1

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

**ERROR CONTROL CODING FOR LAND
MOBILE COMMUNICATIONS**

by

Bernard N.B. McPhail, B.Eng.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment for the degree of
Master of Engineering

Department of Systems and Computer Engineering
Faculty of Engineering
Carleton University
Ottawa, Ontario

July 1986

(c) copyright B. McPhail, 1986

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film. •

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-33416-9

The undersigned recommends to the Faculty of Graduate Studies and Research, acceptance of the thesis .

ERROR CONTROL CODING FOR LAND MOBILE COMMUNICATIONS

submitted by Bernard N.B. McPhail, B.Eng., in ~~partial~~ fulfillment of the requirements for the degree of Master of Engineering.

Thesis Supervisor

Chairman, Department of Systems
and Computer Engineering

Carleton University

October 1986

ABSTRACT

Digital mobile radio communication systems are subject to severe impairments. As a consequence, the bit error rate performance of these systems is usually unsatisfactory. A simple way to reduce the bit error rate is to protect the transmitted message with an error correcting code. The performance of the coded system will depend on the specific coding scheme, the code rate, and the code length. The way in which these three factors influence the bit error rate is investigated. In particular, equations which bound the performance of block coded schemes on mobile channels are given.

Practical coding strategies are evaluated based on field measurements done in the cities of Ottawa and Hull, Canada. The test results agree favourably with the theoretical bounds on performance.

Two notions that are dispelled in the thesis are that block codes that are used on mobile channels must have low rates, and that bit interleaving is beneficial to performance. Interleaving is best viewed as a way to reduce decoder complexity, but it does not improve performance.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis supervisor, Professor H.M. Hafez, for his continued guidance and support. The many stimulating sessions I held with him have contributed immensely to the development of this thesis. It was a privilege and an honour to work under him.

I also wish to thank my graduate colleagues and friends for making my stay at Carleton pleasant. The technical discussions in which we participated have been more enlightening than reading many books.

A heartfelt expression of gratitude goes to Shirley Yik whose constant encouragement has been invaluable during the difficult periods.

A debt of gratitude, which I cannot begin to repay, is owing to my family for their support and encouragement.

Lastly, I must thank Diane Dodds for her effort in typing the manuscript and in bringing the final version up to par.

This thesis is dedicated to the memory of my father.

TABLE OF CONTENTS

Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Figures	viii
List of Symbols	xi

Chapter 1 -- Introduction

1.1	General	1-1
1.2	Mobile Radio Channel Characteristics	1-4
1.3	Preparatory Work: The Transmission Experiments	1-7
	1.3.1 Overview	1-7
	1.3.2 Data Processing	1-11
1.4	Thesis Organization	1-16

Chapter 2 -- The Mobile Channel

2.1	Chapter Summary	2-1
2.2	Qualitative Description of Error Producing Mechanisms	2-2
2.3	Rayleigh Fading	2-5
2.4	Level Crossing Rates, Average Fade Duration, and Fade and Gap Statistics for Rayleigh Fading Channels	2-9
	2.4.1 Theory	2-9
	2.4.2 Experimental Results	2-23
2.5	The Distribution of the Number of Errors Per Block	2-34
2.6	Comparison of Theory and Experiment	2-45

Chapter 3 -- Coding Bounds

3.1	Introduction	3-1
3.2	Bounds on the Error Correction Ability of Block Codes, and on the Achievable Error Probabilities for the Discrete Land Mobile Channel: Outline of Approach	3-4
3.3	Bounds on Error Correcting Ability of Block Codes	3-8
3.3.1	Introduction	3-8
3.3.2	Definitions and Useful Results	3-12
3.3.3	The Elias Upper Bound	3-18
3.3.4	The Gilbert-Varsharmov Lower Bound	3-28
3.3.5	An Upper Bound For Burst Error Correction	3-29
3.3.6	A Lower Bound on the Burst Error Correction Ability of a Code	3-32
3.3.7	Erasure Burst Error Correction	3-33
3.3.8	Limiting Forms for the Elias and the Gilbert Bounds	3-35
3.3.9	Error Correction Ability of Block Codes: Concluding Summary	3-42
3.4	Experimental and Theoretical Results	3-52

Chapter 4 -- Practical Coding Schemes

4.1	Introduction	4-1
4.2	Code Performance Measures	4-2
4.3	The AMPS Coding Scheme	4-6
4.3.1	Outline of the AMPS Data Formats	4-7
4.3.2	Outline of the Simulations	4-11
4.3.3	Equations Used	4-14
4.3.4	Results	4-15
4.3.5	An Alternate Decoding Scheme	4-21
4.3.6	Outline of the Simulations	4-25
4.3.7	Simulation Results	4-28
4.4	Packet Error Rates With Hamming Codes	4-29
4.5	Conclusions	4-35

Chapter 5 -- Conclusions

5.1	Conclusions	5-1
5.2	Non-Optimality of Bounded Distance Decoding	5-2
5.3	Suggestions for Further Research	5-5
5.4	Recommendations	5-9

References

Ref-1

Appendix I

Bit Error Rates at 820 MHz

I-1

Appendix II

Summary of Available Error Pattern Files

II-1

Appendix III

A Short Table of the Elias Bound

III-1

Appendix IVA Short Table of the Gilbert-Varsharmov
Bound

IV-1

Appendix V

Design of the Shortened BCH (40,28) Decoder

V-1

Appendix VISchematic Diagrams for the Packet Scheme
Based on the (8,4) Extended Hamming Code

VI-1

LIST OF FIGURES

1-1	Mobile Radio Communication System	1-6
1-2	Experimental Set-up	1-10
1-3	Data Acquisition Equipment	1-14
1-4	Auto-Correlation Function of the Pseudo-Random Sequence	1-14
2-1	Propagation Mechanism in a Mobile Environment	2-8
2-2	Sample of Rayleigh Faded Envelope	2-8
2-3	Power Spectral Density of the Fading Process	2-8
2-4	Received C.W. Spectrum	2-11
2-5	Model for Error Production on Mobile Channels (On-Off Channel Model)	2-11
2-6	Intercept of the Random Process With the Level	2-11
2-7	The Probability $F(\tau, s)$ that $s(t) < s$ For An Interval Lasting Longer Than τ	2-24
2-8	The Fading Signal Envelope	2-24
2-9	Average Bit Error Rates (BER) for Different Test Parameters	2-26
2-10	Probability That A Burst Will Last For b Bits or More (Downtown)	2-30
2-11	Probability That A Burst Will Last For b Bits or More (Hull)	2-30
2-12	Probability That A Burst Will Last For b Bits or More (Queensway)	2-31
2-13	Average Burst Length Versus BER	2-31
2-13-1	Summary of Vehicle Speed Information	2-35
2-14	Probability That A Burst Will Last For b Bits or More: Comparison Between Theoretical and Experimental Distributions (Downtown)	2-35
2-15	Probability That A Burst Will Last For b Bits or More: Comparison Between Theoretical and Experimental Distributions (Hull)	2-36
2-16	Probability That A Burst Will Last For b Bits or More: Comparison Between Theoretical and Experimental Distributions (Queensway)	2-36
2-17	Data Block Affected By Successive Fades	2-38
2-18	Probability That The Number of Errors In A Block Will Exceed m Bits, Versus m/n. (Experimental) (Downtown)	2-46
2-19	Probability That The Number of Errors In A Block Will Exceed m Bits, Versus m/n. (Experimental) (Hull)	2-46

2-20	Probability That An Error Burst Will Last For b Bits or More, Versus b/n . (Downtown)	2-49
2-21	Probability That An Error Burst Will Last For b Bits or More, Versus b/n . (Downtown)	2-49
2-22	Probability That An Error Burst Will Last For b Bits or More, Versus b/n . (Hull)	2-50
2-23	Probability That An Error Burst Will Last For b Bits or More, Versus b/n . (Hull)	2-50
2-24	Probability That The Number of Errors In A Block Will Exceed m Bits, Versus m/n . (Theoretical) (Downtown)	2-51
2-25	Probability That The Number of Errors In A Block Will Exceed m Bits, Versus m/n . (Theoretical) (Hull)	2-51
3-1	Digital Land Mobile Communication System	3-2
3-2	The Functions $H_2(x)$ and $G_2(x)$ and Their Inverses	3-41
3-3	Asymptotic Forms of the Elias and the Gilbert-Varshamov Bounds	3-43
3-4	Standard Array (Decoding Table) For A Binary Code	3-51
3-6	Block Error Rate Versus Code Rate. (Experimental) Random Error Correction; Lower Bound (Downtown)	3-56
3-7	Block Error Rate Versus Code Rate. (Experimental) Random Error Correction; Upper Bound (Downtown)	3-56
3-8	Block Error Rate Versus Code Rate. (Experimental) Random Error Correction; Lower Bound (Hull)	3-57
3-9	Block Error Rate Versus Code Rate. (Experimental) Random Error Correction; Upper Bound (Hull)	3-57
3-10	Bit Error Rate Versus Code Rate. (Experimental) Random Error Correction; Lower Bound (Downtown)	3-58
3-11	Bit Error Rate Versus Code Rate. (Experimental) Random Error Correction; Upper Bound (Downtown)	3-58
3-12	Bit Error Rate Versus Code Rate. (Experimental) Random Error Correction; Lower Bound (Hull)	3-59
3-13	Bit Error Rate Versus Code Rate. (Experimental) Random Error Correction; Upper Bound (Hull)	3-59
3-14	Block Error Rate Versus Code Rate. Burst Error Correction (Downtown)	3-60
3-15	Block Error Rate Versus Code Rate. Burst Error Correction (Hull)	3-60

3-16	Block Error Rate Versus Code Rate. (Theoretical) Random Error Correction; Lower Bound (Downtown)	3-61
3-17	Block Error Rate Versus Code Rate. (Theoretical) Random Error Correction; Upper Bound (Downtown)	3-61
3-18	Block Error Rate Versus Code Rate. (Theoretical) Random Error Correction; Lower Bound (Hull)	3-62
3-19	Block Error Rate Versus Code Rate. (Theoretical) Random Error Correction; Upper Bound (Hull)	3-62
3-20	Bit Error Rate Versus Code Rate. (Theoretical) Random Error Correction; Lower Bound	3-63
3-21	Bit Error Rate Versus Code Rate. (Theoretical) Random Error Correction; Upper Bound (Downtown)	3-63
3-22	Bit Error Rate Versus Code Rate. (Theoretical) Random Error Correction; Lower Bound (Hull)	3-64
3-23	Bit Error Rate Versus Code Rate. (Theoretical) Random Error Correction; Upper Bound (Hull)	3-64
4-1	Conceptual Schematic of the AMPS Decoder	4-13
4-2	AMPS Simulation Results at 9.6 kbps	4-17
4-3	AMPS Simulation Results at 9.6 kbps	4-17
4-4	AMPS Simulation Results at 9.6 kbps	4-18
4-5	AMPS Simulation Results at 9.6 kbps	4-18
4-6	AMPS Simulation Results at 9.6 kbps	4-19
4-7	AMPS Simulation Results at 9.6 kbps	4-19
4-8	AMPS Simulation Results at 9.6 kbps	4-20
4-9	AMPS Simulation Results at 4.8 kbps	4-20
4-10	Time Varying Erasure and Error Channel	4-22
4-11	Modified On-Off Erasures and Errors Channel	4-22
4-12a	Mobile Transceiver Block Diagram	4-23
4-12b	Functional Block Diagram of Bell Laboratories-Designed Mobile Logic Unit	4-23
4-13	Erasure/Error Correction Algorithm	4-26
4-14	Alternate Decoding Scheme at 9.6 kbps	4-29
4-15	Alternate Decoding Scheme at 9.6 kbps	4-29
4-16	Alternate Decoding Scheme at 9.6 kbps	4-30
4-17	Alternate Decoding Scheme at 9.6 kbps	4-30
4-18	Influence of Interleaving Period on BER	4-33
4-19	Packet Miss Rate	4-33
4-20	Packet False Alarm Rate in Error-Detection Mode	4-33
4-21	Block Error Rate: Results and Theory	4-36
5-1	Bounds on the Capacity of the Binary DLMC As A Function of the BER	5-6
5-2	Best Achievable Zero-Error Rate With Bounded Distance Decoding of Block Codes	5-7.

LIST OF SYMBOLS

Greek Symbols

δ_{ij}	Kronecker delta function
$\delta(\cdot)$	Dirac delta function
$\dot{n}$	slope (time derivative) of the random signal level process
λ	carrier wavelength
μ	reciprocal of average fade duration, $\bar{\tau}(s)$
ν	level crossing rate
ρ	normalized signal level
$\bar{\tau}(s)$	average fade duration below signal level s
$\bar{\tau}_0(s)$	average interfade duration between fades below the level s
$\bar{\tau}_{IF}(s)$	average duration of the interval between successive fades below the level s
ω_0	radian carrier frequency

Roman Symbols

$A_1(n)$	number of vectors of length n and with distance i from an arbitrary reference
$A^{(n)}(z)$	generating function for the $A_1^{(n)} s$
$AB_1^{(n)}$	number of vectors of length n with a burst pattern of length i
B	blocklength measured in time units
BER	bit error rate
BLER	block error rate
BSC	binary symmetric channel

$b_0, b_1, \dots, b_n$	zeroth, first, ..., nth derivative of the autocorrelation function of the fading signal evaluated at zero
b/n	ratio of burst length to blocklength for a code
C_n	nth spectral component of signal process
$\mathbf{c}, \mathbf{c}^T$	a vector (usually a codeword) and its transpose
	a block code
$d_{\min}$	minimum distance of a code
d_{tot}	total distance between all pairs of codewords in a code
d/n	ratio of a code's minimum distance to its block-length
e	number of errors in a block (Chapter 2) number of erasures in a vector of length n (Chapter 3)
e/o	length of an erasure burst in a vector of length n
E	root mean square value of a signal
f_D	maximum Doppler frequency shift
f_0	carrier frequency
F	algebraic field
$GF(q)$	Galois field of q elements
H	parity check matrix for a linear code
$\mathbf{j}_m$	vector of occurrence of field symbols in position m of a code
k	total fade duration measured in bits (Chapter 2) number of information symbols in a code (Chapter 3)
L, l	fade duration
l	average fade duration
LCM	least common multiple
M	number of codewords in an arbitrary code

m	number of errors in a vector of length
$N(s)$	level crossing rate across level s
n	blocklength of a code
O	number of bits affected by a fade in a block of length B
P_m	probability vector of occurrence of field symbols in position m of a code
P_D	probability of detected block errors
P_{uD}	probability of undetected block errors
P_e	probability of error
$p(\geq b/n, n)$	probability of the ratio of the burst length to the blocklength exceeding b/n in a block of n bits
$p(\geq m/n, n)$	probability of the ratio of the number of errors to the number of bits in a block, exceeding m/n in a block of n bits
q	number of symbols in a field F (cardinality of F)
R	code rate
R_T	transmission rate
r	number of parity bits in a code
SINAD	ratio of signal and noise and distortion to noise and distortion
$S, s(t)$	signal level
$S(f)$	power spectral density
t	maximum number of correctable errors in a code
t/n	ratio of maximum number of correctable errors in a code to code length
$V_r(n)$	number of vectors of length n and with distance 1 or less from an arbitrary reference
$VB_b(n)$	number of vectors of length n with a burst pattern of length 1 or less
V	vehicle velocity

x distance travelled by mobile receiver
or transmitter

Other Symbols

$\binom{n}{m}$ binomial coefficient

(E_c, E_s) representation of a signal by its
in-phase and quadrature components

(R, θ) or (s, θ) polar representation of signal

$-$ or $E\{\cdot\}$ average value

$H_q(x), G_q(x), H_q^{-1}(x), G_q^{-1}(x) \dots$

the functions $x \log_q(q-1) - x \log_q(x) - (1-x) \log_q(1-x)$,
 $1 - x \log_q(q-1) + x \log_q(x) + (1-x) \log_q(1-x)$, and their
inverses

$[x]$ greatest integer contained in x

$|F|$ cardinality of F

$\|\bar{x}\|$ norm of $\bar{x}$

CHAPTER 1

INTRODUCTION

1.1 General

New mobile communication systems are being introduced in North America and elsewhere to meet increased demand for analog voice and digital data transmission. In 1983 it was estimated that the growth rate of such systems in North America had averaged twelve percent per year for the past twenty years [1]. Today, with the release of the 850 MHz frequency band by regulatory agencies in Canada and in the U.S. for cellular systems, growth is much more rapid.

In the earlier systems, analog voice communication was predominant and the demand for data transmission had been limited. With the introduction of the cellular system however, the demand for digital data transmission has increased. There are two reasons for the increase. On the one hand, the increased use of the computer and the consequent growth of computer communications has caused a large fraction of radio users to add data transmission capabilities to their systems [2]. On the other hand, the cellular system itself requires large amounts of data to be transmitted for channel assignment and control functions. Fully, up to 11% of the available channels in the cellular system are used for this purpose [3].

In the cellular system, methods are currently being investigated to offer the same data transmission services to the cellular radio users as those offered to normal business users. These include access to standard databases, remote data entry, and interactive computer communications with a remote host.

Other candidate areas for digital transmission are in transmission to portable terminals and in digital voice transmission. Digital voice transmission in particular is a likely contender for implementation in some radio-telephone systems.

Digital messages seldom contain enough natural redundancy to combat the effect of transmission errors. If part of the message is received in error, the meaning of the entire message is likely to change. Suitable protection must be applied to messages sent over noisy channels with sufficient redundancy to ensure satisfactory reception. Protection is usually achieved in one of two ways. In one scheme, usually referred to as automatic repeat request or ARQ, the transmitter adds just enough redundancy to the message to permit the receiver to detect the presence of transmission errors; if the receiver finds any errors, it requests retransmission of the message (or part thereof). This method is simple and can provide high system reliability, but it requires a feedback channel from the receiver to the transmitter. In the second scheme, an error correcting code is used for combating

transmission errors. Redundancy is added to the message to form code words or sequences that are easily distinguishable from each other. Because each sequence is highly unique, the receiver can usually locate and correct any errors that occur during transmission. This scheme is referred to as forward error correction (FEC). For a given level of performance, FEC systems are usually more complex than ARQ schemes, but FEC systems do not require a feedback channel and have constant throughput. Occasionally, both schemes are used together in what is called hybrid ARQ. The objective in a hybrid scheme is to use the FEC subsystem to reduce the frequency of retransmissions, while maintaining the high reliability associated with ARQ.

Although error correcting codes have been advocated for use in land mobile communication systems [4], little research investigating their effectiveness as a means of combating channel errors has been done. Moreover, the existing research seems to have been conducted in an ad-hoc fashion: an author picks one or more codes--more or less arbitrarily--and reports on the improvements achieved in the particular system under test. Not surprisingly, the conclusions reported vary widely; depending on the codes chosen, some authors report little or no improvement [5,6], while others maintain that significant improvement can be achieved at moderate cost [7].

The objective of this thesis is to assess the effectiveness of coding in reducing errors in digital land mobile systems. This thesis does not recommend the adoption of any particular coding scheme, even though some practical alternatives are investigated in Chapter 4. Instead, emphasis is placed on deriving equations that can be used to predict the effectiveness of coded systems operating on the mobile channel. Broadly stated, the relationships between code rate, code length, and desired level of performance--as measured by the decoded bit error rate (BER) or the decoded block error rate (BLER)--are detailed. Throughout, attention is given to block codes only.

1.2 Mobile Radio Channel Characteristics

The mobile radio propagation and reception mechanisms are quite unique. The signal energy at the receiver is the vectorial sum of many components arriving by way of multiple reflections, refractions, and scatterings caused by the topographical features of the terrain. As a consequence, the magnitude of the signal at each location in the mobile's path is highly variable. In some locations the resultant signal is weak whereas in others it is strong enough to saturate the input stage of the receiver. The relative motion between the receiver and the transmitter introduces a frequency shift in the incoming signal with respect to the transmitted one. This

gives rise to frequency and phase distortions. The signal is also affected by the disruptions that are common to all radio frequency communication systems, namely, interference and noise.

Figure 1-1 shows the principal components of a mobile communication system. The analog radio channel is subject to the manifold disruptions described earlier. As a result, many errors are observed on the digital channel between the channel coder/channel decoder pair. The error rates observed by the decoder are substantially higher than those seen in other communication systems, such as the telephone network and satellite systems.

In the investigation for the thesis, the starting point will be the digital channel created by the modulator/analog channel/demodulator chain shown in Figure 1-1. A full characterization of this digital superchannel will be given. To do so, it will be necessary to investigate the analog channel disruptions in detail. The effects of the disruptions, but also those of the modulator and demodulator pair will be embedded in the characterization. Although the characterization is useful by itself, its principal application in the thesis will be to obtain bounds on the performance of coded systems operating in mobile environments.

Because the digital superchannel is the primary object of investigation, the two important and related topics of modulator and demodulator design, as well as the

Figure 1-1 Mobile Radio Communication System

explicitly addressed in the thesis.

1.3 Preparatory Work: The Transmission Experiments

1.3.1 Overview

Experimental data, in the form of error pattern files, are used throughout the thesis to validate and substantiate the theoretical findings. The data used were available in raw form on magnetic tapes and the author and a colleague [8] undertook the task of transforming them into a format suitable for subsequent computer analysis. The data contained much relevant information about the mobile channel. The signal envelope, the modulated signal, the received data, and the recovered data clock were all recorded. In the process of transforming the data, a large data-base of channel information was acquired. This data-base is a valuable source of information that several researchers have already used in their investigations. It is believed that other researchers can benefit from this data. For this reason, the experiment and the data processing are summarized in this section and in the next. Additional and more detailed information about the data may be found in [8,2].

The experimental data used in the research was collected in a 1982/83 study of land mobile data transmission systems. The study was prepared for the federal Department of Communications (DOC) by members of

the Department of Systems and Computer Engineering at Carleton University [2]. The objectives of the experiment were to estimate the characteristics of the digital land mobile channel in the 150, 450, and 850 MHz frequency bands. To this end, a base-station was set up at Carleton University and a mobile unit was equipped with the apparatus necessary to receive, analyze and store the radio signal. The mobile followed six routes in the cities of Ottawa and Hull. The six routes were carefully chosen to reflect the different conditions likely to be encountered by a mobile receiver. The vehicle travelled in urban, suburban, and open terrain. Locations at the centre of the service area as well as near the fringe were chosen. Finally, the locations were chosen to allow for different average vehicle velocities, ranging from slow speeds with frequent stops, to highway speeds. To account for any non-reciprocal behaviour in the channel, transmissions were made in both directions: from base to mobile and from mobile to base. In the report, these transmission paths will be referred to as the forward and reverse channels respectively. Three modulation schemes were evaluated in terms of bit error rate, equipment complexity, compatibility with current systems, and spectral efficiency. All three schemes were of the FM type.

The field data was stored for subsequent processing as follows: at the receiver, the signal was passed

through a power meter and then demodulated. The demodulated signal, along with the received data, the received clock, a stable reference data clock, the signal envelope, a twenty second start/stop signal, and vehicle speed information, were recorded on magnetic tape using a 7-track analog FM tape recorder. Voice comments were stored on an eighth track to identify the parameters and conditions under which the tests were done. Figure 1-2 shows the transmitter and receiver configurations.

The base station equipment consisted of a commercial RF transmitter with thirty Watts of output power and a 100% duty cycle. A fifty ohm coaxial cable with a loss of 2 dB connected the transmitter to an 820 MHz antenna with a gain of 8 dB. The antenna was situated at twenty meters above ground level.

The receiver was a commercial FM transceiver with a noise bandwidth of 11 kHz and a sensitivity of 0.35 V for 12 dB SINAD. It was connected to a quarter wave whip antenna mounted on the roof of the mobile unit. The antenna was at 3 meters above ground level. The receiver and transmitter were synchronized at the start of each experiment by the use of two highly stable rubidium-vapour clocks.

For the research in this thesis, the tapes containing the experimental data for transmission at 4800 bits per second at 820.390 MHz were used. These tapes were chosen because they provide the most interesting data in terms of future mobile radio development. The modem used in these

(a) Transmitter

(b) Receiver

Figure 1-2 Experimental Set-up

tests is a commercial one, manufactured by Mobile Data International, which employs alternate mark inversion (AMI), and which connects directly to the FM modulator, bypassing the preemphasis-deemphasis section of the radio. The terrain types represented on these tapes are the forward urban, reverse urban, forward and reverse residential, and forward open highway types. Most tapes contain about six minutes of recorded data, but some contain nine or twelve minutes. The recordings alternate between twenty seconds of data transmission and twenty seconds of unmodulated carrier. In all, eighty one files—each representing twenty seconds of data transmission--were obtained from the tapes. This represents approximately twenty five minutes of real time data.

1.3.2 Data Processing

A characterization of the channel error process is required in order to evaluate the effect coding has on reducing error occurrences. The experimental data recorded in the 1983 field trial can be processed to yield channel error pattern files. These files are sufficient in characterizing the digital channel error process. An error pattern file is a bit by bit description of the digital channel behaviour where a:

1 corresponds to a bit in error, and a

0 corresponds to a bit received correctly.

The message used in the synchronous data transmission experiments at 4800 bps is a repeating 2047 bit maximum length shift register sequence. The errored bits can be isolated from the sequence by comparing it to a noiseless reference.

Our investigation began by sampling the received data signal recorded on the magnetic tapes and storing it in a format suitable for computer analysis. To this end, a programmable, microprocessor-based, hardware/software unit was built to interface the recorded signals to a microcomputer. Two main tasks are performed by the unit. The incoming pulses are reshaped, amplified, and detected as required, and the timing and protocol required to communicate to the microcomputer is implemented. Figure 1-3 illustrates the unit and the apparatus used to perform the conversion.

The tapes were played back at half speed (15 ips) and the reference clock signal--stored on one of the tracks--provided the timing required for sampling the data signal, which was sampled at $5T/8$, where T is the symbol period. Data for each twenty second interval was stored on computer disc. 96000 bits were stored in each file and the processing yielded 81 files for the transmissions performed at 820 MHz. Data for the other frequencies were also obtained, but those data were not used in the research for this thesis.

From these 81 files of raw unprocessed data, 81 error pattern files are obtained. To find the error

pattern, the received data is compared to a reference 2047 bit pseudo-random sequence: the reference is shifted until more than 75% of the bits in the data sequence and the reference sequence agree. An examination of the autocorrelation function (a.c.f.) of the pseudo-random sequence--depicted in Figure 1-4--shows that the function still has a peak at $\tau=0$ provided 511 or fewer errors occur in any block of 2047 bits. (An error decreases the number of agreements by one and increases the number of disagreements by one in the a.c.f.). If the error pattern for a 2047 bit block exceeds 511 erroneous bits by any amount, the chance of it mimicking the sequence is less than 2^{-2036} . Furthermore, the probability that such an error pattern will cause a false peak of magnitude strictly greater than one-half (incorrectly indicating that 75% or more of the bits are being received correctly) is less than 2^{-376} , which is also negligible. Lastly, if the synchronization procedure is applied successfully to a stream of 4094 bits, the chance of a false synchronization not being detected drops to $(2^{-376})^2$, and in general, if the procedure is successfully applied to n 2047-bit blocks, the probability drops to 2^{-376n} . These observations state that if one searches for peaks whose values are greater than 0.5 (≥ 0.5007), the chance of a false synchronization can be made extremely small regardless of the number of errors that occur.

Figure 1-3 Data Acquisition Equipment

Figure 1-4 Auto-correlation Function of the Pseudo-random Sequence

In cases where no peak exceeding 0.5 is found, the data block has to be discarded because there will be, in general, more than one peak of magnitude less than 0.5, and none of these peaks will be necessarily indicative of an actual synchronization between the reference and data blocks. These results can be derived from an application of the results of Chapters 3 and 4, by noting the close similarity that exists between pseudo-random sequences and cyclic maximum length shift register codes.

Once synchronized, the data file is decoded into the error pattern file.

This procedure fails when synchronization is lost. Occasionally, a bit in the data stream is sampled twice or not at all and the data is falsely decoded from that point onwards. To correct this, the data has to be resynchronized starting at the new position. This is easily done because the autocorrelation function is recomputed at every bit position.

In cases where the error rate remains high for extended periods, approaching 50% for the duration of one or more 2047-bit blocks, only a portion of the file can be decoded. The files are therefore of unequal length. To average the results, a weight w , proportional to the length of the file is associated to each file.

Problems due to high error rates, and insertion and deletion errors, occurred on the average in one out of every five files. There is a marked correlation between the location of the recording and the existence of the

problem. For instance, most of the files recorded downtown on the reverse channel exhibit these problems. Many other researchers using the data have reported the same findings [8,9]. It has been suggested that the source of these anomalies is the low duty cycle of the mobile transmitter.

The data processing yielded 81 error pattern files which are stored in non-control ASCII format on computer disc. The files are divided into:

- MDI01.01B to MDI01.09B: 9 files for transmission from base to mobile in urban terrain. 4 km distance to transmitter. 10 to 25 storey buildings (downtown Ottawa).
- MDI04.01B to MDI04.09B: 9 files for transmission from mobile to base in urban terrain. 4 km distance to transmitter. 10 to 25 storey buildings (downtown Ottawa).
- MDI05.01B to MDI05.13B: 13 files for transmission from mobile to base in an open residential area. 4 km distance to transmitter (Ottawa South).
- MDI09.01B to MDI09.09B: 9 files for transmission from base to mobile in urban terrain. 5 km distance to transmitter. Buildings in the vicinity of the mobile 5 to 15 storeys high (Hull).
- MDI10.01B to MDI10.09B: 9 files for transmission from mobile to base in urban terrain. 5 km distance to transmitter. Buildings in the vicinity of the

mobile 5 to 15 storeys high (Hull).

- MDI122.01B to MDI12.23B: 23 files for transmission from base to mobile on a divided highway. Rural terrain. 10 to 12 km distance to transmitter (Queensway).
- MDI13.01B to MDI13.09B: 9 files for transmission from base to mobile in an open residential area. 4 km distance to transmitter (Ottawa South).

1.4 Thesis Organization

The thesis is divided into five chapters.

To investigate the performance of coded systems, a reliable prediction of the channel error performance is required. This forms the subject of Chapter 2. The salient features of the digital mobile channel are discussed. Major sources of errors are identified. An analysis of the experimental data is then given. Emphasis is on determining the error distribution for blocks of various lengths.

Chapter 3 presents the required background in coding theory. In particular, bounds on the performance of block codes are presented. These bounds are then used to investigate the performance of coded systems over mobile channels.

In Chapter 4, the performance of some practical coding schemes is investigated. The results of this investigation are compared to the theoretical predictions of Chapter 3.

Chapter 5 is a summary of salient results. Information theoretic concepts are introduced to precisely outline the scope and limitations of the research. Finally, suggestions for further work are given.

CHAPTER 2

THE MOBILE CHANNEL

2.1 Chapter Summary

In this chapter the digital mobile communication channel is investigated. Sources of errors are identified and a model for the error producing mechanism is presented. The model is used to derive an expression for the distribution of the number of errors expected in a block of data transmitted over the channel.

The chapter is divided into six sections as follows: In a first section a qualitative description of the various error producing mechanisms in operation on the mobile radio channel is given. Fading is isolated as the main cause of errors. A second part describes fading more closely and presents a model of error production. The model requires the knowledge of fade duration statistics. These form the subject of the third section. Two possible fade duration distributions are conjectured in the third section and in the fourth, comparison is made with experimental data to determine the better fit. The distribution that produces the better fit is used in section five to derive a theoretical expression for the number of errors per block. In a last part, experiment and theory are compared.

2.2 Qualitative Description of Error Producing Mechanisms

In conventional communication systems, receiver front end noise, due to the thermal motion of electrons, causes occasional errors to be made. The frequency of these errors is inversely proportional to the signal power at the receiver, and as the ratio of signal energy to noise power density decreases the number of errors observed per unit time increases. The noise that originates at the receiver front end is sometimes called additive white Gaussian noise (AWGN) because it is broadband and because it corrupts the signal in an additive fashion. In mobile communication systems in contrast, the undesired random fluctuations of the received signal amplitude, which reduce the signal to levels at which front end noise becomes significant, the errors to occur. The direct cause in this case are the signal dropouts, even though the underlying error producing mechanism is the same.

The reduction in signal level can be due to different factors. If the range is too long, path loss becomes significant, and the signal weakens to a point where it is submerged in noise. If the receiver is in a shadow zone behind natural or man-made obstacles, the received signal can become very weak. Finally, the signal fades because of multipath: the transmitted signal energy arrives at the receiver through multiple transmission paths. The transmitted signal is reflected, scattered, and refracted

acute in cellular systems because a large number of channels are allocated in one frequency band, and frequencies are reused in cells in close geographical proximity. Ignition noise is a form of electromagnetic interference which originates in vehicle ignition systems. This form of interference is typically of short duration: only a few bits are corrupted, often only one. Finally, Random FM is caused by the interference between waves that arrive at the mobile from different angular directions and therefore with different Doppler shifts.

The effects of some of the abovementioned sources of errors can be minimized by proper system and equipment design. Satisfactory reception throughout the service area can be provided by accurate propagation fade margin apportionment, and careful system planning. Solutions to the problem of co-channel and adjacent channel interference range from proper system and equipment design, to power level emission control to responsible spectrum management. Similarly, ignition noise can be controlled or even suppressed by proper shielding and careful receiver design. Finally, experimental evidence suggests that random FM and AWGN are of secondary importance as error producing mechanisms when compared to signal fading.

Fading can thus be isolated as the most serious impairment in a properly designed mobile communication system. There are two components to fading. The first is

acute in cellular systems because a large number of channels are allocated in one frequency band, and frequencies are reused in cells in close geographical proximity. Ignition noise is a form of electromagnetic interference which originates in vehicle ignition systems. This form of interference is typically of short duration: only a few bits are corrupted, often only one. Finally, Random FM is caused by the interference between waves that arrive at the mobile from different angular directions and therefore with different Doppler shifts.

The effects of some of the abovementioned sources of errors can be minimized by proper system and equipment design. Satisfactory reception throughout the service area can be provided by accurate propagation fade margin apportionment, and careful system planning. Solutions to the problem of co-channel and adjacent channel interference range from proper system and equipment design, to power level emission control to responsible spectrum management. Similarly, ignition noise can be controlled or even suppressed by proper shielding and careful receiver design. Finally, experimental evidence suggests that random FM and AWGN are of secondary importance as error producing mechanisms when compared to signal fading.

Fading can thus be isolated as the most serious impairment in a properly designed mobile communication system. There are two components to fading. The first is

multipath fading, sometimes referred to by the generic and all-encompassing term of Rayleigh fading even though the signal may not be Rayleigh distributed [8], which gives rise to rapid fluctuations in the received signal envelope. The second is sometimes called shadowing or lognormal fading. It describes the variation in the signal about its mean level--as predicted by path-loss calculations--due to individual objects and general topography in the vicinity of the mobile. Both are often referred to under the general heading of Rayleigh fading.

The next section is devoted to an exposition of the relevant theory about the Rayleigh fading envelope. This theory is used to obtain fade duration and gap duration distributions. These will be required in our prediction of coded system performance.

2.3 Rayleigh Fading

The generally accepted model for mobile transmission is shown in Figure 2-1. If one assumes that the amplitude, phase, and angle of arrival of the waves received at the mobile are random, and that the received signal is the sum of many such waves, it can be shown that the p.d.f. of the resultant signal envelope over local geographic areas is approximately Rayleigh distributed:

$$f_s(s) = \frac{s}{b_0} \exp - \left\{ \frac{s^2}{2b_0} \right\} \quad (1)$$

where s is a linear signal strength (envelope), measured in volts for example, and b_0 is the local mean square

value of s . Furthermore, it is generally assumed that the mean square value, b_0 , is lognormally distributed, provided that the distance to the transmitter remains approximately constant, and that the terrain type remains similar. The process the signal goes through due to local shadowing is thus reflected in the mean value b_0 :

$$f_b(b_0) = \frac{1}{\sqrt{2\pi}\sigma} \exp - \left\{ \frac{(b_0 - \bar{m})^2}{2\sigma^2} \right\} \quad (2)$$

Here σ is a measure of the variability of the local environment, and $\bar{m}$ is a measure of the local mean as predicted by path loss, with compensations for terrain type, antenna heights, etc. Figure 2-2 is a typical plot of Rayleigh fading in a mobile environment. The variations in the local mean value b_0 are slow only when compared to the Rayleigh fades. b_0 is usually taken to be constant over a distance of 50 to 100 wave lengths. In absolute terms however, 5 dB changes in 50 meters of travel are typical at 820 MHz [3].

The mobile radio channel can also be described by the spectrum of the fading process. The power spectral density function (p.s.d.) of the Rayleigh fading--also known as Doppler spectrum--is given by: (at baseband)

$$S(f) = \begin{cases} \frac{E^2}{2\pi f_D} \left[1 - \left(\frac{f}{f_D} \right)^2 \right]^{-1/2} & ; |f| \leq f_D \\ 0 & ; \text{else} \end{cases} \quad (3)$$

and is shown in Figure 2-3. E is the root mean square (RMS) value of the signal. Many channels are characterized by Rayleigh fading, but if the angular distribution of the signals arriving at the mobile is uniform, the spectrum is as shown in Figure 2-3. The HF channel for example is also a Rayleigh fading channel, but it has a Gaussian Doppler spectrum.

If one were to transmit a constant amplitude tone at f Hz on a mobile channel, the spectrum of the received signal would not be a tone at f , but would be smeared as in Figure 2-4. The p.d.f. and the p.s.d. are complementary descriptions of a signal. The first describes the distribution of the amplitude levels, but gives no indication as to how the signal evolves in time, i.e., how fast the signal changes from one level to another. The second description details the time variations of the signal but is equally valid for many amplitude distribution functions. Statistics such as the level crossing rate (LCR), the fade duration distribution, and the interfade duration distribution, can be solved for only if both functions are known.

In the previous section it was concluded that fading was the most serious impairment for digital mobile communication systems. Our model for the digital error process will be based on this important observation. It will be assumed that errors occur when the signal fades

Figure 2-1 Propagation Mechanism in a Mobile Environment

Figure 2-2 Sample of Rayleigh Faded Envelope

Figure 2-3 Power Spectral Density of the Fading Process

below some threshold, in which case the channel is lost. This level can be viewed as an input sensitivity threshold, below which the receiver cannot operate. The thresholding assumption is not a bad one for FM systems which exhibit transitional behaviour. In our model the channel is a noiseless BSC when the signal is above the threshold, and a useless BSC when the signal is below it. This channel model has been used by others [11,12] and is referred to as the on-off channel model. The model is illustrated in Figure 2-5.

Because of the assumption that errors occur when the signal drops below the receiver input sensitivity, the error distributions can be obtained from fade duration statistics for fades below some level S , where S is equated with the input sensitivity threshold. In the next section these distributions are investigated more closely.

2.4 Level Crossing Rates, Average Fade Duration, and Fade and Gap Statistics for Rayleigh Fading Channels

2.4.1 Theory

The fade and gap duration statistics depend on the time variations of the signal as well as on the amplitude distributions of the received envelope. The time variations of interest are conveniently summarized in the Doppler spectrum of the signal.

The level crossing rate across a level S is defined to be the frequency with which the signal crosses this level in a given direction, either upwards or downwards. If a process $z = F(a_1, a_2, \dots, a_n; t)$ depends on several random variables $a_1, a_2, \dots, a_n$, and a parameter t (usually time), then the number of times the signal z crosses the level S in dt is given by Rice as [13,14]:

$$+dt \int_{-\infty}^{\infty} |n| p(S, n; t) dn \quad (4)$$

Here $n = \partial F(a_1, a_2, \dots, a_n; t) / \partial t$ is the derivative of $F(a_1, a_2, \dots, a_n; t)$ and represents the slope of the process z at time t , and $p(S, n; t)$ is the joint p.d.f. of $F(a_1, a_2, \dots, a_n; t)$ and its time derivative. The level crossing rate, $N(S)$, is then given by (if it exists):

$$N(S) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} -dt \int_{-\infty}^0 n p(S, n; t) dn \quad (5)$$

where $\int_{-\infty}^0 n p(S, n; t) dn$ is the integral over all negative slopes. The following argument shows that (5) represents the level crossing rate: If dt is small enough, the random curves observed between time t and time $t+dt$ appear as straight lines. If the process has value S' at t_1 , then referring to Figure 2-6, its intercept at the level S is given by the value $t = t_1 - (S' - S)/n$. If $t_1 \leq t_1 - (S' - S)/n \leq t_1 + dt$, then a level crossing has occurred in dt . Rewriting the inequality above we have $0 \leq -(S' - S)/n \leq dt$, and since n is negative (downwards crossings are being considered) we have $S \leq S' \leq -ndt + S$. Integrating over all

Figure 2-4 Received C.W. Spectrum

Figure 2-5 Model for Error Production on Mobile Channels (On-Off Channel Model)

Figure 2-6 Intercept of the Random Process with the Level d

negative slopes then gives:

$$\int_{-\infty}^{-ndt+S} p(S', n; t) dS' \quad (6)$$

and because dt is assumed small:

$$\int_{-\infty}^{-ndt+S} p(S', n; t) dS' \approx p(S, n; t) \times (-ndt)$$

Thus it can be seen that the number of crossings in one second across a level S with negative slope is given by $N(S)$ in (5).

The average fade duration is defined as the mean length of the interval for which the signal is below some level S . The probability that a signal with amplitude distribution $f(s)$ is below level S is $F(S)$, where $F(S) = p(x \leq S) = \int_{-\infty}^S f(x) dx$ is the cumulative distribution of the amplitude levels. $F(S)$ is also the normalized amount of time the signal spends below the level S . The average fade duration is given by the ratio of the time the signal spends below a level in a one second interval to the number of crossings across that level--either upwards or downwards--in a second. Thus the average fade duration across a level S is given by $\bar{T}(S) = F(S)/N(S)$, where $N(S)$ is as defined above.

We now turn our attentions to calculating these two results for the mobile channel. From Figure 2-4--which illustrates the characteristics of a continuous wave radio carrier at frequency f_0 that has passed through a mobile radio multipath fading medium--the process is

narrowband, centered around f_0 , and symmetric with respect to f_0 . The characteristics of its envelope and phase are not unlike those of thermal noise that is passed through a narrow-band filter having center frequency f_0 . In fact the signal is assumed to be made up of many component waves:

$$s(t) = \sum_{n=1}^N C_n \cos(\omega_n t - \phi_n) \quad (7)$$

Here $C_n = [2W(f_n)\Delta f]^{1/2}$ is the contribution of the components of the power spectrum in the range $[f_n - \Delta f/2, f_n + \Delta f/2]$, and ϕ_n are random phase angles assumed to be uniformly distributed between 0 and 2π . Because the power spectrum is narrow, a center frequency ω_0 may be chosen and (7) may be rewritten:

$$\begin{aligned} s(t) &= \sum_{n=1}^N C_n (\cos(\omega_n t - \phi_n + \omega_0 t - \omega_0 t)) \\ &= \sum_{n=1}^N C_n [\cos(\omega_n t - \phi_n - \omega_0 t) \cos(\omega_0 t) \\ &\quad - \sin(\omega_n t - \phi_n - \omega_0 t) \sin(\omega_0 t)] \quad (8) \end{aligned}$$

and since $\omega_n = \omega_0 + 2\pi f_D \cos(\theta_n) = \omega_0 + 2\pi f_D \cos(\theta_n)$, and f_D is the Doppler frequency due to vehicle motion, the following holds:

$$\begin{aligned} s(t) &= \sum_{n=1}^N \{ C_n [\cos(\omega_D \cos \theta_n t - \phi_n) \cos(\omega_0 t)] \\ &\quad - C_n [\sin(\omega_D \cos \theta_n t - \phi_n) \sin(\omega_0 t)] \} \end{aligned}$$

$$= E_C \cos(\omega_0 t) - E_S \sin(\omega_0 t) \quad (9)$$

Here

$$E_C = \sum_{n=1}^N C_n \cos(\omega_{nt} - \omega_0 t - \phi_n)$$

and

$$E_S = \sum_{n=1}^N C_n \sin(\omega_{nt} - \omega_0 t - \phi_n) \quad (10)$$

are quadrature components of the signal and $\omega_0 = 2\pi f_0$ is the midband frequency. Assuming N to be large enough and the ϕ_n 's to be uniformly distributed between 0 and 2π , it can be shown that E_C and E_S --because they are the sum of many independent random variables--are zero mean independent Gaussian random variables with standard deviation given by:

$$\begin{aligned} \sigma &= \lim_{\Delta f \rightarrow 0} \left[\sum_n C_n^2 \cos^2(\omega_{nt} - \phi_n) \right]^{1/2} \\ &= \lim_{\Delta f \rightarrow 0} \left[\sum_n (2\omega(f_n) \Delta f) \cos^2(\omega_{nt} - \phi_n) \right]^{1/2} \quad (11) \\ &= \lim_{\Delta f \rightarrow 0} \left[\sum_n \omega(f_n) \Delta f \right]^{1/2} = \left[\int \omega(f) df \right]^{1/2} = b_0^{1/2} \end{aligned}$$

For the level crossing rate the joint distribution of S and its time derivative S' must be known. $p(S, S')$ can be determined from the probability density of E_C, E_C', E_S, E_S' . We have:

$$E_C^2 = E_S^2 = b_0 \quad (12)$$

$a_n(t)/\delta t|_{t=t_1}$ with $i=1,2$. If $t_2 = t_1 + \tau$ then the probability of the curve z passing upwards in the interval $t_2, t_2 + dt_2$ given that it passed downwards through $t_1, t_1 + dt_1$ is the same as the probability $N(S) dt_1$ that the curve passes downwards in the interval $t_1, t_1 + dt_1$ times the probability of the interval lasting between τ and $\tau + dt_2$ seconds. Denoting this last probability by $p(\tau)$ we have:

$$N(S) dt_1 p(\tau) dt_2 = -dt_1 dt_2$$

$$\int_{-\infty}^0 \int_0^{\infty} n_1 n_2 p(S, n_1; t_1; S, n_2; t_2) dn_1 dn_2 \quad (38)$$

or

$$p(\tau) = \frac{\int_{-\infty}^0 \int_0^{\infty} n_1 n_2 p(S, n_1; t_1; S, n_2; t_2) dn_1 dn_2}{N(S)}$$

$$= \sqrt{2\pi} \circ f_D(e^{\tau^2} - 1) \int_{-\infty}^0 \int_0^{\infty} n_1 n_2 p(S, n_1; t_1; S, n_2; t_2) dn_1 dn_2 \quad (39)$$

The p.d.f. $p(E_c, E_s, E'_c, E'_s, E_c, E_s, E'_c, E'_s)$ of the eight random variables $E_c \dots E'_s$ can be determined-- since they are all Gaussian--from the p.s.d. . The resulting p.d.f. will be normal in eight dimensions. Using the relations $S_1 = (E_{c1}^2 + E_{s1}^2)^{1/2}$, $\theta_1 = \tan^{-1}(E_{s1}/E_{c1})$, $i = 1, 2$, an expression for the p.d.f. $p(S, S'_1, \theta, \theta'_1, S, S'_2, \theta_2, \theta'_2)$ can be obtained from which $p(S, S'_1, S, S'_2)$ may be obtained by integrating out $\theta_1, \theta'_1, \theta_2, \theta'_2$.

$$\begin{aligned}
\overline{E_c E'_s} &= \overline{E_s E'_c} = E\left\{ \sum_n C_n \cos((\omega_n - \omega_0)t - t_n) \sum_m C_m (\omega_m - \omega_0) \cos((\omega_m - \omega_0)t - t_m) \right\} \\
&= E\left\{ \sum_n C_n^2 (\omega_n - \omega_0) \frac{1}{2} \cos(2(\omega_n - \omega_0)t - 2t_n) \right\} \\
&= 0, \text{ because the spectrum is symmetric} \\
&\quad \text{with respect to the midband} \\
&\quad \text{frequency } f_0 \text{ and } (f - f_0) \text{ is an odd} \\
&\quad \text{function with respect to } f_0
\end{aligned} \tag{16}$$

$$\begin{aligned}
\overline{E'_s E'_c} &= E\left\{ \sum_n C_n (\omega_n - \omega_0) \sin((\omega_n - \omega_0)t - t_n) \times \right. \\
&\quad \left. \sum_m C_m (\omega_m - \omega_0) \cos((\omega_m - \omega_0)t - t_m) \right\} = 0
\end{aligned} \tag{17}$$

In addition, the two random variables,

$$E'_c = - \sum_1^N C_n (\omega_D \cos \theta_n) \sin(\omega_D \cos \theta_n t - t_n) \tag{18}$$

$$\text{and } E'_s = \sum_1^N C_n (\omega_D \cos \theta_n) \sin(\omega_D \cos \theta_n t - t_n) \tag{19}$$

are the sum of several independent random variables and can be shown to be Gaussian. The distribution $P(E_c, E'_c, E_s, E'_s)$ is therefore Gaussian in 4 dimensions. The matrix for the second moments of the four random variables taken in the order E_c, E_s, E'_c, E'_s is:

$$M = \begin{bmatrix} b_0 & 0 & 0 & 0 \\ 0 & b_0 & 0 & 0 \\ 0 & 0 & b_2 & 0 \\ 0 & 0 & 0 & b_2 \end{bmatrix} \tag{20}$$

And the four dimensional Gaussian distribution is given by:

$$P(E_c, E'_c, E_s, E'_s) = \frac{1}{(2\pi)^2} (b_0^2 b_2^2)^{-1/2}$$

$$\exp \left\{ -\frac{1}{2} (E_c, E_s, E'_c, E'_s) M^{-1} (E_c, E_s, E'_c, E'_s)^T \right\}$$

$$= \frac{1}{4\pi^2 b_0 b_2} \exp \left\{ -\frac{1}{2} \left(\frac{E_c^2}{b_0} + \frac{E_s^2}{b_0} + \frac{E'_c{}^2}{b_0} + \frac{E'_s{}^2}{b_2} \right) \right\}$$

(21)

With $S = (E_c^2 + E_s^2)^{1/2}$ and introducing an additional variable $\theta = \tan^{-1}(E_s/E_c)$, the random variables E_c, E_s, E'_c, E'_s are given by:

$$E_c = S \cos \theta - S'(\sin \theta) \theta'; \quad E_s = S \sin \theta + S'(\cos \theta) \theta'$$

(23)

$$\text{and } dE_c dE_s dE'_c dE'_s = S^2 dS dS' d\theta d\theta'$$

(24)

$$\text{Also, } \left\{ \begin{aligned} \frac{E_c^2}{b_0} + \frac{E_s^2}{b_0} &= \frac{S^2}{b_0} \end{aligned} \right.$$

(25)

$$\left\{ \begin{aligned} \frac{E'_c{}^2}{b_2} + \frac{E'_s{}^2}{b_2} &= \frac{S'^2 + S^2 \theta'^2}{b_2} \end{aligned} \right.$$

(26)

$$\text{Thus } p(S, S', \theta, \theta') = \frac{S^2}{4\pi^2 b_0 b_2} \exp \left\{ -\frac{1}{2} \left(\frac{S^2}{b_0} + \frac{S'^2 + S^2 \theta'^2}{b_2} \right) \right\}$$

(27)

$$\text{and } p(S, S') = \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} d\theta' p(S, S', \theta, \theta')$$

$$= \frac{S^2}{4\pi^2 b_0 b_2} \exp \left\{ -\frac{1}{2} \left(\frac{S^2}{b_2} + \frac{R^2}{b_0} \right) \right\} \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} \exp \left\{ -\frac{\theta'^2}{2 \left(\frac{b_2}{S^2} \right)} \right\} d\theta'$$

$$= \frac{1}{\sqrt{2\pi}} \frac{S}{b_0 \sqrt{b_2}} \exp \left\{ -\frac{1}{2} \left(\frac{S^2}{b_2} + \frac{S^2}{b_0} \right) \right\}$$

(28)

distributions:

$$\left\{ \begin{array}{l} F_z(u;S) = e^{-u/\bar{\tau}_c(S)} ; \bar{\tau}_c(S) = \sqrt{2\pi} f_D e^{-u^2} \\ F_y(u;S) = e^{-u/\bar{\tau}_{IF}(S)} ; \bar{\tau}_{IF}(S) = \frac{\bar{\tau}}{\sqrt{2\pi} f_D} \\ F_x(u;S) = \begin{cases} \frac{2\bar{\tau}(S)}{u} I_1 \left(\frac{u^2}{2\bar{\tau}(S)} \right) - \left(\frac{2}{u^2} \bar{\tau}(S)^2 \right) ; \text{ if } S \rightarrow 0 \\ e^{-u/\bar{\tau}(S)} ; \text{ if } S \rightarrow \infty \end{cases} \end{array} \right. \quad (47)$$

where $\bar{\tau}(S) = (e^S - 1) / (\sqrt{2\pi} f_D)$

The distributions are given as a function of the two variables u and S , to show the implicit dependence on the level S .

2.4.2 Experimental Results

In the previous section, formulas for fade duration distributions were obtained. In particular, two quite different laws were shown to apply for shallow and deep fades, and it was noted that actual fade duration distributions were expected to lie between these two extremes. In this section the main goal is to determine which distribution fits actual field data better. A candidate choice is proposed which will be used in the next section to predict the distribution of the number of errors per data block.

The data used in this section is described in Chapter 1. It consists of error pattern files recorded for different locations in the cities of Ottawa and Hull. This section is structured as follows; the experimental data is presented and some of its more important features

azimuths and with random phase angles. An examination of the two results reveals that the rate of occurrence of fades is proportional to the vehicle velocity whereas the mean duration varies inversely with velocity, similarly, the rate of occurrence of fades is proportional to the carrier frequency, and their duration is inversely proportional to the carrier frequency. The product is thus seen to be independent of vehicle speed or frequency, and to depend only on transmitted power for a given threshold level. The product $N(S)\bar{r}(S)$ is the portion of time the signal spends below the level S and is equal to $F(S)$ -- where $F(x)$ is the cumulative distribution of amplitude levels. In our model this will govern the probability of error:

$$P_e = \frac{1}{2} (N(S) \bar{r}(S)) = \frac{1}{2} (1 - e^{-\rho^2}) \quad (36)$$

Turning now to the fade and gap duration statistics, it can be shown, by arguments similar to those used to obtain the level crossing rate $N(S)$, that the probability of the process $z = F(a_1, a_2, \dots, a_n; t)$ passing through level S in time $t_1, t_1 + dt_1$ with negative slope and through the level S in time $t_2, t_2 + dt_2$ with positive slope is given by:

$$-dt_1 dt_2 \int_{-\infty}^0 \frac{dn_1}{dn_1} \int_0^{\infty} \frac{dn_2}{dn_2} n_1 n_2 p(S, n_1; t_1; S, n_2; t_2) \quad (37)$$

As before, $n_1 = \partial F(a_1, a_2, \dots, a_n; t) / \partial t |_{t=t_1}$, and $n_2 = \partial F(a_1, a_2, \dots, a_n; t) / \partial t |_{t=t_2}$, and $p(s_1, n_1; t_1; S_2, n_2; t_2)$ is the joint p.d.f. for the four random variables: $s_1 = F(a_1, a_2, \dots, a_n; t_1)$ and $n_1 = \partial F(a_1, a_2, \dots,$

$a_n(t)/\partial t|_{t=t_i}$ with $i=1,2$. If $t_2 = t_1 + \tau$ then the probability of the curve z passing upwards in the interval $t_2, t_2 + dt_2$ given that it passed downwards through $t_1, t_1 + dt_1$ is the same as the probability $N(S) dt_1$ that the curve passes downwards in the interval $t_1, t_1 + dt_1$ times the probability of the interval lasting between τ and $\tau + dt_2$ seconds. Denoting this last probability by $p(\tau)$ we have:

$$N(S) dt_1 p(\tau) dt_2 = -dt_1 dt_2$$

$$\int_{-\infty}^0 \int_0^{\infty} n_1 n_2 p(S, n_1; t_1; S, n_2; t_2) dn_1 dn_2 \quad (38)$$

or

$$p(\tau) = \frac{\int_{-\infty}^0 \int_0^{\infty} n_1 n_2 p(S, n_1; t_1; S, n_2; t_2) dn_1 dn_2}{N(S)}$$

$$= \sqrt{2\pi} \circ f_D(e^{\tau^2} - 1) \int_{-\infty}^0 \int_0^{\infty} n_1 n_2 p(S, n_1; t_1; S, n_2; t_2) dn_1 dn_2 \quad (39)$$

The p.d.f. $p(E_C, E_S, E'_C, E'_S, E_C, E_S, E'_C, E'_S)$ of the eight random variables $E_C \dots E'_S$ can be determined--since they are all Gaussian--from the p.s.d. The resulting p.d.f. will be normal in eight dimensions. Using the relations $S_1 = (E_{C1}^2 + E_{S1}^2)^{1/2}$, $\theta_1 = \tan^{-1}(E_{S1}/E_{C1})$, $i = 1, 2$, an expression for the p.d.f. $p(S, S'_1, \theta, \theta'_1, S, S'_2, \theta_2, \theta'_2)$ can be obtained from which $p(S, S'_1, S, S'_2)$ may be obtained by integrating out $\theta_1, \theta'_1, \theta_2, \theta'_2$.

The result can in general not be put in closed form, and thus no closed form solution exists for $p(\tau)$ in general. Rice and others [16,17] have obtained bounds on $p(\tau)$. In particular Rice has shown that for many spectral shapes, the possibility of a fade lasting τ seconds or more is independent of the actual spectrum and is given by:

$$\left\{ \begin{array}{ll} F_S(u) = e^{-u} & \text{when } S \rightarrow \infty \end{array} \right. \quad (40)$$

$$\left\{ \begin{array}{ll} F_S(u) = \left(\frac{2}{u}\right) I_1\left(\frac{2}{\tau u}\right) e^{-\frac{2}{\tau u}} & \text{when } S \rightarrow 0 \end{array} \right. \quad (41)$$

Here $u = \tau / \bar{\tau}(S)$, where $\bar{\tau}(S)$ is the average fade duration for level S . The fading spectrum of the mobile channel fits the description and thus we expect the distribution of deep fades (i.e., when $S \rightarrow 0$) to approach:

$$F_S(\tau) = \left(\frac{2 \bar{\tau}(S)}{\tau}\right) I_1\left(\frac{2}{\tau} \left(\frac{\bar{\tau}(S)}{\tau}\right)^2\right) \exp\left[-\frac{2}{\tau} \frac{\bar{\tau}(S)^2}{\tau^2}\right] \quad (42)$$

where $I_1(x)$ = modified Bessel function of the first kind of order one.

This result has been verified experimentally by [18] using a mobile Rayleigh fading simulator. Agreement was good for fades 10 dB or more below the median signal level.

More shallow fades are expected to have a distribution which resembles:

$$F_S(\tau) = e^{-\tau / \bar{\tau}(S)} \quad (43)$$

In Figure 2-7 the two distributions are compared. For values of between the two extremes $S \rightarrow 0$ and $S \rightarrow \infty$, the

curves obtained would fall between those shown in Figure 2-7.

We now turn our attention to the intervals between fades. It can be shown that the intervals between deep fades are exponentially distributed [16]. The exponential law is followed because most of the intervals are so long that the probability of an interval of length x ending in $x + dx$ becomes independent of x . Thus:

$$F_{IF}(x) = e^{-x/\bar{\tau}_{IF}(S)} \quad (44)$$

This result has been confirmed experimentally by [19] using a fading simulator. Agreement is good for interfade intervals as measured between fades 15 dB or more below the median signal level. In the equation $\bar{\tau}_{IF}(S)$ is the mean interfade duration between fades below level S .

Finally, it can be shown that the time between fades is also exponentially distributed, for the same reasons as above. Thus:

$$F_C(x) = e^{-x/\bar{\tau}_C(S)} \quad (45)$$

$\bar{\tau}_C(S)$ is the mean duration of the interval between the beginning of two successive fades below level S . Thus $\bar{\tau}_C(S) = 1/N(S)$. Also, $\bar{\tau}_C(S) = \bar{\tau}(S) + \bar{\tau}_{IF}(S)$, thus:

$$\bar{\tau}_{IF}(S) = \frac{1}{N(S)} - \bar{\tau}(S) = \frac{1}{(\sqrt{2\pi} f_D \rho)} \quad (46)$$

Referring to Figure 2-8 we thus have the following

distributions:

$$\left\{ \begin{array}{l} F_z(u;S) = e^{-u/\bar{\tau}_c(S)} ; \quad \bar{\tau}_c(S) = \sqrt{2\pi} f_D e^{-S^2} \\ F_y(u;S) = e^{-u/\bar{\tau}_{IF}(S)} ; \quad \bar{\tau}_{IF}(S) = \frac{f}{\sqrt{2\pi} f_D e^{-S^2}} \\ F_x(u;S) = \begin{cases} \frac{2\bar{\tau}(S)}{u} I_1 \left(\frac{2}{u} \bar{\tau}(S) \right) e^{-\left(\frac{2}{u} \bar{\tau}(S) \right)^2} ; & \text{if } S \neq 0 \\ e^{-u/\bar{\tau}(S)} ; & \text{if } S = 0 \end{cases} \end{array} \right. \quad (47)$$

where $\bar{\tau}(S) = (e^{-S^2} - 1) / (\sqrt{2\pi} f_D e^{-S^2})$

The distributions are given as a function of the two variables u and S , to show the implicit dependence on the level S .

2.4.2 Experimental Results

In the previous section, formulas for fade duration distributions were obtained. In particular, two quite different laws were shown to apply for shallow and deep fades, and it was noted that actual fade duration distributions were expected to lie between these two extremes. In this section the main goal is to determine which distribution fits actual field data better. A candidate choice is proposed which will be used in the next section to predict the distribution of the number of errors per data block.

The data used in this section is described in Chapter 1. It consists of error pattern files recorded for different locations in the cities of Ottawa and Hull. This section is structured as follows; the experimental data is presented and some of its more important features

Figure 2-7 The Probability $F(\tau, s)$ that $s(t) < s$ for an Interval Lasting Longer than τ .

Figure 2-8 The Fading Signal Envelope

are discussed. The burst distributions are then determined and compared to the theory as outlined above.

The bit error rate (BER) observed for each area is given in Figure 2-9. A distinction is made between transmissions that occur from base to mobile on the forward channel and transmissions from mobile to base on the reverse channel. Appendix I lists the BER for all the decoded files. It is seen that the probability of error stays between 10^{-1} and 10^{-3} in most cases. Very few files show a BER of 10^{-4} or better. There is considerable variation however in the BER for tests done in the same location and under similar conditions. This is attributed to variations in building patterns and to changes in traffic conditions around the mobile. From Figure 2-9 it is seen that in some locations, and particularly when transmission occurs from mobile to base the BER is very high: in the range of 10 to 15 percent. It is believed that this can be attributed to the low antenna height at the base station (30 m) and to the low duty-cycle of the mobile-mounted radio transmitters. These figures are believed to be unrepresentative of the typical mobile channel. For this reason and for others, these particular files have not been extensively analysed. Instead most of the analysis has centered around the transmissions that occurred from base to mobile in the downtown and Hull areas.

In addition to the BER, Appendix I gives the average burst length and variance, for each file. A burst--as

<u>PARAMETERS</u>		<u>BER</u>	<u>BER</u>
LOCATION	: DOWNTOWN		
TERRAIN TYPE	: URBAN	2.3×10^{-3}	2.8×10^{-3}
TRANSMISSION	: FROM BASE TO MOBILE		
LOCATION	: DOWNTOWN		
TERRAIN TYPE	: URBAN	1.6×10^{-1}	0.6×10^{-1}
TRANSMISSION	: FROM MOBILE TO BASE		
LOCATION	: HULL		
TERRAIN TYPE	: URBAN	7.2×10^{-3}	3.5×10^{-3}
TRANSMISSION	: FROM BASE TO MOBILE		
LOCATION	: HULL		
TERRAIN TYPE	: URBAN	4.3×10^{-2}	5.5×10^{-2}
TRANSMISSION	: FROM MOBILE TO BASE		
LOCATION	: QUEENSWAY		
TERRAIN TYPE	: HIGHWAY	6.2×10^{-2}	6.2×10^{-2}
TRANSMISSION	: FROM BASE TO MOBILE		
LOCATION	: WEST END		
TERRAIN TYPE	: URBAN/SUBURBAN	7.3×10^{-2}	3.3×10^{-1}
TRANSMISSION	: FORM MOBILE TO BASE		
LOCATION	: SOUTH-EAST		
TERRAIN TYPE	: SUBURBAN	1.3×10^{-1}	0.5×10^{-1}
TRANSMISSION	: FROM BASE TO MOBILE		

Figure 2-9

Average Bit Error Rates (BER)
for Different Test Parameters.

defined here--starts with an error, ends with an error, and does not contain a run of more than n error free bits between successive errors. A careful selection of the parameter n must be made to ensure that on the one hand, long fade is not counted as several short fades (because n is too small), and that on the other, several independent fades are not regrouped into one (because n is too large). It is assumed that bursts of errors occur when the fading reduces the signal to the level at which errors are caused by front-end noise in the receiver. Because independent fades occur approximately one-half wavelength apart, this sets an upper limit on n . Assuming a maximum vehicle velocity of 25 m/s (90 km/hr), and a transmission rate of 4800 bps, a value of about 40 bits is obtained for the gap between fades one-half wavelength apart. It cannot be insured that independent fades are counted separately if a value for n greater than half this number is chosen. A lower limit of 7 is used because error free runs are not likely to exceed 7 bits very often if the signal is below threshold. Values used and reported in the literature range from 5 to 50.

In Figures 2-10 to 2-12 the probability of a burst being greater than or equal to the value of the abscissa is plotted against the abscissa for various n . In Figure 2-10, which corresponds to transmission in the downtown area at a high signal to noise ratio, the burst distribution is not sensitive to the actual value of n . The situation is quite different in Hull (Figure 2-11)

and on the Queensway, (Figure 2-12) where the signal to noise ratios are not as high. For the Queensway, in particular, it would appear that the choice $n=20$ is inappropriate. The choice of the value of n has been restricted to some value between $n=7$ and $n<20$. The graphs however, do not provide any clues beyond that. An examination of the error pattern files reveals that successive ones, which occur when the signal is faded, always occur in groups of 5 or less. Because of our assumption that the channel is a useless BSC when it is faded, this suggests that successive zeros will also occur in groups of 5 or less. Thus a choice of $n=7$ appears to be quite satisfactory.

From Appendix I it is observed that the average burst length is between 8 and 15 bits most of the time. It is only when the error rate is excessively high that the burst length increases significantly above 15 bits. In Figure 2-13, the average burst length is plotted as a function of the BER. It is seen to increase only very slightly as the BER varies from 10^{-3} to 10^{-1} . Within limits, the average burst length is not critically dependent on the bit error rate. As the BER increases (SNR decreases) the mean inter-burst gap length decreases, but the mean burst length remains approximately constant at around 10 to 15 bits. The value of this observation lies in the insight it provides about choosing a code. It suggests for example, that Reed-Solomon codes with symbols chosen from $GF(2^4)$ or $GF(2^5)$ (16 bit and 32 bit symbols respectively) might adequately protect the data from errors regardless of the conditions in which the vehicle is operating.

In order to compare the experimental distributions obtained in this section with the theoretical ones derived earlier, the latter must be re-expressed as a function of the number of bits affected rather than the fade duration in seconds. It is also more convenient to express the distributions as a function of the vehicle speed and the local bit error rate rather than the Doppler frequency and the threshold to rms signal level.

In our model errors occur with probability 0.5 when the signal fades below a threshold level S . Since fades are usually longer than a bit interval, errors occur in bursts. If a fade lasts for τ seconds it will affect $k = \tau/R_T$ bits, where R_T is the transmission bit rate. The variable τ can thus be replaced by k/R_T . Similarly, $\bar{\tau}(S)$, the average fade duration, is given by the formula:

$$\bar{\tau}(S) = \frac{e^{\rho^2} - 1}{\sqrt{2\pi} f_D \rho} ; \text{ where } \rho = \frac{S}{S_{\text{rms}}} \quad (48)$$

In the model it is assumed that the probability of error and the threshold to rms value are related by the expression $p_e = 1/2(1 - e^{-\rho^2})$ or $\rho = \sqrt{-\ln(1 - 2p_e)}$. Upon substitution in (48) the formula reads:

$$\bar{\tau}(S) = \frac{2p_e}{(1 - 2p_e) \sqrt{-\ln(1 - 2p_e)} f_D \sqrt{2\pi}} \quad (49)$$

Finally, $f_D = v/\lambda$, where v is the vehicle velocity and λ is the carrier wave length. Substituting for the above in the two distributions yields:

Figure 2-10 Probability that a Burst will last for b bits, or more.
(Downtown)

Figure 2-11 Probability that a Burst will last for b bits or more.
(Hull)

Figure 2-12 Probability that a Burst will last for b bits or more. (Queensway)

Figure 2-13 Average Burst Length versus BER

$$F(k, S \rightarrow \infty) = \exp \left\{ - \frac{(1-2p_e) \sqrt{-\ln(1-2p_e)} \sqrt{2\pi}}{2p_e \lambda} \left(\frac{v}{R_f} \right) k \right\} \quad (50)$$

$$F(k, S \rightarrow 0) = \frac{4p_e \lambda R_f}{(1-2p_e) \sqrt{-\ln(1-2p_e)} \sqrt{2\pi} v k} \\ \times I_1 \left(\frac{8p_e^2 \lambda^2 R_f^2}{2\pi^2 (1-2p_e)^2 (-\ln(1-2p_e)) v^2 k^2} \right) \\ \times \exp \left\{ - \left(\frac{8p_e^2 \lambda^2 R_f^2}{2\pi^2 (1-2p_e)^2 (-\ln(1-2p_e)) v^2 k^2} \right) \right\} \quad (51)$$

The above expressions give the probability of a burst lasting longer than k bits as a function of the bit rate R_f , the vehicle velocity v , the probability of error p_e , and the carrier wave length λ .

In actual situations the vehicle velocity v is likely to vary considerably over local geographic areas. If the vehicle speed consists of several component speeds v_i , the equations can be replaced by a summation over the v_i 's with each term weighed by the distance travelled at each speed, X_i . Thus,

$$F(k, S \rightarrow \infty) = \sum_i \frac{X_i}{X} \exp \left\{ - \frac{(1-2p_e) \sqrt{-\ln(1-2p_e)} \sqrt{2\pi}}{2p_e \lambda} \left(\frac{v_i}{R_f} \right) k \right\} \quad (52)$$

where X is the total distance travelled over the period of observation. A similar result hold for the other distribution:

$$F(k, S \rightarrow Q) = \sum_i \frac{X_i}{X} \left(\frac{4p_e \lambda R_f}{(1-2p_e) \sqrt{-\ln(1-2p_e)} \sqrt{2\pi} v_i k} \right)$$

$$x \cdot I_1 \left(\frac{8p_e^2 v^2 R_I^2}{2\pi^2 (1-2p_e)^2 (-\ln(1-2p_e)) v_i^2 k^2} \right)$$

$$x \cdot \exp \left(- \frac{8p_e^2 v^2 R_I^2}{2\pi^2 (1-2p_e)^2 (-\ln(1-2p_e)) v_i^2 k^2} \right)$$

(53)

To compare the distributions with the experimental results, the vehicle velocity v is required. All the other parameters are known or can be calculated. A colleague has analysed the fading envelope of the data. His work is summarized in [8]. One of the tasks he undertook in his investigation was to obtain the vehicle velocity for each file. In particular, he developed a program that reads the speed pulses from the tapes and converts the information acquired into vehicle velocity. His programs and his report were used to obtain information about the vehicle velocity in each area. Our findings are summarized in Figure 2-13-1.

In Figures 2-14 to 2-16 the experimental distributions are compared to the theoretical ones derived above. The dashed portions in Figures 2-14 to 2-16 represent the ranges where there are an insufficient number of bursts to expect reliable averaging. Experiment and theory are in good agreement. The data is seen to lie between the two curves as predicted. Furthermore, the exponential distribution provides a better fit over the range where there is sufficient data.

It is not known however whether the distribution of the longer bursts is exponential or whether it follows the law $F(k, S \rightarrow 0)$. A larger volume of data would be required to establish this point. These two distributions become increasingly dissimilar as the burst length increases. The distribution $F(k, S \rightarrow 0)$ decays as k^{-3} whereas the distribution $F(k, S \rightarrow \infty)$ decays exponentially. In subsequent work the exponential distribution will be used as a model for the field data burst distribution because it provides a better fit in the range of interest.

2.5 The Distribution of the Number of Errors Per Block

A long block of data is likely to be corrupted by several fades. This situation is shown in Figure 2-17. Our aim in this section is to find the distribution of the number of errors in a block when it has been corrupted as shown in Figure 2-17. This distribution will then be used in an investigation of the performance of coded systems.

Assume that the channel is off for a length of L (measured in bits), then the distribution of the number of errors is given by the binomial function:

$$P(e=e_i) = \begin{cases} \binom{L}{e_i} \left(\frac{1}{2}\right)^{e_i} \left(\frac{1}{2}\right)^{L-e_i} & ; 0 \leq e_i \leq L \\ 0 & ; e_i > L \end{cases} \quad (54)$$

Here we have assumed that the channel acts as a useless BSC when it is off (i.e., $p(0|1) = p(1|0) = 0.5$).

	comments	max speed	avg speed
Urban Ottawa	Vehicle infrequently stopped	7.2 m/s C	6.0 m/s
Urban Hull	Vehicle frequently stopped	6.5 m/s	4.2 m/s
Queensway	Markedly different speeds	10 m/s on ramp	8.5 m/s on ramp $X_1/X_T = 0.076$
	Ramp 16.7% of the time Highway, 83.3% of the time	22 m/s on highway	20.8 m/s on highway $X_1/X_T = 0.924$

Figure 2-13-1 Summary of Vehicle speed information

Figure 2-14 Probability that a Burst will last for b bits or more: Comparison Between Theoretical and Experimental Distributions. (Downtown)

Figure 2-15 Probability that a Burst will last for b bits or more: Comparison Between Theoretical and Experimental Distributions. (Hull)

Figure 2-16 Probability that a Burst will last for b bits or more: Comparison Between Theoretical and Experimental Distributions. (Queensway)

For a given distribution of the length (in bits) during which the channel is off, the error distribution will be given by: (by the theorem of total probability)

$$P(e=e_1) = \sum_j p(e=e_1 | l=l_j) p(l=l_j) \quad (55)$$

where $p(e=e_1 | l=l_j)$ is as defined in (54) above.

The next step is to obtain the distribution of the length of the interval during which the channel is off for a given blocklength B . Because we have the distribution function for fades and we know that the time between fades is exponentially distributed, we can--if we assume that successive fades are independent, and that interfade intervals are also independent (over small geographic areas)--obtain the distribution function for the length of the received signal being below a level S assuming that the signal is observed for a total length B . The assumption that successive interfade intervals are independent over small geographic areas is reasonable because in such cases the mean value remains approximately constant. Because the interfade intervals are exponentially distributed and the mean value is constant, the intervals are indeed independent. Because of these assumptions, the point process defined as the time t_i at which a fade starts is a Poisson process and the number of points in an interval of length B is given by:

$$P\{k \text{ points in } B\} = e^{-\nu B} \frac{(\nu B)^k}{k!} \quad (56)$$

Figure 2-17 Data Block Affected by Successive Fades

where $\nu = N(S)$.

A second assumption is required, namely that fades do not overlap. If there are k fades in an interval of length B and if each fade has duration l_i (l_i ; $i=1,2,\dots,k$) then the total length of time that the signal spends below the level S is assumed to be given by the sum of the component l_i 's: $L = \sum l_i$. This assumption is reasonable since $\bar{\tau}(S) \ll \bar{\tau}_0(S) = 1/N(S)$ for deep fades. Thus for each choice of k ($k = 0, 1, \dots, n; n \rightarrow \infty$) the random variable is a sum of independent and identically distributed random variables. Assuming that the distribution of the random variable has been obtained for each choice of k : i.e., we have $p(L \leq L_j | k=k_1)$, then the probability distribution for the interval is obtained from:

$$P\{L \leq L_j\} = \sum_{k=0}^{\infty} p(L \leq L_j | k=k_1) p(k=k_1) \quad (57)$$

Our data suggests that the fade duration distribution follows the law $F(u, S \rightarrow \infty)$. We will calculate $p(L < L_j | k=k_1)$ for this law.

The probability distribution corresponding to $F(u, S \rightarrow \infty)$ is:

$$f(x) = \mu \exp \{-\mu x\} \quad (58)$$

with $\mu = \bar{\tau}(S)^{-1}$. By taking a Laplace transform:

$$F(s) = \mu / (\mu - s) \quad (59)$$

The interval L_j is the sum of k_1 intervals each of which has the distribution in (58). The sum of k_1 independent random variables is a k -fold convolution in

the initial domain or a k-fold multiplication in the transform domain:

$$[L/(L-s)]^k \quad (60)$$

Upon taking an inverse Laplace transform:

$$p(x|k) = \mu^k \frac{x^{k-1} e^{-\mu x}}{(k-1)!} ; \text{ if } k \neq 0 \quad (61)$$

Thus $p(L|k_1 \text{ fades}) p(k_1 \text{ fades in } B) =$

$$\begin{aligned} & e^{-\nu B} \frac{(\nu B)^{k_1}}{k_1!} \mu \frac{x^{k_1-1} e^{-\mu x}}{(k_1-1)!} \quad \text{if } k \neq 0 \\ & e^{\nu B} \delta(L) \quad \text{if } k=0 \end{aligned}$$

Finally,

$$\begin{aligned} p(x) &= \sum_{k=1}^{\infty} p(x|k) p(k) \quad \text{if } k \neq 0 \\ &= \sum_{k=1}^{\infty} \left(e^{-(\nu B + \mu x)} \frac{(\nu \mu B)^k x^{k-1}}{k! (k-1)!} \right) \quad (62) \end{aligned}$$

There is also a contribution at $L=0$ which is given by:

$e^{-\nu B} \delta(L)$. Thus the density is given by:

$$f_x(x) = \frac{e^{-(\nu B + \mu x)} I_1(2\sqrt{\mu \nu x B}) \sqrt{\mu \nu B}}{\sqrt{x}} + e^{-\nu B} \delta(x) \quad (63)$$

where $I_1(x)$ = modified Bessel Function of the first kind of order one. This is indeed a probability density function since:

$$\begin{aligned} \int_0^{\infty} f_x(x) dx &= \sum_{k=1}^{\infty} \frac{e^{-\nu B}}{(k-1)! k!} \nu^k \frac{\Gamma(k)}{\mu^k} (\nu B)^k + e^{-\nu B} \\ &= \sum_{k=0}^{\infty} \frac{(\nu B)^k}{k!} e^{-\nu B} = 1 \quad (64) \end{aligned}$$

To obtain the error distribution, the total fade duration distribution as given by (63) is substituted in (54). Since (54) is discrete but the fade duration is expressed as a function of the continuous variable ℓ , (54) may be evaluated in one of two ways. In a first approach the fade duration can be integrated over the intervals $[\ell_i - \Delta\ell/2, \ell_i + \Delta\ell/2]$, where $\Delta\ell$ is taken as the bit duration, and the result can be substituted in (54). Alternatively, a Gaussian variable with mean $= \ell/2$ and variance $= \ell/4$ can be substituted for the binomial variable. This is an approximation, but a good one. Both approaches however lead to complicated expressions that we prefer to omit. Instead the following simplification is used. With $p=q=0.5$ we know that the expected number of errors per block is $\langle e \rangle = \ell/2$. If we assume that the number of errors e , is a deterministic function of ℓ such that $e = \ell/2$ we can now substitute for e in the above expression. This gives:

$$f_e(\alpha) = 2f_x(2\alpha) \quad (65)$$

The mean value of the length of the faded interval is given by:

$$E(x) = E(\bar{\ell}N(S)B) = \bar{\ell}N(S)B \quad (66)$$

and the variance is given by:

$$\begin{aligned} E(x^2) &= \int_{-\infty}^{\infty} \sum x^2 p(x|k) p(k) dx \\ &= \sum \int_{-\infty}^{\infty} x^2 p(x|k) dx p(k) \\ &= \sum \bar{\ell}^2 kp(k) \\ &= \bar{\ell}^2 N(S)B \end{aligned} \quad (67)$$

where ℓ is the average fade duration, B is the blocklength, and $N(S)$ is the level crossing rate across level S .

Formula (65) will not be accurate for short block lengths. In particular, it is believed that the p.d.f. for short block lengths will have a larger spread than what is indicated by the formula. However, it can be shown that the approximation improves as the block length increases. This involves two steps. As a first step, the Chernoff bound is used to show that as the length ℓ is increased, the binomial distribution with $p=q=0.5$ becomes narrower and more concentrated around $\ell/2$. As a second step, it is shown that as the blocklength is increased, the short intervals become increasingly infrequent, so that where the approximation of the binomial by an impulse at $\ell/2$ is not valid (i.e., for short intervals) the intervals themselves have a negligible probability of occurrence.

With the Chernoff bound we show that:

$$\begin{aligned} P\left[\sum_{i=0}^{\ell} y_i > (1/2 + u)\right] &= P\left\{\sum_{i=0}^{\ell} y_i > \frac{\ell + \epsilon}{2}\right\} \\ &\leq \left[\left(\frac{1/2 + \epsilon}{1/2}\right)^{1/2} \left(\frac{1/2 - \epsilon}{1/2}\right)^{1/2}\right]^{\ell} \\ &= \left[\left(\frac{1/4 - \epsilon^2}{1/4}\right)\right]^{\ell/2} \end{aligned}$$

which tends to zero as ℓ tends to infinity. Here y is an indicator random variable which takes on the values 1 and 0 according to whether an error occurs or not. u , and consequently ϵ , are arbitrarily small positive constants.

Similarly:

$$P\left[\sum_{i=0}^{\ell} y_i < 1/2 - \epsilon\right] \leq \left[\frac{(1/4 - \epsilon)^2}{1/4}\right]^{\ell/2}$$

which also tends to zero as ℓ tends to infinity.

Because of the difficulty of applying the Chernoff bound to the second problem, the weaker law of large numbers is used to show that the probability of the random variable e (representing the number of errors), differing from its mean by a factor of k is bounded by:

$$P\left\{\left|e - \frac{\bar{\ell} N(S) B}{2}\right| \geq \frac{\bar{\ell} N(S) B}{2k}\right\} \leq \frac{\bar{\ell}^2 B^2 N(S)^2}{4k^2} = \frac{k^2}{\bar{\ell} B} \quad (69)$$

and decreases as B is increased. Since the mean value $\bar{\ell} N(S) B / 2$, increases in proportion to B , equation (69) shows that as B is increased short errored intervals become increasingly less probable.

The density function for the total length of the intervals for which the signal is below the threshold level was given by equation (63):

$$f_x(\ell) = \frac{e^{-(\nu B + \mu \ell)} I_1(2\sqrt{\nu \mu B \ell}) \sqrt{\nu \mu B}}{\sqrt{\ell}} + e^{-\nu B} \delta(\ell) \quad (70)$$

where $\nu = N(S)$

$\mu = 1/\bar{r}(S)$

B = blocklength in seconds

ℓ = total duration below threshold level

$$\begin{aligned} \text{But, } \nu &= e^{-\rho} f_D \sqrt{2\pi} = \sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} \frac{\nu}{\lambda} \\ \mu &= (e^{\rho} - 1) / \rho f_D \sqrt{2\pi} \\ &= \sqrt{2\pi} \frac{(1-2p_e)}{2p_e} (\sqrt{-\ln(1-2p_e)}) \cdot \frac{\nu}{\lambda} \end{aligned}$$

$B = n/R_T$; where n is the code length and R_T is the transmission rate

$l = O/R_T$; where O is the total number of bits affected by fades.

The density can therefore be re-expressed as:

$$\begin{aligned}
 f_o(0) = & \exp\left\{-\frac{\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} v}{\lambda R_T} \left[n + \frac{O}{2p_e}\right]\right\} \\
 & \times I_1\left(\frac{(2\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} v)}{\lambda R_T} \left[\sqrt{\frac{nO}{2p_e}}\right]\right) \\
 & \times \left[\frac{\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} v}{\lambda \sqrt{\frac{n}{20p_e}}}\right]^{-1} \\
 & + \delta(0) \exp\left\{-\frac{\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} v}{\lambda R_T} vn\right\} \quad (71)
 \end{aligned}$$

and since the number of errors m , is equal to half the bits affected O :

$$\begin{aligned}
 f_m(m) = 2f_o(2m) = & 2 \exp\left\{-\frac{\sqrt{2\pi} (1-2p_e) (\sqrt{-\ln(1-2p_e)}) v}{\lambda R_T} \left[n + \frac{m}{p_e}\right]\right\} \\
 & \times I_1\left(\frac{(2\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} v)}{\lambda R_T} \sqrt{\frac{nm}{p_e}}\right) \\
 & \times \left(\frac{\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} v}{\lambda \sqrt{\frac{n}{4mp_e}}}\right)^{-1} \\
 & + \exp\left\{-\frac{\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} v}{\lambda R_T} vn\right\} \delta(m) \quad (72)
 \end{aligned}$$

Finally, the cumulative probability is given by:

$$F_m(m) = 1 - \sum_{k=1}^{\infty} \sum_{j=0}^{k-1} \exp\left\{-\frac{\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} v}{\lambda R_T} \left(n + \frac{m}{p_e}\right)\right\}$$

$$x \left[\frac{\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} \sqrt{n}}{R_I} \right]^k / k!$$

$$x \left[\frac{2\sqrt{2\pi} (1-2p_e) \sqrt{-\ln(1-2p_e)} \sqrt{m}}{2p_e R_I} \right]^j / j!$$

(73)

which may be evaluated numerically.

2.6 Comparison of Theory and Experiment

For the channel, the probabilities of finding m errors in a block of length n , $p(m, n)$, can be computed from the error pattern files. For most of our work however, it is not the absolute count of the number of bits in error that is of interest but rather the ratio of incorrect bits to total bits per block. This distribution will be referred to as $p(=m/n, n)$. Furthermore, the probability $p(\geq m/n, n)$ of the ratio exceeding a certain value will be more useful than simply $p(=m/n, n)$. In Figures 2-18 and 2-19 the probability of the fraction of erroneous bits exceeding m/n is plotted as a function of m/n , with n as a parameter. These curves extend from $1/n$ to 1. The intersection with the ordinate, $p(\geq 1/n, n)$, is the probability that the block will be received in error, and hence $1-p(\geq 1/n, n)$ is the probability of receiving a block error free. The values chosen for n lie in the range from 15 to 4095 bits and are of the form $n=2^k-1$, with k an integer. These

Figure 2-18 Probability that the Number of Errors in a Block will exceed m bits, versus $\frac{m}{n}$. (Experimental) (Downtown)

Figure 2-19 Probability that the Number of Errors in a Block will exceed m bits, versus $\frac{m}{n}$. (Experimental) (Hull)

particular values were chosen because many good block codes have a length given by $n=2^k-1$. The distributions become progressively steeper as the blocklength increases, asymptotically tending to a step function. The probability of receiving a block in error also increases as the blocklength increases and is almost 1 when $n=4095$.

It is also useful to know the distribution of error burst lengths within a block. These distributions are useful when considering the use of burst-trapping decoders in conjunction with burst correcting codes. These are some of the simplest decoders to instrument. The definition of an error burst is not the same as the definition of a burst given earlier. An error burst starts with an error, ends with an error, and contains all the erroneous bits in the block. This in contrast with a burst, which contains only the errored bits that are not separated by the guard space or more. An error burst can therefore contain many individual bursts.

Because burst trapping decoders are cyclic, the data block can be viewed as if it were wrapped around a cylinder. Many values for the length of the error burst are obtained as the cylinder is rotated. The burst error of interest will be the shortest one obtained by this process.

The probability $p(\geq b/n, n)$ of the ratio b/n exceeding a certain value is plotted against b/n , with n as a parameter in Figures 2-20 to 2-23. b is the length of

the shortest error burst observed when the data block is rotated as described previously. From the figures, it is observed that the distributions remain fairly constant as the blocklength is increased, suggesting that the burst length increases in proportion to the block length. The probability of receiving a block without a burst error decreases with increasing block size and is small when $n=4095$.

Using the model of equation (73), the theoretical distribution of the number of errors per block is graphed. The results are presented in Figures 2-22 and 2-23 for two locations. The error distributions are seen to be good approximations to the actual field data. It is only when the block is very short (15 bits) that the theoretical distributions differ radically from the experimental ones.

The noteworthy point about the theoretical distributions is the dependence of their slope on the blocklength. The distributions become progressively steeper as the blocklength is increased. This is well illustrated in Figures 2-24 and 2-25. Although the distributions appear to be straight lines when plotted on a logarithmic scale, they are imperceptibly curved, suggesting a decay which is slightly more rapid than exponential.

Figure 2-20 Probability that an Error Burst will last for b bits or more, versus $\frac{b}{n}$ (Downtown) (Experimental)

Figure 2-21 Probability that an Error Burst will last for b bits or more, versus $\frac{b}{n}$ (Downtown) (Experimental)

Figure 2-22 Probability that an Error Burst will last for b bits or more, versus b/n . (Hull) (Experimental)

Figure 2-23 Probability that an Error Burst will last for b bits or more, versus b/n . (Hull) (Experimental)

Figure 2-24 Probability that the Number of Errors in a Block will exceed m bits, versus $\frac{m}{n}$. (Theoretical) (Downtown)

Figure 2-25 Probability that the Number of Errors in a Block will exceed m bits, versus $\frac{m}{n}$. (Theoretical) (Hull)

CHAPTER 3

CODING BOUNDS

3.1 Introduction

The likelihood of the correct reception of a digital message is usually increased if it is protected with redundant check bits. The bits are called redundant because they do not add to the information content of the message. They do, however, increase the immunity of the message to noise. A coder is a device that adds controlled redundancy, in the form of check bits, to a digital message. A decoder is a device that guesses at the message based on the received data and check bits. A decoder may also use other information, such as reliability information provided by the demodulator, to improve its chances at guessing correctly.

In this chapter, the effectiveness of a coder and decoder pair--as a means to increase the reliability of reception in a digital land mobile environment--is investigated. The model of the digital land mobile communication system (DLMCS) which is of interest is shown in Figure 3-1. The channel in the figure consists of the sections between the output of the coder and the input to the decoder. It includes the modulator and demodulator pair as well as the analog channel over which

Figure 3-1 Digital Land Mobile Communication System

the information is transmitted. The aggregate channel defined in such a manner is discrete: its input and output are letters chosen from a countable alphabet. The letters at the channel output are not necessarily from the same alphabet as the letters at the input. The size of the output alphabet, however, is always equal to or larger than that of the input alphabet. In the simplest case, the input is an arbitrary string of 0's and 1's, and the output is a copy that may differ from the original in a few positions.

The coders and decoders investigated will operate on a block by block basis. The coder will accept an information block of k symbols to which it will add r redundant symbols to form an n symbol codeword. Similarly, the decoder will accept an n symbol codeword and decode it into a k symbol information block. The information transmission rate will be k information symbols per n channel uses.

For many channels a limiting capacity on the throughput rate exists. In 1948 Shannon showed that reliable communication was possible if the rate of information transmission was less than the channel capacity. Conversely, he showed that reliable communication was not possible if the rate of information transmission was greater than channel capacity. For the purposes of the present discussion, capacity may be defined in terms of information symbols per channel use. For an uncoded system, a new information symbol is

transmitted at every channel use and the communication system is operating above capacity. Coding provides the option to reduce the information transmission rate--by occasionally sending redundant bits--to below channel capacity while simultaneously increasing the reliability of reception. An optimum coding scheme will operate reliably at any value below channel capacity. It will be of interest to determine how far below the optimum the block coding techniques investigated herein fall. To this end, a brief description of the channel capacity of a DLMCS will be given in the conclusion to the thesis.

3.2 ~~Bounds~~ on the Error Correction Ability of Block Codes, and On the Achievable Error Probabilities for the Discrete Land Mobile Channel: Outline of Approach

For the best block codes, there is a definite relationship between $d_{\min}$, the code's minimum distance, and the number of parity checks, $n-k$. Furthermore, the number of correctable errors also depends heavily on $d_{\min}$. A relation between (t/n) , the ratio of corrected errors per block, and the code rate $R=(k/n)$ can thus be established:

$$\frac{t}{n} = f(R) \quad \text{or} \quad R = f^{-1} \left(\frac{t}{n} \right) \quad (1)$$

The function $f(\cdot)$ depends on the structural properties of the code only. It is therefore independent

of any channel statistics. Many such functions have been proposed in the coding literature. The sphere-packing bound and the Elias bound, when multiplied by appropriate constants, are two such functions.

The channel, on the other hand, can be characterized by the distributions $p(m,n)$ and $p(b,n)$ of the number of errors per block and of the burst lengths per block respectively. The distributions $p(m/n,n)$ and $p(b/n,n)$ can be obtained from the above and can be combined with the function $f(\cdot)$ to relate the block error rate to the code rate as explained in what follows.

The error correction ability of a code as a function of the code rate R , can be upper bounded and lower bounded. If (t/n) represents the fraction of errors correctable by a code of length n , the following relations can be established:

$$\frac{t}{n} \geq f_L(R); \text{ lower bound} \quad (2)$$

$$\frac{t}{n} \leq f_H(R); \text{ upper bound} \quad (3)$$

similarly, if (b/n) represents the fraction of bursts correctable by a code of length n , the inequalities shown below can be established:

$$\frac{b}{n} \geq h_L(R); \text{ lower bound} \quad (4)$$

$$\frac{b}{n} \leq h_H(R); \text{ upper bound} \quad (5)$$

The lower bounds state the achievable. Their proofs usually indicate a method of construction. -Upper bounds

indicate the best that can be achieved under some very general constraints.

For the channel, the distributions $p(\geq m/n, n)$ of errors for blocks of various lengths have been presented and discussed in Chapter 2. $p(\geq m/n, n)$ has been used instead of the distribution function because the probability of block error is related to it, in a straightforward manner. If a decoder corrects up to m errors, only blocks containing more than m errors will cause decoding failures. The probability of block error after decoding is thus given by:

$$\text{BLER} = p(\geq \frac{m+1}{n}, n) \quad (6)$$

Combining (6) with (2) and (3) yields:

$$\text{BLER} = p(\geq \frac{t+1}{n}, n) \leq p(\geq [f_L(R) + \frac{1}{n}], n) \quad (7)$$

$$\text{BLER} = p(\geq \frac{t+1}{n}, n) \geq p(\geq [f_H(R) + \frac{1}{n}], n) \quad (8)$$

These two equations relate the BLER to the code rate R , with n as a parameter. Equation (7) is an upper limit on the BLER. Block coding schemes can, in principle, achieve what is indicated by (7) or better. Equation (8) is a lower limit on the BLER. No block coding technique will do any better than what is indicated by (8). The lower bound on t/n is seen to give an upper bound on the BLER and vice-versa.

Finally, expressions for the BER are obtained from the above as follows [24]:

$$P_e = \frac{d}{n} \text{BLER} \quad (9)$$

Or as a cruder, but more general bound:

$$P_e \leq \frac{1}{2} \text{BLER} \quad (10)$$

Equation (10) merely states that if a block is in error at most 1/2 of all the bits in the block will be in error—on the average. In many situations this is a very poor bound.

Equation (9) can be applied whenever the decoder corrects all patterns of t errors or less and none of greater weight. This is known as bounded distance decoding. The derivation of equation (9) is involved and will be delayed to a later section.

The distributions of burst lengths within a block of length n have been presented in Chapter 2. Again, it is convenient to use $p(\geq b/n, n)$. The decoded BLER is given by, for a decoder that corrects all bursts of length b or less and none of greater length:

$$\text{BLER} = p(\geq \frac{b+1}{n}, n) \quad (11)$$

Combining (11) with (4) and (5) yields:

$$\text{BLER} = p(\geq \frac{b+1}{n}, n) \leq p(\geq [h_L(R) + \frac{1}{n}], n) \quad (12)$$

$$\text{BLER} = p(\geq \frac{b+1}{n}, n) \geq p(\geq [h_H(R) + \frac{1}{n}], n) \quad (13)$$

The two equations relate the decoded BLER to the code rate R and the block length n . In our presentation of results, graphs will be plotted to show the variation of the decoded BLER as a function of R . n will be used as a parameter.

Expressions for the BER are harder to obtain for the case of burst error correction since:

$$P_e = \frac{1}{2} \text{BLER} \quad (14)$$

is a very crude bound. As an example, consider the case where the burst consists of 3 or 4 errors that are widely spaced.

3.3 Bounds on Error Correcting Ability of Block Codes

3.3.1 Introduction

In this section, bounds on the error correction ability of the "best" codes as a function of their rates are derived. These bounds are then used in conjunction with the channel statistics $p(\geq m, n)$ and $P(\geq b, n)$ to yield limits on the decrease in bit error rate (BER) and block error rate (BLER) attainable through coding. The BER and the BLER are both used because they are equally important performance measures in data transmission systems. The BER is generally used as the standard performance measure, but its disadvantage is that it does not distinguish between isolated bit errors and badly garbled contiguous

symbols, even though the latter are usually preferred. The BLER partially remedies the situation since clustered errors result in a lower BLER for a given BER. Neither measure is totally satisfactory and other performance measures, such as the mean error free run length, have been advocated by some authors [26]. Nevertheless, the BER and the BLER will be used because of their widespread acceptability. In cases where the BER is mathematically intractable, only the BLER will be given.

A code's ability to correct random errors, random erasures, a single burst, or an erasure burst is investigated. These are the four basic algebraic techniques that can be applied to the decoding of block codes. Other decoding techniques are either derived from the above or are partly or completely probabilistic and are not investigated in the thesis. The existing multiple burst correcting codes are a special case of the larger class of random error correcting codes and do not require separate treatment. The bounds for random error and erasure correction are obtained from bounds on the code's minimum distance, while those for burst error correction are obtained by relating the burst correcting ability to the number of check bits.

An examination of the coding literature reveals a multiplicity of bounds. The reasons are partly historical: as an understanding of the underlying principles grew, researchers developed tighter bounds. Earlier bounds were retained because of their simplicity

or the insights they offered into code construction. But in addition to historical reasons, the nature of the subject often makes it useful to specialize the bounds according to one or more of the following criteria:

(a) upper versus lower: lower bounds state what is achievable. They are generally constructive, and "good" codes fall on, or above the best lower bounds. They are deficient in one respect: they do not answer the question of how far above the bound the best codes might lie. Upper bounds on the other hand are usually non-constructive: they give an upper limit to the performance achievable by a class of codes. Taken together, the lower and upper bounds define a region where the best possible performance must lie.

(b) asymptotic versus finite block length: all bounds depend on the blocklength as well as the rate. In many cases the dependence on blocklength is weak, for example $O(1/n)$ or $O((1/n) \log(n))$. By letting n tend to infinity, the dependence on n is eliminated and a conceptually simpler bound which is a function of the code rate only is obtained. In some cases, the asymptotic bound remains valid for finite lengths but is weaker than its non-asymptotic counterpart; in others, the bound cannot be used unless the code is moderately long. A reason for considering bounds for finite n , is that for certain combinations of n and $d_{\min}$ they can be made extremely tight; they do not generalize to arbitrary n or $d_{\min}$ however.

(c) linear versus non-linear: define P_{im} as the probability of occurrence of the i th alphabeth letter in position m of a code, ($i=0,1,\dots,q-1$). Let $P_m = \{P_{im}\}$. If the code is constrained to have one of the following forms:

$$P_m = [1, 0, 0, \dots, 0];$$

or

$$P_m = [1/q, 1/q, \dots, 1/q]$$

(A proof of this is given in the next section).

If the code is non-linear, the probabilities P_{im} can be arbitrary. The greater freedom in the selection of the P_{im} 's gives some finite blocklength non-linear codes better error correction abilities than analogous linear codes: this is also reflected in the bounds. It can be shown however, that for large n , non-linear codes do not perform any better than linear ones [23].

In selecting the bounds to use in our research, we are guided by the above observations and we draw the following conclusions. First, asymptotic bounds can often be used even if the blocklength is finite. Secondly, a tight region of achievable performance cannot be defined unless both upper and lower bounds are considered. Lastly, the distinction between linear and non-linear codes is unimportant. The problem can now be reformulated as follows: to find asymptotic upper and lower bounds for the random error, the random erasure, or the burst error correction capability of a code as a function of its rate.

3.3.2 Definitions and Useful Results

The following definitions will be of interest for the remainder of this chapter.

A (n, M) block code of size M and blocklength n is a set of M vectors of length n , with symbols chosen from some field F . A (n, k) linear code is a subspace of dimension k of an n -dimensional vector space over a field F . The elements of F will be called $0, 1, \dots, q-1$.

The Hamming distance between two vectors x and y of length n , denoted $\|x-y\|$, is the number of coordinates in which they differ. The minimum distance of a code, $d_{\min}$, is the largest integer d , such that any two distinct vectors in the code have Hamming distance at least d . Throughout the chapter, the word distance will be taken to mean Hamming distance.

A sphere of radius r and centre Q is the set of all points whose distance from Q is less than or equal to r . The surface area of the sphere, denoted by $A_r^{(n)}$, is the number of points at a distance of exactly r from the centre of the sphere. If $V_r^{(n)}$ denotes the volume of the sphere, then
$$V_r^{(n)} = \sum_{i=0}^r A_i^{(n)}.$$

A burst of length b is an error pattern where all the non-zero components are among b successive components, the first and last of which are non-zero. A cyclic burst of length b is the smallest burst observed when the error-pattern is cyclically shifted by $1, 2, \dots, n$ shifts.

A resolution test chart featuring various patterns of vertical and horizontal lines. The patterns are labeled with numbers: 1.0, 1.1, 1.25, 1.4, 1.6, 1.8, 2.0, 2.2, 2.5, 2.8, 3.2, 3.6, 4.0, 4.5, 5.0, 5.6, 6.3, 7.1, 8.0, 9.0, 10, 11.2, 12.5, 14, 16, 18, 20, 22.5, 25, 28, 32, 36, 40, 45, 50, 56, 63, 71, 80, 90, 100. The lines become progressively smaller and more closely spaced as the numbers increase.

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

A burst sphere of radius b and centre C is the set of all points whose distance from C is a burst of length b or less.

The surface area of this burst sphere is denoted $AB_b^{(n)}$ and its volume is denoted $VB_b^{(n)}$. Clearly $VB_b^{(n)} = \sum_{i=0}^b AB_i^{(n)}$.

In addition, the following lemmas and results will be useful when deriving the bounds. The theorems can be found in any book on coding theory [20, 21, 23]. The proofs are original.

- Lemma 1: If $\{A_i\}$ is the distance distribution for some set S of n -dimensional vectors, then $\{A_i\}$ is invariant under translation. The proof is trivial: if a constant vector c is added to any two vectors x and y of S we have: $\|x-y\| = \|x-y+c-c\| = \|(x-c)-(y-c)\|$. $\square$

- Lemma 2: A code with minimum distance d can correct up to $\lfloor (d-1)/2 \rfloor$ errors or $(d-1)$ erasures, where $\lfloor x \rfloor$ denotes the greatest integer contained in x . The proof is simple; if less than $\lfloor (d-1)/2 \rfloor$ errors occur, the distance between the received word and the transmitted codeword is still less than the distance to any other codeword. Similarly, if $(d-1)$ or fewer erasures occur, then there exists one pattern that can be substituted for the erasures that will result in the transmitted codeword. No other pattern will, when substituted for the erasures, match the received word to any other codeword. $\square$

- Lemma 3 [21]: Let $A^{(n)}(z) = \sum_{i=0}^{\infty} A_i^{(n)} z^i$. $A^{(n)}(z)$ is the generating function for the A_i 's. Then $A^{(n)}(z) = [A^{(1)}(z)]^n = [1+(q-1)z]^n$. Here q is the cardinality of the field F over which the code is defined.

The proof hinges on the fact that the distance is additive across the coordinates. Let $A_1^{(k)}$ denote the number of vectors at distance 1 in a k -dimensional sphere; similarly, let $A_j^{(n-k)}$ denote the number of vectors at distance j in a $(n-k)$ -dimensional sphere. All the vectors at distance 1 in a k -dimensional space can be concatenated to each and every vector at distance j in a $(n-k)$ -dimensional space to yield vectors at distance $(1+j)$ in a $(k+n-k)=n$ -dimensional space. There are $A_1^{(k)}$ vectors at distance 1 and $A_j^{(n-k)}$ at distance j . This yields a total of $A_1^{(k)} A_j^{(n-k)}$ vectors at distance $(1+j)$. It is therefore seen that the total number of points at distance $\ell = (\ell - j + j)$ is given by: $A_\ell^{(n)} = \sum_{j=0}^{\infty} A_{\ell-j}^{(n-k)} A_j^{(k)}$.

From the above, the general result is derived:

$$\begin{aligned}
 A^{(n)}(z) &= \sum_{i=0}^{\infty} A_i^{(n)} z^i \\
 &= \sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} A_{i-j}^{(n-k)} A_j^{(k)} \right) z^i \\
 &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} A_{i-j}^{(n-k)} A_j^{(k)} z^{i-j} z^j \\
 &= \sum_{j=0}^{\infty} \left[\sum_{\ell=-j}^{\infty} A_\ell^{(n-k)} z^\ell \right] A_j^{(k)} z^j; \\
 &\quad \text{where } \ell = i-j \\
 &= \left(\sum_{\ell=0}^{\infty} A_\ell^{(n-k)} z^\ell \right) \left(\sum_{j=0}^{\infty} A_j^{(k)} z^j \right) \\
 &= A^{(n-k)}(z) A^{(k)}(z) \\
 &\quad \text{for all } k: 1 \leq k \leq n
 \end{aligned}$$

In particular then, $A^{(n)}(z) = (A^{(1)}(z))^n$.

If the field F has q symbols, then each of $q-1$ symbols is at distance 1 from any other symbol (in the Hamming sense), whereas this symbol is at distance 0 from itself. Thus $A^{(1)}(z) = [1 + (q-1)z]$.

• Lemma 4: For a linear code, with symbols chosen from a field F , the relative frequency distribution of the q field symbols obeys:

$$P_m = 1/q [1, 1, \dots, 1] \text{ or } P_m = [1, 0, 0, \dots, 0];$$

The proof is by contradiction. Assume that there is some symbol i for which $P_{im} < 1/q$, then there must be some symbol j for which $P_{jm} > 1/q$.

If $j \neq 0$, then it has an inverse, and there exists a field element $k = j^{-1}$ such that multiplying all the codewords which have j in the m th position by the scalar k shows that $P_{im} = P_{jm} > 1/q$, a contradiction, or if all the P_{im} 's = 0 except for one which is 1. This symbol must be the symbol 0, for if it were not, then adding the codeword to itself would show that $P_{2hm}, P_{3hm}, \dots, P_{qhm} = P_{0m}$ were also not equal to zero, which is impossible.

If $j = 0$, then adding all the codewords which have a zero in the m th position to one codeword which has the symbol 1 in the m th position shows that $P_{im} = P_{0m} > 1/q$, a contradiction. $\square$

• Lemma 5: $VB_b^{(n)} = q^{b-1} [(n-b)(q-1) + q]$

$$VB_b^{(n)} = \sum_{i=0}^b AB_i^{(n)} = [1 + n(q-1) + \dots + (n-i+1)(q-1)^2 q^{i-2} + \dots + (n-(b-1))(q-1)^2 q^{b-2}]$$

The first term denotes a burst of length 0. The general term has the form $(n-i+1)(q-1)^2 q^{i-2}$ and is derived as follows. In a burst of length i , the first and last symbols must be non-zero and there are $q-1$ non-zero symbols in an alphabet of size q . The remaining $(i-2)$ symbols can be arbitrary. Hence, there are $(q-1)^2 q^{i-2}$

burst patterns of length 1. Finally, there are $(n-(b-1))$ starting positions for a burst of length 1.

Rearranging gives:

$$\begin{aligned}
 & 1+n(q-1) + (n-1)(q-1)^2 q^0 + (n-2)(q-1)^2 q^1 + \dots \\
 & \quad + (n-(b-1))(q-1)^2 q^{b-2} \\
 & = 1+n(q-1) + \sum_{j=2}^b [(n-(j-1))(q-1)^2 q^{j-2}] \\
 & = 1+n(q-1) + (q-1)^2 \left[n \sum_{j=0}^{b-2} q^j - \sum_{j=0}^{b-1} j q^{j-1} \right] \\
 & = 1+n(q-1) + (q-1)^2 \left[n \frac{q^{b-1}-1}{q-1} - \frac{d}{dx} \left[\sum_{j=0}^{b-1} x^j \right]_{x=q} \right] \\
 & = 1+n(q-1) + (q-1)^2 \left[n \frac{q^{b-1}-1}{q-1} + b \frac{q^{b-1}}{q-1} - \frac{q^{b-1}}{q-1} \right] \\
 & = q^{b-1} [(n-b)(q-1) + q]
 \end{aligned}$$

• Lemma 6: $VB_b^{(n)} \leq VB_{2b}^{(n)}/q$, for all $q \geq 2$ and $2 \leq b \leq (n/2)+1$. One proof which requires some algebra is as follows:

$$VB_b^{(n)} \leq VB_{2b}^{(n)}/q \Leftrightarrow q^{b-1} ((n-b)(q-1)+q) \leq q^{2b-2} ((n-2b)(q-1)+q)$$

But if it can be shown that $q^{b-1}((n-b)q+q) \leq q^{2b-2}((n-2b)(q-1)+q)$, then the first inequality, which is slightly weaker, will certainly hold.

$$q^{b-1}((n-b)+q) \leq q^{2b-2}((n-2b)(q-1)+q)$$

$$\Leftrightarrow q^b(n-b+1) \leq q^{2b-2}(qn-2bq-n+2b+q)$$

$$\Leftrightarrow q^b(n-b+1) \leq q^{2b-1}(n-2b) \frac{n}{q} + \frac{2b}{q} + 1$$

$$\Leftrightarrow q^b(n-b+1) \leq q^{2b-1}((n-2b)(1-\frac{1}{q})+1).$$

Since $q \geq 2$ the following auxiliary inequality holds:

$$q^{2b-1}((n-2b)(1-\frac{1}{q})+1) \geq q^{2b-1}((n-2b)(1-\frac{1}{q})+1)$$

$$= q^{2b-1}(\frac{n}{2} - b + 1),$$

hence if $(q^b(n-b+1) \leq q^{2b-1}(\frac{n}{2} - b + 1))$; then the initial

inequality will certainly be satisfied.

But the new inequality reduces to:

$$(n-b+1) \leq q^{b-1}(\frac{n}{2} - b + 1)$$

which must be solved for b . Since there appears to be no easy solution in terms of b , another approach is to find values of the parameter n for which the inequality holds.

The following results are obtained:

$$(n-b+1) \leq q^{b-1}(\frac{n}{2} - b + 1)$$

$$\Leftrightarrow (n - q^{b-1} \frac{n}{2}) - b + 1 \leq - q^{b-1} b + q^{b-1}$$

$$\Leftrightarrow n(\frac{q^{b-1}}{2} - 1) \geq (b-1)(q^{b-1}-1)$$

$$\Leftrightarrow n \geq (b-1) \frac{(q^{b-1}-1)}{(\frac{q^{b-1}}{2}-1)}$$

In all cases,

$$\left(\frac{q^{b-1}}{2} - 1\right) < \left(\frac{q^{b-1}}{2} - \frac{1}{2}\right) = \frac{1}{2} (q^{b-1} - 1)$$

Hence the original inequality is certainly satisfied if:

$$n \geq \frac{(b-1) (q^{b-1} - 1)}{1/2 (q^{b-1} - 1)} = 2(b-1)$$

$$\Leftrightarrow b \leq \frac{n}{2} + 1$$

The inequality $VB_b(n) \leq VB_2(n)/q$ is thus satisfied for the range of values of q and b of interest. $\square$

3.3.3 The Elias Upper Bound

The Elias upper bound is a bound on the minimum distance of a code. In the next section the bound will be related to the error correction ability of a code. A derivation and proof is given in what follows. The proof is simpler than Berlekamp [21], but is more general than Peterson and Weldon [20] since it is not restricted to binary linear codes but is valid for linear and non-linear codes with symbols from arbitrary (q -ary) fields.

There exists a tighter upper bound on the minimum distance: the McEliece, Rodemich, Rumsey, and Welch bound [28]. It is valid for binary and non-binary codes. It has a limiting form but is also valid for finite blocklengths. Its chief disadvantage is its computational complexity, both for finite and asymptotic n . For finite n the bound is virtually unsolvable. For infinite n , the bound takes on the form:

$$\frac{d}{n} = \min f(R, u) \quad 0 \leq u \leq 1 - \frac{2d}{n}$$

Each pair $(R, d/n)$ requires a separate minimization over u . As an example a table with 100 entries would require 100 separate minimizations over u .

The difference between the Elias and the McEliece bound is slight, particularly for moderate and high rate codes. For this reason the McEliece bound will not be further investigated in the thesis.

The Elias bound is as follows.

Consider a set of k , n -dimensional codewords, $C = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$. The symbols $[c_1^0, c_1^1, \dots, c_1^{n-1}]$ for each codeword α_1 are chosen from some arbitrary field F of q elements: $F = \{0, 1, \dots, q-1\}$.

The total Hamming distance between all possible ordered pairs of codewords is given by:

$$\begin{aligned} d_{\text{tot}} &= \sum_{i=1}^k \sum_{j=1}^k \|\alpha_i - \alpha_j\| \\ &= \sum_{i=1}^k \sum_{j=1}^k \left[\sum_{\ell=0}^{n-1} \|c_i^\ell - c_j^\ell\| \right] \end{aligned}$$

where the second equality follows because the distance is additive across the coordinates.

For $m=0, 1, \dots, q-1$; let J_m^ℓ denote the number of occurrences of the m th alphabet letter among the code symbols, $c_1^\ell, c_2^\ell, c_3^\ell, \dots, c_k^\ell$. (i.e., symbols in the ℓ^{th} position across the codewords). The vector $P^\ell = 1/k [J_0^\ell, J_1^\ell, \dots, J_{q-1}^\ell]$ denotes the probability

distribution for the ℓ^{th} symbol of a codeword selected at random. If the code is linear, with symbols chosen from some field, p^ℓ must have one of the following forms:

$$p^\ell = [1, 0, \dots, 0]$$

or

$$p^\ell = [1/q, 1/q, \dots, 1/q] ; \text{ where } q = |F|$$

If the code is not linear, p^ℓ is any a-priori probability assignment of the q field elements over the ℓ^{th} code symbol.

Let $p^\ell = [p_0^\ell, p_1^\ell, \dots, p_{q-1}^\ell]$. d_{tot} may be rewritten as follows:

$$\begin{aligned} d_{\text{tot}} &= \sum_{\ell=0}^{n-1} \left[\sum_{i=1}^k \sum_{j=1}^k ||c_i^\ell - c_j^\ell|| \right] \\ &= \sum_{\ell=0}^{n-1} \left[k^2 \sum_{i=0}^{q-1} \sum_{j=0}^{q-1} p_i^\ell p_j^\ell (1 - \delta_{ij}) \right] \end{aligned}$$

where δ_{ij} is the Kronecker delta function i.e., $\delta_{ij} = 0, 1$. $\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ otherwise. The last equality arises because the distance between pairs of symbols is 1 unless the symbols are identical, in which case the distance is zero. $k^2 \cdot p_i^\ell$. $p_j^\ell = J_i^\ell \cdot J_j^\ell$, represents the number of ordered pairs out of the total of k^2 pairs which contain elements i and j .

Thus,

$$d_{\text{tot}} = \sum_{\ell=0}^{n-1} \left[k^2 \sum_{i=0}^{q-1} \sum_{\substack{j=0 \\ j \neq i}}^{q-1} p_i^\ell p_j^\ell \right]$$

and since

$$\sum_{j=0}^{q-1} P_j^{\ell} = \sum_{\substack{j=0 \\ j \neq 1}}^{q-1} P_j^{\ell} + p_1^{\ell} = 1$$

then,

$$d_{\text{tot}} = \sum_{\ell=0}^{n-1} [k^2 \sum_{i=0}^{q-1} P_i^{\ell} (1-P_i^{\ell})]$$

$$d_{\text{tot}} \leq k^2 \sum_{\ell=0}^{n-1} [\sum_{i=0}^{q-1} P_i (1-P_i)]$$

where $P = [P_1, P_2, \dots, P_1, \dots, P_{q-1}]$ represents the best possible column composition. Thus, the total Hamming distance for a set of K codewords is bounded by:

$$d_{\text{tot}} \leq nk^2 [\sum_{i=0}^{q-1} P_i (1-P_i)]$$

Maximizing d_{tot} is equivalent to maximizing

$$\sum_{i=0}^{q-1} P_i (1-P_i)$$

subject to certain constraints.

One obvious constraint is $\sum P_i = 1$; however, this constraint alone is not sufficient. The bound it yields, known as the Plotkin bound, is tight for low rates only. A stronger bound is obtained if an argument, first introduced by Elias, is used. The argument states that there must exist a critical sphere which contains many codewords (more than average), in which case, the distance between some pair of codewords in the critical sphere must be small.

More formally, Elias's argument can be broken down into two lemmas: (A) given a code of rate R , and

blocklength n , there exists a critical sphere of radius r which includes M codewords, where $M \geq V_r(n)/q^{n(1-R)}$. The vector at the centre of the sphere can be subtracted from all other vectors in the sphere to centre it at $\mathbf{0}$, without changing the distance distribution. (B) if each of M codewords has weight $\leq (q-1) \cdot x \cdot n/q$, then the distance between some pair of codewords must be less than

$$\left[\frac{q-1}{q} n / (1-M^{-1}) \right] x(2-x),$$

The proof of (A) is straightforward.

Any sphere of radius r will enclose a volume which contains $V_r(n)$ vectors of F^n . Out of a total of q^n vectors in F^n only q^{Rn} are codewords. Thus, on the average, each sphere of radius r encloses $V_r(n) \cdot (q^{Rn}/q^n)$ codewords. Therefore, there exists some sphere for which the number of codewords M , is at least as great as average:

$$M \geq \frac{V_r(n)}{q^{n(1-R)}}$$

By lemma 1, this critical sphere can be translated to $\mathbf{0}$ without affecting the distance distribution.

The proof of (B) is more involved and is as follows. Using the same arguments as before, the total distance between the M^2 ordered pairs of codewords is given by:

$$\begin{aligned} d_{\text{tot}} &= \sum_{i \in M} \sum_{j \in M} \|c_i - c_j\| \\ &= M^2 \sum_{\ell=0}^{n-1} \left[\sum_{i=0}^{q-1} P_i^\ell (1-P_i^\ell) \right] \end{aligned}$$

On the other hand, if each codeword has Hamming weight less than $(q-1) \cdot x \cdot n/q$ -- as per the statement of the lemma -- then the total weight of the M codewords is at most equal to $M (q-1) \cdot x \cdot n/q$. The total weight is given by:

$$\begin{aligned}
 w_{\text{tot}} &= \sum_{i=0}^{M-1} \sum_{\ell=0}^{n-1} |c_{i\ell}| \\
 &= \sum_{\ell=0}^{n-1} \left[M \sum_{i=0}^{q-1} (1 - \delta_{0i}) P_1^\ell \right] \\
 &= \sum_{\ell=0}^{n-1} \left[M \sum_{i=1}^{q-1} P_1^\ell \right] \\
 &\leq \frac{q-1}{q} \cdot x \cdot nM
 \end{aligned}$$

The equalities are derived using arguments similar to those used earlier, and the final inequality follows from the fact that the total weight is at most $M[(q-1)/q] \cdot x \cdot n$. Cancelling the M 's in the last two equations yields:

$$\sum_{\ell=0}^{n-1} \left[\sum_{i=1}^{q-1} P_1^\ell \right] \leq \left[\frac{q-1}{q} x \right] n$$

The minimum distance cannot be greater than the average distance between the M codewords. Since the number of codewords is constant, the average distance is greatest when the total distance is maximized. Maximizing the total distance, d_{tot} , is equivalent to maximizing:

$$M2 \sum_{\ell=0}^{n-1} \left[\sum_{i=0}^{q-1} P_i^{\ell} (1 - P_i^{\ell}) \right]$$

which in turn is equivalent to maximizing

$$\sum_{i=0}^{q-1} P_i^{\ell} (1 - P_i^{\ell})$$

under the constraints:

$$\sum_{i=0}^{q-1} P_i^{\ell} = 1$$

$$\sum_{\ell=0}^{n-1} \left(\sum_{i=0}^{q-1} P_i^{\ell} \right) \leq n \left(\frac{q-1}{q} x \right)$$

The second constraint—due to the constraint on the weight of each individual codeword—can in turn be split into two subconditions. A new variable, x_{ℓ} , is introduced such that:

$$\left(\sum_{i=0}^{q-1} P_i^{\ell} \right) = \left(\frac{q-1}{q} x_{\ell} \right)$$

The second constraint now becomes:

$$\sum_{\ell=0}^{n-1} (x_{\ell}) \leq nx$$

and the maximization is performed in two steps:

(a) maximize $\sum_{i=0}^{q-1} P_i^{\ell} (1 - P_i^{\ell})$ subject to the constraints:

$$\sum_{i=0}^{q-1} P_i^{\ell} = \frac{q-1}{q} x_{\ell}$$

$$\sum_{i=0}^{q-1} P_i^{\ell} = 1$$

Using the technique of Lagrange multipliers, we obtain the objective function:

$$\sum_{i=0}^{q-1} P_i^{\ell} (1-P_i^{\ell}) + A \left(\sum_{i=1}^{q-1} x_i^{\ell} \right) - \sum_{i=1}^{q-1} P_i^{\ell} + B \left(1 - \sum_{i=0}^{q-1} P_i^{\ell} \right)$$

setting $\frac{\partial(\cdot)}{\partial P_i^{\ell}} = 0$, we obtain:

$$\begin{cases} \text{for } i = 0 & -P_0^{\ell} + 1 - P_0^{\ell} - B = 0 \\ \text{for } i \neq 0 & -P_i^{\ell} + 1 - P_i^{\ell} - A - B = 0 \end{cases}$$

or

$$\begin{cases} P_0^{\ell} = (1-B)/2 \\ P_i^{\ell} = (1-A-B)/2, \quad i \neq 0 \end{cases}$$

From the two constraints we then get:

$$B = (1 - 2P_0^{\ell}) = 1 - 2 \left(\sum_{i=0}^{q-1} P_i^{\ell} - \sum_{i=1}^{q-1} P_i^{\ell} \right) = \frac{2(q-1)}{q} x_{\ell} - 1$$

$$\Rightarrow P_0^{\ell} = 1 - x_{\ell} + \frac{x_{\ell}}{q}$$

$$\sum_{i=1}^{q-1} P_i^{\ell} = \sum_{i=1}^{q-1} (1-A-B)/2 = \frac{q-1}{q} x_{\ell}$$

$$\Rightarrow A = 2(1-x_{\ell}) \text{ and } \Rightarrow P_i^{\ell} = x_{\ell}/q, \quad i \neq 0.$$

From which we get:

$$\begin{aligned} d_{\text{tot}} &= M^2 \sum_{\ell=0}^{n-1} \left[\sum_{i=0}^{q-1} P_i^{\ell} (1-P_i^{\ell}) \right] \\ &= M^2 \sum_{\ell=0}^{n-1} [P_0^{\ell}(1-P_0^{\ell}) + (q-1) P_i^{\ell} (1-P_i^{\ell})]; \quad i \neq 0 \\ &= M^2 \sum_{\ell=0}^{n-1} \frac{q-1}{q} [x_{\ell} (2-x_{\ell})] \end{aligned}$$

The second step consists in maximizing d_{tot} subject to

the constraint that $\sum_{\ell=0}^{n-1} x_{\ell} \leq nx$. This is equivalent to:

(b) Maximize $\sum_{\ell=0}^{n-1} x_{\ell} (2-x_{\ell})$ subject to the constraint:

$$\sum_{\ell=0}^{n-1} x_{\ell} \leq nx.$$

The objective function is:

$$\sum_{\ell=0}^{n-1} x_{\ell} (2-x_{\ell}) + A(nx - \sum_{\ell=0}^{n-1} x_{\ell})$$

setting its partial derivatives w.r.t. x to zero yields:

$$\frac{\partial(\cdot)}{\partial x_{\ell}} = (2-x_{\ell}) - x_{\ell} - A = 0$$

$$\Leftrightarrow 2 - 2x_{\ell} = A \quad \Leftrightarrow -\frac{A+2}{2} = x_{\ell}$$

From the constraint $nx \geq \sum_{\ell=0}^{n-1} x_{\ell} = n[\frac{2-A}{2}]$, we then get:

$$x_{\ell} = x, \text{ thus we have } d_{\text{tot}} = nM^2 \cdot x \cdot (2-x) \frac{q-1}{q}.$$

The average distance among the (M^2-M) pairs of distinct codewords is then given by:

$$d_{\text{AV}} = \frac{n M^2 \cdot x \cdot (2-x) \frac{q-1}{q}}{M^2 - M}$$

and since the minimum distance cannot exceed the average, d_{min} is bounded by:

$$d_{\text{min}} \leq \frac{n \cdot x \cdot (2-x) \frac{q-1}{q}}{1 - M^{-1}}$$

As a final step, lemmas (A) and (B) must be combined to yield a bound which relates the minimum distance to the code rate R .

In lemma (A), the existence of a sphere of radius r which contained more than $V_r(n)/q^{n(1-R)}$ codewords was demonstrated. It was also shown that the sphere could be centered at the origin without affecting the distance distribution among the code words. Taken together these two conditions establish the existence of a critical sphere with many codewords, in which the weight of each codeword is no greater than r .

Combining these observations with the results of part (B) yields:

$$r = \frac{q-1}{q} \cdot x \cdot n \quad \Leftrightarrow \quad x = \frac{q r}{(q-1)n}$$

and $M \geq V_r(n)/q^{n(1-R)}$ giving,

$$d_{\min} \leq \frac{r(2 - qr/((q-1)n))}{1 - q^{n(1-R)}/V(n)}$$

The result holds for any value of r which makes $M > 1$; that is, any value of r which makes the critical sphere non-empty.

The tightest possible bound is obtained if we minimize this quantity with respect to r :

$$\frac{d_{\min}}{n} \leq \min_r \frac{r/n(2 - rq/(n(q-1)))}{(1 - M^{-1})}$$

The minimization will be carried out in a subsequent section where asymptotic forms for the Elias and other bounds will be considered.

3.3.4 The Gilbert-Varsharmov Lower Bound

The bound is a lower bound on the minimum distance of a code. It was originally derived by Gilbert, subsequently it was refined by Varsharmov and Sacks.

Let e be a code of blocklength n , and minimum distance $d_{\min}$. If the number of codewords in e satisfies $q^{Rn} \times q^{n/V_{d-1}(n)} < q^n$ then more codewords can be added to the code without reducing $d_{\min}$.

The proof is simple*. If $q^{Rn} V_{d-1}(n) < q^n$, then the interstitial space between the spheres of radius $d-1$ is not empty. All the vectors in the interstitial space are at distance $\geq d$ from any codeword. Hence any one of these vectors may be chosen as an additional codeword without reducing $d_{\min}$.

The above theorem states that codewords can be added to the code—one at a time—until the left hand side and the right hand side of the inequality are arbitrarily close to each other. Rewriting the inequality we have:

$$V_{d-1}(n) < q^{n(1-R)}$$

$$\Leftrightarrow \log V_{d-1}(n) < n(1-R)$$

$$\Leftrightarrow 1 - \frac{1}{n} \log V_{d-1}(n) > R$$

* Surround each codeword with a sphere of radius $d-1$. There are q^{Rn} codewords and hence $q^{Rn} V_{d-1}(n)$ points within the spheres.

and according to the Gilbert bound codewords can be added until the rate R verifies:

$$R \geq 1 - \frac{1}{n} \log V_{d-1}^{(n)}$$

In the next section limiting forms for the expression

$$\frac{1}{n} \log_q V_{d-1}^{(n)} = \frac{\log_q V_{n(\frac{d-1}{n})}^{(n)}}{n}$$

will be considered. Also, the Gilbert bound will be given in a form which relates d/n to R instead of its current form which relates R to d/n .

3.3.5 An Upper Bound For Burst Error Correction

In this section an upper bound for the burst error correction ability of a code is given. The bound is valid for linear and non-linear codes; but it is later shown that linear codes do not lie close to the bound.

Let C be a code of length n with q^{Rn} codewords. If the code is to correct all bursts of length b or less, then all the patterns obtained by adding a burst of length b or less to a codeword must be distinct. This might be possible if:

$$q^{Rn} (1 + n(q-1) + (n-1)(q-1)^2 + \dots + (n-(b-1))(q-1)^{b-2} q^{b-2}) \leq q^n$$

The first term in the parenthesis denotes the codeword; the next represents all bursts of length 1. The general term has the form $(n-i+1)(q-1)^{i-2} q^{i-2}$ and denotes all the bursts of length i : the first and last

symbol of the burst must be non-zero; the middle symbols are arbitrary and there are $(i-2)$ of them; finally, there are $(n-(i-1))$ starting positions for a burst of length i .

Rearranging the above gives:

$$q^{Rn}(1+n(q-1) + (q-1)^2 \left[n \sum_{j=0}^{b-1} q^j - \sum_{j=1}^{b-1} j q^{j-1} \right]) \leq q^n$$

or,

$$q^{Rn}(1+n(q-1) + (q-1)^2 \left[n \sum_{j=0}^{b-2} q^j - \frac{d}{dx} \left[\sum_{j=0}^{b-1} x^j \right]_{x=q} \right]) \leq q^n$$

but the parenthesis was shown to be equal to (lemma 5):

$$q^{b-1}[(n-b)(q-1) + q]$$

thus the inequality reduces to:

$$q^{nR} q^{b-1} [(n-b)(q-1) + q] \leq q^n$$

taking logarithms (to the base q) and rearranging yields:

$$b \leq n(1-R) - \log_q \left[\frac{(n-b)(q-1) + q}{q} \right]$$

finally, dividing by n yields the desired result:

$$\frac{b}{n} \leq (1-R) - \frac{1}{n} \left\{ \log_q \left[\frac{(n-b)(q-1) + q}{q} \right] \right\}$$

This bound is in many ways similar to the sphere packing bound for random-error correcting codes. Hence, the problem of finding good burst correcting codes is not unlike the problem of finding perfect codes (codes that meet the sphere packing bound). Nevertheless, the next theorem shows that good burst correcting codes--codes that lie close to the above bound--are non-linear or very short:

In order to correct all bursts of length b or less a linear code must satisfy $n-k \geq 2b$.

The above theorem, due to Rieger, may be proved as follows: any burst of length $2b$ can be written as the difference between two bursts of length b or less: for example, over $GF(3)$ the burst 0 0 2 1 0 0 1 2 1 0 0 0 of length seven may be written as:

$$\begin{array}{r} 0\ 0\ 2\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\ -0\ 0\ 0\ 0\ 0\ 0\ 0\ 2\ 1\ 2\ 0\ 0\ 0. \end{array}$$

Since the code is linear, a parity-check matrix, H , may be defined. It has $(n-k)$ rows and n columns. Since the code has $n-k$ parity-checks, the rank (dimension) of H is $n-k$. Applying the parity check equations to a vector C , which is a burst of length $2b$, and rewriting C as the difference between two bursts of length b or less gives:

$$HC^T = HC_1^T - HC_2^T$$

But, $HC^T = c_1(h_1) + c_2(h_2) + \dots + c_n(h_n)$, where the (h_i) 's represent the columns of H . Because H has rank $(n-k) \geq 2b$, no linear combination of $2b$ or fewer columns of H is zero. Thus, $HC^T \neq 0$: this implies $HC_1^T \neq HC_2^T \neq 0$, and each burst of length b or less has a different syndrome. Thus a linear code can correct all bursts of length b or less if:

$$n - k \geq 2b$$

$$\Leftrightarrow 1 - (k/n) = 1 - R \geq 2b/n$$

$$\Leftrightarrow b/n \leq (1 - R)/2$$

3.3.6 A Lower Bound on the Burst Error Correction Ability of a Code

The theorem derived in this section is valid for linear codes. If the number of codewords, q^{Rn} , in a code C is chosen such that:

$$q^{Rn} (VB_{2b}^{(n)}/q) \leq q^n$$

then the code can correct all bursts of length b or less.

Rewriting the above we have:

$$q^{(n-k)} \geq VB_{2b} / q.$$

The number of cosets $q^{(n-k)}$ is greater than VB_{2b}/q and is therefore greater than the number of bursts of length b or less because $VB_b < VB_{2b}/q$ for all $q \geq 2$, $b \geq 2$, and $b \leq (n/2+1)$ (lemma 6). All bursts of length b or less can thus be chosen as coset leaders. If the number of parity checks is chosen such that $q^{(n-k)}$ is only slightly larger than VB_{2b}/q then the following inequality is satisfied instead:

$$q^{(n-k)} < VB_{2b}^{(n)}$$

$$q^{n(1-R)} < VB_{2b}^{(n)}$$

Rearranging the above:

$$n(1-R) \leq \log_q VB_{2b}^{(n)}$$

$$n(1-R) \leq 2b-1 + \log_q [(n-2b)(q-1) + q]$$

$$2b \geq n(1-R) - \log_q [(n-2b)(q-1) + q] + 1$$

$$2b \geq n(1-R) - \log_q \left[\frac{(n-2b)(q-1) + q}{q} \right]$$

Thus a burst correcting code can be designed which satisfies the following inequality:

$$\frac{2b}{n} \geq (1-R) - \frac{1}{n} \log_q \left[\frac{(n-2b)(q-1) + q}{q} \right]$$

For finite n , the result presented here is slightly stronger than a similar result, due to C.N. Campopiano, and presented in Peterson and Weldon [20]. Nevertheless, the two results are asymptotically similar.

Little is known about the burst error correcting capability of nonlinear codes in general.

3.3.7 Erasure Burst Error Correction

Decoders for linear burst correcting codes are among the simplest to instrument. Decoders for erasure burst correction are conceptually even simpler. For some channels, the use of erasure burst techniques offers a significant performance advantage. In light of the above it is surprising that little or no mention is made of erasure burst correction in the literature.

Using arguments resembling those in section 3.3.5, it can be shown that the number of parity checks in a linear code that corrects all burst of length eb or less satisfies $n-k \geq eb$.

Choose two distinct vectors that are zero except in eb successive locations. Their syndromes must be distinct because if:

$$HC_1^T = HC_2^T$$

$$\Leftrightarrow H(C_1^T - C_2^T) = 0$$

then $(C_1 - C_2)$ is a code vector and a burst of length eb

or less. But this implies the existence of an erasure pattern of length no greater than eb which is a code vector and is therefore incorrigible. Thus, to correct all erasure bursts of length eb or less the syndromes must be distinct and non-zero. This implies that no linear combination of eb or fewer columns of H may be zero. The rank of H must therefore be at least eb and the number of parity checks must satisfy:

$$n - k \geq eb$$

A method of construction shows that the above bound is achievable. Build a code with $n - eb$ information symbols and eb check symbols. Let each check symbol be chosen to make zero the sum of one of the first eb symbols and every eb th symbol following it. Each location in an erasure of length eb or less affects one check symbol only and can therefore be recovered since it is the only unknown in an equation involving that particular check symbol. An example of such a code is given below (based on [20]).

Let $n = 20$ and $eb = 5$ then $k = 15$.

If the first 15 symbols are information symbols, and the last five symbols are checks:

$$c_{16} = c_1 + c_6 + c_{11}$$

$$c_{17} = c_2 + c_7 + c_{12}$$

$$\vdots$$

$$c_{20} = c_5 + c_{10} + c_{15}$$

If an erasure burst affects $c_3 \dots c_7$ inclusive, then $c_3 = c_{18} - c_8 - c_{13} ; \dots c_7 = c_{17} - c_2 - c_{12}$.

In conclusion, the following equality applies for the erasure burst correction ability of linear codes:

$$\frac{eb}{n} = 1 - R$$

Little is known about the erasure burst correction ability of non-linear codes in general.

3.3.8 Limiting Forms for the Elias and the Gilbert Bounds

In the previous sections, two bounds on the minimum distance of a code were derived. In this section, asymptotic expressions for these bounds will be considered. Also, the bounds will be specified as a function of the parameter R (the independent variable). Recapitulating, the bounds on minimum distance derived earlier were:

the Elias bound:

$$\frac{d_{\min}}{n} \leq \min_r \frac{r}{n} \left(2 - \frac{r}{n} \left(\frac{q}{q-1} \right) \right) \wedge \left(1 - \frac{q^{n(1-R)}}{v_r^{(n)} n^{\left(\frac{r}{n}\right)}} \right)$$

the Gilbert bound:

$$R \geq 1 - \frac{1}{n} \log_q v_r^{(n)} n^{\left(\frac{d-1}{n}\right)}$$

In order to evaluate the bounds, a formula for $\frac{1}{n} \log_q v_r^{(n)}$ is required.

Recall that $A^{(n)}(z) = [1 + (q-1)z]^n = \sum_{i=0}^n A_i^{(n)} z^i$

and that $V_r^{(n)} = \sum_{i=0}^r A_i^{(n)}$. If $r < (q-1)n/q$, then it can be shown that, (Berlekamp, 1968 [21]):

$$\min_{0 \leq z \leq 1} \frac{z^{-r} A^{(n)}(z)}{n+1} \leq A_r^{(n)} \leq V_r^{(n)} \leq \min_{0 \leq z \leq 1} z^{-r} A^{(n)}(z)$$

The middle inequality, $A_r^{(n)} \leq V_r^{(n)}$, requires no proof. Furthermore, if $0 \leq z \leq 1$, then:

$$\begin{aligned} V_r^{(n)} &= \sum_{i=0}^r A_i^{(n)} \leq \sum_{i=0}^r A_i^{(n)} z^{i-r} = z^{-r} \sum_{i=0}^r A_i^{(n)} z^i \\ &= z^{-r} A^{(n)}(z) \end{aligned}$$

proving the last inequality.

Finally, the left inequality may be proved as follows: Assume there exists some value (or values) of z for which: $A_t^{(n)} z^t \geq A_i^{(n)} z^i$ for all $i \neq t$. If this is true, there exists $\bar{z}$, such that:

$$A^{(n)}(\bar{z}) = \sum_{i=0}^n A_i^{(n)} (\bar{z}^i) \leq \sum_{i=0}^n A_t^{(n)} \bar{z}^t = (n+1) A_t^{(n)} \bar{z}^t$$

Upon rearranging

$$\frac{A^{(n)}(\bar{z})}{(n+1)\bar{z}^t} \leq A_t^{(n)}$$

and therefore

$$\min_{0 \leq z \leq 1} \frac{z^t A^{(n)}(z)}{(n+1)} \leq A_t^{(n)}$$

Establishing the left inequality. To complete the proof

we need to show that a value of z satisfying $A_t^{(n)} z^t \geq A_i^{(n)} z^i$ for all $i \neq t$, does indeed exist.

Assume there exists such a value of z . The following can be established:

$$A_t^{(n)} z^i \geq A_{t+1}^{(n)} z^{t+1}$$

and

$$A_t^{(n)} z^t \geq A_{t-1}^{(n)} z^{t-1}$$

which in turn implies:

$$\frac{A_{t-1}^{(n)}}{A_t^{(n)}} \leq z \leq \frac{A_t^{(n)}}{A_{t+1}^{(n)}}$$

but showing the external inequalities is straightforward:

$$\frac{A_{t-1}^{(n)}}{A_t^{(n)}} \leq \frac{A_t^{(n)}}{A_{t+1}^{(n)}} \Leftrightarrow \frac{A_{t-1}^{(n)} A_{t+1}^{(n)}}{A_t^{(n)} A_t^{(n)}} \leq 1$$

$$\Leftrightarrow \frac{(q-1)^{t-1} (q-1)^{t+1} \binom{n}{t-1} \binom{n}{t+1}}{(q-1)^t (q-1)^t \binom{n}{t} \binom{n}{t}} \leq 1$$

$$\Leftrightarrow \frac{t! t! (n-t)! (n-t)!}{(t-1)! (n-t+1)! (t+1)! (n-t-1)!} \leq 1$$

$$\Leftrightarrow \frac{t}{t+1} \times \frac{n-t}{(n-t)+1} \leq 1$$

which is obviously true when t and n are positive and $t \leq n$. Since the inequality is not strict, the interval:

$$\frac{A_{t-1}^{(n)}}{A_t^{(n)}} \leq z \leq \frac{A_t^{(n)}}{A_{t+1}^{(n)}}$$

is therefore non-empty establishing the existence of z .

($0 \leq z \leq 1$). Finally, because $\frac{A_{t-1}^{(n)}}{A_t^{(n)}} \leq \frac{A_t^{(n)}}{A_{t+1}^{(n)}}$

for $0 \leq t \leq (q-1)n/q$, the following sequence of inequalities holds:

$$\frac{A_{t-3}^{(n)}}{A_{t-2}^{(n)}} \leq \frac{A_{t-2}^{(n)}}{A_{t-1}^{(n)}} \leq \frac{A_{t-1}^{(n)}}{A_t^{(n)}} \leq z \leq \frac{A_t^{(n)}}{A_{t+1}^{(n)}} \leq \frac{A_{t+1}^{(n)}}{A_{t+2}^{(n)}} \leq \frac{A_{t+2}^{(n)}}{A_{t+3}^{(n)}}$$

But this sequence implies:

$$z^{t-m} A_{t-m}^{(n)} \leq z^{t-m+1} A_{t-m+1}^{(n)} \leq \dots \leq z^t A_t^{(n)}$$

and

$$z^t A_t^{(n)} \geq z^{t+1} A_{t+1}^{(n)} \geq z^{t+2} A_{t+2}^{(n)} \dots$$

Thus the statement that there exists z such that

$$\frac{A_{t-1}^{(n)}}{A_t^{(n)}} \leq z \leq \frac{A_t^{(n)}}{A_{t+1}^{(n)}}$$

is equivalent to the statement that there exists z such that $A_t^{(n)} z^t \geq A_1^{(n)} z^1$ for all $i \neq t$. This establishes the truth of the last statement, completing the proof.

Having established that the volume, $V_r^{(n)}$, is bounded by:

$$\min_{0 \leq z \leq 1} \frac{z^{-r} A^{(n)}(z)}{n+1} \leq V_r^{(n)} \leq \min_{0 \leq z \leq 1} z^{-r} A^{(n)}(z)$$

the bounds on the minimum distance may be evaluated. Differentiating with respect to z to find the minimum yields:

$$\frac{d(z^{-r} A^{(n)}(z))}{dz} = \frac{d(z^{-r} [1+(q-1)z]^n)}{dz} = 0$$

$$\Leftrightarrow -r z^{-r-1} [1+(q-1)z]^n$$

$$+ n[1+(q-1)z]^{n-1} (q-1) z^{-r} = 0$$

$$\Leftrightarrow z^{-r-1} [1+(q-1)z]^{n-1} [-r[1+(q-1)z] + n(q-1)z] = 0$$

$$\Leftrightarrow z^{-r-1} [1+(q-1)z]^{n-1} [-r+z(q-1)(n-r)] = 0$$

the two solutions are $z=0$ or $z = \frac{r}{(q-1)(n-r)}$.

$$\text{This gives } A^{(n)}(z) = \left(\frac{n}{n-r}\right)^n = \left(\frac{1}{1-r/n}\right)^n$$

Having found the minimum value for z , $1/n \log V_{d-1}^{(n)}$ can be bounded:

$$\frac{[(q-1)^{\frac{d-1}{n}} (\frac{d-1}{n})^{-\frac{d-1}{n}} (1 - \frac{d-1}{n})^{-(1 - \frac{d-1}{n})}]^n}{n+1} \leq V_n^{(n)} (\frac{d-1}{n})$$

$$\leq [(q-1)^{\frac{d-1}{n}} (\frac{d-1}{n})^{-\frac{d-1}{n}} (1 - \frac{d-1}{n})^{-(1 - \frac{d-1}{n})}]^n$$

Since $\log_q(\cdot)$ is a monotonically increasing function, then:

$$\frac{1}{n} \{[(\frac{d-1}{n}) \log_q (q-1) - (\frac{d-1}{n}) \log_q (\frac{d-1}{n}) - (1 - \frac{d-1}{n})$$

$$\log_q (1 - \frac{d-1}{n})]n - \log(n+1)\}$$

$$\frac{1}{n} \log_q V_n^{(n)} (\frac{d-1}{n}) \leq \frac{n}{n} [(\frac{d-1}{n}) \log_q (q-1) - (\frac{d-1}{n})$$

$$\log_q (\frac{d-1}{n}) - (1 - \frac{d-1}{n}) \log_q (1 - \frac{d-1}{n})]$$

or

$$(\frac{d-1}{n}) \log_q (q-1) + H_q (\frac{d-1}{n}) - \frac{1}{n} \log (n+1) \leq$$

$$\frac{1}{n} \log_q V_n^{(n)} (\frac{d-1}{n}) \leq (\frac{d-1}{n}) \log_q (q-1) + H_q (\frac{d-1}{n})$$

where $H_q(\cdot)$ is the entropy function (in base q). For $q=2$ the function is depicted in Figure 3-2. The bound on

$$\frac{1}{n} \log \frac{V_n^{(n)} (\frac{d-1}{n})}{n}$$

is tight as n increases, because the difference between the left and the right sides tends to zero as n tends to infinity.

The Gilbert and Elias bounds can now be evaluated for asymptotic n .

For the Gilbert bound, the expression:

$$R \geq 1 - \left(\frac{d-1}{n}\right) \log_q (q-1) - H_q \left(\frac{d-1}{n}\right)$$

is a bound which is also valid for finite n . Define the function $G_q(x) = 1 - x \log_q (q-1) - H_q(x)$ and define implicitly $G_q^{-1}(x)$, the inverse function of $G_q(\cdot)$, as the smallest solution to:

$$G_q^{-1}(G_q(x)) = x$$

The function $G_q(x)$ is monotonically decreasing for $0 \leq x \leq 1/2$, and so is $G_q^{-1}(\cdot)$. Thus if $a \leq b$, $G_q^{-1}(a) \geq G_q^{-1}(b)$ and in particular since:

$$G_q \left(\frac{d-1}{n}\right) \leq R$$

$$G_q^{-1} \left(G_q \left(\frac{d-1}{n}\right)\right) \geq G_q^{-1}(R)$$

or

$$\frac{d-1}{n} \geq G_q^{-1}(R)$$

This is the asymptotic form of the Gilbert bound. For the Elias bound, it was shown that:

$$\frac{d_{\min}}{n} \leq \min_r \frac{\frac{r}{n} \left(2 - \frac{r}{n} \left(\frac{q}{q-1}\right)\right)}{(1 - M^{-1})}$$

which holds for any choice of r which makes $M > 1$.

Recall

$$M > \frac{V_r(n)}{n(1-R)}$$

(a)

(b)

(c)

(d)

Figure 3-2 The functions $H_2(x)$ and $G_2(x)$ and their inverses

(a) $H_2(x) = -x \log_2 x - (1-x) \log_2 (1-x)$

(b) $H_2^{-1}(x)$

(c) $G_2(x) = 1 + x \log_2 x + (1-x) \log_2 (1-x)$

(d) $G_2^{-1}(x)$

The value of $r = G_q^{-1}(R)n$ has the property that as n tends to infinity,

$$V_{nG_q^{-1}(R)}^{(n)} = q^{n(1-R)}$$

so that any value of $r < G_q^{-1}(R)n$ does not make $M > 1$ in the bound. Conversely, for any value of $r > G_q^{-1}(R)n$ the volume $V_r^{(n)}$ increases so rapidly with increasing n that the term $1/M$ becomes negligible. Thus, the minimum value for the Elias bound occurs if r is made as small as possible provided $r > G_q^{-1}(R)n$, and:

$$\frac{d_{\min}}{n} < G_q^{-1}(R) \left[2 - \frac{G_q^{-1}(R)q}{q-1} \right] + \epsilon$$

The Gilbert and the Elias bounds are shown in Figure 3-3.

3.3.9 Error Correction Ability of Block Codes:

Concluding Summary

There are four algebraic techniques for error correction. Random error correction attempts to correct one or more erroneous symbols in a codeword. Random erasure correction attempts to correct erroneous symbols whose locations are known, but whose values are unknown. Burst error correction attempts to confine all the errors to a short span of the code and to correct the resulting burst. Finally, erasure burst correction tries to correct a burst whose location is known. The bounds derived earlier enable us to relate the error correction ability of a code to the code rate, R , for the four different techniques of error correction. The results

Figure 3-3 Asymptotic Forms of the Elias and the Gilbert-Varshamov Bounds

derived in this section will apply to bounded distance decoding: decoding in which the decoder attempts to correct all patterns up to a certain limit and none of greater value, even though the code might have the ability to do so. A bounded distance random error decoder will correct all pattern of t or fewer errors and none of greater weight, even though the code might be able to correct many patterns of $t+1$ or more errors. Similarly, a bounded distance random erasure decoder will only correct e or fewer erasures per block. Finally, in a bounded distance algorithm for burst correction, the decoder will correct all bursts of length b or less and none of greater length. Most decoding algorithms known, except for those based on exhaustive search methods, are of the bounded-distance type. A code with a minimum distance of d_{min} can correct all patterns of $d_{min}-1$ or fewer erasures. The bounds for random erasure correction are:

$$\frac{e}{n} \geq G_q^{-1}(R)$$

and

$$\begin{aligned} \frac{e}{n} &\leq G_q^{-1}(R) \left[2 - \frac{G_q^{-1}(R)}{q-1} q \right] - \frac{1}{n} \\ &\approx G_q^{-1}(R) \left[2 - \frac{G_q^{-1}(R)}{q-1} q \right] \end{aligned}$$

A code with a minimum distance of d_{min} can correct all patterns of $\lfloor (d_{min}-1/2) \rfloor$ errors or less. The bounds for random error correction are:

$$\frac{t}{n} \geq \left[n \frac{G^{-1}(R)}{2} \right] \times \frac{1}{n} = \frac{G^{-1}(R)}{2}$$

and

$$\begin{aligned} \frac{t}{n} &\leq \left[n \frac{G^{-1}(R)}{2} \left[2 - \frac{G^{-1}(R)}{q-1} q \right] \right] \times \frac{1}{n} - \frac{1}{2n} \\ &\approx \frac{G^{-1}(R)}{2} \left[2 - \frac{G^{-1}(R)}{q-1} q \right] \end{aligned}$$

The bounds for burst correction are given by:

$$\frac{b}{n} \leq (1-R) - \frac{1}{n} \log_q \left[\frac{(n-b)(q-1) + q}{q} \right]$$

for the burst correction ability in general, and by:

$$\frac{b}{n} \leq \frac{1-R}{2}$$

and,

$$\frac{b}{n} \geq \frac{1-R}{2} - \frac{1}{2n} \left[\log_q \left[\frac{(n-2b)(q-1) + q}{q} \right] \right]$$

for the burst correction of linear codes in particular.

The last of these bounds is asymptotically equal to:

$$\frac{b}{n} \geq \frac{(1-R)}{2}$$

For finite n the dependence on the logarithm of b can be eliminated by weakening the bound slightly:

$$\frac{b}{n} \geq \frac{(1-R)}{2} - \frac{1}{2n} \left[\log_q \left[\frac{n(q-1) + q}{q} \right] \right]$$

Finally, the following equality holds for erasure burst correction:

$$\frac{eb}{n} = 1-R$$

Let $P_e(\geq \frac{t}{n}, n)$, $P_x(\geq \frac{e}{n}, n)$, $P_b(\geq \frac{b}{n}, n)$, and $P_{eb}(\geq \frac{eb}{n}, n)$ represent the distributions of channel errors, channel erasures, burst lengths, and erasure burst lengths respectively. If an error correcting code of length n is used to correct channel errors, the decoded block error rate is bounded by:

$$\begin{aligned} \text{BLER} &\leq P_t \left(\geq \frac{q^{-1}(R)}{2} + \frac{1}{n}, n \right) \\ \text{BLER} &\geq P_t \left(\geq \frac{q^{-1}(R)}{2} \left[2 - \frac{q^{-1}(R)}{q-1} \right] + \frac{1}{2n}, n \right) \end{aligned}$$

If the code is used on an erasure channel, the bounds are:

$$\begin{aligned} \text{BLER} &\leq P_x \left(\geq q^{-1}(R) + 1/n, n \right) \\ \text{BLER} &\geq P_x \left(\geq q^{-1}(R) \left[2 - \frac{q^{-1}(R)}{q-1} \right], n \right) \end{aligned}$$

Finally, if the code is used to correct bursts, the bounds are :

For single burst correction:

$$\begin{aligned} \text{BLER} &\leq P_b \left(\geq \frac{1-R}{2} - \frac{1}{2n} \log_q (n(q-1)+q) + \frac{3}{2n}, n \right) \\ \text{BLER} &\geq P_b \left(\geq \frac{1-R}{2}, n \right) \end{aligned}$$

and for single erasure burst correction the following equality holds:

$$\text{BLER} = P_{eb} \left(\geq 1-R, n \right)$$

The expressions for the block error rate presented above are valid for a channel model in which the noise causes either erasures or errors to be made, but not

both. Similarly, the expressions are only valid when the decoder attempts to correct either random errors or bursts but not both. The generalization to combined erasures and errors decoding is straightforward but requires an examination of joint statistics such as

$P_{t,x}(\geq \frac{t}{n}, \geq \frac{e}{n}, n)$. Similarly, the generalization to

combined random error and burst correction is also straightforward but requires the examination of joint

statistics such as $P_{t,b}(\geq \frac{t}{n}, \geq \frac{b}{n}, n)$ and the

specification of a decoding algorithm. Other generalizations are possible too, but they are beyond the scope of the present research.

An incomplete decoder does not attempt to correct every received codeword. A bounded distance decoder in particular, does not correct any pattern of weight larger than t . When more than t errors occur, the decoder sometimes declares an incorrigible pattern or occasionally delivers the wrong message. Hence, the decoder output can be the correct message, an incorrect message, or an erased (a detected) message. It is in general difficult to compute the probabilities associated with these events, even for the simplest channels. For a linear code some simplification is possible, but even here an exact expression can only be obtained for a few known codes [24]. The weight distribution of the

codewords is required for an exact derivation, but the weight distribution for most codes remains unknown.

If some simplifying assumptions are made, approximate results can be obtained. For a linear code, the case conditioned on the all zero codeword being transmitted can be analyzed [24]. Every other codeword will have the same conditional probabilities and the results will therefore be the same as the unconditional probability. Figure 3-4 illustrates the 3 regions into which a received word can fall, assuming that the all zero codeword was transmitted.

Some channel error patterns of $t+1$ or more errors will throw the received pattern into one of the correction columns corresponding to a codeword which differs from the all zero vector, and some will throw the received pattern in the detection portion of the table of Figure 3-4. In the absence of any information on the codeword weight distribution, the best assumption that can be made is that the relative frequency of the two events is proportional to the size of their respective sample spaces.

The BLER derived earlier is the probability that $t+1$ or more channel errors occur. It can therefore be split into two parts:

$$BLER = P_{uD} + P_D = \left(\frac{P_{uD}}{BLER} + \frac{P_D}{BLER} \right)$$

where P_{uD} corresponds to the undetected errors and P_D to the detected ones.

The correctable portion of codeword zero has size

$$\sum_{i=0}^t \binom{n}{i} (q-1)^i$$

Thus the correction portion of all the other codewords

has size $(q^k - 1) \sum_{i=0}^t \binom{n}{i} (q-1)^i$. The size of the entire

array is q^n and the size of the entire array minus the

correctable portion is $q^n - \left(\sum_{i=0}^t \binom{n}{i} (q-1)^i \right)$. The size of

the detected portion is therefore:

$$q^n - \left(\sum_{i=0}^t \binom{n}{i} (q-1)^i \right) - ((q^k - 1) \sum_{i=0}^t \binom{n}{i} (q-1)^i)$$

and the ratios are:

$$\frac{P_D}{BLER} = \frac{q^n - q^k \sum_{i=0}^t \binom{n}{i} (q-1)^i}{q^n - \sum_{i=0}^t \binom{n}{i} (q-1)^i}$$

$$\frac{P_{uD}}{BLER} = \frac{(q^k - 1) \sum_{i=0}^t \binom{n}{i} (q-1)^i}{q^n - \sum_{i=0}^t \binom{n}{i} (q-1)^i}$$

because $\sum_{i=0}^t \binom{n}{i} (q-1)^i$ is small in comparison to q^n ,

the above simplifies to:

$$\frac{P_D}{BLER} \approx 1 - q^{-(n-k)} \sum_{i=0}^t \binom{n}{i} (q-1)^i$$

$$\frac{P_{uD}}{BLER} \approx q^{-(n-k)} \sum_{i=0}^t \binom{n}{i} (q-1)^i$$

Depending on the strategy used at the receiver, the detected channel errors may or may not contribute to the overall bit error rate. If the detected errors do not contribute to the overall BER, it is given by:

$$BER = P_{UD} \times \frac{\text{average number of erroneous bits per word}}{\text{total number of bits per word}}$$

But the average number of erroneous bits is given by:

$$\text{average no. of erroneous bits} = \frac{1}{\sum_{i=t}^n (p(i))} \times$$

$$\{(\text{avg. no. } |t+1 \text{ errors}) p(t+1) + (\text{avg. no. } |t+2 \text{ errors}) p(t+2) + \dots \\ + (\text{avg. no. } |t+(n-t) \text{ errors}) p(t+n-t)\}$$

Again, this value is difficult to compute because it requires the channel probabilities as well as the weight distribution of the codewords for a complete solution. A reasonable estimate is given by the following argument. For a code with minimum distance $d=2t+1$ that is used to correct all patterns of t errors or less, the most likely undetected error will be when the error pattern contains exactly $t+1$ errors. The decoder will erroneously assume that the received sequence contains t errors and will change these so that the sequence has exactly $t+1+t=d$ errors. This process is illustrated in Figure 3-5. The BER rate is given by:

$$BER = P_{UD} \times \frac{d}{n}$$

Figure 3-4 Standard Array (Decoding Table) for a Binary Code

$t+1$ errors occur in a codeword of a code designed to correct t or fewer errors

decoder wrongly assumes that t (or fewer) errors must have occurred elsewhere

by correcting t wrongly assumed locations, the decoder adds t errors to the word delivered to the sink
word now contains $t+1+t = d$ errors

Figure 3-5 A Decoding Error

in all cases this is less than:

$$\text{BER} = \frac{d}{n} \text{BLER}$$

(and sometimes orders of magnitude less).

The last equation can also be applied when the detected words are passed to the decoder unchanged.

For any scheme involving erasures, the BER can be made zero because the patterns that exceed the correction limit of the decoder can be recognized as such. There is, however, a nonzero probability of erasure (or deletion) of the messages.

The only bound for the bit error rate after burst error correction is:

$$\text{BER} = \frac{1}{2} \text{BLER}$$

and in most cases this bound is too weak to be of any value.

In conclusion, in all cases, better estimates on performance can be obtained from the block error rate.

3.4 Experimental and Theoretical Results

When the bounds on the error correction ability of a code are combined with the empirical error distributions of Chapter 2, the graphs of Figures 3-6 to 3-15 are obtained. The two performance measures, the BLER and the BER are plotted as a function of the code rate for different blocklengths. The graphs clearly reveal the

importance of using long block codes to achieve performance advantages. In many cases the values suggested by the graphs are too large to be of practical utility, because system wide constraints limit the tolerable transmission delay or the computational complexity of the decoder. As an example, to achieve a probability of error of 10^{-6} in the downtown area while maintaining a moderately high information rate, a code of length 4095 and minimum distance of 143 might be chosen (Figures 3-20 and 3-21). The rate of such a code would be constrained by the Gilbert and Elias bounds to lie in the range 0.78 - 0.87. The minimum delay, for a bit rate of 4800 bits per second, would be 0.86 seconds, and the decoder would be required to correct all patterns of 71 or fewer errors.

The bounds for bursts error correction reveal that burst correcting codes are a very poor choice for the mobile channel regardless of their lengths. The performance of these codes degrades as the blocklength is increased. This is in stark contrast to the random error correction case where a given error rate can be sustained with progressively higher code rates (up to a limiting value) provided the blocklength is also increased.

The bounds can also be combined with the expression for the block error distribution given by equation (73) of Chapter 2. This yields the graphs of Figures 3-16 through 3-23. The agreement between the two sets of graphs, the ones derived from the empirical distributions

and the theoretical ones, is fair to good. The agreement is better in the downtown core, and worsens slightly for the recordings done in Hull. This is due in part to the erratic velocity profile of the mobile in Hull and the consequent difficulty in estimating the speed. The theoretical distributions usually underestimate the actual values, sometimes by up to an order of magnitude, but usually by much less.

The theoretical bounds on performance are obtained in the following fashion. Equation (73) can be used to relate the BLER to the number of errors corrected by the code. The coding bounds can then be used to relate the latter to the code rate.

If the code corrects all the patterns of m or fewer errors the BLER is given by:

$$\text{BLER} = F(m+1)$$

where $F(m)$ is the probability of the number of errors being equal to or greater than m . But this may be rewritten as:

$$\text{BLER} = F\left(n \frac{(m+1)}{n}\right) = F\left(n \frac{G_2^{-1}(R)}{2} + 1\right)$$

for the Gilbert bound, and:

$$\text{BLER} = F(n(G_2^{-1}(R)) (1-G_2^{-1}(R)))$$

for the Elias bound.

In the case of the Gilbert bound for example the relationship becomes:

$$\sum_{k=1}^{\infty} \sum_{j=0}^{k-1} \exp \left\{ - \frac{\sqrt{2\pi} (1-2p_e) (\sqrt{-\ln(1-2p_e)})^j}{\lambda R_I} \left[n + \frac{nG_2^{-1}(R) + 2}{2p_e} \right] \right\}$$

$$\times \left[\frac{\sqrt{2\pi} (1-2p_e) (\sqrt{-\ln(1-2p_e)})^k}{\lambda R_I} \right] \frac{v_n}{k!}$$

$$\times \left[\frac{2\sqrt{2\pi} (1-2p_e) (\sqrt{-\ln(1-2p_e)})^j}{2p_e \lambda R_I} \left\{ \frac{G_2^{-1}(R)}{2} + 1 \right\} \right] \frac{v_n}{j!}$$

and for the Elias bound the relation is:

$$\sum_{k=1}^{\infty} \sum_{j=0}^{k-1} \exp \left\{ - \frac{\sqrt{2\pi} (1-2p_e) (\sqrt{-\ln(1-2p_e)})^j}{\lambda R_I} \left[n + \frac{n(G_2^{-1}(R) (1-G_2^{-1}(R)))}{p_e} \right] \right\}$$

$$\times \left[\frac{\sqrt{2\pi} (1-2p_e) (\sqrt{-\ln(1-2p_e)})^k}{\lambda R_I} \right] \frac{v_n}{k!}$$

$$\times \left[\frac{2\sqrt{2\pi} (1-2p_e) (\sqrt{-\ln(1-2p_e)})^j}{2p_e \lambda R_I} \{ nG_2^{-1}(R) (1-G_2^{-1}(R)) \} \right] \frac{v_n}{j!}$$

These expressions can be evaluated numerically.

Figure 3-6. Block Error Rate versus Code Rate. (Experimental)
Random Error Correction; Lower Bound (Downtown)

Figure 3-7. Block Error Rate versus Code Rate. (Experimental)
Random Error Correction; Upper Bound (Downtown)

Figure 3-8 Block Error Rate versus
Code Rate. (Experimental)
Random Error Correction; Lower Bound
(Hull)

Figure 3-9 Block Error Rate versus
Code Rate. (Experimental)
Random Error Correction; Upper Bound
(Hull)

Figure 3-10 Bit Error Rate versus Code Rate (Experimental)
Random Error Correction; Lower Bound
(Downtown).

Figure 3-11 Bit Error Rate versus Code Rate (Experimental)
Random Error Correction; Upper Bound
(Downtown)

Figure 3-12 Bit Error Rate versus Code Rate (Experimental)
Random Error Correction; Lower Bound
(Hull)

Figure 3-13 Bit Error Rate versus Code Rate (Experimental)
Random Error Correction; Upper Bound
(Hull)

Figure 3-14 Block Error Rate versus
Code Rate
Burst Error Correction
(Downtown)

Figure 3-15 Block Error Rate versus
Code Rate
Burst Error Correction
(Hull)

Figure 3-16 Block Error Rate versus Code Rate. (Theoretical)
Random Error Correction; Lower Bound (Downtown)

Figure 3-17 Block Error Rate versus Code Rate. (Theoretical)
Random Error Correction; Upper Bound (Downtown)

Figure 3-18 Block Error Rate versus
Code Rate. (Theoretical)
Random Error Correction; Lower Bound
(Hull)

Figure 3-19 Block Error Rate versus
Code Rate. (Theoretical)
Random Error Correction; Upper Bound
(Hull)

Figure 3-20 Bit Error Rate versus Code Rate. (Theoretical)
Random Error Correction; Lower Bound

Figure 3-21 Bit Error Rate versus Code Rate. (Theoretical)
Random Error Correction; Upper Bound
(Downtown)

Figure 3-22 Bit Error Rate versus Code Rate. (Theoretical)
Random Error Correction; Lower Bound
(Hull)

Figure 3-23 Bit Error Rate versus Code Rate. (Theoretical)
Random Error Correction; Upper Bound
(Hull)

CHAPTER 4

PRACTICAL CODING SCHEMES.

4.1 Introduction

To test the validity of the work done in the previous chapters, practical schemes were simulated using the error pattern files as a basis. The basic assumption in the simulations is as follows. The effectiveness of a certain code on the mobile channel is being verified. In most cases the code, or coding scheme, C , is linear. A certain codeword Q from C , corresponding to a message W , is transmitted. The effect of the channel-demodulator pair is to introduce errors randomly into the codeword. At the receiver, the decoder is given the vector $Q+Y$, where Y is an error pattern that has been added to Q on a symbol by symbol basis. The decoder's task is to apply a transformation, F , which recovers the transmitted message W , from the received word. The transformation is assumed to be linear. Applying it to $(Q+Y)$ yields $F(Q+Y) = F(Q)+F(Y) = W+F(Y)$. Hence, the effect of the error pattern is the same regardless of which codeword is transmitted. Applying the transformation to $F(Q+Q) = F(Q) = F(Q)+F(Q)$ shows that $F(Q) = 0$. In the simulations it is therefore convenient to assume that the transmitted codeword is the all-zero codeword, in which case decoding errors are made

only when $F(X) \neq 0$. Hence, all useful information about the performance of the code can be derived by examining the effect the transformation F has on the received error pattern, as given by the error pattern files.

In this chapter, two practical schemes are tested. The first scheme is the protocol used in the AMPS system for transmitting control messages. The second is a simpler scheme based on Hamming codes.

4.2 Code Performance Measures

There are three possible outcomes to any decoding operation. Correct decoding occurs when the word delivered to the information sink is the same as the transmitted word. Erroneous decoding arises when the word delivered to the sink is different from the transmitted word. Lastly, detection of errors occurs when the decoder cannot correct the received word and delivers no information to the sink. The probabilities P_c , P_f and P_d associated with these three outcomes depends on the way in which the code is used. Equations for P_c , P_f and P_d that can be used in conjunction with the block error rate distributions for the channel, $p(\geq m, n)$, are given in what follows. In the notation adopted, n is the block length of the code and k and r denote the number of information and parity bits in a codeword respectively. Finally, the minimum distance is denoted $d_{\min}$. This work is based in part on [7].

If the code is used for error detection only, the relevant equations are:

$$P_c = 1 - P(\geq 1, n)$$

$$P_f = 2^{-r} \times P(\geq d_{\min}, n)$$

$$P_d = 1 - P_c - P_f$$

P_c is identical to the probability that the block is received error free. P_f is derived as follows: decoding errors occur only if the received word is a codeword but is different from the transmitted codeword. For such an event to occur, the channel error pattern must have weight at least $d_{\min}$, because no error pattern of lesser weight can change one codeword into another. Secondly, only $(2^k - 1) / (2^n - 1) = 2^{-r}$ of the possible error patterns will specifically change the transmitted word into another codeword. This is only an approximate result.

If the code is used for random error correction to correct all the patterns of t or fewer errors, but none of greater weight, ($t \leq [(d_{\min} - 1) / 2]$), then (this is known as bounded distance decoding):

$$P_c = 1 - P(\geq t+1, n)$$

$$P_f = P(\geq d_{\min} - t, n) \times 2^{-r} \sum_{j=0}^t \binom{n}{j}$$

$$P_d = 1 - P_c - P_f$$

Before deriving the equations above, an example illustrating the occurrence of a decoding error will be given. The example is intended to clarify the nature of the equations.

Suppose 1011001 and 1000110 are valid codewords in a given code. They would appear on the first line of a standard array for this code:

```
0000000  1011001  ...  1000110  ...
```

On the second line, any minimum weight vector which is not a codeword would be chosen as a coset leader. Let 0000001 be such a vector. The standard array now has the form:

```
0000000  1011001  ...  1000110  ...
0000001  1011000  ...  1000111  ...
```

Let the code be used to correct one error. If the transmitted word is 1011001 and if it is corrupted by the channel error pattern 0011110, the received word becomes 1000111. This is a vector in the first coset, and a maximum likelihood decoding procedure is to find the coset leader (namely 0000001) and add it to the received word. In this case this yields: $0000001 + 1000111 = 1000110$. The transmitted word is 1011001 but the word delivered to the sink is 1000110: a decoding error has been made.

More formally, let an (n, k) code be used to correct all patterns of t or fewer errors. Out of the 2^n possible received words the fraction which is corrected by the decoder is:

$$\sum_{j=0}^t \binom{n}{j} / 2^n$$

For a decoding error, to occur two conditions must be satisfied. Firstly, more than $d_{\min}-t$ errors must occur: this insures that the transmitted codeword and the received word do not lie in the same column of the standard array. If this is the case, any attempt to correct the received word will fail. Secondly, the received word must be one of the words which the decoder will attempt to correct.

There are 2^k codewords and hence the probability of making a decoding error P_f , is given by:

$$P_f = 2^k \times \left[\sum_{j=0}^n 2^{-n} \binom{n}{j} \right] \times P(\geq d_{\min}-t, n)$$

Here $P(\geq d_{\min}-t, n)$ is the probability that the received word has $d_{\min}-t$ or more errors in it.

A more exact expression for P_f , based on the codeword weight distribution, can be used. However, such an expression is cumbersome and there are large classes of codes for which such expressions have not yet been found. It is generally assumed that the simple expression used above gives an accurate estimate of the false rate [7].

Secondly, the probability of correct decoding P_c , is given by:

$$P_c = 1 - P(\geq t+1, n)$$

Here $P(\geq t+1, n)$ is the probability that the received word has $t+1$ or more errors in it.

Finally, the probability of detecting an erroneous codeword P_d , is given by $P_d = 1 - P_f - P_c$.

For a burst trapping decoder in which all bursts of length b or less ($b \leq r/2$) are corrected, and the bursts of length l ($b < l \leq r$) are detected, and the bursts longer than r bits cause errors, the equations are:

$$P_c = 1 - P_b(\lambda = b+1, n)$$

$$P_b(\lambda \geq r+1, n) < P_f < P_b(\lambda = b+1, n)$$

$$P_d = 1 - P_c - P_f$$

Once again, bounds on P_f instead of exact expressions are used. Most good burst correcting codes are cyclic, and many can simultaneously correct bursts whose lengths do not exceed $r/2$ bits and detect bursts whose lengths do not exceed r bits. In this sense, the left inequality, $P_b(\lambda \geq r+1, n) < P_f$, is usually a very good indication of the falsing rate.

The equations can be used with the graphs of block error distributions, $p(\lambda = m, n)$, or burst error distributions, $p(\lambda = b, n)$, to derive results on the performance of different codes.

4.3 The AMPS Coding Scheme

The Bell Labs (currently Bell Communications Research or Bellcore) have developed a system for voice transmission over mobile channels which is meant to accommodate a large number of users and to provide them with telephone quality service. This system is known as the Advanced Mobile Phone Service (AMPS) system.

Although the system is designed for analog voice transmission, extensive control messages must also be transmitted between mobile receivers and a central base-station. In fact, because of the volume of data transmitted, some channels are dedicated solely for the purpose of sending and receiving digital data..

The protocol used by the AMPS system is based on the type of messages to be transmitted (short messages rarely exceeding 200 bits) and it is not necessarily a suitable choice when other types of digital data must be transmitted (computer databases and digital voice are two examples).

In this section, outlines of the coding scheme and of the data formats used in the AMPS system are given. Using the error pattern files, the performance of this scheme is investigated when two different decoding rules are used at the receiver. It is demonstrated that an alternate decoding scheme can increase throughput by 67 percent while maintaining similar performance requirements.

4.3.1 Outline of the AMPS Data Formats.

There are four channels over which data are transmitted in the AMPS system. These are:

- The forward set-up channel
- The forward voice channel
- The reverse set-up channel

- The reverse voice channel.

The set-up channels are dedicated solely for the purpose of data transmission. In the voice channels on the other hand, short digital packets are transmitted between talkspurts using a technique known as "blank-and-burst". The forward channels are used for transmission from base-station to mobile and the reverse channels are used for transmission from mobile to base-station.

In all four cases, the data is Manchester encoded and transmitted at 10 kbps.

To combat the effects of fading and other channel impairments, a combination of 3 out-of 5 majority voting, BCH encoding, and word interleaving is used.*

The BCH codes used are all derived from a basic (63,51) BCH code, which has a minimum distance of 5. These codes can be used either to:

- correct 2 errors, or
- correct 1 error and detect all patterns of 3 errors or less, or
- detect up to four errors.

To keep the receiver design fairly simple and to ensure low falsing rates, (10^{-7} for any given message on the set-up channels and 10^{-5} for messages sent over the voice channels), it was decided that the best compromise was to adopt the second alternative, namely, to correct 1 error and detect two more.

* Additionally diversity reception (switched at the mobile and equal gain at the base-station) is used, but this is not in the realm of coding.

The particular data formats used, although they share the same general features, vary somewhat from channel to channel. Below is a separate description of each. The descriptions are not complete but list the features that are relevant to the processing that was done on the data.

(1) Forward Set Up

- words are 40 bits long (shortened BCH (40,28) code)
- each word is repeated five times
- each repeat is separated from the next one by one word length
- busy-idle bits are inserted after every ten bits. These busy-idle bits do not form part of the codeword and they are treated separately at the receiver.

(2) Forward Voice

- words are 40 bits long (shortened BCH (40,28) code)
- each word is repeated five times
- each codeword is separated from the next by 48 bits of bit and word sync.

Words are repeated eleven times in some cases (e.g., for handoff where the mean SNR is low), but this is done to ensure that at least five of the eleven repeats are "captured". Only five are used by the receiver.

(3) Reverse Set-Up

- words are 48 bits long (shortened BCH (48,36) code)
- each codeword is repeated five times
- codewords are not separated from each other (no word interleaving).

(4) Reverse Voice

- words are 48 bits long (shortened BCH (48,36) code).
- each codeword is repeated five times.
- each codeword is separated from the next by 48 bits of bit and word sync.

Additionally, a fifth channel, which is not part of the AMPS system, has been tested. This channel can best be described as a forward set-up channel without interleaving. That is, the format is identical to the one used on the forward set-up channel, but the repeats come immediately after one another. This channel is referred to as the forward non-interleaved channel.

The performance requirements are comparable to those found in land-line telephone systems:

- (1) The miss rate should be in the range of 10^{-3} at a signal to interference ratio $S/I=15$ dB. Averaged over the entire service area this implies a miss rate of about 10^{-4} .

(2) The falsing rate (probability of erroneous decoding) should be less than 10^{-7} for any given message. For voice channels this requirement is relaxed to 10^{-5} .

4.3.2 Outline of the Simulations

Two data rates were used in the simulations. The first data rate was 9.6 kbps, which corresponded closely to the 10 kbps data rate in the AMPS system. The second data rate was 4.8 kbps. In both cases the experimental data for transmissions at 850 MHz and 4800 bps was used.

The error pattern files were altered to reflect the channel characteristics for transmission rates of 10 kbps. Because operation is well above the reciprocal of the average fade duration, the average probability of error is not expected to be substantially different at 4.8 kbps and at 10 kbps. What is expected however, is that at 10 kbps a fade will affect roughly twice as many bits as at 4.8 kbps. Based on these two observations the scheme adopted was to expand all gaps and all bursts to twice their sizes. The bits added in between bits belonging in a gap were all zeros, but the bits inserted in between the original bits of a burst were randomly assigned the values 0 and 1 with probability one half.

* A burst is defined as beginning and ending with a one and not containing more than seven zeros in succession. A single error is a trivial burst of length one.

As an example, an initial error pattern, reflecting the channel at 4.8 kbps, and an expanded error pattern, reflecting conditions at 9.6 kbps are given below:

```

-gap---|-----burst-----|-----gap-----
0 0 0 0 1 0 0 1 1 0 0 0 0 1 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
00000000011000110100101000110011010000111000000000000000000000

```

In the simulations, sets of 500 bit samples are chosen randomly, and are rearranged to match the data format on the selected channel. For example, busy-idle bits are removed, bit and byte sync bits are removed, and so on. At this stage, five words have been assembled and are ready for use by the decoder. The decoding process is illustrated in Figure 4-1, where a hypothetical hardware decoder for a systematic code is described. The first step is to feed five repeats of any one word to the majority voting circuit (MVC). Each bit in the output word is a majority decision on the five input bits. The voted output is then fed to a BCH decoder which corrects one bit and detects all patterns of 3 bit errors or less. The decoder block diagram is illustrated in Figure 4-1. Any error-pattern with Hamming weight strictly less than 4 will either be detected or corrected. Any error-pattern of weight 4 or more can potentially cause a decoding error. Whether this occurs or not depends on

Figure 4-1 Conceptual Schematic of the AMPS Decoder

the particular error-pattern. For any code, a list of the error-patterns which cause decoding errors can be drawn. Alternatively, an estimate of the proportion of error-patterns which cause decoding errors can be established. For a code this size, exhaustive analysis, while still possible, is not an attractive alternative. Instead an estimate of the number of decoding errors is preferred. This estimate is derived in the next section.

The analysis can therefore be based on the weight of the error-pattern word fed to the BCH decoder. If the word has weight less than or equal to one the output is error-free. If the word has weight between two and three, there will be no output: the word is detected in error. If the word has weight four or more then there exists the possibility of a decoding failure, i.e., the decoder might deliver false information to the receiver.

4.3.3 Equations Used

Equations are needed to relate the simulation results to the miss and false rates.

The miss rate is defined as the probability that the decoder will produce no output and the false rate is the probability that it will produce wrong output.

From the experimental results, the following data can be obtained: the number of words at the output of the MVC that are correctly received N_0 , those that contain one error N_1 , two or three errors N_D , and four or more

errors N_F . Finally, the total number of words is denoted N_T .

From the equations derived above and the results that are obtained experimentally, the false and miss rates are given by:

$$\begin{aligned}
 P_{\text{false}} &= (N_F/N_T) \times \left[2^{-n} \times 2^k \sum_{j=0}^t \binom{n}{j} \right] \\
 &= (N_F/N_T) \times \left[2^{-40} \times 2^{28} \sum_{j=0}^1 \binom{40}{j} \right] \quad \text{for the forward channels} \\
 &= (N_F/N_T) \times \left[2^{-48} \times 2^{36} \sum_{j=0}^1 \binom{48}{j} \right] \quad \text{for the reverse channels}
 \end{aligned}$$

$$P_{\text{miss}} = (N_D/N_T) + [(N_F/N_T) \times (1 - 2^{-n} \sum_{j=0}^t \binom{n}{j})] \times 2^k$$

or alternatively:

$$P_{\text{miss}} = \frac{[N_D + N_1 + N_F]}{N_T} - P_{\text{false}}$$

If P_{false} is much smaller than P_{miss} , then P_{false} can be neglected in the second equation.

Numerically, the factors

$$2^{-40} \times 2^{28} \times \left[\sum_{j=0}^1 \binom{40}{j} \right] \text{ and } 2^{-48} \times 2^{36} \times \left[\sum_{j=0}^1 \binom{48}{j} \right]$$

are equal to 10^{-2} and 1.2×10^{-2} respectively.

4.3.4 Results

The simulation results are given in Figures 4-2 through 4-9. Some conclusions can be readily inferred by examining these tables.

The BCH decoder corrects very few errors over and

above those corrected by the MVC. As an explanation it is suggested that if the output of the MVC is not error-free it is more likely to contain a burst of errors rather than a single error. Therefore attempting to correct single errors at the output of the MVC, as is done in the AMPS decoder, will not result in many words being corrected. Of course the main purpose of the BCH decoder in this scheme is to detect a large fraction of the uncorrected words at the output of the MVC, in order to reduce the false rate to tolerable values. Hence, it might be more useful to use the BCH decoder to detect all patterns of 4 or fewer errors (it presently corrects all single errors and detects two more errors). This will reduce the false rate even further but the word error rate will remain essentially unchanged.

Simple word interleaving improves the false and miss rates, although not by much. More packets are received error free when no interleaving is used but the errors are more correlated and the MVC cannot correct as many codewords as before. The improvements when the interleaving degree (number of codewords interleaved) is increased to 3, 4, ... should be investigated.

The scheme performs well, and its performance remains reasonable even as the bit error rate increases substantially. This is in contrast to simpler schemes where performance degrades very rapidly as the BER increases. All the added redundancy in this scheme insures that it is very robust.

location	Downtown
channel	Forward Set-Up
percent of messages received error free	93.17
percent of messages correct after majority voting	99.978
percent of messages correct after majority voting and BCH error correction	99.983
miss rate	1.7×10^{-4}
false rate	5.5×10^{-7}

Figure 4-2 AMPS Simulation results at 9.6 kbps

location	Downtown
channel	Forward Non-Interleaved
percent of messages received error free	94.90
percent of messages correct after majority voting	99.922
percent of messages correct after majority voting and BCH error correction	99.950
miss rate	5.0×10^{-4}
false rate	3.0×10^{-6}

Figure 4-3 AMPS Simulation Results at 9.6 kbps

location	Downtown
channel	Forward Voice
percent of messages received error free	93.67
percent of messages correct after majority voting	99.978
percent of messages correct after majority voting and BCH error correction	99.983
miss rate	1.7×10^{-4}
false rate	5.5×10^{-7}

Figure 4-4 AMPS Simulation Results at 9.6 kbps

location	Hull
channel	Forward Set-Up
percent of messages received error free	84.57
percent of messages correct after majority voting	99.467
percent of messages correct after majority voting and BCH error correction	99.802
miss rate	2.0×10^{-3}
false rate	1.2×10^{-5}

Figure 4-5 AMPS Simulation Results at 9.6 kbps

location	Hull
channel	Forward Non-Interleaved
percent of messages received error free	88.89
percent of messages correct after majority voting	98.922
percent of messages correct after majority voting and BCH error correction	99.222
miss rate	7.8×10^{-3}
false rate	3.8×10^{-5}

Figure 4-6 AMPS Simulation Results at 9.6 kbps

location	Hull
channel	Forward Voice
percent of messages received error free	85.47
percent of messages correct after majority voting	99.567
percent of messages correct after majority voting and BCH error correction	99.667
miss rate	3.3×10^{-3}
false rate	1.9×10^{-5}

Figure 4-7 AMPS Simulation Results at 9.6 kbps

location	Queensway
channel	Forward Set-Up
percent of messages received error free	45.36
percent of messages correct after majority voting	87.173
percent of messages correct after majority voting and BCH error correction	90.400
miss rate	9.6×10^{-2}
false rate	6.1×10^{-4}

Figure 4-8 AMPS Simulation Results at 9.6 kbps

location	Downtown
channel	All Forward Channels
percent of messages received error free	95.8
percent of messages correct after majority voting	100.0
percent of messages correct after majority voting and BCH error correction	100.0
miss rate	$< 6 \times 10^{-5}$
false rate	$< 1 \times 10^{-8}$

Figure 4-9 AMPS Simulation Results at 4.8 kbps

4.3.5 An Alternate Decoding Scheme

An alternate decoding scheme based on erasure correction is proposed. The demodulator divides the received signal space into three decision regions: 0, 1, and x. The symbol x stands for an erasure, and is output whenever the demodulator cannot make an unambiguous decision on the received bit. The channel becomes a two input, three output channel as illustrated in Figure 4-10. In the figure, the transition probabilities are taken to be time varying. In Figure 4-11, the discrete channel is related to the on-off model of Chapter 2.

From the block diagram of the AMPS receiver shown in Figure 4-12, it is seen that an indication of the instantaneous signal strength at the receiver antennas is available. Diode detectors stationed at the outputs of successive stages of progressively saturating differential amplifiers (at an IF frequency of 10.7 MHz) provide a log output voltage which rises smoothly and monotonically as the input signal ranges from -110 to -30 dBm. This voltage is known as the received signal strength indicator (RSSI) and is available in analog and in digital form--via an 8 bit A/D--to the logic unit which controls the operation of the AMPS receiver. The RSSI is needed to control the gain of the input amplifiers and to control the switched diversity reception circuits.

Figure 4-10 Time Varying Erasure and Error Channel

Figure 4-11 Modified on-off Erasures and Errors Channel

Figure 4-12a - Mobile transceiver block diagram

Figure 4-12b - Functional block diagram of Bell Laboratories-designed mobile logic unit

after [3]

The RSSI can also provide the decoder with side information about the state of the channel. Several schemes which make use of this side information can be devised. A particularly simple scheme which could be readily implemented, even with the present decoder, would be to have the demodulator decode into erasure symbols whenever the RSSI signal fell below a preset threshold. Because most of the erroneous decisions at 850 MHz are due to the fading process, which causes the signal to drop to levels at which it is submerged by front-end noise, the output data stream will contain very few errors, but many erasures. The few errors that would be present would typically be caused by AWGN, and other less important effects.

The results of Chapter 3 indicate that increased performance is possible with erasure decoding, but the incentive for using this alternative decoding technique does not come from a desire to increase performance: the requirements are fixed and do not have to be any better than those found in land-line telephony systems. Instead, it is of interest to examine whether the code rate can be increased with the alternative technique while maintaining the performance requirements. An increase in code rate would imply a more efficient utilization of the set-up channels. This is an important issue because of the set-up channels represent a considerable expense by subtracting from the total reserve of channels available for voice transmission.

4.3.6 Outline of the Simulations

To test the hypothesis that the alternative scheme can be as effective while affording a higher information transmission rate the following modifications are made:

The decoder now consists of a 2 out of 3 majority logic unit followed by a BCH error correction/detection unit. The decoding algorithm is described in the flow chart of Figure 4-13. Briefly, the algorithm operates as follows. For simplicity, a mapping has been defined between the three symbols, 0, x, and 1 and the three numbers -1, 0 and 1. This simplifies the task of the majority voter. The three outcomes are summed and the result is compared to zero. If it is greater than zero then the received bit is taken to be a one, if it is less than zero the received bit is taken to be zero. Otherwise, the received bit which is passed along to the BCH decoder is an erasure. The BCH decoder accepts the word from the majority unit and counts the number of erasures. If there are more than 2 erasures the word is marked as undeliverable. Otherwise, the decoder forms the syndrome, leaving the positions where there are erasures unfilled. The decoding proceeds as explained in the flowchart of Figure 4-13.

The BCH unit can correct either 1 error or 2 erasures (but not both). It can also detect 2 more errors and any number of erasures.

Figure 4-13. Erasure/Error Correction Algorithm
(Continued on next page)

Figure 4-13 Erasure/Error Correction Algorithm for BCH(40,28) Code (continued)

+ corrigible error pattern: 0 or 1 errors

4.3.7 Simulation Results

The error-pattern files contain no indication about the RSSI. It is assumed that bursts of errors occur when the signal fades below some threshold level.. In Chapter 2, it was seen that this assumption gave reasonable agreement with theory. Thus the simulations are performed according to the following rules.

(1) All of the bits in bursts of length 2 or more are replaced by erasures.

(2) All of the bits in bursts of length 5 or more are replaced by erasures.

This second value might be a bit too large but if the results are favorable in this case then the erasure method will be a very promising alternative.

The results are tabulated in Figures 4-14 to 4-17.

The results show that the alternate decoding method is effective in combating channel errors. The miss rates are comparable to those obtained with the scheme currently implemented, and the false rates are considerably better with the alternate scheme because the BCH circuit can now detect any number of erasures.

Throughput with the alternate scheme is increased by about $(5-3)/3$ or 67%. Presented differently, for a given number of customers this scheme requires that fully one third less channels be used for set-up and access requirements.

location	downtown
channel	Forward Set-Up
percent of messages received error free	95.79
percent of messages correct after majority voting	99.989
percent of messages correct after majority voting and BCH error correction	99.989
miss rate	1.1×10^{-4}
false rate	$< 1 \times 10^{-9}$

Figure 4-14
Alternate decoding scheme at 9.6 kbps
(all bursts of 2 or more bits are decoded as erasures)

location	Hull
channel	Forward Set-Up
percent of messages received error free	90.13
percent of messages correct after majority voting	99.739
percent of messages correct after majority voting and BCH error correction	99.878
miss rate	1.2×10^{-3}
false rate	$< 1 \times 10^{-9}$

Figure 4-15
Alternate decoding scheme at 9.6 kbps
(all bursts of 2 or more bits are decoded as erasures)

location	Downtown
channel	Forward Set-Up
percent of messages received error free	95.79
percent of messages correct after majority voting	99.972
percent of messages correct after majority voting and BCH error correction	99.983
miss rate	1.7×10^{-4}
false rate	5.5×10^{-7}

Figure 4-16

Alternate decoding scheme at 9.6 kbps
(all bursts of 5 or more bits are decoded as erasures)

location	Hull
channel	Forward Set-Up
percent of messages received error free	90.13
percent of messages correct after majority voting	99.283
percent of messages correct after majority voting and BCH error correction	99.573
miss rate	94.2×10^{-3}
false rate	1.1×10^{-6}

Figure 4-17

Alternate decoding scheme at 9.6 kbps
(all bursts of 5 or more bits are decoded as erasures)

4.4 Packet Error Rates With Hamming Codes

In this section the performance of the (7,4) simple Hamming code and of the (8,4) extended Hamming code with and without interleaving is examined. As compared to the AMPS scheme this is one of the simpler approaches that can be used to transmit data.

The simple Hamming code can be used to correct one error or alternatively it can be used to detect all patterns of 2 errors and some patterns of 3 errors. The extended Hamming code can be used to correct all single errors and simultaneously detect all patterns of 2 errors and many more. An exhaustive list of detectable patterns is possible because the code is small.

To obtain the results, the error pattern files were used. (They were randomly sampled and each time as many consecutive bits as were required to form a packet were taken. Bit interleaving was then introduced under the constraint that the interleaving period (the total number of bits involved, i.e., ni where n is the codeword length and i the number of codewords that are interleaved) would equal the packet length.

While this procedure is applied, statistics are kept to produce the packet, word and bit error rates before and after decoding. Statistics to determine the false rate are also kept.

The decoder attempted to correct the codeword based on the computed syndrome. This meant that the bit error rate was sometimes higher after the decoding process.

This occurred when no interleaving was present. A (7,4) or an (8,4) Hamming decoder will in fact add an error when attempting to decode an error pattern of weight three. Similarly, other error patterns might also cause the decoder to add errors. The word and packet error rates were, however, always lower.

The performance of the Hamming and extended Hamming codes were evaluated using the error files, and the results are shown in Figures 4-18 to 4-20. In Figure 4-18, the averaged decoded BER for simple and extended codes are calculated, and are plotted against the interleaving interval. The curves show that the average error rates decrease monotonically with an increased interleaving interval, until they level off at a certain irreducible value indicating that the channel has been transformed into a random error channel. The interleaving period that results in a saturated level is comparable to the average no-fade interval. The results indicate that this interval is between 500 and 700 bits. The BER goes from an average value of 2.33×10^{-3} with no coding, down to about 2×10^{-4} for the simple Hamming code and above 10^{-5} in the case of the extended Hamming code.

Figure 4-19 illustrates the packet miss rate against packet length results for the simple and extended Hamming codes, with and without bit interleaving. The miss rate represents the percentage of packets that are not successful. This includes packets wrongly accepted as

Figure 4-18 Influence of interleaving period on BER

- * Average BER without coding = 2.33×10^{-3}
- a Simple code
- b Extended code

Figure 4-19 Packet miss rate

- a No coding
- b Extended (no interleaving)
- c Simple (no interleaving)
- d Extended
- * Simple (used for error correction)

Figure 4-20 Packet false alarm rate in error-detection mode

- a Simple (for error correction without interleaving)
- b Extended (no interleaving)
- c Simple (with interleaving)
- d Extended code

(after [30])

good, as well as those which are rejected for having incorrigible errors. It is clear that without bit interleaving, the improvement over the packet error rate with no coding is very small. In light of what was presented in Chapter 3, this is an expected result. With bit interleaving, there is an order of magnitude improvement in the miss rate for packets of the order of 200 bits or longer. Moreover, the packet error rate curves have wide minima near the 200-bit packet length region. This can be explained by noting that, bit interleaving within short packets is not effective, while very long packets invite the possibility of hitting many fade slots. Figure 4-20 shows the false alarm rate against packet length results for the four cases of coding and interleaving. Without bit interleaving, the false alarm rate is quite high. Simple Hamming with bit interleaving provides two vastly different false rates, depending on whether it is used for error correction or error detection. For the former mode, the false rate stays around the 10^{-2} level, while for the latter mode (error detection), the false rate drops to well below the 10^{-7} level (not shown). With bit interleaving the extended Hamming code shows a considerable improvement, dropping from the 10^{-2} to the 10^{-4} level, with a clear optimum region between 100 and 200 bit packets.

4.5 Conclusions

In Figure 4-21 the simulation results are compared to the theoretical ranges in which they are expected to fall, based on the bounds of Chapter 3. When no interleaving is used, the simulations agree favourably with the theory, and the results generally fall in, or very close to the theoretical ranges. When interleaving is used, the situation is quite different. To form an interleaved packet, i n -bit codewords are assembled together. The new code formed in this manner can be considered to have length n , or nx_i , depending on whether the individual word or the packet is viewed as the codeword. If the former viewpoint is adopted it can be concluded that there is a performance advantage in using interleaving. Furthermore, the advantage increases as the interleaving degree, i , increases. If the latter viewpoint is adopted, it is seen that interleaving causes a loss in performance. It is also observed that the loss increases in proportion to the interleaving degree. The second viewpoint is the natural one to adopt because the bits are in fact being processed nx_i at a time and not in sets of n . The second viewpoint is summarized as follows:

There exists a code of length nx_i and of rate R that will achieve better performance than an interleaved scheme with identical rate R , and whose interleaving period is also nx_i . Additionally, the relative performance advantage of the uninterleaved code over the

Figure 4-21: Block Error Rate: Results and Theory

Scheme	Simulation	Theoretical Range
AMPS, Forward Noninterleaved channel, Hull, 9.6 kbps	7.8×10^{-3}	$1.8 \times 10^{-2} - 8 \times 10^{-3}$
AMPS, Forward Noninterleaved channel, Downtown, 9.6 kbps	5.0×10^{-4}	$1.5 \times 10^{-3} - 3.2 \times 10^{-4}$
AMPS, Forward Noninterleaved channel, Downtown, 4.8 kbps	$< 6 \times 10^{-5}$	$1 \times 10^{-4} - 5 \times 10^{-6}$
Simple Hamming code, No interleaving, 4.8 kbps	5.3×10^{-3}	$4.3 \times 10^{-3} - 3.2 \times 10^{-3}$
Extended Hamming code, No interleaving, 4.8 kbps	6.0×10^{-3}	$3.8 \times 10^{-3} - 3.0 \times 10^{-3}$
Simple Hamming, 35 bit block, with interleaving	* 9.7×10^{-3}	$1 \times 10^{-2} - 7 \times 10^{-3}$
	+ 3.7×10^{-3}	$4.3 \times 10^{-3} - 3.2 \times 10^{-3}$
Simple Hamming, 210 bit block, with interleaving	* 8.6×10^{-3}	$1 \times 10^{-2} - 1.6 \times 10^{-3}$
	+ 6.9×10^{-4}	$4.3 \times 10^{-3} - 3.2 \times 10^{-3}$
Simple Hamming, 420 bit block, with interleaving	* 1.1×10^{-2}	$2 \times 10^{-3} - 1 \times 10^{-4}$
	+ 5.0×10^{-4}	$4.3 \times 10^{-3} - 3.2 \times 10^{-3}$
Extended Hamming, 48 bit block, with interleaving	* 1.0×10^{-2}	$1 \times 10^{-2} - 7 \times 10^{-3}$
	+ 3.3×10^{-3}	$3.8 \times 10^{-3} - 3.0 \times 10^{-3}$
Extended Hamming, 240 bit block, with interleaving	* 8.8×10^{-3}	$4.5 \times 10^{-3} - 4.0 \times 10^{-4}$
	+ 8.2×10^{-4}	$3.8 \times 10^{-3} - 3.0 \times 10^{-3}$
Extended Hamming, 480 bit block, with interleaving	* 1.3×10^{-2}	$6.0 \times 10^{-4} - 4.5 \times 10^{-6}$
	+ 5.7×10^{-4}	$3.8 \times 10^{-3} - 3 \times 10^{-3}$

* The length of the packet is used

+ The Hamming code length is used

interleaved packet will increase as the interleaving degree, i , is increased.

Based on the foregoing, it is therefore conjectured that interleaving is best viewed as a means to decrease decoder complexity rather than as a means to increase performance.

CHAPTER 5

CONCLUSIONS

5.1 Conclusions

In this thesis, an analysis of block code performance on mobile channels has been presented. This was accomplished by introducing a model for the distribution of the number of errors in a data block and then applying coding bounds to the model.

In Chapter 2, the assumption was made that errors were due primarily to the fading process, and a model for the distribution of the number of errors per block was derived. The model applies for different bit rates and transmission frequencies. Testing the general validity of the model is beyond the scope of the thesis, but it provides good agreement with experimental data at 4800 bps and 820 MHz. The model suggests that the ratio of incorrect bits to the total number of bits in a block gradually approaches a step function at the bit error rate as the block length increases. This behaviour is typical of many real channels, and holds for example for the binary symmetric channel (BSC).

It is conjectured that the channel model will remain valid for transmission rates ranging from about 500 bits per second to 200 kbps.

In Chapter 3, upper and lower bounds, valid for bounded distance decoding of block codes, were derived. The bounds were subsequently combined with the channel statistics to study the variation of the block error rate as a function of the code rate and length. Block codes used in the random error correction mode were shown to be an effective way to reduce error occurrences significantly. Single burst correcting codes were observed to be ineffective regardless of their rates and lengths. No proof of this last assertion was given, but the point was well illustrated in the graphs. The proof is conceptually simple but computationally difficult.

In Chapter 4, some practical schemes were evaluated and the results were compared to the predictions made in Chapter 3. Agreement between predicted and simulated results was good. A general conclusion that arose from the work of Chapter 4 was that interleaving was generally detrimental to performance, but was useful in reducing the complexity of the decoder.

5.2 Non-Optimality of Bounded Distance Decoding

In the thesis, only bounded distance decoding of block codes was considered. This is a natural choice because most decoders operate on this principle and because the complexity of schemes which attempt to correct beyond the minimum distance grows exponentially in the number of errors to be corrected. In a bounded

distance scheme, the code is not used to its full error correction ability. A penalty in performance is therefore expected. It is of interest to quantify this loss in performance. To this end, the notion of channel capacity must be introduced.

For most channels, there exists a rate above which information cannot be reliably recovered. This rate, known as channel capacity, can be conveniently expressed in the units of bits per channel use. Finding an explicit formula for the capacity of the binary digital land mobile channel appears quite difficult in general. For the purposes of the present discussion however, it is possible to bound the capacity of the channel.

The following observations and assumptions will be required. Firstly, it will be assumed that the channel is symmetric. As a second observation, it will be noted that the channel is bursty: it has memory. Its capacity must therefore be greater than that of a BSC with the same average probability of error. As a third point, it will be noted that in the channel model of Chapter 2, the bursts and the gaps are exponentially distributed. In that chapter, it was assumed that their average lengths were constant over local geographic areas. Thus over local geographic areas the mobile channel is well described by the Gilbert model [33] for which the capacity is well known [33]. Additionally, it is seen (Chapter 2) that the burst length varies only slightly as the BER increases from 10^{-4} to 10^{-1} . Lastly, it is

noted that in practice the channel also makes random errors, and that its parameters vary as the mobile changes its location. It is therefore not as well structured as a Gilbert channel.

Based on the above observations, the following conjecture is advanced: the capacity of the channel is less than that of either of Channels A or B described below. Channel A is a Gilbert channel [33] for which the average burst length is considerably larger than that of the Land Mobile channel, but for which the BER is the same. Channel B is a model proposed by McEliece and Stark [32]: Δ_0^m and $\Delta_{1/2}^m$ are BSC's with crossover probabilities of 0 and 1/2 respectively. Each block of m bits is transmitted over one of the component channels. The channel used in the transmission is chosen at random according to an a priori probability distribution $p(0) = 1 - \epsilon$ and $p(1/2) = \epsilon$. This is an especially well structured channel because disturbances last in multiples of m -bits.

In Figure 5-1 the capacities of the BSC and of the McEliece channel are plotted. The value chosen for m in the McEliece channel, is 500 bits or roughly 50 times the average burst length on a mobile channel. As long as the BER is below 10^{-2} , the capacities C_{BSC} and $C_{McEliece}$ are not very different. This suggests that the bound:

$$C_{BSC} \leq C_{Mobile} \leq C_{McEliece}$$

is a reasonably tight one.

In Figure 5-2 the maximum error free transmission rate achievable by a bounded distance scheme is compared to $C_{McEliece}$. The maximum rate is obtained from the Elias bound, which relates the code rate to its error correction ability. As the blocklength approaches infinity the ratio of incorrect bits to total bits approaches an impulse function at the bit error rate, that is, $t/n = \text{BER}$ as t and n tend to infinity. Thus code designed with an error correction ability of t' errors will reduce the probability of error to zero as the code length increases, only if $t' > n(t/n)$. The rate of such a code is bounded by the Elias bound to below channel capacity, as is apparent from the figure. However, the loss in performance is negligible if the uncoded BER is less than 10^{-3} .

5.3 Suggestions For Further Research and Recommendations

When operating below channel capacity, it is generally agreed that for most but not all channels, the block probability of error can be reduced further and further as the code length is increased. For the particular channel investigated herein, this is clearly demonstrated by the graphs of Figures 3-6 to 3-9 and 3-16 to 3-19.

For the block coding techniques investigated in the thesis, the blocklength required to achieve a given block error rate has been determined. It is of interest to examine whether other coding techniques can achieve

Figure 5-1 Bounds on the Capacity of the Binary DLMC as a Function of the BER

Figure 5-2 Best Achievable Zero-Error Rate with Bounded Distance Decoding of Block Codes

similar performance objectives with shorter code lengths. This is especially important because the blocklengths suggested by the thesis are sometimes too large to be of practical utility. Some of the techniques that may be investigated include:

- Erasure decoding and more generally soft decision schemes. The effectiveness of erasure based schemes was already demonstrated in Chapter 4 for a particular case. This investigation could be extended. Careful attention should be given to evaluating the complexity required by these schemes.

- Adaptive coding, whereby the decoder tries to guess the state of the channel and adapt its decoding algorithm accordingly.

- Reed-Solomon codes and other codes based on non-binary alphabets.

- Joint design of the modulator and the decoder. In contrast to the other approaches, this approach aims at increasing channel capacity.

In addition to investigating other schemes, a more complete set of expressions relating the bit error rate to the block error rate under the various conditions reported in the thesis would be of interest.

Finally, a more complete investigation of the relative merits of interleaving could be undertaken.

3

OF/DE

3

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

5.4 Recommendations

The main objective of the thesis was to provide a theoretical framework in which the performance of block-coded schemes could be assessed. But the investigation for the thesis has, in addition, provided us with the insight necessary to make some practical recommendations. They are as follows:

It is our belief that Reed-Solomon codes and erasure based techniques are the most promising coding alternatives for the land mobile channel, particularly if they are used together.

Given the length of the codes required to achieve satisfactory performance on the land mobile channel, we do not recommend the use of soft-decision techniques, whose complexity grows exponentially with the number of errors to be corrected.

Finally, interleaving and concatenation--although suboptimum in nature--might always be necessary to reduce the complexity of the decoder to tolerable levels.

REFERENCES

- [1] G.R. Cooper and R.W. Nettleton, "Cellular Mobile Technology: The Great Multiplier", IEEE Spectrum, Vol.20, June 1983, pp.30-37.
- [2] H.M. Hafez, et al., "Experimental Study of Digital Land Mobile Systems", Technical Report, Department of Communications and Carleton University, September 1983.
- [3] "Advanced Mobile Phone Service", Bell System Tech. J., Special Issue, Vol.58, January 1979.
- [4] Theodore Brenig, "Data Transmission for Mobile Radio", IEEE Trans. on Vehicular Technology, Vol.VT-27, No.3, August 1978, pp.78-85.
- [5] A.F.I. Lam, "Forward Error Correction for Land Mobile Channels", Master's Thesis, Carleton University, July 1981.
- [6] J.L. Massey, "Coding and Modulation in Digital Communications", International Zurich Seminar on Digital Communications, 12-15 March 1974, pp.E2-1 to E2-4.
- [7] P.J. Mabey, "Mobile Radio Data Transmission -- Coding for Error Control", IEEE Trans. on Vehicular Technology, Vol.VT-27, August 1978, pp.99-109.
- [8] T.A. Knight, "Fading Characteristics of the Land Mobile Radio Channel at 820 MHz", Master's Thesis, Carleton University, May 1985.
- [9] Francois Letourneau, "Data Transmission Over A Mobile Radio Channel", Graduate Project, Carleton University, 1984.
- [10] W.C. Jakes, Jr., Editor, "Microwave Mobile Communications", John Wiley and Sons, Inc., New York, 1974.
- [11] J.S. Da Silva, H.M. Hafez, and S.A. Mahmoud, "Optimal Packet Length for Fading Land Mobile Data Channels", Proceedings of the International Conference on Communications, June 1980, pp.61.3.1-61.3.5.

- [12] R.A. Comroe and D.J. Costello, Jr., "ARQ Schemes for Data Transmission in Mobile Radio Systems", IEEE Journal on Selected Areas in Communications, Vol.2, July 1984, pp.472-481.
- [13] S.O. Rice, "Mathematical Analysis of Random Noise", Introduction, Bell Syst. Tech. J., Vol.23, July 1944, pp.282-332.
- [14] S.O. Rice, "Mathematical Analysis of Random Noise", Bell Syst. Tech. J., Vol.24, January 1945, pp.46-156.
- [15] S.O. Rice, "Statistical Properties of a Sine Wave Plus Random Noise", Bell Syst. Tech. J., Vol.27, January 1948, pp.109-157.
- [16] S.O. Rice, "Distribution of the Duration of Fades in Radio Transmission: Gaussian Noise Model", Bell Syst. Tech. J., Vol.37, May 1958, pp.581-635.
- [17] N.A. Strakhov and L. Kurz, "An Upper Bound on the Zero-Crossing Distribution", Bell Syst. Tech. J., Vol.47, April 1968, pp.529-547.
- [18] W.F. Bodtmann and H.W. Arnold, "Fade Duration Statistics of a Rayleigh-Distributed Wave", IEEE Transactions on Communications, Vol.30, March 1982, pp.549-553.
- [19] H.W. Arnold and W.F. Bodtmann, "Interfade Interval Statistics of a Rayleigh-Distributed Wave", IEEE Transactions on Communications, Vol.31, September 1983, pp.1114-1116.
- [20] W.W. Peterson and E.J. Weldon, "Error-Correcting Codes", Prentice-Hall, Englewood Cliffs, N.J., 1970.
- [21] E. Berlekamp, "Algebraic Coding Theory", McGraw-Hill Book Co. New York, 1968.
- [22] R.E. Blahut, "Theory and Practice of Error Control Codes", Addison-Wesley Publishing Co., Reading, Mass., 1983.
- [23] F.J. MacWilliams and N.J.A. Sloane, "The Theory of Error Correcting Codes", North-Holland Publishing Co., New York, 1977, Second Printing 1978, (esp.Ch.10 and 18).
- [25] G.C. Clark and J.B. Cain, "Error-Correction Coding for Digital Communications", Plenum Press, New York, 1981.

- [26] R.G. Gallager, "Information Theory and Reliable Communication", John Wiley and Sons, Inc., New York, 1968.
- [27] E.R. Berlekamp, "The Technology of Error-Correcting Codes", Proceedings of the IEEE, Vol.68, May 1980, pp.564-593.
- [28] S. Lin, D.J. Costello, Jr., and M.J. Miller, "Automatic Repeat Request Error-Control Schemes", IEEE Communications Magazine, Vol.22, December 1984, pp.5-16.
- [29] R.J. McEliece, E.R. Rodemich, H. Rumsey, Jr., and L.R. Welch, "New Upper Bounds on the Rate of a Code Via the Delsarte-MacWilliams Inequalities", IEEE Transactions on Information Theory, Vol.23, March 1977, pp.157-166.
- [30] R.C. French, "Error Rate Predictions and Measurements in the Mobile Radio Data Channel", IEEE Transactions on Vehicular Technology, Vol.27, August 1978, pp.110-116.
- [31] H.M. Hafez, A.U.H. Sheikh and B. McPhail, "Experimental Evaluation of Packet Error Rates at 850 MHz", IEE Proceedings, Vol.132, Pt.F, No.5, August 1985, pp.366-374.
- [32] G.D. Forney, Jr., "Burst-Correcting Codes for the Classic Bursty Channel", IEEE Transactions on Communications Technology, Vol.19, October 1971, pp.772-781.
- [33] R.J. McEliece and W.E. Stark, "Channels With Block Interference", IEEE Transactions on Information Theory, Vol.30, January 1984, pp.44-53.
- [34] E.N. Gilbert, "Capacity of a Burst-Noise Channel", Bell Syst. Tech. J., Vol.39, September 1960, pp.1253-1265.
- [35] A.J. Viterbi, "Spread Spectrum Communications -- Myths and Realities", IEEE Communications Magazine, Vol.4, May 1979, pp.11-18.
- [36] "Theorie de l'information", in "Theorie du codage algebrique", Revue du Cethedec, No.57, 1978, particularly pp.12-18.

- [37] L.B. Milstein, S. Davidovici, and D.L. Schilling,
"Coding and Modulation Techniques for Frequency-
Hopped Spread-Spectrum Communications Over a Pulse-
Burst Jammed Rayleigh Fading Channel", IEEE Journal
on Selected Areas in Communications, Vol.3,
September 1985, pp.644-651.

APPENDIX I

BIT ERROR RATES AT 820 MHz

LOCATION : DOWNTOWN

TRANSMISSION : FROM BASE TO MOBILE

file name	P(e)	$\bar{b}$	σ_b	fraction of file recovered w_1
MDI01.01	2.38×10^{-3}	5.5	3.6	1
MDI01.02	3.90×10^{-3}	9.9	14.2	1
MDI01.03	5.61×10^{-3}	8.9	9.1	1
MDI01.04	3.82×10^{-3}	8.3	9.7	1
MDI01.05	1.36×10^{-3}	9.1	8.8	1
MDI01.06	7.51×10^{-4}	9.6	8.7	1
MDI01.07	4.38×10^{-4}	8.6	7.4	1
MDI01.08	1.48×10^{-3}	8.3	5.4	1
MDI01.09	1.26×10^{-3}	8.4	8.8	1

LOCATION : DOWNTOWN

TRANSMISSION : FROM MOBILE TO BASE

file name	P(e)	$\bar{b}$	σ_b	fraction of file recovered w_1
MDI04.01	1.80×10^{-1}	14.5	17.1	10/12
MDI04.02	1.02×10^{-1}	14.5	21.5	1
MDI04.03	1.46×10^{-1}	17.5	25.2	1
MDI04.04	3.03×10^{-1}	22.3	19.8	1/10
MDI04.05	2.94×10^{-1}	22.4	23.6	10/12
MDI04.06	1.76×10^{-1}	13.6	17.4	1
MDI04.07	1.47×10^{-1}	16.0	22.5	1
MDI04.08	8.4×10^{-2}	12.6	15.3	1
MDI04.09	2.0×10^{-1}	15.4	19.9	1

LOCATION: WEST-END

TRANSMISSION : FROM MOBILE TO BASE

file name	P(e)	$\bar{b}$	σ_b	fraction of file recovered w_1
MDI05.01	5.21×10^{-2}	11.3	14.7	1
MDI05.02	6.83×10^{-2}	12.5	16.4	1
MDI05.03	6.10×10^{-2}	10.5	18.7	1
MDI05.04	8.70×10^{-3}	9.1	10.3	1
MDI05.05	4.34×10^{-2}	13.5	18.1	1
MDI05.06	5.95×10^{-2}	13.0	29.5	5/6
MDI05.07	6.56×10^{-2}	12.3	14.3	1
MDI05.08	1.35×10^{-1}	13.5	22.3	1
MDI05.09	7.37×10^{-3}	13.1	18.9	1
MDI05.10	1.50×10^{-1}	13.2	16.6	1
MDI05.11	8.38×10^{-2}	13.1	17.9	1
MDI05.12	3.40×10^{-2}	11.2	15.5	1
MDI05.13	2.01×10^{-1}	14.9	19.3	4/5

LOCATION : HULL

TRANSMISSION : FROM BASE TO MOBILE

file name	P(e)	$\bar{b}$	σ_b	fraction of file recovered w_1
MDI09.01	8.65×10^{-3}	10.8	15.4	1
MDI09.02	7.51×10^{-3}	10.6	11.8	1
MDI09.03	1.16×10^{-2}	12.1	19.3	1
MDI09.04	7.36×10^{-3}	9.4	12.6	1
MDI09.05	5.78×10^{-3}	9.6	11.7	1
MDI09.06	4.80×10^{-3}	11.2	15.3	1
MDI09.07	5.05×10^{-3}	10.8	12.4	1
MDI09.08	6.29×10^{-3}	10.9	15.1	1
MDI09.09	7.50×10^{-3}	8.4	10.0	1

LOCATION : HULL

TRANSMISSION : FROM MOBILE TO BASE:

file name	P(e)	$\bar{b}$	σ_b	fraction of file recovered w_1
MDI10.01	8.25×10^{-2}	10.8	17.4	1
MDI10.02	3.63×10^{-2}	11.2	12.8	1
MDI10.03	2.57×10^{-2}	9.7	12.4	1
MDI10.04	2.08×10^{-2}	11.6	15.1	1
MDI10.05	5.25×10^{-2}	9.9	16.9	1
MDI10.06	4.08×10^{-2}	11.6	14.3	1
MDI10.07	8.16×10^{-2}	9.0	13.2	1
MDI10.08	2.69×10^{-2}	11.8	14.1	1
MDI10.09	2.33×10^{-2}	11.0	13.0	1

LOCATION : SOUTH (RESIDENTIAL AREA)
TRANSMISSION : FROM BASE TO MOBILE
file name

	$P(e)$	$\bar{b}$	σ_b	fraction of file recovered w_1
MDI13.01	1.07×10^{-1}	12.4	16.3	1
MDI13.02	2.18×10^{-2}	7.5	8.5	1
MDI13.03	9.68×10^{-2}	15.0	21.4	1
MDI13.04	1.54×10^{-1}	15.2	21.1	2/3
MDI13.05	1.48×10^{-1}	15.6	20.6	5/6
MDI13.06	2.09×10^{-2}	15.4	20.2	11/12
MDI13.07	8.27×10^{-2}	14.4	20.7	11/12
MDI13.08	3.11×10^{-1}	24.5	30.8	1/4
MDI13.09	1.81×10^{-1}	14.2	19.0	1

LOCATION : QUEENSWAY
TRANSMISSION : FROM BASE TO MOBILE
file name

	$P(e)$	$\bar{b}$	σ_b	fraction of file recovered w_1
MDI12.01	9.98×10^{-2}	11.7	17.2	1
MDI12.02	4.68×10^{-2}	8.6	13.5	11/12
MDI12.03	4.79×10^{-2}	12.3	24.3	1
MDI12.04	1.58×10^{-1}	11.9	16.1	11/12
MDI12.05	2.81×10^{-1}	20.3	25.2	1/2
MDI12.06	1.31×10^{-1}	10.7	13.7	1
MDI12.07	1.40×10^{-1}	12.1	14.8	1
MDI12.08	1.32×10^{-1}	11.9	14.5	1
MDI12.09	7.80×10^{-2}	10.9	14.0	11/12
MDI12.10	2.81×10^{-2}	9.5	12.9	1
MDI12.11	6.37×10^{-2}	13.5	16.9	1
MDI12.12	1.69×10^{-3}	6.9	6.5	1
MDI12.13	4.48×10^{-3}	6.8	9.6	1
MDI12.14	8.31×10^{-3}	6.1	4.9	1
MDI12.15	4.78×10^{-2}	7.5	6.8	1
MDI12.16	5.66×10^{-2}	9.1	9.0	1
MDI12.17	2.12×10^{-3}	7.6	10.3	1
MDI12.18	7.27×10^{-3}	7.7	20.3	1
MDI12.19	2.35×10^{-2}	7.5	8.4	1
MDI12.20	4.79×10^{-2}	8.1	7.9	1
MDI12.21	3.87×10^{-2}	8.1	8.0	1
MDI12.22	1.93×10^{-2}	6.9	7.4	1
MDI12.23	5.47×10^{-2}	10.2	10.9	1

APPENDIX II

SUMMARY OF AVAILABLE ERROR PATTERN FILES

- MDI01 (9 files: MDI01.01 to MDI01.09)
Location: Downtown Ottawa
Transmission: from base to mobile in the 800 MHz band
- MDI02 (13 files: MDI02.01 to MDI02.13)
Location: Downtown Ottawa
Transmission: from base to mobile in the UHF band
- MDI03 (13 files: MDI03.01 to MDI03.13)
Location: Downtown Ottawa
Transmission: from base to mobile in the VHF band
- MDI04 (9 files: MDI04.01 to MDI04.09)
Location: Downtown Ottawa
Transmission: from mobile to base in the 800 MHz band
- MDI05 (9 files: MDI05.01 to MDI05.09)
Location: Ottawa West End (Urban/suburban terrain)
Transmission: from mobile to base in the 800 MHz band
- MDI06 (9 files: MDI06.01 to MDI06.09)
Location: Ottawa West End (Urban/suburban terrain)
Transmission: from base to mobile in the VHF band
- MDI07 (10 files: MDI07.01 to MDI07.10)
Location: Downtown Hull
Transmission: from base to mobile in the VHF band
- MDI08 (9 files: MDI08.01 to MDI08.09)
Location: Downtown Hull
Transmission: from base to mobile in the UHF band
- MDI09 (9 files: MDI09.01 to MDI09.09)
Location: Downtown Hull
Transmission: from base to mobile in the 800 MHz band
- MDI10 (9 files: MDI10.01 to MDI10.09)
Location: Downtown Hull
Transmission: from mobile to base in the 800 MHz band
- MDI11 (23 files: MDI11.01 to MDI11.23)
Location: Queensway (open highway)
Transmission: from base to mobile in the UHF band

- MDI12 (23 files: MDI12.01 to MDI12.23)
Location: Queensway (open highway)
Transmission: from base to mobile in the 800 MHz band
- MDI13 (9 files: MDI13.01 to MDI13.09)
Location: Ottawa South (residential)
Transmission: from mobile to base in the 800 MHz band
- MDI14 (9 files: MDI14.01 to MDI14.09)
Location: Ottawa South (residential)
Transmission: from base to mobile in the VHF band
- MDI15 (9 files: MDI15.01 to MDI15.09)
Location: Ottawa South (residential)
Transmission: from mobile to base in the UHF band
- MDI16 (22 files: MDI16.01 to MDI16.22)
Location: Queensway (open highway)
Transmission: from mobile to base in the VHF band

APPENDIX III

A SHORT TABLE OF THE ELIAS BOUND

[illegible]

APPENDIX IV

A SHORT TABLE OF THE GILBERT-VARSHARMOV BOUND

APPENDIX V

DESIGN OF THE SHORTENED BCH (40,28) DECODER

The shortened (40,28) code is obtained from the BCH (63,51) code by dropping $b=51-28=23$ information bits from each codeword. Because the decoder only corrects one error, a Meggit design is the preferable choice.

The first step in the design of the decoder is to find the generator polynomial for the original BCH (63,51) code. The code has minimum distance five. Its generator polynomial is given by $g(x) = \text{LCM}[f_1(\alpha), f_2(\alpha^2), f_3(\alpha^3), f_4(\alpha^4)] = \text{LCM}[f_1(\alpha), f_3(\alpha^3)]$, where α is some primitive element of $\text{GF}(2^6)$ and $f_1(\alpha^1)$ is the minimal polynomial of α^1 . From a table of primitive polynomials [20], the generator polynomial is therefore found to be $g(x) = x^{12} + x^{10} + x^8 + x^5 + x^4 + x^3 + 1$. The feed-back connections of the syndrome register are thus given by $g(x) = x^{12} + x^{10} + x^8 + x^5 + x^4 + x^3 + 1$.

For an unshortened code, it is convenient to define the syndrome as the remainder of $x^{n-k} v(x)$ divided by the generator polynomial of the code, $g(x)$. Here $v(x)$ is the polynomial which corresponds to the received word. For a shortened code, it is preferable to use $x^{n-k+b} v(x) \bmod g(x)$ as a syndrome because this makes the syndrome of an error in the high-order bit equal to x^{n-k-1} . This also

allows the decoding to proceed on the basis of the shortened rather than the original code length.

Premultiplication of the received word by $x^{n-k+b} = x^{63-51+23} = x^{35} \bmod g(x)$ is equivalent to premultiplication by $x^{11} + x^{10} + x^6 + x^4 + x^3 + x + 1$. The feed-forward coefficients of the syndrome register are thus given by $x^{11} + x^{10} + x^6 + x^4 + x^3 + x + 1$. This completes the design of the syndrome register.

Because the decoder is required to correct single errors only, the syndrome pattern that needs to be checked is $S(x) = x^{11}$, or in vector notation $S = (0,0,0,0,0,0,0,0,0,0,0,1)$. When this pattern is found a one is subtracted from the bit exiting the 40 bit buffer, thus correcting the error. The effect of the error is also removed from the syndrome register. Errors are detected when the syndrome is non-zero and fails to match the pattern $(0,0,0,0,0,0,0,0,0,0,0,1)$ in the 80 shifts required to decode the codeword. The decoder is illustrated in Figure 4-1.

APPENDIX VI

SCHEMATIC DIAGRAMS FOR THE PACKET SCHEME
BASED ON THE (8,4) EXTENDED HAMMING CODE

high level: switch closed
low level: switch open

GENERAL STRUCTURE

high level: switch closed
low level: switch open

C O D E R

notes: * I₁₅ is a reset state (all flip-flops cleared)

high level: switch closed
low level: switch open

DECODER

END

17.08.87

FIN