Turbo Coding in Correlated Fading Channels

by

Danielle A. Hinton

S.B., Massachusetts Institute of Technology (2000)

Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degree of Masters of Engineering Degree

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2002

© Danielle A. Hinton, 2002. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author ....

Department of Electrical Engineering and Computer Science

May 14, 2002

Certified by...

..............................

Vahid Tarokh
Associate Professor
Thesis Supervisor

Accepted by.....

..............................

Arthur C. Smith
Chairman, Department Committee on Graduate Students
Turbo Coding in Correlated Fading Channels

by

Danielle A. Hinton

Submitted to the Department of Electrical Engineering and Computer Science
on May 14, 2002, in partial fulfillment of the
requirements for the degree of
Masters of Engineering Degree

Abstract

Turbo Code research has primarily focused on applying the codes to Additive White Gaussian Noise (AWGN) channels, and to Rayleigh fading channels where the fading parameter was fixed over the duration of a symbol. Turbo Coding has reached within 1dB of the Shannon limit in AWGN channels, and within 2dB of the Shannon limit in slow fading Rayleigh channels. However, in mobile radio – a primary application of Turbo Coding – the channel typically exhibits more rapid fading caused by the reflection of the signals off of scattering objects near the mobile receiver, and the doppler shift caused by the relative motion between mobile and base station. A model that more accurately captures this multipath effect was proposed by Jakes in 1965. In this thesis, Turbo Coding is applied to the fading channel simulated in the Jakes model. The Shannon capacity bounds of the Jakes channel with perfect side information is determined numerically using a Monte-Carlo simulation, and compared to the turbo code simulation results.

The Shannon capacity bounds predicted that the capacity of the fading channel should be less than the bounds in the AWGN channel, and should also be independent of the speed of the fade. The simulation results showed an approximate 12 - 15 dB increase in the SNR required to achieve a BER of $10^{-5}$ dB in time correlated fading channels over results in the AWGN channel. The simulation results also showed a dependency between the speed of the fade, and achievable BER. For a given SNR, as the speed of the fade increased, the BER decreased. This dependence results from the error exponent, which improves with the Doppler Spread since as the fade speed increases, the received symbols become more independent. In conclusion, this work has shown that the performance limiting factor of turbo codes in time correlated fading channels is the error exponent, which improves as the speed of the fade increases.

Thesis Supervisor: Vahid Tarokh
Title: Associate Professor
Acknowledgments

The steps that have been made in understanding the laws the govern the physical world are grand, yet still small in comparison to how much has not yet been understood. And what we can comprehend with our finite minds is minuscule in comparison to all that God has designed, and of little value in comparison to the life that was given so that we can have life eternally (John 3:16, John 17:3).

All glory, honor and praise goes to my Lord and Savior, Jesus Christ. I have been blessed in innumerable ways – with a loving family, wonderful friends, a blessed church, and the opportunity for education at MIT.

Very special thanks is due to Professor Vahid Tarokh for his guidance, encouragement, and patience over the past two years. Professor Tarokh’s deep interest and ability in problem solving, his dedication to education and research, and his profound respect for people continue to inspire me.

To my mother – always a supporter and always and inspiration – I cannot thank you enough for all that you have done, and cannot overstate how much I love and appreciate you. To my aunt Debbie, who keeps me laughing and encourages me to enjoy life – your love and encouragement mean so much to me. To my church family at Pentecostal Tabernacle, thank you for your prayers, support, and for helping me to stay focus on what is most important in life. And to my beloved friends from MIT, members of BWA, BCF, PraiseDance and especially the students of 35-316, thank you for always keeping me smiling!
# Contents

1 Introduction 6

2 Communication System Model 8  
2.1 Elements of the Communication System Model 8  
2.2 Jakes Fading Channel Model 10  
2.3 Capacity of the Jakes Fading Channel 13  
2.3.1 Capacity Derivation 15

3 Turbo Coding Overview 23  
3.1 Turbo Encoder 23  
3.2 Turbo Decoder 24  
3.2.1 The Maximum A Posteriori (MAP) Algorithm 25  
3.2.2 Implementing the Map Algorithm 28  
3.2.3 Map Variants 30  
3.2.4 Iterative Decoding 31

4 Previous Work Done on Turbo Codes 36  
4.1 Turbo Codes in Additive White Gaussian Noise (AWGN) 36  
4.2 Turbo Coding in Fading Channels 37

5 Results 39  
5.1 Discussion of Results 43

6 Conclusion 44
List of Figures

2-1 Standard Communication System ........................................ 9
2-2 Jakes Model ................................................................. 12
2-3 Correlation of Jakes Fading Model ...................................... 14
2-4 Jakes Capacity (SNR in dB) ............................................. 20
2-5 Jakes Capacity ............................................................... 21
3-1 Turbo Encoder ............................................................... 24
3-2 Turbo Decoder ............................................................... 34
5-1 Additive White Gaussian Noise Benchmark 1 ............................. 40
5-2 Jakes Channel BER, g1 10 Scatters .................................... 41
5-3 Jakes Channel BER, g2 10 Scatterers ..................................... 42
Chapter 1

Introduction

Shannon’s monumental formulation of a mathematical model of communication channels provided the bounds on the code rates that could be achieved through communication channels with minimal probability of error. One conclusion of his theory is that as block length increases, the probability of error decreases improving performance. Practically, long block lengths are not computationally feasible. It is well known that the complexity of maximum likelihood decoding algorithms increases with block length, up to a point where decoding becomes physically impossible. Coding theorists have achieved significant gains with error correcting block and trellis codes. They believed, however, that further gains would require increasingly complex decoding algorithms. The development of Turbo Codes in 1993 by Berrou, Glavieux, and Thitimajshima disproved those hypothesis reaching within 0.7 dB of the Shannon limit [1]. Since then, turbo codes have evolved speedily, both as a research area and in applications. Turbo codes were recommended for use in CDMA2000, the 3G standard. They also will be used in NASA’s next generation deep space transponder. Turbo codes have also been compressed successfully in video broadcast systems, where the associated delay is less critical than in wireless systems [2].

The strength of turbo codes is primarily in the independent looks at each bit provided to the decoder through encoding with two convolutional encoders when at least one bit stream is interleaved before it is encoded. At the decoder, soft information is used in an iterative decoding procedure to improve the hypothesis of
the transmitted bit with each iteration.

Research to date has primarily focused on applying turbo codes to Additive White Gaussian Noise (AWGN) channels in which the near Shannon capacity results were achieved, and to Rayleigh fading channels where the fading parameter was fixed over the duration of a symbol. However, in mobile radio – a primary application of Turbo Coding – the channel typically exhibits more rapid fading caused by the reflection of the signals off of scattering objects near the mobile receiver, and the doppler shift caused by the relative motion between mobile and base station. A model that more accurately captures this multipath effect was proposed by Jakes in 1965. In this thesis, Turbo Coding is applied to the fading channel simulated in the Jakes model. The shannon capacity bounds of the Jake channel is determined numerically using a Monte-Carlo simulation, and compared to the turbo code simulation results.

Chapter 2 of this thesis explains the communication system model, focusing specifically on the fading channel model used in this work. The capacity of the Jakes fading channel given perfect knowledge of the channel fade at both the transmitter and the receiver is also determined in this chapter. Chapter 3 reviews turbo encoding and decoding techniques. Chapter 4 reviews the work that has been done with turbo codes in Gaussian and Rayleigh fading channels. Chapter 5 presents and explains the results of this work.
Chapter 2

Communication System Model

This chapter explains the standard communication system model, and the fading channel model which will be used in this thesis.

2.1 Elements of the Communication System Model

The components of the standard communication system are shown in figure 1 [3]. Because of the Information Theoretic Source-Channel coding separation theorem, we know that asymptotically nothing is lost in separating the source coding from the channel coding. The information source could be discrete, like the text of an email message, or continuous like a voice signal. Regardless of whether the signal is analog or digital, the information source is modeled as a sample function of a random process [4]. The output of the information source is fed to the source encoder. At the source encoder, the input is represented as a sequence of symbols from a finite alphabet. The alphabet signals are then coded into fixed lengthed blocks of bits, as efficiently as possible. For an analog source, this means sampling the analog information signal at a rate greater then the Nyquist frequency, and then quantizing the samples, and representing the quantization levels using the symbol alphabet. The compressed data is then fed to the channel encoder to add redundancy in a controlled manner. That redundancy is used at the receiver to mitigate the effects of noise and interference encountered in the transmission of the signal through the channel. The
sequences emitted from the source encoder are called codewords. When the length of the information bits is $n$, and the length of the codework is $k$, the code rate is $R = n/k$. The codewords are then modulated onto signal waveforms for transmission through a communications channel. In this thesis, BPSK modulation will be simulated. Each bit value of 1 is represented as a 1, and each bit value of 0 is represented as $-1$. The modulated signal is then transmitted through a communication channel where it is corrupted by both noise and multipath fading. Multipath fading occurs when different propagation paths of the signal add destructively. The channel model is further explained in the next section.

At the receiver, the digital demodulator processes the channel-corrupted transmitted waveform, and reduces the waveform to a sequence of numbers that represent estimates of the transmitted data symbols. This sequence of numbers is passed to the channel decoder, which attempts to reconstruct the original information sequence from knowledge of the code used by the channel encoder, a model of the channel, and redundancy contained in the received data. The source decoder takes the output sequence from the channel decoder, and with full knowledge of the source encoding, attempts to reproduce the source signal. Because of errors introduced in the channel, and distortion introduced in quantization, only an approximation of the information
signal is constructed.

The two components of a communication system that this work focuses on is channel coding and decoding, using Turbo coding techniques, when the channel is modeled as introducing both a multipath fade, and an additive white Gaussian component.

2.2 Jakes Fading Channel Model

In urban areas, a primary region where mobile communications technology is used, the channel between the base station and the mobile receiver is characterized by many scattering objects (buildings, terrain, billboards), and a relative velocity between base station and mobile (when the mobile receiver is used in a vehicle). When the signals reflect off of nearby scattering objects, several copies of the signals, with varying amplitude and phases, are received at the mobile. When there is a relative velocity between the mobile and the base station, there will also be a doppler shift in the frequency of the signal. The doppler shift is given by:

\[ \omega_d = \frac{2\pi f_c v}{c} \cos(\theta) \] (2.1)

where \( v \) is the vehicle speed, \( c \) is the speed of light, \( f_c \) is the carrier frequency, \( \theta \) is the incident angle with respect to the direction of the vehicle’s motion, and is the \( w_M = \frac{2\pi f_c}{c} \) is the maximum doppler shift.

The traditional model for fading channels is the Rayleigh model, which assumes that many reflected waves (and not the line of sight (LOS) signal) are received at the mobile. Since the number of reflected waves is large, according to the central limit theorem, the two quadrature components of the channel impulse response are uncorrelated Gaussian random processes with zero mean and variance \( \sigma^2 \). The envelope of the channel has a Rayleigh probability distribution and the phases of the channel response is uniformly distributed between \(-\pi\) and \(\pi\). The Rayleigh probability distribution function is:
\[ p(a) = \frac{a}{\sigma^2} e^{-a^2/\sigma^2} \quad a \geq 0 \quad (2.2) \]

The Jakes model also captures the statistical properties of signals in a multipath environment by taking into account the geographic nature of communication between a base station, a moving receiver, and the scatterers. It places scattering objects that produce the multipath effects within a small distance of the mobile, and uniformly located around it, and assumes no line of sight (LOS) signal from the transmitter. It also incorporates the doppler shift of the reflected received waves. This results in a model that produces random phase modulation and a Rayleigh fading envelope. The mathematical model is shown in Figure 2.

The experimental setup is defined as follows. For ease of analysis, we assume that the mobile is fixed and the base moves along the axis with velocity \( \bar{v} \). The scattering objects are located in a circle of radius \( a \) around the transmitter at the mobile unit. The distance \( d \) to the base receiver will be assumed much greater than \( a \), so that the base does not lie within this circle. We also assume that \( d \) is so large that the angle \( \theta \) between \( \bar{v} \) and the direction to the mobile does not change significantly during observation times of interest. This assumption is necessary so that the movement of the base station along the \( x \)-axis is small compared to \( d \).

The model assumes that \( N \) equal-strength signals arrive at the moving receiver with uniformly distributed arrival angles, such that the signals experience a doppler shift \( \omega_d = \omega_M \cos(\alpha_n) \), where \( \omega_M = 2\pi fu/c \) is the maximum doppler shift [5].

Let \( R(t) \) represent the signal received at the mobile, originally transmitted with frequency \( \omega_c \), but subject to multipath effects because of the scatterers near the receiver. \( R(t) \) is the sum of \( N \) signals with arrival angles \( \alpha_n \) uniformly distributed between 0 and \( 2\pi \), with doppler shift \( \omega_d \cos(\alpha_n) \), delay \( \tau_i \), and Gaussian distributed fading coefficient \( A_i \).

\[
R(t) = \sum_{i=1}^{N} A_i e^{j(\omega_c + \omega_M \cos(\alpha_n))t - \omega_c \tau_i)} \quad (2.3)
\]
Figure 2-2: Jakes Model
The time correlation of the Jakes fading channel is derived by taking the expected value over the random phases of \( R(t) \):

\[
R(\tau) = E[R(t)R(t+\tau)]
\]

\[
= \frac{1}{2} ReE[T(t)T(t+\tau)e^{j\omega_c} + T^*(t)T(t+\tau)e^{j\omega_c*}] 
\]

\[
= \cos(\omega_c\tau)[4 \sum_{n=1}^{N} \cos(\omega_d\tau\cos(2\pi/N)) + 2\cos(\omega_d\tau)] 
\]

When \( N \) is large enough, the quantity in brackets is close to a discrete approximation to the integral:

\[
J_o(x) = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \cos(x\cos(\alpha)) d\alpha 
\]

which is the zeroth order Bessel function. The correlation of the Jakes multipath fading signal thus simplifies to:

\[
\frac{1}{2} \cos(\omega_c\tau)J_o(\omega_d\tau) 
\]

Figure (2-3) shows the deterministic correlation computed by Monte Carlo simulation methods. Superimposed on the plot is the zeroth order Bessel function, to which the deterministic correlation is theoretically equal.

The parameter \( \omega_d T \) is the fade rate normalized by the symbol rate \( T \), and is a measure of the channel memory. In this work, we used the a carrier frequency 2.1GHz in the cellular range in the United States, and a sampling period \( T = \frac{1}{24300} \)s. We varied the maximum doppler shifts to vary the parameter \( \omega_d T \) to span \( .001 \leq \omega_d T \leq .05 \), corresponding to mobile speeds ranging \( .55 \leq v \leq 61 \) miles/hour.

### 2.3 Capacity of the Jakes Fading Channel

Shannon’s theoretical bounds on the achievable rates of reliable communication through a given channel give us a upperbound on the rates we could achieve if we had
Figure 2-3: Correlation of Jakes Fading Model
perfectly random codes and infinite block length. Turbo Codes have come remark-
able close to the Shannon limit with block length $N = 65,336$, and in additive white
Gaussian noise. Hall showed in [6] that for a Rayleigh (slow) fading distribution, that
turbo codes could come to within .9 dB of the Shannon Limit. In this section, we
determine the Shannon limit of the Jakes Channel through Monte Carlo simulation,
and in Section 5, we compare the bounds to error rates determined through simulation
of turbo coding in the Jakes Channel.

2.3.1 Capacity Derivation

If we let $L$ be the length of a block of encoded data, to transmit that data through a
channel, we modulate each bit separately onto a waveform and send it independently
through the AWGN channel. Seeking to send $L$ bits of data with $L$ independent
uses of the channel is equivalent to the problem of sending data through $L$ parallel
channels when the noise is independent of noise in other channels. To derive the
capacity of the Jakes Channel, consider first the case of transmitting $\tilde{X}$ through the
Additive White Gaussian Noise (AWGN) channel [7].

$$Y_j = X_j + N_j \quad j = 1, 2, ..., L$$  \hspace{1cm} (2.9)

Given the number of bits to be sent $L$, the covariance matrix of the noise $\bar{N}$,
and the power constraint $P$, the capacity of the channel is:

$$C = \frac{1}{L} \max_{P(X_1, X_2, ..., X_L)} : E[X_i^2] \leq P I(X_1, X_2, ..., X_L; Y_1, Y_2, ..., Y_L)$$  \hspace{1cm} (2.10)
Rewriting these equations in terms of differential entropy yields:

\[
I(X_1, X_2, \ldots, X_L; Y_1, Y_2, \ldots, Y_L) = h(Y) - h(Y | X) \\
= h(X + N) - h(N|X) \\
= h(X + N) - h(N) - \sum_{i=1}^{L} h(N_i) \\
\]

Since differential entropy is maximized when random vector \( \tilde{X} \) (and thus \( \tilde{Y} \)), since \( \tilde{X} \) and \( \tilde{N} \) are independent) has a multivariate normal distribution the entropy of \( \tilde{Y} \), mutual information, and thus the capacity of the parallel channels is also maximized with this choice for the input distribution. Let the covariance matrix of \( \tilde{X} \) be \( K \). The capacity thus reduces to:

\[
h(X + N) - \sum_{i=1}^{L} h(N_i) = \log((2\pi e)^L |K + N|) - \log((2\pi e)^L |N|) \\
\leq \sum_{i=1}^{L} \log\left(1 + \frac{P_i}{N_i}\right) \\
\]

Equation (2.16) follows from the Hadamard Inequality which shows that the determinant of a function is upperbounded by the product of the elements on its diagonal. The maximum capacity is achieved, therefore, when \( \tilde{X} \) is \( L \) independent (and thus uncorrelated) Gaussian random variables.

The bits per channel use capacity is thus:

\[
C = \frac{1}{L} \left[ \sum_{i=1}^{L} \log\left(1 + \frac{P_i}{N_i}\right) \right] \\
\]

The problem has been reduced to determining the power levels \( P_i \) to allot to each channel, subject to the power constraint \( \sum_{i=1}^{L} P_i = P \). This optimization problem is solved using Lagrange multipliers. Letting
\[ J(\bar{P}, \lambda) = \frac{1}{2} \sum_{i=1}^{L} \log(1 + \frac{P_i}{N_i}) + \lambda \sum_{i=1}^{L} P_i \] (2.18)

and differentiating with respect to \( P_i \), yields

\[ P_i = \nu - N_i \] (2.19)

subject to the constraints that all the \( P_i \)'s must be positive, and that \( \sum P_i = P \).

Incorporating the Jakes fading parameter into our model, our model becomes:

\[ Y_j = \alpha_j X_j + N_j \]

\[ j = 1, 2, ..., L \] (2.20)

The quantity we wish to compute now is the capacity of the channel given that both the transmitter and receiver have perfect knowledge of the fading coefficient, or as called in the literature, the capacity given perfect channel state information, where \( A \) is a diagonal matrix of the \( L \) fading coefficients:

\[ C = \operatorname{max} \frac{1}{L} \operatorname{p}(X_1, X_2, ..., X_L) : E[X_i^2] \leq P I(X_1, X_2, ..., X_L; Y_1, Y_2, ..., Y_L | A) \] (2.21)

subject to the constraint:

\[ P = \frac{1}{n} \sum_{i=1}^{L} E[X_i^2] \]

\[ = \operatorname{trace}(K_x) \] (2.22)

(2.23)

where \( K_x \) is the covariance matrix of \( \bar{X} \). Again assuming Additive White Gaussian Noise that is independent of both the signal \( \bar{X} \), and of the fading coefficients, (2.21)
is equivalent to:

\[
I(\alpha_1 X_1, \ldots, \alpha_L X_L; Y_1, \ldots, Y_L|A) = h(\tilde{Y}|A) - h(\tilde{Y}|\tilde{X}, A) = h(A\tilde{X} + \tilde{N}|A) - h(A\tilde{X} + \tilde{N}|\tilde{X}, A) = h(A\tilde{X} + \tilde{N}|A) - h(\tilde{N}) = h(A\tilde{X} + \tilde{N}|A) - \sum_{i=1}^{L} h(N_i)
\] (2.24)

We again want to maximize \(h(A\tilde{X} + \tilde{N}|A)\), which occurs when \(\tilde{Y} = A\tilde{X} + \tilde{N}\) is a multivariate Gaussian distribution. Since the entropy of a Gaussian random vector is a function of the determinant of the covariance, our problem is reduced to maximizing \(K_y\). The covariance of \(\tilde{Y} = A\tilde{X} + \tilde{N}\) given \(A\):

\[
E[\tilde{Y}(\tilde{Y}^*)^T|A] = E[(AX + N)((X^*)^T(A^*)^T + (N^*)^T)|A] = AE[(X^*)^T(A^*)^T + \sigma^2_{1}I] = AK_x(A^*)^T + \sigma^2_{1}I
\] (2.28)

Since the additive Gaussian noise in each channel is independent, \(NN^T\) is a diagonal matrix with the noise variance in the \(i^{th}\) channel equal to \(\sigma^2_{i}\). Decomposing \(K_x\) into a diagonal matrix \(\Lambda\), and orthonormal matrix \(Q\) by a similarity transformation, and taking the determinant of \(K_y\), the following sequence of equalities hold:

\[
|AK_x(A^*)^T + \sigma^2_{1}I| = |AQ\Lambda Q^T(A^*)^T + \sigma^2_{1}I| (2.31)
\]

\[
= |AQ||\Lambda + (AQ)^{-1}\sigma^2_{1}I((A^*Q)^{-1}||AQ)^T| (2.32)
\]

\[
= |AA^T||QQ^T||\Lambda + \sigma^2_{1}(AQQ^T(A^*)^{-1}I) (2.33)
\]

\[
= |AA^T||\Lambda + \sigma^2_{1}(A(A^*)^{-1}I| (2.34)
\]

Since \(A\) is a diagonal matrix, \(|AA^T|\) is the product of the squared magnitude of
the fading coefficients. Similarly, the matrix \((AA^T)^{-1}\) has the reciprocal of the fading coefficients on its diagonals. Using Hadamard’s inequality, the following hold:

\[
|AA^T|\Lambda + \sigma_i^2(A(A^*)^T)^{-1}I = \prod_{i=1}^{L}(|\alpha_i|^2)(P_i + \sigma_i^2) \tag{2.35}
\]

\[
= \prod_{i=1}^{L}(\alpha_i^2 P_i + \sigma_i^2) \tag{2.36}
\]

The maximum entropy is thus:

\[
h(Y) = \sum_{i=1}^{L} \log((2\pi e)^{\frac{L}{2}}(\alpha_i^2 P_i + \sigma_i^2)) \tag{2.37}
\]

where \(P_i\) the power allotment to channel \(i\). Substituting (2.37) into (2.27), the capacity of the Jakes Fading Channel is upperbounded by:

\[
I(\alpha_1X_1, ..., \alpha_LX_L; Y_1, ..., Y_L|A) \leq \sum_{j=1}^{L} \log((2\pi e)(|\alpha_j|^2 P_i + \sigma_i^2)) - \sum_{i=1}^{L} h(\sigma_i^2)\tag{2.38}
\]

\[
= \sum_{i=1}^{L} \log\left(\frac{|\alpha_j|^2 P_i + \sigma_i^2}{\sigma_i^2}\right) \tag{2.39}
\]

\[
= \sum_{i=1}^{L} \log\left(1 + \frac{|\alpha_j|^2 P_i}{\sigma_i^2}\right) \tag{2.40}
\]

The problem has again been reduced to determining the power levels \(P_i\) to allot to each channel, subject to the power constraint \(\sum_{i=1}^{L} P_i = P\). The solution to this optimization problem is to choose the \(P_i's\) such that:

\[
P_i = \nu - \frac{N_i}{|\alpha_i|^2} \tag{2.41}
\]

subject to the constraints that all the \(P_i's\) must be positive, and that \(\sum P_i = P\).

This result is consistent with the literature [8], [9], [10].

The capacity of the Jakes Channel was computed numerically by Monte-Carlo Simulation. The fading parameter \(\omega_dT\) ranged from \(.001 \leq \omega_dT \leq .05\), corresponding to vehicle speeds ranging from \(.5 \leq v \leq 61\) mph. The results of those simulations are
shown in figure 2-4.

As expected, the capacity of the channel was independent of the fading speed. This makes sense because the water-filling solution tells the sender to pour more power into the channel with the least amount of effective noise, where the effective noise is \( \sigma^2 \). Since the mean and variance of the fading distribution envelope are independent of the speed of the fade, we expect the results to be independent of the speed of the fade.

As we also expected, the capacity of the fading channel was less than the capacity of the additive white Gaussian channel at high SNR. This is clear by Jensen’s Inequality. At low SNR, the fading channel capacity surpasses that of the additive white Gaussian noise channel. This is consistent with the results of Verdu and Shamai.
Figure 2-5: Jakes Capacity
in [11] which showed that fading with optimum power control is beneficial relative to the no fading case. The intuition behind this result is that it is possible to transmit any amount of information reliably with as little energy as desired by concentrating all the available transmitter energy at the rare moments when the fading level is very large. He explains that Jensen's Inequality is not violated because, with power control, the received power is higher than the transmitted power times the channel attenuation.
Chapter 3

Turbo Coding Overview

Turbo codes, developed in 1993 by Berrou, Glavieux, and Thitimajshima, are parallel concatenated codes which offer near-capacity performance for deep space and satellite communication channels in the power limited regime. Concatenated codes, first proposed by Forney, have the error-correction capability of much longer codes, and a structure that allows for low complexity decoding algorithms. The strength of turbo coding lies both the simple structure of the encoder, the resistance to error provided by the interleaver, and the iterative and soft decoding procedure. All three of those elements are explained in this chapter.

3.1 Turbo Encoder

A rate 1/3 Turbo encoder is composed of two constituent recursive systematic convolutional (RSC) encoders separated by an interleaver, as shown in figure 3. The outputs \( e_0, e_1, e_2 \), are multiplexed to obtain the rate 1/3 code.

To ensure strongest error performance, recursive systematic convolutional encoders should be used. The systematic form imposes fewer constraints on the code, and results in better BER performance since channel errors are not propagated. A recursive structure is required because it produces an interleaving performance gain, which is crucial for overall performance gain [12].

The Interleaver rearranges the ordering of the data sequence in a one-to-one de-
terministic format. It decorrelates the inputs to the two decoders, effectively providing a second look at the information bits which are uncorrelated from the non-interleaved bits received at the encoder. These two independent looks are used to provide a priori information in the decoding iterations.

A puncturer can be added to the encoder to turn the rate 1/3 encoder into a rate 1/2 encoder. Its role is to periodically delete selected bits to reduce coding overhead. For the case of iterative decoding, it is preferable to only delete parity bits.

### 3.2 Turbo Decoder

Optimal decoding of received signals is achieved using Maximum Likelihood (ML) decoding (assuming bits are equally likely a priori). Since the number of comparisons required grows exponentially with the block length of the code ($2^N$ codewords), direct application of ML decoding is infeasible. Two classes of decoding techniques have been explored for the turbo decoder:

1. Maximum A Posteriori (MAP) Algorithms
2. Sub-Optimal Soft Output Viterbi Algorithms (SOVA)

In a seminal paper by Berrou et al., the Bahl algorithm, which uses MAP decoding, was modified for use in decoding. When used to decode convolutional codes, the algorithm is optimal in terms of minimizing the decoded BER because it examines every possible path through the convolutional decoder trellis. The SOVA algorithm was initially proposed because it is less complex and less computational demanding than the BCJR algorithm. The Viterbi algorithm, from which the SOVA is derived, minimizes the probability of an incorrect path through the trellis being selected by the decoder [2]. It minimizes the number of groups of bits associated with these trellis paths, rather than the actual number of bits, which are decoded incorrectly [13]. In practice, many have shown that this results in approximately a 0.7 dB increase in the BER for decoding turbo codes over decoding with the MAP algorithm. Because of its superior performance, the MAP algorithm will be used in this thesis.

The strength of turbo codes is largely due of the iterative decoding technique used in the decoding algorithm. The basic idea behind the iterative decoding used for turbo codes decoding will also be explained below.

3.2.1 The Maximum A Posteriori (MAP) Algorithm

In 1974 Bahl at al. proposed the MAP algorithm to estimate the a posterior probabilities of the states and transtions of a Markov source observed in memoryless noise on both block and convolutional codes. When used to decode convolutional codes, the algorithm is optimal in terms of minimizing the decoded BER because it examines every possible path through the convolutional decoder trellis. The MAP algorithm provides not only the estimated bit sequence, but also the probabilities for each bit that it has been decoded correctly.

The MAP algorithm computes, for each decoded bit $u_k$, the probability that the bit was $\pm 1$ (BPSK), given the received sequence $\tilde{y}$. In this case, we will also condition on knowledge of the fading coefficients. We will not attempt to take advantage of the correlation properties of the fade coefficients in decoding. Knowledge of
the distribution of the fade will, however, effect the codes we choose for the com-
ponent convolutional encoders. Written in terms of the log-likelihood ratio (LLR), this
probability is:

\[
L(u_k | \overline{y}, \overline{a}) = \ln \left( \frac{P(u_k = +1 | \overline{y}, \overline{a})}{P(u_k = -1 | \overline{y}, \overline{a})} \right) \quad (3.1)
\]

Considering now the state transitions in the trellis of the convolutional code that
produce bit \( u_k \), (3.1) can be rewritten as:

\[
L(u_k | \overline{y}, \overline{a}) = \ln \left( \frac{\sum_{(s', s) \Rightarrow u_k = +1} P(S_{k-1} = s', S_k = s | \overline{y}, \overline{a})}{\sum_{(s', s) \Rightarrow u_k = -1} P(S_{k-1} = s', S_k = s | \overline{y}, \overline{a})} \right) \quad (3.2)
\]

where the previous state is \( S_{k-1} \), and the current state is \( S_k \). Using Bayes rule, this
is equivalent to:

\[
L(u_k | \overline{y}, \overline{a}) = \ln \left( \frac{\sum_{(s', s) \Rightarrow u_k = +1} P(S_{k-1} = s', S_k = s, \overline{y}, \overline{a})}{\sum_{(s', s) \Rightarrow u_k = -1} P(S_{k-1} = s', S_k = s, \overline{y}, \overline{a})} \right) \quad (3.3)
\]

where \((s, s') \Rightarrow u_k = +1\) is the set of transitions from the previous state \( S_{k-1} = s' \)
to the present state \( S_k = s \) that can occur if the input bit \( u_k = +1 \), and likewise for
\((s', s) \Rightarrow u_k = -1\).

By partitioning the received sequence \( \overline{y} \) into three sections: \( y_k, y_{j<k} \) and \( y_{j>k} \),
and by using the assumption that the channel is memoryless, the future received
sequence depends only on the present state \( s \), and not on the previous state \( s' \) or the
previous received channel sequences. The following sequence of equalities hold:

\[
P(s', s, \overline{y}, \overline{a}) = P(s', s, \overline{y}_{j<k}, y_k, \overline{y}_{j>k}, \overline{a}) = P(\overline{y}_{j>k}|s, \overline{a}) \cdot P(s', s, \overline{y}_{j<k}, y_k, \overline{a}) = P(\overline{y}_{j>k}|s, \overline{a}) \cdot P([y_k, s]|s', \overline{a}) \cdot P(s', \overline{y}_{j<k}, \overline{a}) \quad (3.4)
\]

Renaming the terms in (3.6),

\[
\alpha_{k-1}(s') = P(\overline{y}_{j<k}, S_{k-1} = s', \overline{a}) \quad (3.7)
\]
is the probability that the trellis is in state \( s' \) at time \( k - 1 \) and the received channel sequence up to this point is \( \bar{y}_{j<k} \):

\[
\beta_k(s) = P(\bar{y}_{j>k}|S_k = s, \bar{a})
\]  

(3.8)

is the probability that given the trellis is in state \( s \) at time \( k \), the future received channel sequence will be \( \bar{y}_{j>k} \), and

\[
\gamma_k(s', s) = P([y_k, S_k = s]|S_{k-1} = s', \bar{a})
\]  

(3.9)

is the probability that given the trellis is in state \( s \) at time \( k \), it moves to state \( s \) and the received sequence for this transition is \( y_k \).

The conditional LLR of \( u_k \) given the received sequence \( \bar{y} \) can thus be written as:

\[
L(u_k|\bar{y}, \bar{a}) = \ln \left( \frac{\sum_{(s',s)\Rightarrow u_k=+1} P(S_{k-1} = s', S_k = s, \bar{y}, \bar{a})}{\sum_{(s',s)\Rightarrow u_k=-1} P(S_{k-1} = s', S_k = s, \bar{y}, \bar{a})} \right)
\]

(3.10)

\[
= \ln \left( \frac{\sum_{(s',s)\Rightarrow u_k=+1} \alpha_{k-1}(s') \cdot \gamma_k(s', s) \cdot \beta_k(s)}{\sum_{(s',s)\Rightarrow u_k=-1} \alpha_{k-1}(s') \cdot \gamma_k(s', s) \cdot \beta_k(s)} \right)
\]

(3.11)

The MAP algorithm finds \( \alpha_k(s) \) and \( \beta_k(s) \) for all states \( s \) throughout the trellis, and \( \gamma_k(s', s) \) for all possible transitions from state \( S_{k-1} = s' \) to state \( S_k = s \). These values are used to compute (3.11). The decision rule:

\[
\tilde{d}_k = \begin{cases} 
0 & \text{if } L(u_k|\bar{y}, \bar{a}) < 0 \\
1 & \text{if } L(u_k|\bar{y}, \bar{a}) > 0
\end{cases}
\]

is then employed to give the most likely bit values that were sent.
3.2.2 Implementing the Map Algorithm

The probabilities represented by the variable \( \alpha_k(s) \), \( \beta_k(s) \) and \( \gamma_k(s, s') \) are determined as follows. Recall our definition of \( \alpha_k(s) \):

\[
\alpha_k(s) = P(S_k = s, \bar{y}_{j<k+1}, \bar{a})
\]

(3.12)

\[
= \sum_{s'} P(s', s, y_{j<k}, y_k, \bar{a})
\]

(3.13)

where in the second equality, we’ve summed the joint probabilities over all possible previous states (which are mutually exclusive and collectively exhaustive, thus equalling our initial expression for \( \alpha_k(s) \)). Using Bayes Rule, and the memoryless assumption of the channel, the following set of equalities are derived:

\[
\alpha_k(s) = \sum_{s'} P(s', s, y_{j<k}, y_k, \bar{a})
\]

(3.14)

\[
= \sum_{s'} P(s, \bar{y}_k|s', \bar{a})P(s', y_{j<k}, \bar{a})
\]

(3.15)

\[
= \alpha_{k-1}(s')\gamma_k(s', s)
\]

(3.16)

Thus, given the \( \gamma_k(s', s) \) values, \( \alpha_k(s) \) can be calculated recursively. The recursion is begun by assuming that the trellis begins in \( S_0 = 0 \) with probability one, so:

\[
\alpha_0(S_0 = 0) = 1
\]

(3.17)

\[
\alpha_0(S_0 = s') = 0 \quad (\forall s' \neq 0)
\]

(3.18)

The recursion for \( \beta_k(s) \) is derived in a similar fashion. Recall from equation (3.8):

\[
\beta_k(s') = P(\bar{y}_{j>k}, \bar{y}_k|s', \bar{a})
\]

(3.19)

Again using Bayes Rule and the memoryless assumption, the following sequence of
equalities hold:

\[ P(\bar{y}_{j>k}, \bar{y}_k|s', \bar{a}) = \sum_s P(\bar{y}_{j>k}, \bar{y}_k, s|s', \bar{a}) \]
\[ = \sum_s \frac{P(\bar{y}_{j>k}, \bar{y}_k, s, s', \bar{a})}{P(s', \bar{a})} \]
\[ = \sum_s \frac{P(\bar{y}_{j>k}|\bar{y}_k, s, s', \bar{a})P(\bar{y}_k, s, s', \bar{a})}{P(s', \bar{a})} \]
\[ = \sum_s \frac{P(\bar{y}_{j>k}|s, \bar{a})P(\bar{y}_k, s|s', \bar{a})P(s', \bar{a})}{P(s', \bar{a})} \]
\[ = \sum_s \beta_{k+1}(s)\gamma_k(s, s') \quad (3.24) \]

Once the \( \gamma_k(s, s') \) values are known, a backward recursion computes the remaining \( \beta_k(s') \) values. When the code is terminated, the code is in \( S_k = 0 \) in the trellis at time \( k \). So, the initializing values for the Beta recursion are:

\[ \beta_k(S_k = 0) = 1 \quad (3.25) \]
\[ \beta_k(S_k = s') = 0 \quad (\forall s' \neq 0) \quad (3.26) \]

If the code is unterminated, it is reasonable assumed that the trellis is in each state with equal probability \((N = \text{number of states})\):

\[ \beta_k(S_k = s) = 1/N \quad (\forall s) \quad (3.27) \]

The Gamma values are calculated from the received channel sequence, then used both in the iterations to determine \( \alpha_k(s) \) and \( \beta_k(s') \), then again in the final calculation of the likelihood ratio at each time \( k \). Using the definition of \( \gamma_k(s', s) \), Bayes Rule, and the memoryless assumption, we can derive the following equalities:
\[
\gamma_k(s, s') = P(\tilde{y}_k, s|s', \bar{a}) \\
= P(\tilde{y}_k|s, s', \bar{a})P(s|s') \\
= P(\tilde{y}_k|s, s', \bar{a})P(u_k) \\
= P(\tilde{y}_k|\bar{x}_k, \bar{a})P(u_k)
\]

The a-priori probability \(P(u_k)\) is either \(\frac{1}{2}\), or derived from the other decoder when decoding is done iteratively. The conditional received sequence probability \(P(\tilde{y}|\bar{x})\) are calculated, assuming a memoryless Gaussian channel with BPSK modulation, as follows:

\[
P(\tilde{y}_k|\bar{x}_k, \bar{a}) = P(y^s_k|u_k, \bar{a})P(y^p_k|x^p_k, \bar{a}) \\
= \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}[(y^s_k-a^s_ku_k)^2+(y^p_k-a^p_kx^p_k)^2]}
\]

where \(y^s_k\) is the systematic received bit, \(y^p_k\) is the received parity bit (and corresponding parameters for the transmitted sequence \(\bar{x}_k\)), \(a^s_k\) is the fading coefficient of the systematic bit, and \(a^p_k\) is the fading coefficient of the received parity bit— all at time \(k\).

### 3.2.3 Map Variants

The Map algorithm is optimal for the decoding of turbo codes, but because of the multiplications used in the recursive calculations required for implementation, numerical errors tend to propagate. The Log-Map algorithm is theoretically identical to the MAP algorithm, but transfers operations to the log domain where multiplications are turned into additions, dramatically reducing complexity. The Max-Log-MAP algorithm is another alternative to the MAP algorithm which uses an approximation of the log of a sum of numbers. Due to this approximation, it gives a degraded performance compared to the MAP and Log-Map algorithms, and is not used in this
thesis.

The Log-MAP algorithm is implemented as follows. The log of each of the component probabilities are taken:

\[ A_k(s) = \ln(\alpha_k(s)) \]  
\[ B_k(s') = \ln(\beta_k(s')) \]  
\[ \Gamma_k(s) = \ln(\gamma_k(s)) \]

This allows the forward and backward recursions to compute \( A_k(s) \) and \( B_k(s') \) to be additions instead of multiplies.

\[ A_k(s) = \sum_{all s'} \exp[A_{k-1}(s') + \Gamma(s', s)] \]  
\[ B_{k-1}(s') = \sum_{all s} \exp[B_k(s) + \Gamma_k(s', s)] \]

The calculation for \( \Gamma_k(s) \) reduces from an exponential to the sum of three terms:

\[ \Gamma_k(s', s) = \frac{1}{2} u_k L(u_k) + \frac{L_c}{2} (y_k u_k + y_k^p x_k^p) \]

The total expression for the Log Likelihood ratio is thus:

\[ \ln\left( \sum_{(s',s) \Rightarrow u_k = +1} \exp[A_{k-1}(s') + \Gamma_k(s', s) + B_k(s)] \right) - \ln\left( \sum_{(s',s) \Rightarrow u_k = -1} \exp[A_{k-1}(s') + \Gamma_k(s', s) + B_k(s)] \right) \]

3.2.4 Iterative Decoding

Recall the MAP decoder log likelihood ratio of equation (12), and gamma of equation (10). The parameter \( \gamma_k(s, s') \) can be written as:
where $\bar{x}$ is the transmitted symbol, and $\bar{y}$ is the received symbol corresponding to bit $u_k$. Letting

$$L(u_k) = \ln\left(\frac{P(u_k = 1)}{P(u_k = -1)}\right)$$

(3.43)

and noting that $P(u_k = 1) = 1 - P(u_k = -1)$, with a bit of algebra it can be shown that:

$$P(u_k = 1) = \frac{e^{L(u_k)}}{1 + e^{L(u_k)}}$$

(3.44)

$$P(u_k = -1) = \frac{1}{1 + e^{L(u_k)}}$$

(3.45)

Observe that the following manipulation captures $P(u_k)$ when $u_k = 1$ and when $u_k = -1$:

$$P(u_k) = \frac{e^{(L(u_k)/2)}}{1 + e^{L(u_k)}} e^{(u_k L(u_k/2))}$$

(3.46)

$P(u_k)$ is thus equal to $A_k e^{(u_k L(u_k/2))}$, where $A_k$ is independent of $u_k$. The second term of $\gamma_k(s', s)$, $P(\bar{y}_k|\bar{x}_k, \bar{a})$, is:

$$P(\bar{y}_k|\bar{x}_k, \bar{a}) = B e^{-\frac{(y_k - \bar{x}_k)^2}{2\sigma^2} - \frac{(y_k - \bar{y}_k)^2}{2\sigma^2}}$$

$$= B e^{-\frac{u_k y_k^2 + u_k^2 y_k^2 + y_k^2 + \sigma^2}{2\sigma^2} + \frac{u_k y_k^2 + u_k^2 y_k^2}{\sigma^2}}$$

$$= C_k e^{-\frac{u_k y_k^2 + u_k^2 y_k^2}{\sigma^2}}$$

(3.47)

(3.48)

(3.49)

where $C_k$ is independent of the bit $u_k$ (since $u_k^2 = 1$). The resulting equation for
\( \gamma_k(s', s) \) is:

\[
\gamma_k(s', s) = A_k C_k e^{(u_k L(u_k/2))} e^{- \frac{u_k^2 + y_k^2}{\sigma^2}}
\]  

(3.50)

Since we assume BPSK modulation, \( \frac{1}{\sigma^2} = \frac{E_b}{N_0/2} \), we can simplify (3.50) further to be:

\[
\gamma_k(s', s) = e^{\frac{1}{2} u_k (L(u_k) + L_c y_k^2) + \frac{1}{2} L_c y_k^2 s_k^2}
\]  

(3.51)

\[
= e^{\frac{1}{2} u_k (L(u_k) + L_c y_k^2) \gamma_k^E(s', s)}
\]  

(3.52)

where \( L_c = \frac{4E_b}{N_0} \) and

\[
\gamma_k^E(s', s) = e^{\frac{1}{2} L_c y_k^2 s_k^2}
\]  

(3.53)

Substituting (3.53) into equation (3.11) yields:

\[
L(u_k | \overline{y}, \overline{a}) = \ln \left( \frac{\sum_{(s', s) \Rightarrow u_k = +1} \alpha_{k-1}(s') \cdot \gamma_k^E(s', s) D^+_{k} \cdot \beta(s)}{\sum_{(s', s) \Rightarrow u_k = -1} \alpha_{k-1}(s') \cdot \gamma_k^E(s', s) D^-_{k} \cdot \beta(s)} \right)
\]  

(3.54)

\[
L(u_k | \overline{y}, \overline{a}) = L(u_k) + L_c y_k^s + \ln \left( \frac{\sum_{(s', s) \Rightarrow u_k = +1} \alpha_{k-1}(s') \cdot \gamma_k^E(s', s) \cdot \beta(s)}{\sum_{(s', s) \Rightarrow u_k = -1} \alpha_{k-1}(s') \cdot \gamma_k^E(s', s) \cdot \beta(s)} \right)
\]  

(3.55)

where \( D_k = e^{\frac{1}{2} u_k (L(u_k) + L_c y_k^2)} \). Equation (3.55) follows because \( D_k \) can be factored out of the functions in the numerator and denominator. The first term, \( L(u_k) \), is the a-priori LLR. This probability should come from an independent source. In most cases, there will be no independent or a-priori knowledge, so the LLR will be zero, corresponding to the bits being equally likely. But, when iterative decoding is used, each component decoder can provide the other decoder with an estimate of the a-priori LLR.

The second term \( L_c y_k \) is the channel reliability measure and is given by:

\[
L_c = \frac{4a_k}{2\sigma^2}
\]  

(3.56)
where $a_k$ is the $k^{th}$ fading coefficient. The term $y_k^s$ is the received version of the transmitted systematic bit $x_k^n = u_k$. When the SNR is high, the channel reliability value $L_c$ will be high, and have a large impact on the a-posteriori probability. Conversely, when there is a low SNR, it will have less impact on the a-posteriori probability.

The third term, $L_e(u_k)$ is called the extrinsic information provided by a decoder. It is derived using the constraints imposed on the transmitted sequence by the code used, from the a-priori information sequence $L(\bar{u}_k)$, and the received channel information sequence $\tilde{y}$, excluding the received systematic bit and the a-priori information $L(u_k)$ for the bit $u_k$. Equation (3.55) shows that the extrinsic information can be obtained by subtracting the a-priori information $L(u_k)$ and the received systematic channel input from the soft output $L(u_k|\tilde{y}, \bar{a})$ of the decoder.

The structure of the Turbo Decoder is shown in Figure 3-2. The iteration works as follows: The first component decoder receives the channel sequence $L_c\tilde{y}$ containing the received versions of the transmitted bits, and the parity bits from the first en-
coder. It then processes the soft channel inputs and produces its estimate $L(u_k|\tilde{y}, \tilde{a})$ of the data bits $u_k = 1, 2, \ldots, N$. In this first iteration, no a-priori information is available, so $P(u_k) = 1/2$. Next, the second component decoder receives the channel sequence containing the interleaved version of the received bits, and the parity bits from the second encoder. It also uses the LLR $L(u_k|\tilde{y}, \tilde{a})$ provided by the first component decoder to find $L_e$, which it then uses as the a-priori LLRs $L(u_k)$.

For the second iteration, the first component encoder again processes its received channel sequence $\tilde{y}$, but now it has a-priori LLR $L(u_k)$ provided by the extrinsic information $L_e(u_k)$ computed from the a-posteriori LLRs calculated by the second encoder. Using the extrinsic information, an improved a-posteriori LLR can be generated. The iterations continue using the improved a-priori LLR $L(u_k)$ with the received channel sequence. As this iteration continues, the BER of the decoded bits falls. The improvement in performance for each additional iteration carried out falls as the number of iterations increases for SNR (dB) $\geq 0$ [13].
Chapter 4

Previous Work Done on Turbo Codes

This chapter reviews the work done in both Gaussian and Rayleigh fading channels.

4.1 Turbo Codes in Additive White Gaussian Noise (AWGN)

In the monumental paper by Berrou et. al, the performance of Turbo Codes was simulated with block lengths of 65536 bits, and 128 frames. The generator that they found to be optimal was \( g = [11111; 100011] \). The AWGN simulations resulted in BER \( = 10^{-5} \) at \( E_b/N_0 = 0.7 \) dB. The Shannon limit for binary modulation with \( R = 1/2 \) is \( P_e = 0 \) for \( E_b/N_0 = 0 \) dB.

Since the initial publication, a plethora of papers have been written, investigating the use of other coding schemes with the parallel concatenated code structure, and the iterative version of the BCJR algorithm for decoding. Among them:

- Helmut, Hagenauer, and Burkert were able to get closer to the Shannon capacity limit by applying long, but simple Hamming Codes as component codes to an interative turbo-decoding scheme. The complexity of the soft-in/soft-out decoding of binary block codes is high, but they were able to get within 0.27
dB of the Shannon capacity limit [6].

- Yamashita, Ohgane and Ogawa proposed a new transmit diversity scheme implementing space division multiplexing with turbo coding. The scheme uses space-time coding to ensure easy decoding at the receiver. Their simulation results showed that the new scheme performs better than other space-time turbo coding schemes.

- Fan and Xia proposed a joint turbo coding and modulated coding for ISI channels with AWGN. Simulations show that their proposed joint coding method mitigates the ISI and noise. They also showed that with turbo coding, they were able to approach the capacity of a channel with ISI in AWGN.

The results obtained by Berrou et al. will be used as the benchmark for this thesis work.

### 4.2 Turbo Coding in Fading Channels

In [6, 12], a Rayleigh slow-fading channel was considered, with the discrete representation of the channel \( y_i = a_i x_i + n_i \), where the \( x_i \)'s are BPSK symbol amplitudes and the \( n_i \)'s are white Gaussian Noise with zero mean and variance \( N_o/2 \). They assumed the presence of sufficient interleaving of the \( y_i \)'s, so that the fading coefficients are independent and identically distributed random variables with a Rayleigh density. The capacity of the fading channels were computed numerically. Two capacity values of interest for comparison to this work are:

<table>
<thead>
<tr>
<th>Code Rate</th>
<th>Rayleigh ( E_b/N_o )</th>
<th>AWGN ( E_b/N_o )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>1.8dB</td>
<td>0.2dB</td>
</tr>
<tr>
<td>1/3</td>
<td>1.4dB</td>
<td>-0.5dB</td>
</tr>
</tbody>
</table>

In the work of [12], a block size \( N = 1024 \) bits, and 8 decoding iterations were used to simulate Turbo Coding in the Rayleigh fading environment. Simulation results
showed that with perfect channel state information, a BER of $10^{-5}$ was achieved when $E_b/N_o$ was approximately 2.4 dB. This is approximately 1 dB from the Shannon Limit.

In the work of [6], turbo codes of rate 1/3 were simulated at SNRs ranging from 0 to 12 dB. With $N = 5000$ bits, and 8 iterations, He was able to achieve a BER of $10^{-5}$ at .7 dB from the capacity for the fully-interleaved fading channel. He also briefly considered the case of correlated fading channels, and noted that turbo codes performed considerably better in the slower fading process ($\omega_d T = .001$) compared to the faster fading process ($\omega_d T = .1$). In the slower-fading process, a BER = $10^{-5}$ was achieved at approximately $E_b/N_o = 5$ dB. In the faster-fading process, a BER $10^{-5}$ was not achieved with SNR $\leq 12$ dB.
Chapter 5

Results

Recall from Chapter 2 that we calculated by Monte-Carlo simulation that the capacity of a fading channel simulated with the Jakes Model should be independent of the speed of the fade. Our intuition was that since the mean and variance of the envelope were unaffected by the speed of the fade, our waterfilling in time solution should be able to fill the same amount of power at a given signal to noise ratio, regardless of the speed of the fade. To determine how turbo codes will fare when simulated, we've varied the speed of the fade between $0.001 = \omega_d T$ to $0.05 = \omega_d T$ and compared it to the benchmark of the AWGN case, which is well established in the literature (but also simulated here to ensure optimal performance of the code).

With a $1/3$ rate turbo code using the generator matrix $g_1 = [1 1 1 ; 1 0 1]$ for both of the convolutional encoders; and with a frame size of $N = 65536$ bits, 128 frames, and with 5 decoding iterations, the following results were acheived. (recall, the Shannon limit for a rate $1/3$ code is $-0.5$ dB).

These results are approximately $1.5$dB from the Shannon Limit, but this difference has resulted from the smaller constraint length of the two component convolutional encoders, and the comparatively small number of decoding iterations (in the turbo code results of Berrou et al., 18 decoding iterations, and the generator $g_2 = [11111; 10001]$ with constraint length $k = 5$ were used. Results with the optimal generator matrix, and increased decoding iterations are shown later.

To incorporate the Jakes Fading model in to the computation, Jakes coefficients
Figure 5-1: Additive White Gaussian Noise Benchmark 1
were generated, and the BER curves were determined again using a frame length of 50000 and approximately 130 frames. The speed of the fade was varied between $0.001 < \omega_d T < 0.05$, which corresponds to the mobile varying between 0.5 and 61 mph. In the following plots, the presence of 10 scatterers was simulated.

As expected, BER rates of the fading curves are always higher than the BER when no fading is incorporated in the model. But, for the case of 10 scatters, we have obtained the seemingly counter-intuitive result that the BER decreases with the speed of the fade. We will return to the explanation of this result after considering other simulation results.

When the generator matrix $g2 = [1111; 1000]$ was used in an Additive White Gaussian Noise channel simulation using 10 decoding iterations, 130 frames of length
Figure 5-3: Jakes Channel BER, g2 10 Scatterers

N = 50,000 bits, a bit error rate of $10^{-3}$ was achieved at .3 dB. This is .8 dB from the Shannon limit. In the paper by Berrou, 18 decoding iterations were used, which would decrease the BER.

To incorporate the Jakes Fading model into the computation, Jakes coefficients were again generated for 130 frames of length 25000. The speed of the fade was varied between $.001 < \omega_d T < .05$, which corresponds to the mobile varying between .5 and 61 mph. For these plots, the number of scatterers was 10.

As expected, the performance of turbo codes in correlated fading channels is very much degraded compared to the performance of Turbo Codes in an AWGN channel. For a given BER, turbo codes in a correlated fading channel requires an additional 15 dB over turbo codes in a purely Additive White Gaussian Noise Channel. Since the
MAP decoding algorithm depends on the noise process begin Gaussian, the severe degradation is to be expected.

5.1 Discussion of Results

From the simulation plots, there is a clear dependence on the speed of the fade and the SNR required to achieve a given BER. As the speed of the fade increases, the SNR required to achieve a given BER decreases. This discrepancy was not predicted by the capacity calculations done in Chapter 2. In Chapter 2, it was predicted that the capacity of the Jakes Channel should be independent of the speed of the fade. The discrepancy between the capacity derivation in chapter two, and the simulated results, are not, however, contradictory. The dependency between fade speed and BER is related, rather, to the error exponent in a correlated fading channel.

The error exponent, which is also known as the “random coding reliability function,” is an upper bound on the probability of error for transmission of a random code through a given channel. In the proof of the Coding Theorem in [14], Gallager determined that the error exponent bound in a discrete memoryless channel is a function of the transmission rate, and the code length $N$. In [15], the random coding function reliability functions for various fading distributions was determined. To determine the effect of a time correlated fading channel, they modeled the process with the Jakes model, and derived the error exponent numerically.

As shown in [15], the error exponent provides the more stringent bounds on reliability achieved over fading channels. Results by Ahmed and McClane in [15] showed very poor reliability over the time correlated fading channel. They also showed by Monte-Carlo simulation that the error exponent improved with Doppler spread. This result is intuitive because as $\omega d T$ increase, fading time correlation decreases and the received symbols become more independent. This analysis is consistent with the simulation results of this thesis, and has provided an explanation for the fade speed dependent performance of the turbo codes simulations in time correlated fading channel modeled by Jakes.

43
Chapter 6

Conclusion

In this thesis, we’ve used the widely accepted Jakes model for the mobile wireless channel to determine how Turbo Codes, which have done remarkably well in Additive White Gaussian Noise channels, will fare in the more destructive time varying, and time correlated fast fading channel. Because of the nature of the channel and the channel model, closed-form solutions for its capacity and the performance of turbo codes in the channel cannot be found in closed-form. Thus, the capacity and performance were determined numerically by Monte-Carlo simulation. The numerical analysis of the capacity of the time correlated fading channel with perfect state information showed that the capacity of the channel is less than that of the AWGN channel, and that the capacity of the fast fading channel should be independent of the fade speed. Turbo Code performance simulation results showed very poor reliability achieved over the Jakes channel. Simulations also showed a dependency between the speed of the fade and the BER. This dependency was not predicted by the capacity simulation, but are not inconsistent with the capacity results. The simulation results reflect, however, both the limiting behavior of the error exponent in time correlated fading channels, and the dependency of the error exponent on the Doppler Spread that was shown in [15]. As the fading time correlation decreases and the received symbols become more independent, the error exponent improves. In effect, the results show that the error exponent is the performance bounding parameter for turbo codes in fast fading channels, and that that error exponent improves with the speed
of the fade.

These results suggest a few interesting further paths of research in both time correlated fading channels, and in Turbo Coding in time correlated fading channels. Firstly, the results of this work critically depend on the assumption of perfect channel state information. As fade speeds increase, the likelihood of having reliable estimates of the fading speed certainly decrease. Determining performance with less channel state information would be a useful result. Since the Jakes Fading Model is not tractable analytically, bounds on the capacity and error exponent in correlated fading channels seem like a reasonable way to approach the problem. Another avenue of research is to improve the performance of turbo codes in the time correlated fading channel. First keeping the knowledge of the channel state, codes for the two component convolutional encoders to maximize performance in fast fading channels can be researched. Secondly, interleavers that specifically aim at mitigating time correlation could be determined. Once again, a serious limiting factor in performance is the requirement of perfect channel state information. Bounds on performance with varying degrees of channel state information would again be useful. Research into fade characteristic predictive techniques would be very helpful in all of these cases, and when combined with solutions to the above problems, should reduce the gap to capacity in time correlated fading channels.
Bibliography


