INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600
RICE UNIVERSITY

Studies on Capacity and Performance of Digital Transmission over Copper Loops

by

Andrew Sendonaris

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
Master of Science

APPROVED, THESIS COMMITTEE:

Dr. Behnam Aazhang, Chair
Associate Professor
Electrical and Computer Engineering

Dr. Don H. Johnson
Professor
Electrical and Computer Engineering

Dr. Joseph R. Cavallaro
Associate Professor
Electrical and Computer Engineering

Dr. Venu Veeravalli
Visiting Assistant Professor
Electrical and Computer Engineering

Houston, Texas
May, 1995
Studies on Capacity and Performance of Digital Transmission over Copper Loops

Andrew Sendonaris

Abstract

Digital transmission over copper loops is studied. Models are found of main sources of interference, including near-end crosstalk (NEXT), far-end crosstalk (FEXT), impulse noise (IN) and intersymbol interference (ISI). Importance Sampling theory is introduced, and an explanation is given of how it can be used to efficiently estimate the bit error rate (BER) of the above digital transmission system. Results demonstrate the efficiency of Importance Sampling simulations and the validity of the models for the channel interference.

A suboptimal scheme for approaching capacity of the twisted pair is analyzed and its connection to multicarrier modulation is presented. The resulting lower bound on capacity is a significant improvement over what exists in the literature.
Acknowledgments

First, I would like to thank my advisor Dr. Behnaam Aazhang for his guidance and support. Also, the help of Dr. Andrew Deszcky from BNR, Dr. Don Johnson, and Dr. Venu Veeravalli has been invaluable. I would also like to thank Dr. Joseph Cavallaro for serving on my thesis committee, as well as all my fellow graduate students who have helped me on several occasions.
Contents

Abstract ii
Acknowledgments iii
List of Illustrations vii

1 Introduction 1

2 Alpha-mix: A Method for Approaching Capacity on Twisted Pairs 5
2.1 Capacity of Twisted Pairs 5
2.2 Simplified Channel 7
2.3 Generalized Channel 12
2.4 Numerical Results and Discussion 17
2.5 Implementation with Multicarrier Modulation 22
2.6 Possible Extension to More Than Two Users 23

3 Statistical Models for Main Sources of Interference on Copper Loops 25
3.1 NEXT 26
3.2 FEXT 27
3.3 Impulse noise (IN) 28
3.4 Intersymbol Interference (ISI) 29
3.4.1 Numerical calculation of pdf .................................. 31
3.5 Additive white Gaussian noise (WGN) ......................... 33

4 Performance of Digital Transmission over Copper

Loops 35

4.1 Importance Sampling .............................................. 35
  4.1.1 Conditional Importance Sampling .......................... 37
  4.1.2 Example .................................................... 38
4.2 Efficient Simulation of Digital Copper Loops .................. 39
  4.2.1 General Twisted Pair Channel ............................ 39
  4.2.2 An ISDN Example ......................................... 45

5 Conclusions 50

Appendix A 52

A.1 Gaussian Signals and Capacity ................................. 52
A.2 Analysis of Class of $S_1(f)$ and $S_2(f)$ Defined in Terms of $\beta$ in the
  Simplified Channel ............................................. 53
A.3 Analysis of Conjecture 2.1 .................................... 54
  A.3.1 Low SNR ................................................. 54
  A.3.2 Any SNR ................................................ 56
  A.3.3 High SNR ................................................ 57
  A.3.4 Discussion .............................................. 59

Appendix B 61

B.1 Relation Between $K_{NEXT}$ and PSD of $NEXT$ ............... 61
B.2 Derivation of the Density of Impulse Noise .................. 62

Appendix C ............................................. 64
C.1 Bounds on Gain ........................................... 64

Bibliography ........................................... 68
Illustrations

2.1 Simplified channel: how $S_1(f)$ changes with $\alpha$ ......................... 8
2.2 $C_1 + C_2$ as a function of $\alpha_1$ and $\alpha_2$, for four different values of SNR  . 11
2.3 Simplified channel: how $S_1(f)$ changes with $\beta$ .......................... 12
2.4 Channel and crosstalk transfer functions for twisted-pair model from
   paper by Aslanis and Cioffi ........................................... 14
2.5 Capacity calculations for twisted-pair model from paper by Irving Kalet  18
2.6 Capacity calculations for twisted-pair model from paper by Aslanis
   and Cioffi ................................................................. 20
2.7 Graphical illustration of Multicarrier Modulation ........................... 23
2.8 Capacity as a function of SNR when there are $N$ twisted pairs in the
   same binder group ......................................................... 24

3.1 Schematic Representation of NEXT and FEXT ............................... 26
3.2 Illustration of ISI ....................................................... 30
3.3 Transmitted pulse for ISDN (Approximate) .................................. 31

4.1 $\hat{W}$ vs $\alpha$ for the Importance Sampling simulation of the general
   twisted pair channel ..................................................... 43
4.2 Upper and lower bounds on the Importance Sampling Gain $\Gamma$ versus $P_e$ 44
A.1 $C_{ef} - C_{fe}$ vs SNR 

60
Chapter 1

Introduction

Twisted pair (and coaxial) copper cables have long been used to transmit analog and digital data. Although optical fiber has proven to be a superior medium for transmitting data, its costs are prohibitive for commercial use, and will remain so for the foreseeable future. Therefore, it is desirable to use the existing cables for high speed digital transmission, usually within what is called a carrier serving area (CSA), which covers an area of radius up to 12000 ft around the central office [1,2].

In general, about 50 twisted pairs are put together in one binder group. Thus, each twisted pair faces interference from 49 other twisted pairs, each carrying its own signal. Near-end crosstalk (NEXT) is caused by a neighboring twisted pair whose transmitter is physically close to the desired receiver and far-end crosstalk (FEXT) is caused by a twisted pair whose transmitter is physically far from the desired receiver. There is also impulse noise (IN) interference from electromechanical switching devices in the central office [3], and Intersymbol Interference (ISI) which is interference from all previously transmitted symbols and is due to the channel’s non-flat transfer function. These and other sources of interference limit the capability of a twisted pair to carry high bit rate digital data.

There are two major considerations when studying twisted pairs (and any other transmission medium). First, determining channel capacity [4,5], and second, being able to evaluate the performance of given transmission schemes.
Capacity is an indication of the maximum rate at which information can be sent reliably over the channel. That is, the capacity of a channel is the maximum number of bits per second that can be sent over the channel with an arbitrarily small probability of error. Thus capacity is an extremely useful quantity when studying a channel.

When the channel transfer function and the noise Power Spectral Density (PSD) are non-flat, that is, not constant over the transmission frequency range, the form of the signal PSD that achieves capacity is, in general, an open problem. For Gaussian noise, this problem has been solved [4, 6]. The results point to the fact that more signal power should be assigned to the better parts of the spectrum (low attenuation, low noise) than to the other parts of the spectrum. A severe difficulty arises in the case of twisted pairs because, due to the crosstalk, the PSD of the noise is dependent on the PSD of the interfering users. Therefore, to obtain the capacity of a twisted pair, it is necessary to find the signal PSD’s for all the twisted pairs in the binder group, so as to maximize the transmission rate for all users. A desirable property of the solution should be that it results in the same transmission rate for all users.

Unfortunately, the general solution to this problem is analytically intractable, even in the case when there are only two users and even when crosstalk is assumed to be Gaussian. Therefore, to make the problem tractable, constraints have to be placed on the PSD’s. In the literature, it has generally been assumed that the users will have equal PSD’s [7, 1]. This results in an analytically tractable solution when the noise in the channel is assumed to be additive white Gaussian noise (AWGN) plus crosstalk, which is assumed to be also Gaussian. Lechleider in [8] presented the only work where it is not assumed that the users have equal PSD’s. However, Lechleider addresses the issue when there is no AWGN in the channel, and also imposes constraints on the PSD’s that result in a solution that can be improved significantly (see Chapter 2).
In general, the problem with results proving the achievability of capacity is that they guarantee only the existence of a transmission scheme that achieves capacity. No methodology for finding this transmission scheme can be provided.

Transmission schemes consist of encoding the source and then modulating it so that it can be transmitted over the channel. At the receiver end, the signal is demodulated and then decoded to produce the desired information. Since the concept of capacity was introduced \([5]\), a considerable amount of research has gone into finding transmission schemes that approach capacity. Most of the work was concerned with finding the best coding scheme, and this was considered as a separate issue from finding the modulation and demodulation scheme. As a result of this research, block codes and convolutional codes were developed \([9\text{-}12]\) for the coding part of the problem, and such ideas as multicarrier modulation and carrierless amplitude and phase modulation were developed \([13\text{-}16]\) for the modulation part of the problem. Both had varying degrees of success. Relatively recent work \([17]\) has shown that combining the steps of encoding and modulating, thus obtaining coded modulation, can result in transmission schemes that approach the capacity of the continuous-time channel. In all cases, once a transmission system has been proposed, it is desirable to test its performance.

Therefore, the second consideration is, given a transmission system, to be able to evaluate some measure of its performance. A measure of the quality of a digital transmission system is the bit error rate \((BER, or \ P_e)\). This is the probability that the receiver will make an incorrect decision on a received bit. The analytical calculation of the effects of all the interference on the BER of the system is an intractable problem in many practical systems. Therefore, computer simulations are a very common way of obtaining an estimate of the BER for a given configuration.
The most frequently used type of simulation is Monte Carlo simulation which consists of simulating the transmission of a large number of bits, counting the errors at the receiver resulting from this transmission, and then dividing the number of errors by the number of transmitted bits to obtain an estimate of the BER. The main difficulty with Monte Carlo simulations is that to get a fairly accurate estimate of the BER, it is necessary to simulate the transmission of approximately $\frac{10}{P_e}$ to $\frac{100}{P_e}$ bits [18]. When the system has an error rate on the order of $10^{-6}$, this method requires a prohibitively large number of simulation runs. One would like to find a way of estimating the BER of the system with fewer simulation runs than are required with Monte Carlo simulation, while still achieving the same accuracy, that is, the same variance of the estimate.

One way of achieving this goal is by using the theory of Importance Sampling [18–21]. The main concept behind this theory is making rare events, e.g. bit errors when $P_e$ is very small, more frequent. It is necessary, however, when counting the errors, to weigh them in such a way that the estimate is unbiased and has a small variance.

This thesis addresses both considerations described above. First, a suboptimal scheme for approaching channel capacity for twisted pairs is described. Thus, a lower bound on capacity is found. It will be shown that this lower bound is a significant improvement over what exists in the literature.

Second, models for the main sources of interference on twisted pairs are found, so that Importance Sampling can be applied to arbitrary transmission schemes over twisted pairs. Finally, Importance Sampling simulations are performed for two transmission schemes, including the widely-used 2B1Q.
Chapter 2

Alpha-mix: A Method for Approaching Capacity on Twisted Pairs

2.1 Capacity of Twisted Pairs

The capacity of a channel is defined to be the highest rate of information that can be sent through the channel, with no errors. Although error-free transmission has not been achieved on most practical systems, capacity still remains a good indicator of how much information can be sent on a channel, with acceptable probability of error.

This section considers the capacity of twisted pairs. The channel will be assumed to have transfer function $H_c(f)$, and two sources of noise: AWGN (Additive White Gaussian Noise) with spectral density $N_o$, and NEXT (Near-end crosstalk) with crosstalk transfer function $H_x(f)$. For purposes of analytical tractability, it is also assumed that there are only two users. Under such assumptions, the capacity for user 1, given that user 2 transmits a Gaussian signal (see Appendix A), is given by [7,1,8]

$$C_1 = \sup_{S_1(f), S_2(f)} \int_0^\infty \log_2 \left[ 1 + \frac{|H_c(f)|^2 S_1(f)}{N_o + |H_x(f)|^2 S_2(f)} \right] df \text{ (bits/sec)} \quad (2.1)$$

where the supremum is taken over all $S_1(f)$ and $S_2(f)$ such that

$$S_i(f) \geq 0 \ \forall f, \text{ and } \int_0^\infty S_i(f) df \leq P_{\text{max}} \text{ for } i = 1, 2$$

and where $S_i(f)$ is the PSD of the $i^{th}$ user, and the latter inequality corresponds to an average power constraint.
The capacity for user 2 is given by the analogous expression. Of course, when designing the transmission scheme for both twisted pairs, it becomes necessary to consider both users simultaneously. As can be seen from (2.1), when only user 1 is considered, \( C_1 \) is achieved when \( S_2(f) = 0 \ \forall f \), which implies that \( C_2 = 0 \). The same problem arises when only user 2 is considered. Therefore, a joint optimization needs to be performed. There are several ways in which this can be done.

Let

\[
C^*(S_1(f), S_2(f)) = \int_0^\infty \log_2 \left[ 1 + \frac{|H_6(f)|^2 S_1(f)}{N_o + |H_2(f)|^2 S_2(f)} \right] df
\]

Then

\[
C_1 = C^*(S_1^*(f), S_2^*(f)) \quad \text{and} \quad C_2 = C^*(S_2^*(f), S_1^*(f))
\]

where \( S_1^*(f), S_2^*(f) = \arg \max_{S_1(f), S_2(f)} C^*(S_1(f), S_2(f)) \) \( \text{s.t. } C^*(S_1(f), S_2(f)) = C^*(S_2(f), S_1(f)) \)

or \( S_1^*(f), S_2^*(f) = \arg \max_{S_1(f), S_2(f)} [C^*(S_1(f), S_2(f)) + C^*(S_2(f), S_1(f))] \)

or \( S_1^*(f), S_2^*(f) = \arg \max_{S_1(f), S_2(f)} \min[C^*(S_1(f), S_2(f)), C^*(S_2(f), S_1(f))] \)

The problem with the above approach is that the solution is analytically intractable. With few exceptions [8], most authors assume \( S_1(f) = S_2(f) \), thus making the problem tractable [7, 1]. It is desirable not to make this assumption because it is too restrictive and thus will not necessarily yield the global maximum.

Lechleider [8] also points to several reasons why it should not be assumed a priori that \( S_1(f) = S_2(f) \). However, Lechleider arrives at a result which, as will be shown in Section 2.4, is not the global maximum. By assuming \( S_1(f) = S_2(f) \), the problem reduces to [7, 1]

\[
C_1 = \max_{S_1(f)} C^*(S_1(f), S_1(f)) \quad \text{(2.2)}
\]
It would be desirable to reduce the problem to a form similar to (2.2) without the assumption that \( S_1(f) = S_2(f) \). The main idea of this part of the thesis is to make \( S_1(f) \) and \( S_2(f) \) "symmetric", in some sense, thus reducing the problem to

\[
C_1 = \max_{S_1(f)} C^*(S_1(f), S_1^{\text{sym}}(f))
\]

(2.3)

where \( S_1^{\text{sym}}(f) \) is symmetric to \( S_1(f) \). Despite the fact that it is difficult to rigorously define the concept of symmetry, generally speaking it implies that \( S_1^{\text{sym}}(f) \) is small where \( S_1(f) \) is large, and vice versa, so that \( C^*(S_1(f), S_1^{\text{sym}}(f)) = C^*(S_1^{\text{sym}}(f), S_1(f)) \). This concept will become clearer in the following section.

The second idea is to parameterize \( S_1(f) \) so that the maximization is taken over the parameters that now describe \( S_1(f) \).

### 2.2 Simplified Channel

Consider a channel with

\[
|H_1(f)|^2 = \begin{cases} 
H & \text{if } |f| \leq W \\
0 & \text{otherwise}
\end{cases}, \quad |H_2(f)|^2 = \begin{cases} 
X & \text{if } |f| \leq W \\
0 & \text{otherwise}
\end{cases}
\]

Consider the class of \( S_1(f) \) and \( S_2(f) \) defined as follows:

\[
S_1(f) = \begin{cases} 
\alpha \frac{2P_{\text{max}}}{W} & \text{if } |f| \leq \frac{W}{2} \\
(1 - \alpha) \frac{2P_{\text{max}}}{W} & \text{if } \frac{W}{2} < |f| \leq W \\
0 & \text{otherwise}
\end{cases}
\]

(2.4)

\[
S_2(f) = \begin{cases} 
(1 - \alpha) \frac{2P_{\text{max}}}{W} & \text{if } |f| \leq \frac{W}{2} \\
\alpha \frac{2P_{\text{max}}}{W} & \text{if } \frac{W}{2} < |f| \leq W \\
0 & \text{otherwise}
\end{cases}
\]

(2.5)

where \( 0.5 \leq \alpha \leq 1 \). See Figure 2.1. In this configuration, \( \alpha = 0.5 \) corresponds to using equal PSD's for the two users, and \( \alpha = 1 \) corresponds to using Frequency Division Signaling (FDS).
It is easily verified that $S_1(f)$ and $S_2(f)$ satisfy the power constraint

$$
\int_0^\infty S_1(f)df = \frac{W}{2}(\alpha \frac{2P_{\text{max}}}{W}) + \frac{W}{2}((1 - \alpha) \frac{2P_{\text{max}}}{W})
$$

$$
= P_{\text{max}}
$$

The resulting capacity for user 1 becomes

$$
C_1 = \max_{0.5 \leq \alpha \leq 1} \frac{W}{2} \log_2 \left[ 1 + \frac{\alpha \frac{2P_{\text{max}}}{W} H}{N_0 + (1 - \alpha) \frac{2P_{\text{max}}}{W} X} \right] + \frac{W}{2} \log_2 \left[ 1 + \frac{(1 - \alpha) \frac{2P_{\text{max}}}{W} H}{N_0 + \alpha \frac{2P_{\text{max}}}{W} X} \right]
$$

$$
= \max_{0.5 \leq \alpha \leq 1} \frac{W}{2 \ln(2)} \left\{ \ln \left[ 1 + \frac{\alpha SH}{1 + (1 - \alpha) SX} \right] + \ln \left[ 1 + \frac{(1 - \alpha)SH}{1 + \alpha SX} \right] \right\}
$$

where $S = \frac{2P_{\text{max}}}{N_0 W}$.

It is clear that it is only necessary to maximize $C_1$, and this guarantees that $C_2$ will be equal to $C_1$. Taking the derivative with respect to $\alpha$, ...
\[
\frac{\partial C^*}{\partial \alpha} = \frac{W}{2 \ln(2)} \left[ \frac{SH(1 + SX)}{(1 + (1 - \alpha)SX + \alpha SH)(1 + (1 - \alpha)SX)} \right. \\
- \frac{SH(1 + SX)}{(1 + \alpha SX + (1 - \alpha)SH)(1 + \alpha SX)} \right]
\]
\[
= \frac{W}{2 \ln(2)} \frac{(2\alpha - 1)(2X - H + SX^2)SH(1 + SX)}{(1 + (1 - \alpha)SX + \alpha SH)(1 + (1 - \alpha)SX)}\times \frac{1}{(1 + \alpha SX + (1 - \alpha)SH)(1 + \alpha SX)}
\]
\[
= (2\alpha - 1)(2X - H + SX^2)Q
\]

where Q is always a positive quantity.

Therefore, \( \alpha = \frac{1}{2} \) is the only stationary point. If it is a maximum, it is optimum to use equal PSD's for the two users. If it is a minimum, then \( \alpha = 1 \) achieves the maximum, which means that it is optimum to use Frequency Division Signaling for the two users.

For \( \alpha > 0.5 \), the derivative of \( C^* \) with respect to \( \alpha \) will be negative if and only if \( 2X - H + SX^2 < 0 \). This implies that \( \alpha = \frac{1}{2} \) is a maximum if and only if \( S < \frac{H - 2X}{X^2} \). Similarly, \( \alpha = \frac{1}{2} \) is a minimum if and only if \( S > \frac{H - 2X}{X^2} \). Thus, since \( S = \frac{2P_{\text{max}}}{N_0 W} \), it follows that

\[
\frac{2P_{\text{max}}}{N_0 W} \overset{\text{eqpsd}}{\geq} \frac{H - 2X}{X^2} \quad (2.6)
\]

where eqpsd denotes using equal PSD's for the two users and fds denotes using Frequency Division Signaling.

Although the result in (2.6) was obtained under the assumption that \( S_1(f) \) and \( S_2(f) \) have the form described in (2.4) and (2.5), there are indications that achieving
the global maximum always involves using the same threshold test as in (2.6). For example, if the constraint that $S_1(f)$ and $S_2(f)$ have to be symmetric is relaxed, a new class of $S_1(f)$ and $S_2(f)$ can be defined as follows

$$
S_1(f) = \begin{cases} 
\alpha_1 \frac{2P_{mx}}{W} & \text{if } |f| \leq \frac{W}{2} \\
(1 - \alpha_1) \frac{2P_{mx}}{W} & \text{if } \frac{W}{2} < |f| \leq W \\
0 & \text{otherwise}
\end{cases} 
\quad (2.7)
$$

$$
S_2(f) = \begin{cases} 
\alpha_2 \frac{2P_{mx}}{W} & \text{if } |f| \leq \frac{W}{2} \\
(1 - \alpha_2) \frac{2P_{mx}}{W} & \text{if } \frac{W}{2} < |f| \leq W \\
0 & \text{otherwise}
\end{cases} 
\quad (2.8)
$$

That is, both $S_1(f)$ and $S_2(f)$ have the form shown in Figure 2.1, however they are not forced to be coupled, that is, $\alpha_2$ is not set to be equal to $1 - \alpha_1$, as in (2.5). In this case, when maximizing the sum of the capacities of users 1 and 2, it can be shown that $\alpha_1 = \alpha_2 = \frac{1}{2}$ is the only stationary point, and thus can be a maximum, a minimum or a saddle point. If it is a maximum, then using equal PSD's for the two users is optimal. If it is not a maximum, the maximum will be achieved at one of the corner solutions. That is, at $\alpha_1 = 1, \alpha_2 = 0$ or $\alpha_1 = 0, \alpha_2 = 1$, which correspond to using Frequency Division Signaling ($\alpha_1 = \alpha_2 = 0$ and $\alpha_1 = \alpha_2 = 1$ can easily be shown to not achieve the maximum). Therefore, only equal PSD's and FDS can be optimal, which leads to the same threshold test as in (2.6). Figure 2.2 plots the sum of the capacities of users 1 and 2, versus all possible combinations of $\alpha_1$ and $\alpha_2$, for a given channel, and for four different values of Signal-to-Noise Ratio (SNR). Notice how the point $\alpha_1 = \alpha_2 = \frac{1}{2}$ changes from a maximum to a saddle point, as the SNR increases.

As a second example, consider changing the form of $S_1(f)$ and $S_2(f)$, while keeping the constraint that $S_1(f)$ and $S_2(f)$ be symmetric. That is, consider the class of $S_1(f)$
Figure 2.2 $C_1 + C_2$ as a function of $\alpha_1$ and $\alpha_2$, for four different values of SNR.
and $S_2(f)$ defined as follows:

$$S_1(f) = \begin{cases} \frac{2P_{\text{max}}}{W} \frac{1}{1+\exp(\beta(|f|-\frac{W}{2}))} & \text{if } |f| \leq W \\ 0 & \text{otherwise} \end{cases}$$

$$S_2(f) = \begin{cases} \frac{2P_{\text{max}}}{W} \frac{1}{1+\exp(-\beta(|f|-\frac{W}{2}))} & \text{if } |f| \leq W \\ 0 & \text{otherwise} \end{cases}$$

where $0 \leq \beta \leq \infty$. See Figure 2.3.

![Figure 2.3](image)

**Figure 2.3** Simplified channel: how $S_1(f)$ changes with $\beta$

It can be shown that this class of $S_1(f)$ and $S_2(f)$ leads to the same threshold test as in (2.6). (See Appendix A for a proof)

### 2.3 Generalized Channel

Consider a channel that is bandlimited to $|f| \leq W$, however no longer restricted to being constant over this frequency range. The frequency range $0 \leq f \leq W$ is divided into $M$ equal-width bins ($M$ parallel independent subchannels). Within each bin, both $H_e(f)$ and $H_e(f)$ will be approximately constant, given that $M$ is sufficiently large. Also, any capacity-maximizing scheme will have to decide how much power to
allocate for each bin. Therefore, using the results from the previous section, it follows
that in the $i^{th}$ bin

\[
\frac{2P_i}{N_0D} \geq \frac{H_i - 2X_i}{X_i^2}
\]  

(2.9)

where $D = \frac{W}{M}$, $H_i = H_c(f_i)$, $X_i = H_x(f_i)$, $f_i = \frac{D}{2} + (i - 1)D$. ($f_i$ is the center
frequency of the $i^{th}$ bin)

Assume $H_c(f)$ is a decreasing function of $f$ and $H_x(f)$ is an increasing function
of $f$, as is the case for twisted pair channels. This implies that $\frac{H_c(f) - 2H_x(f)}{H_x^2(f)}$ is a
decreasing function of $f$. (See Figure 2.4)

Therefore, $\frac{H_i - 2X_i}{X_i^2}$ is a decreasing function of $i$, which implies that for some $i^*$,
$\frac{H_i - 2X_i}{X_i^2} < 0 \forall i \geq i^*$. Since $P_i \geq 0 \forall i$, it follows that $\frac{2P_i}{N_0D} > \frac{H_i - 2X_i}{X_i^2} \forall i \geq i^*$, which in
turn implies that Frequency Division Signaling is always optimum for all bins after
(and including) the $i^*$-th bin, according to (2.9).

**Conjecture 2.1**  Let

\[
|H_c(f)|^2 = \begin{cases} 
H_1 & \text{if } 0 \leq |f| \leq \frac{W}{2} \\
H_2 & \text{if } \frac{W}{2} < |f| \leq W \\
0 & \text{otherwise}
\end{cases}, \quad |H_x(f)|^2 = \begin{cases} 
X_1 & \text{if } 0 \leq |f| \leq \frac{W}{2} \\
X_2 & \text{if } \frac{W}{2} < |f| \leq W \\
0 & \text{otherwise}
\end{cases}
\]

Assume that there is a power constraint on the transmitted signal. Also
assume that $H_i - 2X_i > 0$ for $i = 1, 2$. Then, if $H_1 > H_2$ and $X_1 < X_2$,
it is always better to use equal PSD's in bin 1 and Frequency Division
Signaling in bin 2, than to use FDS in bin 1 and equal PSD's in bin 2.

Conjecture 2.1 has been proven for low and for high SNR's, and has been verified
computationally for all values of SNR. See Appendix A for a detailed analysis of this
Conjecture.
Figure 2.4 Channel and crosstalk transfer functions for twisted-pair model from paper by Aslanis and Cioffi
Returning to the M-bin scenario, assume that the capacity-maximizing scheme has decided how much power to put in each bin, taking into account (2.9). As was shown above, this scheme must use Frequency Division Signaling $\forall i \geq i^*$. Assume that as $i$ decreases from $i^*$ to 1, the strategy changes from Frequency Division to equal PSD's. That is, $\exists M_0$, in the range $1 \leq M_0 < i^*$, such that equal PSD's are used in the $M_0^{th}$ bin and Frequency Division is used in all bins to the right of the $M_0^{th}$ bin.

Conjecture 2.1 implies, by exchange arguments, that the capacity-maximizing scheme will use equal PSD's for all bins from 1 to $M_0$, and Frequency Division Signaling for bins $M_0 + 1$ to $M$. More formally:

Conjecture 2.2 Let $H_c(f)$ be a decreasing, and $H_e(f)$ an increasing function of $f$. Consider the class of all M-bin transmission schemes, that is, schemes that consist of dividing the given frequency range into $M$ bins and within each bin deciding to perform equal PSD's or Frequency Division Signaling. The capacity-maximizing scheme will use equal PSD's for bins 1 to $M_0$, and Frequency Division for bins $M_0 + 1$ to $M$, for some $0 \leq M_0 \leq M$.

Proof: First, it is convenient to make a few notational definitions. Let $e_i$ indicate that equal PSD's will be used in bin $i$, and $f_i$ indicate that Frequency Division Signaling will be used in bin $i$.

As was argued above, the capacity-maximizing scheme will have the following form

$$\ldots, e_{M_0}, f_{M_0+1}, f_{M_0+2}, \ldots, f_M$$

for some $0 \leq M_0 \leq M$. The transmission strategies for bins $M_0 - 1$ to 1 are not yet known. Assume that as $i$ decreases from $M_0$ to 1, the strategy changes from equal
PSD’s to Frequency Division Signaling. That is, assume that for some \( i_x, 1 \leq i_x < M_o \), the following is true about the capacity-maximizing scheme

\[
\ldots, \int_{i_x}, \epsilon_{i_x+1}, \epsilon_{i_x+2}, \ldots, \epsilon_{M_o-1}, \epsilon_{M_o}, \int_{M_o}, \int_{M_o+1}, \ldots, \int_{M}
\]  

(2.10)

Then, capacity can always be increased by switching the transmission strategy in bins \( i_x \) and \( M_o \), while keeping the total power in both bins the same. That is, the following scheme

\[
\ldots, \epsilon_{i_x}, \epsilon_{i_x+1}, \epsilon_{i_x+2}, \ldots, \epsilon_{M_o-1}, \int_{M_o}, \int_{M_o+1}, \int_{M_o+2}, \ldots, \int_{M}
\]

has a higher capacity due to Conjecture 2.1. However, this leads to a contradiction because the scheme described in (2.10) was assumed to maximize capacity. Thus, there is no \( i_x, 1 \leq i_x < M_o \) such that the transmission strategy changes from equal PSD’s to Frequency Division Signaling, which implies that the capacity-maximizing scheme will have the following form

\[
\epsilon_1, \epsilon_2, \ldots, \epsilon_{M_o-2}, \epsilon_{M_o-1}, \epsilon_{M_o}, \int_{M_o+1}, \int_{M_o+2}, \ldots, \int_{M}
\]

Q.E.D.

This result is intuitively reasonable: Let both users utilize the “good” parts of the channel (where crosstalk is low), and in the parts where the crosstalk becomes a problem use frequency division to eliminate it.

Therefore,

\[
C(P_{\text{max}}) = \max_{M_o, P_1, P_2 \text{ s.t. } P_1 + P_2 = P_{\text{max}}} C_{\text{eqpsd}}(M_o, P_1) + C_{\text{fds}}(M_o, P_2)
\]

(2.11)

where \( C_{\text{eqpsd}}(M_o, P_1) \) is the capacity of the channel when equal PSD’s are used for both users, and when the channel is bandlimited to \( M_oD \) (Hz) and the power constraint is \( P_1 \), and \( C_{\text{fds}}(M_o, P_2) \) is the capacity of the channel when Frequency Division
Signaling is used for both users, and when the channel is bandlimited from \( M_o D \) to \( W \) (Hz) and the power constraint is \( P_2 \). In this configuration, \( C_{\text{fd}}(M_o, P_2) \) corresponds to the capacity of a channel with only AWGN and with a transfer function that is an altered version of \( H_c(f) \), formed by taking away the parts of the spectrum that have been assigned to the other user.

Both \( C_{\text{expd}}(M_o, P_1) \) and \( C_{\text{fd}}(M_o, P_2) \) are well-studied and well-described in the literature [7, 1]. However, they have no closed-form expression. Therefore, the maximization cannot be performed analytically, only numerically. Note that \( C(P_{\text{max}}) \) is only a lower bound on the true capacity of the channel since we are constraining the form of \( S_1(f) \) and \( S_2(f) \).

### 2.4 Numerical Results and Discussion

The capacity \( C(P_{\text{max}}) \) was numerically calculated for two twisted-pair channels; one studied in [7] and one studied in [1].

1. The model of the twisted-pair channel used in [7] was

\[
|H_c(f)|^2 = \exp\left(-\alpha \sqrt{f}\right)
\]

where, \( \alpha = k \frac{l}{l_o} \), \( l \) =length of channel in ft, \( l_o \) =reference length= 18000ft, and \( k \) =constant of the physical channel= 1.158. Also,

\[
|H_z(f)|^2 = \beta f^{\frac{3}{2}}
\]

where, \( \beta = 10^{-9} \). In the above expressions, \( f \) refers to frequency in KHz.

The results obtained (for \( l = 600 \)) are shown in Figure 2.5.
Figure 2.5  Capacity calculations for twisted-pair model from paper by Irving Kalet: solid line corresponds to using the optimum transmission scheme (as described in this chapter), star-dot line corresponds to using equal PSD's for the two users (as assumed in Kalet's paper), and dashed line corresponds to using Frequency Division Signaling in all bins.
2. For the model of the twisted-pair channel used in [1],

\[ |H_c(f)|^2 = \theta \exp\left(-\alpha \sqrt{f}\right) \]  \hspace{1cm} (2.12)

where \( \alpha \) is defined as in the previous model, and \( \theta \) represents the attenuation at \( f = 0 \). Also,

\[ |H_c(f)|^2 = K f^{\frac{3}{2}} \]

where \( K = 10^{-13} \). In the above expressions, \( f \) refers to frequency in Hz.

It is important to note that a closed form expression for \( |H_c(f)|^2 \) was not given in [1]. However, the parameters \( \alpha \) and \( \theta \), in (2.12), were selected so that the resulting \( |H_c(f)|^2 \) matched the plot given in [1] as closely as possible. The fact that the model of \( |H_c(f)|^2 \) used in this example is not exactly equal to the one used in [1] is not important, since it is the approach to achieving capacity that is being tested, and this can accomplished using any model for \( |H_c(f)|^2 \). The results obtained are shown in Figure 2.6.

What is plotted in both Figures 2.5 and 2.6 is the capacity under three different transmission schemes. One uses equal PSD’s for all bins (same as in [7]), one uses Frequency Division in all bins, and the other is the optimal, that is, it uses equal PSD’s for the first \( M_o \) bins (for some \( 0 \leq M_o \leq M \)) and Frequency Division for the remaining bins. These correspond to the plots of \( C_{\text{eqpsd}}(M, P_{\text{max}}), C_{\text{fsd}}(0, P_{\text{max}}) \) and \( C(P_{\text{max}}) \), respectively, versus SNR.

There are a few observations that can be made on both figures. First, \( C(P_{\text{max}}) \) is larger than the other two for all values of SNR, as should be expected. More importantly, \( C(P_{\text{max}}) \) always increases with SNR, unlike \( C_{\text{eqpsd}}(M, P_{\text{max}}) \) which has a maximum value no matter how large the SNR becomes, as was also observed by [7] and [1].
Figure 2.6 Capacity calculations for twisted-pair model from paper by Aslanis and Cioffi: solid line corresponds to using the optimum transmission scheme (as described in this chapter), star-dot line corresponds to using equal PSD's for the two users (as assumed in Aslanis' and Cioffi's paper), and dashed line corresponds to using Frequency Division Signaling in all bins
Also, for low values of SNR $C(P_{\text{max}})$ approaches $C_{\text{eqpsd}}(M, P_{\text{max}})$, and for high values of SNR $C(P_{\text{max}})$ approaches $C_{\text{fds}}(0, P_{\text{max}})$. In fact, for low SNR the optimum $M_0$ in equation (2.11) approaches $M$, and for high SNR it approaches 0. That is, for low SNR it is optimum to use equal PSD's in all bins, and for high SNR it is optimum to use Frequency Division Signaling in all bins. This result should be expected because at low SNR, crosstalk is not an important source of interference, implying that both users can use all the frequency spectrum. On the other hand, at very high SNR, crosstalk becomes a problem at all frequencies, implying that it should be eliminated using Frequency Division.

Finally, if $C(P_{\text{max}})$ is compared to max $\{C_{\text{eqpsd}}(M, P_{\text{max}}), C_{\text{fds}}(0, P_{\text{max}})\}$, it can be seen that it is not significantly larger at any SNR. In fact, the largest improvement that $C(P_{\text{max}})$ offers over max $\{C_{\text{eqpsd}}(M, P_{\text{max}}), C_{\text{fds}}(0, P_{\text{max}})\}$ is roughly 20% to 30%, and this occurs at moderate values of SNR. Therefore, if a 30% difference is not very important to the designer, the transmission system can be designed by calculating only $C_{\text{eqpsd}}(M, P_{\text{max}})$ and $C_{\text{fds}}(0, P_{\text{max}})$ and using the scheme that offers higher capacity, thus saving computational time and effort.

It is important to discuss the results of [8], because it is the only work that questions the assumption that $S_1(f)$ has to be equal to $S_2(f)$ when deriving capacity of a twisted-pair channel. In this paper, it is assumed that $N_0 = 0$, that is, the only source of noise is NEXT. The capacity that is derived is

$$C_1 = W \log_2 \left( 1 + \left( \frac{H_c(f)}{H_x(f)} \right)_{f=f_0} \right)$$

(2.13)

where $f_0 \in (0, W)$ is such that

$$\int_0^W \log_2 \left( 1 + \left( \frac{H_c(f)}{H_x(f)} \right)^2 \right) df = W \log_2 \left( 1 + \left( \frac{H_c(f)}{H_x(f)} \right)_{f=f_0} \right)$$
It is clear that $C_1$, as defined in (2.13), is a finite quantity because $H_2(f) > 0 \forall f \in (0, W)$. According to [8], this formula for capacity holds with no power constraints on the transmitted signals.

However, from (2.9) it follows that if the frequency range $0 \leq f \leq W$ is divided into $M$ bins, then Frequency Division Signaling has to be used in all bins. Thus, each user faces a channel with no noise. Since [8] assumes no power constraints on the transmitted signals, the resulting capacity is infinite.

Therefore, while the true capacity is infinite, [8] finds a capacity that is not only finite, but only slightly better than the capacity when $S_1(f)$ is assumed to be equal to $S_2(f)$.

The reason for this discrepancy is that while [8] does not constrain $S_1(f)$ and $S_2(f)$ to be equal, it does constrain them to have the same support. In fact, it is recognized in the abstract and the introduction of [8] that "... NEXT can be completely eliminated if different bands are used for the two directions of transmission". It is not clear why [8] chose to constrain $S_1(f)$ and $S_2(f)$ to have the same support.

2.5 Implementation with Multicarrier Modulation

The optimum transmission scheme described in Conjecture 2.2 and equation (2.11) can be easily implemented using Multicarrier Modulation (MCM) [13,14,22–25]. The modulation will not be explained in detail here. The reader is referred to the sources cited above.

Very briefly stated, MCM consists of dividing the input data stream into parallel substreams and then modulating each substream on a different carrier frequency and sending the multiplexed signals through the channel. Thus, the transmitter can
shape the PSD of the signal by using the appropriate power for the modulators at each carrier frequency. A graphical illustration is included in Figure 2.7.

![Graphical illustration of Multicarrier Modulation](image)

Figure 2.7 Graphical illustration of Multicarrier Modulation

Therefore, it is straightforward, using Multicarrier Modulation, to implement the proposed scheme to achieve the lower bound on the capacity. The first $M_0$ carriers should be assigned to both users, while the remaining carriers should be assigned alternatingly to each user. Also, the appropriate power should be used for each carrier so that the signal PSD is as desired. It is important to note that the transmitted signal needs to be as Gaussian as possible, so that (2.1) is valid.

2.6 Possible Extension to More Than Two Users

The advantage of the proposed approach is that it is easily extendable to $N$ users ($N > 2$). The definition of equal PSD's remains the same, whereas Frequency Division Signaling now refers to giving each user $\frac{1}{N}$th of the bin. Although the problem becomes more complex than the $N = 2$ case, numerical calculations have shown that the same behavior is exhibited, as is roughly illustrated in Figure 2.8.
Figure 2.8  Capacity as a function of SNR when there are N twisted pairs in the same binder group: solid line represents using equal PSD's for the two users, and dashed line represents using Frequency Division Signaling in all bins

Only the capacity given equal PSD's in all bins and the capacity given Frequency Division in all bins are depicted in Figure 2.8. Finding the $M_0$-scheme is unnecessary for showing the relationship between the $N > 2$ and the $N = 2$ case. It can be seen that $C_{\text{fds}}(0, P_{\text{max}})$ becomes larger than $C_{\text{eqpsd}}(M, P_{\text{max}})$ just as the latter reaches its maximum. More importantly, $C_{\text{fds}}(0, P_{\text{max}})$ continues to increase as SNR increases. Therefore, for SNR's where $C_{\text{eqpsd}}(M, P_{\text{max}})$ is at its maximum, Frequency Division Signaling is a better option.
Chapter 3

Statistical Models for Main Sources of Interference on Copper Loops

Calculating channel capacity involves making some simplifying assumptions on the channel, to make the derivation tractable. The most important is the assumption that the only sources of noise are AWGN and NEXT. However, there are many sources of interference on the twisted pair channel. Some are dominant in one case, some are dominant in another. Some have been studied extensively, while others have not been modeled accurately. Some are neglected because of the assumption that they are insignificant or are reduced to negligible levels (as in the case of echo and echo cancelers). What follows is a list of what are generally considered the main sources of interference.

1. Interference from other twisted pairs in the same binder group
   
   (a) Near-end crosstalk (NEXT)
   
   (b) Far-end crosstalk (FEXT)

2. Impulse noise from electromechanical switching devices in the central office

3. Intersymbol Interference (ISI), caused by previously transmitted bits.

4. Additive White Gaussian noise (AWGN), which includes various sources of noise, such as thermal noise from the electronics used at the receiver.
It would be desirable to obtain statistical models for each source of interference, to enable the analysis of digital transmission systems over copper loops. Therefore, a more detailed description of each type of interference follows.

3.1 NEXT

In general, it is assumed that NEXT is a Wide-Sense Stationary process. The crosstalk transfer function for NEXT is modeled as $|H_{\text{NEXT}}(j\omega)|^2 = K_{\text{NEXT}} \times |\omega|^{1.5}$, where $K_{\text{NEXT}}$ is some constant which depends on the specific binder group [1,26–28]. The resulting Power Spectral Density (PSD) of NEXT is

$$S_{\text{NEXT}}(j\omega) = S_{\text{NS}}(j\omega) \times |H_{\text{NEXT}}(j\omega)|^2$$

where $S_{\text{NS}}(j\omega)$ is the PSD of the neighboring system.

It has been shown that the probability density function of NEXT can be modeled as Gaussian without substantial error [3,29], particularly if good shaping codes are used by the neighboring twisted pairs [1,30].
Therefore, assuming that $K$ consecutive samples of the noise process are taken,

$$p_{\text{NEXT}}(x) = \frac{1}{\sqrt{2\pi K_{\text{NEXT}}}} \exp \left( -\frac{x^T K_{\text{NEXT}}^{-1} x}{2} \right) \quad \forall x \in \mathbb{R}^K \quad (3.1)$$

where $K_{\text{NEXT}}$ is the covariance matrix of the $K$ samples of NEXT and reflects the correlation between samples, which in turn reflects the PSD of NEXT (see Appendix B).

### 3.2 FEXT

In general, FEXT is also assumed to be a Wide-Sense Stationary process. The crosstalk transfer function for FEXT is modeled as $|H_{\text{FEXT}}(j\omega)|^2 = K_{\text{FEXT}} \times |C(j\omega)|^2 \times |\omega|^2$, where $C(j\omega)$ is the loss of the cable [1, 26]. The resulting PSD of FEXT is

$$S_{\text{FEXT}}(j\omega) = S_{\text{NS}}(j\omega) \times |H_{\text{FEXT}}(j\omega)|^2$$

where $S_{\text{NS}}(j\omega)$ is the PSD of the neighboring system.

As with NEXT, FEXT is also modeled as a Gaussian process. Therefore, assuming that $K$ consecutive samples of the noise process are taken,

$$p_{\text{FEXT}}(x) = \frac{1}{\sqrt{2\pi K_{\text{FEXT}}}} \exp \left( -\frac{x^T K_{\text{FEXT}}^{-1} x}{2} \right) \quad \forall x \in \mathbb{R}^K \quad (3.2)$$

where $K_{\text{FEXT}}$ is the covariance matrix of the $K$ samples of NEXT and reflects the correlation between samples, which in turn reflects the PSD of FEXT (see Appendix B).

Usually FEXT is ignored because it has a much smaller effect than NEXT [1, 28]. However, in some cases (e.g. ADSL (Asymmetric Digital Subscriber Lines)), there is no NEXT, so FEXT must be considered [2, 31].
3.3 Impulse noise (IN)

The impulse noise on copper loops is not very well characterized [32,1], even though it has been described as a very important source of noise. One of the most important models is [33–35]

\[
N_{IN}(t) = \sum_{i=-\infty}^{\infty} U_i h(t - t_i)
\]

(3.3)

where \(U_i\)'s are random amplitudes, distributed with hyperbolic distribution, \(h(t)\) is the impulse response of the channel and \(t_i\)'s are the arrival times of a Poisson Process.

Impulse events occur about one to two times per second. Thus, the intensity \(\lambda\) of the Poisson Process is approximately equal to 0.025 events per second. Also, the average duration of an impulse is approximately 100\(\mu\)s. These figures are only approximate and also time-varying [32].

Both [34] and [35] derive the one-dimensional probability density function of IN, which is the density of one sample of the noise at time \(t\). It is of interest to know the \(K\)-dimensional pdf of IN, that is, the pdf of \(K\) consecutive samples of IN at times \(t, t + T_s, \ldots, t + (K-2)T_s, t + (K-1)T_s\). Let \(L\) be the average duration of an impulse (in seconds). Under the assumption that \(\lambda L \ll 1\) and \(KT_s \ll L\), it can be shown that a good model for the \(K\)-dimensional density of IN is

\[
p_{IN}(x) = \alpha p_1(x) + (1 - \alpha) p_2(x)
\]

(3.4)

where \(\alpha = 1 - \lambda L\), \(p_1(x) = \delta(x)\), \(p_2(x) = \frac{1}{\sqrt{2\pi K_{IN}}} \exp \left(-\frac{x^T K_{IN}^{-1} x}{2}\right)\) and \(K_{IN}^{-1}\) is the covariance matrix of an impulse event and depends on \(h(t)\) (see Appendix B).

As mentioned above, \(\lambda \approx 0.025\) and \(L \approx 10^{-4}\). Therefore, \(\lambda L \approx 2.5 \cdot 10^{-6}\), which implies that the assumption \(\lambda L \ll 1\) is valid. Assuming \(L \ll \frac{1}{\lambda}\) corresponds to assuming that the probability of two impulse events overlapping is negligible (see Appendix B). Also, \(KT_s\) is equal to the duration of one symbol (in seconds). Thus the
assumption $KT_s \ll L$ corresponds to assuming that the symbol period is much smaller
than the duration of an impulse event. This, of course, depends on the transmission
system. However, for high data rate systems (which are the ones that are of most
interest), the symbol period is on the order of a microsecond, and therefore much
smaller than the average duration of an impulse. Thus, both assumptions that lead
to equation (3.4) are valid.

3.4 Intersymbol Interference (ISI)

When the channel does not have a flat frequency response, the signals sent through the
channel are distorted and are spread out into the following bits, causing interference.
To see this effect mathematically, assume BPSK transmission and let

$$s(t) = \sum_{n=-\infty}^{\infty} b_n s_p(t - nT)$$  \hspace{0.5cm} (3.5)

where $T$ is the symbol period, $s_p(t)$ is the pulse shape (with support on $t \in [0, T]$),
and $b_n \in \{-1, 1\}$

To better illustrate the effects of ISI, assume that the channel has no other source
of noise. Therefore, given that the channel impulse response is $h(t)$, the received
signal will be

$$r(t) = s(t) * h(t)$$

$$= \sum_{n=-\infty}^{\infty} b_n \left[ s_p(t - nT) * h(t) \right]$$

$$= \sum_{n=-\infty}^{\infty} b_n \tilde{s}_p(t - nT)$$

where $\ast$ denotes convolution. As a result, $\tilde{s}_p(t)$ will have support on $t \in [0, \infty)$.
Therefore, $\tilde{s}_p(t - nT)$ will be non-zero only for $t \in [nT, \infty)$ which implies that for
\[ t \in [iT, (i + 1)T) \]

\[
\begin{align*}
    r(t) &= \sum_{n=-\infty}^{i} b_n \tilde{s}_p(t - nT) \\
    &= b_i \tilde{s}_p(t) + \sum_{n=-\infty}^{i-1} b_n \tilde{s}_p(t - nT)
\end{align*}
\]

Thus, the \( i^{th} \) bit will experience interference from all bits preceding it, and the noise will be equal to \( \sum_{n=-\infty}^{i-1} b_n \tilde{s}_p(t - nT) \). See Figure 3.2 for a graphical illustration of ISI.

![Figure 3.2 Illustration of ISI: signal amplitude as a function of time. Solid line: transmitted signal, dashed line: received signal](image)

What is of interest, for the purposes of analyzing the performance of the loop, is the density of ISI. Unfortunately, this is not readily obtained, particularly when the received vector has length \( K > 1 \) (that is, when \( K \) samples per symbol are used for the detection of that symbol). One of the problems is the fact that the density of ISI is not Gaussian. When \( K > 1 \), a \( K \)-dimensional density is needed, and \( K \)-dimensional non-Gaussian densities are extremely difficult to model. In fact, very few non-Gaussian multidimensional (joint) densities are known.
3.4.1 Numerical calculation of pdf

The problem becomes slightly easier to solve when $K = 1$. That is, when the receiver makes its decision based upon one sample per symbol. This is true for a wide class of systems, including ISDN which uses 2B1Q transmission (4-level PAM) and one sample per symbol to detect which symbol was transmitted. In this case, although in general the density has no closed-form expression, the characteristic function is analytic. The characteristic function of a random variable $X$ is defined as

$$\Phi_X(u) = E[e^{juX}]$$  \hspace{1cm} (3.6)

To demonstrate how the characteristic function of ISI is obtained, ISDN will be used as an example. The transmitted pulse for ISDN is of the form seen in Figure 3.3

![Transmitted pulse for ISDN (Approximate)](image)

At the receiver, the received signal is passed through a low-pass filter which does not change the shape of the pulse. The output of the filter is sampled once every symbol period, at the maximum of the pulse. The receiver then makes its decision based solely upon the value of this sample.
As can be seen from Figure 3.3, the pulse is non-zero at subsequent sampling times, thus causing interference for the following symbols. Let the values of the pulse at subsequent sampling times be $\alpha_1, \alpha_2, \alpha_3, \ldots$

Considering the $i^{th}$ symbol, it is clear that the value of the interference from the $(i - 1)^{st}$ symbol equals $\alpha_1 Y_{i-1}$, where $Y_{i-1}$ is a random variable which corresponds to the value by which the transmitted pulse is multiplied when the transmitted symbol equals the $(i - 1)^{st}$ symbol. For example, in the case of ISDN, the variable $Y_i$ has probability mass function (pmf)

$$p_{Y_i}(y) = \begin{cases} \frac{1}{4} & \text{if } y \in \{-3A, -A, A, 3A\} \\ 0 & \text{otherwise} \end{cases} \quad \forall i$$

Also, the value of the interference from the $(i - 2)^{nd}$ symbol equals $\alpha_2 Y_{i-2}$. In general, the value of the interference from the $(i - n)^{th}$ symbol equals $\alpha_n Y_{i-n}$. Therefore, if $N_{\text{intr}}^i$ denotes the interference during the $i^{th}$ symbol, caused by all previous symbols, then

$$N_{\text{intr}}^i = \sum_{n=1}^{\infty} \alpha_n Y_{i-n} \quad (3.7)$$

The $Y_n$'s are independent and identically distributed (i.i.d.). Therefore,

$$\Phi_{N_{\text{intr}}} (u) = \prod_{n=1}^{\infty} \Phi_{\alpha_n Y_n} (u) \quad (3.8)$$

by an elementary property of characteristic functions. In the case of BPSK transmission,

$$p_{Y_n}(y) = \begin{cases} \frac{1}{2} & \text{if } y \in \{-1, 1\} \\ 0 & \text{otherwise} \end{cases} \quad \forall n$$

$$\Rightarrow \quad \Phi_{\alpha_n Y_n} (u) = \cos(\alpha_n u)$$

$$\Rightarrow \quad \Phi_{N_{\text{intr}}} (u) = \prod_{n=1}^{\infty} \cos(\alpha_n u)$$
In the case of 2B1Q transmission,

\[ p_{Y_n}(y) = \begin{cases} \frac{1}{4} & \text{if } y \in \{-3A, -A, A, 3A\} \\ 0 & \text{otherwise} \end{cases} \quad \forall n \]

\[ \Rightarrow \quad \Phi_{N_{ISI}}(u) = \prod_{n=1}^{\infty} \frac{1}{2} \cos(\alpha_n u A) + \frac{1}{2} \cos(3\alpha_n u A) \]

Therefore, the characteristic function of ISI has a closed-form expression in two very widely used transmission schemes. What is desired is the pdf of ISI, and this is simply the inverse Fourier Transform of the characteristic function. Unfortunately, in most cases the pdf has no closed-form expression. However, it is possible to numerically compute the Inverse Fourier Transform and thus have a numerical approximation for the pdf.

Thus,

\[ p_{isi}(x) \approx \text{IFFT}(\Phi_{N_{ISI}}(u))[i_x] \tag{3.9} \]

where IFFT is the inverse Fast Fourier Transform, and \( i_x \) is the index in the vector \( \text{IFFT}(\Phi_{N_{ISI}}(u)) \) that corresponds to \( x \). See also [36] for an alternative approach for numerically calculating the pdf of ISI.

### 3.5 Additive white Gaussian noise (WGN)

Several sources of noise, including thermal noise from the electronics at the receiver, quantization noise from the A/D converters and residual echo noise from echo cancelers, are treated as a single source of noise, which is modeled as a zero-mean Gaussian process with a flat Power Spectral Density. Briefly stated, the reason for the Gaussian assumption are the many independent sources of noise which, according to the Central Limit Theorem, produce a noise component that is approximately Gaussian. Therefore,

\[ p_{wgn}(x) = \frac{1}{\sqrt{2\pi K_{wgn}}} \exp \left( -\frac{x^T K^{-1}_{wgn} x}{2} \right) \quad \forall x \in \mathbb{R}^K \tag{3.10} \]
where $K_{wgn} = \sigma_{wgn}^2 I$, and $I$ is a $K \times K$ identity matrix.
Chapter 4

Performance of Digital Transmission over Copper Loops

As mentioned in Chapter 1, it is not only important to know the maximum rate at which information can be sent through a channel, it is also important to know which transmission system achieves capacity. Due to the fact that, in general, the capacity-achieving transmission system is not known, it becomes necessary to be able to evaluate the performance of proposed suboptimum systems.

Importance Sampling theory [18, 19, 37, 20, 38, 39, 21] will be applied to digital transmission over twisted pairs, to enable a fast and accurate evaluation of proposed systems. Therefore, a brief description of this technique follows. For a more detailed analysis, see references cited above.

4.1 Importance Sampling

During Monte Carlo simulations, the transmission of a large number of bits is simulated. The errors that occur at the receiver are counted and divided by the total number of bits sent. This gives an estimate of the BER. However, if the system being simulated has a BER of $10^{-6}$, that is, on average one error occurs for every million bits sent, then at least ten million bits have to be transmitted to obtain an accurate estimate. The idea behind Importance Sampling is to emphasize the “important” regions of the channel noise distribution, that is, the regions that produce errors. In other words, errors are made more often, so that it is not necessary to simulate the
transmission of a large number of bits to achieve a statistically significant number of errors. Due to the fact that errors now occur more often, it becomes necessary to weigh each error in such a way that the estimate of the BER is unbiased and has a small variance. The bit error rate estimate from Monte Carlo simulations can be expressed as

\[ \hat{P}_{\text{MC}} = \frac{1}{M_{\text{MC}}} \sum_{i=0}^{M_{\text{MC}}} I(r_i) \]  

(4.1)

where \( r_i \) is the received vector for bit \( i \) and

\[ I(r_i) = \begin{cases} 1 & \text{if } r_i \text{ results in an error} \\ 0 & \text{otherwise} \end{cases} \]

while the Importance Sampling estimate is expressed as

\[ \hat{P}_{\text{IS}} = \frac{1}{M_{\text{IS}}} \sum_{i=0}^{M_{\text{IS}}} W(r_i^*) I(r_i^*) \]  

(4.2)

where the received vector \( r_i^* \) is generated using a probability density function (pdf) that results in more frequent errors. This density is called the biasing density and is denoted by \( p_{R^*}(r^*) \). Also, \( W(r_i^*) \) is a weighting function that makes the estimate unbiased. It has been shown [18] that the correct choice for \( W(r_i^*) \) is

\[ W(r_i^*) = \frac{p_R(r_i^*)}{p_{R^*}(r_i^*)} \]  

(4.3)

where \( p_R(r) \) is the actual pdf of the received vector in the system that is being simulated.

Define \( \Gamma = \frac{M_{\text{MC}}}{M_{\text{IS}}} \) where \( M_{\text{MC}} \) and \( M_{\text{IS}} \) are the number of trials that achieve the same variance of \( \hat{P}_e \) for Monte Carlo and Importance Sampling simulations, respectively. So, for a given accuracy of the estimate of the BER, the gain \( \Gamma \) shows the relation between the number of simulation runs that need to be performed with Importance Sampling versus the number of simulation runs needed with Monte Carlo simulation.
It can be shown [18] that
\[
\text{var} \left( \hat{P}_{\text{MC}} \right) = \frac{P_e - P_e^2}{M_{\text{MC}}}, \quad \text{and} \quad \text{var} \left( \hat{P}_{\text{IS}} \right) = \frac{\overline{W} - P_e^2}{M_{\text{IS}}}
\]
where \( \overline{W} = E_{p_R} [W(R)/I(R)] \). Therefore,
\[
\Gamma = \frac{P_e - P_e^2}{\overline{W} - P_e^2}
\]
Clearly, it is desirable to maximize \( \Gamma \), and the only degree of freedom available is the choice of \( p_{R_\star} \). Thus, the main problem in applying Importance Sampling to digital transmission systems is determining which \( p_{R_\star} \) maximizes \( \Gamma \), or alternatively minimizes \( \overline{W} \).

It has been shown [18] that, for a wide class of systems, letting \( p_{R_\star} \) be a linearly shifted version of \( p_R \) is sufficient to produce a an estimate which has smaller variance than the Monte Carlo estimate. The problem therefore reduces to finding the optimal linear shift that minimizes \( \overline{W} \). Other methods for obtaining a \( p_{R_\star} \) include increased variance and exponential twisting [38].

4.1.1 Conditional Importance Sampling

To implement Importance Sampling as described above, it is necessary to know the probability density function of the noise in the channel. In certain cases though, the noise consists of two parts, one with known density and one with unknown density. That is,
\[
n(t) = n_1(t) + n_2(t)
\]
where \( p_{n_1} \) is known and \( p_{n_2} \) is not known or has no closed-form expression. In such a case, a technique known as Conditional Importance Sampling [37] will be used.

The main idea is to perform a simulation where \( n_1 \) is produced using a biased pdf, and \( n_2 \) is generated using its original pdf, just as in the case of Monte Carlo
simulations. This is equivalent to performing a mixed Importance Sampling and Monte Carlo simulation. Of course, the weighting function $W(r_i^*|n_2)$ has to be modified from what is given in (4.3), so that it will result in an unbiased estimate. Therefore,

$$W(r_i^*|n_2) = \frac{p_{n_2|n_2}(r_i^*|n_2)}{p_{n_2|n_2}(r_i^*|n_2)}$$ (4.4)

and

$$\hat{P}_{cis} = \frac{1}{M_{cis}} \sum_{i=0}^{M_{cis}} W(r_i^*|n_2) I(r_i^*)$$ (4.5)

4.1.2 Example

To illustrate how effective Importance Sampling is, a typical example follows. The system considered is simple BPSK transmission over an AWGN channel. The biasing density is a shifted version of the original density. The simulation uses $K$ samples per bit, white Gaussian noise with variance $\sigma^2$ and a constant power for transmission. Therefore, the only variables are $K$ and $\sigma^2$. It can easily be shown that $P_e = Q(\sqrt{\frac{K}{\sigma^2}})$, where $Q$ is the complementary distribution function for Gaussian noise with zero mean and unit variance. The exact value of $P_e$ is then compared to the estimated value, for several combinations of $K$ and $\sigma^2$. As expected, the results were promising. A small sample of the results follows.

<table>
<thead>
<tr>
<th>$P_e$</th>
<th>$\hat{P}_{cis}$</th>
<th># of bits used in Importance Sampling simulation</th>
<th># of bits required for Monte Carlo simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.69e-6</td>
<td>3.87e-6</td>
<td>500</td>
<td>27.1e6</td>
</tr>
<tr>
<td>4.11e-9</td>
<td>3.88e-9</td>
<td>650</td>
<td>24.4e9</td>
</tr>
<tr>
<td>6.88e-24</td>
<td>7.62e-24</td>
<td>1000</td>
<td>14.5e24</td>
</tr>
</tbody>
</table>
Clearly, Importance Sampling is a very efficient way of obtaining an estimate of the BER. Its application requires knowledge of the statistics of the noise in the system. It is by using this knowledge (and "embedding" it in $W(r^*)$) that it is possible to drastically reduce the number of simulation runs versus Monte Carlo simulation.

### 4.2 Efficient Simulation of Digital Copper Loops

Using the models for the main sources of interference given in Chapter 3, Importance Sampling theory is applied to two cases of digital transmission over copper loops. One case involves a channel with flat frequency response, and with noise that consists of NEXT, impulse noise and AWGN. The other case is a simplified version of ISDN. That is, the channel has non-flat transfer function and the noise consists of ISI and AWGN.

#### 4.2.1 General Twisted Pair Channel

The channel is assumed to have a flat frequency response and noise consisting of NEXT, impulse noise and AWGN. The receiver is assumed to use $K$ samples per symbol and a matched-filter detector. The transmission system being tested is BPSK (Binary Phase Shift Keying).

First, the pdf of the noise needs to be determined.

$$N(t) = N_{\text{NEXT}}(t) + N_{\text{IN}}(t) + N_{\text{WGN}}(t)$$

The various sources of noise are statistically independent. Therefore,

$$p_N(x) = p_{\text{NEXT}}(x) \ast p_{\text{IN}}(x) \ast p_{\text{WGN}}(x)$$
where $*$ denotes convolution. Using the models from Chapter 3, the resulting density is

$$p_n(x) = \alpha \frac{1}{\sqrt{2\pi|K_1|}} \exp \left( -\frac{x^T K_1^{-1} x}{2} \right) + (1 - \alpha) \frac{1}{\sqrt{2\pi|K_2|}} \exp \left( -\frac{x^T K_2^{-1} x}{2} \right) \quad (4.6)$$

where $K_1 = K_{\text{NEXT}} + K_{\text{WGN}}$, $K_2 = K_{\text{NEXT}} + K_{\text{IN}} + K_{\text{WGN}}$, and $K_{\text{NEXT}}, K_{\text{IN}}$ and $K_{\text{WGN}}$ are defined in Chapter 3.

Under BPSK, one of two signals, $S_0(t)$ and $S_1(t)$, is transmitted. These signals correspond to the bits 0 and 1, respectively. Also, $S_0(t) = -S_1(t)$. Therefore, the probability of error is given by

$$P_e = Pr(\text{receiver decides } S_1 | S_0 \text{ was sent}) Pr(S_0 \text{ was sent})$$

$$+ Pr(\text{receiver decides } S_0 | S_1 \text{ was sent}) Pr(S_1 \text{ was sent})$$

where $Pr(\text{event})$ refers to the probability that event will occur. Assuming

$$Pr(S_0 \text{ was sent}) = Pr(S_1 \text{ was sent}) = \frac{1}{2}$$

the above expression for $P_e$ can be re-written as

$$P_e = Pr(\text{receiver decides } S_1 | S_0 \text{ was sent})$$

since the noise density is symmetric around the origin and the detector is symmetric. The extension to non-symmetric detectors is straightforward and will not be discussed here. Given that $S_0(t)$ is transmitted,

$$p_n(r) = p_n(r - S_0)$$

where $S_0$ is a vector whose elements are the $K$ samples of $S_0(t)$ used by the receiver to detect the current bit.

In order to apply Importance Sampling to the above transmission system, a biasing density needs to be chosen. The biasing is two-fold. It consists of changing the mean
and using a different value for \( \alpha \). That is,

\[
p_{R}^{*}(r) = \alpha^{*} \frac{1}{\sqrt{2\pi|K_1|}} \exp \left( -\frac{r^T K_1^{-1} r}{2} \right) + (1 - \alpha^{*}) \frac{1}{\sqrt{2\pi|K_2|}} \exp \left( -\frac{r^T K_2^{-1} r}{2} \right)
\]

(4.7)

The density of the received vector is shifted so that it has a mean that is halfway between \( S_0 \) and \( S_1 \). This choice of linear shift was proven to be asymptotically optimum when \( K = 1 \) and shown to be robust for any value of \( K \) and for small \( P_e \) [18]. The resulting density has zero mean since \( S_0(t) = -S_1(t) \).

The value \( 1 - \alpha \) equals the probability that the current bit is in the middle of an impulse event. Thus, using \( \alpha^{*} \) instead of \( \alpha \), corresponds to changing the probability of occurrence of impulse events. Usually, an impulse is a very rare event, with \( \alpha \approx 1 \). Importance Sampling theory suggests that it is necessary to make rare events more often. Therefore, \( \alpha \) should be decreased. What is desired is the value of \( \alpha^{*} \) that minimizes the variance of the estimate, or equivalently, the value that minimizes \( \bar{W} \).

\[
\bar{W} = \mathbb{E}_{p_{R}^{*}}[W(R)I(R)] = \int_{R^K} \frac{p_{R}(r)}{p_{R}^{*}(r)} p_{R}^{*}(r) dr
\]

(4.8)

The analytical calculation of \( \bar{W} \) is not feasible, given the expressions for \( p_{R} \) and \( p_{R}^{*} \) shown above. However, it is possible to estimate \( \bar{W} \) using Importance Sampling. To see this, let

\[
\hat{W} = \frac{1}{M} \sum_{i=0}^{M} W(r_i^*) I(r_i^*)
\]

(4.9)

It follows that

\[
\mathbb{E} \left[ \hat{W} \right] = \mathbb{E}_{p_{R}^{*}}[W^2(R)I(R)] = \int_{R^K} \left( \frac{p_{R}(r^*)}{p_{R}^{*}(r^*)} \right)^2 p_{R}^{*}(r^*) dr^*
\]

\[
= \int_{R^K} \frac{p_{R}(r^*)}{p_{R}^{*}(r^*)} p_{R}(r^*) dr^*
\]

\[
= \bar{W}
\]
Equation (4.9) implies that to obtain an estimate of $\bar{W}$, it is sufficient to perform a regular Importance Sampling simulation of the system, with the slight modification of weighing the errors by $W^2(r)$, instead of $W(r)$. This method is used to obtain estimates of $\bar{W}$ for various values of $\alpha^*$. Although the optimum value for $\alpha^*$ is different for different channel setups, the value $\alpha^* = 0.5$ was found to be robust in the sense that it results in a $\bar{W}$ that is very close to that given by the optimum $\alpha^*$. Thus, $\alpha^* = 0.5$ is used in all Importance Sampling simulations of this system. Figure 4.1 shows plots of $\hat{W}$ versus $\alpha^*$, for various values of $K$ and $P_e$.

What is desired is an indicator of how much is gained, in terms of reduction in the required number of simulation runs, by using Importance Sampling versus Monte Carlo simulation. The Importance Sampling gain $\Gamma$, as described in Section 4.1, is such an indicator.

$$\Gamma = \frac{P_e - P_e^2}{\bar{W} - P_e^2}$$  \hspace{1cm} (4.10)

where $\bar{W}$ is given by equation (4.8). As mentioned above, the analytical calculation of $\bar{W}$ is not possible, given the noise density and biasing density of this system.

One approach for obtaining an analytical result is to find upper and lower bounds for $\bar{W}$. Fortunately, for the case $K = 1$, it is possible to find upper and lower bounds for $\bar{W}$ that are within a factor of two from each other (see Appendix C). These bounds lead to upper and lower bounds for $\Gamma$, according to equation (4.10). Figure 4.2 shows plots of the bounds for $\Gamma$ as a function of $P_e$, for different channel setups. It is clear that Importance Sampling is a far better option than Monte Carlo simulation, especially for small values of $P_e$.

Simulations of BPSK transmission were performed using the channel model given above. It was desirable to test the validity of the models for the probability density functions of the various sources of noise, as described in Chapter 3. The best option
Figure 4.1 \( \tilde{W} \) vs \( \alpha^* \) for the Importance Sampling simulation of the general twisted pair channel. The four plots correspond to different channel setups.
Figure 4.2 Upper and lower bounds on the Importance Sampling Gain, $\Gamma$, versus $P_e$. The four plots correspond to different setups of the channel.
would have been to compare estimated BER's with actual real-world data. However, this data was not available. Due to the fact that an analytical evaluation of the BER is not readily obtained, the only remaining option is to perform Monte Carlo simulations of the same system. It is essential to note that to test the validity of the models for the densities, it was necessary to perform Monte Carlo simulations in which the noise was produced in a way that emulated the real-world situation, and not according to the models in question. A sample of results is included in Table 4.2.

<table>
<thead>
<tr>
<th>$P_{cis}$</th>
<th>$P_{cmc}$</th>
<th># of bits used in Importance Sampling simulation</th>
<th># of bits used in Monte Carlo simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.055</td>
<td>0.0433</td>
<td>100</td>
<td>3000</td>
</tr>
<tr>
<td>0.0173</td>
<td>0.0157</td>
<td>100</td>
<td>3000</td>
</tr>
<tr>
<td>2.98e-4</td>
<td>2.33e-4</td>
<td>100</td>
<td>30000</td>
</tr>
<tr>
<td>3.19e-4</td>
<td>3.67e-4</td>
<td>100</td>
<td>30000</td>
</tr>
<tr>
<td>2.16e-5</td>
<td>2.45e-5</td>
<td>100</td>
<td>100000</td>
</tr>
</tbody>
</table>

### 4.2.2 An ISDN Example

The channel is assumed to have a non-flat frequency response and noise consisting of ISI and AWGN. The receiver is assumed to use 1 sample per symbol, which is compared to threshold values to detect the transmitted symbol. The transmission system used is 2B1Q.

This modulation refers to 4-level PAM (Pulse Amplitude Modulation). That is, the transmitted signals, each corresponding to one of four symbols, are $-3Ag(t)$, $-Ag(t)$, $Ag(t)$ and $3Ag(t)$, where $g(t)$ is the unit pulse. It is easily verified that, under the assumption of channel noise with a density that is symmetric around the
origin,

\[ P_e^{2B1Q} = \frac{3}{2} P_e^{BPSK} \]  

(4.11)

when \( S_1(t) = A g(t) \). Therefore, the performance of this transmission system is estimated by performing a modified BPSK simulation, with \( S_1(t) = A g(t) \) and with a correction factor of \( \frac{3}{2} \) for \( P_e \).

Due to the fact that ISI does not have a closed-form expression for its pdf, two approaches are used. Importance Sampling with numerical calculation of pdf, and Conditional Importance Sampling.

As shown in Chapter 3, the characteristic function of ISI, in the case of 2B1Q transmission, is given by

\[ \Phi_{N_{ISI}}(u) = \prod_{n=1}^{\infty} \left( \frac{1}{2} \cos(\alpha_n u A) + \frac{1}{2} \cos(3\alpha_n u A) \right) \]

where \( \alpha_n \) corresponds to the rate of decay of the received pulse.

Thus,

\[ p_{ISI}(x) \approx \text{IFFT}(\Phi_{N_{ISI}}(u))[i_x] \]  

(4.12)

where IFFT is the inverse Fast Fourier Transform, and \( i_x \) is the index in the vector IFFT\( (\Phi_X(u)) \) that corresponds to \( x \). The pdf of the total noise may now be numerically obtained.

\[ N(t) = N_{ISI}(t) + N_{WGN}(t) \]

The two sources of noise are statistically independent. Therefore,

\[ p_N(x) = p_{ISI}(x) * p_{WGN}(x) \]  

(4.13)

where \( * \) denotes convolution. Using equation (4.12) and the expression for \( p_{WGN}(x) \) given in Chapter 3, \( p_N(x) \) is obtained by numerical convolution.

As discussed above, a slightly modified BPSK simulation was used to obtain estimates of the BER of this system. In subsection 4.2.1 it is shown that, under BPSK,
the probability of error can be written as

\[ P_e = Pr(\text{receiver decides } S_1 | S_0 \text{ was sent}) \]

Given that \( S_0(t) \) is transmitted,

\[ p_N(r) = p_N(r - S_0) \]

The biasing density \( p_N^*(r) \) is chosen to be a linearly shifted version of the original density. That is,

\[ p_N^*(r) = p_N(r) \]

Obtaining \( p_N(r) \) numerically may not be satisfactory because of possible numerical problems, or may not be applicable (as in the case \( K > 1 \)). Therefore, a second approach is taken, where the density of ISI is assumed to be unknown, resulting in the need to use Conditional Importance Sampling (see section 4.1). Thus,

\[ \hat{P}_{\text{cIS}} = \frac{1}{M_{\text{cIS}}} \sum_{i=0}^{M_{\text{cIS}}} W(r_i^*|N_{\text{ISI}})I(r_i^*) \]  \hspace{1cm} (4.14)

where,

\[ W(r_i^*|N_{\text{ISI}}) = \frac{p_{R|\text{ISI}}(r_i^*|N_{\text{ISI}})}{p_{N^*|\text{ISI}}(r_i^*|N_{\text{ISI}})} \]

and

\[ p_{N^*|\text{ISI}}(r_i^*|N_{\text{ISI}}) = p_{\text{WON}}(r) \]

This choice of \( p_{N^*|\text{ISI}}(r_i^*|N_{\text{ISI}}) \) resulted from applying the ideas of [37] to the case under consideration.

Simulations of 2B1Q transmission were performed using the channel model given above. Both approaches to Importance Sampling were used. That is, one where the numerical density of ISI is used, and another where Conditional Importance Sampling is used. Various setups of the channel were simulated, and for each setup, three
Table 4.3 Simulation results for 2B1Q transmission over a non-flat channel with ISI and AWGN

<table>
<thead>
<tr>
<th>$\hat{P}_{\text{cis}}$</th>
<th>$\hat{P}_{\text{cis}}$</th>
<th>$\hat{M}_{\text{MC}}$</th>
<th>$M_{\text{cis}}$</th>
<th>$M_{\text{MC}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3e-3</td>
<td>2.1e-3</td>
<td>4.8e-3</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>4.6e-3</td>
<td>4.71e-3</td>
<td>4.5e-3</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>1.4e-3</td>
<td>3.0e-3</td>
<td>2.5e-3</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>6.78e-4</td>
<td>5.6e-4</td>
<td>6.0e-4</td>
<td>200</td>
<td>400</td>
</tr>
<tr>
<td>7.1e-5</td>
<td>1.18e-58</td>
<td>9.0e-5</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>7.8e-5</td>
<td>5.28e-5</td>
<td>8.0e-5</td>
<td>200</td>
<td>4000</td>
</tr>
<tr>
<td>1.3e-5</td>
<td>1.6e-5</td>
<td>2.0e-5</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>

estimates were obtained: $\hat{P}_{\text{cis}}$, $\hat{P}_{\text{cis}}$, and $\hat{P}_{\text{MC}}$. Table 4.3 includes a selection of results from various simulations.

It should be noted that Importance Sampling using a numerical approximation of the noise density, and Conditional Importance Sampling need to be used with care. The former may result in numerical errors when the noise with known density dominates the noise with unknown density. On the other hand, the latter results in an estimate with extremely large variance when the noise with unknown density dominates the noise with known density. Although not shown explicitly in Table 4.3, examples of such cases are shown on rows 3, 5 and 6. Therefore, when simulating a system which has a source of noise with unknown density, it is necessary to establish which source of noise is dominant, and then use the appropriate simulation technique. That is, Importance Sampling using a numerical density, or Conditional Importance Sampling.

The results from Sections 4.2.1 and 4.2.2 indicate that Importance Sampling is indeed a very efficient way of estimating the BER of a transmission system. They also indicate that the models for the densities of the noise sources in the twisted
pair channel are reliable and can be used to apply Importance Sampling to any given
transmission system over copper loops.
Chapter 5

Conclusions

This thesis has considered two of the main problems related to the study of twisted pairs. Obtaining the channel capacity and developing a method for evaluating the performance of proposed transmission systems.

A transmission scheme has been found that approaches capacity in the case of two twisted pairs in the same binder cable. This capacity is larger than what was previously reported in the literature. In fact, under no power constraints, it was believed that the channel has finite capacity, whereas it has been shown in this thesis that in such a case the capacity is infinite. Also, this transmission scheme is easily extendable to $N$ users ($N > 2$). Even though optimality is not guaranteed, it results in capacities which range from equal to far better than the capacities reported in the literature.

It should be noted that all capacity calculations were made under the assumption of an average power constraint. An alternative approach is to assume a peak-power constraint. This would imply that the signals cannot be Gaussian, making all calculations in Chapter 2 not applicable. This approach has been studied in [40], where upper and lower bounds on capacity were derived. The results in [40] indicate that, under the assumption of equal PSD's for the two users, the capacity under peak-power constraint does not differ significantly from the capacity under an average power constraint. However, the case where the two users do not necessarily have the same PSD has not been addressed and is a possible direction for future work.
Calculating channel capacity involves making some simplifications to the channel. The most important is the assumption that the only sources of noise are AWGN and NEXT. It also involves assuming that the transmission system is using the optimal coding technique. Due to the fact that real twisted pairs have many other sources of noise, and the fact that the coding technique actually used may not be the optimal one, it is of interest to know how well a proposed transmission system performs.

To this effect, Importance Sampling and its application to digital transmission over copper loops have been studied. Models for simulating the main sources of interference (NEXT, FEXT, impulse noise and ISI) have been found. Importance Sampling has been applied to BPSK transmission systems over a channel with a flat frequency response and with noise that consists of NEXT, impulse noise and AWGN. It has also been applied to 2B1Q transmission over a channel with non-flat frequency response and with noise that consists of ISI and AWGN.

Simulation results show that Importance Sampling provides an accurate estimate of the BER of the above system while substantially reducing the required number of simulation runs versus the number required using Monte Carlo simulation. Results also demonstrate that the models for the sources of interference on copper loops are reliable, and thus can be used for the simulation of any given transmission scheme over the twisted pair channel.
Appendix A

A.1 Gaussian Signals and Capacity

The problem can be stated as

\[ Y_1 = X_1 + \hat{X}_2 + Z_1 \]
\[ Y_2 = X_2 + \hat{X}_1 + Z_2 \]

where \( Z_1 \) and \( Z_2 \) are i.i.d. (independent and identically distributed), zero-mean Gaussian random variables. That is, both users receive their own signal, a filtered version of the other user's signal and an AWGN component which is independent from the AWGN component of the other user. To achieve capacity, it is necessary to maximize the mutual information between \( Y_1 \) and \( X_1 \), and between \( Y_2 \) and \( X_2 \).

This is an interference channel. A simplified version is given in [4], where it is noted that "this channel has not been solved". That is, it is not known what type of signals the two users must have to achieve capacity.

Although there is no proof, there is strong indication that approaching capacity involves using Gaussian signals for both users. The intuition behind this is based on two facts. First, the multiple access channel [4] can be viewed as a special case of the interference channel, with \( \hat{X}_2 = X_2 \) and \( \hat{X}_1 = X_1 \), and capacity on the multiple access channel is achieved when both users transmit Gaussian signals, as shown in [4]. Second, \( \hat{X}_1 \) and \( \hat{X}_2 \) are filtered versions of \( X_1 \) and \( X_2 \), respectively, which tends to make them approximately Gaussian. Therefore, the total noise in each sub-channel is approximately Gaussian, which implies that Gaussian signals should be used for
both users, because when the pdf of the noise on a channel is given, the best choice is a Gaussian signal [4].

In any case, letting the two users transmit Gaussian signals is a valid option, and will result in a certain value for the mutual information between $Y_1$ and $X_1$, and between $Y_2$ and $X_2$. Since capacity is the maximum possible value of the mutual information, it will necessarily be equal or larger than the value obtained by assuming that the signals are Gaussian. Therefore, making the two users transmit Gaussian signals gives a lower bound on capacity. A lower bound is all that can be obtained at this moment, since, as mentioned above, the solution to the interference channel is not known.

A.2 Analysis of Class of $S_1(f)$ and $S_2(f)$ Defined in Terms of $\beta$ in the Simplified Channel

Given the expressions for $S_1(f)$ and $S_2(f)$, the resulting capacity for user 1 becomes,

$$C_1 = \max_\beta \int_0^W \log_2 \left( 1 + \frac{2P_{\text{max}}}{W} \frac{1}{1+e^{(f-f_0)(f-f_0)}} H \right) \frac{1}{N_0 + \frac{2P_{\text{max}}}{W} X} \, df$$

Therefore,

$$\frac{\partial C_1}{\partial \beta} = \frac{2P_{\text{max}} H}{\ln(2)} \int_0^W \frac{(W/2)^e(f-W/2)}{(1+e^{(f-W/2)})^2} \left( \frac{N_0 + \frac{2P_{\text{max}} X}{1+e^{(f-W/2)}}}{N_0 + \frac{2P_{\text{max}} X}{1+e^{(f-W/2)}}} \right)^2 \frac{2P_{\text{max}} X}{1+e^{(f-W/2)}} \frac{(W/2)^e(f-W/2)}{(1+e^{(f-W/2)})^2} \, df$$

$$= \frac{GH}{\ln(2)} \int_{-W/2}^{W/2} \frac{f^2}{(1+e^{(f-W/2)})^2} \left[ \frac{N_0(1+e^{-\beta f})+GX+Xe^{-\beta f}}{[N_0(1+e^{-\beta f})+GX][N_0(1+e^{-\beta f})+GX+GHe^{-\beta f}]} \right] df$$

$$= \frac{GH}{\ln(2)} \int_0^{W/2} \left[ \frac{N_0(1+e^{-\beta f})+GX+Xe^{-\beta f}}{[N_0(1+e^{-\beta f})+GX][N_0(1+e^{-\beta f})+GX+GHe^{-\beta f}]} - \frac{N_0(1+e^{\beta f})+GX+Xe^{\beta f}}{[N_0(1+e^{\beta f})+GX][N_0(1+e^{\beta f})+GX+GHe^{\beta f}]} \right] df$$
where \( G = \frac{2P_{\max}}{W} \).

Since \( \frac{BC}{\beta} \) has to equal zero, the only two choices for \( \beta \) are \( \beta = 0 \) and \( \beta = \infty \). These two choices correspond to equal PSD’s and Frequency Division Signaling, respectively. Therefore either equal PSD’s are optimum or Frequency Division, just as in the alphanmix case, thus leading to the same threshold test.

### A.3 Analysis of Conjecture 2.1

First, let \( C_{ef} \) be the capacity of the channel given that equal PSD’s are used in bin 1 and FDS is used in bin 2. Let \( C_{fe} \) be defined in a similar manner. Then

\[
C_{ef} = \max_{P_1, P_2 \text{ s.t. } P_1 + P_2 = P_{\max}} \left[ D \log_2 \left( 1 + \frac{P_1 H_1}{N_0 D + P_1 X_1} \right) + \frac{D}{2} \log_2 \left( 1 + \frac{2P_2 H_2}{N_0 D} \right) \right]
\]

and

\[
C_{fe} = \max_{P_1, P_2 \text{ s.t. } P_1 + P_2 = P_{\max}} \left[ \frac{D}{2} \log_2 \left( 1 + \frac{2P_1 H_1}{N_0 D} \right) + D \log_2 \left( 1 + \frac{P_2 H_2}{N_0 D + P_2 X_2} \right) \right]
\]

where \( D = \frac{W}{2} \). Therefore,

\[
C_{ef} = \frac{D}{2 \ln(2)} \max_{S_1, S_2 \text{ s.t. } S_1 + S_2 = S_{\max}} \left[ 2 \ln \left( 1 + \frac{S_1 H_1}{1 + S_1 X_1} \right) + \ln \left( 1 + 2S_2 H_2 \right) \right] \tag{A.1}
\]

and

\[
C_{fe} = \frac{D}{2 \ln(2)} \max_{S_1, S_2 \text{ s.t. } S_1 + S_2 = S_{\max}} \left[ \ln \left( 1 + 2S_1 H_1 \right) + 2 \ln \left( 1 + \frac{S_2 H_2}{1 + S_2 X_2} \right) \right] \tag{A.2}
\]

where \( S_i = \frac{P_i}{N_0 D} \).

#### A.3.1 Low SNR

When \( S_{\max} \ll , S_1 \ll \) and \( S_2 \ll \) because \( S_1 + S_2 = S_{\max} \) and \( S_i \geq 0, i = 1,2 \). Also, \( 0 \leq H_i \leq 1 \), due to the physical attributes of the channel, and \( 0 \leq X_i \leq 1 \),
because of the assumption that $H_1 - 2X_1 > 0$. Thus $S_i H_i \ll$ and $S_i X_i \ll$, $i = 1, 2$.

This implies that

$$C_{ef} \approx \frac{D}{2 \ln(2)} \max_{s.t. \ S_1, S_2 \ s.t. \ S_1 + S_2 = S_{max}} [2S_1 H_1 + 2S_2 H_2]$$

and

$$C_{fe} \approx \frac{D}{2 \ln(2)} \max_{S_1, S_2 \ s.t. \ S_1 + S_2 = S_{max}} [2S_1 H_1 + 2S_2 H_2]$$

In both cases we have,

$$S_1 = \arg \max_{0 \leq S_1 \leq S_{max}} [S_1 H_1 + (S_{max} - S_1) H_2]$$

$$= \arg \max_{0 \leq S_1 \leq S_{max}} [S_1 (H_1 - H_2) + S_{max} H_2]$$

$$= S_{max} \text{ because } H_1 - H_2 > 0$$

$$\Rightarrow \lim_{S_{max} \to 0} S_1 = S_{max} \text{ both for ef and fe}$$

$$\Rightarrow \lim_{S_{max} \to 0} C_{ef} = \frac{D}{2 \ln(2)} 2 \ln \left(1 + \frac{S_{max} H_1}{1 + S_{max} X_1}\right)$$

and

$$\lim_{S_{max} \to 0} C_{fe} = \frac{D}{2 \ln(2)} \ln \left(1 + 2S_{max} H_1\right)$$

Therefore, as $S_{max}$ goes to zero,

$$C_{ef} > C_{fe} \iff 2 \ln \left(1 + \frac{S_{max} H_1}{1 + S_{max} X_1}\right) > \ln \left(1 + 2S_{max} H_1\right)$$

$$\iff \left(1 + \frac{S_{max} H_1}{1 + S_{max} X_1}\right)^2 > 1 + 2S_{max} H_1$$

$$\iff S_{max} < \frac{H_1 - 2X_1}{2X_1^2}$$

Since $H_1 - 2X_1 > 0$, by assumption, it must be the case that for $S_{max} \to 0$,

$$S_{max} < \frac{H_1 - 2X_1}{2X_1^2}, \text{ which implies } C_{ef} > C_{fe}. \text{ Thus,}$$

$$C_{ef} > C_{fe} \text{ as } S_{max} \to 0 \quad (A.3)$$
A.3.2 Any SNR

Equations (A.1) and (A.2) can be used for finding $C_{ef}$ and $C_{fe}$ for any SNR. Letting $S_{2ef} = S_{max} - S_{1ef}$ and $S_{1fe} = S_{max} - S_{2fe}$ leads to

$$C_{ef} = \frac{D}{2 \ln(2)} \max_{0 \leq S_{1ef} \leq S_{max}} \left[ 2 \ln \left( 1 + \frac{S_{1ef} H_1}{1 + S_{1ef} X_1} \right) + \ln \left( 1 + 2 (S_{max} - S_{1ef}) H_2 \right) \right]$$ (A.4)

and

$$C_{fe} = \frac{D}{2 \ln(2)} \max_{0 \leq S_{2fe} \leq S_{max}} \left[ \ln \left( 1 + 2 (S_{max} - S_{2fe}) H_1 \right) + 2 \ln \left( 1 + \frac{S_{2fe} H_2}{1 + S_{2fe} X_2} \right) \right]$$ (A.5)

Taking the derivatives with respect to $S_{1ef}$ and $S_{2fe}$ in the above equations and setting them equal to zero, results in

$$\frac{H_1}{(1 + S_{1ef} X_1)(1 + S_{1ef} (H_1 + X_1))} = \frac{H_2}{1 + 2 (S_{max} - S_{1ef}) H_2}$$

and

$$\frac{H_1}{1 + 2 (S_{max} - S_{2fe}) H_1} = \frac{H_2}{(1 + S_{2fe} X_2)(1 + S_{2fe} (H_2 + X_2))}$$

which in turn lead to

$$S_{1ef} = \frac{-3H_1 - 2X_1 + \sqrt{H_1 (9H_1 + 8X_1) + 4X_1 H_1 (X_1 + H_1)(\frac{1}{H_2} + 2S_{max})}}{2X_1 (X_1 + H_1)}$$ (A.6)

and

$$S_{2fe} = \frac{-3H_2 - 2X_2 + \sqrt{H_2 (9H_2 + 8X_2) + 4X_2 H_2 (X_2 + H_2)(\frac{1}{H_1} + 2S_{max})}}{2X_2 (X_2 + H_2)}$$ (A.7)

Thus, $C_{ef}$ and $C_{fe}$ could be compared by substituting equation (A.6) into equation (A.4), and equation (A.7) into equation (A.5). Unfortunately, the resulting expressions are too complex to compare. It is, however, possible to make some approximations that would simplify the resulting expressions for $C_{ef}$ and $C_{fe}$, thus making the comparison feasible.
A.3.3 High SNR

The approximations can be made when \( S_{\text{max}} \gg 1 \). In this case,

\[
S_{1_{\text{ef}}} \approx \frac{\sqrt{4X_1H_2(X_1+H_1)(2S_{\text{max}})}}{2X_1(X_1+H_1)}
= \sqrt{\frac{2H_1}{X_1(X_1+H_1)}} S_{\text{max}}
\]

and

\[
S_{2_{\text{ef}}} \approx \sqrt{\frac{2H_2}{X_2(X_2+H_2)}} S_{\text{max}}
\]

Therefore, \( S_{1_{\text{ef}}} \gg 1 \) and \( S_{2_{\text{ef}}} \gg 1 \), which implies that

\[
C_{\text{ef}} \approx \frac{D}{2\ln(2)} \left[ 2\ln \left( 1 + \frac{H_1}{X_1} \right) + \ln \left( 1 + 2 \left( S_{\text{max}} - \sqrt{\frac{2H_1}{X_1(X_1+H_1)}} S_{\text{max}} \right) H_2 \right) \right]
\]

and

\[
C_{\text{fe}} \approx \frac{D}{2\ln(2)} \left[ \ln \left( 1 + 2 \left( S_{\text{max}} - \sqrt{\frac{2H_2}{X_2(X_2+H_2)}} S_{\text{max}} \right) H_1 \right) + 2\ln \left( 1 + \frac{H_2}{X_2} \right) \right]
\]

Hence,

\[
C_{\text{ef}} - C_{\text{fe}} \approx \frac{D}{2\ln(2)} \left[ 2\ln \left( 1 + \frac{H_1}{X_1} \right) + \ln \left( \frac{2H_2 S_{\text{max}} - 2H_2 \sqrt{\frac{2H_1}{X_1(X_1+H_1)}} S_{\text{max}} + 1}{2H_1 S_{\text{max}} - 2H_1 \sqrt{\frac{2H_2}{X_2(X_2+H_2)}} S_{\text{max}} + 1} \right) \right]
\]

Therefore,

\[
C_{\text{ef}} - C_{\text{fe}} \rightarrow \frac{D}{2\ln(2)} \left[ 2\ln \left( 1 + \frac{H_1}{X_1} \right) + \ln \left( \frac{H_2}{H_1} \right) \right] \quad \text{as \( S_{\text{max}} \rightarrow \infty \)} \quad (A.8)
\]

Now, \( X_1 < X_2 \Rightarrow \ln \left( \frac{X_1}{X_2} \right) < 0 \), which implies

\[
2\ln \left( \frac{1+\frac{H_1}{X_1}}{1+\frac{H_1}{X_2}} \right) + \ln \left( \frac{H_2}{H_1} \right) > 2\ln \left( \frac{1+\frac{H_1}{X_1}}{1+\frac{H_1}{X_2}} \right) + \ln \left( \frac{H_2}{H_1} \right) + \ln \left( \frac{X_1}{X_2} \right)
= \ln \left( \frac{1+\frac{H_1}{X_1}}{1+\frac{H_1}{X_2}} \right) + \ln \left( \frac{1+\frac{H_1}{X_1}}{1+\frac{H_1}{X_2}} \right) + \ln \left( \frac{X_1}{X_2} \right)
= \ln \left( \frac{1+\frac{H_1}{X_1}}{1+\frac{H_1}{X_2}} \right) + \ln \left( \frac{1+\frac{H_1}{X_1}}{1+\frac{H_1}{X_2}} \right)
= \ln \left( \frac{1+\frac{H_1}{X_1}}{1+\frac{H_1}{X_2}} \right) + \ln \left( \frac{X_1+1}{X_2+1} \right)
\]
Therefore,

\[2 \ln \left( \frac{1 + \frac{H_1}{X_1}}{1 + \frac{H_2}{X_2}} \right) + \ln \left( \frac{H_2}{H_1} \right) > \ln \left( \frac{2 + \frac{1}{\kappa} + \lambda}{2 + \frac{1}{\lambda} + \lambda} \right) \quad \text{where} \quad \kappa = \frac{H_1}{X_1} \quad \text{and} \quad \lambda = \frac{H_2}{X_2} \quad \text{(A.9)}\]

Also,

\[H_2 - 2X_2 > 0 \quad \text{(by assumption)}\]
\[\Rightarrow \frac{H_2}{X_2} > 2\]
\[\Rightarrow \lambda > 2\]
\[\Rightarrow \lambda > \frac{1}{\lambda}\]

and,

\[H_1 > H_2 \quad \text{and} \quad X_1 < X_2 \quad \text{(by assumption)}\]
\[\Rightarrow \frac{H_1}{X_1} > \frac{H_2}{X_2}\]
\[\Rightarrow \kappa > \lambda\]
\[\Rightarrow \kappa > \frac{1}{\lambda}\]
\[\Rightarrow (\kappa - \lambda) > 0 \quad \text{and} \quad (\kappa - \frac{1}{\lambda}) > 0\]
\[\Rightarrow (\kappa - \lambda)(\kappa - \frac{1}{\lambda}) > 0\]
\[\Rightarrow \kappa^2 - \kappa \frac{1}{\lambda} - \lambda \kappa + 1 > 0\]
\[\Rightarrow \kappa - \frac{1}{\lambda} - \lambda + \frac{1}{\kappa} > 0\]
\[\Rightarrow \kappa + \frac{1}{\kappa} > \lambda + \frac{1}{\lambda}\]
\[\Rightarrow 2 + \kappa + \frac{1}{\kappa} > 2 + \lambda + \frac{1}{\lambda}\]
\[\Rightarrow \frac{2 + \frac{1}{\kappa} + \kappa}{2 + \frac{1}{\lambda} + \lambda} > 1\]
\[\Rightarrow \ln \left( \frac{2 + \frac{1}{\kappa} + \kappa}{2 + \frac{1}{\lambda} + \lambda} \right) > 0\]

Thus, by equation (A.9),

\[2 \ln \left( \frac{1 + \frac{H_1}{X_1}}{1 + \frac{H_2}{X_2}} \right) + \ln \left( \frac{H_2}{H_1} \right) > 0\]

and by equation (A.8),

\[C_{ef} - C_{fe} > 0 \quad \text{as} \quad S_{\text{max}} \rightarrow \infty\]
It therefore follows that

\[ C_{ef} > C_{fe} \text{ as } S_{\text{max}} \rightarrow \infty \]  \hspace{1cm} (A.10)

\section*{A.3.4 Discussion}

Equations (A.3) and (A.10) demonstrate that \( C_{ef} \) is better than \( C_{fe} \) for very low and for very high SNR's. It remains to be shown that this is the case for all values of SNR. Although the expression for \( C_{ef} - C_{fe} \) obtained as described in subsection A.3.2 is too complex to enable determination of its sign for all values of SNR, computer calculations can be used to determine its behavior under various scenarios. Even though computer calculations do not constitute a proof, they provide an indication of how \( C_{ef} - C_{fe} \) changes with SNR, and may also lead to some insight that will enable a future analytical proof.

Figure A.1 shows the plots of \( C_{ef} - C_{fe} \) vs SNR under four different scenarios. One where \( H_2 \approx H_1 \) and \( X_2 \approx X_1 \), one where \( H_2 \approx H_1 \) and \( X_2 \gg X_1 \), another with \( H_2 \gg H_1 \) and \( X_2 \approx X_1 \), and finally another with \( H_2 \gg H_1 \) and \( X_2 \gg X_1 \). These four cases are representative of all situations that can occur. It can be seen that in all cases \( C_{ef} - C_{fe} > 0 \) for all values of SNR.

It should also be noted that the approximation for \( C_{ef} - C_{fe} \) given by equation (A.8) agrees with the calculated value in all four cases mentioned above.
Figure A.1 $C_{ef} - C_{fe}$ vs SNR.
Appendix B

B.1 Relation Between $K_{\text{NEXT}}$ and PSD of NEXT

It is assumed that the receiver uses $K$ samples per symbol. These are taken at times $t^*, t^* + T_s, \ldots, t^* + (K-2)T_s, t^* + (K-1)T_s$. Thus, assuming that NEXT is a WSS (wide-sense stationary) process,

$$K_{\text{NEXT}_{i,j}} = E \left[ N_{\text{NEXT}_{t^*+(i-1)T_s}} N_{\text{NEXT}_{t^*+(j-1)T_s}} \right]$$

$$= R_{\text{NEXT}} (t^* + (i - 1)T_s - t^* - (j - 1)T_s)$$

$$= R_{\text{NEXT}} ((i - j)T_s)$$

where $R_{\text{NEXT}}(\tau)$ is the autocorrelation function of NEXT, and

$$R_{\text{NEXT}}(\tau) = \mathcal{F}^{-1} \{ S_{\text{NEXT}}(f) \}$$

$\mathcal{F}^{-1}$ denotes the inverse Fourier transform, and $S_{\text{NEXT}}(f)$ is the PSD of NEXT.

Similarly, for FEXT,

$$K_{\text{FEXT}_{i,j}} = R_{\text{FEXT}} ((i - j)T_s)$$

where $R_{\text{FEXT}}(\tau) = \mathcal{F}^{-1} \{ S_{\text{FEXT}}(f) \}$
B.2 Derivation of the Density of Impulse Noise

Using an intuitive approach, the problem can be viewed as follows. In one second there are, on average, λ events. Let L denote the average duration of an impulse, in seconds, and assume that the probability of two impulses overlapping is negligible, that is, \( \frac{1}{\lambda} \gg L \). Then, the average duration of all impulse events, in a given second, is \( \lambda L \). The number of samples taken during a period of length \( \lambda L \) seconds, is equal to \( \frac{N}{T_s} \). Therefore, the average number of symbols, in one second, affected by an impulse event is \( \frac{N}{K T_s} \), assuming that the duration of a symbol is much smaller than the average duration of an impulse event, that is, \( K T_s \ll L \).

Let \( R \) denote the symbol rate, that is, the number of symbols per second. Then, in one second, \( R \) symbols are transmitted and an average of \( \frac{N}{K T_s} \) symbols are affected by an impulse event. This implies that if one of the symbols in a given second is chosen at random, the probability that it will be in an impulse event is \( \frac{N}{K T_s R} = \frac{N}{T_s} \lambda L \). Therefore, the probability that a given symbol is not in an impulse event is equal to \( \alpha = 1 - \lambda L \).

The pdf of \( K \) samples of impulse noise can be written as

\[
p_{IN}(x) = Pr(\text{symbol is not in an impulse event})p_{X|\text{no impulse}}(x|\text{no impulse})
+ Pr(\text{symbol is in an impulse event})p_{X|\text{impulse}}(x|\text{impulse})
= \alpha \delta(x) + (1 - \alpha)p_2(x)
\]

\( p_2(x) \) is assumed to be Gaussian with large variance. This is justified by the fact that, given that the symbol under consideration in in an impulse event, the \( K \) samples corresponding to that symbol will take on very large values. Thus, it is irrelevant, for the detection of that symbol, whether these large values were a result of a Gaussian or non-Gaussian pdf. Hence, a Gaussian pdf is assumed for analytical convenience.
The validity of this assumption is also shown by the results of Chapter 4. Therefore,

\[ p_2(x) = \frac{1}{\sqrt{2\pi K_{ii}}} \exp \left( -\frac{x^T K_{ii}^{-1} x}{2} \right) \]

The covariance matrix \( K_{ii} \) reflects the correlation between the \( K \) samples when an impulse has occurred, and as will be shown, depends on \( h(t) \), the impulse response of the channel. \( h(t) \) is assumed to have support only on \( t \in [0, L] \). Therefore,

\[
K_{ii,j} = E \left[ N_i^{\text{in}} N_j^{\text{in}} \right] = E \left[ U \ h \left( t^* + (i - 1)T_s \right) \ U \ h \left( t^* + (j - 1)T_s \right) \right] = E \left[ U^2 \right] \frac{1}{L} \int_0^L h(t) h \left( t + (i - j)T_s \right) dt
\]

where \( U \) is the random amplitude of an impulse event (see Chapter 3).
Appendix C

C.1 Bounds on Gain

For $\alpha^* = 0.5$, $K = 1$, and BPSK transmission,

$$W = \int_{0}^{\infty} \frac{\alpha}{\sqrt{2\pi}\sigma_1^2} e^{-\frac{(r-S_0)^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sqrt{2\pi}\sigma_2^2} e^{-\frac{(r-S_0)^2}{2\sigma_2^2}} \frac{r^2}{2\sigma_1^2} + \frac{0.5}{\sqrt{2\pi}\sigma_1^2} e^{-\frac{r^2}{2\sigma_1^2}} - \frac{0.5}{\sqrt{2\pi}\sigma_2^2} e^{-\frac{r^2}{2\sigma_2^2}} p_n(r)dr$$

$$= 2 \int_{0}^{\infty} \frac{\alpha}{\sigma_1} e^{-\frac{r^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sigma_2} e^{-\frac{r^2}{2\sigma_2^2}} \frac{r^2}{2\sigma_1^2} + \frac{1}{\sigma_1} e^{-\frac{r^2}{2\sigma_1^2}} + \frac{1}{\sigma_2} e^{-\frac{r^2}{2\sigma_2^2}} p_n(r)dr$$

$$= 2 \int_{0}^{\infty} \frac{\alpha}{\sigma_1} e^{-\frac{r^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sigma_2} e^{-\frac{r^2}{2\sigma_2^2}} \frac{r^2}{2\sigma_1^2} + \frac{1}{\sigma_1} e^{-\frac{r^2}{2\sigma_1^2}} + \frac{1}{\sigma_2} e^{-\frac{r^2}{2\sigma_2^2}} p_n(r)dr$$

$$+ 2 \int_{0}^{\infty} \alpha \sigma_1 e^{-\frac{(r-S_0)^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sigma_2} e^{-\frac{(r-S_0)^2}{2\sigma_2^2}} \frac{r^2}{2\sigma_1^2} + \frac{1}{\sigma_1} e^{-\frac{r^2}{2\sigma_1^2}} + \frac{1}{\sigma_2} e^{-\frac{r^2}{2\sigma_2^2}} p_n(r)dr$$

where $\theta = \sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 - \sigma_1^2}} \ln\left(\frac{\sigma_2^2}{\sigma_1^2}\right)$. The significance of $\theta$ is that

$$\frac{1}{\sigma_1} e^{-\frac{r^2}{2\sigma_1^2}} \geq \frac{1}{\sigma_2} e^{-\frac{r^2}{2\sigma_2^2}} \text{ iff } r \leq \theta$$
Therefore,

\[
\bar{W} \geq 2 \int_0^\theta \frac{\alpha}{\sigma_1} e^{-\frac{(r-S_0)^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sigma_2} e^{-\frac{(r-S_0)^2}{2\sigma_2^2}} p_n(r) dr
\]

\[
+ 2 \int_0^\infty \frac{\alpha}{\sigma_1} e^{-\frac{(r-S_0)^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sigma_2} e^{-\frac{(r-S_0)^2}{2\sigma_2^2}} p_n(r) dr
\]

\[
= \bar{W}_{\text{min}}
\]

(C.1)

and

\[
\bar{W} \leq 2 \int_0^\theta \frac{\alpha}{\sigma_1} e^{-\frac{(r-S_0)^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sigma_2} e^{-\frac{(r-S_0)^2}{2\sigma_2^2}} p_n(r) dr
\]

\[
+ 2 \int_0^\infty \frac{\alpha}{\sigma_1} e^{-\frac{(r-S_0)^2}{2\sigma_1^2}} + \frac{1-\alpha}{\sigma_2} e^{-\frac{(r-S_0)^2}{2\sigma_2^2}} p_n(r) dr
\]

\[
= \bar{W}_{\text{max}}
\]

(C.2)

It is clear that \( \bar{W}_{\text{min}} \leq \bar{W} \leq \bar{W}_{\text{max}} \), and \( \bar{W}_{\text{max}} = 2\bar{W}_{\text{min}} \). The equation for the Importance Sampling gain is

\[
\Gamma = \frac{p_e - p_e^2}{\bar{W} - p_e^2}
\]

Thus,

\[
\frac{p_e - p_e^2}{2\bar{W}_{\text{min}} - p_e^2} \leq \Gamma \leq \frac{p_e - p_e^2}{\bar{W}_{\text{min}} - p_e^2}
\]

and, for small \( p_e \),

\[
\frac{1}{2 \bar{W}_{\text{min}}} \leq \Gamma \leq \frac{p_e}{\bar{W}_{\text{min}}}
\]

It is therefore necessary to calculate \( \bar{W}_{\text{min}} \), in order to obtain the bounds on \( \Gamma \).
\[ \bar{W}_{\text{mlm}} = A + B \]

where, \[ A = \int_0^\theta \left[ ae^{-\frac{s_1^2 - 2rs_0}{2\sigma_1^2}} + \left( 1 - \alpha \right) \frac{\sigma_1}{\sigma_2} e^{-\frac{s_1^2 + s_2^2 - 2rs_0 + \frac{1}{2}r^2}{2\sigma_2^2}} \right] \rho_n(r)dr \tag{C.3} \]

and, \[ B = \int_\theta^\infty \left[ \frac{\sigma_2}{\sigma_1} e^{-\frac{s_1^2 + s_2^2 - 2rs_0 + \frac{1}{2}r^2}{2\sigma_1^2}} + \left( 1 - \alpha \right) e^{-\frac{s_1^2}{2\sigma_2^2}} \right] \rho_n(r)dr \]

After extensive calculations and manipulations, the following result is obtained,

\[ \bar{W}_{\text{mlm}} = \alpha^2 e^{\frac{s_2^2}{2\sigma_1^2}} \left[ Q \left( \frac{2s_1}{\sigma_1} \right) - Q \left( \frac{\sigma_1}{\sigma_1} + 2 \frac{s_1}{\sigma_1} \right) \right] \]

\[ + \alpha (1 - \alpha) e^{\frac{1}{2} \frac{s_2^2}{\sigma_1^2}} \left[ \frac{s_2^2}{2\sigma_1^2} \right] \left[ Q \left( \frac{1 + \frac{s_2^2}{\sigma_1^2} \frac{s_1}{\sigma_1} \right) - Q \left( \frac{\sigma_1}{\sigma_1} + \left( 1 + \frac{s_2^2}{\sigma_1^2} \right) \frac{s_1}{\sigma_1} \right) \right] \]

\[ + \alpha (1 - \alpha) \frac{\sigma_1}{\sqrt{\sigma_1^2 + 2\sigma_2^2}} e^{-\frac{s_2^2}{2\sigma_1^2 + \frac{1}{2}r^2}} \left[ Q \left( \frac{\sigma_1}{\sigma_1} + \frac{s_1}{\sigma_2} \right) \frac{s_1}{\sqrt{\sigma_1^2 + 2\sigma_2^2}} \right] \]

\[ - Q \left( \frac{\sigma_1}{\sigma_1} + \left( 1 + \frac{s_2^2}{\sigma_1^2} \right) \frac{s_1}{\sigma_1} \right) \right] \]

\[ + (1 - \alpha)^2 \frac{\sigma_1^2}{\sigma_2 \sqrt{\sigma_1^2 + 2\sigma_2^2}} e^{-\frac{s_2^2}{2\sigma_1^2 + \frac{1}{2}r^2}} \left[ Q \left( 2 \frac{s_1}{\sigma_2} \frac{s_1}{\sqrt{\sigma_1^2 + 2\sigma_2^2}} \right) \right] \]

\[ - Q \left( 2 \frac{s_1}{\sigma_2} + 2 \frac{s_1}{\sigma_1} \frac{s_1}{\sqrt{\sigma_1^2 + 2\sigma_2^2}} \right) \]

\[ + (1 - \alpha)^2 e^{\frac{s_2^2}{\sigma_1^2}} Q \left( \frac{\sigma_1}{\sigma_1} + \frac{s_1}{\sigma_2} \right) + \alpha (1 - \alpha) e^{\frac{1}{2} \frac{s_2^2}{\sigma_1^2}} \left[ \frac{s_2^2}{2\sigma_1^2} \right] \left[ Q \left( \frac{\sigma_1}{\sigma_1} + \left( 1 + \frac{s_2^2}{\sigma_1^2} \right) \frac{s_1}{\sigma_1} \right) \right] \]

\[ + \alpha (1 - \alpha) \frac{\sigma_1}{\sqrt{2\sigma_1^2 + \sigma_2^2}} e^{-\frac{s_2^2}{2\sigma_1^2 + \frac{1}{2}r^2}} \left[ Q \left( \frac{\sigma_1}{\sigma_1} + \left( 1 + \frac{s_2^2}{\sigma_1^2} \right) \frac{s_1}{\sigma_1} \right) \right] \]

\[ + \alpha^2 \frac{\sigma_2^2}{\sigma_1 \sqrt{2\sigma_1^2 + \sigma_2^2}} e^{-\frac{s_2^2}{2\sigma_1^2 + \frac{1}{2}r^2}} \left[ Q \left( \frac{s_1}{\sigma_1} + \frac{s_1}{\sigma_2} \right) \frac{s_1}{\sqrt{2\sigma_1^2 + \sigma_2^2}} \right] \right) \tag{C.4} \]
where $Q$ is the complementary distribution function of a Gaussian random variable with zero mean and unit variance.
Bibliography


