Adaptive Coded Modulation: Design and Simulation with Realistic Channel State Information

Ola Jetlund

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOKTOR INGENIØR



Department of Electronics and Telecommunications
Norwegian University of Science and Technology

2005

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

Doktor ingeniøravhandling 2005:38
Rapportnr. 420502

ISBN 82-471-6943-6 (printed version)
ISBN 82-471-6942-8 (electronic version)
ISSN 1503-8181

Abstract

Spectrally efficient transmission schemes are becoming a more common requirement for digital communication systems. Especially in wireless communication since the bandwidth of available frequencies is a shared resource. In addition, wireless communication systems suffer from transmission media with varying conditions. Adaptive coded modulation (ACM) has been suggested as a bandwidth-efficient transmission technique in wireless fading environments. The use of ACM is motivated by its ability to improve spectral efficiency (SE) by adapting the transmission rates to the variations in channel signal-to-noise ratio. Any ACM scheme rely on being able to predict future states of the transmission medium. Under idealized conditions, such as the prediction being perfect, an ACM scheme can be configured to maximize the SE under the condition of the bit error rate (BER) being below a specified target BER. Here, computer simulations of an example system show that such systems in some cases fail to achieve the target BER, since the idealized conditions used in the design process do not hold in a realistic setup.

By limiting the number of transmission modes, introducing imperfect prediction, and other practical conditions such as delay in the communication system and probability of outage, a more practical ACM scheme can be considered. We show that it is still possible to optimize the performance of such schemes. A wireless communication channel with a Rayleigh fading envelope is assumed here since most results then can be presented in closed form expressions. For other distributions of the fading, results can be found numerically. By optimizing the performance of an idealized ACM scheme using capacity achieving channel codes, we have been able to upper bound the SE of practical ACM schemes. The results also provide us with a technique to control the average BER in the case of imperfect knowledge of future channel states. Simulation results for a modified ACM scheme that uses this technique is shown to have an average BER that is less than the target BER.

Preface

This dissertation is submitted in partial fulfillment of the requirements for the degree of *doktor ingeniør* at the Department of Electronics and Telecommunications, Norwegian University of Science and Technology (NTNU). My advisors have been Professor Geir E. Øien at Department of Electronics and Telecommunications, NTNU and Professor Kjell J. Hole at the Selmer Center, Department of Informatics, University of Bergen (UiB).

The studies have been carried out in the period from January 2000 to December 2004. The work includes the equivalent of a year of full-time course studies, 18 months of being a teaching assistant in various undergraduate and graduate courses, and five months teaching two graduate classes.

The work was funded by a scholarship from Telenor, via the project Turbo codes, Access and Network Technology (TURBAN) which is a collaboration between NTNU and Telenor. The teaching assistantship was funded by the Department of Electronics and Telecommunications, NTNU.

Acknowledgments

It is a pleasure to finally be able to thank my supervisors, Professor Geir E. Øien and Professor Kjell J. Hole. Their knowledge in the art of writing technical documents is genuinely superb. The ideas and suggestions made by Professor Øien have been very helpful during my studies. Professor Hole has been very eager to teach me the skills of scientific research.

Other researchers at the Signal Processing Group at NTNU must also be acknowledged, especially fellow doktor ingeniør student Bengt Holter, and Doktor Henrik Holm. The Signal Processing Group as a whole has provided a pleasant working environment that none should be without. Other researches that I have discussed both scientific and non-scientific issues with are also thanked, most notably Sivilingeniør Bård Myhre at Sintef in Trondheim.

I am most of all grateful for the unconditional love and support from my family, my incredible patient fiancé Safiye, and for my son Falk Markus making every day more joyful then the previous!

OLA JETLUND
Trondheim, Norway
January 2005

Contents

Abstract **i**

Preface **iii**

Acknowledgments **v**

Contents **vii**

Figures **xi**

Tables **xv**

Nomenclature **xvii**

Abbreviations **xxvii**

1 Introduction to Adaptive Coding and Modulation **1**
1.1 Spectrally Efficient Transmission Techniques 2
1.2 Problem Formulation 6
1.3 System Overview 7
1.4 Mobility Constraints in Adaptive Coding and Modulation 11
1.5 Outline of this Thesis 14

2 Theoretical Spectral Efficiency Bounds for ACM with Outage Probability Constraints **17**
2.1 Introduction 17
2.2 Problem Formulation 18
2.3 Maximum Average Spectral Efficiency for Adaptive Coded Modulation 18
2.4 Optimal CSNR Thresholds and Optimal MASA 22
2.5 Concluding Remarks 26

3 ACM Design under Imperfect Channel State Information **29**
3.1 Introduction 29
3.2 MAP-optimal Estimation with PSAM 30
3.3 Predictor Performance 34
3.4 Obtaining the Switching Thresholds: Problem Formulation 38
3.5 Obtaining the Switching Thresholds: The Solution 39
3.6 An Upper Bound on the Probability of Codec Mismatch . . 43
3.7 An Example System and Optimal Thresholds 48
3.8 Concluding Remarks 51

4 MASA for the Case of Imperfect Channel Knowledge **55**
4.1 Introduction 55
4.2 System Model and Problem Formulation 55
4.3 A Theoretical Upper Bound On the MASA 56
4.4 Results and Discussion 58
4.5 Concluding Remarks 61

5 Simulation of Adaptive Gallager Coded Modulation on Correlated Fading Channels **65**
5.1 Introduction 65
5.2 System Model 66
5.3 Theoretical ASE Performance - Selecting Codecs for an ACM Scheme 70
5.4 Predictor Performance 74
5.5 Codec Selection Strategies 76
5.6 ACM Performance 77
5.7 Concluding Remarks 86

6 Conclusions and Further Research **89**
6.1 Contributions 89
6.2 Conclusions 91
6.3 Suggested Topics for Further Research 92

A Special Functions **95**

B Statistics and Probability Density Functions **99**
B.1 Rayleigh distribution 99
B.2 Exponential distribution 99
B.3 Bivariate Exponential Distribution 100
B.4 Nakagami distribution 102

C Simulating Correlated Fading **103**
C.1 Filtering Gaussian Random Variables 103
C.2 A Note on the Generation of Random Numbers 104

D Gallager based codecs **107**
D.1 Introduction 107
D.2 System Model 107
D.3 Modulation and Demodulation 108
D.4 The Concept and Properties of Gallager Codes 111
D.5 Example Codecs 114

References **119**

Figures

1.1 A classic transmission system with coding and decoding of binary information, symbol mapping and demapping, and transmit and receive filters. 3
1.2 Baseband model of a frequency-flat multipath fading channel. 7
1.3 Transmission system utilizing an ACM scheme with N codecs, and using channel prediction to obtain the necessary information for choosing the appropriate codec to use for transmission. 8
1.4 BER-versus-CSNR relationship and corresponding CSNR thresholds for N codecs employed by an ACM scheme. 9
1.5 Envelope correlation versus normalized time lag. 13
1.6 Maximum velocity of the relative transmitter-receiver motion as a function of demands on envelope correlation and block length, when carrier frequency $f_c = 5.4$ GHz, and transmission time of one channel symbol is $T_s = 4\mu$s. 15

2.1 MASA plotted as a function of γ_1 for $N \in \{1, 2, 4\}$ and $\overline{\gamma} \in \{10, 20\}$ dB. 24
2.2 MASA plotted as a function of $\mathrm{P_{out}}$ for $N \in \{1, 2, 4\}$ and $\overline{\gamma} \in \{10, 20\}$ dB. 25
2.3 Unconstrained MASA of an ACM system on a Rayleigh fading for $N \in \{1, 2, 4, 8, 16\}$ codecs. 26
2.4 The percent wise reduction in SE, in terms of difference between MASE and MASA, of an ACM system on a Rayleigh fading for $N \in \{1, 2, 4, 8, 16\}$ codecs. 27
2.5 Outage Probability for $N = \{1, 2, 4\}$ codecs when the optimal CSNR thresholds are employed. 27

3.1 Pilot symbol assisted modulation used in both channel prediction and channel estimation. 31
3.2 The conditional PDF for γ conditioned on $\widehat{\gamma}$, for $\widehat{\gamma} = 10$ dB. . . 37

3.3 The expectation of γ conditioned on $\widehat{\gamma}$ and ρ plotted as function of $\overline{\gamma}$, for $\widehat{\gamma} = 10$ dB. 37
3.4 Plot of γ_n versus s_n for $\overline{\gamma} = 10$ dB, $\rho = 0.95$, and different values of ϵ. 42
3.5 Plot of γ_n versus s_n for $\overline{\gamma} = 10$ dB, $\epsilon = 2 \cdot 10^{-3}$, $K = 500$, $L = 10$, $j = 200$ symbols, and different values of jf_mT_s. 43
3.6 Plot of γ_n versus s_n for $\epsilon = 2 \cdot 10^{-3}$, $K = 500$, $L = 10$, $j = 200$ symbols, $jf_mT_s = 7.9 \cdot 10^{-2}$, and different values of $\overline{\gamma}$. . . 44
3.7 CSNR, temporal, and switching threshold for codec n for a practical codec. The two curves show the BER-versus-CSNR performance of a typical practical codec and for a theoretical codec achieving capacity at γ_n. 45
3.8 Plot of the difference $s_n - \gamma_n$ versus $\text{BER}_{n,\tau}$ for the $N = 6$ Gallager codecs. 49
3.9 Plot of the difference $s_n - \gamma_n$ versus ϵ_n for the $N = 6$ Gallager codecs. 50
3.10 Optimal values for $\text{BER}_{n,\tau}$ and ϵ_n and the corresponding (minimum) distance $s_n - \gamma_n$ plotted for $n \in \{1, 2, \ldots, 6\}$ and $\text{BER}_0 = 10^{-3}$ for the Gallager codecs. 51
3.11 Optimal values for $\text{BER}_{n,\tau}$ and ϵ_n and the corresponding (minimum) distance $s_n - \gamma_n$ plotted for $n \in \{1, 2, \ldots, 6\}$ and $\text{BER}_0 = 10^{-4}$ for the Gallager codecs. 52

4.1 MASA(ϵ) plotted as a function of P_{out} for $N \in \{1, 2, 4\}$, $\overline{\gamma} \in \{10, 20\}$ dB, $\rho = 0.99$, and $\epsilon = 2 \cdot 10^{-3}$. 58
4.2 MASA(ϵ) plotted as a function of s_1 for $N \in \{1, 2, 4\}$, $\overline{\gamma} \in \{10, 20\}$ dB, $\rho = 0.99$, and $\epsilon = 2 \cdot 10^{-3}$. 59
4.3 MASA(ϵ) plotted as a function of P_{out} for $N \in \{1, 2, 4\}$, $\overline{\gamma} \in \{10, 20\}$ dB, $\rho = 0.90$, and $\epsilon = 2 \cdot 10^{-3}$. 60
4.4 The optimal values for s_1 and γ_1 plotted as a function of $\overline{\gamma}$ for $N = 4$ and different values of the normalized prediction lag. . 61
4.5 Resulting SE for all codecs resulting from the optimal CSNR thresholds plotted as a function of $\overline{\gamma}$ for $N = 4$ and different values of the normalized prediction lag. 62
4.6 MASA(ϵ) for different values of the normalized prediction lag plotted against average CSNR, for $\epsilon = 2 \cdot 10^{-3}$. 62

5.1 Simulation system. 67
5.2 Theoretical ASE on a Rayleigh MFC when perfect prediction is assumed. The first N codecs in Table D.1 are employed in order to obtain the ASE curves for $N \in \{1, 2, \ldots, 6\}$. 71

5.3 The upper plot shows the gained ASE by using permutations of the six available codecs other than the first $N = 5$ codecs. The lower plot shows the outage probability when the first and the second of the available codecs govern the outage. 72

5.4 The ASE obtained by using the first five codecs available, the MASE, and the MASA for $N = 5$ and $\rho = 1$ plotted against $\overline{\gamma}$. The reduced ASE due to a pilot spacing of $L = 10$ symbols and the SE of the fifth codec R_5 are also plotted. 73

5.5 Scatter plots of the predicted CSNR versus (in the upper plots) actual CSNR and (in the lower plots) averaged actual CSNR, for varying normalized prediction lags. 74

5.6 Simulation results for the normalized correlation ρ between predicted CSNR and actual CSNR are plotted against average expected CSNR and normalized prediction lag. 75

5.7 Switching thresholds for the codecs used in the simulations; a) CSNR thresholds used as switching thresholds, b-c) CSNR thresholds plus a constant (in dB) used as switching thresholds; d-f) switching thresholds found for $\overline{\gamma} \in \{5, 15, 25\}$ dB using $\epsilon = \text{BER}_0$. 76

5.8 Switching thresholds for the codecs used in the simulations; a) CSNR thresholds used as switching thresholds, b-c) CSNR thresholds plus a constant (in dB) used as switching thresholds; d-f) switching thresholds found using $\epsilon = 200 \cdot \text{BER}_0$ with $\overline{\gamma} \in \{5, 15, 25\}$ dB. 78

5.9 The average BER of CSS 1 plotted for $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$ and $\overline{\gamma} \in [1, 25]$ dB. The circle-markers show simulation points with average BER equal to the target BER. 79

5.10 The ASE of CSS 1 plotted for $\overline{\gamma} \in [1, 25]$ dB and $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$. Dash-dot lines indicate that the target BER is not attained while solid lines show the region where the average BER is below BER_0. 80

5.11 The outage probability of CSS 1 plotted for $\overline{\gamma} \in [1, 25]$ dB and $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$. Dash-dot lines indicate that the target BER is not attained while solid lines show the region where the average BER is below BER_0. 81

5.12 The average BER of CSS 2 plotted for $\delta_s = 0.5$ dB, average CSNRs $\overline{\gamma} \in [1, 25]$ dB, and $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$. The circle-markers show simulation points with average BER equal to the target BER. 82

5.13 The average BER of CSS 2 plotted for $\delta_s = 1.0$ dB, average CSNRs $\overline{\gamma} \in [1, 25]$ dB, and $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$. The circle-markers show simulation points with average BER equal to the target BER. 83
5.14 Plots of the ASE versus average CSNR for $\overline{\gamma} \in [1, 25]$ dB and $jf_mT_s = 7.9 \cdot 10^{-2}$. Solid lines indicate that the average BER is less than BER_0. In a) the ASE of CSS 1, CSS 2, and CSS 3 is plotted for different values of δ_s and ϵ. In b) the difference in the ASE of CSS 3 with $\epsilon = 200 \cdot \text{BER}_0$ and CSS 2 with $\delta_s = 1.0$ dB is plotted, and here the dotted line indicate the area where CSS 2 with $\delta_s = 1.0$ dB does not satisfy the target BER. 84
5.15 Plots of the outage probability versus average CSNR for all three strategies with different values of δ_s and ϵ, $\overline{\gamma} \in [1, 25]$ dB, and $jf_mT_s = 7.9 \cdot 10^{-2}$. Solid lines indicate that the average BER is less than BER_0. 85
5.16 Plots of the average BER versus average CSNR for $\overline{\gamma} \in [1, 25]$ dB and $jf_mT_s = 7.9 \cdot 10^{-2}$. Solid lines indicate that the average BER is less than BER_0. In a) the average BER of CSS 1, CSS 2, and CSS 3 is plotted for different values of δ_s and ϵ. In b) the average BER-versus-CSNR plot is zoomed in on CSS 2 with $\delta_s = 1.0$ dB, and CSS 3 with $\epsilon = 200 \cdot \text{BER}_0$ for $\overline{\gamma} \in [7.5, 22.5]$ dB. 86

C.1 Smith's method for generating K samples of a fading channel. 104

D.1 System model for Gallager coded information transmitted on an AWGN channel using complex valued channel symbols. . 107
D.2 16 QAM constellation with Gray mapped binary sequences. . 109
D.3 Example of a bipartite graph built from the parity check matrix. Message passing is done in both directions of an edge. 114
D.4 Simulated and approximated BER-versus-CSNR for the Gallager based codecs. 116

Tables

3.1 Codec indices, code rates, SEs, and constellations sizes and types for the 6 example codecs employing Gallager coding. 48

5.1 Simulation parameters. 70

A.1 Approximate solutions to a = $g_a(b, \epsilon)$. 97
A.2 Approximate solutions to b = $g_b(a, \epsilon)$. 98

D.1 Simulated Gallager codes with block length $M = 200$ channel symbols. 115
D.2 Values of the parameters used in the estimation of the BER-versus-CSNR relationship, and the CSNR thresholds for $\text{BER}_0 \in \{10^{-3}, 10^{-4}\}$. 117
D.3 CSNR thresholds for the $N = 6$ codecs for varying block lengths M. 117

Nomenclature

Definitions of a probability density function (PDF) are written as

$$f_{X_1,X_2,\ldots X_N}(x_1, x_2, \ldots x_N)$$

where X_n denotes a random variable (RV) and x_n is the realization of the RV (this notation rule is not followed for Greek letters). Subscripts I and Q are used to represent the real and the imaginary parts (in-phase and quadrature components) of a complex variable: $X = X_I + jX_Q$. In addition to the meaning in the table below, letters a, b, x, and z and the Greek letters α, β, μ, λ, and γ are used as both arguments and parameters in the definitions of different functions and PDFs in Appendices A and B.

Function arguments indicate whether the function is continuous or discrete; the continuous (time) variables t and τ are used for continuous functions, while discrete functions have a discrete time index as argument (e.g. k, l, and m). Also, in some cases the variable x is used as the integration variable, and then indicating a continuous function. Vectors are written in bold lowercase, matrices in bold uppercase. An exception is made for this rule in Appendix C where bold uppercase letters are used for vectors consisting of filter coefficients and random numbers that are separated in frequency rather than time. If not otherwise defined, vectors are column vectors.

$\lvert x \rvert$	Absolute value
x^*	Complex conjugate
$(\cdot)^{-1}$	1. Matrix inversion: $\mathbf{X}^{-1}$ 2. Function inversion: $f^{-1}(x)$
$[\mathbf{x}]_k$	Element number k in vector $\mathbf{x}$
$[\mathbf{X}]_{km}$	Element in row k and column m in matrix $\mathbf{X}$
$\mathbf{x}^T$	Transpose

$\{x_k\}_{k=1}^{K}$	The sequence of numbers x_k: $\{x_k\}_{k=1}^{K} = \{x_1, x_2, \ldots, x_K\}$
$\frac{\delta}{\delta x_k}$	Coordinate differentiation with respect to x_k
$\nabla_{\{x_k\}_{k=1}^{K}}$	Coordinate differentiation with respect to all coordinates (variables) in $\{x_k\}_{k=1}^{K}$
$\mathbf{X} \circ \mathbf{Y}$	The Hadamard product: componentwise multiplication of two matrices having the same dimension
$\alpha_{i,k}$	Bit number i in the tuple corresponing to the kth channel symbol in the signal constellation
$\gamma, \gamma(t)$	Instantaneous received CSNR at time t
γ_n	CSNR threshold for codec n, i.e. the lowest CSNR attaining the target BER when codec n is used
$\gamma_{n,\tau}$	Temporal threshold for codec n
$\overline{\gamma}$	Expected CSNR: $\overline{\gamma} = \mathrm{E}[\gamma]$
$\overline{\gamma}_{M'}(k)$	Empirical block-wise average of the CSNR for an entire block of M' channel symbols
$\widehat{\gamma}, \widehat{\gamma}(t), \widehat{\gamma}(k)$	Predicted/estimated CSNR
$\Gamma(\cdot)$	Gamma function
$\Gamma(\cdot, \cdot)$	Complementary incomplete gamma function
δ_s	Constant used to obtain the switching thresholds (in dB) as a linear function of the CSNR thresholds (in dB)
Δt	Time separation [s]
Δf	Frequancy separation [Hz]
ϵ, ϵ_n	Probability of selecting a codec that does not attain the target BER (and index n is used when ϵ is codec dependent)
ζ_k	Channel symbol k in a signal constellation
κ	Size of the tuple (sequence of bits) associated with each channel symbol in a signal constellation

λ	Lagrange multiplier
λ_c	Carrier wave length [m]
μ_X	Mean or expectation of the random variable X: $\mu_x = \mathrm{E}[X]$
ρ	Correlation coefficient
ρ_0	Sufficient correlation between samples of the magnitude of the fading envelope on a frequency-flat fading channel
σ	Delay spread [s]
σ_w^2	Variance of additive white Gaussian Noise
σ_X^2	Variance of the random variable X
τ	Time lag
$\phi_{zz}(\tau)$	Autocorrelation of the complex fading envelope $z(t)$ at two time instances separatet by the time lag τ
$\Phi_{zz}(f)$	The Fourier transform of $\phi_{zz}(\tau)$
$\Phi(\alpha, \gamma; z)$	The degenerate hypergeometric function (also commonly denoted ${}_1F_1\left(\alpha; \gamma; z\right)$)
Ω_p	Variance of the fading, (also referred to as the average power gain)
$\widehat{\Omega}_p$	Average power gain for the predicted/estimated fading samples
$\Psi_s(\gamma_n, \epsilon)$,$\Psi_s(\gamma_n)$	Function relating switching threshold s_n to CSNR threshold γ_n through ϵ (ϵ is ommited when chosen as a constant)
$\Psi_\gamma(s_n, \epsilon)$,$\Psi_s(\gamma_n)$	Function relating CSNR threshold γ_n to switching threshold s_n through ϵ (ϵ is ommited when chosen as a constant)
a_n	Constant used in the approximation of BER-versus-CSNR for Gallager codes
a_p	Pilot symbol amplitude value

$\mathbf{A}$	Vector of samples from real-valued zero-mean Gassian random process with variance σ_w^2
ASE, ASE$(\overline{\gamma})$	Average Spectral Efficiency of an ACM scheme at expected CSNR $\overline{\gamma}$
ASE_L	Average Spectral Efficiency of an ACM scheme as a function of the pilot spacing L
b_n	Constant used in the approximation of BER-versus-CSNR for Gallager codes
B	One-sided transmission bandwidth [Hz]
BER, BER(γ)	Bit error rate at CSNR γ
BER_0	Target BER
$\text{BER}_n(\gamma)$	BER of codec n at CSNR γ
$\text{BER}_{n,\tau}$	BER of codec n at the temporal threshold $\gamma_{n,\tau}$
$\overline{\text{BER}}_n(\gamma)$	Average BER of codec n a function of CSNR γ
$\text{BER}(\gamma\|\widehat{\gamma})$	BER of codec n as a function of γ and conditioned on $\widehat{\gamma}$
$\mathbf{B}$	Vector of samples from real-valued zero-mean Gassian random process with variance σ_w^2
c	Speed of light [m/s]
c_n	1. Constant used in the approximation of BER-versus-CSNR for Gallager codes 2. Check node n
$\mathbf{c}$	Syndrome vector
C	Channel capacity [bits/s/Hz]
C_{AWGN}	Channel capacity (of an AWGN channel) [bits/s]
$\text{Cov}(x, y)$	Covariance operator
$\mathbf{C}$	Matrix representing redundant (parity) information in a linear block code
$d(k)$	Information carrying channel symbol

d_n	Constant used in the approximation of BER-versus-CSNR for Gallager codes
$\tilde{d}(k)$	Received channel symbols carrying information
$\widehat{d}(k)$	Received channel symbols carrying information after detection
$\mathbf{d}$	Vector of M channel symbols output from the modulator
$\widehat{\mathbf{d}}$	Vector holding M complex channel symbols after detection
$\mathcal{D}(\cdot,\cdot)$	Euclidean distance between two complex channel symbols
$E_i(x)$	The exponential integral function for negative arguments
$\mathrm{E}[\cdot]$	Expectation operator
f	Frecuency [Hz]
f_c	Carrier frequency [Hz]
f_m	Maximum Doppler frequency [Hz]
$f_X(x)$	PDF for the random variable X
$f_{X_1,X_2}(x_1,x_2)$	Bivariate PDF
$f_{X_1\|X_2}(x_1,x_2)$	Conditional PDF
$F_X(x)$	CDF for the random variable X
${}_1F_1\left(\alpha;\gamma;z\right)$	See $\Phi(\alpha,\gamma;z)$
$\mathbf{F}$	Vector holding filter coefficients used in the simulation of the Rayleigh fading channel
$\mathcal{F}[\cdot]$	The Fourier transform
g, g_n	Number of information bits in one information word input to the channel coder (in codec n)
$g(\cdot,\cdot)$	Function producing a metric based on the Euclidean distance between two complex channel symbols

$\mathbf{G}$	Generator matrix
$h_j(k)$	Filter coefficient
$\mathbf{h}_j$	Vector of filter coefficients
$\mathbf{h}_{j,\mathrm{MAP}}$	Maximum a posteriori optimal filter coefficients
$\mathrm{H}(.,.)$	The H-function defined by the author to simplify notation
$\mathbf{H}$	Parity check matrix
i	Discrete index
$\mathbf{I}_{K\times K}$	Identity matrix with K rows and columns
$I_n(\cdot)$	nth-order Bessel function of the first kind
I_N	Maximum number of iterations in an iterative decoder
j	1. Prediction delay (lag) 2. The complex unit ($j = \sqrt{-1}$) 3. Discrete index
$J_n(\cdot)$	nth-order modified Bessel function of the first kind
k	Discrete (time) index
K	1. Filter length 2. Length of vectors in the simulation model of the Rayleigh fading channel
K_e	Length of estimation filter
K_p	Length of prediction filter
l	Discrete (time) index
L	Pilot spacing ($L-1$ is the number of information carrying channel symbols between two subsequent pilot symbols)
m	1. Nakagami parameter 2. Discrete (time) index
M	Length of a channel symbol block [symbols]

M'	Length of a channel symbol block including deterministic pilots [symbols]
$\mathrm{M}_{\lambda,\mu}(\cdot)$	The Whittaker function
MASE, MASE(γ)	Maximum average spectral efficiency at CSNR γ
MASA, MASA(ϵ)	Maximum ASE for an ACM scheme for a probability of choosing wrong code ϵ
$\max(\cdot)$	The maximum operator, i.e. the input value that maximizes the argument
mod	The modulo operator
$\min(\cdot)$	The minimum operator, i.e. the input value that minimizes the argument
n	Discrete index
$n(t), n(k)$	Complex valued additive white Gaussian noise at time t (or k)
$\mathbf{n}$	Vector holding M' complex AWGN samples
N	Number of codecs in an ACM scheme
N_0	One-sided noise power spectral density [W/Hz]
$\mathcal{N}(\mu, \sigma_w)$	A random process with a Gaussian PDF with mean μ and variance σ_w^2
p_n^0, p_n^1	Probability of the decision of bit n after the demodulation being wrong (superscript 0) or correct (superscript 1)
$p(k)$	Deterministic pilot symbol (pilot)
$\tilde{p}(k)$	Received pilot symbols
$\mathbf{p}$	Pilot vector
$\widehat{\mathbf{p}}$	Received pilot vector
$\mathbf{p}^0$	Vector holding probabilities of the decisions in $\widehat{\mathbf{t}}$ being incorrect
$\mathbf{p}^1$	Vector holding probabilities of the decisions in $\widehat{\mathbf{t}}$ being correct

P	Average transmit power [W]
P_n	Probability of codec n being selected for transmission
P_0	Probability of outage
$\mathrm{P}(\cdot)$	Probability of event
$\mathrm{P}_{\mathrm{out}}$	Outage probability when a demand on outage probability is used
q, q_n	Number of coded bits in a code word from the channel coder (in codec n)
$q_a(b, \epsilon)$	The inverse of the complementary Marcum-Q function with respect to its first argument
$q_b(a, \epsilon)$	The inverse of the complementary Marcum-Q function with respect to its second argument
$Q(a, b)$	The Marcum-Q function
r	1. Ratio between predicted and actual CSNR 2. Rate of a channel code [information bits per code bit]
r_n	Code rate of channel coder in codec n
$\mathbf{r}_{j,K}$	Vector with K elements representing the normalized covariance between the fading to be predicted/estimated and the fading at the pilot symbol time instants
$\mathbf{R}_K$	Normalized covariance matrix
$R(k)$	Sampled version of the normalized autocorrelation function
R_n	Spectral efficiency of codec n [bits/s/Hz]
s_n	Switching threshold for codec n
$\mathbf{s}$	Information vector input to a channel coder
$\widehat{\mathbf{s}}$	Vector holding the decoded information
S, S_n	Constellation size (of codec n)

$S_{zz}(f)$	Power density spectrum of the fading process (Doppler power spectrum)
t	Continuous time variable
$\mathbf{t}$	Vector holding coded information output from a channel coder
$\tilde{\mathbf{t}}$	Most likely codeword after iterative decoding of a Gallager code
$\widehat{\mathbf{t}}$	Vector holding hard decision on the coded information output from a demodulator
T	Time period used to transmit a block of M channel symbols [s]
T_s	Time period used to transmit a single channel symbols [s]
u	Difference between length of a codeword and the corresponding information sequence
v	Relative transmitter-receiver velocity [m/s]
$v_{\max}$	Maximum relative transmitter-receiver velocity when a demand on the correlation between samples of the magnitude of the fading envelope is employed [m/s]
$\mathrm{Var}(x)$	Variance operator
w_k	Noise node k
$\mathbf{w}$	Binary column vector representing additive noise
$\widehat{\mathbf{w}}$	The most likely binary noise vector after iterative decoding of a Gallager code
$x, x(t), x(k)$	Complex valued transmitted channel symbol
$\mathbf{x}$	Vector holding M' transmitted complex channel symbols
$\widehat{\mathbf{x}}$	Vector holding M' complex channel symbols after detection
$y, y(t), y(k)$	Complex valued received channel symbol

$\mathbf{y}$	Vector holding M' received complex channel symbols
$z, z(t), z(k)$	Complex valued channel gain
$\mathbf{z}$	Vector holding complex fading samples
$\tilde{z}(k)$	Maximum likelihood estimate of the fading
$\widehat{z}(k)$	Prediction/estimate of the fading
$\tilde{\mathbf{z}}$	Vector holding buffered maximum likelihood estimates of the fading
$\mathbf{Z}$	Vector holding the Fourier transformed version of the values in $\mathbf{z}$

Abbreviations

ACM	Adaptive Coding and Modulation
ASE	Average Spectral Efficiency
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CDF	Cumulative Density Function
Codec	An encoder and decoder pair of a channel code and corresponding channel symbol modulation and demodulation
CSI	Channel State Information
CSNR	Carrier Signal-to-Noise Ratio
CSS	Codec Selection Strategy
dB	Decibel
Hz	Hertz
IID	Independent and Identically Distributed
MAP	Maximum a Posteriori
MASA	Maximum Average Spectral efficiency for Adaptive coded modulation
MASE	Maximum Average Spectral Efficiency
MFC	Multipath Fading Channel
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
OFDM	Orthogonal Frequency Division Multiplexing
PDF	Probability Density Function
PSAM	Pilot Symbol Assisted Modulation
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
SE	Spectral Efficiency
WSS	Wide Sense Stationary

Chapter 1

Introduction to Adaptive Coding and Modulation

Any new electronic device which includes sufficiently new or improved technology is often adopted by a high number of users. To become popular, electronic devices must process larger amounts of information at higher information rates to satisfy new and "greedier" applications. The increased mobility among users results in an increased popularity for wireless applications where the information processed by the device can be updated (downloaded) on the fly. Popular devices are also usually very small and with a low weight, and it is commonly required that the battery life-time (between rechargings) is long.

Examples of popular *portable* devices are personal computers, game consoles, music players, and mobile (cellular or satellite) phones. As a result, the radio spectrum is being used by an increasing number of systems and users. The available radio spectrum is a finite resource that is shared between all users. In order to avoid in-band interference, the power consumption should be kept to a minimum (allowing the use of small long-life batteries). Keeping both the bandwidth and power used by communication systems low limits the amount of information that can be communicated. The increasing use of wireless transmission media therefore calls for more spectrally efficient transmission schemes, where *spectral efficiency* (SE) can be defined as the amount of information bits transmitted per time unit per Hertz available bandwidth.

Evidently, SE or information throughput is a measurement of the performance of a wireless transmission scheme. In addition, the amount of information that is received in error may reduce the quality of the communication. For applications such as voice communication relatively large

amounts of errors can be tolerated. For some pure data communication applications, such as downloading a text document to a computer, no errors can be tolerated since one error may affect the entire amount of information transmitted. However, voice communication and other applications, such as streaming media, often require constant transmission of information, thus putting a very strong demand on the so called *outage probability*, which is the probability that no reliable communication is possible. Thus, outage probability is another performance measure for a wireless communication system.

A technique commonly referred to as rate-adaptive modulation or *adaptive coding and modulation* (ACM) can be used to improve the SE in wireless communications. In traditional investigation of ACM schemes an idealized channel model is assumed and the technique promises the user the highest possible *average spectral efficiency* (ASE). In this thesis the practical limitations and possibilities of ACM schemes are investigated by removing some of the idealized assumptions. The remainder of this introductory chapter is organized as follows: First, in Section 1.1 existing transmission techniques that can be used to increase the SE are described. Also, in this section, an overview of some characteristics of a mobile radio environment is given. Subsequently, the problems addressed in this thesis are formulated in Section 1.2. In Section 1.3 an overview of the system model assumed in this thesis is given. Then, in Section 1.4 a first step towards explaining the constraints imposed on an ACM system in the case of mobile wireless transmission is discussed. This section was also presented in [Jetlund, Øien, Hole, Markhus, and Myhre, 2002]. An outline of the remainder of the thesis is given in Section 1.5.

1.1 Spectrally Efficient Transmission Techniques

In this subsection a selection of established techniques that can be used in the design of spectrally efficient communication systems are described. The "classical" transmission system in Figure 1.1 shows how (possibly non-redundant) information is coded into a redundant bit stream. The coded information is subsequently modulated onto the (communication) channel by performing a mapping of coded information and filtering the resulting symbols. The receive filter is usually built as a matched filter taking into account the stationary characteristics of the channel [Barry, Lee, and Messerschmitt, 2004]. After filtering the received information carrying signal, the information is detected (or sliced) into a set of symbols such that all symbols are elements in the alphabet used at the transmitter. Subsequently,

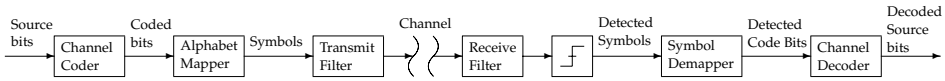


FIGURE 1.1: A classic transmission system with coding and decoding of binary information, symbol mapping and demapping, and transmit and receive filters.

the demapper produces a detected version of the coded information bits which is then decoded producing the decoded source bits. The system in Figure 1.1 outlines the use of a so called "hard" decoder, i.e. decisions are made before decoding. If a "soft" decoder is used the received information from the channel is employed directly in the decoding process.

Non-adaptive Techniques

Spectrally efficient transmission techniques, using modulation constellations with a high number, S, of different channel symbols, and thus many information bits per symbol ($\log_2 S$ in the uncoded case), can be shown to be very effective as long as the *channel signal-to-noise ratio* (CSNR) is sufficiently high and constant. Also, employing error correcting codes may help reduce the amount of errors introduced by the communication channel by adding redundant information to the information signal to be transmitted. The expenses paid using an error correcting code are reduced SE and increased processing delay, complexity, and power consumption in both transmitter and receiver. The error correcting properties of such codes are also dependent on the CSNR. However, fluctuations in the CSNR are among the most severe restrictions in wireless communication systems [Hanzo, Wong, and Yee, 2002].

Characteristics of mobile radio environments

A signal traveling in an environment with both natural and man-made objects is scattered, reflected, and diffracted, resulting in *multipath* transmission, and thus a composite received signal [Stüber, 2001]. Scattering occurs when radio waves reach objects which are smaller than the radio wave length, reflection occurs when radio waves interact with objects which are much larger than the wave length, and diffraction occurs when radio signals are "bent" around an object. In addition the channel varies with shadowing (long-term variations in the radio environment) and path loss (varying distance between transmitter and receiver) [Rappaport, 1996]. Changes

in the environment due to movement of objects, receiver, and transmitter introduce *time-varying fading* on the communication channel. How fast the communication channel—and thus the amplitude and phase of the received signal—vary depends on the velocity of objects, receiver, and transmitter. Depending on the system parameters, a wireless channel can be modeled as a *multipath fading channel* (MFC) with either frequency-selective or frequency-flat fading distribution.

When the *delay spread* [Greenwood and Hanzo, 1999] of the channel is short compared to the symbol duration, the channel is said to be frequency-flat, i.e. all frequency components of a transmitted signal are affected (attenuated or amplified) in the same way by the channel. In other words, the *coherence bandwidth* is large compared to the signal bandwidth [Stüber, 2001; Greenwood and Hanzo, 1999]. By employing a multi-carrier technique like *orthogonal frequency division multiplexing* (OFDM), and when the channel in question is a sub-channel in the OFDM scheme, the fading can be assumed to be frequency-flat [Stüber, 2001]. In addition to being frequency-flat it is assumed in this thesis that the duration of a transmitted channel symbol is much shorter than the *coherence time* of the channel, i.e. a slowly varying fading channel is assumed.

Frequency-flat fading is usually modeled by a complex channel gain with a certain probability distribution. In addition the signal is typically disturbed by *additive white Gaussian noise* (AWGN). If a line-of-sight (LoS) component is present between transmitter and receiver, the received signal is said to exhibit *Ricean* fading [Stüber, 2001]. *Rayleigh* fading is often used to describe a radio environment with no such component [Stüber, 2001]. The *Nakagami m*-distribution [Nakagami, 1960] can approximate both the Rayleigh and the Rice distribution (e.g. when choosing the Nakagami parameter $m = 1$ the distribution becomes Rayleigh). The distribution is often used since it in many cases is a closer fit to empirical data [Stüber, 2001], and is also easier to manipulate analytically than the Ricean distribution.

Due to these variations in the CSNR, any combination of *channel code and modulation constellation* (codec)[1] can only guarantee reliable transmission, in terms of *bit error rate* (BER), for a fraction of the time. In the remaining time period the system experiences outage. By reducing the SE of a codec it can guarantee a low BER also for lower CSNR values. Thus, reducing the probability of outage reduces the possible information throughput of a wireless communication system.

[1]The term "codec" usually refers to the encoder and decoder pair of a specific channel code, but in this thesis the term also includes the corresponding modulation and demodulation of coded information.

Adaptive Techniques

A low outage probability, high SE, and low BER can not be achieved simultaneously by a traditional system using a single codec. Hayes suggested in 1968 an "Adaptive Feedback Communications" system to combat the variations introduced by a MFC with a Rayleigh distributed envelope [Hayes, 1968]. The technique is based on some parameter(s) in the communication system being dependent on the state of the channel. This state is communicated from receiver to transmitter on a separate feedback channel. Then the transmitter can adapt the transmission according to the variations in the CSNR. In [Hayes, 1968] a *power-adaptation* scheme was proposed. The disadvantages of this scheme are an increased average transmit power, increased co-channel interference, and strict requirements on the linearity of amplifiers used by the transmitter and the receiver. The increased average transmit power requires larger batteries and result in unpredictable battery life-time.

Another approach is to adjust the *data rate* of the communication system. In [Cavers, 1972] this was done by adjusting the duration of a single channel symbol on the channel. Evidently, this will result in a variable bandwidth. In order to use such a scheme a relatively wide frequency band (compared to the average bandwidth usage) must be assigned, and thus, reducing the overall SE in a multiuser environment. Varying the size of the *modulation constellation* used ("adaptive modulation") according to variations in CSNR [Steele and Webb, 1991] is more attractive. In adaptive modulation, a small constellation is used when the CSNR is low and a large constellation when the CSNR is high, leading to increased ASE.

The motivation of adaptive schemes is in general to be able to transmit with an ASE as close to the capacity as possible, at a BER which fulfills the desired quality requirements. The channel capacity, or the *maximum average spectral efficiency* (MASE), for a communication system utilizing perfect *channel state information* (CSI) was presented in [Goldsmith and Varaiya, 1997]. An optimal variable-power and variable-rate transmission scheme can reach the channel capacity under idealized conditions [Goldsmith and Varaiya, 1997; Alouini and Goldsmith, 2000]. In [Goldsmith and Varaiya, 1997] it was also shown that a variable-rate and constant power scheme has a performance close to the capacity under the same idealized conditions. That is, only a small fraction of the possible ASE is traded away by keeping the transmit power constant.

In [Chua and Goldsmith, 1997; Goldsmith and Chua, 1997] adaptive modulation was combined with channel coding leading to so-called ACM. In ACM, not only does a channel symbol from a large signal constella-

tion carry more code bits compared to a small signal constellation, but also the amount of redundant information in the transmitted information stream is varied by adjusting the code rate of the channel coder according to variations in the communication channel. Further investigation of ACM schemes on flat fading channels using coded modulation, with multidimensional trellis codes, was presented in [Hole, Holm, and Øien, 2000]. Adaptive systems can also be designed in conjunction with other techniques such as *multiple-input multiple-output* (MIMO) systems [Zhou and Giannakis, 2004].

An adaptive coding and/or modulation scheme also rely on the CSI communicated from receiver to transmitter on a separate feedback channel. The CSI itself must be obtained through prediction of future values of one or several channel metrics. *Pilot symbol assisted modulation* (PSAM) is a well established technique in which deterministic channel symbols (pilot symbols) are multiplexed into the transmitted stream of information carrying channel symbols [Cavers, 1991]. PSAM or *data aided channel estimation* outperforms more traditional systems employing pilot tones since these are more complex, sensitive to frequency shifts, use more bandwidth, and require highly linear amplifiers in the transmitter [Meyr, Moeneclaey, and Fechtel, 1998]. Introducing pilot symbols into the stream of transmitted information reduces the theoretically available SE somewhat since a fraction of the channel symbols transmitted does not carry information. However, using the pilot symbols help providing information on the state of the channel that can be used to improve symbol detection, channel estimation, channel prediction, and ultimately increase the practically available SE [Meyr et al., 1998; Baltersee, Fock, and Meyr, 2001].

1.2 Problem Formulation

Traditional theoretical analysis of adaptive communication systems assume perfect knowledge of the CSI and a return channel that is error-free and with a zero delay (see e.g. [Hayes, 1968; Goldsmith and Varaiya, 1997; Alouini and Goldsmith, 2000; Hole et al., 2000]). In [Goeckel, 1999] an adaptive coding scheme using outdated fading estimates was considered, and it was shown that information on the fading process should be included in the system design since variations in the channel characteristics highly affects the BER performance of an ACM scheme (similar conclusions were made in [Alouini and Goldsmith, 1997a]).

The objective of this thesis is to explore some limitations and advantages of an ACM scheme after removing the following idealized conditions:

- Perfect prediction.
- Infinitely many codecs.
- Arbitrary length channel codes in the component codecs.
- Zero-delay return channel.
- Constant fading between two subsequent codec updates.

In particular, we shall take into account the prediction errors resulting from using a practical channel predictor. An imperfect channel estimator will also be incorporated in the detection and demodulation of received information. The predictor performance varies with delays introduced by the system; return channel delay, communication channel delay, and delays introduced by processing in the transmitter and receiver. Thus, the predictor used is investigated for various amounts of delay as well as for different degrees of mobility in the communication channel.

1.3 System Overview

The Communication Channel

The communication channel under consideration in this thesis is a time-varying *wide-sense stationary* (WSS),[2] frequency-flat, slow[3] MFC. In the com-

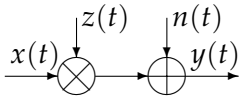


FIGURE 1.2: Baseband model of a frequency-flat multipath fading channel.

plex baseband model of the MFC in Figure 1.2 the received signal can be written as

$$y(t) = z(t) \cdot x(t) + n(t), \tag{1.1}$$

where $x(t)$ is the transmitted complex-valued symbol, $n(t)$ is complex-valued AWGN, and $z(t)$ is the complex fading gain. The instantaneous

[2]The statistical properties; expectation, autocorrelation, and cross correlation, of a wide-sense stationary process do not depend on the observation time.

[3]"Slow" refers to time.

and average CSNR are defined as

$$\gamma = \gamma(t) = \frac{|z(t)|^2 \cdot P}{N_0 B} \tag{1.2}$$

and

$$\overline{\gamma} = \mathrm{E}[\gamma(t)] = \frac{\Omega_p \cdot P}{N_0 B} \tag{1.3}$$

respectively, where P [W] is the average transmit power, N_0 [W/Hz] is the one-sided noise power spectral density, B [Hz] is the one-sided transmission bandwidth, and $\Omega_p = \mathrm{E}[|z(t)|^2]$ is the average power gain [Stüber, 2001].

Experimental and analytical results in this thesis are obtained assuming Rayleigh fading. That is, it is assumed that the received complex fading envelope $|z(t)|$ has a Rayleigh distribution. Then, the CSNR has an exponential distribution with expectation $\overline{\gamma}$ [Stüber, 2001]:

$$f_\gamma(\gamma) = \frac{1}{\overline{\gamma}} e^{-\frac{\gamma}{\overline{\gamma}}}. \tag{1.4}$$

Rate-adaptive Coding and Modulation: System Description

The ACM system under study in this thesis is shown in Figure 1.3. The system consists of a transmitter and receiver employing a set of N component codecs (each with a different SE), a communication channel, and a separate return channel which is assumed free of errors, but with a non-zero delay. The codecs have different SEs and are used to transmit information within

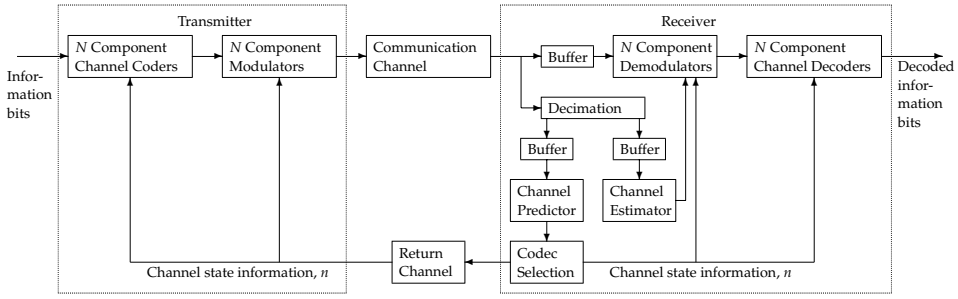


FIGURE 1.3: Transmission system utilizing an ACM scheme with N codecs, and using channel prediction to obtain the necessary information for choosing the appropriate codec to use for transmission.

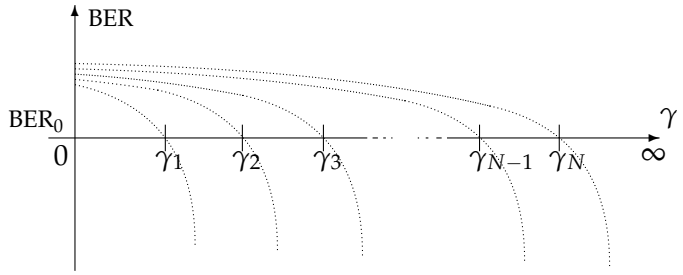


FIGURE 1.4: BER-versus-CSNR relationship and corresponding CSNR thresholds for N codecs employed by an ACM scheme.

different CSNR ranges on the channel. If N is large the CSNR range each codec operates in is commonly approximated by an AWGN channel. Thus, if the BER-versus-CSNR characteristics for each codec on an AWGN channel is known, a CSNR threshold can be defined for each codec, such that it guarantees a BER below the target BER, denoted BER_0, when the CSNR is above the CSNR threshold. The CSI at a future time instant may be obtained by predicting the future CSNR at the receiver. The range of possible CSNR values on the MFC, $\gamma \in [0, \infty\rangle$, is thus divided into $N+1$ CSNR regions by $N+2$ CSNR thresholds, $\{\gamma_n\}_{n=0}^{N+1}$, as outlined in Figure 1.4, with $\gamma_0 = 0$ and $\gamma_{N+1} = \infty$. The CSNR is said to fall into CSNR region n when $\gamma_n \leq \gamma < \gamma_{n+1}$. The CSNR thresholds $\{\gamma_n\}_{n=1}^{N}$ are selected such that the BER of codec n is less than or equal to BER_0 when $\gamma \geq \gamma_n$ on an AWGN channel. It is assumed that $R_1 < R_2 < \cdots < R_N$, where R_n denotes the SE of codec n. Codec n is the codec with highest SE that guarantees a BER below the target BER when $\gamma \in [\gamma_n, \gamma_{n+1}\rangle$. The zeroth region, the *outage* region, is the range $[0, \gamma_1\rangle$ where none of the employed codecs can guarantee a BER below the target BER.

The ASE of the ACM scheme is defined as [Goldsmith and Chua, 1997]

$$\text{ASE} = \sum_{n=1}^{N} R_n \cdot P_n \ \ [\text{bits/s/Hz}], \tag{1.5}$$

where P_n is the probability of codec n being used, and can be found from the probability distribution of the predicted fading [Goldsmith and Chua, 1997]. The scheme will use a low SE, R_n, when the predicted CSNR is low and a high SE when the predicted CSNR is high. A well-designed ACM system will therefore typically outperform a "traditional" system (i.e. a system employing a single-rate codec), since it may allow for transmission,

with reduced SE, at CSNRs below the outage level of the traditional system, and at higher SEs as the CSNR increases above this outage level.

Assuming perfect prediction of future CSNR values and assuming the fading to be approximately constant during a given time interval, the CSNR thresholds can be used as switching thresholds, that is, to select the appropriate codec to be used for transmission. In practice however, the channel prediction is not perfect. The system can be made more robust towards channel prediction errors by increasing the switching thresholds so as to be more conservative in the choice of codecs. However, we do not want to be more conservative than necessary, as this will reduce the ASE more than needed. The trade-off between ASE and the overall BER of an ACM scheme, for a given set of codecs, can be seen from Figure 1.4; increasing the ASE the CSNR thresholds are moved to the left, and on average codecs with higher SE are used at lower CSNRs. Moving the thresholds $\{\gamma_n\}_{n=1}^{N}$ to the left is equivalent of moving the horizontal line showing the intersection between the BER_0 and the BER-versus-CSNR curves, upwards and thus increasing BER_0.

Channel Prediction and Estimation

In Figure 1.3 the predictor operates on a decimated sequence of the sequence of received channel symbols. Likewise a channel estimator aiding the demodulation/detection process operates on a decimated sequence. Unlike the channel predictor (which only uses channel symbols received in the past) the channel estimator may use channel symbols received both in the past and in the future to estimate the fading envelope. In this case all channel symbols must be stored while the system waits for the channel estimation. Thus, the receiver has three buffers; one for holding the two decimated sequences and one storing the received information carrying channel symbols while waiting for the channel estimation process to be finished. The CSI is assumed communicated to the transmitter via the separate feedback channel. The transmitter uses this information to select the codec with the highest SE, among the available codecs, that satisfies the BER demand.

The transmission system considered introduces several delays (processing delays in both transmitter and receiver and transmission delays in both the communication channel and the return channel). As a result, the CSNR value used to select the appropriate codec has to be predicted at the receiver for some time instant in the future. The relevant CSI for the ACM scheme in the previous subsection is the appropriate codec index n: A codec update is performed when this index is changed. To provide the

ACM scheme with information on the fading envelope the system utilizes PSAM [Cavers, 1991]. That is, deterministic channel symbols (pilot symbols) are multiplexed into the stream of information carrying channel symbols to be transmitted. It is assumed that both the receiver and the transmitter knows both the pilot symbols and the pilot symbol instants. For each received pilot symbol a *maximum likelihood* (ML) estimate of the fading envelope can be obtained. The ML-estimates are buffered at the receiver, and filtered through a linear filter with *maximum a posteriori* (MAP) optimal filter coefficients (see e.g. [Øien, Holm, and Hole, 2004]) to find a predicted value of the CSNR at some time instant in the future.

1.4 Mobility Constraints in Adaptive Coding and Modulation

The underlying assumption used when modeling a fading channel by time multiplexed AWGN channels, is that the channel is slowly varying in time [Hole and Øien, 2001]. As described in the first part of this chapter, this is the case when the coherence time of the channel is much larger than the time slot used to transmit a channel symbol.

Coherence time is commonly defined as the reciprocal of the Doppler Spread [Proakis, 2000] which is defined in terms of the Doppler spectrum (the Fourier transform of the autocorrelation of the fading envelope). The Doppler spectrum varies with both frequency and time separation, and the Doppler spread is defined as the time separation in which the Doppler spectrum is essentially nonzero. In other words the coherence time can be defined in terms of the autocorrelation [Proakis, 2000]. The degree of correlation which is sufficient will vary with characteristics of the transmitter and receiver [Greenwood and Hanzo, 1999]. Since the autocorrelation varies with the Doppler frequency, increased mobility in terms of faster relative motion between receiver and transmitter decreases the coherence time of a wireless channel.

In a practical ACM scheme it is not realistic to switch codec after each transmitted channel symbol. Especially, this is the case when the codecs employ block codes. Also, in practice, packetized transmission is used. Thus, it is assumed here that a codec update can only be done after transmitting M consecutive channel symbols. As a result the channel must be slow enough to ensure that the CSNR stays within the same CSNR region during transmission of M channel symbols. That is, the correlation between fading values must be high enough for lags less than or equal to

M. This puts in effect an upper bound on the relative transmitter-receiver velocity.

The period used to transmit a single channel symbol is, for the case of Nyquist transmission, given as $T_s = 1/B$ [s]. The total transmission delay for a codeword of length M is then given as $T = M \cdot T_s$ [s]. The channel variation due to the relative movement of transmitter and receiver manifests itself as Doppler shift [Rappaport, 1996]. The Maximum Doppler shift (or maximum Doppler frequency) is given as

$$f_m = \frac{v}{\lambda_c} \quad [\text{Hz}] \tag{1.6}$$

where v [m/s] is the relative velocity of the receiver-transmitter movement, and λ_c [m] is the wavelength of the arriving plane wave. Since carrier frequency can be written as

$$f_c = \frac{c}{\lambda_c} \tag{1.7}$$

where c [m/s] is the speed of light, this results in the following well-known relationship between maximum Doppler shift and carrier frequency:

$$f_m = \frac{v f_c}{c}. \tag{1.8}$$

Clearly, mobility in a transmission environment changes the state of a wireless channel both with and without obstructing objects in the transmission path.

The criterion for assuming both slowly varying and frequency-flat fading channels is that the correlation between received samples is "high enough." Assuming isotropic scattering, the correlation coefficient for two signal components[4] separated by Δf Hz and Δt seconds is equal to [Greenwood and Hanzo, 1999]

$$\rho(\Delta f, \Delta t) = \frac{J_0^2(2\pi f_m \Delta t)}{1 + (2\pi \Delta f)^2 \sigma^2}\,, \tag{1.9}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind (the nth-order Bessel function is given in Equation (A.2a)) and σ is the delay spread of the channel. In our case we may set $\Delta f = 0$ since we always assume frequency-flat fading [Greenwood and Hanzo, 1999]. The channel can be viewed as approximately constant over each codeword if

$$\rho(0, T) \geq \rho_0 \tag{1.10}$$

[4] "Signal component" in this context refers to the magnitude of the complex fading gain, $|z(t)|$.

where the value of ρ_0 must chosen sufficiently high. Since the parameter ρ_0 controls the variations in the fading envelope it also controls at least three parts of an ACM scheme; channel estimation for channel symbol detection, channel prediction for determining future channel states, and the error correcting properties of the channel codes used. The performance of these three parts will all degrade as ρ_0 decreases.

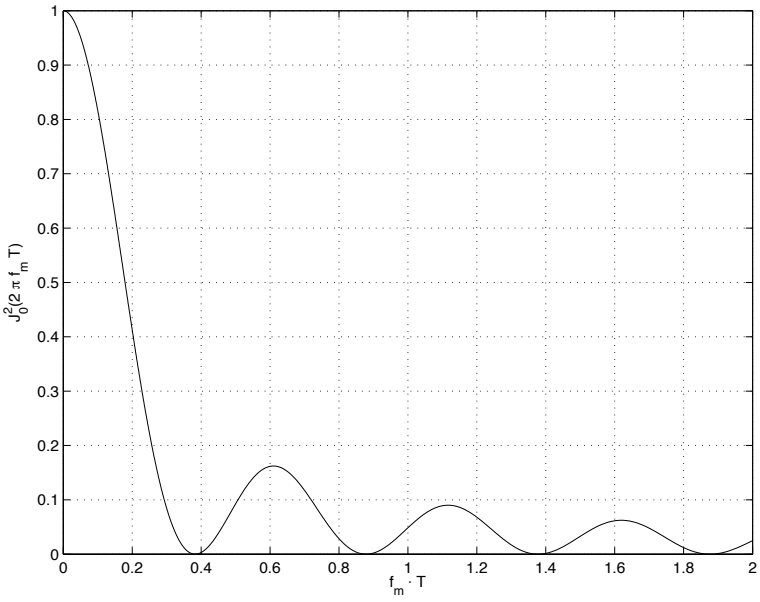


FIGURE 1.5: Envelope correlation versus normalized time lag.

From the assumptions, the only phenomenon that has to be restricted to make sure that the channel does not vary to much over each codeword, is the velocity of relative movement between the transmitter and the receiver. Note, that the assumption of a slowly varying channel should ensure not only that the channel stays within a single fading region during transmission of a sequence or block of channel symbols, but also that the predicted channel quality will be sufficiently accurate.

Using Equations (1.8)–(1.10), the maximum velocity for an almost constant channel quality over a codeword is given, for a given choice of "acceptable" correlation ρ_0, as

$$v \leq v_{\max} = \frac{c \cdot J_0^{-1}(\sqrt{\rho_0})}{T 2\pi f_c}, \tag{1.11}$$

were $J_0^{-1}(\cdot)$ is the inverted zero-order Bessel function of the first kind. The

function is not one-to-one, but from Figure 1.5 it can be seen that the smallest solution of $J_0^{-1}(\cdot)$ is the desired value. A realistic example will show how allowed mobility is reduced with increased block lengths.

Example 1.1 (Maximum velocity under a constraint on the coherence time)
The ETSI standard HIPERLAN/2 [ETSI, 2000] is specified to use carrier frequencies around $f_c = 5.4$ GHz, and a symbol period of $T_s = 4\ \mu$s. If the codeword length is $M = 200$ symbols, and the correlation demand is set to $\rho_0 = 0.99$ the maximum velocity according to Equation (1.11) becomes

$$v_{\max} = 1.56 \text{ m/s (or } 5.62 \text{ km/hr)}$$

which is about walking speed. If the correlation is chosen to be $\rho_0 = 0.90$ the maximum velocity increases to

$$v_{\max} = 5.04 \text{ m/s.}$$

In Figure 1.6 the maximum velocity is plotted against different values of the envelope correlation demand, $\rho_0 \in [0.90, 0.99]$, and for block lengths $M \in [1, 450]$ symbols. The figure shows how increased block lengths and increased correlation demand reduce the upper bound on the velocity of the relative movement of transmitter and receiver.

Note, that the results from this example only indicate that there is an upper bound on the maximum velocity of a wireless (ACM) scheme. The correlation function in Equation (1.9) yields the magnitude of the complex envelope and not the complex fading envelope or the correlation between predicted and actual CSNR. The results showed here can be used as a rule-of-thumb regarding how high the velocity in an ACM scheme can be before the system breaks down. It should also be noted; the choice of ρ_0 depends on the transmission scheme used and will govern the BER and the ASE. It is therefore up to the designer to choose the appropriate ρ_0. One of the main topics in this thesis is how the codecs are selected by the ACM scheme. Modifying strategies for selecting codecs are here done mainly to control the BER. When such a modification reduces the BER, the choice of ρ_0 may be reduced and allowing a larger degree of mobility.

1.5 Outline of this Thesis

The main objective in this thesis is to investigate the practical possibilities and limitations of ACM schemes. We shall do this by means of both simulations and theoretical results. Results presented are utilized when suggesting new techniques that can be used in the design of ACM schemes.

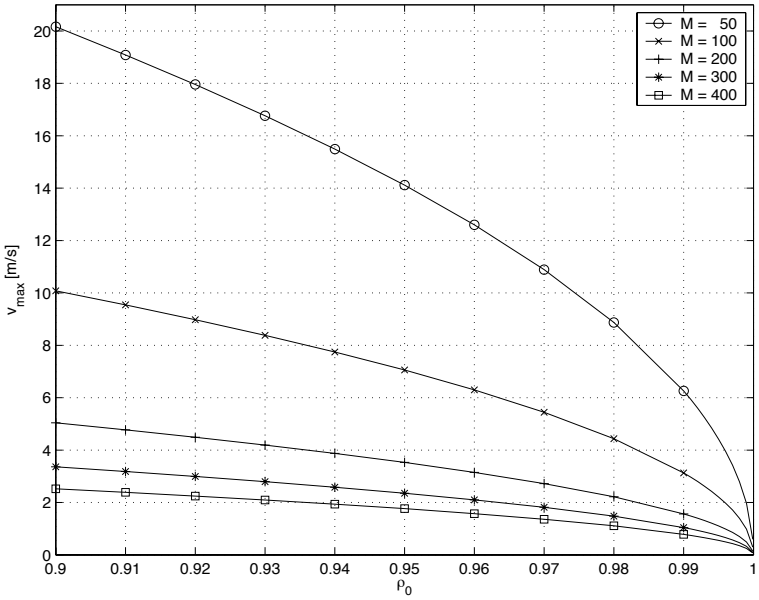


FIGURE 1.6: Maximum velocity of the relative transmitter-receiver motion as a function of demands on envelope correlation and block length, when carrier frequency $f_c = 5.4$ GHz, and transmission time of one channel symbol is $T_s = 4\mu$s.

The theoretical part of the thesis focus on maximizing the ASE under a constraint on the BER. As argued previously, the outage probability is of practical concern in wireless communication systems. Thus, the maximization of the ASE in this thesis therefore includes an upper constraint on the outage probability.

When evaluating an ACM scheme in terms of BER and ASE, imperfect channel prediction should be taken into account since this in fact reduces system performance. We shall design techniques that can be used when maximizing the ASE under both a BER and outage probability constraint also for the case of imperfect channel prediction. By means of simulations, the performance of an example ACM scheme designed under idealized conditions is compared to the performance of an ACM scheme designed using the techniques developed in this thesis.

In Chapter 2, theoretical capacity achieving codecs of arbitrary length, are assumed. This assumption is made to obtain the maximum achievable ASE for an ACM scheme using a limited number of codecs. An existing ultimate upper bound on the ASE of a practical ACM scheme is extended to yield maximum ASE under a constraint on the outage probability.

The specific predictor used in this thesis is investigated in Chapter 3. The investigation includes deriving the expectation of the actual CSNR conditioned on predicted CSNR, average CSNR, and correlation between actual and predicted CSNR. In this chapter a technique that can be used to obtain switching thresholds from CSNR thresholds is outlined. The method is based on controlling the event of codec mismatch, i.e. choosing a codec that does not guarantee the target BER. In this method the probability of codec mismatch is used as a parameter. Then, it is shown that there exists optimal switching thresholds both for the capacity achieving codecs and a set of more practical codecs, in the sense of maximizing ASE under an constraint on the BER.

In Chapter 4 the optimal thresholds are found for an example value of the probability of codec mismatch. This is done both for the case of imperfect channel prediction and under a constraint on the outage probability.

In Chapter 5 simulations of an ACM scheme on Rayleigh fading are presented for different configurations (a theoretically optimal configuration and more practical configurations).

Finally, in Chapter 6 the concluding remarks in this thesis are summed up.

In addition to the content of the thesis outlined above there are four appendices. The first, Appendix A, lists all the special functions used in this thesis, as well as tables of values for two expressions that are used in the numerical calculations in Chapter 4. In Appendix B a set of relevant statistical distributions are listed, and the derivation of the above mentioned expectation of the actual CSNR is shown. The method used for simulating the Rayleigh fading is described in Appendix C. Appendix D describes the component channel codes and the modulation technique used in our codecs.

Chapter 2

Theoretical Spectral Efficiency Bounds for ACM with Outage Probability Constraints

2.1 Introduction

The performance of a rate-adaptive transmission scheme can be measured in terms of BER and ASE. A communication system with a single codec can only achieve an ASE that is equal to the SE of this codec. Likewise, the ASE of an ACM scheme is upper bounded by the codec with highest SE. The ASE can be compared to the channel capacity [Goldsmith and Varaiya, 1997], but since a practical ACM scheme employs a limited number of codecs a more realistic reference is the *maximum ASE for an ACM scheme* (MASA) [Holm, Øien, Alouini, Gesbert, and Hole, 2003]. In [Holm et al., 2003] a method that can be used to obtain the MASA with a limited number of component codecs was described. A drawback of this method in its pure form is that, although the resulting MASA is maximum, it may result in an unacceptable high outage probability. The method is therefore generalized here to optimize the MASA under an outage constraint.

In the following section the problems addressed in this chapter are formulated. In Section 2.3 expressions for the channel capacity and the MASA of an ACM scheme are given. The method in [Holm et al., 2003] is extended such that the MASA can be optimized also under an outage constraint in Section 2.4. This work was also presented in [Jetlund, Øien, Hole, and Holm, 2004a]. Finally, Section 2.5 presents some concluding remarks regarding the work presented in this chapter.

2.2 Problem Formulation

In the case of perfect channel prediction the probability of outage is given as [Choi and Hanzo, 2002]

$$P_0 = \int_0^{\gamma_1} f_\gamma(\gamma) d\gamma, \tag{2.1}$$

where $f_\gamma(\gamma)$ is the *probability density function* (PDF) of the CSNR. Note that P_0 is a function of γ_1. As a result, maximizing the MASA may increase the outage probability above a desirable value chosen by the designer of an ACM scheme. From an information theoretical point of view, maximizing the ASE is maximizing the information throughput over infinitely long time periods. Thus, long periods of outage is allowed as long as the average throughput of the system is maximized. In a practical system where *quality of service* (QoS) depends on e.g. continuous transmission, the outage probability will be of practical concern. It is therefore also desirable to obtain the MASA under an upper constraint on the outage probability.

Obtaining the constrained MASA is the fist objective of this chapter. This is done by assuming that we have perfect channel prediction and that each of the codecs employed perform with BER $= 0$ when the CSNR is above a certain switching threshold. The second objective is to obtain the N thresholds $\{\gamma_n\}_{n=1}^{N}$ that maximize the ASE, with a possible constraint on the outage probability, and the resulting MASA which is the maximum achievable SE for an adaptive scheme employing N codecs. The thresholds are not intended as a direct guideline for the designer, but rather the resulting bounds could be used to evaluate the performance of ACM schemes in terms of ASE. However, the results can be used (as a starting point) to indicate how a practical ACM scheme should be configured. The case of imperfect channel knowledge will be treated in the subsequent chapters of this thesis.

2.3 Maximum Average Spectral Efficiency for Adaptive Coded Modulation

Channel Capacity

The *maximum SE*, or the *channel capacity*, of a memoryless AWGN channel with CSNR γ is defined as:

$$C_{\text{AWGN}}(\gamma) = B \cdot \log_2(1+\gamma) \quad \text{[bits/s]}, \tag{2.2}$$

where B [Hz] is the bandwidth of the channel, and $\gamma = \gamma(t)$ is the CSNR of the channel at time t (the time index t is omitted in order to simplify notation). The channel capacity is often defined in terms of information throughput per Hz available bandwidth and can then be written as

$$C(\gamma) = \frac{C_{\text{AWGN}}(\gamma)}{B} = \log_2(1+\gamma) \quad \text{[bits/s/Hz]}. \tag{2.3}$$

In [Goldsmith and Varaiya, 1997] the capacity of a frequency-flat fading channel with AWGN was investigated under the assumptions of perfect channel knowledge at both transmitter and receiver, and that a communication system is able to instantaneously adapt to changes in the CSNR. When the transmitter only adapts the SE, and not the transmit power, the capacity of the channel in Figure 1.2 was given in [Goldsmith and Varaiya, 1997] as

$$C = \int_0^\infty \log_2(1+\gamma) f_\gamma(\gamma) d\gamma \text{ [bits/s/Hz]}. \tag{2.4}$$

To achieve the capacity in Equation (2.4) by means of an ACM scheme it is required that the system can instantaneously switch between $N = \infty$ capacity achieving codecs designed for AWGN channels, and that codecs for all CSNRs are available ("continuous" switching between codecs).

For adaptive transmission schemes the maximal possible SE will vary with the instantaneous channel conditions, in our case the CSNR. The resulting ASE is the expected SE over all available adaptation modes. The MASE for the Rayleigh fading channel is the expected channel capacity in the Shannon sense, i.e., the ultimate upper theoretical limit of the ASE in Equation (2.4), and for the case of constant power it is given by [Alouini and Goldsmith, 1997b, Eq. (23)][1]

$$\text{MASE}(\overline{\gamma}) = -e^{\frac{1}{\overline{\gamma}}} \log_2(e)\, E_i\left(-\frac{1}{\overline{\gamma}}\right) \qquad \text{[bits/s/Hz]}, \tag{2.5}$$

where $E_i(\cdot)$ is the *exponential integral function* defined in Equation (A.3).

The formula in Equation (2.5) holds under the assumption that the receiver is able to estimate the CSI perfectly and transmit it back to the transmitter on a noiseless zero-delay return channel. Also, achieving the MASE requires infinitely long codewords in the channel coder, infinitely many codecs, and the ability to instantly switch codec when the CSI changes.

[1] In [Alouini and Goldsmith, 1997b] the MASE was presented for the Nakagami case. Rayleigh fading, which is treated in this thesis, is a special case of the Nakagami fading distribution with $m = 1$ [Stüber, 2001]. Thus, the expression here (in Equation (2.5)) is found by substituting $m = 1$ in [Alouini and Goldsmith, 1997b, Eq. (23)].

Thus, for a practical system with limited number of codecs, non-perfect prediction, and delay both in the transmission system and the channels, the MASE is an impractical bound on the SE.

Maximum ASE for ACM (MASA)

A *discrete approximation* to the formula in Equation (2.4) can be written as

$$C = \sum_{n=1}^{N} R_n P_n \text{ [bits/s]}, \tag{2.6}$$

where R_n is the SE of codec n and P_n is given as

$$P_n = \mathrm{P}(\gamma_n \leq \gamma < \gamma_{n+1}) = \int_{\gamma_n}^{\gamma_{n+1}} f_\gamma(\gamma) d\gamma. \tag{2.7}$$

That is, P_n is the probability of codec n being chosen for transmission.[2] When N goes towards infinity and $R_n = \log_2(1+\gamma_n)$ [bits/s] the approximation in Equation (2.6) goes towards the capacity expression in Equation (2.4).

To find the largest value of R_n we make the following observations; The *sample space* of the CSNR is given as $\gamma \in [0, \infty\rangle$. In rate-adaptive coding, as explained in Section 1.3, the sample space is divided in $N+1$ *subsets*: $[\gamma_n, \gamma_{n+1}\rangle$ for $n \in \{0, 1, 2, \ldots, N\}$. The subsets are disjoint,

$$[\gamma_n, \gamma_{n+1}\rangle \cap [\gamma_m, \gamma_{m+1}\rangle = \emptyset \text{ for } n \neq m, \tag{2.8}$$

and the sample space can be written as

$$\cup_{n=0}^{N} [\gamma_n, \gamma_{n+1}\rangle = [0, \infty\rangle \tag{2.9}$$

where $\gamma_{N+1} = \infty$.

Now, it follows that

$$\max R_n = C|_{\gamma \in [\gamma_n, \gamma_{n+1}\rangle}, \tag{2.10}$$

where $C|_{\gamma \in [\gamma_n, \gamma_{n+1}\rangle}$ denotes the capacity conditioned on the CSNR being in subset n, $[\gamma_n, \gamma_{n+1}\rangle$. Using the *law of total expectations* (see e.g. [Råde and Westergren, 1990]) the capacity in Equation (2.6) can be written as

$$C = \sum_{n=1}^{N} C|_{\gamma \in [\gamma_n, \gamma_{n+1}\rangle} P_n. \tag{2.11}$$

[2] Note, that in this chapter perfect prediction is assumed. This means that the predicted CSNR is equal to the actual CSNR. In this case, P_n depends on the actual CSNR. However, when this assumption fails the predicted CSNR is used to choose the codec index n and Equation (2.7) is not valid.

This formula is however independent on the choice of $\{\gamma_n\}_{n=1}^{N}$, as long as Equations (2.8) and (2.9) are satisfied. It is also irrelevant how N is chosen. In fact, if capacity achieving fixed codes for fading channels were available, we could choose to use only one codec ($N = 1$). However, note that in order to achieve this capacity we need a codec designed for the particular fading channel.

Following the idea in [Holm et al., 2003], it is from now on instead assumed that the communication system contains $N < \infty$ capacity-achieving codecs designed for AWGN channels. This implies modeling the MFC as time-multiplexed AWGN channels, which is appropriate if the fading is (more or less) constant between two successive codec updates. Now, the SE for each codec n is given by the capacity of some AWGN channel, i.e.

$$R_n(\gamma') = \log_2(1+\gamma') \quad \text{where} \quad \gamma' \in [\gamma_n, \gamma_{n+1}\rangle \tag{2.12}$$

and γ' is the CSNR of the AWGN channel in question. The BER for a capacity achieving codec on an AWGN channel with CSNR γ' can be tightly upper bounded as

$$\text{BER}(\gamma) \leq \begin{cases} \frac{1}{2} & \gamma < \gamma' \\ 0 & \gamma \geq \gamma'. \end{cases} \tag{2.13}$$

Since codec n is selected when $\gamma \in [\gamma_n, \gamma_{n+1}\rangle$, we know that γ' must be in the same region (or subset): $\gamma' \in [\gamma_n, \gamma_{n+1}\rangle$. However, choosing $\gamma' > \gamma_n$ would result in a BER up to $1/2$ when $\gamma \in \langle \gamma_n, \gamma' \rangle$ and since

$$\text{P}(\gamma_n < \gamma < \gamma') = \int_{\gamma_n}^{\gamma'} f_\gamma(\gamma) d\gamma > 0 \quad \text{if } \gamma' > \gamma_n, \tag{2.14}$$

we must require $\gamma' = \gamma_n$ for error-free transmission. Thus, the maximum SE for the codec to be used in region n is given as

$$\max R_n = \log_2(1+\gamma_n). \tag{2.15}$$

The MASA is obtained when each of the N codecs employed have the maximum SE in the corresponding CSNR regions. Thus, the MASA defined as (see e.g. [Holm et al., 2003])

$$\text{MASA} = \sum_{n=1}^{N} \log_2(1+\gamma_n) P_n \tag{2.16}$$

is an upper bound on the ASE when assuming block wise multiplexed AWGN channels and capacity achieving codecs. Note, the MASA is not the capacity for an ACM scheme unless we restrict the individual codecs to be optimal for AWGN channels with CSNR levels equal to the CSNR thresholds $\{\gamma_n\}_{n=1}^{N}$ (see Figure 1.4).

2.4 Optimal CSNR Thresholds and Optimal MASA

In order to optimize Equation (2.16) with respect to the CSNR thresholds $\{\gamma_n\}_{n=1}^N$, the gradient with respect to the CSNR thresholds is first obtained. Setting the result to zero produces N equations:

$$\nabla_{\{\gamma_n\}_{n=1}^N} \text{MASA} = \begin{bmatrix} \frac{\delta \text{MASA}}{\delta \gamma_1} \\ \frac{\delta \text{MASA}}{\delta \gamma_2} \\ \vdots \\ \frac{\delta \text{MASA}}{\delta \gamma_n} \\ \vdots \\ \frac{\delta \text{MASA}}{\delta \gamma_N} \end{bmatrix} = \mathbf{0}. \tag{2.17}$$

Solving these equations with respect to the CSNR thresholds involves inverting the *cumulative distribution function* (CDF) of the CSNR. Assuming Rayleigh fading the CSNR has an exponential PDF (see Equation (1.4)) with corresponding CDF:

$$F_\gamma(\gamma) = \int_0^\gamma f_\gamma(x)dx = \int_0^\gamma \frac{1}{\overline{\gamma}} e^{-\frac{x}{\overline{\gamma}}} dx = 1 - e^{-\frac{\gamma}{\overline{\gamma}}}. \tag{2.18}$$

Optimal MASA under a Constraint on the Outage Probability

By expressing γ_1 in terms of P_0 (using Equations (2.1) and (2.18)),

$$\gamma_1 = \overline{\gamma} \ln(P_0 - 1), \tag{2.19}$$

the optimal MASA can be constrained by an outage probability demand: $P_0 \leq \text{P}_{\text{out}}$. By selecting the outage probability before obtaining the optimal CSNR thresholds, the optimal MASA is not necessarily found for $\nabla\text{MASA} = \mathbf{0}$. Introducing a single Lagrange multiplier, λ, the optimization procedure under an outage constraint can be written as

$$\nabla_{\{s_n\}_{n=1}^N} \text{MASA} + \lambda \text{P}_{\text{out}} = \nabla_{\{s_n\}_{n=1}^N} \text{MASA} + \lambda \int_0^{\gamma_1} f_\gamma(\gamma)d\gamma = \mathbf{0}. \tag{2.20}$$

Thus, obtaining the optimal thresholds involves finding the Lagrange multiplier that satisfies Equation (2.20). To simplify the notation Equation (2.20)

is multiplied by the constant $\ln 2$, resulting in the following N equations:

$$\begin{bmatrix} \frac{1}{(1+\gamma_1)}P_1 - \ln(1+\gamma_1) f_\gamma(\gamma_1) + \lambda \ln(2) f_\gamma(\gamma_1) \\ \ln(1+\gamma_1) f_\gamma(\gamma_2) + \frac{1}{(1+\gamma_2)}P_2 - \ln(1+\gamma_2) f_\gamma(\gamma_2) \\ \vdots \\ \ln(1+\gamma_{n-1}) f_\gamma(\gamma_n) + \frac{1}{(1+\gamma_n)}P_n - \ln(1+\gamma_n) f_\gamma(\gamma_n) \\ \vdots \\ \ln(1+\gamma_{N-1}) f_\gamma(\gamma_N) + \frac{1}{(1+\gamma_N)}P_N - \ln(1+\gamma_N) f_\gamma(\gamma_N) \end{bmatrix} = \mathbf{0}, \quad (2.21)$$

where P_n is given by Equation (2.7). Assuming Rayleigh fading and using Equation (2.18) the finite integrals P_n are given as

$$P_n = F_\gamma(\gamma_{n+1}) - F_\gamma(\gamma_n) = e^{-\frac{\gamma_n}{\overline{\gamma}}} - e^{-\frac{\gamma_{n+1}}{\overline{\gamma}}}. \quad (2.22)$$

Now, the first row in Equation (2.21) can be used to express γ_2 in terms of γ_1. Likewise, using the second row for expressing γ_3, and the nth row for expressing γ_{n+1} results in the complete set of equations to find the optimal CSNR thresholds $\{\gamma_n\}_{n=1}^N$:

$$\gamma_1 = -\ln(1-\mathrm{P_{out}})\,\overline{\gamma} \quad (2.23)$$

$$\gamma_2 = \gamma_1 - \overline{\gamma} \ln\left(1 - \frac{1}{\overline{\gamma}}(1+\gamma_1)(\ln(1+\gamma_1) - \lambda\ln(2))\right) \quad (2.24)$$

$$\gamma_3 = \gamma_2 - \overline{\gamma} \ln\left(1 + \frac{1}{\overline{\gamma}}(1+\gamma_2)\ln\left(\frac{1+\gamma_1}{1+\gamma_2}\right)\right) \quad (2.25)$$

$$\vdots$$

$$\gamma_n = \gamma_{n-1} - \overline{\gamma} \ln\left(1 + \frac{1}{\overline{\gamma}}(1+\gamma_{n-1})\ln\left(\frac{1+\gamma_{n-2}}{1+\gamma_{n-1}}\right)\right) \quad (2.26)$$

$$\vdots$$

$$\gamma_N = \gamma_{N-1} - \overline{\gamma} \ln\left(1 + \frac{1}{\overline{\gamma}}(1+\gamma_{N-1})\ln\left(\frac{1+\gamma_{N-2}}{1+\gamma_{N-1}}\right)\right). \quad (2.27)$$

Note, for $\lambda = 0$, which is the case of no constraint on the outage probability, the equations for γ_n, $n \geq 2$, is equal to the result presented in [Holm et al., 2003]. Equations (2.23) through (2.27) can in principle also be used

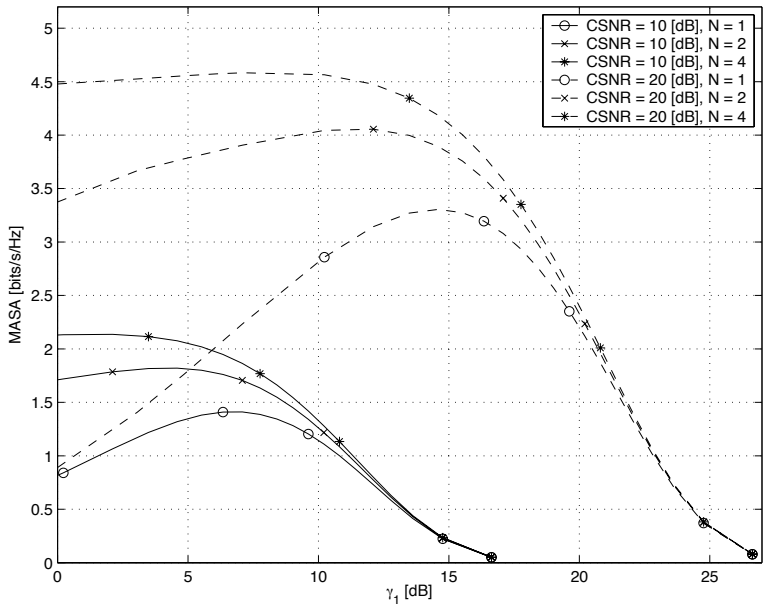


FIGURE 2.1: MASA plotted as a function of γ_1 for $N \in \{1,2,4\}$ and $\overline{\gamma} \in \{10,20\}$ dB.

to obtain the optimal thresholds for this case, by first obtaining the outage probability that maximize the ASE (for $\lambda = 0$).

When the outage probability is chosen according to some design criteria, the optimal thresholds are found by first obtaining the λ that maximizes the ASE.[3] In the following the MASA for the Rayleigh fading channel is plotted for different values of the CSNR, number of codecs, and for different given values of γ_1 or, correspondingly, the outage probability.

In Figure 2.1 the MASA is plotted for different given values of γ_1, $\overline{\gamma} \in \{10,20\}$dB, and for $N \in \{1,2,4\}$. From the curves it can be seen that, for a given N, the MASA can always be maximized by choosing the appropriate γ_1 if there are no outage constraints. It can also be observed that the MASA is always increasing for increasing N for any given value for γ_1. However, increasing γ_1 beyond an optimal value will reduce the MASA, and in the extreme there is almost no gain in increasing the number of codecs (see Figure 2.1 for $\gamma_1 > 15$ dB when $\overline{\gamma} = 10$ dB).

In Figure 2.2, the plots were again generated for $\overline{\gamma} \in \{10,20\}$ dB and

[3] In either case ($\lambda = 0$ or $\lambda \neq 0$) a search must be made to obtain the optimal value of the outage probability or λ. For each search all thresholds must be obtained and the resulting MASA calculated.

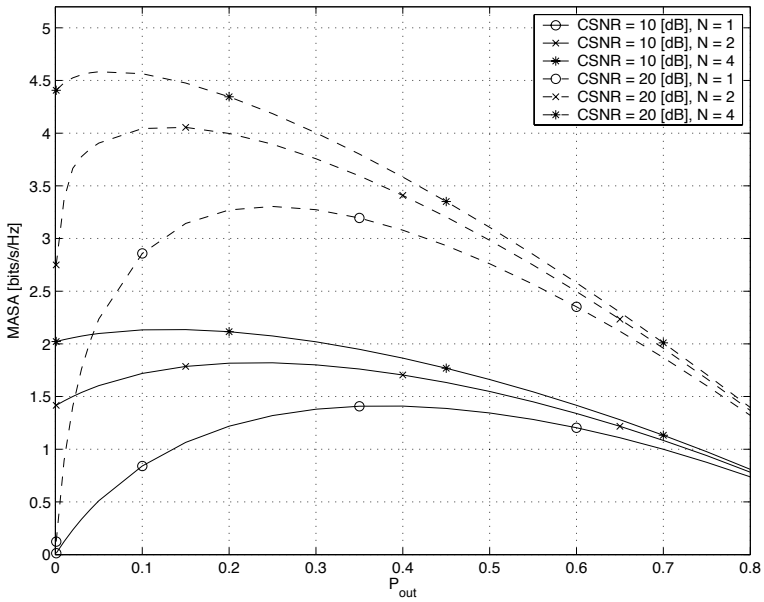


FIGURE 2.2: MASA plotted as a function of P_{out} for $N \in \{1,2,4\}$ and $\overline{\gamma} \in \{10,20\}$ dB.

$N \in \{1,2,4\}$, but the MASA is plotted for varying values of the designer-chosen outage probability. Assuming that P_{out} is never chosen to be larger than the value that maximize MASA, for a given N, the figure shows that there is a trade-off between MASA and P_{out}. For small values of P_{out} and high values of N the optimal MASA varies less with P_{out}, compared to low values of N. That is, the trade-off between high ASE and low P_{out} is not as crucial for higher values of N. From Figure 2.1 it can also be observed that as $\overline{\gamma}$ increases the optimal γ_1 increases. As a consequence the MASA is more sensitive to an outage demand for high average CSNRs (as seen in Figure 2.2).

The resulting MASA for the case of no outage demand is plotted for $N \in \{1,2,4,8,16\}$, along with the MASE of the Rayleigh fading channel, in Figure 2.3. As the results in the figure show, increasing the number of codecs increases the optimal MASA towards the MASE. The increasing gap between the MASA curves and the MASE curve for increasing CSNR is somewhat misleading, as can be seen in Figure 2.4. In this figure the loss in SE (measured in percent of the MASE), in terms of difference between MASE and optimal MASA, is plotted for the choices of N in Figure 2.3. The figure shows that the percent wise reduction in SE is decreased as CSNR increases. For a high value of N the difference between the MASE and the

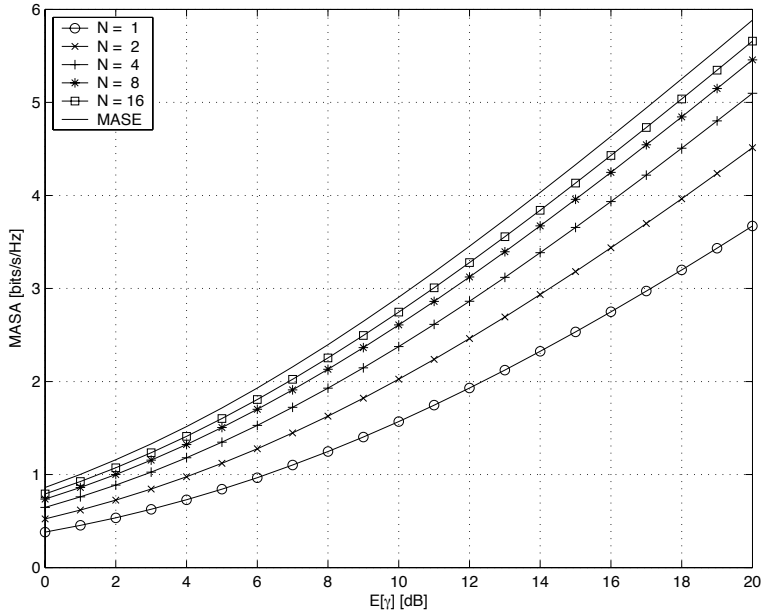


FIGURE 2.3: Unconstrained MASA of an ACM system on a Rayleigh fading for $N \in \{1, 2, 4, 8, 16\}$ codecs.

MASA is lower than for low values of N. And also, the reduction is not as sensitive to variations in the average CSNR for systems with a high number of codecs.

The outage probability is plotted for $N \in \{1, 2, 4\}$ and $\overline{\gamma} \in [0, 20]$ dB in Figure 2.5. As N and/or $\overline{\gamma}$ increase the outage probability is reduced. This implies that for a system with an outage demand, increasing the number of codecs is one obvious solution. However, adding an extra codec might be too expensive or impractical for the designer of an ACM scheme. Instead an ACM scheme with N codecs can be specified to have a given outage probability, at the price of a certain loss in ASE.

2.5 Concluding Remarks

In this chapter a unified tool for obtaining the SE bounds for a ACM scheme and the corresponding optimal CSNR was presented. The bounds can be used to evaluate the performance of realistic adaptive schemes in the sense that only $N < \infty$ codecs are used. The optimization procedure outlined results in a recursive procedure such that all CSNR thresholds can be represented as functions of the first CSNR threshold. For the case of Rayleigh

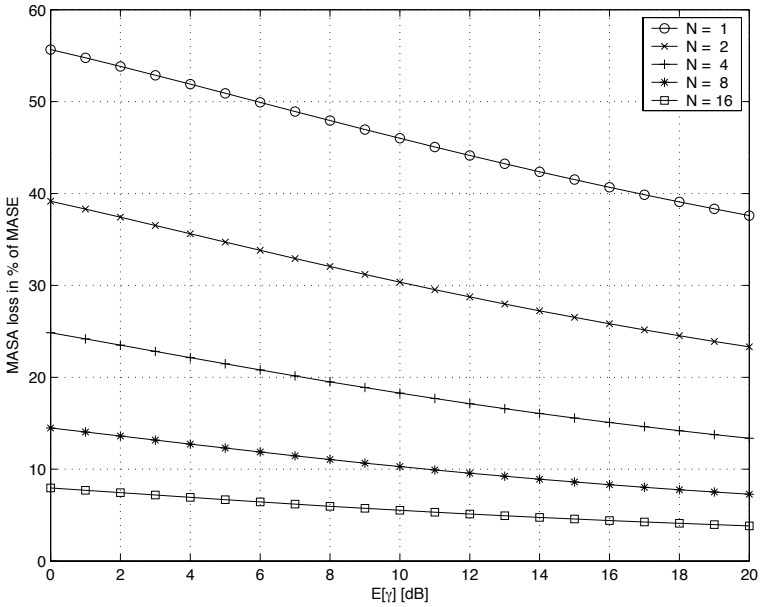


FIGURE 2.4: The percent wise reduction in SE, in terms of difference between MASE and MASA, of an ACM system on a Rayleigh fading for $N \in \{1, 2, 4, 8, 16\}$ codecs.

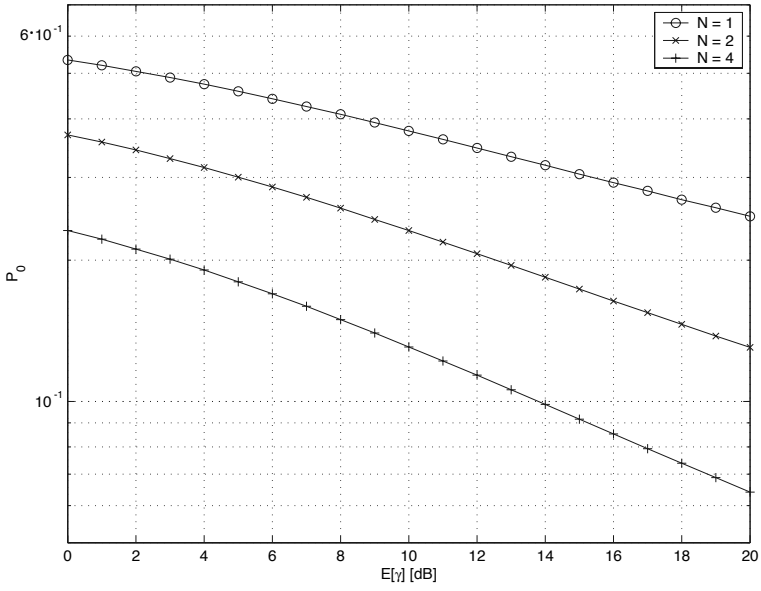


FIGURE 2.5: Outage Probability for $N = \{1, 2, 4\}$ codecs when the optimal CSNR thresholds are employed.

fading the PDF and CDF of the CSNR are of such beneficial forms that the equations used in the recursion can be solved analytically. For the case of e.g. the Nakagami fading channel it is necessary to optimize the system using numerical approximations [Holm et al., 2003]. The resulting MASA is a function of the CSNR threshold corresponding to the codec with the lowest SE. Then, optimizing MASA is done by searching for the optimal value of this threshold. The resulting MASA, which is also a function of this threshold, is then an upper bound on the ASE.

The tool also include optimization under an outage constraint. The resulting method is a powerful tool that can be used to improve certain QoS measures (that are defined in terms of outage probability). The results show that an ACM scheme using a high number of codecs can allow for a reduced outage probability without trading large proportions of the ASE. However, one should note that certain idealized conditions are still assumed during this chapter; perfect prediction and zero-delay return channel (these assumptions will be removed in the following chapters).

Since the method also produce the thresholds of the capacity achieving codecs employed it can be used as a guide in system design. That is, practical system design should strive towards finding codecs that perform with a BER close to the target BER at the optimal thresholds. Or equivalently, the optimal SE of each codec can be calculated (from the optimal thresholds) and system design should include searching for codecs with SEs close to these optimal codecs. There will of course be a loss, in terms of either ASE or BER or both, since a practical codec has a BER that is larger than zero also when the CSNR is above the corresponding threshold. It is shown how the optimal MASA increases as N increases, and that the gained MASA achieved by increasing $\overline{\gamma}$ also increases for increased N.

Chapter 3

ACM Design under Imperfect Channel State Information

3.1 Introduction

The time-varying nature of an MFC results in several issues which need to be addressed by the designer of a wireless communication system. Since the fading gain varies there is a need for *channel estimation* to improve detection and demodulation of the information carrying channel symbols. For an ACM scheme the selection of the appropriate codec for transmission is based on predicted CSI at some point in the future. Thus, some sort of *channel prediction* is required in order for an ACM scheme to operate properly.

Traditionally, and in the previous chapter, ACM schemes have been designed assuming *perfect* prediction of future CSNR values. Thus, the true CSNR thresholds can be used as *switching thresholds*. The codec with highest SE among those with switching thresholds smaller than the predicted CSNR, is selected for transmission.

In practice, the prediction is not perfect, and the switching thresholds should be chosen such that they ensure (with a high probability) that the ACM scheme does not select a codec that cannot guarantee the target BER. That is, it is desirable to control the probability (denoted ϵ) of the actual CSNR being above a certain CSNR threshold when the predicted CSNR is above the corresponding switching threshold.

In the case of perfect prediction, an ACM scheme can be designed such that the instantaneous BER is always smaller than the target BER. Assuming imperfect channel prediction, the probability ϵ of selecting a codec that cannot guarantee the target BER will always be larger than zero. A more practical requirement is to configure the ACM scheme such that the *average*

BER (over all predicted CSNRs) for each codec is less than the target BER. Then it is also ensured that the overall average BER for the ACM scheme is below the target BER.

In this chapter, we shall provide the necessary tools for designing switching thresholds for different values of the (designer chosen) probability ϵ. Choosing switching thresholds higher than the CSNR thresholds will clearly reduce both the average BER and the ASE. Therefore, a method for upper bounding the choice of ϵ under the constraint of the average BER being below the target BER will also be shown.

Most of the results presented in this chapter were presented in [Jetlund et al., 2004a; Jetlund, Øien, and Holm, 2004b]. The remainder of this chapter is organized as follows; In Section 3.2 the system model and a MAP-optimal filter that can be used in both channel prediction and channel estimation is briefly outlined. The main objective of this section is the resulting normalized correlation ρ between predicted and actual CSNR, which can be used to relate switching thresholds to CSNR thresholds taking into account the variations in the communication channel.

In Section 3.3 the theoretical performance of the predictor is investigated by means of the PDF for γ conditioned on $\widehat{\gamma}$. The expectation of this PDF is derived (the derivation is shown in Appendix B) in order to gain more insight in the relationship between predicted and actual CSNR.

In Section 3.4 the problem of obtaining the switching thresholds by means of the probability of codec mismatch ϵ is described. The Resulting expressions for the case of Rayleigh fading channels are outlined in Section 3.5.

In Section 3.6 expressions for the upper bound on ϵ are found for both capacity achieving codecs and for practical codecs that can be realized for use in an ACM scheme. The BER-versus-CSNR relationship for a set of Gallager codes is used in Section 3.7 to show the practical limitations and properties of ϵ and how choices of ϵ will affect the ASE. Concluding remarks on the contributions made in this chapter are given in Section 3.8.

3.2 MAP-optimal Estimation with PSAM

In this and the following section we shall consider the system outlined in Figure 3.1. The information-carrying channel symbols at the transmitter are denoted $d(k)$ where k is a discrete time index. Deterministic pilot symbols, $p(k)$, are multiplexed into the information stream at pilot time instants known at both transmitter and receiver. The resulting stream of channel symbols, $x(k)$, are transmitted on the communication channel described in

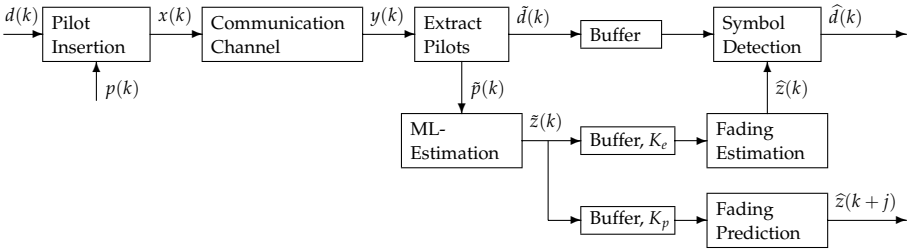


FIGURE 3.1: Pilot symbol assisted modulation used in both channel prediction and channel estimation.

Figure 1.2 and Equation (1.1). That is, for the case of discrete time indexes the received signal is given as

$$y(k) = z(k) \cdot x(k) + n(k). \tag{3.1}$$

It is here assumed that the pilot symbols are inserted after equal spaced intervals, such that $x(k)$ is the information signal except at the pilot instants $k = lL$ where l is an integer and L is a constant integer larger than zero.

For simplicity, although it has recently been shown not to be optimal [Chai and Giannakis, 2005; Duong and Øien, 2004], it is also assumed here (as in [Holm, 2002]) that all pilot symbols have the same amplitude value

$$x(lL) = a_p, \tag{3.2}$$

and that the transmit power for pilots is equal to the average transmit power for information carrying channel symbols, that is,

$$|a_p| = \sqrt{P}. \tag{3.3}$$

Since every Lth channel symbol $x(k)$ is a pilot symbol, the SE of codec n is reduced by a factor $(L-1)/L$. The resulting ASE for an ACM scheme using this configuration then becomes

$$\text{ASE}_L = \sum_{n=1}^{N} \frac{L-1}{L} R_n \cdot P_n \text{ [bits/s/Hz]}. \tag{3.4}$$

At the receiver the channel symbols $y(k)$ from the channel are separated into channel symbols carrying information, $\tilde{d}(k)$, and pilot symbols, $\tilde{p}(k)$. At each pilot symbol instant $k = lL$, the following ML-estimate at the pilot instants (based on one received observation [Meyr et al., 1998]) of $z(lL)$ is calculated by the receiver:

$$\tilde{z}(k) = \frac{\tilde{p}(k)}{a_p} = \frac{y(k)}{a_p} = z(k) + \frac{n(k)}{a_p}, \quad k = lL. \tag{3.5}$$

The channel predictor uses the ML-estimates to predict the fading at time instant $k+j$, where $j > 0$ represents the prediction lag. Throughout this thesis, it is assumed that the prediction lag takes into account both transmission delays and processing delays. The ML-estimates are also used by the channel estimator which produces estimates of the fading, $\widehat{z}(k)$, at every time instant k between two consecutive pilot symbol time instants. The predictor and the estimator are linear filters of length K_p and K_e, respectively.

MAP-optimal Filter Coefficients

Obtaining the filter coefficients for the MAP-optimal filter requires knowledge of the autocorrelation function of the fading envelope. As described in Section 1.4 a WSS MFC can be characterized by its maximum Doppler frequency $f_m = v \cdot f_c/c$ [Hz]. Under the assumption of isotropic scattering, the fading is said to have a Jakes spectrum, and the autocorrelation of the complex fading envelope can be written as [Stüber, 2001]

$$\phi_{zz}(\tau) = \frac{1}{2}\mathrm{E}[z^*(t)z(t+\tau)] = \frac{\Omega_p}{2}J_0(2\pi f_m \tau), \tag{3.6}$$

where τ is the time delay between samples. The normalized autocorrelation function sampled at a period equal to T_s is then defined as

$$R(l) = \frac{1}{\Omega_p}\mathrm{E}[z^*(k)z[k+l]] = J_0(2\pi f_m T_s l) = \frac{2}{\Omega_p}\phi_{zz}(T_s l). \tag{3.7}$$

The higher v, and thus f_m, the less correlated the fading typically is for a given l. Also, the correlation is reduced as l increases.

A sequence of ML-estimates represents estimated values of a decimated sequence of the actual fading. In this case, a channel predictor can be implemented as a Wiener filter with precomputed filter coefficients for all ML-estimates $\tilde{z}(k)$ [Meyr et al., 1998]. The realization of the channel predictor in Figure 3.1 uses K_p ML-estimates on received pilot symbols in the past of the fading sample to be predicted. The ML-estimates are shifted into a buffer of size K_p, where K_p is an even integer. The buffered estimates, at time instant k, can be written as

$$\tilde{\mathbf{z}}_k = [\tilde{z}(k), \tilde{z}(k-L), \ldots, \tilde{z}(k-(K_p-1)L)]^T, \tag{3.8}$$

where $k = lL$. The predicted value can be written as (see e.g. [Meyr et al., 1998])

$$\widehat{z}(k+j) = \mathbf{h}_j^T \tilde{\mathbf{z}}_k, \tag{3.9}$$

where

$$\mathbf{h}_j = [h_j(0), h_j(1), \ldots, h_j(K_p - 1)]^T \tag{3.10}$$

is the vector holding the estimation filter coefficients which correspond to estimating the symbol at time instant $k + j$.

The MAP-optimal filter coefficient vector of length K on a Rayleigh fading channel is given as (see e.g. [Meyr et al., 1998; Øien, Holm, and Hole, 2002c])

$$\mathbf{h}_{j,\mathrm{MAP}} = \mathbf{r}_{j,K}^T \left(\mathbf{R}_K + \frac{1}{\overline{\gamma}} \mathbf{I}_{K \times K} \right)^{-1}, \tag{3.11}$$

where $\mathbf{I}_{K \times K}$ is a $K \times K$ identity matrix. The vector $\mathbf{r}_{j,K}$ is of length K and the elements represent the normalized covariance between the fading to be predicted at time $k + j$ and the fading at the pilot instants $k, k - L, k - 2L, \ldots, k - (K - 1)L$. $\mathbf{R}_K$ is the normalized covariance matrix, but can be viewed as a $K \times K$-matrix holding the normalized autocorrelation of the fading process sampled at a period equal $L \cdot T_s$. With the assumption of Jakes spectrum in the fading process, the elements of $\mathbf{r}_{j,K}$ and $\mathbf{R}_K$ can be calculated from the following equations (using the normalized autocorrelation in Equation (3.7)):

$$\left[\mathbf{r}_{j,K}\right]_l = \frac{1}{\Omega_p} \mathrm{E}\left[z^*(k+j)\, z(k - lL)\right] = \frac{2}{\Omega_p} \phi_{zz}\left((j + lL)\, T_s\right) \tag{3.12}$$

and

$$[\mathbf{R}]_{lm} = \frac{1}{\Omega_p} \mathrm{E}\left[z^*(k - lL)\, z(k - mL)\right] = \frac{2}{\Omega_p} \phi_{zz}\left(|l - m| L T_s\right) \tag{3.13}$$

where l and m are integers.

The channel predictor can now be implemented using the filter coefficients in Equation (3.11) with $K = K_p$. The filter coefficients in Equation (3.11) can also be used in the channel estimator (with $K = K_e$), but it is here assumed that the channel estimator is allowed to use ML-estimates from pilot symbols both in the past and the future of the fading sample to be estimated. That is, elements of $\mathbf{r}_{j,K}$ are calculated for both negative and positive values of j.

Channel State Information

The predicted CSNR at time instant $k + j$, $\widehat{\gamma}(k + j)$, must be estimated from the predicted fading $\widehat{z}(k + j)$. Defining

$$\mathrm{E}\left[|\widehat{z}(k+j)|^2\right] = \widehat{\Omega}_p, \tag{3.14}$$

then there exists a constant r such that [Tang, Alouini, and Goldsmith, 1999; Øien et al., 2002c]

$$\widehat{\Omega}_p = r \cdot \Omega_p. \tag{3.15}$$

Since the predicted fading envelope is a linear combination of complex Gaussians, it is itself a complex Gaussian. It follows that the predicted fading is Rayleigh distributed, and that the corresponding predicted CSNR,

$$\widehat{\gamma}(k+j) = \frac{|\widehat{z}(k+j)|^2 P}{N_0 B}, \tag{3.16}$$

is exponentially distributed [Øien et al., 2004],

$$f_{\widehat{\gamma}}(\widehat{\gamma}) = \frac{1}{r\overline{\gamma}} e^{-\frac{\widehat{\gamma}}{r\overline{\gamma}}}, \tag{3.17}$$

with expectation

$$\mathrm{E}[\widehat{\gamma}] = \frac{\widehat{\Omega}_p P}{N_0 B} = r \cdot \overline{\gamma}. \tag{3.18}$$

The *normalized correlation* between the actual and predicted CSNR given as

$$\rho = \frac{\mathrm{Cov}(\widehat{\gamma}, \gamma)}{\sqrt{\mathrm{Var}(\widehat{\gamma})\mathrm{Var}(\gamma)}} \leq 1, \tag{3.19}$$

where $\mathrm{Cov}(\cdot,\cdot)$ and $\mathrm{Var}(\cdot)$ are the covariance and variance operators, respectively. It is shown in [Holm, 2002; Øien, Holm, and Hole, 2002a] that for the MAP-optimal predictor, the ratio r and the correlation coefficient ρ are both equal to

$$\rho = r = \mathbf{r}_{j,K}^T \left(\mathbf{R}_K + \frac{1}{\overline{\gamma}} \mathbf{I}_{K \times K} \right)^{-1} \mathbf{r}_{j,K}, \tag{3.20}$$

with $K = K_p$.

3.3 Predictor Performance

The predictor performance (i.e. the accuracy of the predictor) depends on the ML-estimates employed in the prediction, the variations in the communication channel, and the bias of the predictor. The ratio r between the average actual CSNR and the average predicted CSNR lies between zero and unity since the ratio is equal to the normalized correlation (Equation (3.20)). Thus, the MAP-optimal prediction filter has a negative bias (i.e. the predicted CSNR will *on average* be smaller than the average CSNR).

The variation of the fading envelope is described by the normalized correlation which is controlled by $\phi_{zz}(\tau)$ or equivalently by $R(m)$ (through Equations (3.20), (3.12), and (3.13)), which again varies with the maximum Doppler frequency f_m and the time delay τ (or $j \cdot T_s$). Also, the normalized correlation in Equation (3.20) depends on the filter length K_p, the pilot spacing L, and the average CSNR $\overline{\gamma}$ through Equations (3.12) and (3.13).

In order to relate general results to a more practical case the following example can be considered, both here and in later chapters.

Example 3.1 (Correlation properties of the MAP-optimal predictor)
Using a carrier frequency $f_c = 5.4$ GHz and a symbol time duration $T_s = 4\ \mu$s (collected from the ETSI HIPERLAN/2 standard [ETSI, 2000]), a normalized Doppler frequency $f_m T_s \in [0, 8 \cdot 10^{-4}]$ corresponds to a Doppler frequency and a relative transmitter-receiver velocity of

$$f_m \in [0, 200] \text{ Hz}$$

and

$$v \in [0, 11.11] \text{ m/s},$$

respectively. Likewise the normalized prediction lag $j f_m T_s \in [0, 0.2]$ corresponds to a prediction lag measured in channel symbols and seconds of

$$j \in [0, 555] \text{ symbols}$$

and

$$\tau \in [0, 2.22] \text{ ms},$$

respectively, when the normalized Doppler frequency is set to $f_m T_s = 3.6 \cdot 10^{-4}$ (or equivalently $v = 5$ [m/s]).

The normalized correlation between the actual and predicted CSNR was studied in [Holm, 2002]. It was shown that ρ was severely reduced when the pilot spacing L was increased. Increasing the number of filter coefficients K_p showed an improvement in terms of increased ρ. However, the results in [Holm, 2002] indicated that (at least for large values of K_p) the correlation was not very sensitive to changes in K_p. In the subsequent chapters it is assumed that the MAP-optimal predictor is used, that the correlation between actual and predicted CSNR is known, and that the correlation is high, typically $0.9 < \rho < 1$. It is then effectively assumed that the normalized prediction lag is kept low (in [Holm, 2002] it was shown that the normalized correlation for a predictor with $K_p \geq 500$ and $L = 10$ is above 0.9 as long as the average CSNR is above 10 dB and the normalized prediction delay is less than 0.25 seconds).

Properties of the Fading Envelope

In [Falahati, Svensson, Sternad, and Mei, 2003] the BER performance of an adaptive trellis coded modulation system was compared to the performance of an adaptive uncoded *quadrature amplitude modulation* (QAM) scheme. The BER was evaluated as a function of $\widehat{\gamma}$. The results in [Falahati et al., 2003] indicate that the CSNR thresholds should be increased when the average CSNR increases. This may seem somewhat counter intuitive since the correlation of a channel predictor increases with $\overline{\gamma}$, and yields a more accurate prediction.

The performance of an ACM scheme depends on the accuracy of the predicted CSNR $\widehat{\gamma}$, and the probability of this prediction being sufficiently close to the actual CSNR. For the case of Rayleigh fading the two CSNRs are exponentially distributed. Using Equation (B.4) the PDF of γ conditioned on $\widehat{\gamma}$ can be written as

$$f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) = \frac{1}{\overline{\gamma}(1-\rho)} I_0\left(\frac{2\sqrt{\gamma\widehat{\gamma}}}{\overline{\gamma}(1-\rho)}\right) e^{-\left(\frac{\gamma}{\overline{\gamma}(1-\rho)}+\frac{\widehat{\gamma}}{\overline{\gamma}(1-\rho)}\right)} \tag{3.21}$$

where the expectation of γ and $\widehat{\gamma}$ is $\overline{\gamma}$ and $\rho\overline{\gamma}$, respectively and ρ is the normalized correlation between γ and $\widehat{\gamma}$. The expectation of γ conditioned on $\widehat{\gamma}$ and ρ is given as

$$\mathrm{E}\left[\gamma|\widehat{\gamma},\overline{\gamma},\rho\right] = (1-\rho)\overline{\gamma} + \widehat{\gamma} \tag{3.22}$$

(for the derivation of the expectation of the conditional exponential distribution see Section B.3).

In Figure 3.2 the conditional PDF in Equation (3.21) is plotted against γ for a predicted CSNR $\widehat{\gamma} = 10$ dB, different values of the normalized prediction lag, $\overline{\gamma} \in \{0, 10, 20\}$ dB, $K_p = 500$, and $L = 10$. In the upper and lower plot the normalized prediction lag was set to $jf_mT_s = 4 \cdot 10^{-2}$ and $jf_mT_s = 10^{-1}$, respectively. The PDF has its maximum at (or very close to) $\widehat{\gamma}$, indicating the most likely value for γ. As $\overline{\gamma}$ increases the PDF is flattened (i.e. the uncertainty in the actual value for γ increases with $\overline{\gamma}$). This behavior can be explained by observing that $\widehat{\gamma} \ll \overline{\gamma}$ corresponds to predicting a deep fade in the fading envelope. Any channel symbols transmitted during a deep fade is severely degraded by additive noise. Since this also holds for the pilot symbols a deep fade reduces the accuracy of the channel predictor. This can also be deduced from the plots in Figure 3.3 where the expected value of γ conditioned on $\widehat{\gamma}$, $\overline{\gamma}$, and ρ, is plotted against $\overline{\gamma}$; the expected value of γ increases with $\overline{\gamma}$ and increased normalized prediction lag (reduced ρ). That is, as $\overline{\gamma}$ increases a larger proportion of the probability

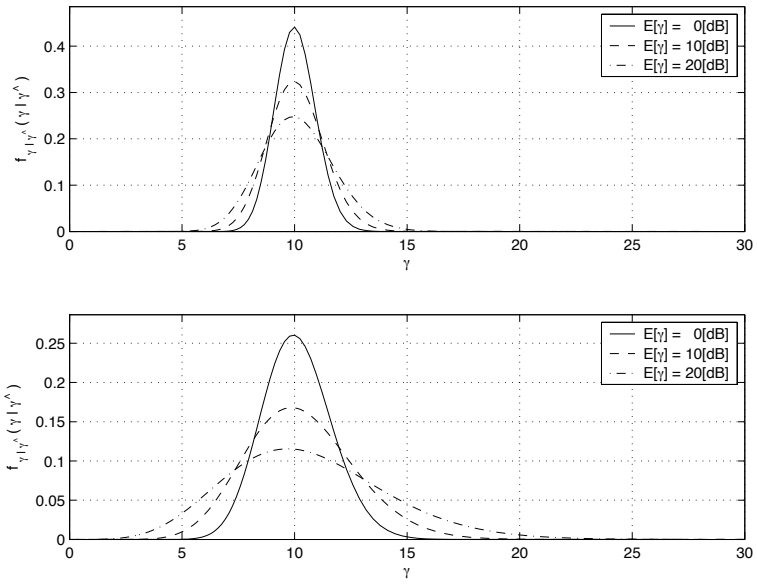


FIGURE 3.2: The conditional PDF for γ conditioned on $\widehat{\gamma}$, for $\widehat{\gamma} = 10$ dB.

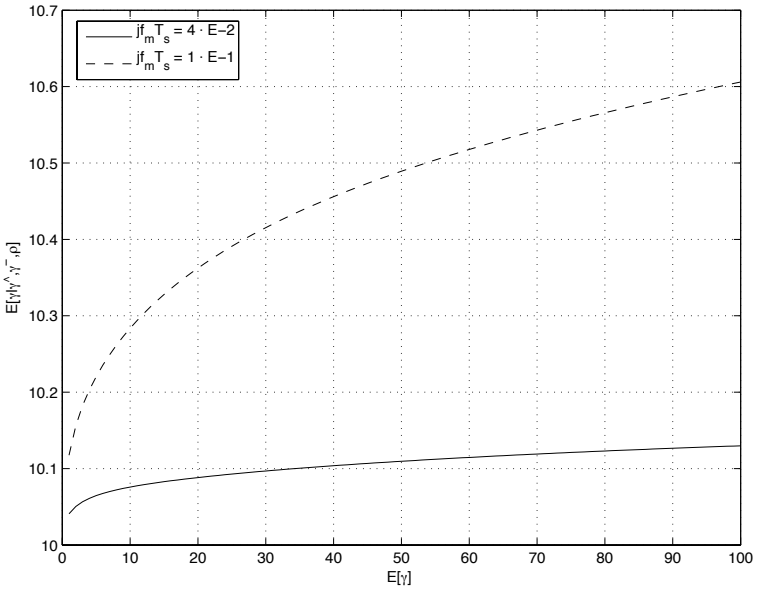


FIGURE 3.3: The expectation of γ conditioned on $\widehat{\gamma}$ and ρ plotted as function of $\overline{\gamma}$, for $\widehat{\gamma} = 10$ dB.

mass is moved to the right of $\widehat{\gamma}$ in Figure 3.2, reducing the probability of the actual CSNR being equal to the predicted CSNR. The increased expectation of γ when ρ is reduced is to be expected since a reduced ρ reduces the performance of a predictor. The switching thresholds should then be increased to maintain a constant probability of a codec mismatch.

3.4 Obtaining the Switching Thresholds: Problem Formulation

An imperfect channel predictor as described in previous sections will increase the average BER of an ACM scheme. The predicted CSNR might be higher than the actual CSNR and falling in a higher indexed CSNR region than the actual CSNR. As discussed previously the *switching thresholds* used to select codecs should be increased to reduce the probability of error.[1] These thresholds take into account the probability of the predictor making an error that result in choosing an inappropriate codec for transmission (a codec mismatch). The switching thresholds should then be found by taking into account predictor performance. However, in order to do this the correlation between actual and predicted CSNR must be known. We shall assume that the MAP-optimal estimator is used. The necessary expressions for the properties of the predicted CSNR were given in Section 3.2.

Since the actual CSNR is unknown at the receiver we define a *codec switching strategy* (CSS) by means of a new set of thresholds, for the predicted CSNR, that are used to select a codec based on the predicted CSNR. The switching thresholds are denoted

$$\{s_n\}_{n=0}^{N+1} = \{s_0, s_1, s_2, \ldots, s_N, s_{N+1}\}, \tag{3.23}$$

with $s_0 = 0$ and $s_{N+1} = \infty$. Codec n is selected when $\widehat{\gamma}$ falls in *switching* region n, defined as $[s_n, s_{n+1}\rangle$. The probability of selecting codec n can, in the case of imperfect channel prediction, be written as (using Equations (3.17) and (3.20))

$$P_n = \int_{s_n}^{s_{n+1}} f_{\widehat{\gamma}}(\widehat{\gamma}) d\widehat{\gamma} = \int_{s_n}^{s_{n+1}} \frac{1}{\rho\overline{\gamma}} e^{-\frac{\widehat{\gamma}}{\rho\overline{\gamma}}} d\widehat{\gamma}. \tag{3.24}$$

[1] Recall that, in this thesis, switching thresholds defines the regions used when the *predicted* CSNR is used for codec selection. CSNR thresholds (which can be used as switching thresholds when the normalized correlation between predicted and actual CSNR is $\rho = 1$) are defined as the CSNR level at which the corresponding codec performs with a BER equal to BER_0 on an AWGN channel. The set of switching thresholds should be found such that the average BER is less than or equal to BER_0 when $\rho < 1$.

When $\widehat{\gamma} \in [s_0, s_1\rangle$ the system experiences an outage and only pilot information is transmitted. The probability of outage is thus given by P_0 (and is found by substituting $n = 0$ into Equation (3.24)).

When perfect prediction is assumed ($\rho = 1$), the CSNR thresholds in Figure 1.4 can be used directly as switching thresholds: That is, $s_n = \gamma_n$ for $n \in \{1, \ldots, N\}$. As the normalized correlation ρ is reduced (corresponding to either a lower $\overline{\gamma}$, a larger prediction channel delay, or a higher terminal velocity) there is an increasing probability of mismatch between the predicted and actual CSNR. This will result in an increased BER, since the actual CSNR sometimes will fall into a lower indexed region than the predicted CSNR. Although this cannot be completely avoided with any CSS it is desirable to control the probability of this event. It is then natural to demand

$$\mathrm{P}(\gamma < \gamma_n | \widehat{\gamma} = s_n) = \int_0^{\gamma_n} f_{\gamma|\widehat{\gamma}}(\gamma | \widehat{\gamma} = s_n) d\gamma = \epsilon, \tag{3.25}$$

where ϵ is some (small) constant chosen by the designer [Øien, Holm, and Hole, 2002b]. Thus, by increasing the switching thresholds $\{s_n\}_{n=1}^N$ to obtain a certain desired (sufficiently small) ϵ the probability of a codec mismatch can be reduced in a more controlled manner.

3.5 Obtaining the Switching Thresholds: The Solution

Assuming that the actual CSNR is constant, an increase in the predicted CSNR reduces the probability of a codec mismatch. If s_n is chosen to fulfill Equation (3.25) this can be expressed as

$$\mathrm{P}(\gamma < \gamma_n | \widehat{\gamma} \geq s_n) \leq \epsilon. \tag{3.26}$$

One can in principle imagine ϵ chosen such that the BER always stays below BER_0, or so that it stays below BER_0 with a certain (sufficiently high) probability. The choice of ϵ can also be made codec dependent (if some codecs are used more frequently than others, a smaller ϵ might be appropriate for these codecs). Minimizing the probability of γ falling into a lower indexed region than $\widehat{\gamma}$ involves solving Equation (3.25), with respect to s_n, $n \in \{1, 2, \ldots, N\}$. Equivalently the equation for the probability of the complementary event, given as

$$1 - \epsilon = \mathrm{P}(\gamma > \gamma_n | \widehat{\gamma} = s_n), \tag{3.27}$$

can be solved with respect to s_n. In the following subsection a closed form expression that relates γ_n, s_n, and ϵ for the Rayleigh fading case is obtained.

Rayleigh Fading Case

Assuming a Rayleigh faded envelope the actual and predicted CSNRs are (as discussed previously in this chapter) correlated with normalized correlation ρ, and have expectations $\overline{\gamma}$ and $\rho\overline{\gamma}$, respectively. The two then follow a joint exponential PDF and the probability in Equation (3.27) can be obtained as follows (using the conditional PDF in Equation (B.4)):

$$\begin{aligned}
\mathrm{P}(\gamma > \gamma_n|\widehat{\gamma} = s_n) &= \int_{\gamma_n}^{\infty} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma} = s_n)d\gamma \\
&= \int_{\gamma_n}^{\infty} \frac{1}{\overline{\gamma}\,(1-\rho)} \times I_0\left(\frac{2\sqrt{\gamma s_n}}{\overline{\gamma}\,(1-\rho)}\right) \\
&\quad \times e^{-\left(\frac{\gamma}{\overline{\gamma}(1-\rho)} + \frac{s_n}{\overline{\gamma}(1-\rho)}\right)} d\gamma \\
&= \int_{\gamma_n}^{\infty} \frac{1}{\overline{\gamma}\,(1-\rho)} I_0\left(\sqrt{\frac{2\gamma}{\overline{\gamma}\,(1-\rho)}}\sqrt{\frac{2s_n}{\overline{\gamma}\,(1-\rho)}}\right) \\
&\quad \times e^{-\frac{1}{2}\left(\frac{2\gamma}{\overline{\gamma}(1-\rho)} + \frac{2s_n}{\overline{\gamma}(1-\rho)}\right)} d\gamma.
\end{aligned}$$

Substituting

$$a = \sqrt{\frac{2s_n}{\overline{\gamma}\,(1-\rho)}},$$

$$b = \sqrt{\frac{2\gamma_n}{\overline{\gamma}\,(1-\rho)}},$$

and

$$x = \sqrt{\frac{2\gamma}{\overline{\gamma}\,(1-\rho)}}$$

and changing the integration variable results in

$$\mathrm{P}(\gamma > \gamma_n|\widehat{\gamma} = s_n) = \int_b^{\infty} x I_0\,(ax)\, e^{-\frac{1}{2}\left(a^2+x^2\right)} dx = Q\,(a,b) \tag{3.28}$$

where $Q(\cdot,\cdot)$ is the Marcum-Q function [Marcum, 1948]. This closed form expression was also presented in [Øien et al., 2002b] with a different version of the Marcum-Q function,[2] but the derivation was not shown. In [Øien

[2]Note, there exist multiple definitions of the Q-function, the generalized Q-function, and the mth order Q-function that are all referred to as the Marcum Q-function. In [Øien et al., 2002b] the generalized (mth order) Marcum Q-function from [Temme, 1996] was used. The relationship between the Q-function used in [Øien et al., 2002b] and the Q-function used here can be found by substitution.

et al., 2002b] this result was actually presented for a Nakagami-m fading channel.

The new switching thresholds can now be found by solving

$$Q\left(\sqrt{\frac{2s_n}{\overline{\gamma}\,(1-\rho)}}, \sqrt{\frac{2\gamma_n}{\overline{\gamma}\,(1-\rho)}}\right) = 1-\epsilon \tag{3.29}$$

with respect to every s_n, $n \in \{1,2,\ldots,N\}$ for given values of $\overline{\gamma}$, ρ, and $\{\gamma_n\}_{n=1}^{N}$, and a chosen $\epsilon \in \langle 0,1\rangle$. To the author's knowledge there does not exist a closed form expression for the inverse of the Marcum-Q function. But the inverse of one of the arguments can be obtained, with a given accuracy, using e.g. Ridders' method [Ridders, 1979]. The inverse of the complementary Marcum-Q function, $1-Q(a,b)$, with respect to its first and second argument can be defined as

$$a = q_a(b,\epsilon) \tag{3.30}$$

and

$$b = q_b(a,\epsilon), \tag{3.31}$$

respectively. Tables of values of $a = q_a(b,\epsilon)$ and $b = q_b(a,\epsilon)$ are given in Appendix A. The inverse of Equation (3.29) with respect to s_n can then be written as (for $n \in \{1,2,\ldots,N\}$)

$$s_n(\epsilon) = \Psi_s\left(\gamma_n,\epsilon\right) = \begin{cases} \left(\sqrt{\frac{\overline{\gamma}(1-\rho)}{2}}\cdot q_a\left(\frac{\sqrt{s_n}}{\sqrt{\frac{\overline{\gamma}(1-\rho)}{2}}},\epsilon\right)\right)^2 & \rho<1 \\ \gamma_n & \rho=1 \end{cases} \tag{3.32}$$

and with respect to γ_n as

$$\gamma_n(\epsilon) = \Psi_\gamma\left(s_n,\epsilon\right) = \begin{cases} \left(\sqrt{\frac{\overline{\gamma}(1-\rho)}{2}}\cdot q_b\left(\frac{\sqrt{s_n}}{\sqrt{\frac{\overline{\gamma}(1-\rho)}{2}}},\epsilon\right)\right)^2 & \rho<1 \\ s_n & \rho=1. \end{cases} \tag{3.33}$$

Note that, the case of perfect prediction ($\rho = 1$) is not included in Equation (3.29), but is included in Equations (3.32) and (3.33) for completeness.

In the following curves s_n versus γ_n are plotted for different values of ϵ, $\overline{\gamma}$, and the normalized prediction lag (or equivalently ρ). The correlation properties results from using a MAP-optimal predictor with $K_p = 500$ coefficients and a pilot spacing of $L = 10$ symbols. Figure 3.4 shows s_n versus γ_n for different values of ϵ, with $\overline{\gamma} = 10$ dB and $\rho = 0.95$. The switching

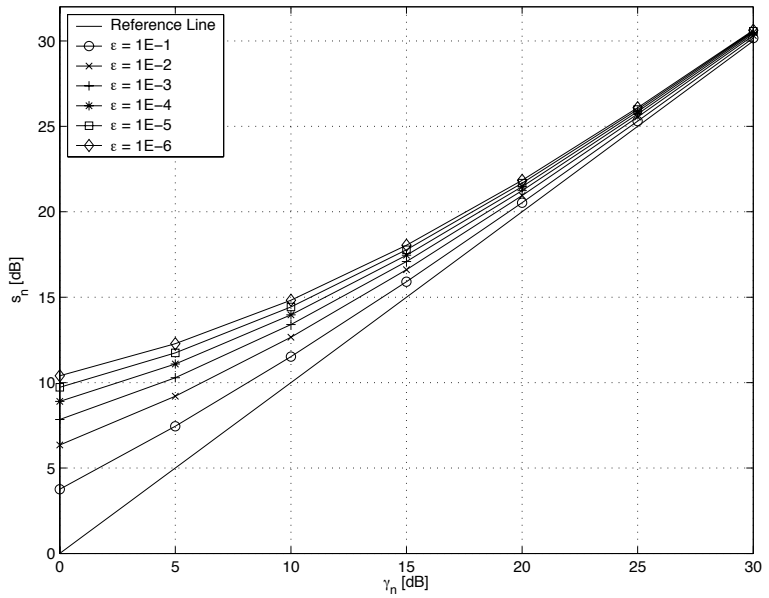


FIGURE 3.4: Plot of γ_n versus s_n for $\overline{\gamma} = 10$ dB, $\rho = 0.95$, and different values of ϵ.

thresholds are always greater than or equal to the CSNR thresholds, implying a more conservative choice of codecs. From the curves it can however be seen that the difference between the CSNR threshold and the new switching threshold is reduced as the CSNR threshold increases. This indicates that the ASE loss will be smaller for the high-rate codecs as long as $\overline{\gamma}$, ρ and ϵ are kept constant. The reason is that the prediction is typically more accurate at high CSNRs.

In Figure 3.5 the s_n versus γ_n curves are plotted for $\epsilon = 2 \cdot 10^{-3}$, $\overline{\gamma} = 10$ dB, and different values of the normalized prediction lag jf_mT_s. Since the correlation is reduced for increasing values of jf_mT_s the results show that s_n increases with reduced ρ. This is an expected result since the amount of predictor errors increase as ρ is reduced, and to keep the probability of codec mismatch constant the value of s_n must be increased accordingly.

The s_n versus γ_n curves in Figure 3.6 were found for $jf_mT_s = 7.9 \cdot 10^{-2}$, $\epsilon = 2 \cdot 10^{-3}$, and $\overline{\gamma} \in \{0, 10, 20, 30, 40\}$ dB. It can be observed that the switching thresholds are increased when $\overline{\gamma}$ is increased, as long as the normalized correlation is kept constant. This may seem counter intuitive, but can be explained by the properties of the conditional PDF—as discussed earlier in this chapter. That is, keeping the predicted CSNR constant and increasing the average CSNR also increases the probability of a higher ac-

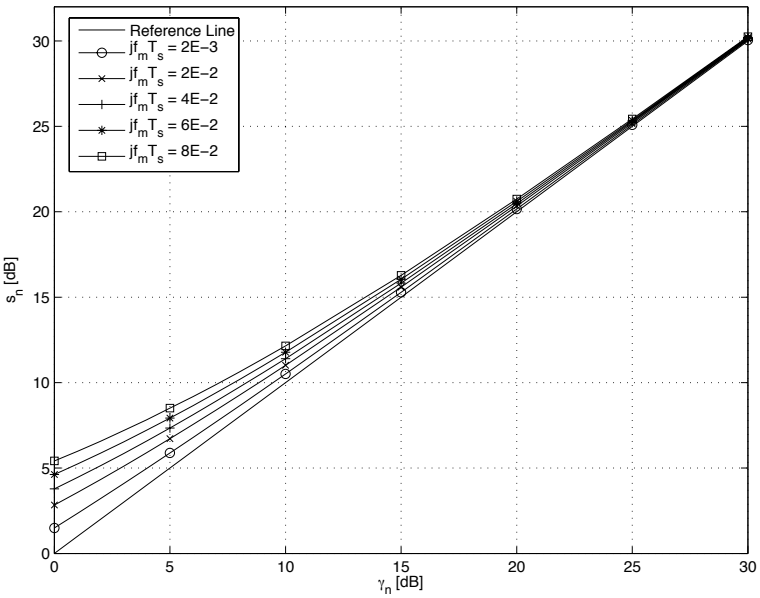


FIGURE 3.5: Plot of γ_n versus s_n for $\overline{\gamma} = 10$ dB, $\epsilon = 2 \cdot 10^{-3}$, $K = 500$, $L = 10$, $j = 200$ symbols, and different values of jf_mT_s.

tual CSNR, and vice versa. Note, that the choice of ϵ can be made codec dependent, and it is not yet clear whether ϵ should be increased or reduced as $\overline{\gamma}$ increases. This will be investigated in Section 3.7.

3.6 An Upper Bound on the Probability of Codec Mismatch

In order to find maximum values for the choice of ϵ the following assumptions are made; a switching threshold is always larger than the corresponding CSNR threshold, a codec n is chosen for a period of time when the fading is constant enough for BER performance to be closely approximated by the BER-versus-CSNR performance on an AWGN channel, and the prediction of future channel states is assumed to be imperfect.

Both capacity achieving codecs and practical codecs are considered as components in the ACM scheme. The BER-versus-CSNR performance of a capacity achieving codec and the typical *waterfall* curve for a practical codec is shown in Figure 3.7. The CSNR threshold γ_n is the CSNR level at which the practical codec performs with a BER equal to the target BER, and at which the capacity achieving codec achieves capacity.

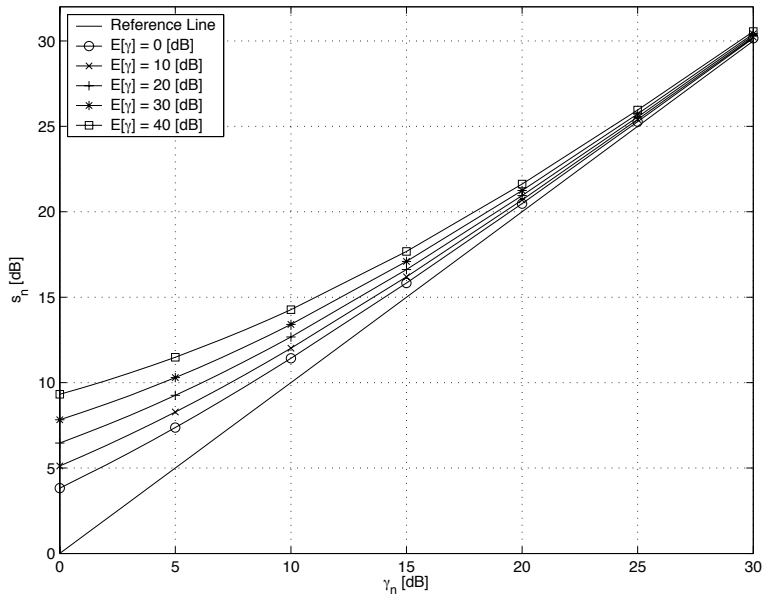


FIGURE 3.6: Plot of γ_n versus s_n for $\epsilon = 2 \cdot 10^{-3}$, $K = 500$, $L = 10$, $j = 200$ symbols, $jf_mT_s = 7.9 \cdot 10^{-2}$, and different values of $\overline{\gamma}$.

The BER of a capacity achieving codec was given in Equation (2.13). Thus, for codec n the actual BER is given as

$$\mathrm{BER}_n(\gamma) = \begin{cases} \frac{1}{2} & \gamma < \gamma_n \\ 0 & \gamma \geq \gamma_n. \end{cases} \tag{3.34}$$

In this case the event of choosing a wrong codec controls the overall BER. That is, only a codec mismatch may result in bit errors. However, for a practical codec contributions to the overall BER will occur at any CSNR. The BER of the practical codec can be approximated as

$$\mathrm{BER}_n(\gamma) \leq \begin{cases} \frac{1}{2} & \gamma < \gamma_n \\ \mathrm{BER}_0 & \gamma_n \leq \gamma < \gamma_{n,\tau} \\ \mathrm{BER}_{n,\tau} & \gamma \geq \gamma_{n,\tau}. \end{cases} \tag{3.35}$$

The *temporal threshold*, $\gamma_{n,\tau}$, utilized in the approximation is used for calculating the switching threshold s_n. How much a practical codec contributes to the overall BER depends on how rapid the BER falls as a function of the CSNR. The temporal threshold can be imagined as a the equivalent of the CSNR threshold of a capacity achieving codec. That is, the temporal threshold, or rather the BER for codec n at $\gamma_{n,\tau}$ denoted $\mathrm{BER}_{n,\tau}$, should be chosen

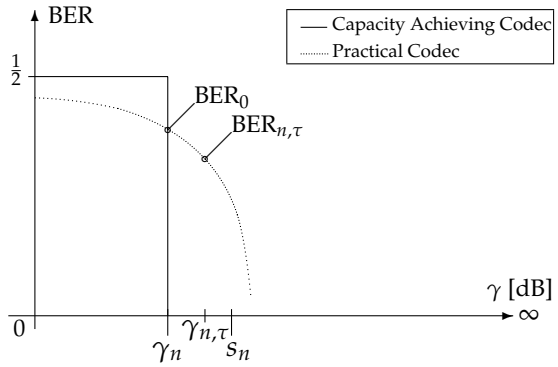


FIGURE 3.7: CSNR, temporal, and switching threshold for codec n for a practical codec. The two curves show the BER-versus-CSNR performance of a typical practical codec and for a theoretical codec achieving capacity at γ_n.

such that the average BER as a function of the predicted CSNR is always less than the target BER. A closer bound for the BER of a practical codec can be found by obtaining the BER for more CSNR values. Especially, for the part of the waterfall curve that decreases slowly ($\gamma < \gamma_n$) the bound will typically not be very close to the true BER-performance.[3]

In the following two subsections bounds for ϵ, for the capacity achieving and practical codecs respectively, are found. The bounds are found requiring that the average BER of codec n as a function of the predicted CSNR, denoted $\overline{\mathrm{BER}}_n(\widehat{\gamma})$, is less than the target BER:

$$\overline{\mathrm{BER}}_n(\widehat{\gamma}) \leq \mathrm{BER}_0. \tag{3.36}$$

Since codec n is only used when $\widehat{\gamma} \in [s_n, s_{n+1}\rangle$ the actual BER of codec n as a function of γ and conditioned on $\widehat{\gamma}$ can be written as

$$\mathrm{BER}_n(\gamma|\widehat{\gamma}) = \begin{cases} \mathrm{BER}_n(\gamma) & \widehat{\gamma} \in [s_n, s_{n+1}\rangle \\ 0 & \text{otherwise.} \end{cases} \tag{3.37}$$

where $\mathrm{BER}_n(\gamma)$ is the BER of codec n on an AWGN channel with CSNR γ.

[3]The resulting upper bound for ϵ, presented in the following, can be closer to the true upper bound using a tighter bound on the BER of codec n. The method will still be same as what is outlined here, and since the bound will vary from codec to codec depending on the shape of the waterfall curve we only use the three values $\frac{1}{2}$, BER_0, and $\mathrm{BER}_{n,\tau}$ in the outline.

The average BER of codec n conditioned on $\widehat{\gamma}$ can now be written as

$$\begin{aligned}\overline{\mathrm{BER}}_n(\widehat{\gamma}) &= \int_0^\infty \mathrm{BER}_n(\gamma|\widehat{\gamma}) f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) d\gamma \\ &= \int_0^\infty \mathrm{BER}_n(\gamma) f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) d\gamma \end{aligned} \tag{3.38}$$

when $\widehat{\gamma} \in [s_n, s_{n+1}\rangle$.

Capacity Achieving Codecs

Using Equation (3.34), the integral over γ in Equation (3.38) can be splitted in two parts and written as

$$\begin{aligned}\overline{\mathrm{BER}}_n(\widehat{\gamma}) &= \int_0^{\gamma_n} \mathrm{BER}_n(\gamma) f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) d\gamma + \int_{\gamma_n}^\infty \mathrm{BER}_n(\gamma) f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) d\gamma \\ &= \int_0^{\gamma_n} \frac{1}{2} \cdot f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) d\gamma + \int_{\gamma_n}^\infty 0 \cdot f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) d\gamma \\ &= \frac{1}{2} \int_0^{\gamma_n} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) d\gamma. \end{aligned} \tag{3.39}$$

Using Equation (3.26) the following inequality is obtained

$$\begin{aligned}\int_0^{\gamma_n} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma}) d\gamma &= \mathrm{P}(\gamma < \gamma_n|\widehat{\gamma}) \\ &\leq \epsilon \end{aligned} \tag{3.40}$$

when $\widehat{\gamma} \in [s_n, s_{n+1}\rangle$. The average BER of codec n is now upper bounded by

$$\overline{\mathrm{BER}}_n(\widehat{\gamma}) \leq \frac{\epsilon}{2}. \tag{3.41}$$

Using the BER demand in Equation (3.36), the maximum value of ϵ is found found from

$$\begin{aligned}\frac{\epsilon}{2} &\leq \mathrm{BER}_0 \\ &\Downarrow \\ \epsilon &\leq 2 \cdot \mathrm{BER}_0. \end{aligned} \tag{3.42}$$

Practical Codecs

In this subsection, practical codecs with a BER performance that can be described by the approximation in Equation (3.35) are considered. As in the

previous subsection, the expression for the average BER in Equation (3.38) can be splitted using a bound on the actual BER for codec n. Using the approximations in Equation (3.35), the average BER can be written as

$$\begin{aligned}\overline{\mathrm{BER}}_n(\widehat{\gamma}) &= \int_0^{\gamma_n} \overbrace{\mathrm{BER}_n(\gamma)}^{\leq \frac{1}{2}} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma + \int_{\gamma_n}^{\gamma_{n,\tau}} \overbrace{\mathrm{BER}_n(\gamma)}^{\leq \mathrm{BER}_0} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma \\ &\quad + \int_{\gamma_{n,\tau}}^{\infty} \overbrace{\mathrm{BER}_n(\gamma)}^{\leq \mathrm{BER}_{n,\tau}} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma \\ &\leq \frac{1}{2}\int_0^{\gamma_n} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma + \mathrm{BER}_0 \int_{\gamma_n}^{\gamma_{n,\tau}} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma \\ &\quad + \mathrm{BER}_{n,\tau} \int_{\gamma_{n,\tau}}^{\infty} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma. \end{aligned} \tag{3.43}$$

From Equations (3.26) and (3.27), we can upper bound the integrals as

$$\int_0^{\gamma_n} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma = \mathrm{P}(\gamma < \gamma_n|\widehat{\gamma}) \leq \epsilon, \tag{3.44}$$

$$\int_{\gamma_n}^{\gamma_{n,\tau}} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma = \mathrm{P}(\gamma_n < \gamma < \gamma_{n,\tau}|\widehat{\gamma}) \leq \epsilon, \tag{3.45}$$

and

$$\int_{\gamma_{n,\tau}}^{\infty} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma = \mathrm{P}(\gamma > \gamma_{n,\tau}|\widehat{\gamma}) \leq 1 - \epsilon < 1. \tag{3.46}$$

The $\overline{\mathrm{BER}}_n(\widehat{\gamma})$ in Equation (3.43) can now be further bound as follows

$$\begin{aligned}\overline{\mathrm{BER}}_n(\widehat{\gamma}) &\leq \frac{1}{2}\overbrace{\int_0^{\gamma_n} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma}^{\leq \epsilon_n} + \mathrm{BER}_0 \overbrace{\int_{\gamma_n}^{\gamma_{n,\tau}} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma}^{\leq \epsilon_n} \\ &\quad + \mathrm{BER}_{n,\tau} \overbrace{\int_{\gamma_{n,\tau}}^{\infty} f_{\gamma|\widehat{\gamma}}(\gamma|\widehat{\gamma})d\gamma}^{<1} \\ &< \frac{\epsilon_n}{2} + \epsilon_n \cdot \mathrm{BER}_0 + \mathrm{BER}_{n,\tau}. \end{aligned}$$

Here the choice of ϵ is made codec dependent and is upper bounded using Equation (3.36):

$$\begin{aligned} &\tfrac{\epsilon_n}{2} + \epsilon_n \cdot \mathrm{BER}_0 + \mathrm{BER}_{n,\tau} \leq \mathrm{BER}_0 \\ &\qquad \Downarrow \\ &\epsilon_n \leq 2\tfrac{\mathrm{BER}_0 - \mathrm{BER}_{n,\tau}}{1+\mathrm{BER}_0} < 2 \cdot \mathrm{BER}_0. \end{aligned} \tag{3.47}$$

Increasing $\gamma_{n,\tau}$ result in a higher bound for ϵ_n (since $\text{BER}_{n,\tau}$ is reduced). Reducing both BER_0 and $\text{BER}_{n,\tau}$ (or equivalently increasing γ_n and $\gamma_{n,\tau}$ the bound approaches $2 \cdot \text{BER}_0$ which is the bound in the case of capacity achieving codecs. That is, for a codec with a steep waterfall curve the bound will be closer to that of a capacity achieving codec then for a codec with a slowly decreasing waterfall curve. Considering the ASE of an ACM scheme it is therefore favorable to use component codecs with very low BERs for CSNRs above the CSNR thresholds. A code class that has very steep waterfall curves are the Gallager codes discussed in the following section and Appendix D.

3.7 An Example System and Optimal Thresholds

In this section the codecs are composed of Gallager codecs and subsequent high level modulation. All codecs have a block length of $M = 200$ symbols. The code rate, r_n, SE R_n, and constellations size and type for each of the $N = 6$ codecs indexed with $n \in \{1, 2, \ldots, N\}$ are listed in Table 3.1. Simulated BER against average CSNR is plotted in Figure D.4. A more comprehensive treatment of the component codecs, how the BER-versus-CSNR relationship is obtained, and the approximation used to describe this relationship are given in Appendix D. For now the codecs are used as codecs with known SEs and BER performance (the approximation of the BER-versus-CSNR is given in Equation (D.22), and the parameters used in the approximation for the BER performance in Figure D.4 is given in Table D.2).

The upper bound on ϵ_n in Equation (3.47) is a function of the BER at the CSNR threshold and the BER at the temporal threshold. In this section the

n	r_n	R_n	Constellation size and type
1	1/2	1	4 QAM
2	2/3	2	8 PSK
3	3/4	3	16 QAM
4	4/5	4	32 QAM
5	5/6	5	64 QAM
6	6/7	6	128 QAM

TABLE 3.1: Codec indices, code rates, SEs, and constellations sizes and types for the 6 example codecs employing Gallager coding.

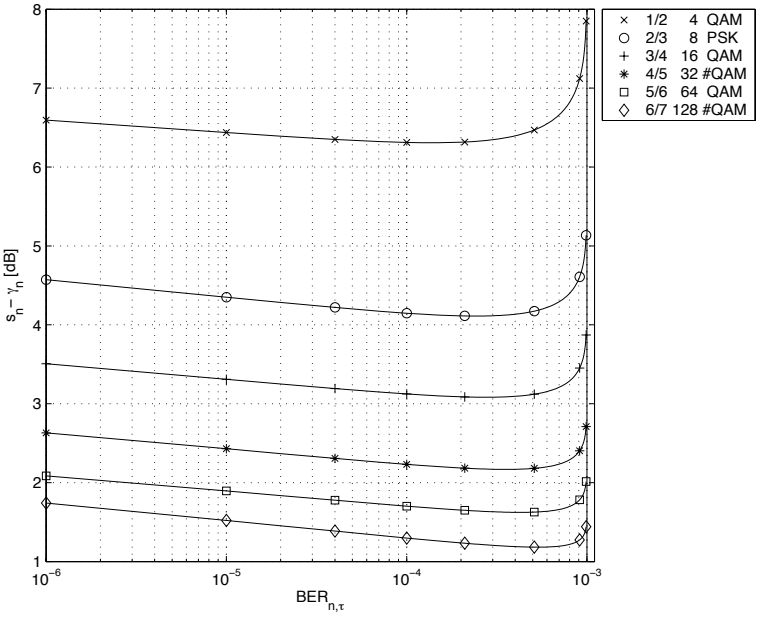


FIGURE 3.8: Plot of the difference $s_n - \gamma_n$ versus $\text{BER}_{n,\tau}$ for the $N = 6$ Gallager codecs.

upper bound is used to obtain the value of ϵ_n. That is,

$$\epsilon_n = 2\frac{\text{BER}_0 - \text{BER}_{n,\tau}}{1 + \text{BER}_0}. \tag{3.48}$$

The CSNR thresholds $\{\gamma_n\}_{n=1}^{N}$ are found from the simulated data. Then, by setting the normalized correlation to $\rho = 0.95$, $\text{BER}_0 = 10^{-3}$, and the average CSNR to $\overline{\gamma} = 10$ dB, Figures 3.8, 3.9, and 3.10 are obtained. The difference between the resulting switching thresholds and the CSNR thresholds are plotted against $\text{BER}_{n,\tau}$ and ϵ_n in Figure 3.8 and Figure 3.9, respectively. As can be seen from the figures there exists an optimal $\text{BER}_{n,\tau}$ (or equivalently ϵ_n) that gives the lowest switching threshold that guarantee an average BER that is less than or equal to BER_0. That is, optimality in the sense of maximizing the ASE.

From the figures it can also be seen that the optimal ϵ_n is reduced and the optimal $\text{BER}_{n,\tau}$ is increased as the codec index increases. This is plotted in the upper and middle sub plot of Figure 3.10. In the lower sub plot in Figure 3.10 the difference $s_n - \gamma_n$ is plotted against n for values of s_n obtained by using the optimal ϵ_n. The simulations were replicated for $\text{BER}_0 = 10^{-4}$ resulting in the stem-plots in Figure 3.11. In both simulations the optimal

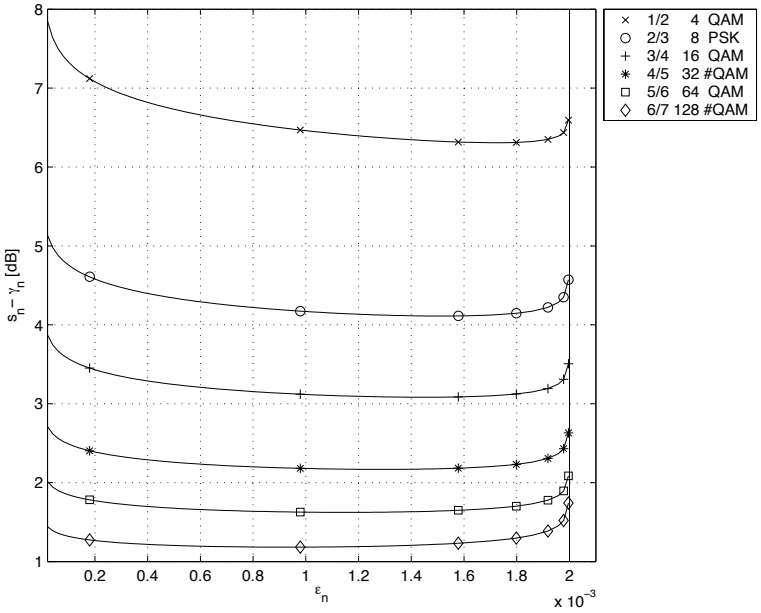


FIGURE 3.9: Plot of the difference $s_n - \gamma_n$ versus ϵ_n for the $N = 6$ Gallager codecs.

ϵ_n is always less than $2 \cdot \text{BER}_0$ and is reduced as n increases. For the example codecs choosing $\epsilon_n = \text{BER}_0$ does not produce thresholds that are significantly higher than the optimal values. However, high rate codecs are more sensitive to the choice in ϵ_n (around the optimal value) than low rate codecs.

Although the results found here are empirical results from a specific set of codecs the results indicate that the choice of ϵ_n should be made codec dependent. It might seem counter intuitive that the optimal value of ϵ_n is reduced for increased n since increased optimal $\text{BER}_{n,\tau}$ (or equivalently reduced $s_n - \gamma_n$) indicates that less protection against predictor errors is needed. It was earlier argued that an *increased* ϵ_n provides less protection against predictor errors. However, this is only true when γ_n is kept constant. From the curves in Section 3.5 (see e.g. Figure 3.4) it can be seen that as γ_n increases s_n is approaching γ_n when ϵ_n is held constant. That is, although the difference $s_n - \gamma_n$ is reduced it might be necessary for a lower ϵ_n, since ϵ_n merely controls the probability of prediction errors resulting in codec mismatch.

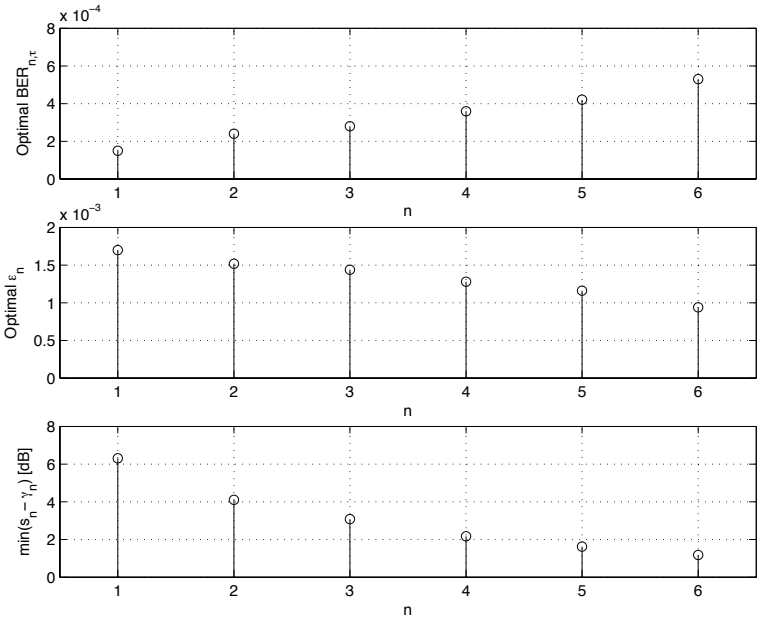


FIGURE 3.10: Optimal values for $\text{BER}_{n,\tau}$ and ϵ_n and the corresponding (minimum) distance $s_n - \gamma_n$ plotted for $n \in \{1, 2, \ldots, 6\}$ and $\text{BER}_0 = 10^{-3}$ for the Gallager codecs.

3.8 Concluding Remarks

In this chapter a technique that can be used to find a set of switching thresholds (from the CSNR thresholds) is outlined. The technique requires knowledge of the expected CSNR and the correlation between predicted and actual CSNR. This knowledge is, in our study, provided by a MAP-optimal predictor, by assuming a known Doppler spectrum of the MFC, and known average CSNR. Clearly, increasing the switching thresholds will improve the BER performance. However, any increase of the switching thresholds will reduce the ASE, since the higher indexed codecs (the ones with highest SE) then have lower probability of being selected. For any combination of ρ and $\overline{\gamma}$ it is also seen that the threshold increase is larger the smaller ϵ is. Keeping ϵ and $\overline{\gamma}$ constant, the reduction in SE will be smaller for higher values of ρ, which again corresponds to lower terminal velocity and/or feedback delay.

The next step towards finding the appropriate switching thresholds was to obtain the maximum allowed probability of the predicted CSNR falling in a higher indexed region than the actual CSNR. This corresponds

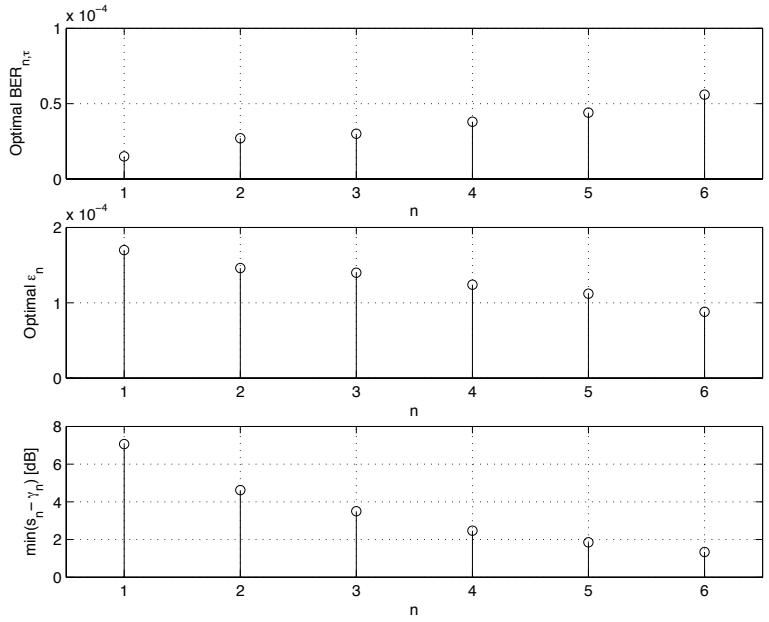


FIGURE 3.11: Optimal values for $\text{BER}_{n,\tau}$ and ϵ_n and the corresponding (minimum) distance $s_n - \gamma_n$ plotted for $n \in \{1, 2, \ldots, 6\}$ and $\text{BER}_0 = 10^{-4}$ for the Gallager codecs.

to choosing reasonable values for ϵ, securing the necessary robustness level. The probability should ideally be chosen such that the ASE is maximized under the constraint of an upper bound on the BER, for any given values of the expected CSNR, the prediction lag, and the degree of mobility. The appropriate value for ϵ in Equation (3.29) can also be made codec dependent. Equivalently, the sufficient value of $\text{BER}_{n,\tau}$ for $n \in \{1, 2, \ldots, N\}$ can be obtained. For the capacity achieving codecs it is the possible to find a close bound on the sufficient protection against prediction errors resulting in a codec mismatch. Using this bound the minimum value of the switching thresholds can be found, resulting in the maximum ASE under a constraint on the average BER. In the case of practical codecs this bound will vary with the BER-versus-CSNR performance of the codecs on AWGN channel. Especially, a practical codec does not perform with zero BER for CSNR values larger than the CSNR threshold. Thus, simply modifying the CSNR threshold according to the technique in Section 3.5 may produce an average BER that is higher than the target BER. Instead, a temporal threshold (which is larger than the CSNR threshold) is utilized to find the switching thresholds that provides an average BER smaller than the target BER. Thus, if the BER

at the temporal threshold can be used to bound ϵ it is again possible to find the minimum switching thresholds that provide an average BER satisfying the BER demand.

The maximum value for the probability of a codec mismatch ϵ has been found for both theoretical codecs achieving capacity and for practical codecs with a more realistic BER performance. It is also shown how the optimal switching thresholds, in the sense of maximizing ASE, for a given set of practical codecs could be found. For the practical codecs it is shown that the choice of ϵ should be made codec dependent, but also that as a rule of thumb (for these codecs) the choice of ϵ should be around BER_0. Note, that codecs with slower falling waterfall curves might require a lower value for ϵ since CSNRs above the CSNR threshold then contribute more bit errors to the overall average BER.

The optimal switching thresholds for the capacity achieving codecs, resulting from the optimization, can be used to find a new upper bound on the MASA. That is, the switching thresholds should be minimized such that high rate codecs have a higher probability of being chosen for transmission and thus increasing the resulting MASA.

Chapter 4

MASA for the Case of Imperfect Channel Knowledge

4.1 Introduction

It was indicated in Chapter 2 that the optimal CSNR thresholds for perfect CSI could be used as a starting point in the design of an ACM scheme. That is, one could attempt to find codecs with SE close to the SE of the capacity achieving codecs, and with BER close to the target BER at these thresholds. However, since imperfect channel knowledge was not taken into account, the ASE performance of a more realistic ACM scheme might not be close to the MASA obtained for the perfect CSI case. Therefore, it is desirable to obtain the MASA also for the case of imperfect channel prediction. The resulting CSNR thresholds can now be used as a more reliable guideline for the designer when assessing system performance.

This work was also presented in [Jetlund et al., 2004a]. The remainder of this chapter is organized as follows: In Section 4.2 an overview of the system model is given, and the expression for the MASA is revised after taking into account imperfect channel knowledge. Section 4.3 develops a unified tool for obtaining the MASA. Results in Section 4.4 show how imperfect channel knowledge affect the MASA and how an outage constraint will affect the design of an ACM scheme. In Section 4.5 the contributions in this chapter are summed up.

4.2 System Model and Problem Formulation

The system model in Figure 1.3 is considered in this chapter. The codecs are (as in Chapter 2) assumed to be capacity achieving codecs—for the pur-

pose of obtaining the MASA. To do this, the correlation properties of the channel predictor, the tool for designing switching thresholds from CSNR thresholds, and the upper bound on the probability of a codec mismatch found in Chapter 3 are all utilized. Now, it is possible to find an upper bound on the ASE of an ACM scheme for the case of imperfect channel knowledge. The method used here is similar to that in Chapter 2. The MASA in Equation (2.16) can now be written as

$$\mathrm{MASA}(\epsilon) = \frac{1}{\ln(2)} \sum_{n=1}^{N} \left[\ln\left(1 + \Psi_\gamma\left(s_n, \epsilon\right)\right) \int_{s_n}^{s_{n+1}} f_{\widehat{\gamma}}(\widehat{\gamma}) d\widehat{\gamma} \right] \tag{4.1}$$

using Equations (3.24) and (3.33). This formula is found by observing that the SE efficiency of the optimal codecs relates to the actual CSNR, that $\gamma_n = \Psi_\gamma\left(s_n, \epsilon\right)$, and that $R_n = \log_2(1 + \Psi_\gamma\left(s_n, \epsilon\right))$.

The MASA can now be optimized with respect to the switching thresholds $\{s_n\}_{n=1}^{N}$ and the corresponding CSNR thresholds can be found using Equation (3.33). Both the MASA and each of the switching thresholds $\{s_n\}_{n=1}^{N}$ will be a function of ϵ. To find the true MASA, ϵ should be as large as possible. In the numerical examples in the following ϵ is set to $\epsilon = 2 \cdot \mathrm{BER}_0$, with $\mathrm{BER}_0 = 10^{-3}$. The reader should however note that this is shown empirically to be a rather conservative (small) value of ϵ and that the true MASA may be increased beyond the numerical results presented here. How to optimize ϵ is still a partially open problem, but it is worth notifying that the theoretical analysis here is independent of the value of ϵ.

4.3 A Theoretical Upper Bound On the MASA

Finding the MASA in Equation (4.1) is done by first obtaining the gradient of the MASA expression with respect to the switching thresholds for a given ϵ. Setting the gradient equal to zero results in N equations that can be used to find expressions for the optimal switching thresholds. Again, optimizing MASA with a condition on the outage probability $P_0 = \mathrm{P}_{\mathrm{out}}$ is done by introducing a Lagrange multiplier, λ, into the optimization procedure as follows:

$$\nabla_{\{s_n\}_{n=1}^{N}} \left(\mathrm{MASA} + \lambda \cdot \mathrm{P}_{\mathrm{out}}\right) = \mathbf{0}. \tag{4.2}$$

Setting $\lambda = 0$ the solution to Equation (4.2) yields the optimal switching thresholds that produce the MASA without any outage constraint. Solving Equation (4.2) (and multiplying the result with $\ln(2)$) produce the follow-

ing set of equations:

$$\begin{bmatrix} f_{\widehat{\gamma}}(s_1)\left(\ln(2)\lambda - \ln\left(1+\Psi_\gamma(s_1)\right)\right) \\ \qquad + \frac{\frac{\delta}{\delta s_1}\Psi_\gamma(s_1)}{1+\Psi_\gamma(s_1)} P_1 \\ \ln\left(\frac{1+\Psi_\gamma(s_1)}{1+\Psi_\gamma(s_2)}\right) f_{\widehat{\gamma}}(s_2) + \frac{\frac{\delta}{\delta s_2}\Psi_\gamma(s_2)}{1+\Psi_\gamma(s_2)} P_2 \\ \vdots \\ \ln\left(\frac{1+\Psi_\gamma(s_{n-1})}{1+\Psi_\gamma(s_n)}\right) f_{\widehat{\gamma}}(s_n) + \frac{\frac{\delta}{\delta s_n}\Psi_\gamma(s_n)}{1+\Psi_\gamma(s_n)} P_n \\ \vdots \\ \ln\left(\frac{1+\Psi_\gamma(s_{N-1})}{1+\Psi_\gamma(s_N)}\right) f_{\widehat{\gamma}}(s_N) + \frac{\frac{\delta}{\delta s_N}\Psi_\gamma(s_N)}{1+\Psi_\gamma(s_N)} P_N \end{bmatrix} = \mathbf{0}, \tag{4.3}$$

where P_n is given in Equation (3.24). Then, by setting $P_0 = \mathrm{P}_{\text{out}}$, solving Equation (3.24) (with $n = 0$) for s_1, and solving equation n in Equation (4.3) for s_{n+1} the following N equations are obtained:

$$s_1 = -\ln(1-\mathrm{P}_{\text{out}})\,\overline{\gamma}\rho \tag{4.4}$$

$$\begin{aligned} s_2 =& s_1 - \overline{\gamma}\rho \\ &\times \ln\left(1 - \frac{1+\Psi_\gamma(s_1)}{\overline{\gamma}\rho\frac{\delta}{\delta s_1}\Psi_\gamma(s_1)} \ln\left(\frac{1+\Psi_\gamma(s_1)}{2^\lambda}\right)\right) \end{aligned} \tag{4.5}$$

$$\begin{aligned} s_n =& s_{n-1} - \overline{\gamma}\rho \\ &\times \ln\left(1 - \frac{1+\Psi_\gamma(s_{n-1})}{\overline{\gamma}\rho\frac{\delta}{\delta s_{n-1}}\Psi_\gamma(s_{n-1})} \ln\left(\frac{1+\Psi_\gamma(s_{n-1})}{1+\Psi_\gamma(s_{n-2})}\right)\right) \\ &\quad \text{for } n \in \{3,4,\ldots,N\}. \end{aligned} \tag{4.6}$$

Now, every switching threshold s_n (for $n > 1$) can be found recursively, and as a result the MASA is a function of s_1 and λ. In order to obtain the optimal MASA ($\lambda = 0$) a search through all possible values of s_1 (or equivalently $\mathrm{P}_{\text{out}} \in \langle 0, 1\rangle$) must be made, and subsequently the s_1 that maximize Equation (2.16) must be chosen. When the outage constraint is employed the value of s_1 is given from Equation (4.4), and a similar search must be made through all possible values of λ.

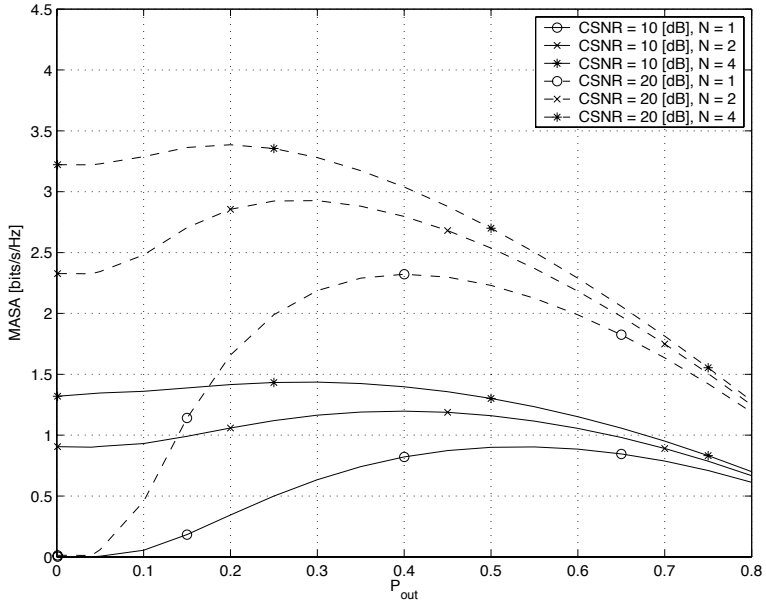


FIGURE 4.1: MASA(ϵ) plotted as a function of P_{out} for $N \in \{1,2,4\}$, $\overline{\gamma} \in \{10,20\}$ dB, $\rho = 0.99$, and $\epsilon = 2 \cdot 10^{-3}$.

4.4 Results and Discussion

In this section the MAP-optimal predictor is assumed used with $K_p = 500$ and $L = 10$. Inserting pilot symbols at equally spaced intervals reduces the MASA(ϵ) by a factor $1 - (L-1)/L = 10\%$. This reduction is taken into account for all plots of MASA(ϵ) as function of s_1, P_{out}, and $\overline{\gamma}$.

In Figure 4.1 the MASA(ϵ) is plotted as a function of P_{out} for or $\overline{\gamma} \in \{10,20\}$ dB, $\rho = 0.99$, and $N \in \{1,2,4\}$. For each of the curves the maximum value corresponds to the unconstrained MASA(ϵ) ($\lambda = 0$). The MASA(ϵ) for a given ϵ increases with both N and $\overline{\gamma}$. That is, for higher values of N an optimally designed ACM system will use some codecs with higher SEs, resulting in an overall increased ASE. Increasing $\overline{\gamma}$ increases the probability of actually having a higher instantaneously CSNR; thus codecs with higher SEs can be used by an ACM scheme and the resulting optimal switching thresholds increase. In Figure 4.2 the same result is plotted as a function of s_1 (for the values of P_{out} in Figure 4.1). In this figure it can be seen that for low values of s_1 and $N > 1$ the MASA reduction from reducing s_1 is negligible. Thus, it can then be concluded that an ACM scheme with a relatively high number of codecs does not trade much of the MASA

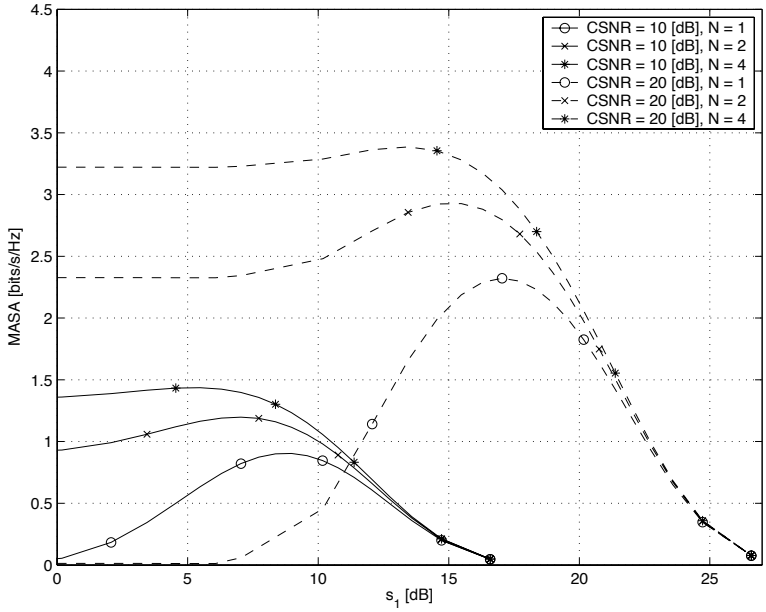


FIGURE 4.2: MASA(ϵ) plotted as a function of s_1 for $N \in \{1, 2, 4\}$, $\overline{\gamma} \in \{10, 20\}$ dB, $\rho = 0.99$, and $\epsilon = 2 \cdot 10^{-3}$.

when introducing an outage constraint (this can also be observed from Figure 4.2).

Increasing the value of N also reduces the unconstrained outage probability. That is, increasing N also allows a capacity maximizing ACM scheme to support codecs with lower SEs, and thus the switching threshold of the first codec s_1 is reduced producing a lower P_{out}. Increasing the average CSNR increases all the optimal thresholds since the probability of predicting a higher CSNR increases. Thus, s_1 increases with $\overline{\gamma}$. Then reducing s_1 according to a decrease in P_{out} results in a larger decrease in MASA(ϵ) as the average CSNR increases.

In Figure 4.3 the correlation was reduced to $\rho = 0.90$. Comparing this figure and Figure 4.1 it can easily be seen that the MASA(ϵ) decreases with decreased ρ for a given ϵ. Also, the value of P_{out} that result in the unconstrained MASA(ϵ) increases with decreased ρ. The sensitivity to changes in P_{out} is less for lower values of ρ. This can be explained by the plots of s_1 and γ_1 against $\overline{\gamma}$ in Figure 4.4 which were generated for normalized prediction lags $jf_mT_s \in \{2 \cdot 10^{-2}, 4 \cdot 10^{-2}\}$. As can be observed from this figure, the switching thresholds increase for reduced ρ. However, the CSNR thresholds are reduced and thus so are the SE of each codec and thus the

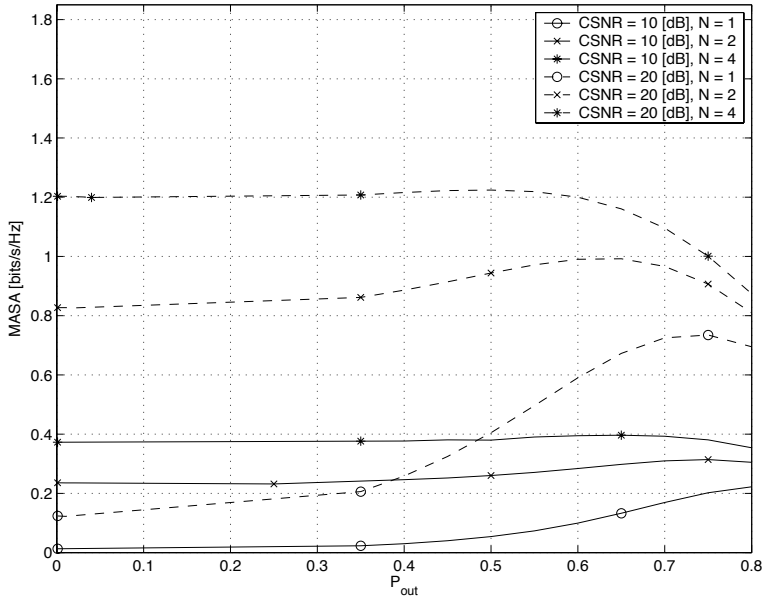


FIGURE 4.3: MASA(ϵ) plotted as a function of P_{out} for $N \in \{1, 2, 4\}$, $\overline{\gamma} \in \{10, 20\}$ dB, $\rho = 0.90$, and $\epsilon = 2 \cdot 10^{-3}$.

overall MASA. Then by reducing s_1 according to a reduced P_{out} the SE of the first code is already very low and does not contribute much to the overall ASE. This is confirmed in Figure 4.5 where the SE of all codecs in an ACM scheme with $N = 4$ codecs are plotted against $\overline{\gamma}$ for the normalized prediction lags in Figure 4.4. In both these figures, results for $\rho = 1$ are plotted for reference. From both figures it can be seen that as $\overline{\gamma}$ increases the difference $s_n - \gamma_n$ is decreased. Also, all thresholds increase with $\overline{\gamma}$, switching thresholds are increased when jf_mT_s increases while the CSNR thresholds are reduced.

Comparing Figures 4.1 and 4.3 it can be seen that the MASA(ϵ) increases more for increasing values of N when γ and ρ increases. That is, when the channel conditions are good and the channel variations are slow the increased gain in terms of SE by adding another codec is higher than for the case of a severely noisy and fast varying channel. Since the MASA is not sensitive to an outage constraint for low values of ρ and $\overline{\gamma}$ it might therefore be sufficiently to employ only a small number of codecs.

Finally, in Figure 4.6 the optimal MASA(ϵ) is plotted against $\overline{\gamma}$ for different values of jf_mT_s with both the MASE and MASA(ϵ) with $\rho = 1$ for reference. It can now be seen how the MASA is reduced as the correlation is reduced. When the CSNR is high (close to 20 dB) the MASA for $\rho < 1$

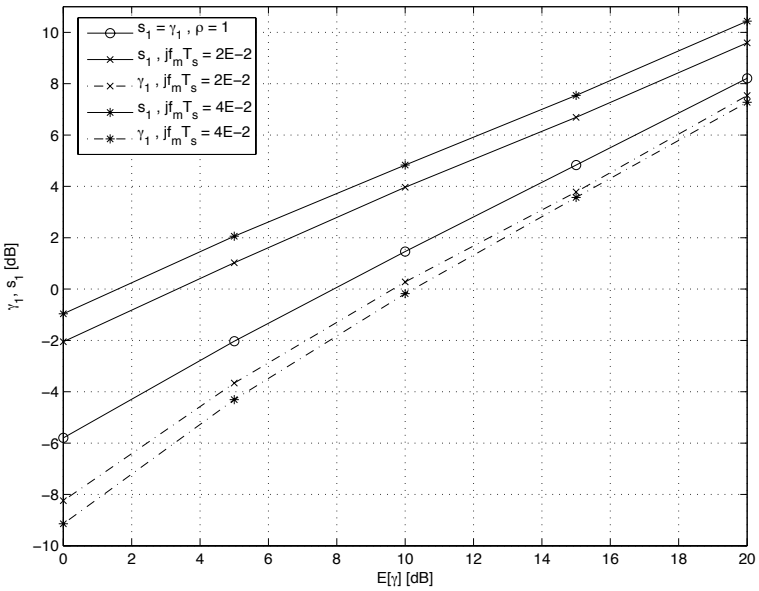


FIGURE 4.4: The optimal values for s_1 and γ_1 plotted as a function of $\overline{\gamma}$ for $N = 4$ and different values of the normalized prediction lag.

is closer to the MASA for $\rho = 1$ than for low CSNRs. Note that in Figure 4.6, both the MASE and the MASA (for all values of ρ) are affected by the ASE reduction caused by the inserted pilots.[1] The results presented here show that the MASE is not a very good reference for the ASE of an ACM scheme since it is not possible to get close to the MASE even for high CSNRs. Note, this conclusion is based on the results presented for a given ϵ. Choosing a less conservative (larger) ϵ would increase the MASA towards the MASE, since the switching thresholds are reduced. However, even for the case of switching thresholds equal to the CSNR thresholds, the MASE will be much higher than the MASA unless N is chosen very high (as seen in Chapter 2).

4.5 Concluding Remarks

We have found bounds for the ASE of a rate-adaptive communication scheme employing a limited number of codecs when imperfect channel knowledge is taken into account. The correlation between predicted and actual

[1] It can be argued that it is unfair to reduce the MASE. However, this is done to be able to have a fair comparison between the MASE and the MASA at the given channel quality, which is a result of the information provided by the pilots.

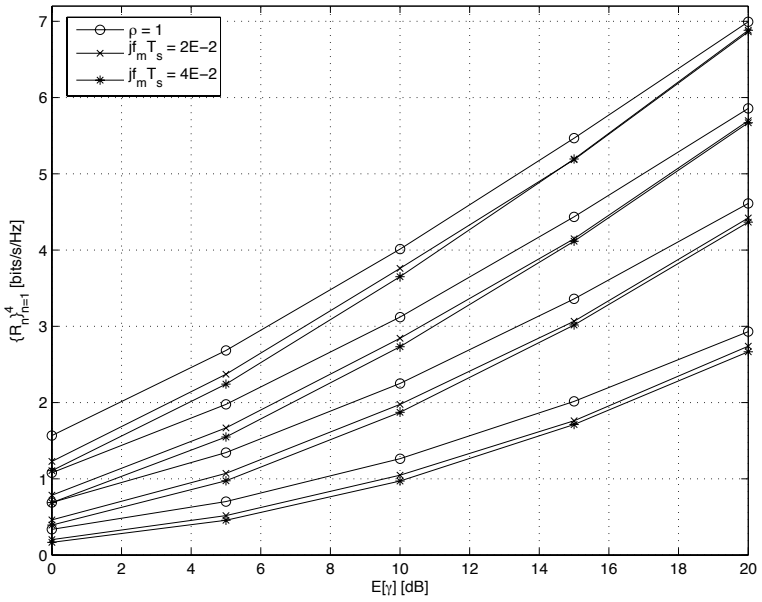


FIGURE 4.5: Resulting SE for all codecs resulting from the optimal CSNR thresholds plotted as a function of $\overline{\gamma}$ for $N = 4$ and different values of the normalized prediction lag.

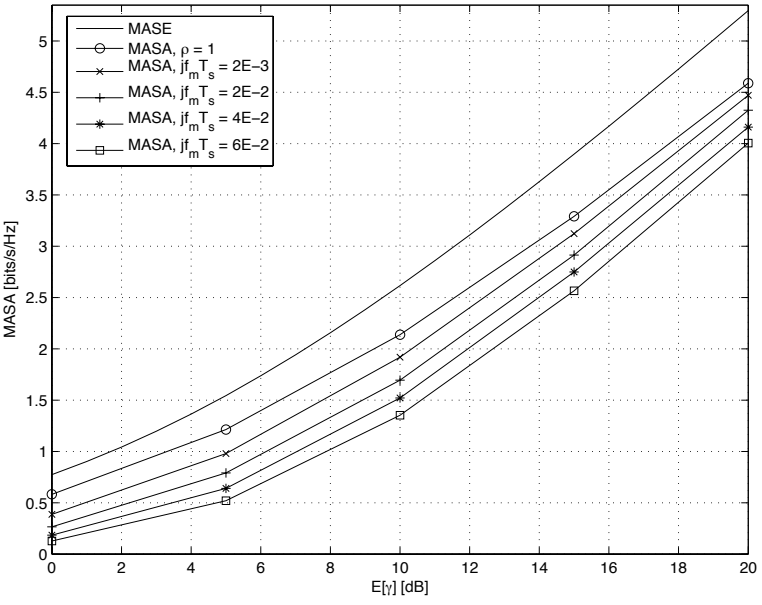


FIGURE 4.6: MASA(ϵ) for different values of the normalized prediction lag plotted against average CSNR, for $\epsilon = 2 \cdot 10^{-3}$.

values of the fading envelope affects the bound in the sense that the bound is reduced as the correlation is reduced. Likewise, the bound increases with increasing average CSNR. It should be noted that the correlation between predicted and actual correlation itself depends on the average CSNR as well as the choice and complexity of the predictor employed by the ACM scheme.

All the numerical results presented in this chapter are obtained using a specific ϵ. It has been argued in this and the previous chapter that a low ϵ result in conservative design which reduces the MASA. Thus, in order to find the true MASA the optimal ϵ should be found. However, as was shown in the previous chapter the choice of ϵ is codec dependent. For the chosen ϵ the results presented here show that as long as the average CSNR and the correlation between predicted and actual CSNR is sufficiently high there is a substantial gain in using a high number of codecs in an ACM scheme. Finally, it is observed that the MASE is an impractical upper bound, and that the MASA (at least for $\rho = 1$) should be used as an ultimate upper bound on the ASE.

Chapter 5

Simulation of Adaptive Gallager Coded Modulation on Correlated Fading Channels

5.1 Introduction

In this chapter simulation results for a practical ACM scheme performing on a frequency-flat MFC with a Rayleigh distributed envelope are presented. The system employs PSAM, MAP-optimal channel estimation, and MAP-optimal channel prediction. Results from this chapter regarding the theoretical ASE performance for the transmission scheme were investigated in [Jetlund et al., 2002]. The first simulation results for the average BER and ASE of our scheme were presented in [Jetlund, Øien, and Hole, 2003a], for the case of both perfect channel prediction and perfect channel estimation. The correlation properties of the predicted and actual CSNR, as well as the average BER and ASE performance of the system after introducing imperfect channel prediction, were investigated in [Jetlund, Øien, Holter, and Hole, 2003c]. In [Jetlund, Øien, Hole, and Holter, 2003b] imperfect channel estimation was introduced and the resulting performance in terms of average BER and ASE presented.

As concluded in the Chapter 3, it is desirable to have codecs with rapidly decreasing BER-versus-CSNR curves (for CSNRs higher than the CSNR threshold) in order to maximize the ASE. It is also desirable to employ codecs that perform as close as possible to the channel capacity. As indicated by Shannon's Channel Coding Theorem [Shannon, 1948], the best performance is achieved when using a code consisting of very long codewords. In practical systems the codewords have to be short enough for complex-

ity, buffering, and delay not to represent a problem. For ACM schemes, the codeword length must be restricted to the time over which we can reasonably assume that the CSNR stays within one CSNR region. Gallager codes[1] are well known for their extremely good performance on AWGN channels. Results from e.g. [MacKay and Hesketh, 1997] show that Gallager codes perform quite close to the capacity even for relatively short (and practical) block lengths using a *low-latency* decoder. The Gallager codes in Section 3.7 (also see Section D.5) will be used as components in the codecs in our simulation system.

In this chapter, simulation results for the performance of the predictor presented are used both to confirm the theoretical properties of the channel predictor, and to gain more insight into the behavior of our ACM scheme. The simulated performance of the ACM scheme is measured through average BER, outage probability, and ASE. The remainder of the chapter is organized as follows: First, in Section 5.2 the simulation model is outlined. In Section 5.3 the theoretical ASE performance of the simulated scheme is shown, and by means of an example it is shown how the choice of codecs affect the system performance. The correlation properties of the predictor used are investigated in Section 5.4. The strategies used for selecting codecs, in our simulations, and the resulting switching thresholds are discussed in Section 5.5. In Section 5.6 simulation results for the ASE, average BER, and outage probability are presented for three different methods (codec selection strategies) of obtaining the switching thresholds. As a starting point, a CSS in which the CSNR thresholds are used as switching thresholds is employed. Next, the switching thresholds are modified by simply adding a constant (in dB) to the existing thresholds. The last CSS uses the technique outlined in Chapter 3 to obtain the switching thresholds. Finally, the contributions from this chapter are summed up in Section 5.7.

5.2 System Model

We shall consider the baseband system in Figure 5.1. All bold-faced symbols in the figure are column vectors.

The source generates uniformly distributed information bits. Codec n receives a binary information vector $\mathbf{s}$ of length g_n. The g_n information bits are encoded by a binary block code (a Gallager code) into a binary vector $\mathbf{t}$

[1]Gallager codes were presented as *Low Density Parity Check Codes* (LDPC codes) in [Gallager, 1963]. The change of name is simply a tribute to the inventor R. J. Gallager. The Gallager codes was not a subject for further research until recently when they were rediscovered by D. J. C. MacKay [MacKay and Hesketh, 1997].

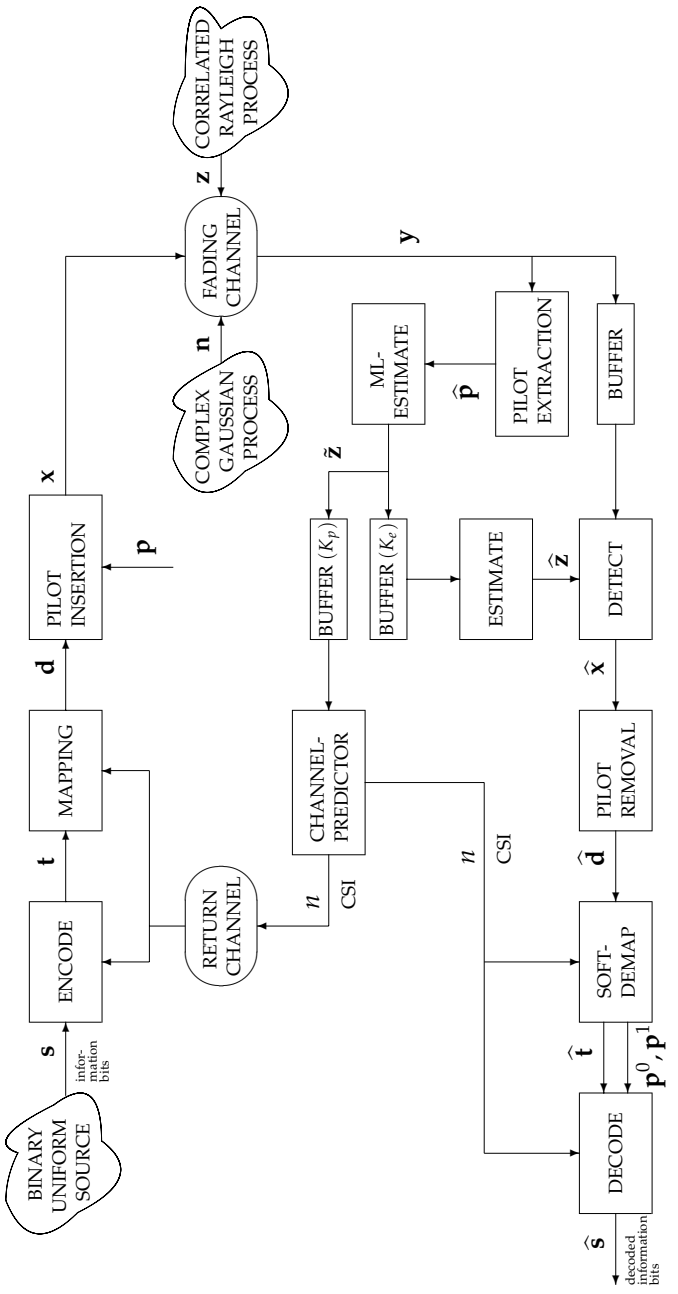


FIGURE 5.1: Simulation system.

of length q_n. The code rate of codec n in the simulation setup is chosen to be

$$r_n = \frac{g_n}{q_n} = \frac{n}{n+1} \text{ [information bits per code bit].} \tag{5.1}$$

The size of the QAM/PSK constellation used by the modulator is denoted S_n and is here defined to be $S_n = 2^{n+1}$. The modulator is designed such that the resulting channel symbol block, $\mathbf{d}$, is always of length M. That is, the constellation size S, channel symbol block length M, code vector length, and information vector length satisfy

$$\frac{q_n}{\log_2(S_n)} = \frac{g_n}{r_n \log_2(S_n)} = M, \quad \text{for } n \in \{1, 2, \ldots, N\}. \tag{5.2}$$

The length of the information and code vectors, g_n and q_n, and the constellation size S_n will vary with the specific code, but to simplify notation the subscript index n is omitted in the following unless there is a need to distinguish between the different codecs.

Each channel symbol can uniquely represent $\kappa = \log_2 S$ bits. That is, each channel symbol $\zeta_i, i \in \{1, 2 \ldots, S\}$ corresponds to a κ-tuple of code bits, and channel symbols (in $\mathbf{d}$) are generated by mapping consecutive sub-vectors of length κ from $\mathbf{t}$ to channel symbols in the constellation. The mapping to channel symbols is done using Gray mapping [Stüber, 2001], i.e. the κ-tuple corresponding to two horizontal or vertical neighboring symbols in the complex plane differ at only one position (for an example see Figure D.2). The modulation/demodulation process is further described in Section D.3 in Appendix D.

The piloting is implemented such that pilots (the pilot vector is denoted $\mathbf{p}$) are inserted periodically into a block of channel symbols of length M, producing a piloted block, $\mathbf{x}$, of length M'. Note that, both $\mathbf{d}$ and $\mathbf{x}$ are complex vectors representing complex channel symbols. An additional restriction used here is that M' should be constant for a given M and a given pilot spacing L. This is fulfilled when

$$M \bmod L = 0. \tag{5.3}$$

For simplicity all pilot symbols have the same value $[\mathbf{x}]_k = a_p$ and equal power is used for pilots and information symbols, i.e. $|a_p| = \sqrt{P}$.

Communication is simulated by multiplying the transmitted complex-valued vector $\mathbf{x}$ by a complex fading envelope vector, $\mathbf{z}$, and adding a complex valued additive white Gaussian noise vector, $\mathbf{n}$. The received vector becomes

$$\mathbf{y} = \mathbf{z} \circ \mathbf{x} + \mathbf{n} \tag{5.4}$$

where the Hadamard product "$\circ$" represents component-wise multiplication.

The average power gain of the fading envelope is set to $\Omega_p = 1$, the transmit power is $P = 1$ W, and the variance of the AWGN $\sigma_w^2 = N_0 B$ is varied such that the expected CSNR is $\overline{\gamma} \in [1, 25]$ dB. The vector $\mathbf{z}$ represents a sampled correlated random process, where the correlation is dependent on the maximum Doppler frequency and thus terminal velocity. The fading simulator is described in greater detail in Appendix C.

At the receiver the $M' - M$ noisy pilots, $\widehat{\mathbf{p}}$, are extracted form the received symbol vector $\mathbf{y}$. For each of the received pilot symbols a ML-estimate of the fading (see Equation (3.5) in Section 3.2) is calculated producing a vector $\tilde{\mathbf{z}}$ of length $M' - M$. The result is shifted into two buffers. One of the buffers stores values to be used in prediction of the fading envelope, which is used in the codec selection. The other buffer stores values to be used in estimation of the fading envelope, which is used in the symbol detection. The complete received vector $\mathbf{y}$ also has to be buffered since channel estimation and detection are performed after pilot symbols in the future are received. Detection is performed by simply dividing the received channel symbols with the MAP-estimated fading. After the detection the symbols at the pilot symbol time instants are removed from the detected sequence, $\widehat{\mathbf{x}}$, producing an estimated vector $\widehat{\mathbf{d}}$ of the information carrying channel symbols transmitted.

Using MAP-prediction an estimate of the CSNR at the appropriate future time instant is obtained. The codec index n, the relevant CSI for our system, is chosen by finding which CSNR region the predicted CSNR falls in. This information is passed on to the demodulator and decoder in the receiver. The CSI is also transmitted to the transmitter on the error-free return channel where it is used to choose the correct codec.

Since $\mathbf{x}$ and $\mathbf{n}$ are uncorrelated and the modulation is memoryless, the received vector can be demodulated symbol-by-symbol producing a hard decision, $\widehat{\mathbf{t}}$, on the received coded information (hard-demodulation). The Gallager decoder uses this vector and the probabilities of each of the bits being demapped with error, $\mathbf{p}^0$, or without error, $\mathbf{p}^1$. These probabilities are calculated by the demodulator (this "soft-demodulation" is described in Section D.3). The Gallager decoder uses the iterative decoding described Section D.4. The decoding algorithm ends after either successfully decoding a legal codeword or after reaching a maximum number of iterations, I_N. The decoded information is represented by the vector $\widehat{\mathbf{s}}$ in Figure 5.1.

Relevant simulation parameters are listed in Table 5.1. The terminal velocity (equivalently normalized Doppler frequency), prediction lag and the

Parameter	Symbol	Value(s)
Carrier frequency	f_c	5.4 GHz
Symbol duration	T_s	4 μs
Pilot spacing	L	10 symbols
Codec block length	M	200 symbols
Block length (including pilots)	M'	220 symbols
Number of codecs	N	5
Average Power Gain	Ω_p	1
Average Transmit Power	P	1 W
Pilot transmit power	$\lvert a_p \rvert$	1 W
Average CSNR	$\overline{\gamma}$	$[0, 25]$ dB
Prediction lag	j	$\{0, M', 2M', \ldots\}$ symbols
Terminal velocity	v	$\{1, 2, \ldots\}$ m/s
Target BER	BER_0	10^{-3}
Predictor length	K_p	500
Estimator length	K_e	1000
Decoder iterations	I_N	100
Constellation size	S	{4, 8, 16, 32, 64}

TABLE 5.1: Simulation parameters.

average CSNR are varied to obtain simulation results under different channel conditions. It is assumed that each block in the system model knows these parameters.

5.3 Theoretical ASE Performance - Selecting Codecs for an ACM Scheme

The theoretical ASE for an ACM system can be found from Equation (1.5). Assuming that the prediction of future channel states is perfect and that the CSNR thresholds are used as switching thresholds the ASE is given as

$$\text{ASE}(\overline{\gamma}) = \sum_{n=1}^{N} R_n \cdot P_n = \sum_{n=1}^{N} R_n \cdot \int_{\gamma_n}^{\gamma_{n+1}} f_{\gamma|\overline{\gamma}}(\gamma|\overline{\gamma}) d\gamma \quad \text{[bits/s/Hz]}, \qquad (5.5)$$

where $f_{\gamma|\overline{\gamma}}(\gamma|\overline{\gamma})$ denotes the PDF of γ for a fixed expected CSNR, $\text{E}[\gamma] = \overline{\gamma}$. Using the first N codecs in Table D.2 as components in the ACM scheme, the ASE versus $\overline{\gamma}$ plots for $N \in \{1, 2, \ldots, 6\}$ in Figure 5.2 are obtained. It can be seen that the ASE approaches the SE of the Nth codec as $\overline{\gamma}$ increases

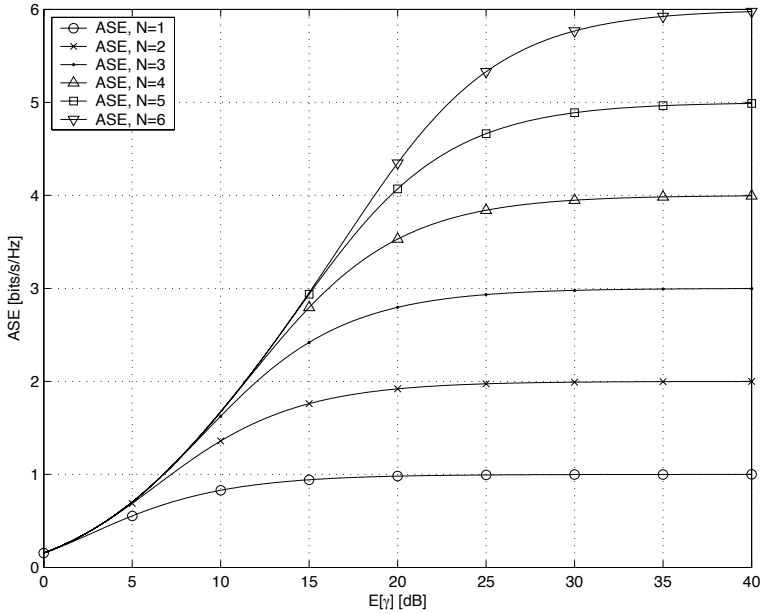


FIGURE 5.2: Theoretical ASE on a Rayleigh MFC when perfect prediction is assumed. The first N codecs in Table D.1 are employed in order to obtain the ASE curves for $N \in \{1, 2, \ldots, 6\}$.

(since the SE of codec n in the results presented here is $R_n = n$). The simulation system is specified to use only $N = 5$ codecs, and a choice has to be made regarding which codecs to employ. In the following four different subsets of the available codecs are used in an example to indicate some of the practical implications of choosing a subset of available codecs.

The ASE obtained by using the $N = 5$ first codecs is given by the curve in Figure 5.2 with the square labels. Using these first five codecs will be referred to as configuration 0. Any other configuration is found by removing one of the codecs and adding the last codec in Table D.2 (codec $n = 6$). In the upper plot in Figure 5.3 the extra gains in terms of ASE from using three other configurations are plotted. Configuration 1 uses the last 5 codecs, i.e. $n \in \{2, 3, 4, 5, 6\}$, configuration 2 uses $n \in \{1, 2, 4, 5, 6\}$, and configuration 3 employs every codec except the fifth, $n \in \{1, 2, 3, 4, 6\}$.

As can be seen from the figure, all configurations have gains approaching 1 bit/s/Hz for high values of $\overline{\gamma}$. However, for low values of the average CSNR the ASE is reduced. For configuration 1 the reduction in ASE is approximately 0.4 bits/s/Hz at $\overline{\gamma} \approx 5$ dB. The ASE of configurations 2 and 3 are approximately equal to configuration 0 up to $\overline{\gamma} = 5$ dB and $\overline{\gamma} = 10$ dB,

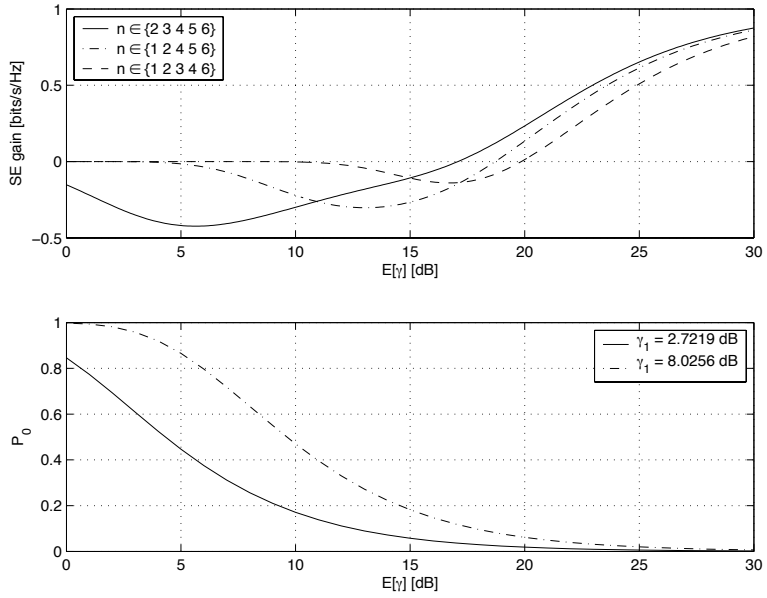


FIGURE 5.3: The upper plot shows the gained ASE by using permutations of the six available codecs other than the first $N = 5$ codecs. The lower plot shows the outage probability when the first and the second of the available codecs govern the outage.

respectively. When the average CSNR is sufficiently high (above approximately 15 dB) configuration 1 has the highest ASE, which is as expected since this configuration uses all the codecs with the highest SEs.

However, since the first codec in an ACM scheme also governs the outage probability, configuration 1 (which does not use codec $n = 1$) will have a higher outage probability than the other configurations. This is shown in the lower plot in Figure 5.3, where the outage probability is plotted against $\overline{\gamma}$ for outage levels (the CSNR threshold of the first codec used) of 2.7219 dB and 8.0256 dB (the CSNR thresholds for our codecs are listed in Table D.3). The outage probability is increased very much for low values of $\overline{\gamma}$ when the first of the available codecs is not used (e.g. for $\overline{\gamma} = 5$ dB the outage probability is approximately doubled). Only for very high average CSNRs does the outage probability of configuration 1 become approximately as low as the outage probability for the other three configurations (0, 2, and 3).

The results show the trade-off between maximizing the ASE and managing the outage probability, but also the ASE trade-off when using different codec sets at different average CSNR values. Throughout this chapter configuration 0 will be used from now on. This choice is made to keep

the outage probability low and to have the highest ASE for most average CSNRs.[2]

In Figure 5.4 the ASE of the chosen configuration is plotted along with the MASE of a Rayleigh fading channel, the MASA for $N = 5$ codecs (assuming perfect prediction), and the SE of the highest indexed codec ($n = 5$). For any value of $\overline{\gamma}$ there is a reduction in SE when comparing the ASE

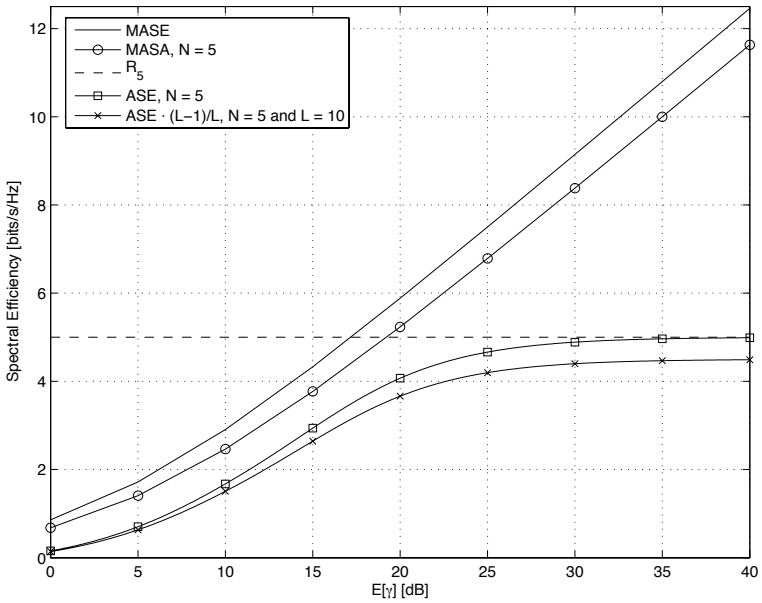


FIGURE 5.4: The ASE obtained by using the first five codecs available, the MASE, and the MASA for $N = 5$ and $\rho = 1$ plotted against $\overline{\gamma}$. The reduced ASE due to a pilot spacing of $L = 10$ symbols and the SE of the fifth codec R_5 are also plotted.

to either the MASA or the MASE. However, for CSNR values lower than the CSNR at the intersection between the SE of the fifth codec R_5 and the MASA, the difference between the MASA and the ASE is approximately equal to or less than 1 bit/s/Hz. The reduced ASE resulting from using PSAM with $L = 10$ symbols is also plotted in Figure 5.4 such that the simulated performance of the ACM scheme (presented in Section 5.6) can be compared to the theoretical ASE. In this case, the maximum spectral efficiency for the ACM scheme is reduced to

$$R_5 \cdot \frac{L-1}{L} = 4.5 \text{ bits/s/Hz}.$$

[2]Note, there exists other configurations, but in the discussion here we have assumed that there are only four configurations.

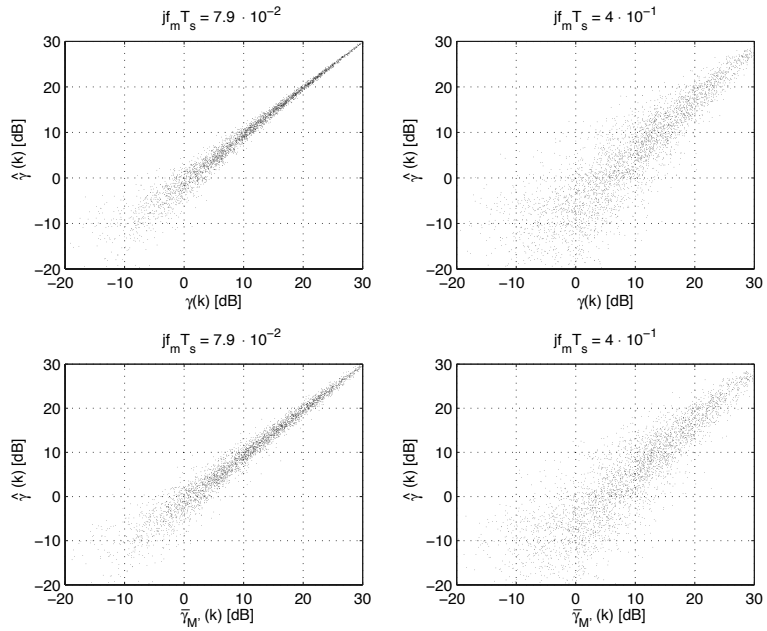


FIGURE 5.5: Scatter plots of the predicted CSNR versus (in the upper plots) actual CSNR and (in the lower plots) averaged actual CSNR, for varying normalized prediction lags.

5.4 Predictor Performance

The upper plots in Figure 5.5 show scatter plots of outcomes of the predicted CSNR versus the actual CSNR. In this experiment, $\gamma(k)$ is the actual CSNR for the first symbol in a block of M' channel symbols, and $\widehat{\gamma}(k)$ is the corresponding predicted value.[3] The scatter plots show 100 simulated points per $\overline{\gamma} \in \{1, 2, \ldots, 25\}$ dB. In the lower plots the same values of the predicted CSNR are plotted as function of the empirical block wise average of the CSNR for the entire block of channel symbols ($\widehat{\gamma}(k)$ versus $\overline{\gamma}_{M'}(k)$),

$$\overline{\gamma}_{M'}(k) = \frac{1}{M'} \sum_{l=k}^{k+M'-1} \gamma(l). \tag{5.6}$$

In both the upper and lower plots the normalized prediction lag is varied, between the sub plots.

[3] k is a discrete time index. In Figure 5.5 the values of $\gamma(k)$, $\widehat{\gamma}(k)$, and $\overline{\gamma}_{M'}(k)$ are found for $k \in \{1 \cdot M', 2 \cdot M', \ldots\}$.

Perfect prediction would yield $\widehat{\gamma}(k) = \gamma(k)$, thus all points in the scatter plots showing $\widehat{\gamma}(k)$ versus $\gamma(k)$ would be located on a *reference line* with a positive slope equal to one (going through the point $(0, 0)$). If constant fading during a block of channel symbols was assumed the upper and lower plots (for a constant value of jf_mT_s) would be identical.

For a relatively small normalized prediction lag (see the plot in Figure 5.5 for $jf_mT_s = 7.9 \cdot 10^{-2}$), the points in the scatter plot are quite close to the reference line when $\gamma(k) > 0$ dB. As we could expect, the largest prediction errors occur at the deepest fades of the fading envelope $\gamma(k) < 0$ dB. By increasing jf_mT_s, larger prediction errors can be observed from the scatter plots.

Comparing the upper and lower plots we observe that the prediction error increases somewhat, although not dramatically. This is explained by the fact that the fading is highly correlated (not far from constant over a block) due to the low normalized prediction lags used. For larger prediction lags (caused by either longer delays or higher degree of mobility) it can be observed (see Figure 5.5 for $jf_mT_s = 4 \cdot 10^{-1}$) that the prediction tends to be lower than the actual CSNR, i.e. the predictor tends to underestimate the CSNR.

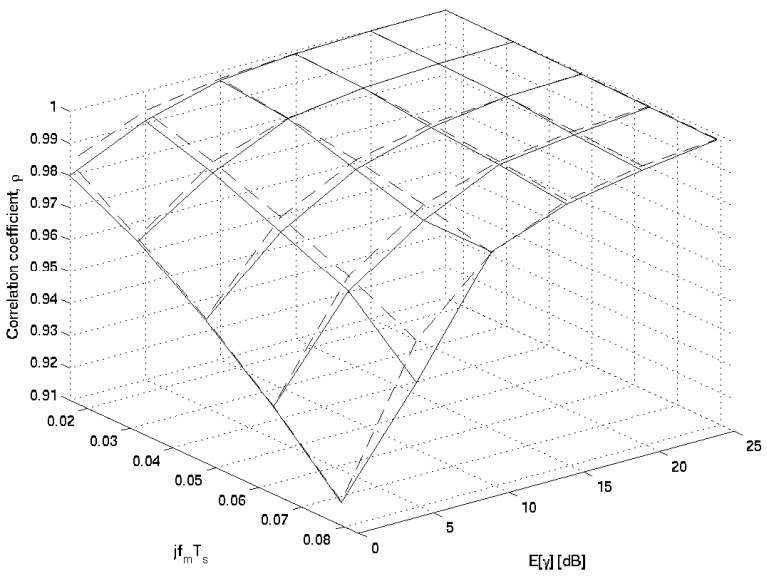


FIGURE 5.6: Simulation results for the normalized correlation ρ between predicted CSNR and actual CSNR are plotted against average expected CSNR and normalized prediction lag.

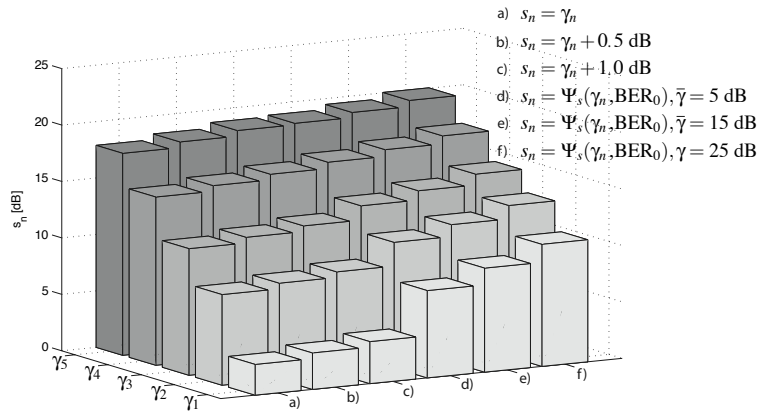


FIGURE 5.7: Switching thresholds for the codecs used in the simulations; a) CSNR thresholds used as switching thresholds, b-c) CSNR thresholds plus a constant (in dB) used as switching thresholds; d-f) switching thresholds found for $\overline{\gamma} \in \{5, 15, 25\}$ dB using $\epsilon = \mathrm{BER}_0$.

Simulation results for the normalized correlation ρ between predicted CSNR and actual CSNR are plotted against average expected CSNR and normalized prediction lag in Figure 5.6. The theoretical correlation is plotted for reference. The figure shows how ρ increases as $\overline{\gamma}$ increases and jf_mT_s decreases. As can be seen from the figure, the simulated correlation coefficient is less or equal to the theoretical correlation in Equation (3.20). From this and Equation (3.18), we could expect that the predicted CSNR on average is lower than the actual CSNR, as was also observed from the scatter plots.

The minor differences between the simulated and theoretical ρ can be explained as follows; The simulated ρ was estimated using the normalized covariance between predicted CSNR values and the corresponding actual CSNR values. The results in Figure 5.6 was generated using 500 pairs of predicted and actual CSNR values per simulation point. This might not be a high enough number to provide high confidence for each simulation point. We still have a high confidence in the results since the shape of the simulated ρ is very close to the shape of the theoretical ρ.

5.5 Codec Selection Strategies

In the following section, the resulting ACM performance is presented for different CSSs, that is, different choices of the switching thresholds:

CSS 1: $s_n = \gamma_n$, $n \in \{1, 2, \ldots, N\}$—original thresholds are used as switching thresholds.

CSS 2: $s_n = \gamma_n + \delta_s$, $n \in \{1, 2, \ldots, N\}$—A constant δ_s (in dB) is added to the CSNR thresholds to obtain the switching thresholds.

CSS 3: $s_n(\epsilon) = \Psi_s(\gamma_n, \epsilon)$—the switching thresholds are obtained from the CSNR thresholds with the probability ϵ of a codec mismatch.

It will be shown, later in this chapter, that using the value for ϵ obtained in Chapter 3 results in a conservative design, but also that the performance of the ACM scheme can be improved by using a higher value of ϵ. This is indicated in Figures 5.7 and 5.8 which show the resulting switching thresholds, for our chosen codecs, using CSS 1, 2, and 3. In these bar-plots, the switching thresholds are plotted in rows for one specific codec (marked with $\gamma_1, \gamma_2, \ldots \gamma_5$), and the columns (marked with letters a-f) represent the CSS used; In both figures the bars in column a) are obtained using CSS 1. Bars in columns b) − c) are obtained using CSS 2 with δ_s equal to 0.5 and 1.0 dB, respectively. The last three columns, d) − f), are obtained for CSS 3 with $\overline{\gamma} \in \{5, 15, 25\}$. In Figure 5.7 CSS 3 with $\epsilon = \text{BER}_0$ is used while $\epsilon = 200\text{BER}_0$ is used in Figure 5.8.

When ϵ is set to the value in Chapter 3, the switching thresholds are always larger than or equal to the CSNR thresholds (see Figure 5.7). Especially, the lower switching thresholds are increased much when CSS 3 is used. This will reduce the average BER, but the resulting ASE may be reduced too much if the average BER is much lower than the target BER. In Figure 5.8 the choice of ϵ is increased to 200BER_0 (this choice of ϵ will become evident at the end of this chapter). It can now be observed that the switching thresholds obtained using CSS 3 are reduced.

5.6 ACM Performance

Monte Carlo Simulations

The average BER, ASE, and outage probability are performance measures for an ACM scheme. In this section simulation results for the ACM scheme with $N = 5$ component codecs (utilizing Gallager codes) are presented. The simulation results were obtained using Monte Carlo simulations. The simulation software uses the parameters in Table 5.1, and the average BER, ASE and outage probability has been simulated for varying normalized prediction lags and average CSNRs. In order to terminate the simulations the following definition of events were used:

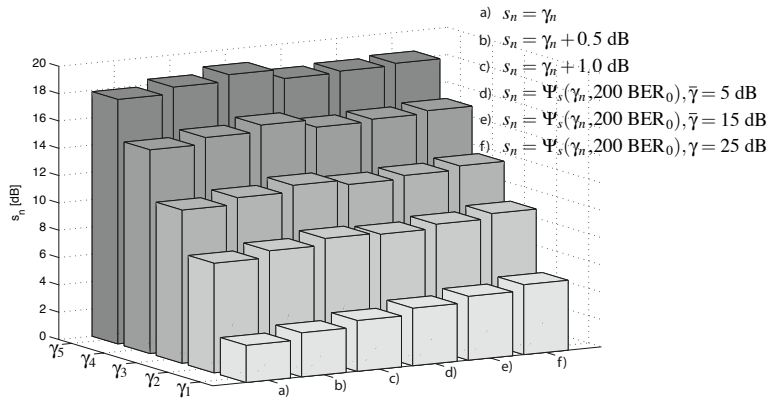


FIGURE 5.8: Switching thresholds for the codecs used in the simulations; a) CSNR thresholds used as switching thresholds, b-c) CSNR thresholds plus a constant (in dB) used as switching thresholds; d-f) switching thresholds found using $\epsilon = 200 \cdot \mathrm{BER}_0$ with $\overline{\gamma} \in \{5, 15, 25\}$ dB.

- A *block error*: The event of a block of information received at the decoder being decoded with error. There are two events leading to a block error; the decoder reaches the maximum number of iterations I_N and is unable to find a legal codeword, or the decoder finds a legal codeword which is different from the transmitted codeword.[4]

- An *outage*: The event of the channel not being used and that the ACM scheme waits $M' \cdot T_s$ seconds[5] before checking if the CSNR on the channel is high enough to be used for transmission.

- BER above the *minimum average BER*: The experienced average BER is larger than or equal to the minimum average BER (which is set to a low value much smaller than the target BER).

Using these definitions the Monte Carlo simulations are terminated for each simulation point (given values of jf_mT_s and $\overline{\gamma}$) when one of the following conditions are met:

I. The total number of block errors is larger than or equal to 10^3.

II. The number of outages experienced is larger than or equal to $5 \cdot 10^6$.

[4] A codeword consists of the coded bits transmitted in a block of channel symbols of length M.

[5] Recall that M' is the number of channel symbols in the transmitted block (including pilots).

III. The experienced average BER is larger than zero and less than or equal to the minimum average BER, which is set to 10^{-6}, after a minimum of 10^6 channel blocks are transmitted. That is, the simulated average BER is set to the minimum average BER when at least one bit error has occurred after transmitting 10^6 blocks of channel symbols.

The first condition guarantees a high confidence in the simulated ASE, average BER, and outage probability. The two last conditions are introduced to reduce simulation times. If one of the two last conditions are met only the simulation results for the ASE and the outage probability are presented here and the average BER is treated as unknown, but smaller than the minimum average BER.

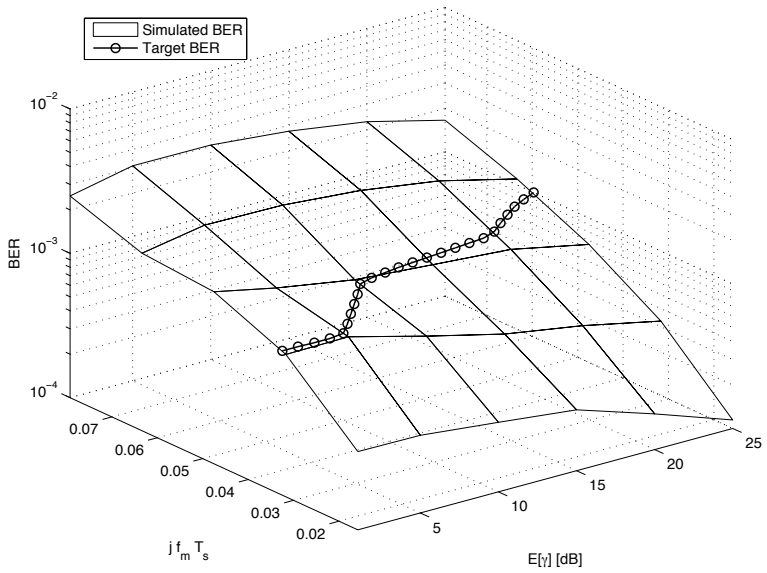


FIGURE 5.9: The average BER of CSS 1 plotted for $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$ and $\overline{\gamma} \in [1, 25]$ dB. The circle-markers show simulation points with average BER equal to the target BER.

Average BER, ASE, and outage probability for CSS 1

In Figure 5.9 the average BER is plotted as a function of normalized prediction lag for varying CSNRs. Simulation points resulting in a simulated average BER equal to the target BER are included and plotted with circle-markers. The figure shows that there is an upper limit on the prediction lag resulting in an average BER equal to the target BER, but also that this upper limit increases with increased CSNR. Thus, there is a trade-off between

jf_mT_s and $\overline{\gamma}$. In most cases the average BER is reduced when increasing the CSNR.[6]

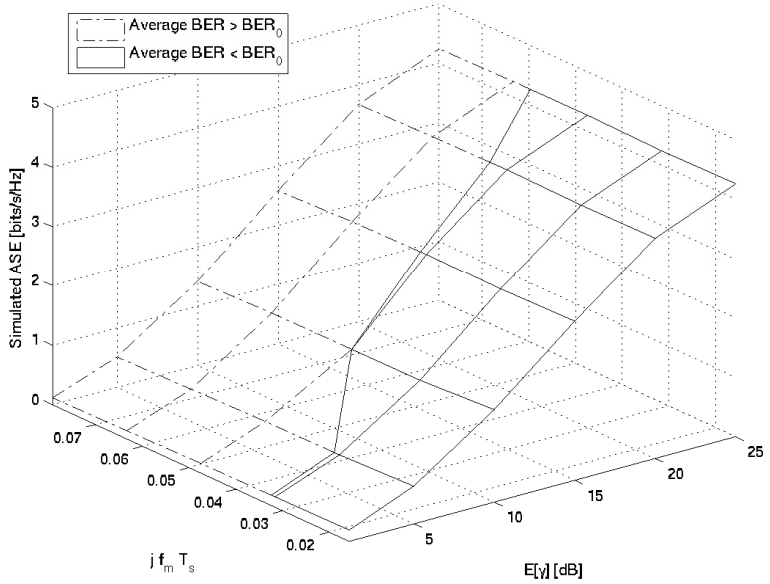


FIGURE 5.10: The ASE of CSS 1 plotted for $\overline{\gamma} \in [1, 25]$ dB and $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$. Dash-dot lines indicate that the target BER is not attained while solid lines show the region where the average BER is below BER_0.

For the average BER plot in Figure 5.9 the corresponding ASE and outage probability are plotted in Figures 5.10 and 5.11, respectively. Note that, in these figures solid lines mean that the target BER is attained while dashed lines mean that the average BER is above the target BER. In Figure 5.10 it can be observed that the ASE is almost constant for varying normalized prediction lags, except for at low CSNR values (see e.g. Figure 5.10 for $\overline{\gamma} = 0$ dB). By comparing the ASE in Figure 5.10 to the theoretical ASE in Figure 5.4 it can be observed that the simulated ASE for low values of jf_mT_s is equal (with only minor deviations) to the theoretical ASE. Since the spectral efficiencies are fixed for the $N = 5$ codecs used, and the switching thresholds are kept constant, the only parameter that varies with the normalized prediction lag is ρ. Reducing the correlation also reduces the

[6]For high values of the normalized prediction lag and low values of the CSNR it can be observed that the average BER increases with increased CSNR. This effect is a result of the outage probability behavior of the ACM scheme when the normalized correlation and the average CSNR is kept low.

average predicted CSNR value, thus it might have been expected that the ASE was reduced more notably as the prediction lag increases. However, it can be observed from Figure 5.6 that only for very low CSNR values does ρ drop notably. Also, for all the simulation results presented in this chapter we have $\rho \in [0.9, 1\rangle$, i.e. on average the expected predicted CSNR will only deviate a maximum of 10% compared to the expectation of the actual CSNR.

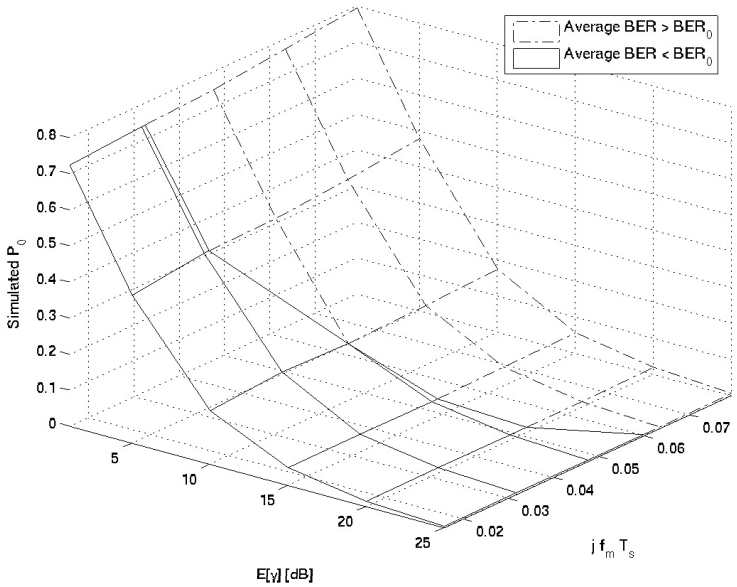


FIGURE 5.11: The outage probability of CSS 1 plotted for $\overline{\gamma} \in [1, 25]$ dB and $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$. Dash-dot lines indicate that the target BER is not attained while solid lines show the region where the average BER is below BER_0.

The plot in Figure 5.11 shows how the probability of outage is reduced as CSNR increases. Also, the outage probability increases somewhat with increased prediction lag, especially for low CSNRs. This can again be explained by observing that ρ only varies much for increasing jf_mT_s when the CSNR is low.

Average BER performance for CSS 2

In this subsection the simulated average BER is presented after increasing the switching thresholds by a constant δ_s. The BER plots in Figures 5.12 and

5.13 were obtained for $\delta_s = 0.5$ dB and $\delta_s = 1.0$ dB, respectively. As can be seen by comparing these figures and Figure 5.9, there is an increased tolerance for increasing prediction lag after increasing the switching thresholds. As an example, for $jf_mT_s = 7.9 \cdot 10^{-2}$ the average BER of CSS 1 and CSS 2 with $\delta_s = 0.5$ dB is above BER_0, but for CSS 2 with $\delta_s = 1.0$ dB the average BER is below BER_0 for all simulated CSNRs except in a small region: $\overline{\gamma} \in \langle 8.7, 12.2 \rangle$ dB.

The reduced average BER comes to the expense of a reduced ASE. Increasing δ_s will also increase the outage probability since the first switching threshold is increased. Results for the ASE and outage probability will be shown in the next subsection where the performance of all three strategies are compared.

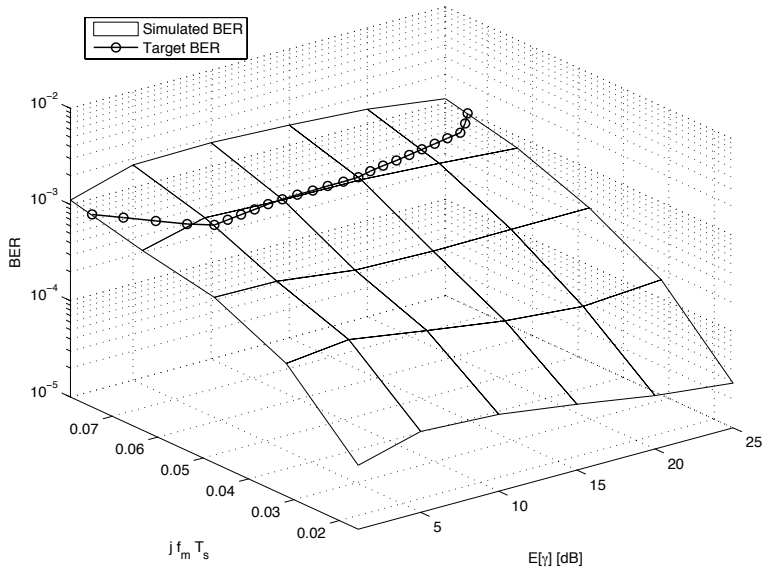


FIGURE 5.12: The average BER of CSS 2 plotted for $\delta_s = 0.5$ dB, average CSNRs $\overline{\gamma} \in [1, 25]$ dB, and $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$. The circle-markers show simulation points with average BER equal to the target BER.

Performance of CSS 3 and comparison of all the three strategies

In CSS 3, the switching thresholds are given as $s_n = \Psi_s(\gamma_n, \epsilon)$, where $\Psi_s(\cdot, \cdot)$ is defined in Equation (3.32), and are controlled by the parameter ϵ. However, in this case the switching thresholds also vary with the normalized correlation ρ (through normalized prediction lag and average CSNR). In this subsection simulation results are presented for $jf_mT_s = 7.9 \cdot 10^{-2}$ and

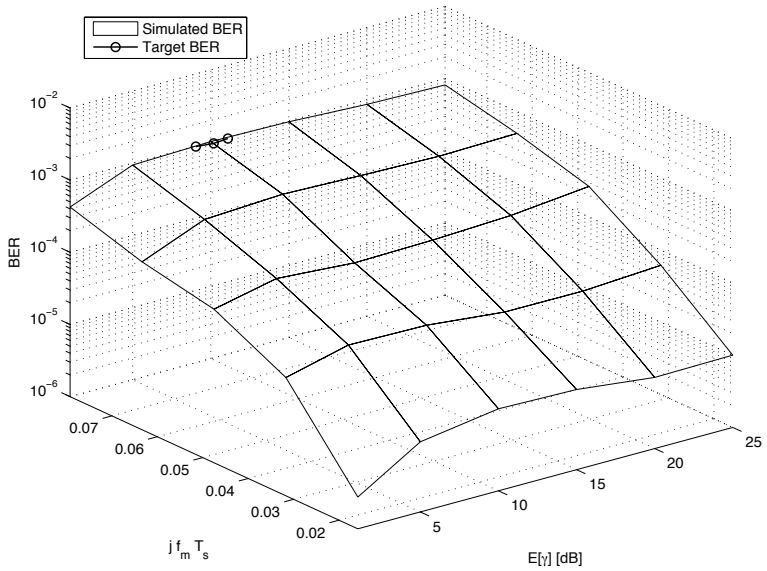


FIGURE 5.13: The average BER of CSS 2 plotted for $\delta_s = 1.0$ dB, average CSNRs $\overline{\gamma} \in [1, 25]$ dB, and $jf_mT_s \in [1.6 \cdot 10^{-2}, 7.9 \cdot 10^{-2}]$. The circle-markers show simulation points with average BER equal to the target BER.

$\overline{\gamma} \in [1, 25]$. The performance of CSS 3 is presented for five different values of ϵ:

$$\frac{\epsilon}{\text{BER}_0} \in \{1, 2, 10, 100, 200\},$$

and as in the previous section the performance of CSS 2 is evaluated for $\delta_s \in \{0.5, 1.0\}$ dB. The resulting ASE, outage probability, and average BER are plotted in Figures 5.14, 5.15, and 5.16, respectively. In all three figures dashed lines are utilized to indicate where the target BER is not attained.

The ASE plots for CSS 1 and CSS 2 in Figure 5.14 show that the ASE is reduced when adding the constant δ_s to the CSNR thresholds to obtain the new switching thresholds. It can be observed that the reduction in ASE is approximately a linear function of the constant δ_s (e.g. at $\overline{\gamma} = 10$ dB the reduction in ASE by using CSS 2 with $\delta_s = 1.0$ dB compared to CSS 1 is 0.22 bits/s/Hz and the reduction in ASE using CSS 2 with $\delta_s = 0.5$ dB compared to CSS 1 is 0.11 bits/s/Hz). The same behavior can be observed in Figure 5.15 where the outage probability increases (approximately linearly) with δ_s.

As argued previously in this chapter; using switching thresholds obtained from CSS 3 with $\epsilon = \text{BER}_0$ would result in a very conservative design. This is confirmed in Figure 5.16; Using a low value for ϵ (less than

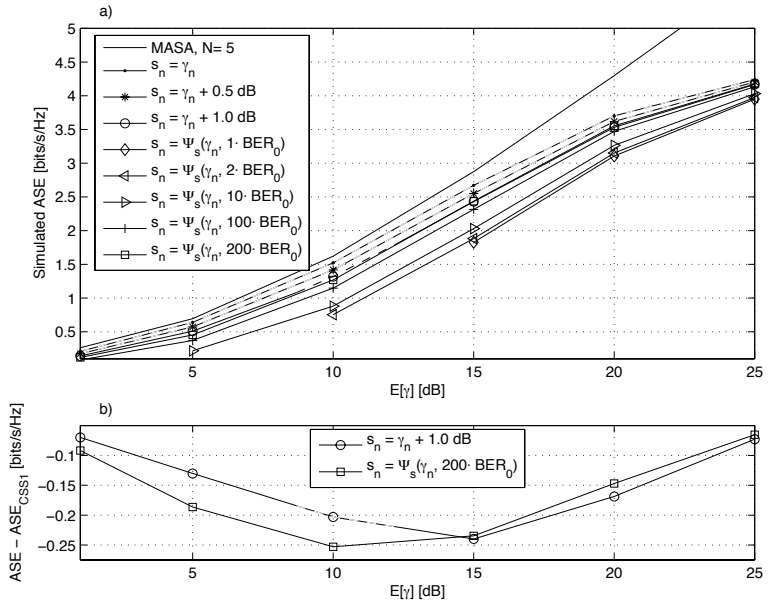


FIGURE 5.14: Plots of the ASE versus average CSNR for $\overline{\gamma} \in [1, 25]$ dB and $jf_mT_s = 7.9 \cdot 10^{-2}$. Solid lines indicate that the average BER is less than BER_0. In a) the ASE of CSS 1, CSS 2, and CSS 3 is plotted for different values of δ_s and ϵ. In b) the difference in the ASE of CSS 3 with $\epsilon = 200 \cdot BER_0$ and CSS 2 with $\delta_s = 1.0$ dB is plotted, and here the dotted line indicate the area where CSS 2 with $\delta_s = 1.0$ dB does not satisfy the target BER.

$100 \cdot BER_0$) results in a very low average BER, especially at low CSNRs. As a result of this the ASE is reduced more than necessary and the outage probability becomes very high. Using $\epsilon = 200 \cdot BER_0$ results in an average BER that is quite close to the average BER of CSS 2 with $\delta_s = 1.0$ dB, but always below BER_0. It should be noted that this choice of ϵ does not guarantee that the average BER stays below BER_0 for increasing normalized prediction lags.

When referring to CSS 2 and CSS 3 in the following it is implied that $\delta_s = 1.0$ dB and that $\epsilon = 200 \cdot BER_0$. In the lower sub plot of Figure 5.16 the average BER-versus-CSNR plot is zoomed in on the average BER of CSS 2 and CSS 3 for this case. The resulting ASE for the two strategies can be compared in the lower sub plot of Figure 5.14 where the plots shows the difference in ASE between CSS 2 and CSS 1, and between CSS 3 and CSS 1 as a function of $\overline{\gamma}$.

It can now be observed that for $\overline{\gamma} < 15$ dB the average BER of CSS 2 is higher than that of CSS 3, and lower for higher CSNRs. Likewise, the ASE

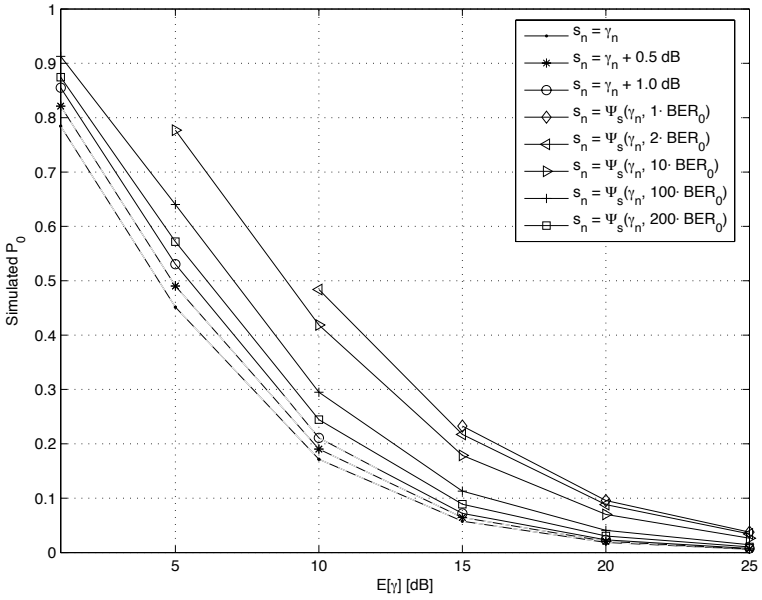


FIGURE 5.15: Plots of the outage probability versus average CSNR for all three strategies with different values of δ_s and ϵ, $\overline{\gamma} \in [1, 25]$ dB, and $jf_mT_s = 7.9 \cdot 10^{-2}$. Solid lines indicate that the average BER is less than BER_0.

of CSS 3 is larger than the ASE of CSS 2 for $\overline{\gamma} > 15$ dB, and lower for CSNRs below 15 dB. Thus, for this example we may conclude as follows; CSS 3 is the best candidate strategy for CSNRs above 15 dB since it provides the highest ASE, and for lower CSNRs since CSS 2 does not attain the average BER demand for all CSNRs below 15 dB.

In Figure 5.15 it can be seen that the outage probability of CSS 3 is always larger than that of CSS 2. This might seem strange since the average BER of CSS 3 in some instances is higher than that of CSS 2. This behavior can easily be explained by the fact that the ASE and the average BER will depend on all the switching thresholds while the outage probability depend only on s_1. In this figure it can also be observed that the outage probability (as expected) is reduced with increasing ϵ.

5.7 Concluding Remarks

A simulation tool has been developed for the purpose of evaluating transmission of information using an ACM scheme, in terms of ASE, average BER, and outage probability. In this chapter we have presented simula-

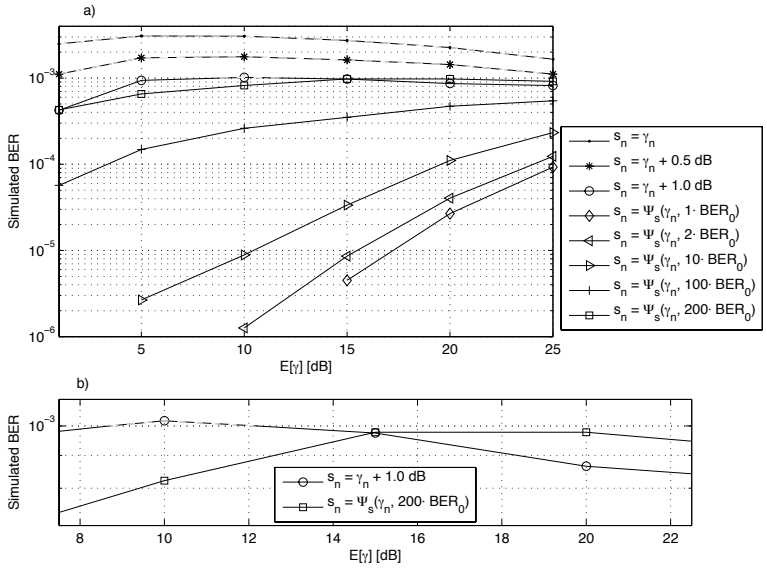


FIGURE 5.16: Plots of the average BER versus average CSNR for $\overline{\gamma} \in [1, 25]$ dB and $jf_mT_s = 7.9 \cdot 10^{-2}$. Solid lines indicate that the average BER is less than BER_0. In a) the average BER of CSS 1, CSS 2, and CSS 3 is plotted for different values of δ_s and ϵ. In b) the average BER-versus-CSNR plot is zoomed in on CSS 2 with $\delta_s = 1.0$ dB, and CSS 3 with $\epsilon = 200 \cdot \text{BER}_0$ for $\overline{\gamma} \in [7.5, 22.5]$ dB.

tion results for one particular ACM system, performing on a flat Rayleigh fading channel with correlated fading gain. The system has not been developed for the sole purpose of simulating ASE, average BER, and outage probability, but also to gain more insight in the components in a ACM system. The simulation tool has been used to investigate the performance of the MAP-optimal channel predictor as well as the overall performance of the ACM scheme.

One of the first steps in designing an ACM scheme is to select the codecs to be used as components. In this chapter it was shown that there is not only a trade-off between a high ASE and a low outage probability, but also a trade-off between the ASE at low and high average CSNRs.

The average BER will vary with the switching thresholds used, the normalized correlation ρ between predicted and actual CSNRs, variations in the fading envelope (equivalently variations in the correlation between samples of the fading envelope), and the bit error correcting properties of the codecs. The codecs used in our simulation tool are designed for AWGN channels, and thus, the bit error correcting properties of the codecs may

vary with the correlation between the fading samples. A possible way of improving BER performance which would not reduce the ASE might be to include an interleaver. But as indicated by the results in [Guo, Ng, and Hanzo, 2003], the BER performance of a Gallager code on a correlated Rayleigh fading channel yields only very small improvements by introducing an interleaver.

The normalized correlation ρ also affects the ASE and the outage probability since these to quantities depends on the PDF of the predicted CSNR. The system uses predictions of the fading for first symbols in each block of M' symbols to select the appropriate codec. That is, the system is designed assuming that the CSNR is approximately constant during transmission of one block. Since the fading is time varying also within one block, the instantaneous CSNR may fall below the corresponding CSNR threshold during the block transmission. From the scatter plots of $\widehat{\gamma}(k)$ versus $\gamma(k)$ it can be observed that the prediction error is largest when the actual CSNR is low. By comparing the scatter plots for $\widehat{\gamma}(k)$ versus $\gamma(k)$ to the $\widehat{\gamma}(k)$ versus $\overline{\gamma}_{M'}(k)$, we observe that the assumption of approximately constant fading during transmission of a block of channel symbols is slightly optimistic, but still quite accurate for the simulations presented here.

In our simulations the filter length and pilot spacing are kept constant while the correlation between fading samples are varied by varying the normalized prediction lag. Simulations of ρ show that the implemented predictor performs quite close to the theoretical performance of the predictor outlined in Chapter 3. Both the scatter plots and the plots of ρ as a function of jf_mT_s confirms that the MAP-optimal predictor has a negative bias.

In the simulation results presented in this chapter three different methods have been used to obtain the switching thresholds; in CSS 1 the CSNR thresholds are used as switching thresholds, in CSS 2 switching thresholds are obtained by adding a constant (in dB) to the CSNR thresholds, and in CSS 3 the method in Chapter 3 is used to obtain the switching thresholds. In CSS 2 the switching thresholds vary with the constant δ_s added, and in CSS 3 the switching thresholds are controlled by the probability of a codec mismatch ϵ, but are also dependent on the average CSNR and on ρ.

We have shown by means of simulation results that; Obtaining the switching thresholds by adding the same constant to all the CSNR thresholds, in order to attain the target BER, result in a too large ASE reduction for some average CSNRs. To maximize the ASE the constant added to the thresholds should be dependent on the correlation properties of the channel and the predictor. This can be done by designing the switching thresholds such that the probability of a codec mismatch ϵ is kept below a cer-

tain bound. Then the overall average BER will be below the target BER. However, the upper bounds found in Chapter 3 are too low and result in a conservative design. It has been shown that the technique could still be used by using higher values for ϵ (obtained by manually testing different values). The results from these simulations showed that using the technique (CSS 3) can provide an ACM scheme with a BER very close to (and below) the target BER, and thus a maximized ASE.

Chapter 6

Conclusions and Further Research

This chapter gives a summary of the work presented in this thesis. The work is separated into three categories; The contributions made (see Section 6.1), general concluding remarks (in Section 6.2), and topics for further research (described in Section 6.3).

6.1 Contributions

The main contributions in this thesis are:

- A method for upper bounding the probability of codec mismatch (the event of choosing a codec that cannot guarantee the target BER) in the case of imperfect channel prediction.
- New upper bounds for the ASE of an ACM scheme (the MASA) with outage constraint and imperfect channel knowledge.
- Development of an extensive software tool for simulating communication utilizing ACM on a correlated Rayleigh fading channel.
- Simulation results in which the above tool is used to confirm and illustrate the theory from previous works.
- New theoretical results have been obtained by investigating the PDF of the actual CSNR conditioned on the predicted CSNR, the correlation between the two, and on the average CSNR. The knowledge gained from these results can be used in the design of thresholds (for switching between codecs) to be used in a practical ACM scheme.

In this thesis the practical possibilities and limitations of ACM schemes were investigated by means of both simulations and theoretical results. New techniques that can be used in the design of ACM schemes were presented. As a starting point, the MASA was optimized under an outage constraint. The optimization is important since the existing method for optimizing the MASA in some cases result in very high outage probabilities. A high probability of outage might conflict with certain requirements for the QoS in a communication system.

The assumption of perfect channel prediction in the design of ACM schemes has been addressed since this is unlikely to be true for realistic communication scenarios. Imperfect channel prediction may lead to codec mismatch. We have described a method for designing ACM schemes using the probability of codec mismatch as a parameter. We have presented a method that can be used to obtain optimal values for this parameter. Using the new method, the resulting ACM scheme provides an average BER always below the target BER while still keeping the ASE as high as possible. Using the optimal values of the probability of codec mismatch for capacity achieving codecs we have found a unified tool for optimization of the MASA that take into account both imperfect channel prediction and outage constraints.

In order to investigate the performance of a practical ACM scheme utilizing a limited number of transmission modes, pilot symbols, channel estimation, and channel prediction a software tool has been developed and simulation results are presented in this thesis. The simulation software includes MAP-optimal filtering for both channel estimation (used in detection of channel symbols) and channel prediction, piloting of transmitted information, a correlated Rayleigh fading simulator, Gallager coding, and soft iterative Gallager decoding. The results include the average BER, ASE, and outage probability for different methods of selecting the switching thresholds as well as simulated performance of the channel predictor. The results are shown for an example system using a subset of the available codecs in the software. By means of an example, we have illustrated some practical implications, in terms of ASE and outage probability, of selecting a subset of codecs.

Most of the contributions in the thesis has been previously documented in [Jetlund et al., 2002, 2003a,b,c, 2004a,b].

6.2 Conclusions

The list below gives a summary of the conclusions in this thesis.

- The optimal MASA can be found for the case of imperfect channel prediction and for the case of a constraint on outage probability.
- The methods presented here (to do the optimization of the MASA) can be used as an aid or as a starting point in system design.
- System design involves finding the maximum value for the probability of a codec mismatch, and this value is codec dependent.
- The target BER is only attained for low degrees of mobility and prediction lags when the CSNR thresholds are used as switching thresholds.
- Increasing all switching thresholds by a constant (in dB) allows for increased mobility and/or prediction lags. However, such an increase reduces the ASE unnecessary.
- The approximations used to find the optimal value of the probability of codec mismatch result in a conservative designed ACM scheme. However, simulation results show that increasing the probability of codec mismatch results in an average BER very close to (but below) the target BER.

The optimized MASA under an outage constraint shows that increasing the number of codecs in an ACM scheme not only may increase the ASE, but also reduce the outage probability. That is, for an optimally designed system adding another codec allows for stricter outage probability constraints without trading away large amounts of the ASE.

There is a trade off between ASE and outage probability and a trade off between ASE at different CSNRs when choosing a subset of available codecs to be used in an ACM scheme. It is not necessarily always best to use the codecs with the highest SEs since this may increase the outage probability, and since the ASE may become too small at lower average CSNRs.

Comparing the MASA to the MASE (the channel capacity) it can be observed that the MASA is always below the MASE. This result is as expected since achieving the MASE would require infinitely many codecs. It can then be concluded that the MASA for a given number of codecs is a

much better (closer) upper bound on the ASE of an ACM scheme with the same number of codecs.

Simulation results show a trade off between allowed terminal velocity and allowed prediction lags in an ACM system. The tolerance towards prediction lags and mobility could easily be improved by adding a constant (in dB) to all the CSNR thresholds. As expected (and also shown by simulation results) this will reduce not only the average BER, but also the ASE. In fact, the reduction in ASE is approximately a linear function of the constant added. Increasing the switching thresholds will also increase the outage probability. From the simulation results of this strategy and from simulation results of the channel predictor it can be concluded that the lower CSNR thresholds should be increased more than the switching thresholds for the higher indexed codecs. This conclusion can also be seen from the optimization of the probability of a codec mismatch where it is seen that the difference between the switching thresholds and the CSNR thresholds is highest for the lowest indexed codecs.

Also shown by simulation results of average BER and by the behavior of the simulated channel predictor; the optimal value for the probability of a codec mismatch found using the approximate description of the BER-versus-CSNR relationship for our codecs result in a conservative design. That is, we conclude that although we have found a method for obtaining the optimal values of the probability of codec mismatch, the resulting codec mismatch probabilities found in this thesis are too low.

6.3 Suggested Topics for Further Research

There are many directions in which to go from the results provided in this thesis. Even more practical considerations could be taken into account in the design of ACM schemes. Alternative techniques could be treated in the same fashion as the ACM scheme is treated here, and the performance could be compared by means of simulation. In addition some of the topics considered here could be investigated further. The list below gives some of the possible directions for future research.

- Extensions of the simulation software:
  - More advanced channel models, such as e.g. unknown average CSNR, time varying power density spectrum, and frequency-selective fading channels.
  - Other code families.
  - Multiuser scenarios.

  - System employing OFDM and/or multiple antennas.
  - Packet based transmission, in which the packet error rate is a more relevant performance measure than the BER.

- Alternatives to ACM based on forward error correction:
  - Automatic repeat request protocols.
  - Incremental redundancy schemes.
  - Hybrid schemes.

- Extension to topics addressed in this thesis:
  - More accurate bounds on the probability of codec mismatch.
  - System performance as a function of other system parameters.
  - An extensive search for codecs that result in better system performance.

Currently, the software system assumes a flat Rayleigh fading channel with a Jakes spectrum. The software can easily be extended to take into account e.g. frequency selectivity, co-channel interference, shadowing, other fading statistics, non-stationary channel statistics, and synchronization errors. Estimation of other parameters of the channel besides the instantaneous CSNR should also be taken into account. Examples of such parameters are the average CSNR and the normalized Doppler frequency. The return channel is, in this thesis, assumed to be error-free. If a noisy return channel is included, the system performance could be evaluated for different amounts of errors in this channel. By introducing more advanced channel models it is also natural to incorporate other techniques such as OFDM and the use of multiple antennas (MIMO systems).

The performance of the ACM scheme is dependent on the codecs employed. There also exist other practical and popular types of channel codes (besides the Gallager codes used in this thesis) e.g. trellis codes (trellis coded modulation), turbo codes, and bit-interleaved coded modulation. Including other code families is a natural extension of the software.

Employing ACM in a wireless network/communication system with multiple users introduces many new possibilities and restrictions on the ACM scheme. The multi-user scenario should take into account cellular networks with link adaptation, feedback (return channel) load, different traffic classes and QoS, and overall system aspects such as multi-user access protocols.

The main idea behind ACM is to vary the information throughput with the variations in the CSNR. However, techniques such as automatic repeat request protocols and incremental redundancy schemes, can also vary the information throughput. The ACM scheme presented here could be compared to other techniques with regard to system performance, power consumption, complexity, and latency.

By using more accurate approximations of the performance of each of the codecs or by developing a more accurate method values of the probability of codec mismatch closer to the true optimal values could be found. The design strategies presented here can then be further improved. One could easily extend the simulation results presented (with the existing simulation software) to include e.g. varying predictor lengths and other configurations for the pilot symbols.

Appendix A

Special Functions

Definition 1 (The gamma function)
The gamma function, also known as Euler's integral of the second kind, is defined as follows [Gradshteyn and Ryzhik, 1980, Equation 8.310.1]:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad \text{for } x > 0. \tag{A.1}$$

Definition 2 (The *n*th-order Bessel functions of the first kind)
The nth-order Bessel function of the first kind and the nth-order *modified* Bessel function of the first kind are defined as follows [Gradshteyn and Ryzhik, 1980, Equations 8.402, 8.406.3]:

$$J_n(x) = \frac{x^n}{2^n} \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2^{2k} k! \Gamma(n+k+1)} \tag{A.2a}$$

$$\begin{aligned} I_n(x) &= j^{\,n} J_n(jx) \\ &= \frac{x^n}{2^n} \sum_{k=0}^{\infty} \frac{x^{2k}}{2^{2k} k! \Gamma(n+k+1)}. \end{aligned} \tag{A.2b}$$

Definition 3 (The exponential integral)
The Exponential Integral is defined as [Gradshteyn and Ryzhik, 1980, Equations 8.211.1]:

$$E_i(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt \qquad \text{for } x < 0. \tag{A.3}$$

Definition 4 (The degenerate hypergeometric function)
The degenerate hypergeometric function can be defined in terms of a series [Gradshteyn and Ryzhik, 1980, Eq. 9.210.1]

$$\Phi(\alpha,\gamma;z) = 1 + \frac{\alpha}{\gamma}\cdot\frac{z}{1!} + \frac{\alpha}{\gamma}\cdot\frac{\alpha+1}{\gamma+1}\cdot\frac{z^2}{2!} + \frac{\alpha}{\gamma}\cdot\frac{\alpha+1}{\gamma+1}\cdot\frac{\alpha+2}{\gamma+2}\cdot\frac{z^3}{3!} + \frac{\alpha}{\gamma}\cdot\frac{\alpha+1}{\gamma+1}\cdot\frac{\alpha+2}{\gamma+2}\cdot\frac{\alpha+3}{\gamma+3}\cdot\frac{z^4}{4!} + \cdots . \quad \text{(A.4)}$$

This function is also commonly denoted ${}_1F_1(\alpha;\gamma;z)$.

Definition 5 (The Whittaker function)
The Whittaker function can be defined in terms of the degenerate hypergeometric function as [Gradshteyn and Ryzhik, 1980, Eq. 9.220.1]

$$\mathrm{M}_{\lambda,\mu}(z) = z^{\mu+\frac{1}{2}}e^{-\frac{z}{2}}\Phi\left(\mu-\lambda+\frac{1}{2}, 2\mu+1; z\right). \quad \text{(A.5)}$$

Definition 6 (The Marcum Q-function)
The Q-function was defined by Marcum as (see [Marcum, 1947, 1948]:

$$Q(a,b) = \int_b^\infty x e^{-1/2\,x^2 - 1/2\,a^2} I_0(ax)\,dx. \quad \text{(A.6)}$$

Definition 7 (The inverse of the complementary Marcum Q-function)
The inverse of the complementary Marcum Q-function

$$\epsilon = 1 - Q(a,b) \quad \text{(A.7)}$$

with respect to its first and second argument is here denoted as:

$$a = q_a(b,\epsilon) \quad \text{(A.8)}$$

and

$$b = q_b(a,\epsilon) \quad \text{(A.9)}$$

respectively.

The example values of $a = q_a(b,\epsilon)$ and $b = q_b(a,\epsilon)$ for different arguments ϵ, a, and b in Tables A.1 and A.2 were found using an iterative method[1] commonly known as Ridders' method [Ridders, 1979] (there is a C-implementation available in [Press, Teukolsky, Vetterling, and Flannery, 2002]).

[1] The values were originally found with an accuracy of 10^{-10}, but is here presented with only 3 decimals, i.e. the round-off error is much larger than the error in the values found from the iterative procedure.

$b \setminus \epsilon$	10^{-6}	10^{-5}	10^{-4}	10^{-3}	$b \setminus \epsilon$	10^{-6}	10^{-5}	10^{-4}	10^{-3}
0.1	4.132	3.529	2.800	1.795	15	19.724	19.236	18.689	18.060
0.2	4.470	3.916	3.268	2.456	16	20.726	20.237	19.691	19.062
0.3	4.675	4.141	3.528	2.782	17	21.727	21.239	20.692	20.063
0.4	4.833	4.311	3.717	3.005	18	22.729	22.240	21.694	21.065
0.5	4.971	4.456	3.872	3.182	19	23.730	23.241	22.695	22.066
0.6	5.097	4.587	4.011	3.334	20	24.731	24.242	23.696	23.067
0.7	5.218	4.711	4.139	3.471	21	25.732	25.243	24.697	24.068
0.8	5.334	4.829	4.261	3.599	22	26.733	26.244	25.698	25.069
0.9	5.447	4.944	4.379	3.721	23	27.734	27.245	26.699	26.070
1	5.558	5.057	4.494	3.840	24	28.734	28.246	27.700	27.071
2	6.622	6.127	5.573	4.933	25	29.735	29.246	28.700	28.071
3	7.652	7.159	6.609	5.973	26	30.736	30.247	29.701	29.072
4	8.670	8.179	7.630	6.996	27	31.736	31.248	30.702	30.073
5	9.683	9.192	8.644	8.012	28	32.737	32.248	31.702	31.073
6	10.692	10.202	9.654	9.023	29	33.737	33.249	32.703	32.074
7	11.699	11.209	10.661	10.031	30	34.738	34.249	33.703	33.074
8	12.704	12.215	11.667	11.037	40	44.742	44.253	43.707	43.078
9	13.709	13.219	12.672	12.042	50	54.744	54.255	53.709	53.081
10	14.712	14.223	13.676	13.047	60	64.745	64.257	63.711	63.082
11	15.716	15.226	14.680	14.050	70	74.747	74.258	73.712	73.083
12	16.718	16.229	15.683	15.053	80	84.747	84.259	83.713	83.084
13	17.721	17.232	16.685	16.056	90	94.748	94.259	93.714	93.085
14	18.723	18.234	17.687	17.058	100	104.749	104.260	103.714	103.085

TABLE A.1: Approximate solutions to a = $g_a(b, \epsilon)$.

$a \setminus \epsilon$	10^{-6}	10^{-5}	10^{-4}	10^{-3}	$a \setminus \epsilon$	10^{-6}	10^{-5}	10^{-4}	10^{-3}
0.100	0.001	0.004	0.014	0.045	16.000	11.284	11.771	12.317	12.944
0.200	0.001	0.005	0.014	0.045	17.000	12.281	12.769	13.314	13.942
0.300	0.001	0.005	0.014	0.046	18.000	13.279	13.767	14.312	14.940
0.400	0.001	0.005	0.015	0.047	19.000	14.277	14.765	15.310	15.938
0.500	0.002	0.005	0.015	0.048	20.000	15.275	15.763	16.309	16.937
0.600	0.002	0.005	0.015	0.049	21.000	16.274	16.762	17.307	17.936
0.700	0.002	0.005	0.016	0.051	22.000	17.272	17.760	18.306	18.934
0.800	0.002	0.005	0.017	0.052	23.000	18.271	18.759	19.305	19.933
0.900	0.002	0.005	0.017	0.055	24.000	19.270	19.758	20.304	20.932
1.000	0.002	0.006	0.018	0.057	25.000	20.269	20.757	21.303	21.931
2.000	0.004	0.012	0.038	0.121	26.000	21.268	21.756	22.302	22.930
3.000	0.013	0.042	0.133	0.397	27.000	22.267	22.755	23.301	23.929
4.000	0.077	0.233	0.592	1.139	28.000	23.266	23.754	24.300	24.929
5.000	0.522	0.949	1.460	2.063	29.000	24.265	24.754	25.299	25.928
6.000	1.408	1.878	2.410	3.026	30.000	25.265	25.753	26.299	26.927
7.000	2.365	2.844	3.382	4.004	40.000	35.260	35.748	36.294	36.923
8.000	3.341	3.824	4.365	4.989	50.000	45.257	45.746	46.291	46.920
9.000	4.325	4.810	5.352	5.978	60.000	55.255	55.744	56.290	56.918
10.000	5.314	5.800	6.343	6.969	70.000	65.254	65.742	66.288	66.917
11.000	6.306	6.793	7.336	7.963	80.000	75.253	75.742	76.287	76.916
12.000	7.300	7.787	8.331	8.958	90.000	85.252	85.741	86.287	86.915
13.000	8.294	8.782	9.326	9.954	100.000	95.252	95.740	96.286	96.915
14.000	9.290	9.778	10.322	10.950	200.000	195.249	195.738	196.284	196.912
15.000	10.287	10.774	11.319	11.947	300.000	295.248	295.737	296.283	296.911

TABLE A.2: Approximate solutions to b = $g_b(a, \epsilon)$.

Appendix B

Statistics and Probability Density Functions

B.1 Rayleigh distribution

A random variable $X = X_I + jX_Q$, where X_I and X_Q are *independent and identically distributed* (IID) zero-mean Gaussian random processes with variance β, has a Rayleigh distribution [Råde and Westergren, 1990]

$$f_X(x) = \frac{x}{\beta} e^{-\frac{x^2}{2\beta}}, \quad x \geq 0 \tag{B.1}$$

with mean and variance:

$$\mu_X = \sqrt{\frac{\pi\beta}{2}}$$
$$\sigma_X^2 = \frac{4-\pi}{2}\beta.$$

B.2 Exponential distribution

An exponentially distributed random variable X with expectation β has a PDF [Stüber, 2001]

$$f_X(x, \beta) = \frac{1}{\beta} e^{-\frac{x}{\beta}} \tag{B.2}$$

with mean and variance:

$$\mu_X = \beta$$
$$\sigma_X^2 = \beta^2.$$

B.3 Bivariate Exponential Distribution

Two correlated exponentially distributed random variables X_1 and X_2 with expectations β_1 and β_2 respectively, and correlation denoted ρ has a joint PDF [Nagao and Kadoya, 1971]

$$f_{X_1,X_2}(x_1, x_2, \beta_1, \beta_2, \rho) = \frac{e^{-\left(\frac{x_1}{\beta_1(1-\rho)} + \frac{x_2}{\beta_2(1-\rho)}\right)} I_0\left(2\frac{\sqrt{\rho}}{1-\rho}\sqrt{\frac{x_1 x_2}{\beta_1\beta_2}}\right)}{\beta_1\beta_2(1-\rho)} \tag{B.3}$$

with mean and variance:

$$\begin{aligned}\mu_{X_1} &= \beta_1\\ \mu_{X_2} &= \beta_2\\ \sigma^2_{X_1} &= \beta_1^2\\ \sigma^2_{X_2} &= \beta_2^2.\end{aligned}$$

A conditional exponential PDF

The conditional PDF of X_1 and X_2 can be obtained using Bayes rule as follows:

$$f_{X_1|X_2}(x_1|x_2, \beta_1, \beta_2, \rho) = \frac{f_{X_1,X_2}(x_1, x_2, \beta_1, \beta_2, \rho)}{f_{X_2}(x_2\beta_2)} \tag{B.4}$$

$$\begin{aligned} &= \frac{1}{\beta_1(1-\rho)} \\ &\quad \times I_0\left(\frac{2\sqrt{\rho}}{1-\rho}\sqrt{\frac{x_1 x_2}{\beta_1\beta_2}}\right) \\ &\quad \times e^{-\left(\frac{x_1}{\beta_1(1-\rho)} + \frac{\rho x_2}{\beta_2(1-\rho)}\right)}. \end{aligned} \tag{B.5}$$

Expectation of the conditioned exponential variable

The expectation of X_1 conditioned on X_2 is defined as

$$\mathrm{E}[X_1|X_2] = \int_0^\infty x_1 \cdot f_{X_1|X_2}(x_1|x_2, \beta_1, \beta_2, \rho)\, dx \tag{B.6}$$

Thus, obtaining the expectation yields solving the following integral:[1]

$$\mathrm{H}(a, b) = \int_0^\infty x \mathrm{I}_0\left(a\sqrt{x}\right) e^{-x/b} dx \tag{B.7}$$

[1] For simplicity we refer to the function in Equation (B.7) as the H-function.

From [Gradshteyn and Ryzhik, 1980, Eq. 6.643.2] we have

$$\int_0^\infty x^{\mu-\frac{1}{2}} \mathrm{I}_{2v}\left(2\beta\sqrt{x}\right) e^{-\alpha x} dx = \frac{\Gamma\left(\mu+v+\frac{1}{2}\right)}{\Gamma\left(2v+1\right)} \beta^{-1} e^{\frac{\beta^2}{2\alpha}} \alpha^{-\mu} \mathrm{M}_{-\mu,v}\left(\frac{\beta^2}{\alpha}\right) \quad \text{(B.8)}$$

when

$$\mathrm{Re}\left(\mu+v+\frac{1}{2}\right) > 0, \quad \text{(B.9)}$$

where $\mathrm{M}_{\lambda,\mu}(z)$ is the Whittaker function defined in Equation (A.5) and $\Gamma(\cdot)$ is the gamma function defined in Equation (A.1). Thus, by substituting

$$\begin{aligned} \mu &= \tfrac{3}{2}, \\ v &= 0, \\ \beta &= \tfrac{a}{2}, \\ \alpha &= \tfrac{1}{b} \end{aligned} \quad \text{(B.10)}$$

into Equation (B.8) the H-function can be written as

$$\begin{aligned} \mathrm{H}(a,b) &= \frac{\Gamma(2)}{\Gamma(1)} \frac{2}{a} e^{\frac{a^2 b}{8}} \sqrt{b^3} \mathrm{M}_{-\frac{3}{2},0}\left(\frac{a^2 b}{4}\right) \\ &= \frac{2}{a} e^{\frac{a^2 b}{8}} \sqrt{b^3} \mathrm{M}_{-\frac{3}{2},0}\left(\frac{a^2 b}{4}\right). \end{aligned} \quad \text{(B.11)}$$

Substituting the definition of the Whittaker function (in Equation (A.5)) this result can be written as

$$\mathrm{H}(a,b) = \frac{2}{a} e^{\frac{a^2 b}{8}} \sqrt{b^3} \left(\frac{a^2 b}{4}\right)^{\frac{1}{2}} e^{-\frac{a^2 b}{8}} \Phi\left(2,1;\frac{a^2 b}{4}\right), \quad \text{(B.12)}$$

where $\Phi(\alpha,\gamma;z)$ is the *degenerate hypergeometric function* which is defined as a series in Equation (A.4). The following relationship for $\Phi(\alpha,\gamma;z)$ can be used to reduce the number of terms in the series to a finite number [Gradshteyn and Ryzhik, 1980, Equation 9.212.1]:

$$\Phi(\alpha,\gamma;z) = e^z \Phi(\gamma-\alpha,\gamma;-z). \quad \text{(B.13)}$$

In our case, the series now reduces to a sum with only two terms since

$$\begin{aligned} \Phi(2,1;z) &= e^z \Phi(-1,1;-z) \\ &= e^z \left(1 - \frac{1}{1}\cdot\frac{(-z)}{1!} - \frac{1}{1}\cdot\frac{0}{2}\cdot\frac{(-z)^2}{2!} - \frac{1}{1}\cdot\frac{0}{2}\cdot\frac{1}{3}\cdot\frac{(-z)^3}{3!} - \cdots\right) \\ &= e^z (1+z). \end{aligned} \quad \text{(B.14)}$$

The H-function in Equation (B.7) can now be written with the following closed form expression

$$\begin{aligned} \mathrm{H}(a,b) &= \frac{2}{a} e^{\frac{a^2 b}{8}} \sqrt{b^3} \left(\frac{a^2 b}{4}\right)^{\frac{1}{2}} e^{-\frac{a^2 b}{8}} e^{\frac{a^2 b}{4}} \left(1 + \frac{a^2 b}{4}\right) \\ &= b^2 \left(1 + \frac{a^2 b}{4}\right) e^{\frac{a^2 b}{4}}. \end{aligned} \tag{B.15}$$

The expectation in Equation (B.6) can now be written as

$$\mathrm{E}[X_1 | X_2] = \beta_1 \left(1 - \rho\right) \left(1 + \frac{\rho}{1-\rho} \frac{x_2}{\beta_2}\right) \tag{B.16}$$

B.4 Nakagami distribution

A Nakagami distributed random variable X with Nakagami parameter m and average power $\mathrm{E}[X^2] = \Omega$ has a PDF [Nakagami, 1960]

$$f_X(x) = \frac{2 m^m x^{2m-1}}{\Gamma(m) \Omega^m} e^{-\frac{m x^2}{\Omega}}, \quad m = \frac{\Omega^2}{\mathrm{E}[(X^2 - \Omega)^2]} \geq \frac{1}{2}. \tag{B.17}$$

Appendix C

Simulating Correlated Fading

C.1 Filtering Gaussian Random Variables

The Doppler power spectrum of the fading envelope is found by taking the Fourier transform of the autocorrelation in Equation (3.6):

$$\Phi_{zz}(f) = \mathcal{F}\left[\phi_{zz}(\tau)\right] = \begin{cases} \frac{\Omega_p}{2\pi f_m} \frac{1}{\sqrt{1-\left(\frac{f}{f_m}\right)^2}} & |f| \leq f_m \\ 0 & \text{otherwise.} \end{cases} \tag{C.1}$$

It is possible to generate samples from the Rayleigh fading channel model having the correlation in Equation (3.6) from a sampled version of the Doppler power spectrum. The model was first introduced by Bell Labs in the 1970s, and in [Arredondo, Chriss, and Walker, 1973], results from a hard wired simulator was presented. Smith presented an example of a similar model for computer based simulations of Rayleigh fading in [Smith, 1975]. The model has later been generalized as described in e.g. [Rappaport, 1996]. There also exist several other models, e.g the *sum of sinusoids method* [Stüber, 2001]), but these do not always satisfy the stationary behavior demanded or other statistical criteria [Young and Beaulieu, 2000]. A comprehensive description of both the statistical properties of Smith's method and a modified version were presented by Young and Beaulieu in [Young and Beaulieu, 2000]. In our computer simulations we have used the modified version, which reduces the number of calculations needed to generate fading samples.

A block diagram of this method is shown in Figure C.1. The elements of the K-dimensional column vectors $\mathbf{A}$ and $\mathbf{B}$ are real-valued zero-mean identical and independent distributed Gaussian random variables with variance σ_w^2 (the random process is denoted $\mathcal{N}(0, \sigma_w)$). The filter coefficients in

the real-valued column vector **F** are found by sampling the Doppler power spectrum in Equation (C.1). The two boxes marked **F** perform element wise multiplication between vectors **A** and **F** (on the upper branch), and vectors **B** and **F** (lower branch). The resulting vectors are added together component-wise such that

$$\mathbf{Z} = \mathbf{F} \circ \mathbf{A} - i\mathbf{F} \circ \mathbf{B}. \tag{C.2}$$

Thus, **Z** is a complex vector. The sampled version of the complex fading envelope **z** of length K is computed by taking the K-point *inverse discrete Fourier transform* (efficient implementation if the discrete Fourier Transform and its inverse can be achieved by using the *fast Fourier transform* in e.g. [Press et al., 2002]).

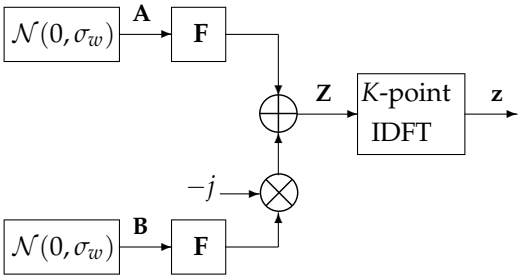


FIGURE C.1: Smith's method for generating K samples of a fading channel.

C.2 A Note on the Generation of Random Numbers

There are three components in the simulation model in Figure 5.1 using random numbers; the source, the fading gain and the additive noise. The source is specified to have a uniform distribution, the AWGN of the channel is complex with IID imaginary and real components, and the fading simulator described above also uses Gaussian random numbers. It is necessary to use a *random number generator* (RNG) for each of these operations that is independent of the others such that the results produced from Monte Carlo simulations of the system have high confidence. Also, since the fading simulator produces correlated vectors with a specific autocorrelation it is important that the random variables used are uncorrelated.

Any RNG implemented in a computer will not be "truly" random, since the output of any algorithm used can be predicted. In fact any RNG will have a given period. An excellent description of the limitations of practical

RNGs is available in [Press et al., 2002]. The simulation model used here is implemented in the ANSI C programming language, and the `rand()` function that returns an uniformly distributed integer in a range of (usually) 2^{16} numbers could have been used. The outcome of this function can easily be transformed into any other uniform distribution or a Gaussian distribution. However, the number of evaluations performed in our system might reach a much higher number then the period of this specific RNG (the period depends on the architecture of the processor used). As an example: calculations in a Monte Carlo simulation obtaining 100 block errors for a Gallager codec with block length $M = 200$ performing on a complex AWGN channel at a BER of 10^{-6} requires at least $100 \cdot 200 \cdot 2 \cdot 10^6 > 2^{35}$ random numbers. Thus, confidence in simulation results are reduced since simulation may evaluate the same outcomes of the RNGs multiple times. Marsaglia and MacLaren proposed a series of RNGs in the 1960's (see e.g. [Marsaglia, 1963; MacLaren and Marsaglia, 1965]). Results presented in this thesis were obtained using an implementation of Marsaglia's so called "Mother of all random number generators" (the C-code is available freely on the world wide web e.g. in [Wheeler]). This RNG produces uniformly distributed random variables and has a period of approximately 2^{250} which is more than sufficient for the simulations in this thesis. Gaussian random variables are generated from uniform variables using the Box-Muller algorithm [Press et al., 2002] (available in [Carter]).

Appendix D

Gallager based codecs

D.1 Introduction

In this appendix the design of codecs using Gallager codes with subsequent modulation is outlined. In order to obtain N codecs with different SEs Gallager based codecs with varying code rates and different modulation constellations are used. Note, the combination of a specific Gallager code and corresponding constellation is not optimized. The set of codecs presented here is an example set of codecs that *can* be used in a practical ACM scheme.

First, in this appendix a model for the use of a Gallager based codec on an AWGN channel is outlined in Section D.2. The modulation and demodulation technique and the Gallager encoder/decoder are described in Section D.3 and Section D.4, respectively. In Section D.5 the example codecs, used in simulations in Chapters 3 and 5, are presented.

D.2 System Model

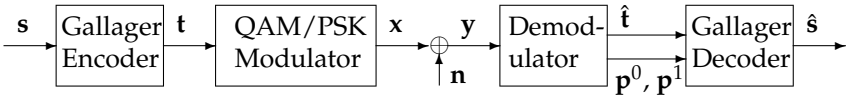


FIGURE D.1: System model for Gallager coded information transmitted on an AWGN channel using complex valued channel symbols.

Figure D.1 shows the system model for transmission of information encoded with a Gallager code and subsequently QAM or PSK modulation on an AWGN channel. A g-dimensional column vector of information bits, $\mathbf{s}$,

is coded into a q-dimensional codeword, $\mathbf{t}$, (column vector) by the Gallager encoder. The vector of channel symbols, $\mathbf{x}$ is generated by mapping sub-vectors of length $\log_2(S)$ from $\mathbf{t}$ to channel symbols by the QAM/PSK modulators, here S is the size of the constellation and it is assumed that $q = M\log_2(S)$ where M is an integer. The received vector becomes

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \tag{D.1}$$

where $\mathbf{n}$ is a complex AWGN vector. The vectors $\mathbf{y}$, $\mathbf{x}$, and $\mathbf{n}$ are all column vectors of length $M = \frac{q}{\log_2(S)}$. Since $\mathbf{x}$ and $\mathbf{n}$ are uncorrelated and the modulation is memoryless, the received vector can be demodulated symbol-by-symbol producing a hard decision, $\widehat{\mathbf{t}}$, on the received coded information (hard-demodulation). The Gallager decoder uses this vector and the probabilities p_l^0 and p_l^1 of bit l in $\widehat{\mathbf{t}} = \{\widehat{t}_l\}_{l=1}^q$ being demapped without or with error, respectively. Soft-demodulation (calculation of the probabilities of the decisions in $\widehat{\mathbf{t}}$) is also performed by the demodulator. The Gallager decoder produces the decoded information vector $\widehat{\mathbf{s}}$.

The noise added by the AWGN channel affects the modulated channel symbols. However, it is in the following useful to express the relationship between the binary transmitted and received information vectors, $\widehat{\mathbf{t}}$ and $\mathbf{t}$. This relationship can be modeled as

$$\widehat{\mathbf{t}} = \mathbf{t} + \mathbf{w} \tag{D.2}$$

where $\mathbf{w}$ is a binary column vector of length q representing additive noise as a result of demodulating noisy channel symbols. In the following, $\mathbf{t}$ and $\mathbf{w}$ are assumed independent.

D.3 Modulation and Demodulation

Each channel symbol in a constellation of size S, ζ_i ($i \in \{1, 2, \ldots, S\}$), can uniquely represent $\kappa = \log_2 S$ bits. The κ-tuple corresponding to channel symbol ζ_i is denoted $\{\alpha_{i,n}\}_{n=1}^{\kappa} = \{\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,\kappa}\}$. The modulation procedure is done by mapping tuples from $\mathbf{t}$ to channel symbols in the constellation. The mapping to channel symbols is done using Gray mapping [Stüber, 2001]. That is, the κ-tuple corresponding to two horizontal or vertical neighbors—in the complex plane—differ at only one position. In Figure D.2 a QAM constellation with 16 symbols is shown with the corresponding 4-tuples used in the Gray mapping.

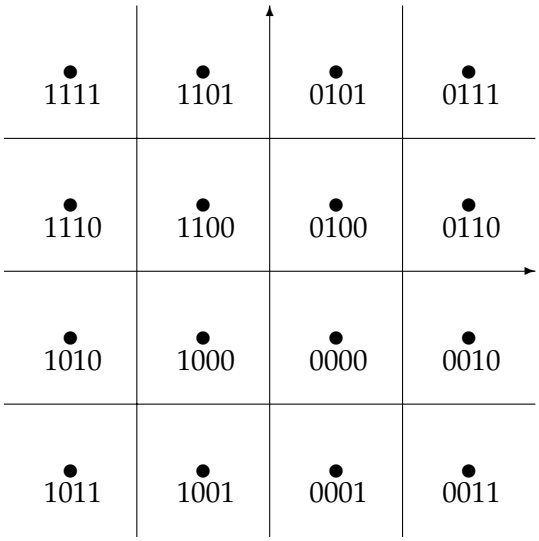


FIGURE D.2: 16 QAM constellation with Gray mapped binary sequences.

Symbol-by-Symbol Hard-Demodulation

Hard-demodulation of the kth received channel symbol $[\mathbf{y}]_k$ is defined as finding the symbol in the constellation with the smallest Euclidean distance to the received channel symbol. Define the Euclidean distance between a received channel symbol $[\mathbf{y}]_k$ and ζ_i as

$$\mathcal{D}([\mathbf{y}]_k, \zeta_i) = \sqrt{([\mathbf{y}]_{k,I} - \zeta_{i,I})^2 + ([\mathbf{y}]_{k,Q} - \zeta_{i,Q})^2}. \tag{D.3}$$

where the indexes I and Q are used to separate the real and imaginary parts of channel symbols. Then, symbol-by-symbol demodulation can be done by first solving

$$[\widehat{\mathbf{x}}]_k = \left\{ \zeta_i : \mathcal{D}([\mathbf{y}]_k, \zeta_i) = \min_{l \in \{1,2,\ldots,S\}} \{\mathcal{D}([\mathbf{y}]_k, \zeta_l)\} \right\}, \tag{D.4}$$

where $[\widehat{\mathbf{x}}]_k$ denotes the hard detected channel symbol.[1] Demapping is then done by using the κ-tuple corresponding to ζ_i:

$$\{\widehat{t}_l\}_{l=m}^{m+\kappa-1} = \{\alpha_{i,n}\}_{n=1}^{\kappa}, \qquad m = (k-1)\kappa + 1 \tag{D.5}$$

Symbol-by-Symbol Soft-Demodulation

Soft demodulation is here defined as obtaining the prior probability of each bit in $\widehat{\mathbf{t}}$ being correct. Again, this is done on a symbol-by-symbol basis. There are two obvious approaches that can be used to obtain this probability. The probability of $[\widehat{\mathbf{x}}]_k = \zeta_i$ being the channel symbol transmitted can be calculated from

$$\mathrm{P}(\zeta_i|[\mathbf{y}]_k) = \frac{g(\zeta_i, [\mathbf{y}]_k)}{\sum_{l=1}^{S} g(\zeta_l, [\mathbf{y}]_k)}, \tag{D.6}$$

where

$$g(\zeta_l, [\mathbf{y}]_k) = \frac{1}{\sqrt{2\pi N_0 B}} \exp\left(-\frac{\mathcal{D}^2([\mathbf{y}]_k, \zeta_i)}{2N_0 B}\right), \tag{D.7}$$

and the standard deviation $\sqrt{N_0 B}$ of the samples in the AWGN vector $\mathbf{n}$ is assumed known. A straight forward approach is now to assign the probability in Equation (D.6) to each code bit $\{\widehat{t}_l\}_{l=m}^{m+\kappa-1}$ resulting from detecting $[\widehat{\mathbf{x}}]_k$ as channel symbol ζ_i:

$$p_l^0 = \mathrm{P}(\zeta_i|[\mathbf{y}]_k) \tag{D.8}$$

and

$$p_l^1 = 1 - p_l^0. \tag{D.9}$$

A more sophisticated approach, which is used here, is to calculate the probability of each bit in the κ-tuple found from the demapping of ζ_i being correct based on the κ-tuples for the other symbols in the constellation as follows; assume that the lth bit in the demapping of a hard-demodulated channel symbol $[\mathbf{y}]_k$ is $\alpha_{i,l}$. Now, the prior probability of $\alpha_{i,l}$ being correct can be found by summing the probability $\mathrm{P}(\zeta_m|[\mathbf{y}]_k)$ for each of the symbols

[1] Using the Euclidean distance in detecting channel assumes that a QAM constellation is employed. For PSK constellations channel symbols can be separated into amplitude and phase. Then, instead of the Euclidean distance the absolute value of the difference between phase of each symbol in the constellation and the phase of the received signal is used as a distance measure. Subsequently hard-detection is done by choosing the ζ_i that minimizes this distance measure.

in the constellation weighted by the probability of bit l, $\alpha_{i,m}$, in the tuple being equal to $\alpha_{i,l}$:

$$p_l^0 = \sum_{m=1}^{S} \mathrm{P}(\zeta_m|[\mathbf{y}]_{k,l})\mathrm{P}(\alpha_{m,l} = \alpha_{i,l}) \tag{D.10}$$

where

$$\mathrm{P}(\alpha_{m,l} = \alpha_{i,l}) = \begin{cases} 1 & \text{when } \alpha_{m,l} = \alpha_{i,l} \\ 0 & \text{otherwise} \end{cases} \tag{D.11}$$

and again $p_l^1 = 1 - p_l^0$.

D.4 The Concept and Properties of Gallager Codes

Linear Block Codes

An error-correcting code can be used to correct information that has changed (e.g. communicated on a distorted channel or stored on a lossy medium) based on the redundancy in the code. A specific class of error-correcting codes are called "linear block codes."[2] A (q, g) linear block code receives a block of information (an *information word*) of finite length g and produce a codeword of finite length q. The rate of the code is then defined as

$$r = g/q \text{ [information bits per code bit].} \tag{D.12}$$

Considering only binary codes there are 2^g valid codewords. A linear block code (q, g) can be described by its parity check matrix $\mathbf{H}$ and the generator matrix $\mathbf{G}$. The codeword $\mathbf{t}$ can be found from

$$\mathbf{t} = \mathbf{G}^T\mathbf{s} \tag{D.13}$$

where $\mathbf{G}$ is a matrix of dimension $g \times q$, and $\mathbf{s}$ is the information word.

The parity check matrix of a (q, g) linear block code of dimension $(q - g) \times q$ is defined such that

$$\mathbf{H}\mathbf{G}^T = \mathbf{0}. \tag{D.14}$$

Then, it follows that

$$\mathbf{H}\widehat{\mathbf{t}} = \mathbf{H}(\mathbf{t} + \mathbf{w}) = \mathbf{H}(\mathbf{G}^T\mathbf{s} + \mathbf{w}) = \mathbf{H}\mathbf{w} \tag{D.15}$$

where $\mathbf{c} = \mathbf{H}\widehat{\mathbf{t}} = \mathbf{H}\mathbf{w}$ is commonly referred to as a "syndrome" vector.

[2] A linear combination of two valid codewords produced by a linear block code is also a valid codeword.

The optimal decoder for a block code returns the information vector $\widehat{\mathrm{s}}$ that maximize the posteriori probability [MacKay and Hesketh, 1997]

$$\mathrm{P}(\mathbf{s}|\widehat{\mathbf{t}},\mathbf{G}) = \frac{\mathrm{P}(\widehat{\mathbf{t}}|\mathbf{t},\mathbf{G})\mathrm{P}\mathbf{s}}{\mathrm{P}(\widehat{\mathbf{t}}|\mathbf{G})}. \tag{D.16}$$

However, this decoding problem is NP-complete and can not implemented in practice [MacKay and Hesketh, 1997]. Instead suboptimal algorithms can be used to obtain a solution.

Gallager Codes

Gallager's invention was to construct the parity check matrix $\mathbf{H}$, which is used in the decoding, such that it has a low density of ones (i.e. it is *sparse*). This ensures that the code can be decoded with a relatively low complexity decoder (with simple implementation), while also allowing good error correcting properties [MacKay and Hesketh, 1997].

A binary Gallager parity check matrix $\mathbf{H}$ has q columns (where q is the codeword length) and $u = q - g$ rows. In order for the matrix to be sparse the number ones in each column must much less than the number of rows, and the number ones in each row must be much less than the number of columns.

A regular parity check matrix has an equal number of ones in each row and an equal number of ones in each column. In this thesis a brute force approach is used to generate the parity check matrix. This is done by generating the rows randomly one-by-one and for each row checking that the row generated last is not a linear combination of one or more of the other rows. Then it follows, that the matrix is of full rank. In order to simplify the creation of the generator matrix the parity check matrix is made systematic by performing Gauss elimination. That is, $\mathbf{H}$ can be written as a concatenation between a matrix $\mathbf{C}$ representing the redundant information and an identity matrix representing the source information to be transmitted:

$$\mathbf{H} = [\mathbf{C}|\mathbf{I}_{u\times u}]\,. \tag{D.17}$$

The generator matrix can now be written as

$$\mathbf{G}^T = \left[\frac{\mathbf{I}_{g\times g}}{\mathbf{C}}\right]. \tag{D.18}$$

Decoding of Gallager Codes by the Message Passing Algorithm

Several soft decoding algorithms have been developed for Gallager codes. Examples of such algorithms are the message passing propagation (also

known as the sum-product algorithm) and an a posteriori probability decoding algorithm [Lin and Costello, Jr., 2004]. In this thesis message passing has been used in the decoder. In this section the main idea behind the message passing algorithm is described. The mathematical expressions for the algorithm (i.e. for the messages passed) is text book material and out of scope for this thesis. Excellent descriptions of decoding algorithms can be found in [MacKay and Hesketh, 1997; Davey, 1999; Lin and Costello, Jr., 2004].

The task of the decoder can be formulated as finding a noise vector, $\widehat{\mathbf{w}}$, that satisfies

$$\mathbf{H}\widehat{\mathbf{w}} = \mathbf{H}\widehat{\mathbf{t}} = \mathbf{H}\mathbf{w}. \tag{D.19}$$

In other words, the task of the decoder is to obtain the most likely noise vector. When this noise vector is found it can be subtracted from received codeword producing a new codeword, $\tilde{\mathbf{t}}$. When $\tilde{\mathbf{t}} = \mathbf{t}$ the coded information can be reproduced without error. For a binary systematic code with rate g/q the first g bits of the codeword is equal to the information encoded and the remaining $q - g$ bits are parity bits, i.e. decoding information from a legal codeword is done by extracting the first g bits.

An iterative decoding algorithm can be designed such that the posterior probability of the value of each noise sample in the noise vector is estimated, given $\widehat{\mathbf{t}}$ and the prior probabilities calculated in the previous section. A Gallager code can be described by a bipartite graph with two sets of nodes[3] defined by $\mathbf{H}$ [Lin and Costello, Jr., 2004]. One set represents the noise samples in $\widehat{\mathbf{w}}$. The nodes in the second set are denoted check nodes and represent the symbols in the syndrome vector $\mathbf{c}$. The syndrome vector contains g samples:

$$\mathbf{c} = [c_1, c_2, \ldots, c_g]^T. \tag{D.20}$$

For simplicity, the noise samples in $\widehat{\mathbf{w}}$ are denoted w_l resulting in the column vector

$$\widehat{\mathbf{w}} = [w_1, w_2, \ldots, w_q]^T. \tag{D.21}$$

An example bipartite graph is shown in Figure D.3. Nodes c_m and w_l are now connected if $[\mathbf{H}]_{ml} = 1$. The message passing is done in two steps. First each noise node w_l passes two messages to each of the check nodes it is connected to (i.e. a message is passed from w_l to c_m if and only if $[\mathbf{H}]_{ml} = 1$). The messages represent the belief the node w_l has of being zero and one, respectively.[4] After all check nodes have received messages from all

[3] A bipartite graph is a graph with two disjoint set of nodes; noise nodes and check nodes.

[4] Recall that all vectors are assumed to be binary vectors. Thus, each node can only take on two values: zero and one.

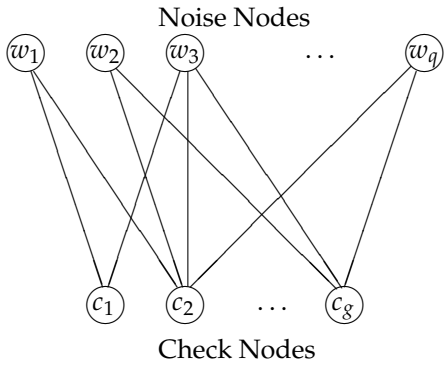


FIGURE D.3: Example of a bipartite graph built from the parity check matrix. Message passing is done in both directions of an edge.

noise nodes they are connected to, a belief in whether the noise nodes are in state zero or one from the view of the check node can be calculated. Then, each check node c_m in $\mathbf{c}$ passes on this belief.[5] When these two steps are performed the most likely noise vector, based on all received messages at each noise node, is selected and tested. That is, if $\mathbf{H}\widehat{\mathbf{w}} = \mathbf{c}$ the decoding ends by subtracting $\widehat{\mathbf{w}}$ from $\widehat{\mathbf{t}}$. If a proper noise vector is not found the algorithm repeats the two steps until such a vector is found or a maximum number of iterations is reached.

The probabilities produced by the demodulator (see Equation (D.10)) can be used to initialize the message passing algorithm. That is, the messages passed from the noise nodes in the first initialization are set equal to the prior probabilities p_l^0 and p_l^1.

D.5 Example Codecs

In this section an example codec set with $N = 6$ codecs are described. Codec n is defined to use a constellation of size $S_n = 2^{n+1}, n \in \{1, 2, \ldots, 6\}$. The number of information bits input to Gallager code in codec n is denoted g_n and the number of bits in the codeword produced is denoted q_n. The code rate is defined as $r_n = g_n/q_n = n/(n+1)$ information bits

[5]Note, both the check nodes and the corresponding syndrome samples are denoted c_m. Likewise, both the noise nodes and the noise samples are denoted w_l. The value for noise sample w_l is updated after each iteration, while the syndrome sample c_m is kept constant. Thus, each node has two functions; holding the value of its noise or syndrome sample and passing messages.

per code bit. This is done to keep the length M of a block of channel symbols produced by the codecs constant. For codec n there are $M \cdot n = M \cdot (\log_2(S_n) - 1)$ information bits per block of M channel symbols, and for all codecs there are M parity bits in a block of M channel symbols of length M. The SE of codec n now becomes $R_n = \log_2(S_n) - 1 = n$ information bits. Fixing the choice of the block length as $M = 200$ symbols produce the parameters for the Gallager codecs in Table D.1. The coded bits are mod-

n	r_n	g_n	S_n	R_n
1	1/2	200	4	1
2	2/3	400	8	2
3	3/4	600	16	3
4	4/5	800	32	4
5	5/6	1000	64	5
6	6/7	1200	128	6

TABLE D.1: Simulated Gallager codes with block length $M = 200$ channel symbols.

ulated to channel symbols using Gray mapping. The iterative decoder in simulation results presented in this thesis is set to terminate if no errors are detected or if the number of iterations reached 100.

Note, there exists more than one Gallager code for a given rate and block length. Since the parity check matrices are generated in a stochastic manner, it is not claimed here that the codecs presented are the best codecs; rather a set of codecs that can be employed in a rate-adaptive scheme.

The BER-versus-CSNR performance for the codecs is shown in Figure D.4. Points in the figure were obtained using Monte Carlo simulations terminating when a minimum of 200 errors were found. An error is here defined as a block of $M = 200$ channel symbols decoded in error (not just one single information bit decoded in error). The points on the curves represent actual simulation results while the solid line represents an approximation of the BER-versus-CSNR performance for each of the codecs.

For the BER-versus-CSNR relationship of Gallager codes, MacKay *et.al.* have found a closed form approximation [MacKay and Hesketh, 1997]. In this thesis, a generalized version of the approximation is used (the generalization being the constant a_n, which in [MacKay and Hesketh, 1997] was set to 1):

$$\mathrm{BER}_n(\overline{\gamma}) = \frac{a_n}{(1 + e^{c_n(\overline{\gamma} - b_n)})^{d_n}} \,. \tag{D.22}$$

The parameters a_n, b_n, c_n, and d_n can be found using nonlinear curve fitting (results presented here were obtained employing the least squares method `lsqcurvefit` in the Optimization Toolbox of MATLAB).

The CSNR thresholds $\{\gamma_n\}_{n=1}^{N}$ depend on the target BER and can be calculated by demanding that $\mathrm{BER}(\gamma_n) = \mathrm{BER}_0$ for code n, and then using the inverse of the approximation describing the simulated data points

$$\gamma_n = \frac{1}{c_n} \ln\left(\left(\frac{a_n}{\mathrm{BER}_0}\right)^{1/d_n} - 1\right) + b_n. \tag{D.23}$$

Then, by setting $\mathrm{BER}_0 = 10^{-3}$ and 10^{-4} the CSNR thresholds $\{\gamma_n\}_{n=1}^{N}$ in Table D.2 were obtained. In this table the values found for the constants a_n, b_n, c_n, and d_n for all of the codecs are also shown. In Table D.3 the simulated CSNR thresholds for the $N = 6$ codecs where obtained for block lengths up to $M = 2000$ symbols.

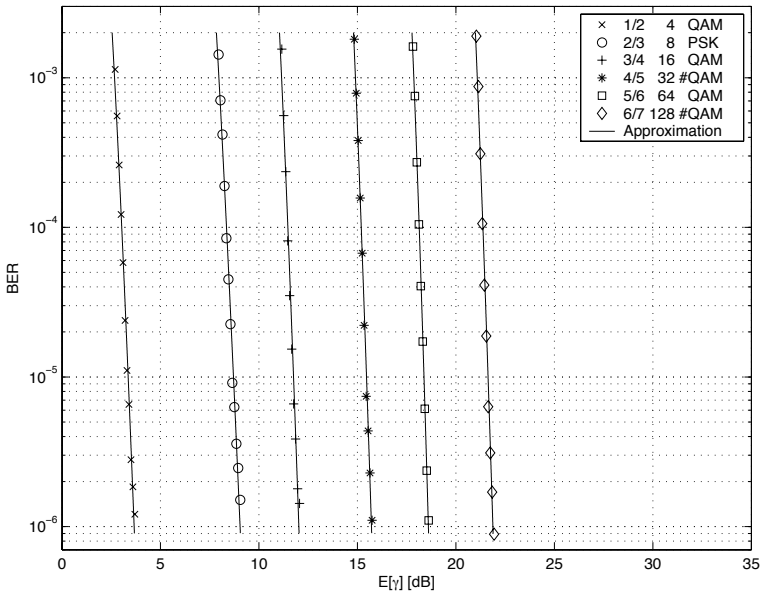


FIGURE D.4: Simulated and approximated BER-versus-CSNR for the Gallager based codecs.

n	a_n	b_n	c_n	d_n	γ_n (in dB)	
					$\text{BER}_0 = 10^{-3}$	$\text{BER}_0 = 10^{-4}$
1	14.0168	1.0620	4.2335	1.4013	2.7	3.1
2	13.0823	6.1563	4.0004	1.3176	8.0	8.4
3	13.4064	9.7041	4.3240	1.3151	11.4	11.8
4	13.4105	13.6210	4.5389	1.5632	14.9	15.3
5	0.0272	18.0941	6.5885	2.0968	18.3	18.5
6	0.0513	21.0407	4.5350	2.8911	21.3	21.5

TABLE D.2: Values of the parameters used in the estimation of the BER-versus-CSNR relationship, and the CSNR thresholds for $\text{BER}_0 \in \{10^{-3}, 10^{-4}\}$.

Block length M	CSNR thresholds (in dB)					
	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6
200	2.7219	8.0256	11.2111	14.9086	17.8923	21.0884
300	2.5665	7.8967	11.1120	14.8267	17.8091	21.0249
400	2.4910	7.8228	11.0699	14.7725	17.7721	20.9808
500	2.4356	7.7734	11.0349	14.7462	17.7479	20.9555
600	2.3992	7.7461	11.0075	14.7241	17.7338	20.9477
700	2.3690	7.7092	10.9962	14.7108	17.7276	20.9443
800	2.3469	7.7010	10.9717	14.6881	17.7120	20.9313
900	2.3307	7.6872	10.9619	14.6839	17.7149	20.9371
1000	2.3105	7.6672	10.9491	14.6653	17.7034	20.9283
1500	2.2545	7.6132	10.9145	14.6411	17.6863	20.9139
2000	2.2305	7.5991	10.8976	14.6076	17.6600	20.8903

TABLE D.3: CSNR thresholds for the $N = 6$ codecs for varying block lengths M.

References

M.-S. Alouini and A. J. Goldsmith.
Adaptive M-QAM modulation over Nakagami fading channels.
In *Proceedings of the 6th Communication Theory Mini-Conference (CTMC VI) in conjunction with IEEE Global Communications Conference (GLOBECOM'97)*, pages 218–223, Phoenix, AZ, USA, November 1997a.

M.-S. Alouini and A. J. Goldsmith.
Capacity of Nakagami multipath fading channels.
In *Proceedings of Vehicular Technology Conference, 1997, IEEE 47th*, volume 1, pages 358–362, Phoenix, AZ, USA, May 1997b.

M.-S. Alouini and A. J. Goldsmith.
Capacity of Rayleigh Fading Channels Under Different Adaptive Transmission and Diversity-Combining Techniques.
Wireless Personal Communications, 13:119–143, 2000.

G. Arredondo, W. Chriss, and E. Walker.
A multipath fading simulator for mobile radio.
IEEE Transactions on Vehicular Technology, VT-22:241–244, November 1973.

J. Baltersee, G. Fock, and H. Meyr.
An information theoretic foundation of synchronized detection.
IEEE Transactions on Communications, 49(12):2115–2123, 2001.

J. R. Barry, E. A. Lee, and D. G. Messerschmitt.
Digital Communication.
Kluwer Academic Publishers, third edition, 2004.

E. F. Carter.
Implements the Polar form of the Box-Muller Transformation.
Available at: ftp://ftp.taygeta.com/pup/c/boxmuller.c.

J. K. Cavers.
Variable-Rate Transmission for Rayleigh Fading Channels.
IEEE Transactions on Communications, COM-20(1):15–22, February 1972.

J. K. Cavers.
An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels.
IEEE Transactions on Vehicular Technology, 40(4):686–693, November 1991.

X. Chai and G. B. Giannakis.
Adaptive PSAM Accounting for Channel Estimation and Prediction Errors.
IEEE Transactions on Wireless Communications, 4(1):246–256, January 2005.

B. J. Choi and L. Hanzo.
Adaptive Modulation Mode Switching Optimization.
In L. Hanzo, C. H. Wong, and M. S. Yee, editors, *Adaptive Wireless Transcievers*, chapter 6. John Wiley & Sons, LTD, 1st edition, 2002.

S.-G. Chua and A. J. Goldsmith.
Adaptive Coded Modulation for Fading Channels.
In *Proceedings of IEEE International Conference on Communications*, pages 1488–1492, Montreal, Canada, June 1997.

M. C. Davey.
Error-correction using Low-Density Parity-Check Codes.
PhD thesis, University of Cambridge, 1999.

D. V. Duong and G. E. Øien.
Adaptive trellis-coded modulation with imperfect channel state information at the receiver and transmitter.
In *Proceedings of The Nordic Radio Symposium*, Oulu, Finland, August 2004.

ETSI.
HIPERLAN Type 2 Physical (PHY) layer.
Technical report, ETSI Broadband Radio Access Networks, 2000.

S. Falahati, A. Svensson, M. Sternad, and H. Mei.
Adaptive Trellis-Coded modulation over Predicted Flat Fading Channels.
In *Proceedings of Vehicular Technology Conference, 2003, IEEE 58th*, Orlando, Florida, USA, October 2003.

R. G. Gallager.
Low-Density Parity-Check Codes.
M.I.T. Press, Cambridge, Massachusetts, 1963.

D. L. Goeckel.
Adaptive Coding for Time-Varying Channels Using Outdated Fading Estimates.
IEEE Transactions on Communications, 47(6):844–855, June 1999.

A. J. Goldsmith and S.-G. Chua.
Variable-Rate Variable-Power MQAM for Fading Channels.
IEEE Transactions on Communications, 45(10):1218–1230, October 1997.

A. J. Goldsmith and P. P. Varaiya.
Capacity of Fading Channels with Channel Side Information.
IEEE Transactions on Information Theory, 43(6):1986–1992, November 1997.

I. S. Gradshteyn and I. M. Ryzhik.
Table of Integrals, Series, and Products.
Academic Press, San Diego, CA, 4th edition, 1980.

D. Greenwood and L. Hanzo.
Characterization of Mobile Radio Channels.
In R. Steele and L. Hanzo, editors, *Mobile Radio Communications*. John Wiley & Sons, LTD, 2nd edition, July 1999.

S. Guo, S. X. Ng, and L. Hanzo.
LDPC Assisted Block Coded Modulation for Transmission Over Rayleigh Fading Channels.
In *Proceedings of Vehicular Technology Conference, 2003, IEEE 57th*, JEJU, Korea, April 2003.

L. Hanzo, C. H. Wong, and M. S. Yee.
Prologue.
In L. Hanzo, C. H. Wong, and M. S. Yee, editors, *Adaptive Wireless Transcievers*, chapter 1. John Wiley & Sons, LTD, 1st edition, 2002.

J. F. Hayes.
Adaptive Feedback Communications.
IEEE Transactions on Communication Technology, 16(1):29–34, 1968.

K. J. Hole, H. Holm, and G. E. Øien.

Adaptive Multidimensional Coded Modulation Over Flat Fading Channels.
IEEE Journal on Selected Areas in Communications, 18(7):1153–1158, July 2000.

K. J. Hole and G. E. Øien.
Spectral Efficiency of Adaptive Coded Modulation in Urban Microcellular Networks.
IEEE Transactions on Vehicular Technology, 50(1):205–222, January 2001.

H. Holm.
Adaptive Coded Modulation Performance and Channel Estimation Tools for Flat Fading Channels.
PhD thesis, NTNU, 2002.

H. Holm, G. E. Øien, M.-S. Alouini, D. Gesbert, and K. J. Hole.
Optimal design of adaptive coded modulation schemes for maximum average spectral efficiency.
In *Proceedings of IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Rome, Italy, June 2003.

O. Jetlund, G. E. Øien, and K. J. Hole.
Adaptive Gallager Coded Modulation on Rayleigh Fading Channels: Comparison of Simulated and Theoretical Performance.
Temporary Document TD(03) 040, Cooperation européenne dans le domaine de la recherche Scientifique et Technique - 273 (03), Barcelona, Spain, January 2003a.

O. Jetlund, G. E. Øien, K. J. Hole, and H. Holm.
Spectral Efficiency Bounds for Adaptive Coded Modulation with Outage Probability Constraints and Imperfect Channel Prediction.
In *Proceedings of The Nordic Radio Symposium*, Oulu, Finland, August 2004a.

O. Jetlund, G. E. Øien, K. J. Hole, and B. Holter.
Adaptive Gallager coded modulation scheme on Rayleigh fading channels: Comparison of simulated and theoretical performance.
In *Proceedings of NORSIG*, pages 67–72, Bergen, Norway, October 2003b.

O. Jetlund, G. E. Øien, K. J. Hole, V. Markhus, and B. Myhre.
Rate-Adaptive Coding and Modulation with LDPC Component Codes.
Temporary Document TD(02) 108, Cooperation européenne dans le domaine de la recherche Scientifique et Technique - 273 (02), Lisbon, Portugal, September 2002.

O. Jetlund, G. E. Øien, and H. Holm.
Switching thresholds for robust adaptive modulation with predicted channel state information: A step towards an analytical solution.
Temporary Document TD(04) 034, Cooperation européenne dans le domaine de la recherche Scientifique et Technique - 273 (04), Athens, Greece, January 2004b.

O. Jetlund, G. E. Øien, B. Holter, and K. J. Hole.
Optimal Channel Prediction in an Adaptive Gallager Coded Modulation Scheme on Rayleigh Fading Channels.
Temporary Document TD(03) 127, Cooperation européenne dans le domaine de la recherche Scientifique et Technique - 273 (03), Paris, France, May 2003c.

S. Lin and D. J. Costello, Jr.
Error Control Coding.
Pearson Prentice Hall, 2nd edition, 2004.

D. J. C. MacKay and C. P. Hesketh.
Performance of Low Density Parity Check Codes as a Function of Actual and Assumed Noise Levels.
Available at: http://www.inference.phy.cam.ac.uk/mackay/sensit.pdf, 1997.

M. D. MacLaren and G. Marsaglia.
Uniform Random Number Generators.
Journal of the ACM, 12(1):83–89, 1965.

J. I. Marcum.
A statistical theory of target detection by pulsed radar.
Research Memo RM-754, RAND Corporation, Santa Monica, California, December 1947.

J. I. Marcum.
A statistical theory of target detection by pulsed radar: Mathematical appendix.
Research Memo RM-753, RAND Corporation, Santa Monica, California, July 1948.

G. Marsaglia.
Generating discrete random variables in a computer.
Communications of the ACM, 6(1):37–38, 1963.

H. Meyr, M. Moeneclaey, and S. A. Fechtel.
Digital Communication Receivers.
John Wiley & Sons, Inc, 1998.

M. Nagao and M. Kadoya.
Two-variate exponential distributions and its numerical table for engineering applications.
Technical Report 20(178), Bull. Disas. Prev. Res. Inst., Kyoto Univ., March 1971.

M. Nakagami.
The *m*-distribution—A general formula of intensity distribution of rapid fading.
In W. C. Hoffman, editor, *Statistical Methods in Radio Wave Propagation*, pages 3–36. Pergamon Press (London, Oxford, New York, Paris), June 1960.
Proceedings of a Symposium held at the University of California, Los Angeles, June 18–20, 1958.

W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery.
Numerical Recipes in C - The Art of Scientific Computing.
Cambridge University Press, second edition, 2002.

J. G. Proakis.
Digital Communications.
Electrical and Computer Engineering. McGraw-Hill, 4th edition, 2000.

T. S. Rappaport.
Wireless Communications: Principles & Practice.
Prentice Hall PTR, 1st edition, 1996.

C. Ridders.
A new algorithm for computing a single root of a real continuous function.
IEEE Transactions on Circuits and Systems, 26(11):979–980, November 1979.

L. Råde and B. Westergren.
Beta - Mathematics Handbook.
Studentlitteratur, Lund, 2nd edition, 1990.

C. E. Shannon.
A Mathematical Theory of Communication.
The Bell System Technical Journal, 27:379–423, 623–656, July, October 1948.

J. I. Smith.
A Computer Generated Multipath Fading Simulation for Mobile Radio.
IEEE Transactions on Vehicular Technology, VT-24(3):39–40, August 1975.

R. Steele and W. T. Webb.
Variable Rate QAM for Data Transmission Over Mobile Radio Channels.
In *Proceedings of WIRELESS '91*, pages 1–14, Calgary, Alberta, Canada, June 1991.

G. L. Stüber.
Principles of Mobile Communication.
Kluwer Academic Publishers, 2nd edition, 2001.

X. Tang, M.-S. Alouini, and A. J. Goldsmith.
Effect of Channel Estimation Error on M-QAM BER Performance in Rayleigh Fading.
IEEE Transactions on Communications, 47(12):1856–1864, December 1999.

N. M. Temme.
Special Functions: An Introduction to the Classical Functions of Mathematical Physics.
John Wiley & Sons, Inc, 1996.

B. Wheeler.
George Marsaglia's the mother of all random number generators.
Available at: ftp://ftp.taygeta.com/pup/c/mother.c.

D. J. Young and N. C. Beaulieu.
The Generation of Correlated Rayleigh Random Variates by Inverse Discrete Fourier Transform.
IEEE Transactions on Communications, 48(7):1114–1127, July 2000.

S. Zhou and G. B. Giannakis.
Adaptive Modulation for Multiantenna Transmissions With Channel Mean Feedback.
IEEE Transactions on Wireless Communications, 3(5):1626–1636, September 2004.

G. E. Øien, H. Holm, and K. J. Hole.
Adaptive Coded Modulation Enhanced by Channel Prediction and Antenna Diversity.
Temporary Document TD(02) 054, Cooperation européenne dans le domaine de la recherche Scientifique et Technique - 273 (02), Espoo, Finland, May 2002a.

G. E. Øien, H. Holm, and K. J. Hole.
Adaptive coded modulation with imperfect channel state information: System design and performance analysis aspects.
In *Proceedings of IEEE International Symposium on Advances in Wireless Communications (ISWC-2002)*, Victoria, BC, Canada, September 2002b.

G. E. Øien, H. Holm, and K. J. Hole.
Channel prediction for adaptive coded modulation in Rayleigh fading.
In *Proceedings of European Signal Processing Conference (EUSIPCO)*, Toulouse, France, September 2002c.

G. E. Øien, H. Holm, and K. J. Hole.
Impact of Imperfect Channel Prediction on Adaptive Coded Modulation Performance.
IEEE Transactions on Vehicular Technology, 53(3):758–769, May 2004.