (solid), and equal power for all blocks (dash-dotted). Estimated optimal system for no Tx CSI (solid*). For $r_{\text{avg}} = 0.5$ (upper group) and $r_{\text{avg}} = 0.1$ (lower group).

The performance for the "Lena" image is plotted in Figure 5.5 for two different r_{avg} values. As a reference, a theoretical ideal system[3] transmitting at channel capacity, equation (2.23), is plotted. The PSNR values given in the figure, are the average values of the different PSNR_t values

[2] 80 regions are used, where the regions are equally spaced from the lowest to the highest PSNR value in the statistical data set for a given parameter setting.

[3] This system is discussed in Section 4.1 and simulate a system using an infinite number of mappings performing according to OPTA.

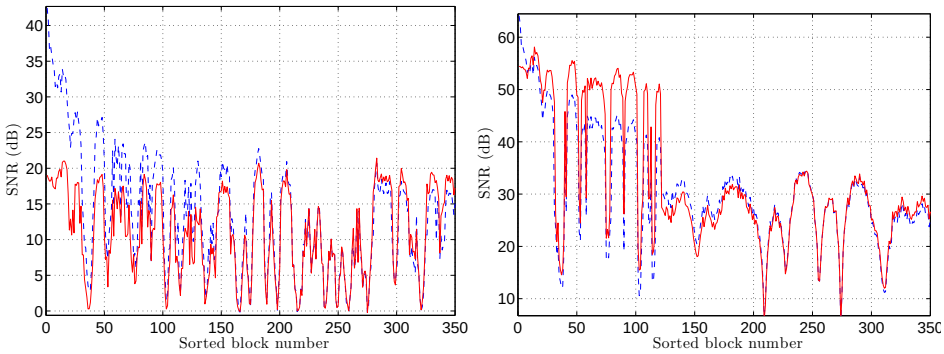
from D simulations for a target $\bar{\gamma}_e$ value. As mentioned above, when the transmitter has no CSI, the transmission power will not change for different transmissions. For this reason, the average CSNR value, $\bar{\gamma}$, in Figure 5.5 will be equal to the true CSNR value $\bar{\gamma}_t$.

In Figure 5.5 there are a few things worth mentioning. The first is that the performance is further away from the ideal system for the case were $r_{\text{avg}} = 0.5$ than when $r_{\text{avg}} = 0.1$. The reason for this is that there are more source-blocks being transmitted when r_{avg} is higher. Or in other words, there are fewer blocks coded with a mapping of rate $\hat{r}_j = 0$. This mapping performs according to OPTA, so for the case where $r_{\text{avg}} = 0$, the system would be performing optimally, this case is however of very little practical interest. This behaviour can be seen throughout the thesis. Another interesting thing in Figure 5.5, is that if the transmission power is above a certain threshold, it does not pay off to compensate for mapping-rate mismatch. The reason for this can be found in the perturbation from the channel. When the transmitter has no CSI, the reduced distortion when compensating for rate-mismatch, is lost in the perturbation from the channel. This can be seen in Figure 5.6, where the SNR of the first 350 blocks are plotted at 10 dB and 30 dB. Since there is no CSI, the blocks with largest variance are transmitted first. In Figure 5.6(a) the blocks with the highest variance should ideally be coded with a mapping with rate higher than $\hat{r} = 2$, but since it is the highest mapping-rate available, more power is allocated to these blocks to reduce the distortion. When power allocation for each block is enabled, extra power leads in addition to extra robustness towards the channel fades for these blocks. In Figure 5.6(b), the average power is high enough so that the blocks with high variance do not need the full protection that the mapping of rate $\hat{r} = 2$ provides. The power for these blocks are then reduced, but at the cost of poorer robustness towards deep channel fades.

In Figure 5.7, the estimated PSNR distribution is plotted for target $\bar{\gamma}_e = \{10, 30\}$ dB for $r_{\text{avg}} = 0.5$. Through Figure 5.7(a) it can be seen how using power to compensate for mapping-rate mismatch will help to narrow the distribution of the received PSNR. As the average CSNR increases, the PSNR distribution will also be narrower when not compensating for mapping rate-mismatch as shown in Figure 5.7(b).

5.1.1.3 Tx and Rx CSI

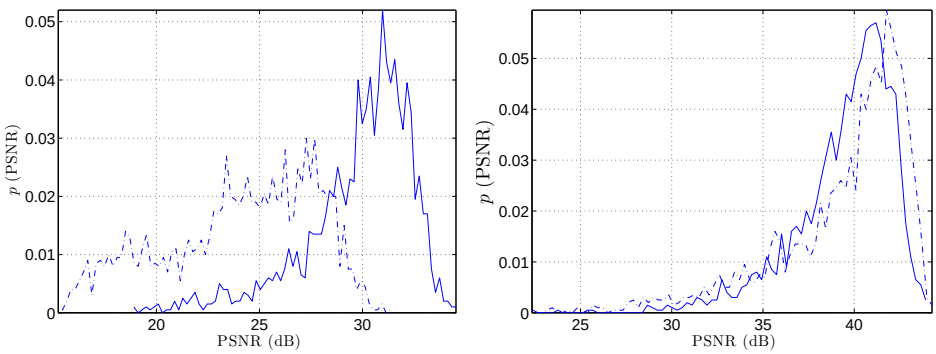
When the transmitter has CSI, there is an added possibility of choosing not to transmit if the channel condition is too poor. Outage is declared and the transmitter chooses not to transmit. With CSI the transmitter can also start adapting the information to the channel in a more sophisticated



(a) $\bar{\gamma} = 9.7$ dB. $\mathfrak{NB}$: PSNR = 29.1 dB. $\mathfrak{B}$: PSNR = 32.3 dB

(b) $\bar{\gamma} = 29.9$ dB $\mathfrak{NB}$: PSNR = 42.8 dB. $\mathfrak{B}$: PSNR = 42.1 dB

Figure 5.6: Example of SNR values in dB for the first 350 transmitted blocks for the "Lena" image. No Tx CSI, full Rx CSI. $r_{\text{avg}} = 0.5$. $\mathfrak{NB}$ (solid), $\mathfrak{B}$ (dashed).



(a) $\bar{\gamma} = 9.7$ dB

(b) $\bar{\gamma} = 29.9$ dB

Figure 5.7: Estimated pdf for the received PSNR values for the "Lena" image. No Tx CSI, full Rx CSI. $r_{\text{avg}} = 0.5$. $\mathfrak{NB}$ (dash-dotted), $\mathfrak{B}$ (solid).

way. Up till now, all the results have been given for the case when the transmitted symbols are not adapted to the channel gain. Now the option of adapting to the channel gain will be used. Assume that the CSI only has two states: transmit mode and outage, meaning that $M = 1$. The amount of CSI will in that case be small, but it will also reduce the CAS to either $\mathcal{N}$ or $\mathcal{S}$. Allowing continuous CSI will give the additional option of $\mathcal{C}$. Of these different techniques $\mathcal{C}$ will naturally give the best PSNR value for a given average CSNR. With $\mathcal{C}$ as CAS, the performance for "Lena" is given in Figure 5.8 with and without mapping-rate mismatch compensation. The figure shows that allowing compensation for mapping-rate mismatch can give a significant gain for low CSNR values, and that this gain reduces as the average CSNR increases. Allowing the transmitter to have full information about the channel at all times, will remove the crossing of the $\mathfrak{B}$ and $\mathfrak{NB}$ curves, as seen in Figure 5.5. The reason why this does not happen when using $\mathcal{C}$ is that the need for robustness is removed. The received signal will always have the CSNR value the mappings were designed for. One interesting thing to notice is the threshold that seems to be at $\bar{\gamma} = 10$ dB. This threshold is due to the outage level γ_{T_0}. For $M = 1$ there is only one representation point γ_{C_1} that the different source-blocks are coded for. Since no power is used in the case of outage, increasing the outage threshold γ_{T_0} will increase the CSNR value γ_{C_1} for which the blocks can be coded. Even though an increased γ_{T_0} will mean less time to transmit, it will also mean that the rate of a mapping can be decreased, leading to fewer channel samples per source-block. Fixing γ_{T_0} for $M = 1$ means that there are actually no parameters to adjust, as γ_{C_1} will be determined by the choice of CAS, and power constraint. As mentioned before, the optimal γ_{T_0} level will not be found. Due to the simplicity of $M = 1$, a few examples showing how important the outage level seems to be for low average CSNR values when $M = 1$, is however included in Figure 5.9. It is not until there are multiple representation points there is a need for a more complex algorithm, so for now all results for $M = 1$ are plotted when using the *simple* algorithm to find the representation point γ_{C_1}. Figure 5.9(a) show that significant gain can be achieved by adjusting the outage level. This should be investigated more in future research.

In Figure 5.10 the performance for "Lena" of the system when using no adaptation ($\mathcal{N}$), and a single scaling factor ($\mathcal{S}$) as CAS is shown. Figure 5.10(a) show that the performance with less CSI will give poorer performance than full CSI, especially in the case when mapping-rate mismatch compensation is used. For $r_{\text{avg}} = 0.5$ the system seems to loose performance due to poor choice of outage level, as shown in Figure 5.9(a). For the case when mapping-rate mismatch is not used, the loss is not quite as much compared to $\mathcal{C}$ when comparing the curves in Figure 5.10(b).

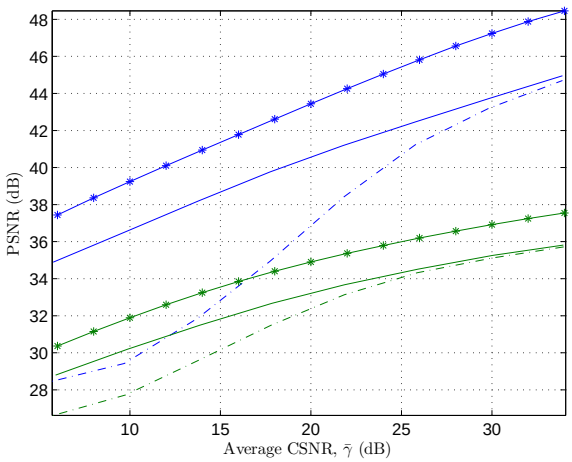
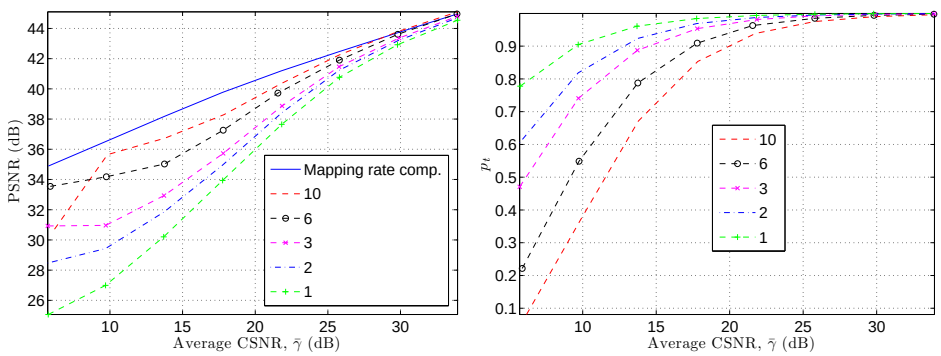
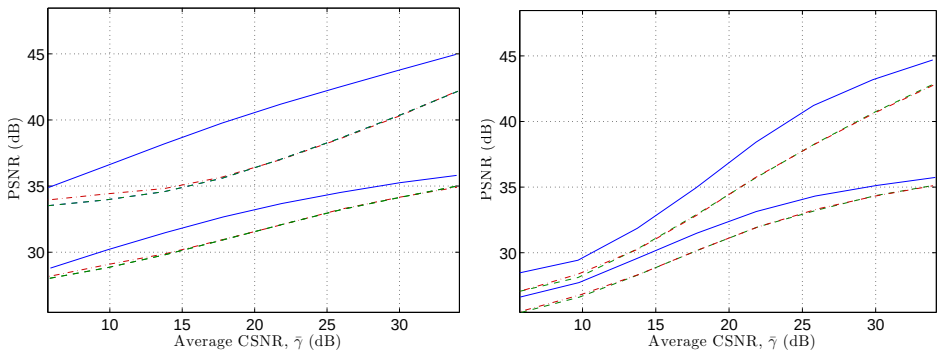


Figure 5.8: Comparison between using $\mathfrak{B}$ (solid), $\mathfrak{NB}$ (dash-dotted) and estimated optimal system (solid*). For $r_{\text{avg}} = 0.5$ (upper curves) and $r_{\text{avg}} = 0.1$ (lower curves). CAS is $\mathcal{C}$. "Lena" image. $M = 1$.



(a) Performance (PSNR) (b) Probability of transmission p_{t}

Figure 5.9: Performance and transmission probability of system for different γ_{T_0} values. For the case when $M = 1$, CAS is $\mathcal{C}$, and for $\mathfrak{NB}$. The performance when $\mathfrak{B}$ is used, and $\gamma_{T_0} = 2$ is included as a reference (Mapping rate compensation).



(a) $\mathfrak{B}$ (b) $\mathfrak{NB}$

Figure 5.10: Comparison for different channel adaptation strategies, $\mathcal{C}$ (solid-blue), $\mathcal{N}$ (dash-dotted-red) and $\mathcal{S}$ (dashed-green). $r_{\text{avg}} = 0.5$ (upper curves) and $r_{\text{avg}} = 0.1$ (lower curves) "Lena" image. $M = 1$.

5.1.2 Multiple channel regions with transmission

Allowing $M > 1$ will give the transmitter multiple channel regions so that it is possible to adjust power and rate more accurate to the channel state. For $\mathcal{N}$ and $\mathcal{S}$ multiple γ_{C_m} values will give a smaller mismatch between the CSNR value that a block is coded for, and the CSNR seen at the receiver. In addition, the protection given by the mappings can be more accurately set according to the target distortion-level μ. The effect of this will be addressed later, but first the impact of the different channel region optimization methods presented in Section 3.2.3 will be addressed.

5.1.2.1 Effect of optimization method

In Section 3.2.3.1 some problems regarding the optimization of the channel region representation points $\{\gamma_{C_m}\}_{m=1}^{M}$, and channel region thresholds $\{\gamma_{T_m}\}_{m=1}^{M-1}$ were mentioned. The different approaches were:

- The use of a complex numerical optimization to find optimum, denoted *complex*.
- Using a simpler optimization algorithm, initialized from equiprobable regions, denoted *simple*.

The different approaches will give different results, but how different will they be?

Figure 5.11 shows the comparison between the *complex* and *simple* algorithms for the "Lena", "Goldhill" and "Bridge" images. It is seen that for all cases, except for high $\bar{\gamma}$ values when CAS is set to $\mathcal{C}$, the *simple* algorithm

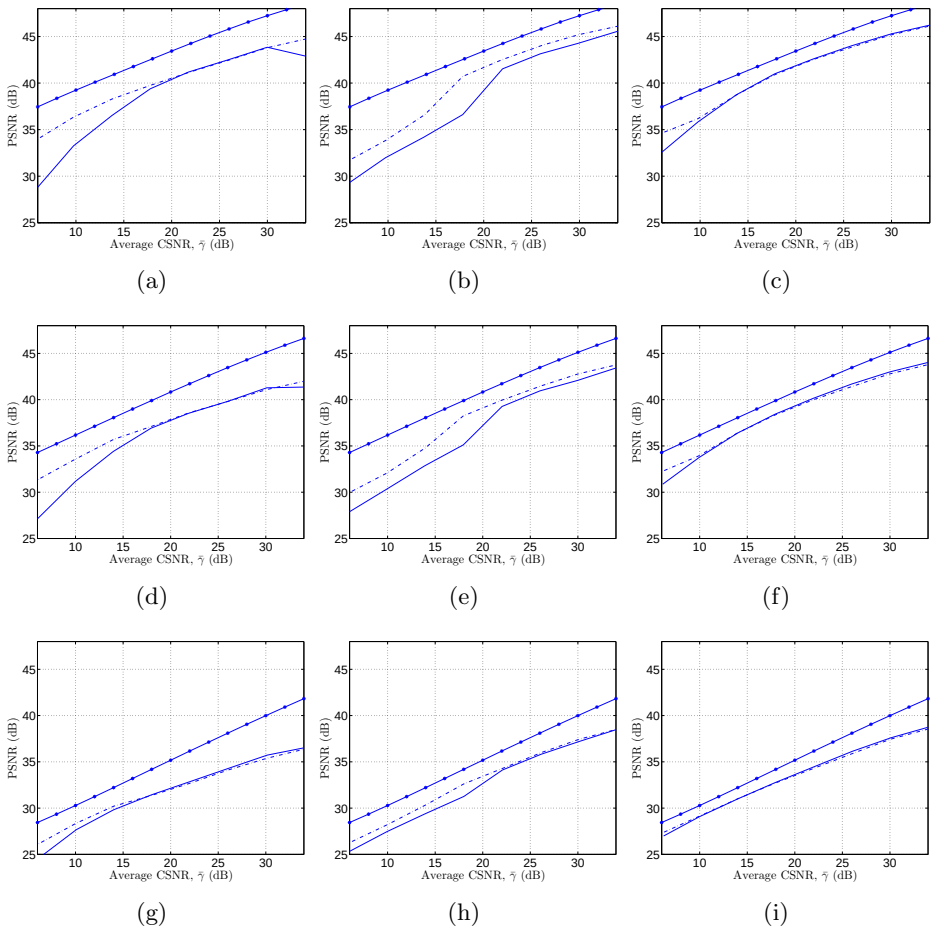


Figure 5.11: Image comparison for the received "Lena" (top row), "Goldhill" (middle row), and "Bridge" (bottom row) images for $M = 4$, $r_{\text{avg}} = 0.5$. $\mathcal{N}$ (left), $\mathcal{S}$ (middle) and $\mathcal{C}$ (right). $\mathfrak{NB}$, *simple* algorithm (dash-dot), *complex* algorithm (solid), estimated optimum (solid*).

outperforms the *complex* algorithm. There are, however, some interesting features that should be explained further. In Figures 5.11(a) to 5.11(c) it is seen how the performance is different depending on the CAS. A natural question to ask is: why is the difference in performance so large? For the *simple* algorithm, there is a difference of about 2 dB for low average CSNR values between the $\mathcal{N}$ and $\mathcal{S}$ scenarios. For the *complex* algorithm, the performance is comparable for the $\mathcal{N}$ and $\mathcal{S}$ scenarios, but $\mathcal{C}$ is approximately 3 dB better. To find the explanation for the differences, it is possible to look at the representation values $\{\gamma_{C_m}\}_{m=1}^{M}$. In Figures 3.11(a), 3.13(a) and 3.15(a), these values are given for the *simple* algorithm. It is seen how both $\mathcal{N}$ and $\mathcal{C}$ has $\gamma_{C_4} \approx 15$ dB, while $\mathcal{S}$ has $\gamma_{C_4} \approx 12.5$ dB for $\bar{\gamma} = 6$ dB. The effect of this can be seen in Figure 5.12, where examples for the SNR value of each block is plotted. The blocks are sorted decreasingly according to their variance, to make the effect of the γ_{C_m} values easier to see. From looking at Figure 5.12(b), it is seen how the blocks with largest variance have approximately 20 dB SNR. For $\gamma_C = 12.5$ dB, the performance is 20 dB SNR for the mapping with $\hat{r}_5 = 2$, see Figure 3.9. Compared to the SNR for each block for the case when the distortion is equal to the distortion-level μ used in the simulation. It is seen that even though the blocks with large variance are coded with the highest dimensional mapping, the loss compared to the distortion-level μ is large. The blocks with large variance will have large distortion, contributing to a large loss as seen in Figure 5.11(b). Looking at Figure 5.11(a), the performance is much better for the *simple* algorithm. The reason why this is happening can be found from Figure 5.12(a), and Figure 5.13(a). In the latter of the two figures, it is shown how the source-blocks are preallocated, both to mapping rate and representation value γ_{C_m}. Figure 5.13(a), shows that the blocks with largest variance are allocated to $\gamma_{C_4} = 15.3$ dB, and $\hat{r}_5 = 2$. The result is clearly seen in Figure 5.12(a): the blocks preallocated to the best channel region get an SNR value around 30 dB, which is clearly better than for the case of $\mathcal{S}$. One thing that is strange with this, is that from Figure 3.9, a mapping of $\hat{r}_5 = 2$ optimized for 15 dB should give an SNR value at around 24 dB. To find the 6 dB missing, the channel-gain mismatch filter in the receiver has to be considered as well. For $\mathcal{N}$, this filter is given by equation (3.20). When transmitting in the best channel state, the gain of the channel is large. In the case of $\mathcal{N}$, *simple*, $\bar{\gamma} = 3.98$(6 dB), using the lower threshold, $\gamma_{T_3} = 11.48$, (10.59 dB), as γ, for $\mathcal{R}_4$, and assuming $\sigma_N^2 = 1$, the

resulting scaling factor in the receiver is given by

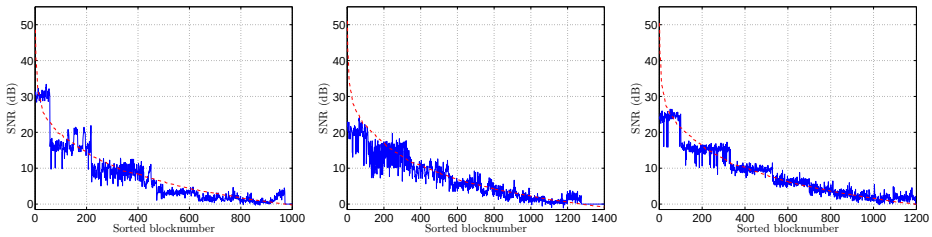
$$\begin{aligned} w_4 &= \frac{\sqrt{\alpha(k)}}{\alpha(k) + \frac{\sigma_N^2}{\sigma_{Sm}^2(\gamma)}} = \frac{\sqrt{\gamma/\bar{\gamma}}}{\gamma/\bar{\gamma} + 1/\sigma_{Sm}^2(\gamma)} \\ &= \frac{\sqrt{11.48/3.98}}{11.48/3.98 + 1/33.96} = 0.58, \end{aligned} \tag{5.3}$$

which means that the AWGN noise power, σ_N^2, will be reduced by at least a factor of $1/0.58^2 = 2.97$ in $\mathcal{R}_4$. The nature of the mappings will then give an improved performance compared to the CSNR level they are designed for, since they do not suffer from the leveling-off effect. The full analysis will have to be done when considering the effect the receiver-filter has on the information signal as well. Relating the scaling factor, w_m, to a SNR, is, however, difficult due to the nonlinearities of the mappings. Further analysis of the effect of the different receiver filters should be done in the future, as this has a big impact on performance. So why is the same performance gain not appearing for the $\mathcal{S}$ case? Since the transmitter scales the transmitted signal with the expected gain of the channel, the mismatch is smaller, and the noise-reduction will be less dramatic. For the case of $\mathcal{C}$, this gain is not visible at all, so from Figure 5.12(c) and 5.13(c) the performance given by Figure 3.9, is followed.

In Figure 5.13 it is also interesting to see that when a given channel state m is fully occupied with source-blocks, the following blocks might need to use a mapping with higher rate to compensate for the lower γ_C value.

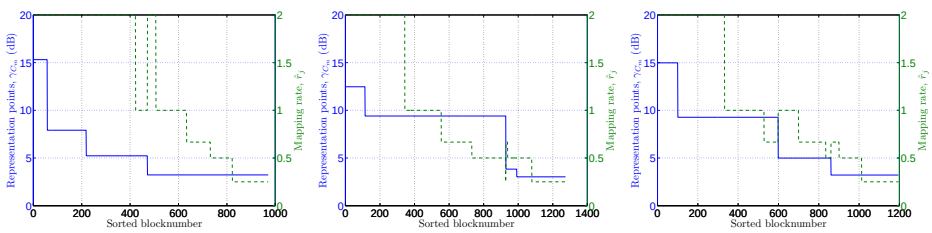
In Figure 5.11, it is shown how the CSNR regions and representation points found by the *complex* algorithm, result in a system performance that is outperformed by the system performance when using the *simple* algorithm. Because of this, and for being a simpler approach, only the *simple* algorithm will be used in the remainder of the thesis.

In addition to the differences in performance for transmission of the same image for different choice of CAS, it is seen in Figure 5.11, that the performance of the proposed system compared to the theoretical performance is different for different images. The performance for the "Bridge" image is closer to the theoretical performance, compared to the "Lena" image. Since the representation points and CSNR regions are chosen independently of the image for a given CAS, the difference in performance has to be due to differences in image statistics. In Figure 5.14, an example of the SNR values for the different source-blocks for the "Bridge" image is given. Comparing with the same parameters for the "Lena" image in Figure 5.12, shows that for the same mapping $\hat{r}_j$, and same representation



(a) $\mathcal{N}$, $\bar{\gamma}$= 5.78 dB (b) $\mathcal{S}$, $\bar{\gamma}$= 5.70 dB (c) $\mathcal{C}$, $\bar{\gamma}$= 5.73 dB

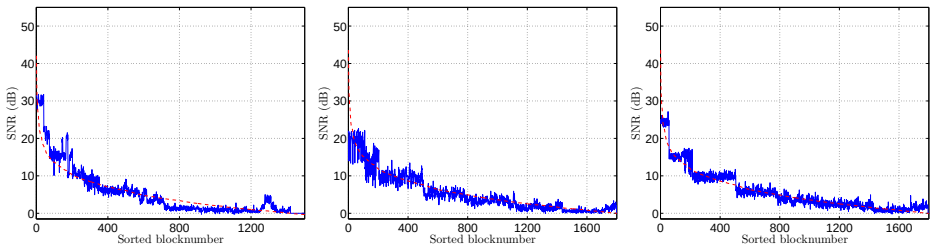
Figure 5.12: Example of SNR for each block (solid), block variance sorted descending. SNR for block with target distortion-level μ (dashed). $r_{\text{avg}} = 0.5$, $\mathfrak{NB}$, $M = 4$, *simple*, "Lena".



(a) $\mathcal{N}$, $\bar{\gamma}$= 5.78 dB (b) $\mathcal{S}$, $\bar{\gamma}$= 5.70 dB (c) $\mathcal{C}$, $\bar{\gamma}$= 5.73 dB

Figure 5.13: Preallocation of sorted source-blocks to channel region representation point $\{\gamma_{C_m}\}_{m=1}^{4}$ (solid, left axis), and mapping-rate $\hat{r}_j$ (dashed, right axis). $\mathfrak{NB}$, *simple*, "Lena".

value γ_{C_m}, the SNR value for a block is similar. The use of the mappings are not equal, though. In Figure 5.15(a) the mapping of rate $\hat{r}_5 = 2$ is preallocated to fewer blocks than in Figure 5.13(a). The reason for this lies in the source-block variances. In Figure 4.3, the dB values for the variance of the source-blocks for the different images are shown. It is seen that the variance of the source-blocks for the "Lena" image is smaller than for the "Bridge" image[4]. There is more activity in the "Bridge" image: it is not as lowpass as the "Lena" image. For this reason, the "Lena" image is easier to code. For the same parameters, transmission power and r_{avg}, "Lena" will have better quality as shown in Figure 4.2. Since the variance of the source-blocks for "Lena" drops faster than the "Bridge" image, many of the source-blocks can be coded with a mapping of low $\hat{r}_j$, which then means that the distortion-level μ can be set low to include enough rate. Low distortion level will then lead to high SNR for each block, a mapping with high rate, $\hat{r}_j$. So for being easier to code, a lower distortion-level can be set, resulting in the need for higher dimensional mappings for the blocks with highest variance. Since there are no mappings with rate higher than $\hat{r}_5 = 2$, the actual distortion will be larger than expected when preallocating.



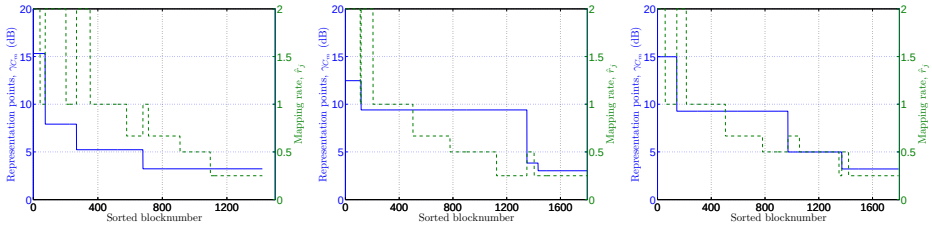
(a) $\mathcal{N}$, $\bar{\gamma}$= 6.14 dB (b) $\mathcal{S}$, $\bar{\gamma}$= 5.99 dB (c) $\mathcal{C}$, $\bar{\gamma}$= 5.99 dB

Figure 5.14: Example of SNR for each block (solid), block variance sorted descending. SNR for block with target distortion-level μ (dashed). $r_{\text{avg}} = 0.5$, $\mathfrak{NB}$, $M = 4$, *simple*, "Bridge".

5.1.2.2 Effect of $\mathfrak{B}$, mapping-rate mismatch compensation

As shown previously, the mismatch of mapping-rate can lead to a great loss of performance. For $\mathcal{S}$, there was a need for a mapping with rate larger than $\hat{r}_5 = 2$, which was the largest available. In Figure 5.16 the average PSNR values are plotted for the "Lena", "Goldhill" and "Bridge" images, with and without mapping-rate mismatch compensation for each block. When using $\mathcal{N}$, the average PSNR is lifted about one dB regardless

[4]Except the 9 blocks with largest variance where they are similar.



(a) $\mathcal{N}$, $\bar{\gamma}$= 6.14 dB (b) $\mathcal{S}$, $\bar{\gamma}$= 5.99 dB (c) $\mathcal{C}$, $\bar{\gamma}$= 5.99 dB

Figure 5.15: Preallocation of sorted source-blocks to channel region representation point $\{\gamma_{C_m}\}_{m=1}^{4}$ (solid, left axis), and mapping-rate $\hat{r}_j$ (dashed, right axis). $\mathfrak{NB}$, *simple*, "Bridge".

of $\bar{\gamma}$ value. For $\mathcal{S}$ it is seen in Figure 5.16(b), 5.16(e) and 5.16(h) how the lack of a high dimensional mapping for low $\bar{\gamma}$, has been compensated for by distributing the power differently for each block. As the need for such a mapping reduces when $\bar{\gamma}$ increases, the gain of using mapping-rate compensation is reduced. Since $\mathcal{S}$ compensates for the channel gain more accurately compared to $\mathcal{N}$, the performance is better.

As seen earlier in Figure 5.3 for the no-outage case, using $\mathfrak{B}$ will make the spread of the PSNR values narrower. This is also seen in Figure 5.17, where the estimated pdf is plotted for the "Lena" image with and without mapping-rate mismatch compensation for $\bar{\gamma} = 6$ when $M = 4$.

The effect of $\mathfrak{B}$ for the SNR of each block is shown in Figure 5.18 for the "Lena" image. It is seen that the distortion $\sigma^2_{D_n}$ for each block becomes more equal, resulting in an SNR for each block similar to the case when the distortion is μ for all blocks. The preallocation of the blocks for different power levels is shown in Figure 5.19.

5.1.2.3 Spread of parameters

As mentioned earlier, the channel will not be fully ergodic for the transmission of a single image due to a limited number of channel samples. The true transmission power $\bar{\sigma}^2_{S_t}$, true source symbol/channel symbol ratio r_{avg_t}, and true PSNR_t value, will then vary around an average value. In Figure 5.20, r_{avg_t} and PSNR_t are plotted as functions of $\bar{\gamma}_t$ for the "Goldhill" image for different CASs when $M = 1$. Figure 5.20(a) and 5.20(b) shows how r_{avg_t} are directly connected with $\bar{\gamma}_t$ for $\mathcal{N}$ and $\mathcal{S}$ for $M = 1$. This is due to the fact that the transmission power is either constantly on or off, and the switching depends on the outage. So when the channel is in outage less than expected, the power during transmission will be larger than expected, and r_{avg_t} will be smaller, as the time the channel is in outage is included in

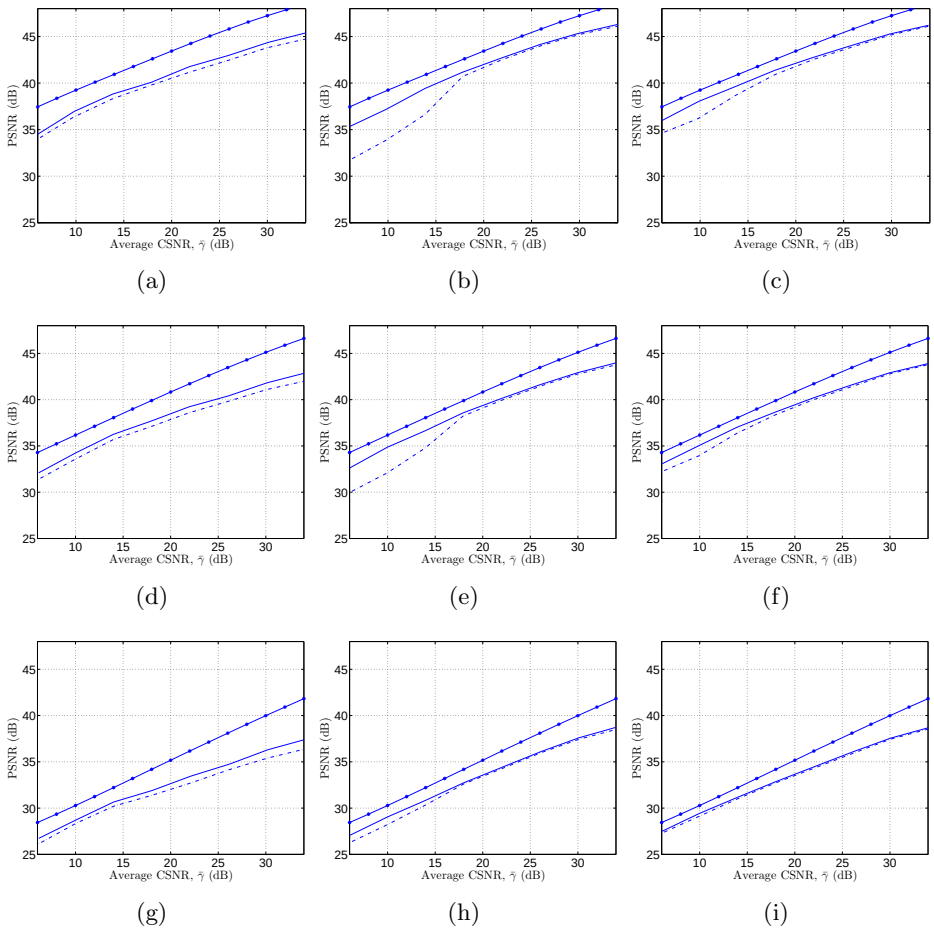
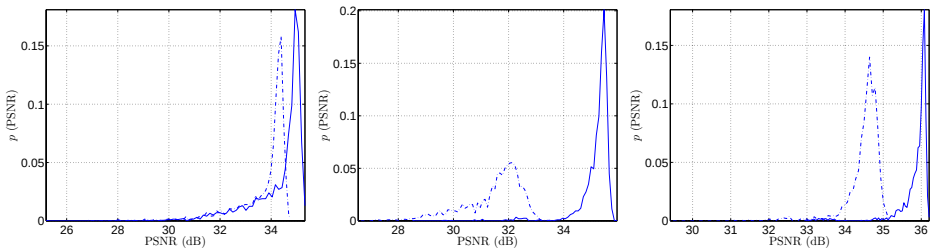
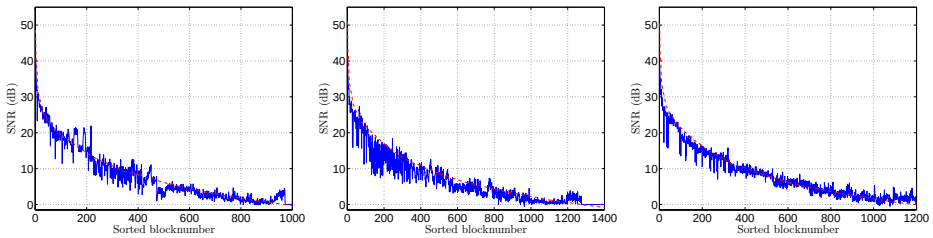


Figure 5.16: PSNR comparison for the received "Lena" (top row), "Goldhill" (middle row), and "Bridge" (bottom row) images. $r_{\text{avg}} = 0.5$, $M = 4$. $\mathcal{N}$ (left column), $\mathcal{S}$ (middle column) and $\mathcal{C}$ (right column). $\mathfrak{NB}$ (dash-dot), $\mathfrak{B}$ (solid), estimated optimum (solid*).



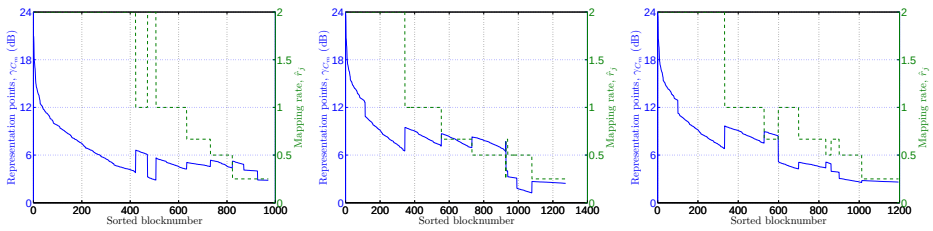
(a) $\mathcal{N}$ (b) $\mathcal{S}$ (c) $\mathcal{C}$

Figure 5.17: Estimated pdf for PSNR values for the "Lena" image, $\bar{\gamma}_e = 6$ dB, $r_{\text{avg}} = 0.5$, $M = 4$. $\mathfrak{B}$ (solid), $\mathfrak{NB}$ (dashed).



(a) $\mathcal{N}$ (b) $\mathcal{S}$ (c) $\mathcal{C}$

Figure 5.18: Example of SNR for each block (solid), block variance sorted descending. SNR for block with target distortion-level μ (dashed). $\bar{\gamma}_e = 6$ dB, $r_{\text{avg}} = 0.5$, $\mathfrak{B}$, $M = 4$, "Lena".



(a) $\mathcal{N}$ (b) $\mathcal{S}$ (c) $\mathcal{C}$

Figure 5.19: Preallocation of sorted source-blocks to channel region representation point $\{\gamma_{C_n}\}_{n=1}^{N}$ (solid, left axis), and mapping-rate $\hat{r}_j$ (dashed, right axis). $\mathfrak{B}$, "Lena", $\bar{\gamma}_e = 6$ dB, $M = 4$.

the calculation of r_{avg}. As $\bar{\gamma}_e$ increases, the probability of being in outage decreases, and the variation of r_{avg_t} and $\bar{\gamma}_t$ will decrease. When using $\mathcal{C}$, the situation is not so simple. For low $\bar{\gamma}$, r_{avg_t} still vary around the target value, but r_{avg_t} and $\bar{\gamma}_t$ are not connected one-to-one anymore. Depending on the state of the channel through the transmission, the transmission power will vary as the transmitter will invert the channel at all times. As the probability of being in outage decreases, the variation of r_{avg_t} goes to zero, but the variation of the transmission power will increase. The reason for this, is that the transmission power will change over a larger region as $\bar{\gamma}_e$ increases. As the number of symbols will increase, this variation will also converge to the target value. These variations for $\mathcal{C}$ can be seen in Figure 5.20(c). In Figure 5.20(f), it can be seen that the variation of $\bar{\gamma}_t$ decreases as r_{avg_t} increases. Comparing Figure 5.20(f) with Figure 5.20(d) and Figure 5.20(e), the effect of inverting the channel can be seen. The spreading of the PSNR values for the case of $\mathcal{N}$ and $\mathcal{S}$ is larger than for $\mathcal{C}$. For $\mathcal{N}$ and $\mathcal{S}$, PSNR_t is dependent on the channel during transmission, even the state of the channel when the transmission starts, as that is when the most important image information is transmitted. This is the reason why r_{avg_t} and PSNR_t is independent for the case of $M = 1$, as seen in figure 5.20(d), and 5.20(e).

In Figure 5.21 the variations of r_{avg_t}, PSNR_t and $\bar{\gamma}_t$ for the case when $M = 4$ are plotted. For $\mathcal{N}$ when $M = 1$, the variations of r_{avg_t} is purely dependent of the outage, but when multiple transmission power levels are used, the state of the channel during transmission will also influence both r_{avg_t} and $\bar{\gamma}_t$ when $M = 4$, as seen in Figure 5.21(a). The combination of outage probability and transmission power will give the same trends as in Figure 5.20(a), but where the spread over $\bar{\gamma}_t$ is larger. The same is valid for Figure 5.21(b), the spread of r_{avg_t} is equal to the $M = 1$ case, but the spread of $\bar{\gamma}_t$ is less compared to Figure 5.21(a). The reason for this is that for $\mathcal{S}$, the range of the transmission power is smaller than for $\mathcal{N}$. For $\mathcal{S}$, the mapped symbols are normalized by the region centroid before transmission, which leads to a down-scaling when the channel is good, and an up-scaling when the channel is poor, resulting in a tighter distribution of the power levels. Figure 5.21(c) shows that for $\mathcal{C}$, the spread of $\bar{\gamma}_t$ has become tighter compared to when $M = 1$. This is due to the fact that the CSNR range of each region is smaller when $M = 4$. The power variation is smaller within a region. As for the PSNR, Figure 5.21(d) shows that the spread of the PSNR_t values is reduced since the most important blocks are being transmitted on a good channel state. From the figure it is also seen that the spread of PSNR_t is smaller when $\bar{\gamma}_t$ is larger than the targeted $\bar{\gamma}_e$ as well. The reason for this is that when the channel is in a good channel state more often than expected, blocks allocated for a poor channel state

are reallocated to a better channel state and sent with a higher power. The needed mapping-rate is then reduced, so the overall r_{avg} is smaller, as seen in Figure 5.20(a). If the channel is less often in a good channel state than expected, some blocks pre-allocated for a good channel state, needs to be transmitted at a worse channel state, and a mapping of higher rate $\hat{r}_j$ is needed. This can be seen in Figure 5.22, where the probabilities of the different channel states seen by the transmitter are plotted. The best state with the best channel condition is given in Figure 5.22(d). It can be seen that the cases for which the PSNR value is low is when the best state shows a lower probability than the theoretical. $\mathcal{R}_3$ shows similar trends in Figure 5.22(c). When using $\mathcal{S}$, the spread of PSNR_t is shown in Figure 5.21(e). The spread of PSNR_t is not similar to the $\mathcal{N}$, as was the case for $M = 1$. This is, however, mostly due to the low representation value for the best channel region as explained earlier, especially for low $\bar{\gamma}$ values. For higher $\bar{\gamma}$ values, the loss from low representation values is gone, and the spread of PSNR_t and $\bar{\gamma}_t$ is low. The low difference in transmission power for the different channel states, results in low spread of $\bar{\gamma}_t$. The pre-scaling does in addition give a low spread of the PSNR values. For $\mathcal{C}$ the spread of PSNR_t is low, but results in a large spread for $\bar{\gamma}_t$ for high $\bar{\gamma}$, as seen in Figure 5.21(f). For low $\bar{\gamma}$, the spread of $\bar{\gamma}_t$ is due to the outage.

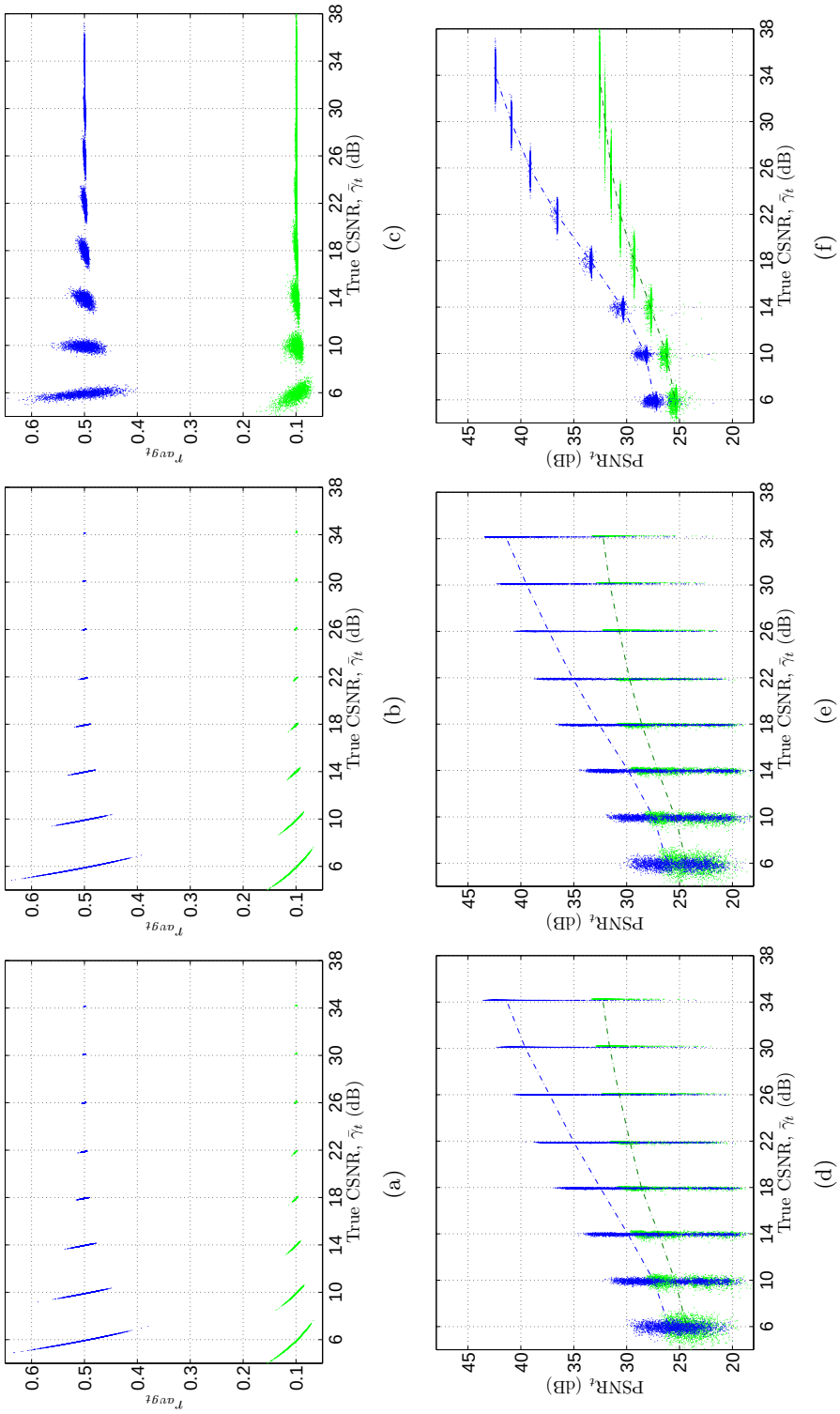


Figure 5.20: Scatterplot of r_{avg_t} (top row), and PSNR_t (lower row) for the "Goldhill" image. $r_{\text{avg}} = \{0.5(\text{blue}), 0.1(\text{green})\}$. $\mathcal{N}$ (left column), $\mathcal{S}$ (middle column) and $\mathcal{C}$ (right column). $\mathfrak{MB}$, $M = 1$. Plot for the average PSNR and $\bar{\gamma}_t$ is included in the lower row (dash-dotted).

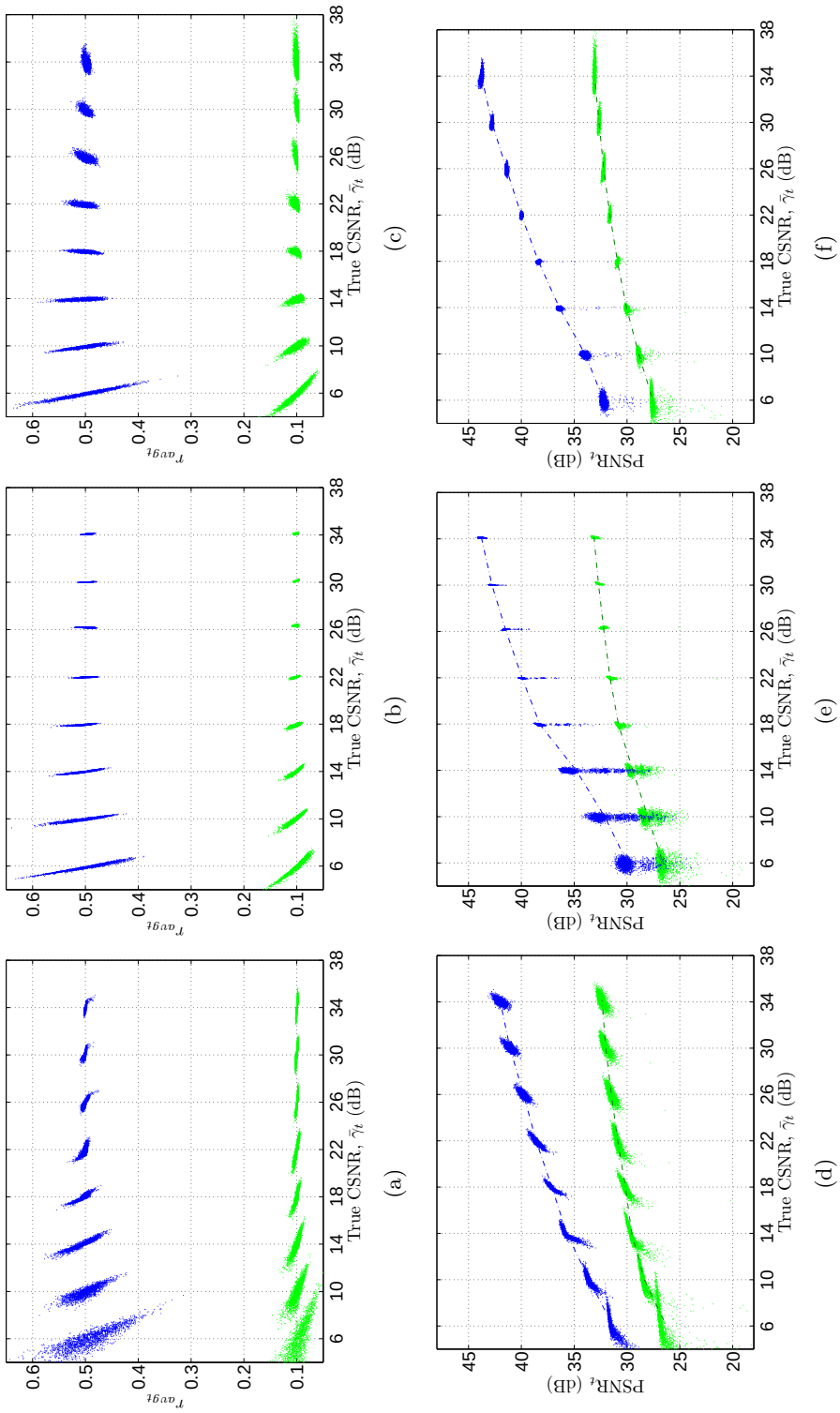
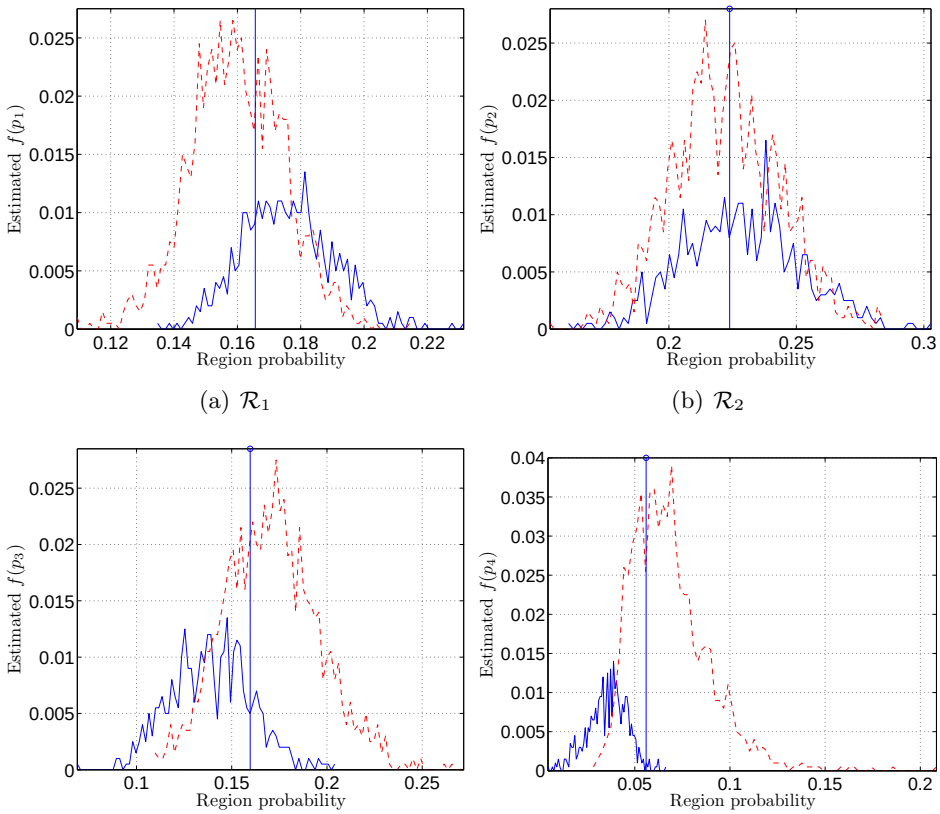


Figure 5.21: Scatterplot of r_{avg_t} (top row), and PSNR$_t$ (lower row) for the "Goldhill" image, $r_{\text{avg}} = \{0.5(\text{blue}), 0.1(\text{green})\}$. $\mathcal{N}$ (left column), $\mathcal{S}$ (middle column) and $\mathcal{C}$ (right column). $\mathfrak{MB}$, estimated optimum (solid*), $M = 4$.



(a) $\mathcal{R}_1$ (b) $\mathcal{R}_2$

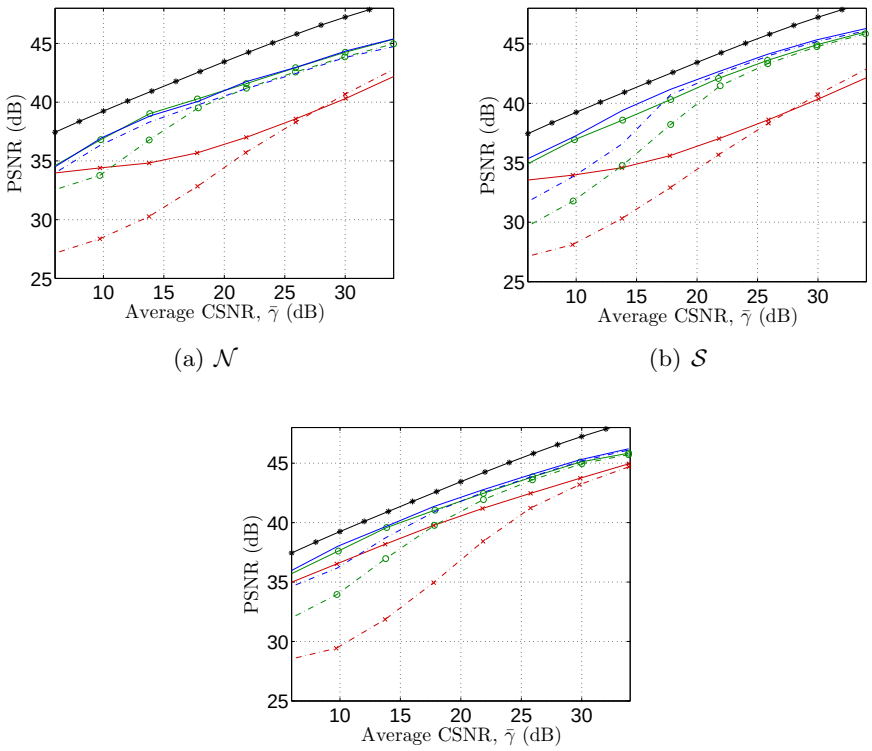
(c) $\mathcal{R}_3$ (d) $\mathcal{R}_4$

Figure 5.22: Estimated pdfs of the probabilities an image transmission for each channel region $\mathcal{R}_m$. Instances where the PSNR is greater than average (dashed), lower than average (solid), normalized with the total number of instances. Theoretical probability for region is given by solid line. $\bar{\gamma} = 6.06$ dB, "Goldhill", $\mathfrak{NB}$, $\mathcal{N}$.

5.1.2.4 Performance gain by increasing the number of regions

In Figure 5.23, the performance of the system is given for $\mathfrak{NB}$ and $\mathfrak{B}$, for $M = \{1, 2, 4\}$. For $\mathfrak{NB}$, the performance is steadily increasing as M increases. In the case of $\mathcal{N}$, it is seen that for $M = 2$ and $M = 4$, the performance curves comes together. This shows that adapting in one sense without thinking about other sources of inaccuracies does not help. By adding a bit more sophisticated channel adaptation for the channel symbols as done in the case of $\mathcal{S}$, there will be an extra gain by increasing M. For $M = 1$ there is a big gain by using $\mathfrak{B}$. In Figure 5.9 it was shown that by adjusting the level of outage, the difference between $\mathfrak{B}$ and $\mathfrak{NB}$ can be

reduced by a large factor. Based on that observation, one can say that it is likely that the gap between $\mathfrak{B}$ and $\mathfrak{NB}$ can be reduced for the other values of M as well by adjusting the outage-level, regions and representation points. Thus giving only a small gain through the extra complexity added by using $\mathfrak{B}$.



(a) $\mathcal{N}$

(b) $\mathcal{S}$

(c) $\mathcal{C}$

Figure 5.23: Average PSNR for the "Lena" image, $r_{\text{avg}} = 0.5$ $\mathfrak{NB}$ (dashed-dotted), $\mathfrak{B}$ (solid), $M = 1$ (x-red), $M = 2$ (o-green) and $M = 4$ (blue). Estimated upper limit (solid*-black).

5.2 Impact of Doppler shift

The impact of the Doppler shift f_m is that the channel will change faster. Since it is assumed that the transmitter instantly knows when the channel moves to another state among the $M+1$ different possible, the Doppler shift will not have a big impact on the whole system. In Figure 5.24 the PSNR performance of the system is plotted for the case when $f_m = [100, 300, 500]$

Hertz (Hz) for the case of $\mathcal{N}$. For low $\bar{\gamma}$ values, the difference between $f_m = 100$ Hz and $f_m = 500$ Hz, is at most 0.2 dB for $r_{\text{avg}} = 0.5$. This might seem strange, but for the system proposed in this thesis, a higher f_m means that the channel stays a shorter time in one CSNR region than for lower f_m values. This leads to a channel that is more ergodic, and the spread of PSNR_t becomes smaller than for $f_m = 100$ Hz. A bigger gain can be seen for the case when $r_{\text{avg}} = 0.1$, since the number of channel symbols is smaller, a more ergodic channel will tighten the spread of PSNR_t more. A gain of most 0.4 dB is achieved at low $\bar{\gamma}$ values.

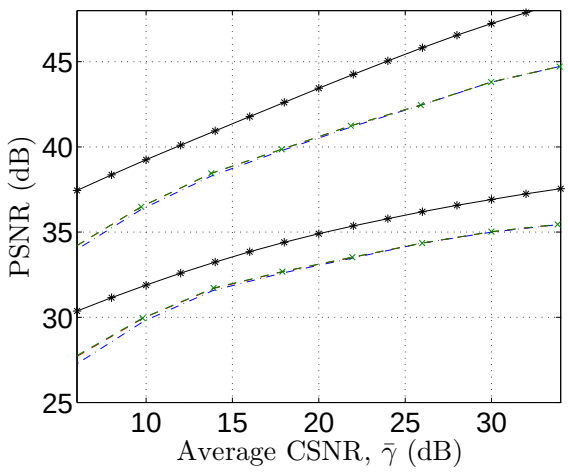


Figure 5.24: Performance for different Doppler-shift f_m, $f_m = 100$ (dash-dotted-blue), $f_m = 300$ (dashed-red) and $f_m = 500$ (dashed-x-green). The results for $f_m = 300$ and $f_m = 500$ are very similar. For the case when $M = 4$, "Lena", $\mathcal{N}$, $\mathfrak{MB}$. Estimated upper limit (solid*-black).

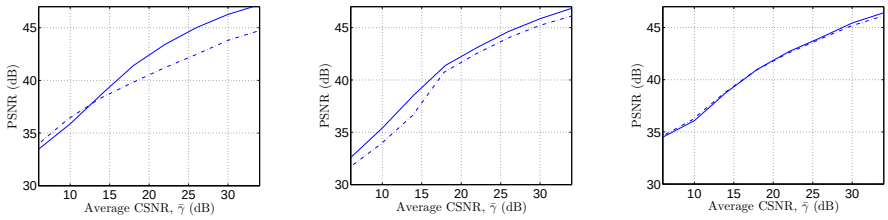
5.3 Comparison to reference systems

5.3.1 Theoretical system

5.3.1.1 Estimating system performance

In Section 4.2 the estimation of the performance for nonlinear mappings was discussed. In the estimation it is assumed that the mappings always operate at the CSNR level they are designed for, which for $\mathcal{N}$ and $\mathcal{S}$ is not the case. In Figure 5.25 the estimated performance for the "Lena" image is compared to the average PSNR values of the implemented system. It is

seen in Figure 5.25(c) that for $\mathcal{C}$ the estimation comes close, but as seen in Figure 5.25(a) and 5.25(b), $\mathcal{N}$ and $\mathcal{S}$ will give an inaccurate approximation. Figure 5.25(a) confirms what was discussed earlier in Section 5.1.2.1, that for low $\bar{\gamma}$ values, the mappings will operate on γ_C levels giving lower distortion on average compared to only operating on the CSNR level they are designed for. The effect of pre-scaling the transmitted signal with the expected gain in each region for $\mathcal{S}$, can be seen by comparing Figures 5.25(a) and 5.25(b) for high $\bar{\gamma}$ values. The mappings will operate in a smaller region around the designed γ_{C_m} level in the $\mathcal{S}$ case, and hence come closer to the estimated performance. For the above mentioned reasons, only $\mathcal{C}$ will be considered for theoretical estimation.



(a) $\mathcal{N}$ (b) $\mathcal{S}$ (c) $\mathcal{C}$

Figure 5.25: Comparison of performance for proposed coder (dash-dotted) and theoretical system estimating distortion for each block (solid), using distortion from implemented mappings for theoretical system. $\mathfrak{NB}$, "Lena", $M = 4$, $r_{\text{avg}} = 0.5$.

5.3.1.2 Loss due to imperfect mappings

None of the implemented mappings, except the mapping of rate $\hat{r}_4 = 1$, will perform according to the optimal limit. To see the loss due to this, the simulated system-performance when assuming perfect mappings, and the system-performance when using the implemented mappings, are plotted in Figure 5.26 for the "Lena" image. It is seen that the loss due to imperfect mappings is in the range $0.5-1.5$ dB when using $\mathfrak{B}$. For the $\mathfrak{NB}$ case, it can be seen how the performance of the implemented $\hat{r}_5 = 2$ mapping makes the overall performance poor. By using perfect mappings the performance is increased by almost 2 dB for low $\bar{\gamma}$ values. It can further be seen that the performance for the implemented system drops off at high $\bar{\gamma}$ values compared to the theoretical system. The reason for this can be found in Figure 3.7. The performance of these mappings drops off gradually as the $\bar{\gamma}$ increases. Especially the mapping with $\hat{r}_3 = 2/3$ contributes as it saturates around $\bar{\gamma} = 30$ dB.

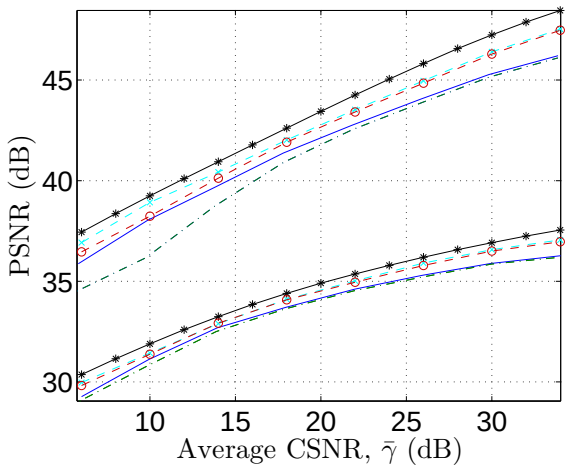


Figure 5.26: Performance in PSNR for proposed system with $\mathfrak{B}$ (solid-blue), $\mathfrak{NB}$ (dash-dotted-green) compared to theoretical system using perfect mappings with $\mathfrak{B}$ (dashed-x-cyan), $\mathfrak{NB}$ (dashed-o-red), $M = 4$, $\mathcal{C}$, "Lena", $r_{\text{avg}} = \{0.1, 0.5\}$ from below. The theoretical optimum is included for reference (solid*-black).

5.3.1.3 Extra mappings

Since it was shown that the imperfectness of the mappings available contributes to a significant loss in some cases, it could be interesting to see how much the performance could increase by adding another mapping. In Figure 5.27 the performance for the proposed system is shown when a mapping of rate $\hat{r} = 3$ is added. This is shown both for the case when the extra mapping performs according to the theoretical optimum, and when the extra mapping has a CSNR loss of 3.5 dB from the optimum. 3.5 dB is assumed to be a reasonable estimate for the performance. For high γ values, a 3.5 dB shift leads to a performance approximately 10 dB SNR below the optimum. Figure 5.27 shows that adding a mapping of rate $\hat{r} = 3$, will lead to a performance for the system when using $\mathfrak{NB}$ that is comparable to the case when using $\mathfrak{B}$ for the implemented system. A loss of 3.5 dB CSNR does not contribute significantly to the performance-loss of the system, in the presented case, a loss of max 0.3 dB PSNR is found.

In Figure 5.28 a mapping of rate $\hat{r} = 3/2$ is added. The systems performance is given both for the case when the performance of the mapping with rate $\hat{r} = 3/2$ is optimal, and when there is a loss of 3.5 dB CSNR. It is seen in Figure 5.28 that adding a mapping of rate $\hat{r} = 3/2$ will not help at

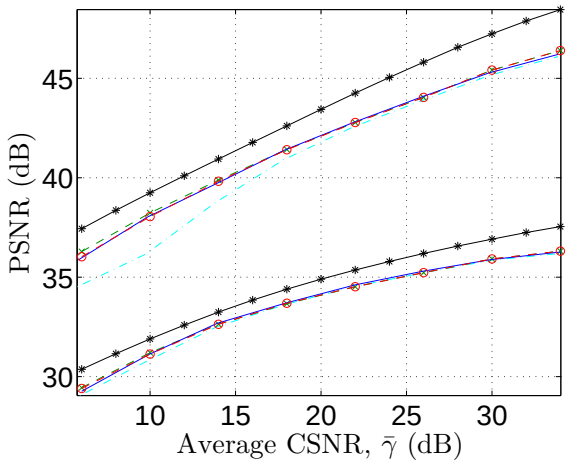


Figure 5.27: Estimated performance when introducing extra mapping of rate $\hat{r} = 3$ with optimal performance (dashed-x-green), and with a 3.5 dB CSNR loss (dashed-o-red) for $\mathfrak{NB}$. The performance of the implemented system is given for $\mathfrak{NB}$ (dash-dotted-cyan) and $\mathfrak{B}$ (solid-blue). "Lena", $M = 4$, $\mathcal{C}$, $r_{\text{avg}} = \{0.1, 0.5\}$ from below. The theoretical optimum is included for reference (solid*-black).

low $\bar{\gamma}$ values, but will help at most 0.5 dB PSNR at high $\bar{\gamma}$ values. It might seem strange that systems performance is actually poorer with the extra mapping for low $\bar{\gamma}$ values. This is due to the way the different mappings are being chosen. In the implemented coder, the mapping with distortion closest to the target distortion-level μ is chosen. In a real system, the rate of the mapping will have to be included in the selection.

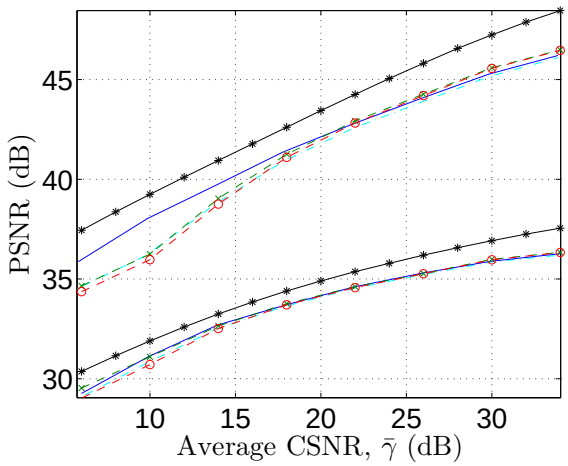


Figure 5.28: Estimated performance when introducing extra mapping of rate $\hat{r} = 3/2$ with optimal performance (dashed-x-green), and with a 3.5 dB CSNR loss (dashed-o-red) for $\mathfrak{NB}$. The performance of the implemented system is given for $\mathfrak{NB}$ (dash-dotted-cyan) and $\mathfrak{B}$ (solid-blue). "Lena", $M = 4$, $\mathcal{C}$, $r_{\text{avg}} = \{0.1, 0.5\}$ from below. The theoretical optimum is included for reference (solid*-black).

The increased performance of the proposed system by adding an extra mapping does not, as shown, in the case of a fading channel with several regions, give a huge gain. As long as there is a representation point γ_{C_m} which has large value, the need for a high dimensional mapping is reduced. In the case of a AWGN channel, the potential gain is much higher, as all mappings are designed for the average CSNR. An example of this is given in Figure 5.29 for an extra mapping of $\hat{r} = 3$. It is seen that the impact of $\mathfrak{B}$ is larger compared to the fading case. The impact of an extra mapping is larger as well. By adding an extra mapping, the blocks with largest variance can get a much lower distortion, thus increasing the total image quality by a large amount.

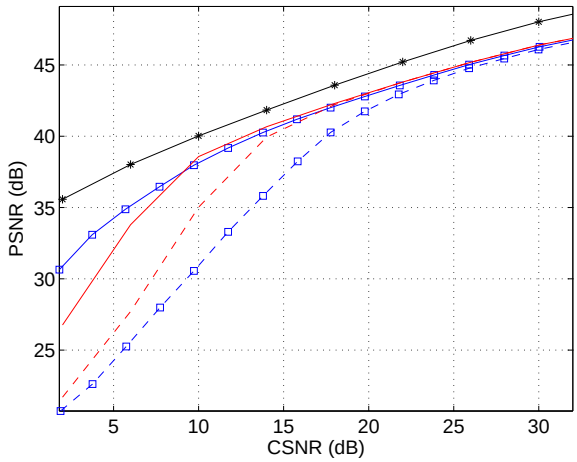


Figure 5.29: Performance of proposed system with $\mathfrak{B}$ (solid-squares-blue), and $\mathfrak{NB}$ (dashed-squares-blue). Simulated performance of system with mapping of rate $\hat{r} = 3$ with ideal performance (solid-red), and a 3.5 dB CSNR loss (dashed-red). $r_{\text{avg}} = 0.5$, "Lena". The theoretical optimum for an AWGN channel is included for reference (solid*-black).

5.3.2 Comparison to practical schemes

In Figure 5.30 the performance of the proposed system is compared to two reference systems described in Section 3.3. Both uses JPEG2000 as source coder, but uses different channel transmission schemes. The first denoted ACM, uses a set transmission rates achieving the capacity for an AWGN channel with CSNR $\{\gamma_{C_m}\}_{m=1}^{M}$. The other scheme denoted TuCM, varies turbo-code, constellation and power to combat fading. The ACM scheme is a theoretical model, and comes very close to the channel capacity. The TuCM scheme is a more practical scheme, and is assumed to follow the channel capacity with a shift of 3 dB in $\bar{\gamma}$. This shift can be seen in Figure 5.30 as the performance for the ACM and TuCM scheme is approximately parallel with a $\bar{\gamma}$ separation of 3 dB.

For the "Lena" image, the proposed system has poorer performance than the ACM system, but better performance than the TuCM system, as seen in Figure 5.30(a). For the "Goldhill" and "Bridge" images, in Figures 5.30(b) and 5.30(c), the proposed coder is comparable with the ACM system, and performs better than the TuCM system for $\bar{\gamma}$ values shown.

Regardless of the specific image coded, there are a couple of charac-

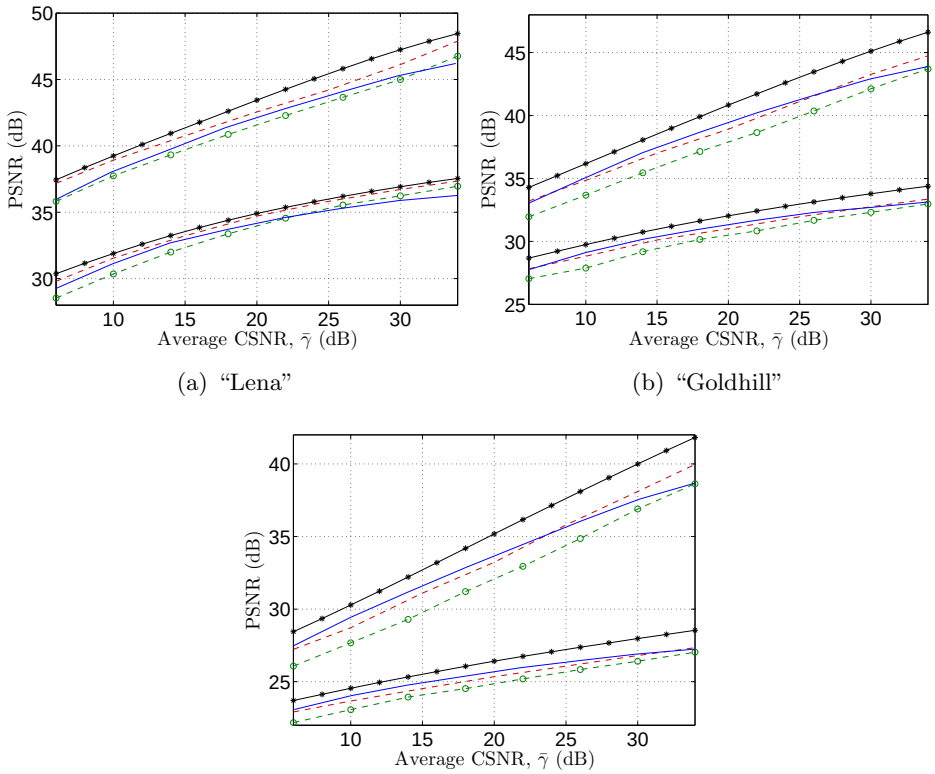
teristics that are typical for the proposed coder. For low $\bar{\gamma}$ values, the performance drops off as there is a need for higher dimensional mappings than the ones available. For high $\bar{\gamma}$ values, the proposed system performance drops of since the performance of the mappings used drops off with respect to the optimal performance.

For visual comparison, examples of the images from the different schemes are given in Figures 5.31, 5.32 and 5.33. Since both the ACM and TuCM scheme use JPEG2000, the image artifacts seen for these cases are similar. The difference for these two schemes is the bit-rate the image is coded for. Comparing Figure 5.31(a) with Figure 5.31(b) it can be seen that the skin of "Lena" is smoother in the case of the JPEG2000 images, which gives a more pleasant image to look at, but this is the same as In the case of the "'Goldhill' image in Figure 5.32, the same tendency can be seen in the roof tiles. For the JPEG2000 coded images, the details are more smeared out compared to the image coded with the proposed system. Some artifacts appear due to high compression, these can be clearly seen in Figure 5.33(c) for the "Bridge" image.

When comparing the images, it should be taken into consideration that for the image quality for the proposed system is an average. So the received image quality would differ depending on the channel. If the channel is good, the image quality is good, and if the channel is poor, the quality will be poorer. The system performance tracks the channel changes, which indicates a certain robustness. So if the CSI at the transmitter is not perfect, e.g. it might be delayed and give an inaccurate estimate, the receiver would still receive an image containing information. In the case of the reference systems, a breakdown will occur in the system if the channel is poorer than expected.

5.4 Discussion

The proposed coder has shown good performance compared to the reference systems. This is quite interesting when it is taken into consideration that JPEG2000 is an advanced standard, developed by hundreds of researchers. For the channel transmission part of the reference systems, one theoretical scheme was used, which would in practice require a complex channel coder to work. The other channel transmission scheme does not show the same performance as the ACM scheme. The use of turbo codes together with adaptive modulation makes it more practical than the ACM scheme. It is assumed that the TuCM scheme comes within 3 dB $\bar{\gamma}$ of the channel capacity. This is however at $\bar{\gamma} = 6$ dB, and at $\bar{\gamma} = 14$ dB the system is almost 4 dB away from the capacity [Vishwanath and Goldsmith, 2003].



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

(c) "Bridge"

Figure 5.30: Comparison of proposed system (solid-blue), system using ACM (dashed-red) and TuCM (dashed-o-green). Estimated upper limit (solid*-black). Proposed system uses $\mathfrak{B}$, $M = 4$, $\mathcal{C}$.

(a) PSNR = 31.1 dB

(b) PSNR = 31.5 dB, ACM

(c) PSNR = 30.4 dB, TuCM

Figure 5.31: Image comparison between proposed coder, ACM and TuCM. Extract from "Lena" for $r_{\text{avg}} = 0.1$, $\bar{\gamma}_e = 10$ dB. Target rate in bits/pixels is 0.1468 for ACM and 0.1131 for TuCM. PSNR value calculated for whole image.

(a) PSNR = 29.1 dB

(b) PSNR = 28.8 dB, ACM

(c) PSNR = 27.9 dB, TuCM

Figure 5.32: Image comparison between proposed coder, ACM and TuCM. Extract from "Goldhill" for $r_{\text{avg}} = 0.1$, $\bar{\gamma}_e = 10$ dB. Target rate in bits/pixels is 0.1468 for ACM and 0.1131 for TuCM. PSNR value calculated for whole image.

(a) PSNR = 24.0 dB

(b) PSNR = 23.7 dB, ACM

(c) PSNR = 23.1 dB, TuCM

Figure 5.33: Image comparison between proposed coder, ACM and TuCM. Extract from "Bridge" for $r_{\text{avg}} = 0.1$, $\bar{\gamma}_e = 10$ dB. Target rate in bits/pixels is 0.1468 for ACM and 0.1131 for TuCM. PSNR value calculated for whole image.

In the TuCM scheme a set of 4 regions, included outage, is used leading to $M = 3$.

The proposed system uses $M = 2$ and $M = 4$, so an accurate comparison is not possible. Considering the difference in performance for the proposed system between $M = 2$ and $M = 4$ for $\mathcal{C}$ and $\mathfrak{B}$, shown in Figure 5.23(c) for the "Lena" image, it is seen that the loss is about 0.5 dB CSNR, which is valid for the other images as well. Another issue for the TuCM scheme is the fact that the block lengths used by the turbo coder are constrained by the time the channel spends in each region. Since it is not possible to stop transmission in the middle of a codeword, a higher Doppler shift f_m will lead to shorter codewords, hence poorer protection. The proposed system will on the other hand have a maximum of 2 consecutive channel symbols that are dependent on each other. Few consecutive dependent channel symbols means that the transmitter can adapt fast to a changing channel. A faster change between channel regions results in the need for more CSI, but no other change.

The delay of the turbo-coded system might also be a constraining factor. Long block lengths will lead to a large delay, but by reducing the block lengths, protection is lost. For the proposed system there is also a delay, but here the largest delay comes from the analysis of the image and transmission of the side information. Since the decoder needs to know the side-information to be able to decode the main information, transmission can not start until the side information is correctly received.

It should be noted that JPEG2000 was not only designed for image quality. There are also other features in the JPEG2000 standard, progressive transmission, *region of interest* (ROI) and the possibility of lossless coding, to mention a few. Through the use of source-blocks, the proposed coder also has the possibility of progressive decoding. All the source blocks do not have to be received before the received image can be shown. The proposed system was, however, not designed with this in mind, it is more a side-effect of the system architecture.

The choice of optimization algorithm used to find the representation points and thresholds did not seem to be very important. The choice of regions and representation points did, however, show to be very important. Especially the representation point for the best channel region. From the results found, it looks like there could be a gain by including the source statistics when finding the regions and representation points. The results showed that correct representation of the most important source-blocks is crucial for the total image quality. For low $\bar{\gamma}$, this is important since high dimensional mappings are needed, but not available. An optimization problem for finding the best representation points and thresholds should take this mapping mismatch into account. As a simpler approach, one can

think that it is possible to assume a general image-model. By designing the representation points from such a model, the computational load on the receiver can be reduced. To reduce the amount of side-information, a general image-model can also be used for the source-blocks. The different source-block variances can be quantized to a smaller set. This could be a topic for further research.

Compensating for the mapping-rate mismatch showed to give a large improvement, about 5 dB PSNR for low $\bar{\gamma}$ values for $M = 4$. On the other side, the complexity is increased, both with respect to design, and transmission. Using a different setting for the mappings for each block, requires extra computational power at the transmitter. It is also beneficial to keep the dynamic range of the power amplifier at the transmitter as small as possible, to avoid saturating the power amplifiers. Both $\mathcal{N}$ and $\mathcal{S}$ when using $\mathfrak{NB}$ have a limited range over which an amplifier needs to work. $\mathcal{S}$ even moves the possible transmission powers closer together during pre-scaling of the symbols. Such pre-scaling showed to be beneficial for high $\bar{\gamma}$ values, but at low $\bar{\gamma}$ values the performance gain given by the noise-reduction in the filter in the receiver for the $\mathcal{N}$ case, seemed to give greater gain. To find the dynamics behind this is a strong candidate for future research. Using the robustness of the mappings can be used to get good performance.

The difference in performance from $\mathcal{C}$ and $\mathcal{S}$ can not be said to be dramatic. As $\mathcal{S}$ is a quantization of $\mathcal{C}$, it is possible to picture a system between $\mathcal{S}$ and $\mathcal{C}$. By allowing more CSI, the pre-scaling could be done for several levels within each channel region. Considering that the performance for $M = 2$ and $M = 4$ when using $\mathfrak{B}$ is quite similar, the amount of CSI needed is already very low. By using the robustness of the mappings, the amount of CSI can be low, and still have a good performing system. For a low Doppler shift, the channel variations will be slow and thus reducing the need for CSI. The amount of CSI needed is another very interesting topic for further research. A traditional way of estimating CSI is through the use of pilots, see e.g. [Cavers, 1991]. Special pilot symbols are multiplexed into the data-stream, which the receiver uses to estimate the channel. Few information channel symbols between each pilot will result in a good estimate, but the pilots do not contain any information, so it is desirable to keep the number of pilots as low as possible. By using the robustness of the proposed system, the density of pilots could most probably be kept smaller than a traditional ACM system. For a broadcasting scenario, one might think of a scheme where, since the transmitter uses constant power at all times, the receiver uses a sliding window to estimate the channel from the power of the received symbols. This information may then be used to compensate for the channel fades.

Setting the outage level in an optimal way should also, from the indication of the results for $M = 1$, give a large gain. The results for $M > 1$ have shown that when using a limited set of mappings, it is crucial for the image quality that the most important blocks are coded with correct mapping-rate. For low values of M, this can be compensated through the use of outage.

Traditional systems use discrete amplitude channel symbols. In doing so, the receiver can use an eye-diagram to find the optimal timing of sampling the received signal. For a system using continuous amplitude channel symbols, as the proposed system, such synchronization techniques are not possible. To synchronize in the case of continuous amplitude systems, one might think of special synchronization symbols sent in a frequent manner. For a fading channel, such symbols must however also compensate for the fading. This is an open problem, and needs to be solved to make a practical system.

5.4.1 Extending the system to other scenarios

In the thesis, a framework for transmitting an image over a fading channel has been presented. The results have been shown for a Rayleigh flat fading channel, but can be extended to another flat fading distributions without significant changes. In the following, a brief discussion not necessarily limited to fading channel is given.

5.4.1.1 Received image quality

For the results shown in Chapter 5, a time constraint[5], r_{avg}, has been put on the transmission. It might seem strange to use a time constraint on still image transmission as for most image transmission scenarios, a constraint on the image quality is used. The reason why a time constraint is used in this thesis, is to demonstrate the robustness of the system. The full advantages of robustness in such a system would be apparent if the system is extended to a video scenario. When streaming a video sequence to a mobile device, keeping the transmission delay below a certain threshold, yet at the same time getting graceful degradation and improvement, is essential. For two way live communication, the delay would be even more critical.

The amount of modification needed in the proposed system for a scenario where a target image quality is wanted, is relatively small. Through the distortion level μ, the transmitter can set the target distortion in the

[5] A constraint for the number of channel symbols can be translated into a time constraint through the duration of the channel symbols.

received image. Focusing more on the image quality than the transmission delay, would mean that there should be put a larger emphasis on making sure that each block is mapped with a mapping with a correct rate $\hat{r}_j$. This might include not transmitting the blocks preallocated for the best channel state on poorer channel states even if this means not transmitting for a period of time. If a mapping with sufficient rate is not available, it has been shown that the impact on image quality is substantial. This could, however, be compensated for by using more power, either on a block basis, or by shifting the representation point for the best channel state high enough. If a certain PSNR value is targeted, it might even be a possibility to add the theoretical distortion for each block and test the image PSNR by decoding a test image in the transmitter. By doing so the right distortion-level can be found for a given image and PSNR value. This would, however, increase the computational load on the transmitter.

5.4.1.2 OFDM channel

The system in this thesis was demonstrated for a flat fading channel. Another possible scenario is when using *orthogonal frequency division multiplex* (OFDM). Assume that the OFDM sub-channels are time invariant, and that each sub-channel can be assumed to be an AWGN channel. In the combination of image transmission using nonlinear source-channel mappings with OFDM transmission, the different sub-channels might be seen as different channel states used in this thesis. The different sub-channels would then have different probabilities, but it would be beneficial to utilize each sub-channel as much as possible to minimize overall transmission time. A possible strategy for transmission might be to start by distributing power for the different sub-channels according to the water filling principle to maximize channel transmission rate [Cover and Thomas, 1991]. The source blocks can then be allocated to the sub-channels according to the same principle as used in this thesis. Source blocks with the largest variance is allocated to the sub-channel with best CSNR to reduce the need for mapping with higher dimension than the ones implemented. Power can be reallocated after blocks have been allocated to sub-channels to minimize distortion. Special care might, in the sense of power, also be taken of the source blocks with largest variance in case a large enough mapping-rate is not available.

In the case that each of the OFDM sub-channels are time invariant, a different strategy is needed. Since the sub-channels are time invariant, preallocation needs to be done. For multiple sub-channels the preallocation has to be done across the available channels.

Chapter 6

Conclusion and Further research

In this thesis a framework for image transmission over flat fading channels using nonlinear dimensional changing mappings has been presented.

The nonlinear mappings offer the possibility of different level of protection that is matched to the channel condition and source importance. Through the inherent robustness of these mappings, the overall system becomes robust. For traditional channel codes, the channel gain is inverted to achieve a fixed CSNR level at the receiver. Using the robustness of the system, new schemes for adapting to the channel variations has been proposed. These schemes require less CSI, at the cost of only a small performance loss.

Through the use of the distribution of the channel condition, preallocation was done to be able to utilize the channel in an efficient manner, and to adjust the transmitted power and plan the transmission time.

The proposed system was not designed with the goal of achieving progressive decoding, but due to the small dependencies between the channel symbols, and the fact that the most important channel symbols are planned to be transmitted first, progressive decoding becomes possible.

Through simulations, it was shown how the system adapts on the fly to the channel seen, and that the true used time and power might not be exactly equal to the assumed for each transmission, but will be right on average.

As a reference, a theoretical optimum was estimated. This reference is equal to the proposed system with an infinite number of perfect nonlinear dimensional changing mappings. Since previous systems have designed mappings without knowing the possible gain, a scheme to test a mapping of arbitrarily rate and performance was proposed. Through this scheme it was found that a dimensional expanding mapping of rate $\hat{r} = 3$ would be

beneficial for low $\bar{\gamma}$.

Since the system is robust, it means that the quality of the received image will vary around a mean value. For the case of video, this would give a gradual change from frame to frame similar to analog film. One traditional way of dealing with detected bit-errors in a video stream, is to freeze the current frame. For a poor channel, that leads to a noncontinuous video stream which might feel tiring for the user, as the human visual system is very good at detecting changes. A gradual degradation might, on the other hand, yield a smoother stream.

6.1 Contributions of the thesis

The main contribution of this thesis is the proposed an image transmission system for flat fading channels using nonlinear dimensional changing mappings. Through this work, other contributions have been made, and are summarized in the following.

- Preallocation was used to be able to use the channel in an efficient manner, and to able to plan the average transmitted power and transmission time. Efficient utilization of the channel comes at the cost of deviations from the set transmission power and transmission time.

- It was shown how the transmitted signal can be adapted to the channel variations on multiple levels. The first is through the known technique of splitting the CSNR range into separate regions, and use different settings for each region. In connection with this a known technique of inverting the channel within each region to get constant received CSNR was used. In addition to this, two ideas requiring less channel feedback was shown to work due to the robustness of the mappings. One where no scaling of the mapped signal was done, and one where the expected channel gain within a channel region was used.

- Two simple receive filters were developed to compensate for the channel gain mismatch assumed in the transmitter.

- Deeper understanding of the role of the mapping-rate in a system using nonlinear Shannon mappings was obtained through analysis of the results. This includes the effect of compensating mapping-rate mismatch through power allocation.

- A theoretical system was proposed. An estimation of an upper limit for a given image, average rate and transmission power was found.

- A theoretical system was developed to see the potential benefits of introducing new mappings with an arbitrarily performance.

6.2 Future research, ideas and thoughts

In the following list, some ideas and thoughts for future research are given.

- It has been shown that the probability of outage has a big impact on performance for the case when the CSNR range is split into an outage region and one region with transmission. This has, however, not been fully analyzed. For few CSNR regions, setting the correct outage level should increase the performance. The results further indicate that the number of channel regions could be kept low, $M \leq 2$, and still get good performance. The difference between the $M = 4$ and $M = 2$ will probably be smaller by setting the outage level optimally.

- What is the performance of the reference systems with less CSI? Should be compared with the proposed system for $\mathcal{S}$ and $\mathcal{N}$.

- An algorithm for finding the optimal representation points and region thresholds including the distortion of rate mismatch, should be developed. The results in this thesis have shown that a mismatch in mapping-rate for the source-blocks with largest variance has a big impact for the image quality for low $\bar{\gamma}$ values.

- The amount of side-information should be studied further. What is the loss in performance for different images when using a general image source-block distribution? Using such a general distribution may reduce the amount of side-information drastically. Another scenario is to quantize the different source-block variances into a discrete set of values. The increased rate for protection of the side-information should also be studied.

- The results show that the effect of the channel gain mismatch filter in the receiver has a large impact on performance. It was also shown that it was actually beneficial not to pre-scale the mapped signal to the channel gain in a given region, as the receive filter would reduce the power of the noise. This effect should be looked further into, maybe a combination of pre-scaling and no pre-scaling within a region should be used, or maybe different strategies for different regions?

- The proposed system can be extended to video transmission. In video transmission the robustness of the system can be fully utilized. It might be beneficial to allow preallocation over several frames, which

results in the need for a buffer in the receiver. The problem with correlation in time in video should then be further studied as well. A solution might be to use $3D$ filter banks, that is, however, a huge research topic.

- The system should be extended to OFDM channels, both time varying and time constant.
- The channel estimation has been assumed to be perfect in this thesis. In practice that is not the case, so the impact of inaccurate channel estimation at the receiver, and delayed CSI at the transmitter should be studied further. For a OFDM system with time varying sub-channels, such estimation must be done in frequency as well.
- As most of the transmitted symbols have continuous amplitude, it can be a problem to synchronize the receiver as there is no eye-diagram to use. This is an fundamental problem with this type of transmission that needs to be solved for an implemented system.

Appendix A

Allocating source-blocks to channel states

This appendix gives the proof of the combination of source-block and channel state.

Looking at equation (3.7) and noticing that $\sigma^2_{X_n}/\mu \geq 1$, and that $\gamma_{C_m} > 0$ for all practical situations, equation (3.7) can be rewritten into

$$r_{n,m} = \frac{a_n}{b_m}, \text{ for } a_n, b_m \in [0, \infty) \tag{A.1}$$

where $a_n = \log_2\left(\sigma^2_{X_n}/\mu\right)$ and $b_m = \log_2\left(1 + \gamma_{C_m}\right)$. This is possible since $\log_2(x)$ is an increasing function for positive values of x.

Without loss of generality b_m can be ordered so that a set with N elements where,

$$b_1 \geq b_2 \geq \ldots \geq b_N, \tag{A.2}$$

is obtained. Denote $\mathbf{a}_k$ as a random permutation

$$\mathbf{a}_k = \{a_{k_1}, a_{k_2}, \ldots, a_{k_N}\}. \tag{A.3}$$

The goal is to find the permutation k that minimizes

$$\sum_{n=1}^{N} \frac{a_{k_n}}{b_n}. \tag{A.4}$$

The solution to this, is the permutation where

$$a_1 \geq a_2 \geq \ldots \geq a_N. \tag{A.5}$$

The proof for $N = 2$ will be given first.

Proof. Let $b_1 \geq b_2$, and assume that

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} \leq \frac{a_2}{b_1} + \frac{a_1}{b_2}. \tag{A.6}$$

Rearranging the terms yields

$$\frac{a_1 - a_2}{b_1} < \frac{a_1 - a_2}{b_2} \tag{A.7}$$

which is only true if and only if $a_1 \geq a_2$. □

General case

Proof. Let a set $\mathbf{b}$ with N elements be organized as

$$b_1 \geq b_2 \geq \ldots \geq b_N. \tag{A.8}$$

Take another set $\mathbf{a}_k$ with N elements in a random order and combine it with set $\mathbf{b}$ in the following way

$$S(k) = \frac{a_{k_1}}{b_1} + \frac{a_{k_2}}{b_2} + \cdots + \frac{a_{k_N}}{b_N} \tag{A.9}$$

By finding the largest element a_{k_n} and swapping it with a_{k_1}, a new permutation of $\mathbf{a}$ with only two elements different than $\mathbf{a}_k$ is found. Let this new permutation be denoted $\mathbf{a}_{k+1}$. From the proof for the case when $N = 2$, it is clear that

$$S(k) \geq S(k+1) = \frac{a_{k+1_1}}{b_1} + \frac{a_{k+1_2}}{b_2} + \cdots + \frac{a_{k+1_N}}{b_N} \tag{A.10}$$

The next step is to find the second largest element in $\mathbf{a}_{k+1}$ and swap it with a_{k+1_2} generating $\mathbf{a}_{k+2}$, where $S(k+1) \geq S(k+2)$. Continuing this for N steps, results in all the elements in $\mathbf{a}_{k+N}$ being sorted in a descending order. □

Appendix B

Channel gain mismatch filter

To minimize the perturbation of the channel, the difference between the mapped signal $g(k)$, and the signal, $\hat{g}(k)$, going into the demapper, should be minimized. In this thesis this is done through a single tap channel gain mismatch filter $w_m(k)$. The difference between $g(k)$ and $\hat{g}(k)$ is minimized by using the MSE, and the difference, ϵ, is given by

$$\epsilon = E\left[(\hat{g}(k) - g(k))^2\right]. \tag{B.1}$$

ϵ is minimized over $w(k)$, where $w(k)$ is such that $\hat{g}(k)$ is given by

$$\hat{g}(k) = w(k)y(k) = w(k)\left(s(k)\sqrt{\alpha(k)} + n(k)\right). \tag{B.2}$$

B.1 No pre-scaling of channel symbols

Not scaling the output before transmission means that the mapped signal $g(k)$ is the same as the transmitted signal $s(k)$

$$g(k) = s(k). \tag{B.3}$$

$\hat{g}(k)$ is given by

$$\hat{g}(k) = w(k)\left(g(k)\sqrt{\alpha(k)} + n(k)\right). \tag{B.4}$$

An optimal one tap equalizer is found through the minimization of

$$\epsilon = E\left[\left(\left(s(k)\sqrt{\alpha(k)} + n(k)\right)w(k) - s(k)\right)^2\right] \tag{B.5}$$

over $w(k)$. Writing this out yields

$$\epsilon = E\left[\left(s(k)\sqrt{\alpha(k)}w(k) + n(k)w(k) - g(k)\right)^2\right] \tag{B.6}$$

$$\begin{aligned} =& E\left[s^2(k)\alpha(k)w^2(k) + n^2(k)w^2(k) + s^2(k)\right. \\ &+ 2s(k)\sqrt{\alpha(k)}w^2(k)n(k) - 2s^2(k)\sqrt{\alpha(k)}w(k) \\ &\left.- 2n(k)w(k)s(k)\right] \end{aligned} \tag{B.7}$$

Assuming that the transmitted signal $s(k)$ and the noise $n(k)$ are independent, (B.7) can be simplified as

$$\epsilon = \sigma_S^2\alpha(k)w^2(k) + w^2(k)\sigma_N^2 + \sigma_S^2 - 2\sigma_S^2\sqrt{\alpha(k)}w(k) \tag{B.8}$$

Differentiating ϵ with respect to filter tap $w(k)$, and setting equal to zero

$$\frac{d\epsilon}{dw(k)} = 2\sigma_S^2\alpha(k)w(k) + 2w(k)\sigma_N^2 - 2\sigma_X^2\sqrt{\alpha(k)} \tag{B.9}$$

$$\frac{d\epsilon}{dw(k)} = 0$$

$$\Downarrow$$

$$w(k)\left(\alpha(k)\sigma_S^2 + \sigma_N^2\right) = \sqrt{\alpha(k)}\sigma_S^2 \tag{B.10}$$

$$w(k) = \frac{\sqrt{\alpha(k)}}{\alpha(k) + \frac{\sigma_N^2}{\sigma_S^2}}. \tag{B.11}$$

The transmitted power is dependent on the power in the mapped signal $g(k)$. The setting of a mapping can either be chosen from a channel state m, or for a specific power level for a given block n, when using mapping-rate mismatch compensation. The channel gain mismatch factor will then be given by

$$w_m(k) = \frac{\sqrt{\alpha(k)}}{\alpha(k) + \frac{\sigma_N^2}{\sigma_{S\,m}^2}}, \tag{B.12}$$

for $\mathfrak{NB}$, And

$$w_n(k) = \frac{\sqrt{\alpha(k)}}{\alpha(k) + \frac{\sigma_N^2}{\sigma_{S\,n}^2}}, \tag{B.13}$$

for $\mathfrak{B}$.

B.2 Using state centroid as scaling factor

The transmitter scales the mapped signal $g(k)$ down with the channel gain α_{A_m} corresponding to the centroid CSNR value in a region. The transmitted signal $s(k)$ is then given by

$$s(k) = \frac{g(k)}{\sqrt{\alpha_{A_m}}}. \tag{B.14}$$

Resulting in an expression for $g(k)$ given by

$$g(k) = s(k)\sqrt{\alpha_{A_m}}. \tag{B.15}$$

An optimal one tap equalizer $w(k)$ is found by minimizing ϵ with respect to $w(k)$. The expression for ϵ is found through inserting equation (B.15) and (B.2) into equation (B.1), giving

$$\epsilon = E\left[\left(\left(s(k)\sqrt{\alpha(k)} + n(k)\right) w(k) - s(k)\sqrt{\alpha_{A_m}}\right)^2\right]. \tag{B.16}$$

Writing this out results in

$$\epsilon = E\left[\left(s(k)\sqrt{\alpha(k)}w(k) + n(k)w(k) - s(k)\sqrt{\alpha_{A_m}}\right)^2\right] \tag{B.17}$$

$$\begin{aligned} =& E\Big[s^2(k)\alpha(k)w^2(k) + n^2(k)w^2(k) + s^2(k)\alpha_{A_m} \\ &+ 2s(k)\sqrt{\alpha(k)}w^2(k)n(k) - 2s^2(k)\sqrt{\alpha(k)}w(k)\sqrt{\alpha_{A_m}} \\ &- 2n(k)w_m(k)s(k)\sqrt{\alpha_{A_m}}\Big] \end{aligned} \tag{B.18}$$

Assuming that the transmitted signal $s(k)$ and the noise $n(k)$ are independent, (B.18) can be simplified as

$$\epsilon = \alpha(k)w^2(k)\sigma_S^2 + w^2(k)\sigma_N^2 + \alpha_{A_m}\sigma_S^2 - 2\sqrt{\alpha(k)}\sqrt{\alpha_{A_m}}w(k)\sigma_S^2 \tag{B.19}$$

Differentiating with respect to the filter tap $w(k)$, yields

$$\frac{d\epsilon}{dw(k)} = 2\alpha(k)w(k)\sigma_S^2 + 2w(k)\sigma_N^2 - 2\sqrt{\alpha(k)}\sqrt{\alpha_{A_m}}\sigma_S^2 \tag{B.20}$$

$$\frac{d\epsilon}{dw(k)} = 0$$

$$\Downarrow$$

$$w(k)\left(\alpha(k)\sigma_S^2 + \sigma_N^2\right) = \sqrt{\alpha(k)}\sqrt{\alpha_{A_m}}\sigma_S^2 \tag{B.21}$$

$$w(k) = \frac{\sqrt{\alpha(k)}\sqrt{\alpha_{A_m}}}{\alpha(k) + \frac{\sigma_N^2}{\sigma_S^2}} \tag{B.22}$$

For channel region m, the channel mismatch filter is given by

$$w_m(k) = \frac{\sqrt{\alpha(k)}\sqrt{\alpha_{A_m}}}{\alpha(k) + \frac{\sigma_N^2}{\sigma_{S\,m}^2}}. \tag{B.23}$$

And when using mapping-rate mismatch compensation, the channel mismatch filter is given by

$$w_n(k) = \frac{\sqrt{\alpha(k)}\sqrt{\alpha_{A_m}}}{\alpha(k) + \frac{\sigma_N^2}{\sigma_{S\,n}^2}}. \tag{B.24}$$

Appendix C

Finding channel regions using the *simple* algorithm

Maximizing the channel transmission rate R_c, when assuming a fixed received CSNR, is done by maximizing

$$R_c = \max_{\sigma^2_{S\,m}:\sum_m \int_m \sigma^2_{S\,m}(\gamma) f_\gamma(\gamma) d\gamma = \bar{\sigma}^2_S} \frac{1}{2} \sum_{m=1}^{M} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \log_2 \left(1 + \frac{\sigma^2_{S\,m}(\gamma)}{\bar{\sigma}^2_S} \gamma\right) f_\gamma(\gamma) d\gamma, \tag{C.1}$$

under power constraint

$$\sum_{m=1}^{M} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \sigma^2_{S\,m}(\gamma) f_\gamma(\gamma) d\gamma \leq \bar{\sigma}^2_S. \tag{C.2}$$

In this thesis, the *GNU scientific library* (GSL) [Galassi et al., 2006] is used for optimization. The numerical optimization done there does not support constraints, so these have to be included in the object function. To optimize this without using constrained optimization, *quadratic penalty terms* for the power constraint must be used, and *logarithmic barrier functions* [Nocedal and Wright, 1999] for the non-negative constraints of the variables. Since the GSL methods used minimize a given function, the final expression to

minimize is given by

$$\begin{aligned}\mathcal{B}(\gamma_{\mathbf{C}}, \gamma_{\mathbf{T}}, \mu) = &- \frac{1}{2}\sum_{m=1}^{M} \log_2\left(1 + \frac{\sigma_{Sm}^2(\gamma)}{\bar{\sigma}_S^2}\gamma\right) f(\gamma)d\gamma \\ &+ \frac{1}{2\varkappa}\left(\sum_{m=1}^{M} \frac{\sigma_{Sm}^2(\gamma)}{\bar{\sigma}_S^2} \int_{\gamma_{T_{m-1}}}^{\gamma_{T_m}} \frac{1}{\gamma} f(\gamma)d\gamma - 1\right)^2 \\ &- \varkappa \sum_{m=1}^{M} \log(\gamma_{T_m} - \gamma_{C_m}) - \varkappa \sum_{m=1}^{M} \log(\gamma_{C_m} - \gamma_{T_{m-1}}),\end{aligned} \tag{C.3}$$

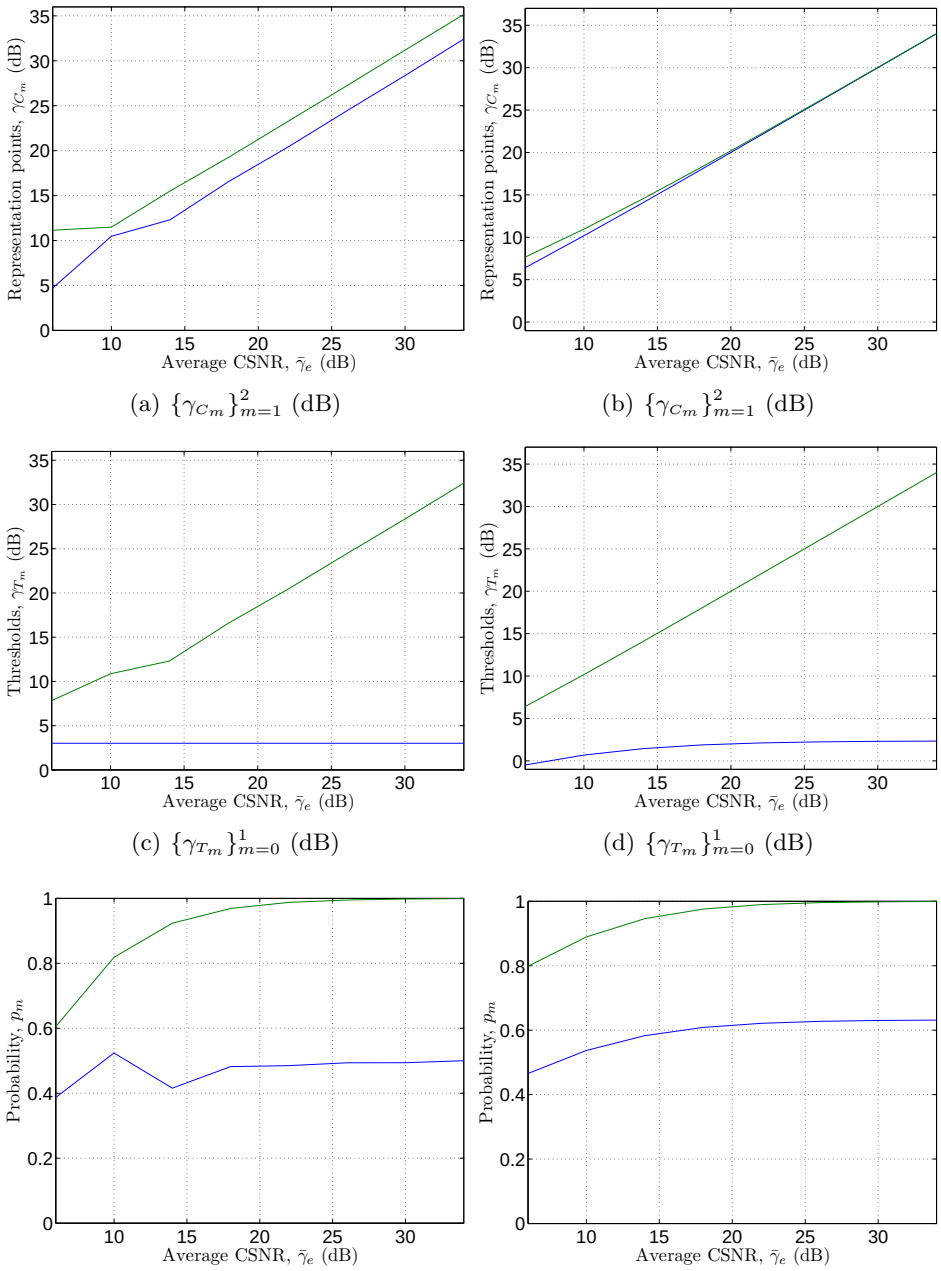
where $\varkappa < 1$ is the barrier parameter. The second term is to make sure the power is correct, and the third and fourth terms are to make sure the representation points stay between the region thresholds.

Using GSL will not guarantee the global minimum, but will most likely end up in a local minimum.

Appendix D

Regions and representation points

In Section 3.2.3 the thresholds $\{\gamma_{T_m}\}_{m=0}^{M-1}$, representation points $\{\gamma_{C_m}\}_{m=1}^{M}$ and region probabilities was shown for the *simple* algortithm and *complex* algorithm, for $M = 4$. In this appendix the thresholds and representation points are given for $M = \{1, 2\}$.



(a) $\{\gamma_{C_m}\}_{m=1}^{2}$ (dB)

(b) $\{\gamma_{C_m}\}_{m=1}^{2}$ (dB)

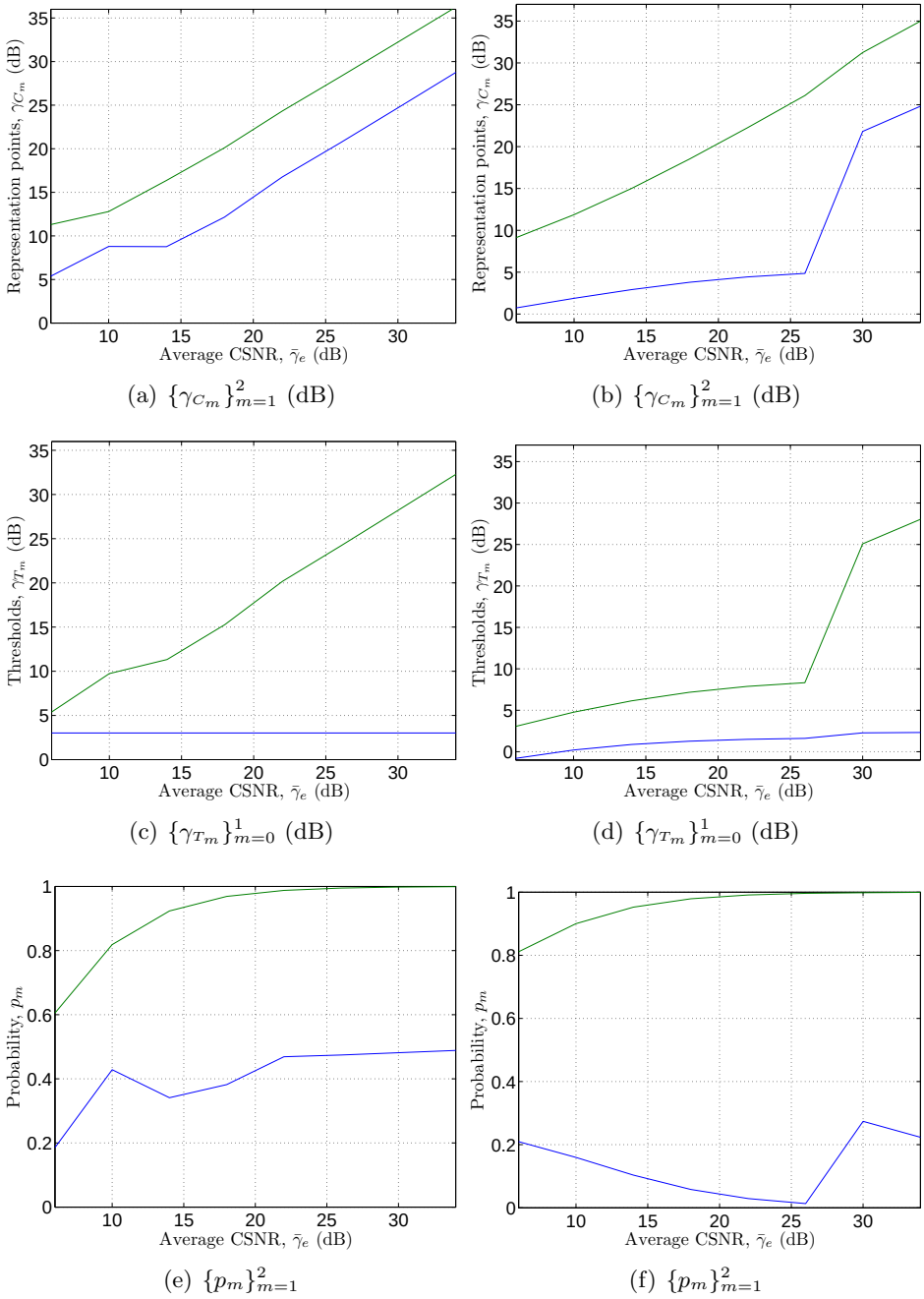
(c) $\{\gamma_{T_m}\}_{m=0}^{1}$ (dB)

(d) $\{\gamma_{T_m}\}_{m=0}^{1}$ (dB)

(e) $\{p_m\}_{m=1}^{2}$

(f) $\{p_m\}_{m=1}^{2}$

Figure D.1: Representation points, γ_{C_m}, thresholds, γ_{T_m} for CAS set to $\mathcal{N}$, and accumulated region probability p_m. Found by *simple* algorithm(left column), *complex* algorithm(right column), $M = 2$.



(a) $\{\gamma_{C_m}\}_{m=1}^{2}$ (dB)

(b) $\{\gamma_{C_m}\}_{m=1}^{2}$ (dB)

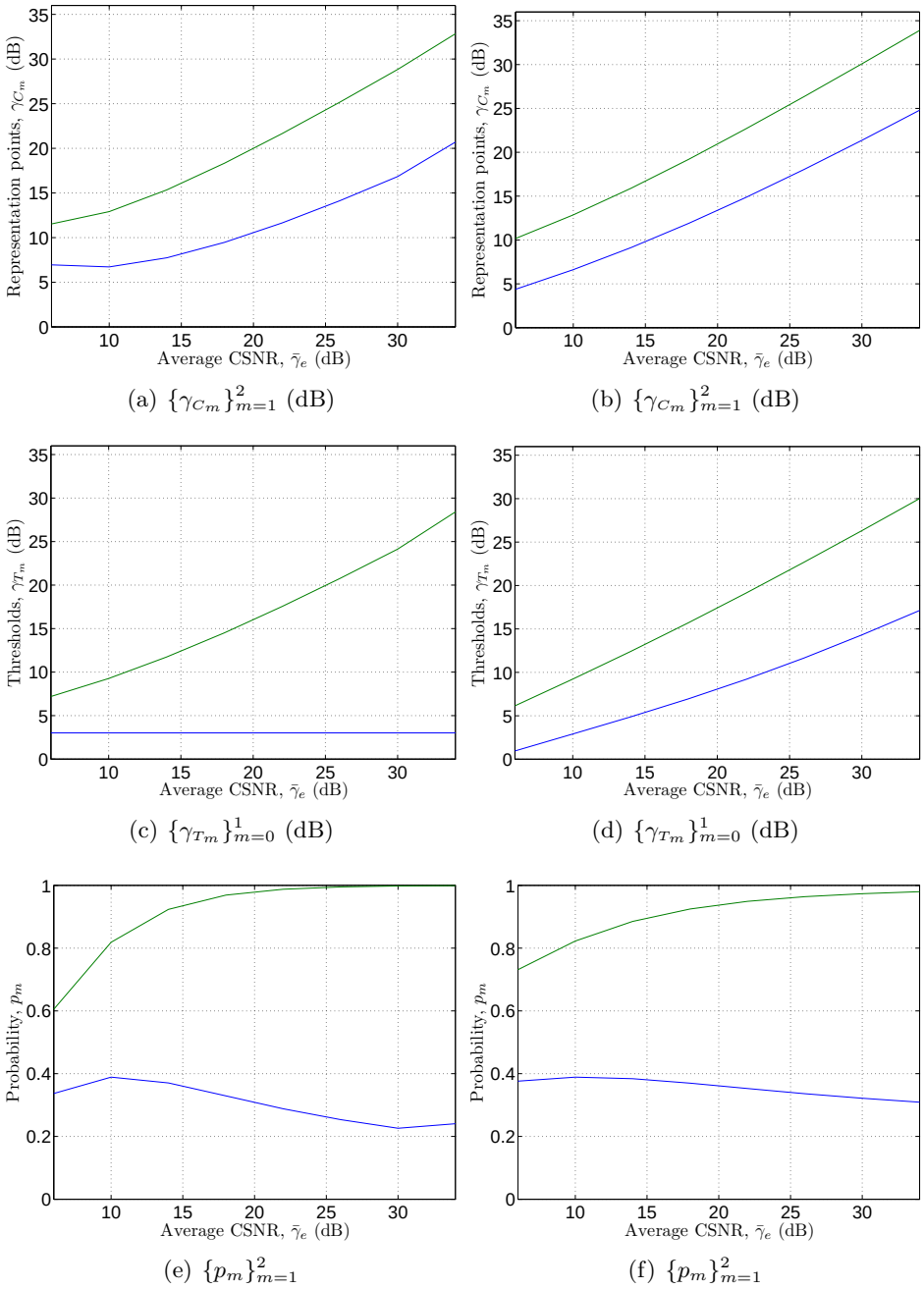
(c) $\{\gamma_{T_m}\}_{m=0}^{1}$ (dB)

(d) $\{\gamma_{T_m}\}_{m=0}^{1}$ (dB)

(e) $\{p_m\}_{m=1}^{2}$

(f) $\{p_m\}_{m=1}^{2}$

Figure D.2: Representation points, γ_{C_m}, thresholds, γ_{T_m} for CAS set to $\mathcal{S}$, and accumulated region probability p_m. Found by *simple* algorithm(left column), *complex* algorithm(right column), $M = 2$.



(a) $\{\gamma_{C_m}\}_{m=1}^{2}$ (dB)

(b) $\{\gamma_{C_m}\}_{m=1}^{2}$ (dB)

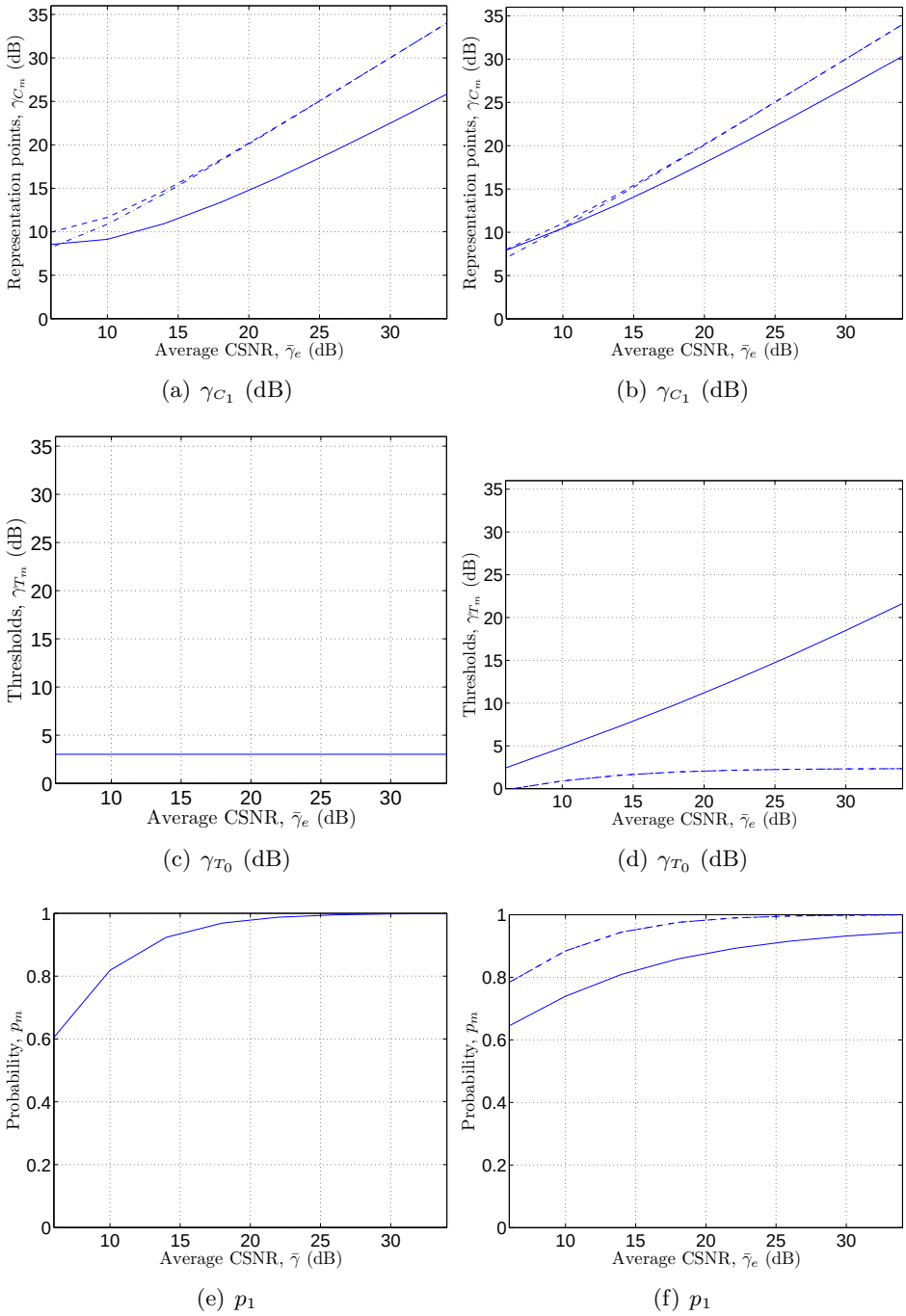
(c) $\{\gamma_{T_m}\}_{m=0}^{1}$ (dB)

(d) $\{\gamma_{T_m}\}_{m=0}^{1}$ (dB)

(e) $\{p_m\}_{m=1}^{2}$

(f) $\{p_m\}_{m=1}^{2}$

Figure D.3: Representation points, γ_{C_m}, thresholds, γ_{T_m} for CAS set to $\mathcal{C}$, and accumulated region probability p_m. Found by *simple* algorithm(left column), *complex* algorithm(right column), $M = 2$.



(a) γ_{C_1} (dB) (b) γ_{C_1} (dB)

(c) γ_{T_0} (dB) (d) γ_{T_0} (dB)

(e) p_1 (f) p_1

Figure D.4: Representation points, γ_{C_m}, thresholds, γ_{T_m} for CAS set to $\mathcal{N}$(dash-dotted), $\mathcal{S}$(dashed) and $\mathcal{C}$(solid), and accumulated region probability p_m. Found by *simple* algorithm(left column), *complex* algorithm(right column), $M = 1$.

Appendix E

Original images

Three images where used during the simulations. The original 512×512 images are presented in Figure E.1, Figure E.2 and Figure E.3.

Figure E.1: Original image "Lena"

Figure E.2: Original image “Goldhill”

Figure E.3: Original image "Bridge"

Bibliography

Alouini, M. S. and Goldsmith, A. J. (1999). Capacity of rayleigh fading channels under different adaptive transmission and diversity-combining techniques. *IEEE Transactions on Vehicular Technology*, 48(4):1165–1181.

Alouini, M. S. and Goldsmith, A. J. (2000). Comparison of fading channel capacity under different CSI assumptions. In *Vehicular Technology Conference, 2000. IEEE VTS-Fall VTC 2000. 52nd*, volume 4, pages 1844–1849, Boston, MA.

Balasingham, I. (1998). *On Optimal Perfect Reconstruction Filter Banks for Image Compression*. PhD thesis, Norwegian University of Science and Technology, Norway.

Berger, T. (1971). *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice Hall, Englewood Cliffs, NJ.

Berger, T. and Tufts, D. (1967). Optimum pulse amplitude modulation–i: Transmitter-receiver design and bounds from information theory. *IEEE Transactions on Information Theory*, 13(2):196–208.

Berrou, C., Glavieux, A., and Thitimajshima, P. (1993). Near shannon limit error-correcting coding and decoding:turbo-codes. 1. In *Communications, 1993. ICC 93. Geneva. Technical Program, Conference Record, IEEE International Conference on*, volume 2, pages 1064–1070, Geneva, Switzerland.

Blahut, R. (1972). Computation of channel capacity and rate-distortion functions. *IEEE Transactions on Information Theory*, 18(4):460–473.

Boeglen, H. and Chatellier, C. (2006). On the robustness of a joint source-channel coding scheme for image transmission over non frequency selective rayleigh fading channels. In *Information and Communication Technologies, 2006. ICTTA '06. 2nd*, volume 2, pages 2320–2324.

Bystrom, M. and Modestino, J. W. (1998). Recent advances in joint source-channel coding of video. In *Signals, Systems, and Electronics, 1998. ISSSE 98. 1998 URSI International Symposium on*, pages 332–337, Pisa.

Caire, G. and Shamai, S. (1999). On the capacity of some channels with channel state information. *IEEE Transactions on Information Theory*, 45(6):2007–2019.

Cavers, J. K. (1991). An analysis of pilot symbol assisted modulation for rayleigh fading channels [mobile radio]. *IEEE Transactions on Vehicular Technology*, 40(4):686–693.

Chung, S.-Y., Forney, G. D., J., 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 Communications Letters*, 5(2):58–60.

Cover, T. M. and Thomas, J. A. (1991). *Elements of Information Theory.* Wiley, New York.

Coward, H. (2001). *Joint source-channel coding: Development of methods and utilization in image communication.* PhD thesis, Norwegian University of Science and Technology, Norway.

Coward, H. and Ramstad, T. (1999). Hybrid digital-analog transmission of analog source signals. pages 17–21, Asker, Norway. Proc. Norwegian Signal Processing Symposium and Workshop (NORSIG).

Coward, H. and Ramstad, T. (2000a). Bandwidth doubling incombined source-channel coding of memoryless gaussian sources. volume 1, pages 571–576, Honolulu, HI, USA. Proc. IEEE Int. Symp. Intell. Signal Processing Commun. Syst. (ISPACS).

Coward, H. and Ramstad, T. (2000b). Robust image communication using bandwidth reducing and expanding mappings. volume 2, pages 1384–1388, Pacific Grove, CA, USA. Conf. Record 34th Asilomar Conf. Signals, Syst., Comput.

Coward, H. and Ramstad, T. A. (2000c). Quantizer optimization in hybrid digital-analog transmission of analog source signals. In *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, volume 5, pages 2637–2640, Istanbul.

DAB (1997). European telecommunication standard institute, ETSI: Radio broadcast systems; digital audio broadcasting (DAB) to mobile, portable and fixed receivers, ETS 300 401.

EIA/TIA-95 (1989). Mobile station-base station compatibility standard for dual-mode wideband spectrum cellular system.

Farvardin, N. (1990). A study of vector quantization for noisy channels. *IEEE Transactions on Information Theory*, 36(4):799–809.

Farvardin, N. and Vaishampayan, V. (1987). Optimal quantizer design for noisy channels: An approach to combined source - channel coding. 33(6):827–838.

Farvardin, N. and Vaishampayan, V. (1991). On the performance and complexity of channel-optimized vector quantizers. *IEEE Transactions on Information Theory*, 37(1):155–160.

Fine, T. (1964). Properties of an optimum digital system and applications. 10(4):287–296.

Floor, P. and Ramstad, T. (2006a). Noise analysis for dimension expanding mappings in source-channel coding. Cannes, France. Proc. Signal Processing Advances in Wireless Communications (SPAWC).

Floor, P. and Ramstad, T. (2006b). Noise immunity for 1:n and m:1 nonlinear mappings for source-channel coding. In *Data Compression Conference, 2006. DCC 2006. Proceedings.*

Floor, P. A. and Ramstad, T. A. (2006c). Dimension reducing mappings in joint source-channel coding. In *Signal Processing Symposium, 2006. NORSIG 2006. Proceedings of the 7th Nordic*, pages 282–285, Reykjavik, Iceland.

Fuldseth, A. (1997). *Robust subband video compression for noisy channels with multilevel signaling.* PhD thesis, Norwegian University of Science and Technology.

Fuldseth, A. and Ramstad, T. A. (1997). Bandwidth compression for continuous amplitude channels based on vector approximation to a continuous subset of the source signal space. In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, volume 4, pages 3093–3096.

Galassi, M., Theiler, J., and Davies, J. (2006). *GNU Scientific Library Reference Manual.* Network Theory Limited; 2Rev Ed edition (1 Oct 2006).

Gallager, R. G. (1963). *Low-Density Parity-Check Codes.* MIT Press, Cambridge, MA.

Gersho, A. and M. Gray, R. (1992). *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Boston, MA, USA.

Gjendemsjø, A., Øien, G., and Holm, H. (2005). Optimal power control for discrete-rate link adaptation schemes with capacity-approaching coding. In *Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE*, volume 6.

Goblick, T., J. (1965). Theoretical limitations on the transmission of data from analog sources. *IEEE Transactions on Information Theory*, 11(4):558–567.

Goldsmith, A. (2005). *Wireless Communications*. Cambridge University Press, first edition.

Goldsmith, A. and Chua, S.-G. (1997). Variable-rate variable-power MQAM for fading channels. *IEEE Transactions on Communications*, 45(10):1218–1230.

Goldsmith, A. J. and Effros, M. (1998). Joint design of fixed-rate source codes and multiresolution channel codes. *IEEE Transactions on Communications*, 46(10):1301–1312.

Goldsmith, A. J. and Varaiya, P. P. (1997). Capacity of fading channels with channel side information. *IEEE Transactions on Information Theory*, 43(6):1986–1992.

Hagenauer, J. (1988). Rate-compatible punctured convolutional codes (RCPC codes) and their applications. *IEEE Transactions on Communications*, 36(4):389–400.

Håkonsen, G. and Ramstad, T. (2006a). Image transmission over flat fading channels using joint source channel coding. In *Data Compression Conference, 2006. DCC 2006. Proceedings*.

Håkonsen, G. H. and Ramstad, T. A. (2005). Joint source channel coder for image transmission over flat fading channels. Proc. Norwegian Signal Processing Symposium and Workshop (NORSIG).

Håkonsen, G. H. and Ramstad, T. A. (2006b). On losses of performance in a joint source channel coder. In *Signal Processing Symposium, 2006. NORSIG 2006. Proceedings of the 7th Nordic*, pages 278–281, Reykjavik, Iceland.

Håkonsen, G. H., Ramstad, T. A., and Gjendemsjø, A. (2006). Image transmission with adaptive power and rate allocation over flat fading channels

using joint source channel coding. In *International Conference on Wireless Information Networks and Systems, WINSYS, 2006, Proceedings*, pages 158–163, Setubal, Lisboa.

Hekland, F., Øien, G., and Ramstad, T. (2005). Using 2:1 shannon mapping for joint source-channel coding. In *Data Compression Conference, 2005. Proceedings. DCC 2005*, pages 223–232.

Hekland, F. and Ramstad, T. (2006a). Optimal rate-constrained transcoding for a 2:1 bandwidth reducing shannon mapping. Cannes, France. Proc. Signal Processing Advances in Wireless Communications (SPAWC).

Hekland, F. and Ramstad, T. A. (2006b). Digitising the 2:1 shannon mappings for transport over heterogeneous networks. *2006. DCC 2006. Proceedings Data Compression Conference*.

Hole, K., Holm, H., and Øien, G. (2000). Adaptive multidimensional coded modulation over flat fading channels. *IEEE Journal on Selected Areas in Communications*, 18(7):1153–1158.

Holm, H., Øien, G., Alouini, M.-S., Gesbert, D., and Hole, K. (2003). Optimal design of adaptive coded modulation schemes for maximum average spectral efficiency. In *Signal Processing Advances in Wireless Communications, 2003. SPAWC 2003. 4th IEEE Workshop on*, pages 403–407.

ISO/IEC (1991). Digital compression and coding of continuous-tone still images (JPEG).

ISO/IEC (2000). Information technology - JPEG2000 image coding system - part 1: Core coding system.

Jafarkhani, H., Farvardin, N., and Lee, C. C. (1994). Adaptive image coding based on the discrete wavelet transform. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 3, pages 343–347, Austin, TX.

Jakes, W. C. (1974). *Microwave Mobile Communication*. Wiley, New York.

Joshi, R. L. and Fischer, T. R. (1995). Comparison of generalized gaussian and laplacian modeling in DCT image coding. *IEEE Signal Processing Letters*, 2(5):81–82.

Kafedziski, V. (1998). Joint source channel coding of images over frequency selective fading channels with feedback using DCT and multicarrier block pulse amplitude modulation. In *Signals, Systems and Computers, 1998.*

Conference Record of the Thirty-Second Asilomar Conference on, volume 1, pages 37–41, Pacific Grove, CA.

Kang, D. W., Kang, S. W., and Lee, C. W. (1994). Entropy reduction of symbols by source splitting and its application to video coding. *Signal Processing: Image Communication*, 6(5):471–478.

Kleider, J. E. and Abousleman, G. P. (2000). Orthogonal source adaptive modulation for image transmission over noisy channels. In *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, volume 6, pages 2099–2102, Istanbul.

Kotel'nikov, V. A. (1959). *The theory of optimum noise immunity.* McGraw-Hill Book Company.

Kozintsev, I. and Ramchandran, K. (1998). Robust image transmission over energy-constrained time-varying channels using multiresolution joint source-channel coding. *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, 46(4):1012–1026.

Kurtenbach, A. and Wintz, P. (1969). Quantizing for noisy channels. *Communications, IEEE Transactions on [legacy, pre - 1988]*, 17(2):291–302.

Lee, K.-H. and Petersen, D. (1976). Optimal linear coding for vector channels. *Communications, IEEE Transactions on [legacy, pre - 1988]*, 24(12):1283–1290.

Lervik, J. M. (1996). *Subband Image Communication over Digital Transparent and Analog Waveform Channels.* PhD thesis, Norwegian University of Science and Technology.

Lervik, J. M. and Fischer, T. R. (1997a). Robust subband image coding for waveform channels with optimum power- and bandwidth- allocation. In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, volume 4, pages 3089–3092, Munich.

Lervik, J. M. and Fischer, T. R. (1997b). Robust subband image coding for waveform channels with optimum power- and bandwidth- allocation. volume 4, pages 3089 – 3092, Mnchen, Germany. Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP).

Lervik, J. M. and Ramstad, T. A. (1996). Optimality of multiple entropy coder systems for nonstationary sources modeled by a mixture distribution. In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, volume 4, pages 875–878, Atlanta, GA, USA. IEEE.

Lin, L., Yates, R. D., and Spasojevic, P. (2006). Adaptive transmission with finite code rates. *IEEE Transactions on Information Theory*, 52(5):1847–1860.

Liu, W. and Daut, D. G. (2005). Joint source-channel decoding for JPEG2000 codestream transmission over flat fading channels. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 3, pages 900–3.

MacKay, D. J. C. and Neal, R. M. (1996). Near shannon limit performance of low density parity check codes. *Electronics Letters*, 32.

McEliece, R. and Stark, W. (1984). Channels with block interference. *IEEE Transactions on Information Theory*, 30(1):44–53.

McRae, D. (1971). Performance evaluation of a new modulation technique. *Communications, IEEE Transactions on [legacy, pre - 1988]*, 19(4):431–445.

Mittal, U. and Phamdo, N. (2002). Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications. *IEEE Transactions on Information Theory*, 48(5):1082–1102.

Modestino, J. and Daut, D. (1979). Combined source-channel coding of images. *Communications, IEEE Transactions on [legacy, pre - 1988]*, 27(11):1644–1659.

Modestino, J., Daut, D., and Vickers, A. (1981). Combined source-channel coding of images using the block cosine transform. *Communications, IEEE Transactions on [legacy, pre - 1988]*, 29(9):1261–1274.

Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization*. Springer.

Nosratinia, A., Lu, J., and Aazhang, B. (2003). Source-channel rate allocation for progressive transmission of images. *IEEE Transactions on Communications*, 51(2):186–196.

Pan, X., Banihashemi, A. H., and Cuhadar, A. (2006). Progressive transmission of images over fading channels using rate-compatible LDPC codes. *IEEE Transactions on Image Processing*, 15(12):3627–3635.

Ramamurthi, B. and Gersho, A. (1986). Classified vector quantization of images. *Communications, IEEE Transactions on [legacy, pre - 1988]*, 34(11):1105–1115.

Ramstad, T. A. (2006). On joint source-channel coding for the non-white gaussian case. Cannes, France. Proc. Signal Processing Advances in Wireless Communications (SPAWC).

Rappaport, T. S. (1996). *Wireless Communcations: Principles & practice.* Prentice Hall.

Ruf, M. J. and Modestino, J. W. (1999). Operational rate-distortion performance for joint source and channel coding of images. *IEEE Transactions on Image Processing*, 8(3):305–320.

Said, A. and Pearlman, W. A. (1996). A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Transactions on Circuits and Systems for Video Technology*, 6(3):243–250.

Shannon, C. E. (1948a). A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423 and 623–656.

Shannon, C. E. (1948b). A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423.

Shannon, C. E. (1949). Communication in the presence of noise. 37(1):10–21.

Shannon, C. E. (1959). Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.*, pages 142–163.

Shapiro, J. M. (1993). Embedded image coding using zerotrees of wavelet coefficients. *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, 41(12):3445–3462.

Sherwood, P. G. and Zeger, K. (1997a). Progressive image coding for noisy channels. *IEEE Signal Processing Letters*, 4(7):189–191.

Sherwood, P. G. and Zeger, K. (1997b). Progressive image coding on noisy channels. In *Data Compression Conference, 1997. DCC '97. Proceedings*, pages 72–81, Snowbird, UT.

Skoglund, M., Phamdo, N., and Alajaji, F. (2002). Design and performance of VQ-based hybrid digital-analog joint source-channel codes. *IEEE Transactions on Information Theory*, 48(3):708–720.

Skoglund, M., Phamdo, N., and Alajaji, F. (2006). Hybrid digitalanalog sourcechannel coding for bandwidth compression/expansion. *IEEE Transactions on Information Theory*, 52(8):3757–3763.

Srinivasan, M. and Chellappa, R. (1997). Joint source-channel subband coding of images. In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, volume 4, pages 2925–2928, Munich.

Stüber, G. L. (2001). *Principles of Mobile Communication.* Kluwer Academic Publishers.

Sun, Y. and Xiong, Z. (2006). Progressive image transmission over space-time coded OFDM-based MIMO systems with adaptive modulation. *IEEE Transactions on Mobile Computing*, 5(8):1016–1028.

Tanabe, N. and Farvardin, N. (1992). Subband image coding using entropy-coded quantization over noisy channels. *IEEE Journal on Selected Areas in Communications*, 10(5):926–943.

Taubman, D. and Marcellin, M. (2001). *JPEG2000 Image compression fundamentals, standards and practice.* Kluwer Academic Publishers.

Thomas, C., May, C., and Welti, G. (1975). Hybrid amplitude-and-phase modulation for analog data transmission. *Communications, IEEE Transactions on [legacy, pre - 1988]*, 23(6):634–645.

Thomos, N., Boulgouris, N. V., and Strintzis, M. G. (2005). Wireless image transmission using turbo codes and optimal unequal error protection. *IEEE Transactions on Image Processing*, 14(11):1890–1901.

Tie, H.-X., Goldsmith, A., and Effros, M. (1998). Joint design of fixed-rate source codes and UEP channel codes for fading channels. In *Signals, Systems and Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference on*, volume 1, pages 92–96, Pacific Grove, CA.

Vaidyanathan, P. P. (1993). *Multirate systems and filter banks.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Vaishampayan, V. (1989). *Combined Source-Channel Coding for Bandlimited Waveform Channels.* PhD thesis, University of Maryland.

Vaishampayan, V. A. and Farvardin, N. (1990). Optimal block cosine transform image coding for noisy channels. *IEEE Transactions on Communications*, 38(3):327–336.

Vishwanath, S. and Goldsmith, A. (2000). Exploring adaptive turbo coded modulation for flat fading channels. In *Vehicular Technology Conference, 2000. IEEE VTS-Fall VTC 2000. 52nd*, volume 4, pages 1778–1783, Boston, MA.

Vishwanath, S. and Goldsmith, A. (2003). Adaptive turbo-coded modulation for flat-fading channels. *IEEE Transactions on Communications*, 51(6):964–972.

Wang, Y., Alajaji, F., and Linder, T. (2005). Design of VQ-based hybrid digital-analog joint source-channel codes for image communication. *2005. Proceedings. DCC 2005 Data Compression Conference*, pages 193–202.

Westerink, P., Biemond, J., and Boekee, D. (1988a). An optimal bit allocation algorithm for sub-band coding. In *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, pages 757–760, New York, NY.

Westerink, P. H., Boekee, D. E., Biemond, J., and Woods, J. W. (1988b). Subband coding of images using vector quantization. *IEEE Transactions on Communications*, 36(6):713–719.

Zhang, Y., Yuan, D., and Zhang, H. (2004a). Adaptive LDPC in image transmission under rayleigh fading channel. In *Advanced Communication Technology, 2004. The 6th International Conference on*, volume 1, pages 192–196.

Zhang, Y., Yuan, D., Zhang, H., and Zhang, H. (2004b). The application of LDPC codes in image transmission. In *Communications, Circuits and Systems, 2004. ICCCAS 2004. 2004 International Conference on*, volume 1, pages 29–33.

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 Transactions on Circuits and Systems for Video Technology*, 9(5):823–838.