1

Information, Channel Capacity and Channel Modelling

1.1 Introduction

In this chapter, an introduction to the fundamental aspects of information theory is given, with particular attention given to the derivation of the capacity of different channel models. This is followed by an explanation of the physical properties of fading channels and descriptions of different channel models for fixed wireless access, universal mobile telecommunication systems (UMTS) for single-input–single-output (SISO) and multiple-input–multiple-output (MIMO), and finally magnetic recording channels. Therefore, this chapter presents the prerequisites for evaluating many binary and non-binary coding schemes on various channel models.

The purpose of a communication system is, in the broadest sense, the transmission of information from one point in space and time to another. We shall briefly explore the basic ideas of what information is and how it can be measured, and how these ideas relate to bandwidth, capacity, signal-to-noise ratio, bit error rate and so on.

First, we address three basic questions that arise from the analysis and design of communication systems:

- Given an information source, how do we determine the 'rate' at which the source is transmitting information?
- For a noisy communication channel, how do we determine the maximum 'rate' at which 'reliable' information transmission can take place over the channel?
- How will we develop statistical models that adequately represent the basic properties of communication channels?

For modelling purposes we will divide communication channels into two categories: analogue channels and discrete channels. We wish to construct a function that

Non-Binary Error Control Coding for Wireless Communication and Data Storage Rolando Antonio Carrasco and Martin Johnston

measures the amount of information present in an event that occurs with probability p. A lower probability means that we obtain more information about the event, whereas if the event is entirely predictable ($p = 1$) then we obtain no information about the event since we already have knowledge of the event before we receive it.

1.1.1 Information Theory [1]

- How can the symbols of communication be transmitted? Technical or Syntactic level?
- How precisely do the transmitted symbols carry the desired meaning? Semantic problem.
- How effectively does the received meaning affect conduct in the desired manner?

1.1.2 Definition of Information [1, 2]

Information: Knowledge not precisely known by the recipient, or a measure of unexpectedness.

Given an information source, we evaluate the *rate* at which the source is emitting information as:

$$\text{Rate} = \frac{\text{symbols}}{\text{second}} \quad \text{OR} \quad \text{Rate} = \frac{\text{bits}}{\text{second}}.$$

However, given a noisy communication channel, how do we determine the maximum rate at which *reliable* information transmission can take place over the channel?

1.2 Measure of Information [1–3]

We should briefly explore basic ideas about what information is, how it can be measured, and how these ideas relate to bandwidth and signal-to-noise ratio. The amount of information about an event is closely related to its probability of occurrence. Messages containing knowledge of a high probability of occurrence (i.e. those indicating very little uncertainty in the outcome) convey relatively little information. In contrast, those messages containing knowledge of a low probability of occurrence convey relatively large amounts of information. Thus, a measure of the information received from the knowledge of occurrence of an event is inversely related to the probability of its occurrence.

Assume an information source transmits one of nine possible messages M_1, M_2, $\ldots$, M_9 with probability of occurrence $P_1, P_2, \ldots, P_9$, where $P_1 + P_2 + \cdots + P_9 = 1$, as shown in Figure 1.1.

According to our intuition, the information content or the amount of information in the ith message, denoted by $I(M_i)$, must be inversely related to P_i. Also, to satisfy

Figure 1.1 Messages and their associated probabilities.

our intuitive concept of information, $I(M_i)$ must satisfy:

$$I(M_i) > I(M_j) \text{ if } P_i < P_j$$
$$I(M_i) \to 0,\ P_i \to 1$$
$$I(M_i) \geq 0 \text{ when } 0 \leq P_i \leq 1.$$

We can explain the concept of independent messages transmitting from the same source. For example, the received message 'It will be cold today and hot tomorrow' is the same as the sum of information received in the two messages 'It will be cold today' and 'It will be hot tomorrow' (assuming that the weather today does not affect the weather tomorrow).

Mathematically, we can write this as:

$$I(M_i \text{ and } M_j) \triangleq I(M_i, M_j) = I(M_i) + I(M_j),$$

where M_i and M_j are the two independent messages. We can define a measure of information as the logarithmic function:

$$I(M_i) = \log_x \left(\frac{1}{P_i}\right), \quad \text{where } x \text{ is the base } 2, e, 10, \ldots \tag{1.1}$$

The base x for the logarithm in (1.1) determines the unit assigned to the information content:

$$x = e \quad \text{nats}$$
$$x = 2 \quad \text{bits}$$
$$x = 10 \quad \text{Hartley},$$

$P_1 + P_2 + P_3 + \cdots + P_q = 1$ where P_q is the probability of the message occurring and q is the index value, that is $1, 2 \ldots$

The information content or the amount of information $I(M_k)$ in the kth message with the set kth probability (P_k) boundary values is:

1. $I(M_k) \to 0 \quad \text{as} \quad P_k \to 1.$
   Obviously, if we are absolutely certain of the outcome of an event, even before it occurred.

2. $I(M_k) \geq 0$ when $0 \leq P_k \leq 1$.
   That is to say, the occurrence of an event $M = M_k$ either provides some or no information content.
3. $I(M_k) > I(M_i)$ if $P_k < P_i$.
   The less probable an event is, the more information we gain when it occurs.
4. $I(M_k M_i) = I(M_k) + I(M_i)$.
   If $I(M_k)$ and $I(M_i)$ are statistically independent messages.
5. Mathematically we can prove that:
6. $I(M_k \text{ and } M_j) = I(M_k) + I(M_j)$, where M_k and M_j are two statistically independent messages. The logarithmic measure is convenient.

Example 1.1: A source puts out five possible messages. The probabilities of these messages are:

$$P_1 = \frac{1}{2} \quad P_2 = \frac{1}{4} \quad P_3 = \frac{1}{8} \quad P_4 = \frac{1}{16} \quad P_5 = \frac{1}{16}.$$

Determine the information contained in each of these messages.

Solution:

$$I(M_1) = \log_2 \left(\frac{1}{(1/2)} \right) = 1 \text{ bit}$$

$$I(M_2) = \log_2 \left(\frac{1}{(1/4)} \right) = 2 \text{ bits}$$

$$I(M_3) = \log_2 \left(\frac{1}{(1/8)} \right) = 3 \text{ bits.}$$

$$I(M_4) = \log_2 \left(\frac{1}{(1/16)} \right) = 4 \text{ bits}$$

$$I(M_5) = \log_2 \left(\frac{1}{(1/16)} \right) = 4 \text{ bits}$$

Information total = 14 *bits*.

1.2.1 Average Information

Text messages produced by an information source consist of sequences of symbols but the receiver of a message may interpret the entire message as a single unit. When we attempt to define the information content of symbols, we are required to define the average information content of symbols in a long message, and the statistical dependence of symbols in a message sequence will change the average information content of symbols.

Let a transmitter unit consist of U possible symbols, $s_1, s_2, \ldots, s_u$, in a statistically independent sequence. The possibility of occurrence of a particular symbol during a symbol time does not depend on the symbol transmitted by the source previous in time. Let $P_1, P_2, \ldots, P_u$ be the probability of occurrence of the U symbols, in a long message having N symbols. The symbol s_1 will occur on average P_1N times. The symbol s_2 will occur P_2N times, and the symbol s_i will occur P_iN times. Assuming an individual symbol s to be a message of length 1, we can define the information content of the ith symbol as $\log_2(\frac{1}{P_i})$ bits. The P_iN occurrence of s_i contributes an information content of $P_i N \log_2(\frac{1}{P_i})$ bits.

The total information content of message is then the sum of each of the U symbols of the source:

$$I_{\text{total}} = \sum_{i=1}^{U} NP_i \log_2 \left(\frac{1}{P_i}\right) \text{ bits.} \tag{1.2}$$

The average information per symbol is measured by dividing the total information by the number of symbols N in the message:

$$H = \frac{I_{\text{total}}}{N} = \sum_{i=1}^{U} P_i \log_2 \left(\frac{1}{P_i}\right) \text{ bits/symbol.} \tag{1.3}$$

The average information H is called the source entropy (bits/symbol).

Example 1.2: Determine the entropy of a source that emits one of three symbols, A, B, C, in a statistically independent sequence, with a probability of $\frac{1}{2}$, $\frac{1}{4}$ and $\frac{1}{4}$, respectively.

Solution: The information contents of the symbols are:

one bit for A

two bit for B

two bit for C

$$H = \frac{1}{2}\log_2\left(\frac{1}{\left(\frac{1}{2}\right)}\right) + \frac{1}{4}\log_2\left(\frac{1}{\left(\frac{1}{4}\right)}\right) + \frac{1}{4}\log_2\left(\frac{1}{\left(\frac{1}{4}\right)}\right).$$

So the average information content per symbol on the source entropy is:

$$H = \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + \frac{1}{4}\log_2(4).$$
$$H = 1.5 \text{ bits/symbols}$$

If we have a fixed time, say r_s symbols/s, then by definition the average source of information rate R in bits per second is the *product* of the average information content per symbol and the symbol rate r_s, as shown below:

$$R = r_s \cdot H \text{ bits/s}.$$

1.2.2 The Entropy of a Binary Source

For 'unbiased' results, the probability of logic 0 or logic 1 is:

$$P(1) = P(0) = 0.5$$
$$\text{Entropy} = H = -\sum_{i=1}^{M} P_i \log_2 P_i$$
$$H = -P_1 \log_2 (P_1) - P_2 \log_2 (P_2)$$
$$H = -P_1 \log_2 \left(\frac{1}{2}\right) - P_2 \log_2 \left(\frac{1}{2}\right)$$
$$H = 0.5 + 0.5 = 1.$$

This agrees with expectation. Now consider a system $P_1 \neq P_2$. We note that, for a two-state system, since $P_1 + P_2 = 1$ we have $P_2 = 1 - P_1$. We can then write:

$$H = P_1 \log_2 \left(\frac{1}{P_1}\right) + (1 - P_1)\ \log_2 \left(\frac{1}{1 - P_1}\right) \text{ bits}. \tag{1.4}$$

Table 1.1 shows how the entropy H varies for different values of P_1 and P_2.

Table 1.1 The variation of entropy for different values of P_1 and P_2.

P_1	$P_2 = 1 - P_1$	$P_1 \log_2 \left(\frac{1}{P_1}\right)$	$(1 - P_1) \log_2 \left(\frac{1}{1-P_1}\right)$	H
0	1	0	0	0
0.2	0.8	0.46	0.25	0.72
0.4	0.6	0.528	0.44	0.97
0.5	0.5	0.5	0.5	1
0.6	0.4	0.44	0.528	0.97
0.8	0.2	0.25	0.46	0.72
1	0	0	0	0

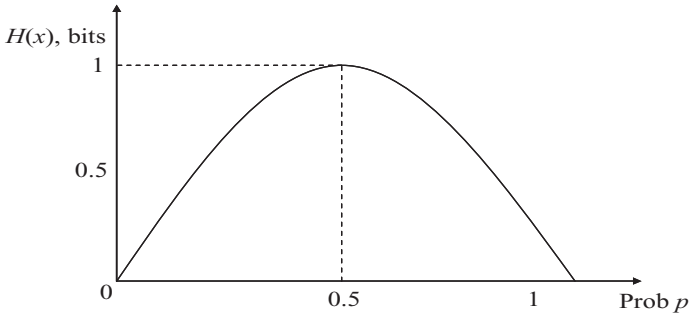

Figure 1.2 Entropy function of a binary source.

From Table 1.1 the following observations can be made:

The entropy $H(x)$ attains its maximum value, $H_{\max} = 1$ bit, when $P_1 = P_2 = 1/2$; that is, symbols 1 and 0 are equally probable.

When $P_1 = 0$ or 1, the entropy value of $H(x) = 0$, resulting in no information.

The entropy is plotted in Figure 1.2.

The next two examples show how to determine the entropy functions for an event with a uniform and Gaussian probability distribution.

Example 1.3: Determine the entropy of an event x with a uniform probability distribution defined by:

$$P(x) = \frac{1}{2.X_0}, -X_0 < x < X_0$$
$$P(x) = 0, \quad \text{elsewhere}$$

The uniform distribution is illustrated in Figure 1.3.

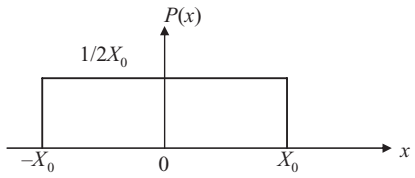

Figure 1.3 A uniform distribution.

Determine the entropy sources [1, 3]:

$$H(x) = -\int_x P(x) \log P(x)\, \mathrm{d}x$$

$$H(x) = \int_{-X_0}^{X_0} \frac{1}{2X_0} \log(2X_0)\ \mathrm{d}x$$

$$H(x) = \frac{\log(2X_0)}{2X_0} \int_{-X_0}^{X_0} \mathrm{d}x \qquad .$$

$$H(x) = \frac{\log(2X_0)}{2X_0} (X_0 - (-X_0))$$

$$H(x) = \frac{\log(2X_0)}{2X_0} (2X_0) = \log(2X_0)$$

Example 1.4: Determine the entropy of an event x with a Gaussian probability distribution defined by:

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad \text{where } \sigma \text{ is the standard deviation.}$$

The distribution is normalized so that the area under the pdf is unity

$$\int_{-\infty}^{\infty} P(x)\,\mathrm{d}x = 1.$$

The variance σ^2 of the distribution is given as:

$$\int_{-\infty}^{\infty} x^2 P(x)\,\mathrm{d}x = \sigma^2.$$

Therefore, the entropy function is:

$$H(x) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{x^2}{2\sigma^2}} \left(\frac{x^2}{2\sigma^2} + \ln \sigma\sqrt{2\pi}\right) \mathrm{d}x$$

$$H(x) = \frac{1}{2} + \ln\left(\sigma\sqrt{2\pi}\right) \qquad .$$

$$H(x) = \ln\ e^{\frac{1}{2}} + \ln\left(\sigma\sqrt{2\pi}\right)$$

$$H(x) = \ln\left(\sigma\sqrt{2\pi e}\right) \text{ nats or } H(x) = \log_2\left(\sigma\sqrt{2\pi e}\right) \text{ bits}$$

1.2.3 Mutual Information

Given that we think of the channel output y as a noisy version of the channel input x value, and the entropy H is a measure of the prior uncertainty about x, how can we measure the uncertainty about x after observing the y value?

The conditional entropy of x is defined as [4]:

$$H(X|Y = y_k) = \sum_i P\,(x_i|y_k) \log_2 \left[\frac{1}{P\,(x_i|y_k)}\right], \quad k = 0, 1, 2, \ldots,\ i = 0, 1, 2, \ldots \tag{1.5}$$

This quantity is itself a random variable that takes on values:

$$H(X|y = y_0), \ldots, H(Y|y = y_{k-1})$$

with probabilities of:

$$P(y_0), \ldots, P(y_{k-1}) \text{ respectively.}$$

So the mean entropy is:

$$\overline{H} = \sum\nolimits_k H\,(X|y = y_k)\,P\,(y_k)$$

$$\overline{H} = \sum\nolimits_k \sum\nolimits_i P\,(x_i|y_k)\,P\,(y_k) \log_2 \left[\frac{1}{P\,(x_i|y_k)}\right].$$

$$\text{Or} \quad \overline{H} = \sum\nolimits_k \sum\nolimits_i P\,(x_i, y_k) \log_2 \left[\frac{1}{P\,(x_i|y_k)}\right] \tag{1.6}$$

This quantity $\overline{H}$ or $H\ (X|Y)$ is called conditional entropy. It represents the amount of uncertainty remaining about the input after the channel output has been observed. Since the entropy $H(X)$ represents our uncertainty about the channel input before observing the channel output, it follows that the difference is:

$$H(X) - \overline{H(X|Y)}.$$

This must represent our uncertainty about the channel input, which is resolved by observing the channel output. This important quantity is called the mutual information of the channel. Denoting the mutual information by $I\ (X, Y)$, we may thus write:

$$I(X, Y) = H(X) - \overline{H(X|Y)}. \tag{1.7}$$

Similarly:

$$I(Y, X) = H(Y) - \overline{H(Y|X)}, \tag{1.8}$$

where $H(Y)$ is the entropy of the channel output and $H(Y|X)$ is the conditional entropy of the channel output given the channel input.

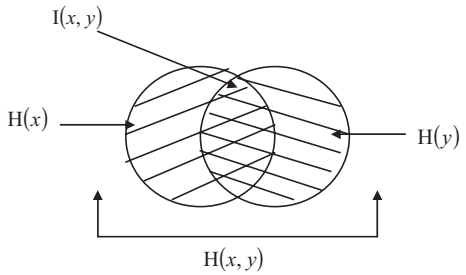

Figure 1.4 A Venn diagram illustrating the relationship between mutual information and entropy.

Property 1 The mutual information of a channel is symmetric, that is:

$$I(x, y) = I(y, x).$$

Property 2 The mutual information is always nonnegative, that is:

$$I(x, y) \geq 0.$$

Property 3 The mutual information of a channel is related to the joint entropy of the channel input and output by:

$$I(x, y) = H(x) + H(y) - H(x, y)$$

and is shown in Figure 1.4.

Example 1.5: Determine the mutual information of the binary symmetric channel in Figure 1.5 where the associated probabilities are $P(x = 1) = 0.7$, $P(x = 0) = 0.3$, $P(y = 1|x = 1) = 0.8$, $P(y = 1|x = 0) = 0.2$, $P(y = 0|x = 1) = 0.2$ and $P(y = 0|x = 0) = 0.8$.

Solution:

$$H(x) = P\,(x = 0)\log_2 \frac{1}{P(x = 0)} + P(x = 1)\log_2 \frac{1}{P(x = 1)}$$
$$H(x) = 0.3 \times \log_2 \frac{1}{0.3} + 0.7 \times \log_2 \frac{1}{0.7}$$
$$H(x) = \log_2 10 - \left(0.3 \times \log_2 3 + 0.7 \times \log_2 7\right)$$
$$H(x) = 0.881 \text{ bits}$$

So, the conditional entropy of x:

$$H(x|y = 0) = \frac{24}{38}\log_2 \frac{38}{24} + \frac{14}{38}\log_2 \frac{38}{24}$$

$$H(x|y=1) = \frac{3}{31}\log_2\frac{31}{2} + \frac{28}{31}\log_2\frac{31}{28}$$

$$\begin{aligned} H(x|y) &= 0.38\ H(x|y=0) + 0.62\ H(x|y=1) \\ &= P(y=0)H\ (x|y=0) + P(y=1)H(x|y=1) \\ &= 0.645. \end{aligned}$$

Therefore the mutual information of the binary symmetric channel shown in Figure 1.5 is:

$$I(x, y) = H(x) - \overline{H(X|y)} = 0.881 - 0.645 = 0.236 \text{ bits/symbol.}$$

1.3 Channel Capacity

Successful electrical/optical communication systems depend on how accurately the receiver can determine the transmitted signal. Perfect identification could be possible in the absence of *noise*, but noise is always present in communication systems. The presence of noise superimposed on signal limits the receiver ability to correctly identify the intended signal and thereby limits the rate of information transmission. The term 'noise' is used in electrical communication systems to refer to unwanted electrical signals that accompany the message signals. These unwanted signals arise from a variety of sources and can be classified as man-made or naturally occurring. Man-made noise includes such things as electromagnetic pickup of other radiating signals. Natural noise-producing phenomena include atmospheric disturbances, extraterrestrial radiation and internal circuit noise.

Definition: The capacity C of the channel is the maximum mutual information, taken over all input distribution of x. In symbols [3, 4]:

$$C = \max_{P(x_i)} I(x, y) \quad \text{where} \quad P(0) = P(1) = \frac{1}{2}. \tag{1.9}$$

The units of the capacity C are bits per channel input symbols (bits/s).

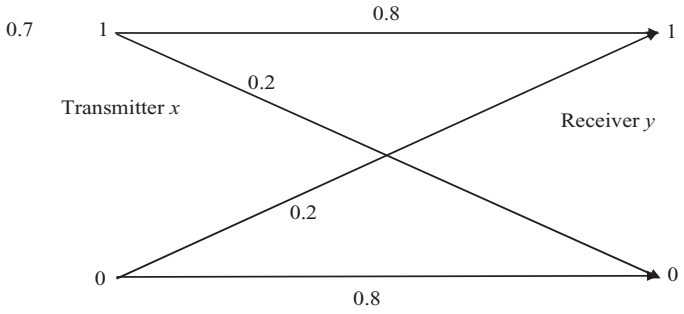

Figure 1.5 Binary symmetric channel of Example 1.5.

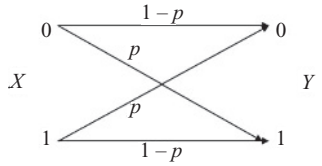

Figure 1.6 A binary symmetric channel.

1.3.1 Binary Symmetric Input Channel

Figure 1.6 shows the binary symmetric channel (BSC).

If we assume a uniform input distribution $P(x=0)=P(x=1)=1/2$ then $H(x)=1$. Furthermore, we have the set of transition probabilities [5]:

$$P(y=0\,|\,x=0)=P(y=1\,|\,x=1)=1-p$$
$$P(y=0\,|\,x=1)=P(y=1\,|\,x=0)=p\,.$$

The conditional entropy is given by:

$$H(x\,|\,y)=-\left[\sum_k\sum_i P(y_i\,|\,x_k)P(x_k)\log_2 P(x_k\,|\,y_i)\right]$$
$$H(x\,|\,y)=-\left[\frac{1}{2}(1-p)\log_2 P\,(x=0\,|\,y=0)+\frac{1}{2}p\log_2 P\,(x=0\,|\,y=1)\right.$$
$$\left.+\frac{1}{2}p\log_2 P\,(x=1\,|\,y=0)+\frac{1}{2}\,(1-p)\log_2 P\,(x=1\,|\,y=1)\right].$$

The distribution of y is determined as follows:

$$P(y=0)=P(y=0\,|\,x=0)P(x=0)+P(y=1\,|\,x=1)P(x=1)$$
$$\therefore\ P(y=0)=\frac{1}{2}(1-p)+\frac{1}{2}p=\frac{1}{2}\qquad,$$

and $P(y=1)$ is:

$$P(y=1)=P(y=1\,|\,x=0)P(x=0)+P(y=1\,|\,x=1)P(x=1)$$
$$=\frac{1}{2}\qquad.$$

The result is not surprising since the channel is symmetric. The joint distribution is determined as follows:

$$P(x=0\,|\,y=0)=\frac{P(x=0,\,y=0)}{P(y=0)}=\frac{P(y=0\,|\,x=0)p(x=0)}{P(y=0)}$$
$$P(x=0\,|\,y=0)=P(y=0\,|\,x=0)=1-p\qquad.$$
$$P(x=0\,|\,y=1)=P(x=1\,|\,y=0)=p$$

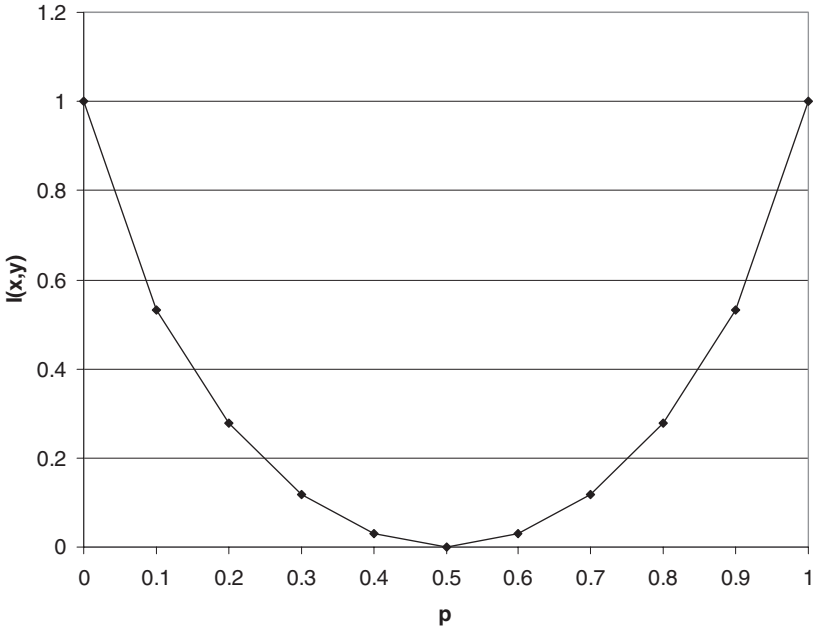

Figure 1.7 The mutual information of the BSC as a function of symbol error probability.

Thus:

$$\begin{aligned} H(x \mid y) &= -\left[\frac{1}{2}(1-p)\log_2(1-p) + \frac{1}{2}p\log_2 p + \frac{1}{2}(1-p)\log_2(1-p)\right] \\ &= -p\log_2 p - (1-p)\log_2(1-p) \\ I(x, y) &= 1 + p\log_2 p + (1-p)\log_2(1-p) \end{aligned} .$$

The mutual information as a function of the symbol error probability p is shown in Figure 1.7.

1.3.2 Binary Erasure Channel (BEC)

The BEC output alphabets are 0 or 1, plus an additional element, denoted as e, called the erasure. This channel corresponds to data loss. Each input bit is either transmitted correctly with probability $1 - p$ or is erased with probability p. The BEC is shown in Figure 1.8.

The channel probabilities are given by [1, 5]:

$$\begin{aligned} P(y = 0 \mid x = 0) &= P(y = 1 \mid x = 1) = 1 - p \\ P(y = e \mid x = 0) &= P(y = e \mid x = 1) = p. \end{aligned}$$

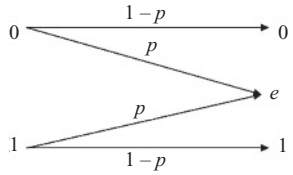

Figure 1.8 The binary erasure channel.

The conditional entropy is given by:

$$H(x \mid y) = -\left[\sum_k \sum_i P(x_k \mid y_i) P(x_k) \log_2 P(x_k \mid y_i)\right]$$

$$H(x \mid y) = -\left[\frac{1}{2}(1-p)\log_2 P\,(x=0 \mid y=0) + \frac{1}{2}p \log_2 P\,(x=0 \mid y=e) + \frac{1}{2}p \log_2 P\,(x=1 \mid y=1) + \frac{1}{2}(1-p)\log_2 P\,(x=1 \mid y=e)\right],$$

$$H(x \mid y) = -\frac{1}{2}(1-p)\log_2 P\,(x=0 \mid y=0) - \frac{1}{2}p \log_2 P\,(x=0 \mid y=e)$$

where we used the fact that X is equiprobably distributed. The distribution of y is determined as follows:

$$P(y=0) = P(y=0/x=0)P(x=0) = \frac{1}{2}(1-p)$$

and similarly for $P(y=e)$, we have

$$P(y=0) = \frac{1}{2},\ P(y=e) = p$$

and:

$$P(x=0 \mid y=0) = \frac{P(x=0,\, y=0)}{P(y=0)} = \frac{P(y=0 \mid x=0)p(x=0)}{P(y=0)} = \frac{\frac{1}{2}(1-p)}{\frac{1}{2}(1-p)}$$

and similarly:

$$P(x=0 \mid y=0) = 1$$

$$P(x=0 \mid y=e) = \frac{1}{2}.$$

Thus, $H(x \mid y) = -(1-p)\log_2(1) - p\log_2\left(\frac{1}{2}\right) = p$ and $I(x, y) = H(x) - H(x \mid y) = 1 - p$.

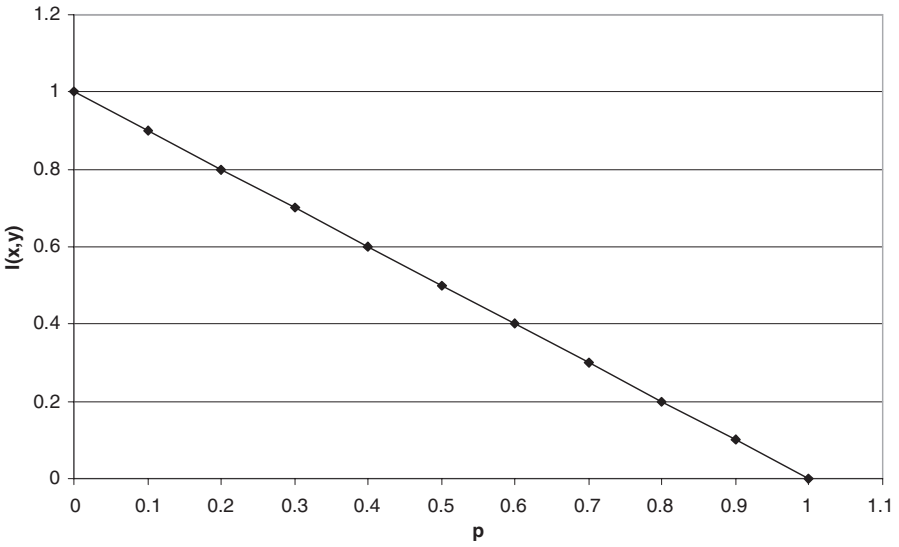

Figure 1.9 Mutual information of the binary erasure channel.

The mutual information of the binary erasure channel as a function of symbol error probability is plotted in Figure 1.9.

1.3.3 The 4-ary Symmetric Channel [1, 5]

The 4-ary Symmetric Channel (SC) is shown in Figure 1.10.

Let us consider the case of the 4-SC, where both the input X and the output Y have four possible values from the alphabet $A = [\alpha_0 \ldots \alpha_3]$. Since we are sending 4-ary symbols over the channel, we take the logarithm in equations $[H(x), I(x, y)]$ to the

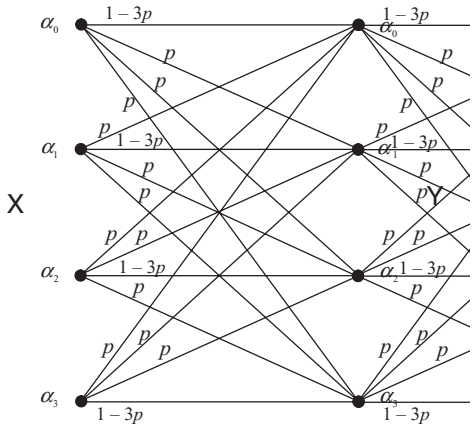

Figure 1.10 The 4-ary symmetric channel.

base 4. The same holds for the capacity defined previously. Assume a uniform input distribution where $P(x = a) = \frac{1}{4}$ for all $a \in A$. Hence, $H(x) = 1$, and we can set the transition probabilities:

$$P(y = a \mid x = a) = 1 - p,\ P(y = a \mid x = b) = \frac{p}{3} \text{ for } a \neq b, \text{ where } a, b \in A.$$

The conditional entropy is given by:

$$\begin{aligned} H(x \mid y) &= -\left[\sum_k \sum_i P(y_i \mid x_k)\ P(x_k) \log_4 P(x_k \mid y_i)\right] \\ &= -4\left[\frac{1}{4} p \log_4 P\ (x = a \mid y = b) + \frac{1}{4}(1 - p) \log_4 P\ (x = a \mid y = a)\right]. \end{aligned}$$

For $a \neq b$, where $a, b \in A$, the distribution of y is determined as follows:

$$\begin{aligned} P(y = a) &= 3P(y = a \mid x = b)P(x = b) + P(y = a \mid x = a)P(x = a) \\ &= 3 \times \frac{1}{4}\frac{p}{3} + \frac{1}{4}(1 - p) = \frac{1}{4}. \end{aligned}$$

The joint distribution is determined as follows:

$$P(x = a \mid y = a) = \frac{P(x = a, y = a)}{P(y = a)} = \frac{P(y = a \mid x = a)P(x = a)}{P(y = a)} = 1 - p.$$

Similarly:

$$P(x = a \mid y = b) = \frac{P(x = a, y = b)}{P(y = b)} = \frac{P(y = b \mid x = a)P(x = a)}{P(y = b)} = \frac{p}{3}.$$

For all $a, b \in A$, where $a \neq b$, we have:

$$\begin{aligned} H(x \mid y) &= -4\left[\frac{1}{4} p \log_4 \frac{p}{3} + \frac{1}{4}(1 - p) \log_4(1 - p)\right] \\ &= -p \log_4 \frac{p}{3} - (1 - p) \log_4(1 - p), \end{aligned}$$

and:

$$I(x, y) = 1 + p \log_4 \frac{p}{3} + (1 - p) \log_4(1 - p).$$

The mutual information as a function of the symbol error probability p is shown in Figure 1.11.

Notice that when the capacity is 1, the 4-ary symbol per channel use for $p = 0$. The mutual information is zero for $p = \frac{3}{4}$.

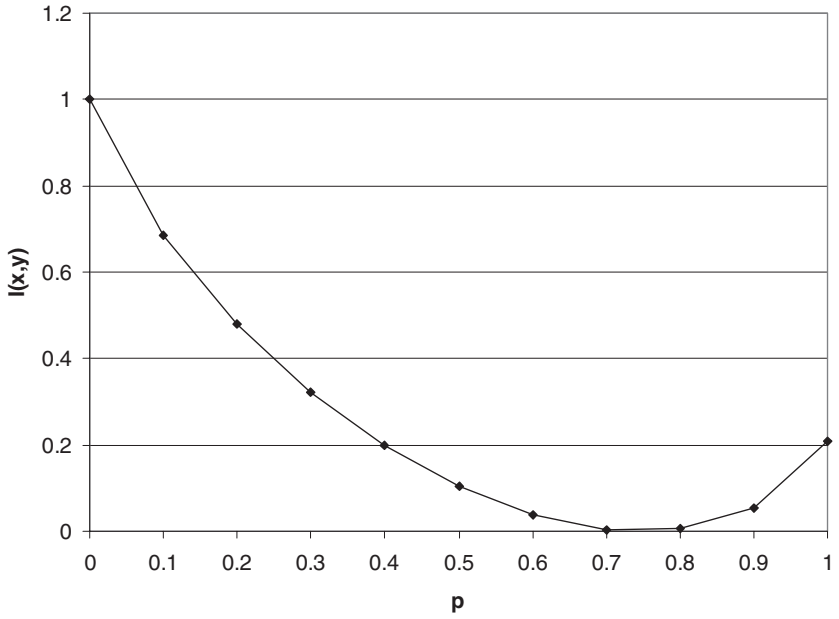

Figure 1.11 The mutual information of the 4-ary SC as function of symbol error probability.

1.3.4 Binary Input Capacity of a Continuous Channel (Channel Capacity)

Definition: The capacity C of a discrete memoryless channel is defined as the maximum mutual information $I(x, y)$ that can be transmitted through the channel:

$$\begin{aligned} I(X, Y) &= H(Y) - H(Y \mid x) = H(x) - H(x \mid y) \\ &= H(x) + H(y) - H(x, y). \end{aligned}$$

The capacity is:

$$\begin{aligned} C &= \max\ I(x, y) \\ C &= \max \left[\sum \sum P(x, y) \log_2 \frac{P(x, y)}{P(x)P(y)} \right]. \end{aligned}$$

If $P(x) = \frac{1}{\sqrt{2\pi S}} e^{-\frac{x^2}{2S}}$ is a Gaussian distribution for the signal S then $P(y) = \frac{1}{\sqrt{2\pi (S+N)}} e^{\frac{y^2}{2(S+N)}}$ is a Gaussian distribution for the signal S and the noise N.

We will make use of the expression for information $I(x, y) = H(y) - H(y \mid x)$ and hence we require to find $H(y \mid x)$. $P(y \mid x)$ will be Gaussian distributed about a particular value of x, as seen in Figure 1.12.

$$P(y \mid x) = \frac{1}{\sqrt{2\pi N}} e^{\frac{(y-x)^2}{2N}}.$$

Figure 1.12 The Gaussian distribution of the conditional probability $P(y \mid x)$.

It has been shown that the entropy is independent of the DC value and thus:

$$H(y) = \log_2 \sqrt{2\pi eN}.$$

Since this is independent of x:

$$\begin{aligned} H(y) &= \log_2 \sqrt{2\pi e(S+N)} \\ I(x, y) &= \log_2 \sqrt{2\pi e\,(S+N)} - \log_2 \sqrt{2\pi eN} \\ &= \log_2 \sqrt{\frac{2\pi e\,(S+N)}{2\pi eN}}. \\ &= \log_2 \sqrt{\frac{S+N}{N}} = \log_2 \sqrt{1+\frac{S}{N}} \end{aligned}$$

This is the information per sample. Thus in time T information is:

$$\begin{aligned} \text{Total information} &= 2BTI \\ &= 2BT \log_2 \sqrt{1+\frac{S}{N}} = BT \log_2 \left(1+\frac{S}{N}\right). \end{aligned}$$

Maximum information rate is:

$$C = B \log_2 \left(1+\frac{S}{N}\right) \text{ bits/s},$$

where B is the bandwidth of the channel. The practical and non-practical channel capacities of the AWGN channel are shown in Figure 1.13.

1. When the channel is noise free, permitting us to set $p = 0$, the channel capacity C attains its maximum value of one bit per channel use.

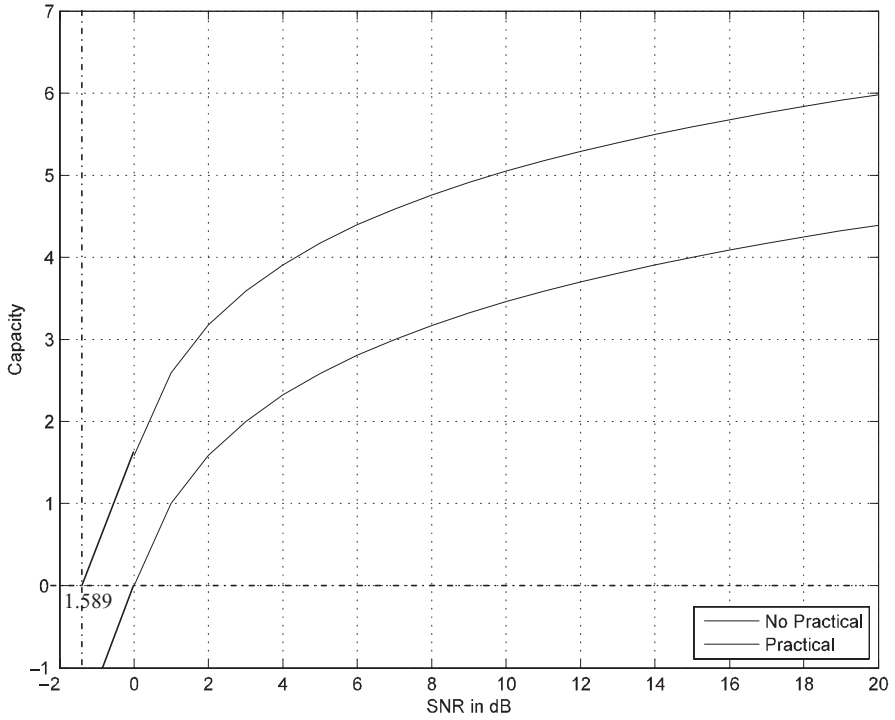

Figure 1.13 Channel capacity of the AWGN channel.

2. When the conditional probability of error $p = 1/2$ due to noise, the channel capacity C attains its minimum value of zero.

$$S = \text{Average Power (Symbols)}$$
$$N = \text{Average Noise Power.}$$

Example 1.6: Determine the capacity of a low-pass channel with usual bandwidth of 3000 Hz and $S/N = 10$ dB (signal/noise) at the channel output. Assume the channel noise to be Gaussian and white.

Solution:

$$C = B \log_2 \left(1 + \frac{S}{N}\right)$$
$$= 3000 \ \log_2 (1 + 10) \cong 10\,378 \text{ bits/s}.$$

Signalling at rates close to capacity is achieved in practice by error correction coding. An error correction code maps data sequences of k bits to code words of n

symbols. Because $n > k$, the code word contains structured redundancy. The code rate, $r = k/n$, is a measure of the spectral efficiency of the code. In order to achieve reliable communications, the code rate cannot exceed the channel capacity ($r \leq c$). The minimum theoretical signal-to-noise ratio (S/N) required to achieve arbitrarily reliable communications can be found by rearranging the equation for the capacity of the AWGN channel [2].

$$S/N \geq \frac{1}{2r}\left(2^{2r} - 1\right).$$

This is the minimum S/N required for any arbitrary distinction for the input signal. Shannon's proof of the channel coding theorem [1, 2] used a random coding argument. Shannon showed that if one selects a rate $r < c$ codes at random, the bit error probability approaches zero as the block length n of the code approaches infinity. However, random codes are not practically feasible. In order to be able to encode and decode with reasonable complexity, codes must possess some sort of a structure.

1.3.5 Channel Capacity in Fading Environments

In the case of a single input and single output (SISO) fading channel, the received signal at the kth symbol instant is $y(k) = h(k)x(k) + n(k)$, where $h(k)$ is the impulse response of the channel, $x(k)$ is the input signal to the channel and $n(k)$ is additive white Gaussian noise. To ensure a compatible measure of power, set $E(|h(k)|^2) = 1$ and $E\{|x(k)|^2\} \leq E_s$. The capacity in the case of the fixed fading channel with random but unchanging channel gain is given below [6]:

$$C = B\log_2\left(1 + |h|^2\,\rho\right), \quad \text{where } \rho = \frac{E_s}{\sigma_n^2}, \tag{1.10}$$

where E_s is the symbol energy, σ_n^2 is the n-dimensional variance and h is the gain provided by the channel. An interesting point to note is that in a random but fixed fading channel the theoretical capacity can be zero, when the channel gain is close enough to zero to make data rate impossible. In this case, the possible scenario is determining what the chances are that a required capacity is available. This is defined by Outage probability P_{out} as the probability that the channel is above a threshold capacity C_{thres} given by following equation [6]:

$$P_{\text{out}} = P(C > C_{\text{thres}}) = P\left(|h|^2 > \frac{2C_{\text{thres}} - 1}{\rho}\right) \tag{1.11}$$

$$P_{\text{out}} = 1 - \mathrm{e}^{-\frac{2C_{\text{thres}}-1}{\rho}}, \tag{1.12}$$

where (1.13) is valid for Rayleigh fading.

In the case of a time-varying channel, the channel is independent from one symbol to the next and the average capacity of K data symbols is:

$$C_K = \frac{1}{K}\sum_{k=1}^{K}\left\{\log_2\left(1+|h|^2\,\rho\right)\right\}. \tag{1.13}$$

Based on the law of large numbers, as $K \to \infty$ the term on the right converges to the average or expected value, therefore:

$$C = E_h\left\{\log_2\left(1+|h|^2\,\rho\right)\right\}, \tag{1.14}$$

where the expectation is taken over the channel values h. (1.14) is nonzero; therefore with a fluctuating channel it is possible to guarantee the existence of an error-free data rate.

Increasing the transmitted and received antenna, the capacity of the channel increases by $N_T\ N_R$-fold, where N_T is number of transmit antenna and N_R is number of receive antenna. The capacity of the AWGN channel in case of multiple-input–multiple-output is approximately given by (1.15):

$$C \approx B\log_2\left(1+N_T\cdot N_R\cdot\rho\right) \quad \text{where } \rho = E_s/\sigma_n^2. \tag{1.15}$$

Figure 1.14 shows the channel capacity with the increase in the number of transmitted and received antennas plotted using (1.15).

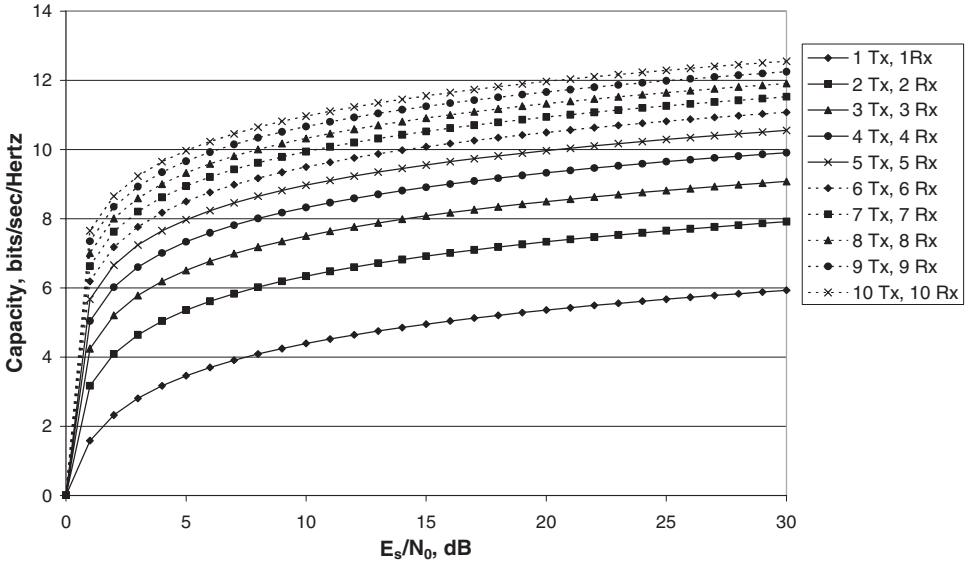

Figure 1.14 Channel capacity in bits/s/Hz with increasing number of transmit and receive antennas.

1.4 Channel Modelling

This section begins by describing the fundamentals and characteristics of the channel propagation models used in WiMAX or IEEE 802.16d (Fixed Broadband) systems and UMTS cellular communications. Propagation models are the fundamental tools for designing any wireless communication system. A propagation model basically predicts what will happen to the transmitted signal while in transit to the receiver. In general, the signal is weakened and distorted in particular ways and the receiver must be able to accommodate for these changes if the transmitted information is to be successfully received. The design of the transmitting and receiving equipment and the type of communication service that is being provided will be affected by these signal impairments and distortions. The role of propagation modelling is to predict the performance of the system with these distortions and to determine whether it will successfully meet its performance goals and service objectives.

The section ends with a description of magnetic storage channel modelling for longitudinal and perpendicular storage devices. In this situation, information is retrieved not in a different *location* from where it was originally sent, as in radio communications, but at a different *time* from when it was originally stored.

1.4.1 Classification of Models

In wireless channels, a narrowband implies that the channel under consideration is sufficiently narrow that the fading across it is flat (i.e. constant). It is usually used as an idealizing assumption; no channel has perfectly flat fading, but the analysis of many aspects of wireless systems is greatly simplified if flat fading can be assumed. Early communication systems were narrowband systems in which median signal level prediction and some description of the signal level variability (fading) statistics were the only model parameters needed to adequately predict the system performance. However, modern communications systems use wideband, hence achieving higher data rates.

In communications, wideband is a relative term used to describe a wide range of frequencies in a spectrum. A system is typically described as wideband if its bandwidth is much greater than its centre frequency or carrier frequency. For such systems, narrowband prediction of signal levels and fading alone does not provide enough information to predict system performance. In fact the concept of propagation models is now enlarged to include the entire transfer function of the channel. These new propagation models, known as channel models, now include parameters such as signal time dispersion information and Doppler effect distortion arising from motion of the mobile device. Time dispersion causes the signal fading to vary as a function of frequency, so wideband channels are often known as frequency-selective fading channels.

These propagation channel models are broadly classified into three main categories: Theoretical, Empirical and Physical.

1.4.1.1 Theoretical Models

The channel models in this group are based on theoretical assumptions about the propagation environment. The channel parameters based on these model assumptions do not define any particular environment or model any specific channel conditions. The theoretical models cannot be used for planning and developing any communication systems as they do not reflect the exact propagation medium the signal would be experiencing. The theoretical models can be nontime dispersive or time dispersive. Nontime-dispersive channel models are those in which the duration of the transmitted signal is the same on arriving at the receiving end. However, in time dispersion the signal extends in time so that the duration of the received signal is greater than that of the transmitted signal. The theoretical modelling of the time-dispersive channel has been presented in [6–8]. The theoretical time-dispersive channel can also be modelled by the tapped delay line structure, in which densely-spaced taps, multiplying constants and tap-to-tap correlation coefficients are determined on the basis of measurements or some theoretical interpretation of how the propagation environment affects the signal [7, 8].

1.4.1.2 Empirical Models

Empirical models are those based on observations and measurements alone. These models are mainly used to predict the path loss, but models that predict rain fade and multipath have also been proposed [7]. The problem can occur when trying to use empirical models in locations that are broadly different from the environment in which the data is measured. For example, the Hata Model [9] is based on the work of Okumura, in which the propagation path loss is defined for the urban, suburban and open environments. But models like Hata and ITU-R are widely used because they are simple and allow rapid computer calculations.

Empirical models can be subclassified in two categories, namely nontime dispersive and time dispersive, as described. The time dispersive provides information relating to the time-dispersive characteristics of the channel, that is the multipath delay spread of the channel. Examples of this type are channel models developed by Stanford University Interim (SUI) for use in setting up the fixed broadband systems. These types of channel model are extensively used for WiMAX or IEEE 802.16 system development [10]. This chapter is mainly concerned with the time-dispersive empirical channel model (WiMAX specification), which will be explained in detail later on.

1.4.1.3 Physical Models

A channel can be physically modelled by attempting to calculate the physical processes which modify the transmitted signal. These models rely on basic principles of physics rather than statistical outcomes from the experiments. Physical channel models are also known as deterministic models, which often require a complete 3D map of the

propagation environment. These models are not only divided into nontime dispersive and time dispersive, but are also modelled with respect to the site specification.

The works published in [6, 8] are examples of site-specific, time-dispersive channel models. These models are principally identified as ray-tracing, a high-frequency approximation approach that traces the route of electromagnetic waves leaving the transmitter as they interrelate with the objects in the propagation environment. A deterministic ray-tracing propagation model was used to predict the time delay and fading characteristics for the channel in a hypothetical urban area. Using this propagation model, the channel response throughout the urban area was described in terms of the signal level, root-mean-square (RMS) delay spread, and the fading statistics at each point in the service area. These time-dispersive models provide not only multipath delay spread of the channel but also information related to the angle of arrival (AoA) of the signal.

However, those physical models which do not have signal time delay information are known as nontime-dispersive channel models. These propagation models are specifically applicable to propagation prediction in the fixed broadband wireless systems.

1.5 Definition of a Communications Channel and its Parameters

The main factor affecting the design of a fixed wireless access system is the nature of the channel available, which affects the behaviour of electromagnetic waves propagating through it. The previous subsection presented the different types of channel modelling procedures and their subcategories. Before dealing with the specification and technical parameters of this model, a general background of the channel is presented along with the definitions of parameters like delay spread, path loss, fading, Doppler effect and so on.

The main processes of a communication system consist of a source, a transmitter, a channel, a receiver and a destination, as shown in Figure 1.15, where b_k are the message bits, x_n are the modulated symbols, r_n are received symbols, $\hat{x}_n$ are the demodulated symbols and $\hat{b}_k$ are the detected message bits. The transmitter takes information from the source and converts it into a form suitable for transmission.

The wireless communication channel consists of the medium through which the RF signal passes when travelling from the transmitting antenna to the receiving antenna. The medium causes the transmitted signal to be distorted in several ways, as previously mentioned.

In the absence of a line-of-sight between the transmitting antenna and the receiving antenna, some of the transmitted signal finds a path to the receiving antenna by

Figure 1.15 Simple communication system.

reflecting or refracting from whatever is blocking the direct line-of-sight between the two antennas. This action is known as a *multipath* signal scenario as the many transmitted signals undergo different degrees of dispersion as they traverse multiple paths to reach the receiving antenna.

Eventually some of the signal paths recombine vectorially at the receiving antenna, producing a signal the amplitude of which is dependent upon the phases of the individual component waveforms. This is the concept of signal fading, which is a purely spatial phenomenon. If a mobile receiving antenna is moving relative to the environment and/or the transmitting antenna, the incoming phases of the signals will vary, producing a signal whose amplitude varies with spatial movement of the mobile relative to the environment. Although fading is a spatial phenomenon, it is often perceived as a temporal phenomenon as the mobile device moves through the multipath signal field over time.

To determine a channel model, mathematical descriptions of the transmitter, the receiver and the effect of the environment (walls) on the signal must be known. A linear channel can be totally described by its impulse response, that is by what the received signal would look like if the transmitted signal were an impulse. The impulse response is the response of the channel at all frequencies, that is, once the impulse response of the channel is known, one can predict the channel response at all frequencies.

Let $x(t)$ be the signal transmitted from an antenna through the channel $h(t)$, and $y(t)$ be the signal received at the receiving side. Assuming no delay, multipath signals and no other noise present in the system, the channel can be considered as a linear system with $x(t)$ as an input and $y(t)$ as an output. This relationship between the input and output in the time domain is represented in (1.16) and shown in Figure 1.16.

$$y(t) = h(t) \otimes x(t). \tag{1.16}$$

The channel impulse response $h(t)$ is obtained by applying the impulse function to the channel and can be represented as:

$$h(t) = \sum_{i=0}^{\infty} A_i e^{j\phi_i} \delta(t - \tau_i), \tag{1.17}$$

where A_i is the magnitude of the impulse response at delay τ_i with phase ϕ_i, and $\delta(t)$ is the Dirac delta function [5]. The system in Figure 1.16 is modelled as a linear

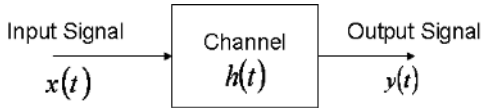

Figure 1.16 Channel input-output in time domain.

time variant filter with impulse response $h(t, \tau)$ and frequency response $H(f, t)$. The frequency response can be obtained by taking the Fourier transform of $h(t)$:

$$H(f, t) = \int_{-\infty}^{\infty} h(t) e^{-j2\pi f t} \mathrm{d}t. \tag{1.18}$$

The channel output $y(t)$ at a particular time $t = \tau$ is the convolution of the impulse response $h(t)$ with the input signal $x(t)$, that is:

$$y(t) = x(t) \otimes h(t) = \sum_{\tau=0}^{\infty} h(\tau) x(t - \tau), \tag{1.19}$$

where $\otimes$ represents the convolution operation. The final output $y(t)$ can be obtained by adding independent noise $n(t)$ to (1.19):

$$y(t) = x(t) \otimes h(t) + n(t) = \sum_{\tau=0}^{\infty} h(\tau) x(t - \tau) + n(t). \tag{1.20}$$

However, in a wireless environment the transmitted signals are affected in different ways, for example by the reflection that occurs when the signal hits a surface. At the receiving side the signal arrives from different paths, which add delay at the receiver, and it can be affected by moving objects in terms of a Doppler shift.

1.5.1 Doppler Shift

Doppler frequency effects occur whenever there is relative movement between the transmitting and receiving antennas. It manifests itself as a change in the frequency of the received signal. For a receiver tuned to a particular frequency, this phenomenon has the effect of reducing the received signal energy as the receiver's front end is no longer tuned to the centre frequency of the signal and is therefore not operating efficiently.

The Doppler frequency shift in each propagation path is caused by the rate of change of signal phase (due to motion). Referring to Figure 1.17 below, if a mobile device, moving from point A to point B at a velocity v, is receiving a signal from the signal source, the distance travelled, d, can be found by $v\Delta t$, where Δt is the change in time from point A to point B.

From this it can be shown geometrically that the extra distance that the wave has to travel to get from the signal source to point B(Δl) with respect to point A is: $d \cos(\alpha)$.

The change in phase of the received signal at point B relative to point A is given by [8]:

$$\Delta\phi = -\frac{2\pi}{\lambda} \Delta l = -\frac{2\pi v \Delta t}{\lambda} \cos\alpha. \tag{1.21}$$

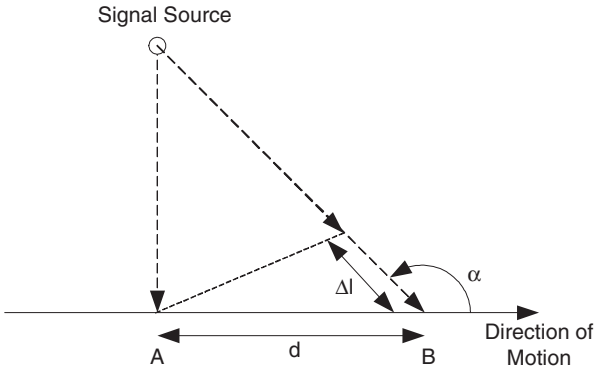

Figure 1.17 Example of the effects of Doppler shift [8].

With the change in frequency of the received signal at point B relative to point A given by:

$$f = -\frac{1}{2\pi}\frac{\Delta\phi}{\Delta t} = \frac{v}{\lambda}\cos\alpha. \tag{1.22}$$

it can be seen that the change in path length is governed by the angle between the direction of motion and the received wave. It should also be noted that when the mobile antenna is moving closer to the signal source, a positive change in frequency (Doppler) is caused, and conversely, if the mobile antenna is moving further away from the signal source it causes a negative change in frequency.

If the mobile antenna is moving in the same plane as the signal source (either to or from it) then the frequency shift is given by:

$$f = \pm v/\lambda. \tag{1.23}$$

This information is required for modelling the channel using the ray-tracing method, which calculates all the angles of the received signals in order to calculate the received signal strength over time. This method becomes impractical for large numbers of received signals due to the computational burden. The number of signal reflections also limits this method, and in practice only signals with two reflections are considered.

The movement of the receiving antenna relative to the transmitting antenna, coupled with the large number of reflections and received signal paths, causes the resultant RF signal envelope at the receiving antenna to appear random in nature. Therefore, statistical methods must be employed in order to produce a mathematically tractable model, which produces results in accordance with the observed channel properties.

1.5.2 Scattering

At the receiving antenna, the incoming signal consists of a superposition of individual waves. Each individual wave has the following characteristics: amplitude, frequency, phase, polarization angle, vertical angle of arrival and horizontal angle of arrival, all relative to the signal at the transmitter.

If we assume that both the transmitting and receiving antennas are similarly polarized and operating on the same frequency, only the amplitude (A_n), phase (ϕ_n), vertical angle of arrival (β_n) and horizontal angle of arrival (α_n) need to be considered. Figure 1.18 shows an individual wave (n) relative to the point of reception.

The values for A_n, ϕ_n, β_n and α_n are all random and statistically independent. The mean square value for the amplitude (A_n) is given by:

$$E\left\{A_n^2\right\} = \frac{E_0}{N}, \tag{1.24}$$

where E_0 is a positive constant and N is the number of received waves at the point of reception.

Clarke [7] makes a generalization which assumes that the height of the transmitting antenna is approximately the same as that of the receiving antenna, so (assuming Clarke's model) the vertical angles of arrival (β_n) can be set to zero. In practice this is found to be a good approximation.

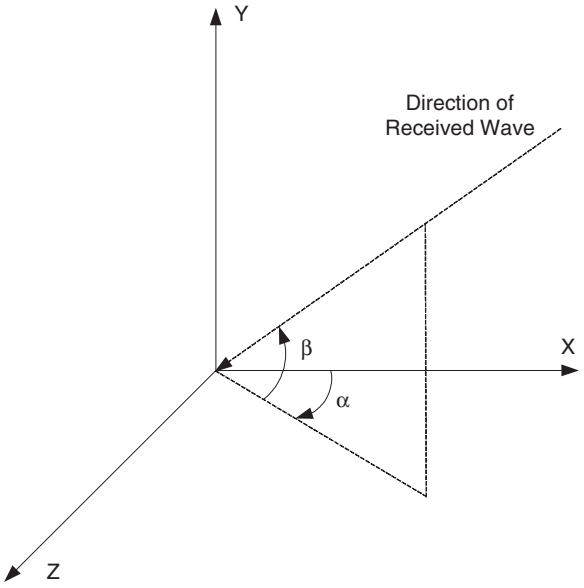

Figure 1.18 An individual wave relative to the point of reception.

The phase angles (ϕ_n) are uniformly distributed within the range of 0 to 2π, but the probability density functions for the angles of arrival α_n and β_n are generally not specified.

At the point of reception, the field resulting from the superposition of the incoming waves (n) is given by:

$$E(t) = \sum_{n=0}^{N} E_n(t). \tag{1.25}$$

If an unmodulated carrier is transmitted, the received signal ($E_n(t)$) can be expressed at the point of reception (x_0, y_0, z_0) as follows [8]:

$$E(t) = A_n \cos\left(\omega_0 t - \frac{2\pi}{\lambda}(x_0 \cos(\alpha_n)\cos(\beta_n) + y_0 \sin(\alpha_n)\cos(\beta_n) + z_0 \sin(\beta_n)) + \phi_n\right). \tag{1.26}$$

If the receiving antenna moves in the xy plane at an angle γ (relative to the x-axis) with a velocity v, then after a unit time, the coordinates of the received signal can be expressed as:

$$v\cos\gamma, v\sin\gamma, z_0. \tag{1.27}$$

Which means that the received signal $E(t)$ can be expressed as:

$$E(t) = I(t)\cos(\omega_c t) - Q(t)\sin(\omega_c t), \tag{1.28}$$

where $I(t)$ and $Q(t)$ represent the in-phase and quadrature components of the signal respectively, and can be expressed as:

$$I(t) = \sum_{n=1}^{N} A_n \cos(\omega_n t + \theta_n) \tag{1.29}$$

and:

$$Q(t) = \sum_{n=1}^{N} A_n \sin(\omega_n t + \theta_n), \tag{1.30}$$

where ω_n equals $2\pi f_n$, and f_n represents the Doppler shift in frequency experienced by the individual wave n. The terms ω_n and θ_n can be expressed as:

$$\omega_n = \frac{2\pi v}{\lambda}\cos(\gamma - \alpha_n)\cos(\beta_n) \tag{1.31}$$

and:

$$\theta_n = \frac{2\pi z_0}{\lambda}\sin(\beta_n) + \phi_n. \tag{1.32}$$

The previous equations reduce to Clarke's two-dimensional model if β is taken to be zero.

If the number of received waves (N) is very large (in practice 6, but theoretically much larger), then according to the central limit theorem the components $I(t)$ and $Q(t)$ are independent Gaussian processes and are completely characterized by their mean and autocorrelation functions. Because $I(t)$ and $Q(t)$ both have zero mean values, it follows that $E\{E(t)\}$ is also zero. Also, $I(t)$ and $Q(t)$ have variance values ( ) that are the same as the mean square value (average power), thus the probability density function of $I(t)$ and $Q(t)$ can be expressed as:

$$p_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{1.33}$$

where $x = I(t)$ or $Q(t)$, and $\sigma^2 = E\{A_n^2\} = \dfrac{E_0}{N}$.

1.5.3 Angle of Arrival and Spectra

For a system in motion (that is, receiver antenna movement relative to the transmitting antenna), the individual components of the received signal will experience a change in frequency due to the Doppler effect, depending upon their individual angles of arrival. This change in frequency is given by [8]:

$$f_n = \frac{\omega_n}{2\pi} = \frac{v}{\lambda}\cos(\gamma - \alpha_n)\cos(\beta_n). \tag{1.34}$$

All frequency components within the received signal will experience this Doppler shift in frequency and, as long as the signal bandwidth is relatively small (compared to the receive signal path bandwidth), it can be assumed that all the individual component waves will be affected in the same way.

The RF spectrum of the received signal can be obtained by using the Fourier transform of the temporal autocorrelation function in terms of the time delay (τ):

$$\begin{aligned} E\{E(t)E(t+\tau)\} &= E\{I(t)I(t+\tau)\}\cos(\omega_c\tau) - E\{I(t)Q(t+\tau)\}\sin(\omega_c\tau) \\ &= a(\tau)\cos(\omega_c\tau) - c(\tau)\sin(\omega_c\tau). \end{aligned} \tag{1.35}$$

Aulin [11] showed that the correlation properties can therefore be expressed by $a(\tau)$ and $c(\tau)$:

$$\begin{aligned} a(\tau) &= \frac{E_0}{2} E\{\cos(\omega\tau)\} \\ c(\tau) &= \frac{E_0}{2} E\{\sin(\omega\tau)\}. \end{aligned} \tag{1.36}$$

In order to simplify further we make the assumption that all the signal waves arrive in the horizontal plane (α) with equal probability, so that:

$$p_\alpha(\alpha) = \frac{1}{2\pi}. \tag{1.37}$$

Which means that by Fourier transforming the following, the power spectrum can be obtained:

$$a(\tau) = \frac{E_0}{2} \int_{-\pi}^{+\pi} J_0(2\pi f_m \tau \cos\beta) p_\beta(\beta)\, \mathrm{d}\beta. \tag{1.38}$$

1.5.4 Multipath Channel and Tapped Delay Line Model

The fixed or mobile radio channel is characterized by the multipath propagation. The signal offered to the receiver contains not only a direct line-of-sight radio wave, but also a large number of reflected radio waves. This can be even worse in urban areas, where obstacles often block the line-of-sight and a mobile antenna receives a collection of various delayed waves. These reflected waves interfere with the direct wave, which causes significant degradation of the performance of the link. If the mobile user moves, the channel varies with the location and time, because the relative phase and amplitude of the reflective wave change. Multipath propagation seriously degrades the performance of the communication system. The adverse effects produced by the medium can be reduced by properly characterizing the medium in order to design the transmitter and receiver to fit the channel. Figure 1.19 shows the signal arriving at the receiver from different paths, which include the direct line-of-sight (LOS) path and nonline-of-sight (NLOS) paths.

As the different variants of the same signal arrive at different times, some are delayed relative to one another. This time dispersion is known as multipath delay spread. This delay spread is an important parameter in channel modelling and is commonly measured as root mean square (RMS) delay spread. For reliable communication over these channels without any interference reduction techniques, the transmitted data rate should be much smaller than the coherence bandwidth. This type of channel, when

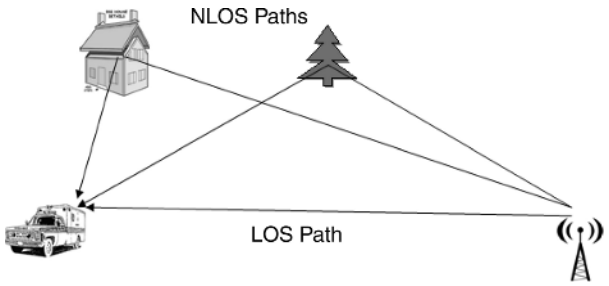

Figure 1.19 Signal arriving at mobile station from different paths.

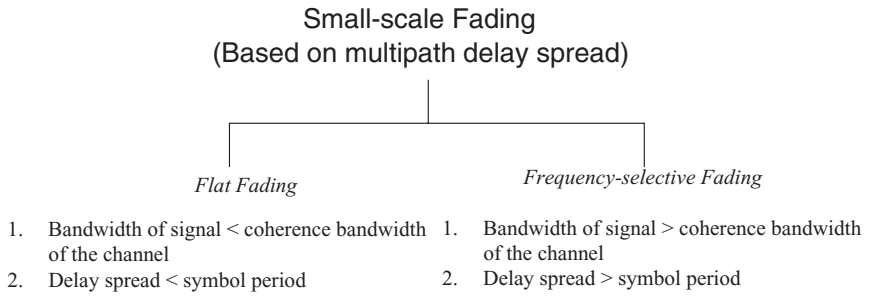

Figure 1.20 The two classes of fading channels.

the transmitted bandwidth of the signal is much smaller than the coherent bandwidth, is known as a flat channel or narrowband channel. When the transmitted bandwidth of the signal is nearly equal to or larger than the coherent bandwidth, the channel is known as a frequency-selective channel or broadband channel. The relationship between flat fading and frequency-selective fading is shown in Figure 1.20.

The multipath delay profile is characterized by τ_{rms}, which is defined as [6]:

$$\tau_{\text{rms}}^2 = \sum\nolimits_j P_j \tau_j^2 - (\tau_{\text{avg}})^2, \tag{1.39}$$

where $\tau_{\text{avg}} = \sum_j P_j \tau_j$, τ_j is the delay of the jth delay component of the profile and $P_j =$ (power in the jth delay component)/(total power in all components).

The channel output $y(t)$ of (1.22) can be realized as a tapped delay line, as shown in Figure 1.21, where $A_i(t)$ is the fading amplitude of the ith tap and T_i are the delays of the ith tap. Each tap represents a scattered ray multiplied by a time-varying fading profile coefficient $A_i(t)$. The relative tap delays are dependent on the type of channel

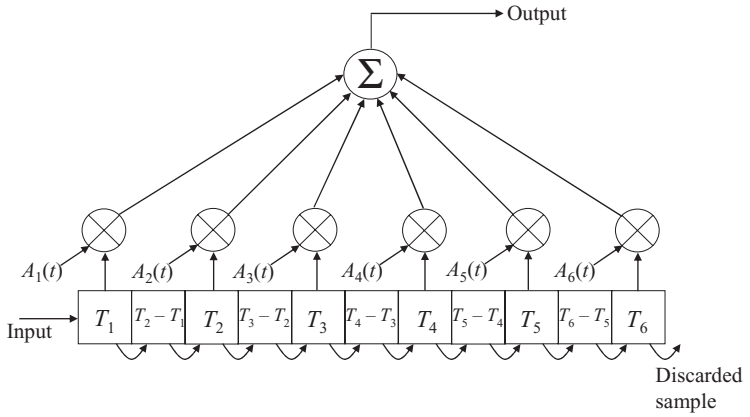

Figure 1.21 An example of a tapped delay line.

model. These values are given later for the fixed broadband wireless-access channel model and the urban UMTS mobile radio channel models.

1.5.5 Delay Spread

Because of the multipath reflections, the channel impulse response of a wireless channel looks like a series of pulses. In practice the number of pulses that can be distinguished is very large and depends on the time resolution of the communication or measurement system.

The system evaluations process typically prefers to address a class of channels with properties that are likely to be encountered, rather than one specific impulse response. Hence it defines the (local-mean) average power, which is received with an excess delay that falls within the interval $(T, T + dt)$. Such characterization for all T gives the 'delay profile' of the channel. The delay profile determines the frequency dispersion; that is, the extent to which the channel fading at two different frequencies f_1 and f_2 is correlated (Figure 1.22).

The maximum delay time spread is the total time interval during which reflections with significant energy arrive. The RMS delay spread T_{RMS} is the standard deviation (or root-mean-square) value of the delay of reflections, weighted proportionally to the energy in the reflected waves. For a digital signal with high bit rate, this dispersion is experienced as frequency-selective fading and intersymbol interference (ISI). No serious ISI is likely to occur if the symbol duration is longer than, say, ten times the RMS delay spread. The RMS delay spread model published in [14] follows a lognormal distribution and the median of this distribution grows as a power of distance. This model was developed for rural, urban and mountainous environments (Figure 1.23).

$$\tau_{\mathrm{rms}} = T_1 d^e y, \tag{1.40}$$

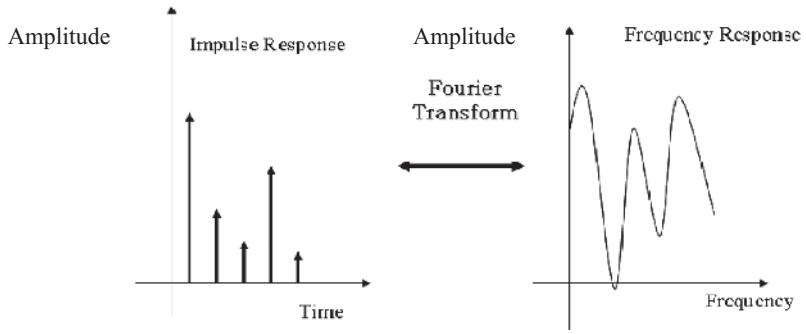

Figure 1.22 Example of the impulse response and frequency transfer function of a multipath channel.

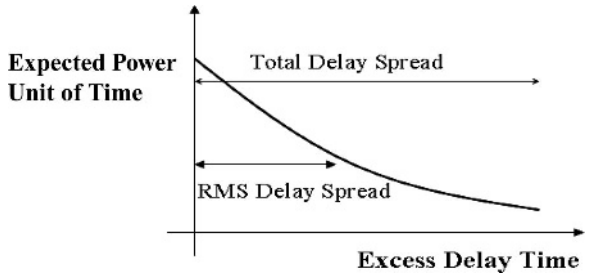

Figure 1.23 Expected power per unit of time.

where τ_{rms} is the RMS delay spread, d is the distance in km, T_1 is the median value of the $\tau_{\text{rms at}}$ $d = 1$ km, e is an exponent that lies between 0.5 and 1.0, and y is the lognormal variable.

The narrowband received signal fading can be characterized by a Ricean distribution. The K factor is the ratio between the constant component powers and the scattered power. The model that represents the K factor by extensive experimental data taken at 1.9 GHz is given below [12]:

$$K = F_s F_h F_b K_o d^{?}, \tag{1.41}$$

where F_s is a seasonal factor: $F_s = 1.0$ in summer (leaves) and 2.5 in winter (no leaves); F_h is the receive antenna height factor, $F_{\text{h}} = (h/3)^{0.46}$ (h is the receive antenna height); F_b is the beamwidth factor, $F_b = (b/17)^{-0.62}$ (b in degrees); K_o and γ are regression coefficients, $K_o = 10$; $\gamma = -0.5$; u is a lognormal variable with 0 dB mean and a standard deviation of 8.0 dB.

1.5.6 The Fixed Broadband Wireless Access Channel

The Doppler spectrum of the fixed wireless system is different from that in the mobile wireless access systems. In fixed wireless systems the Doppler spectrum of the scattered component is mainly scattered around $f = 0$ Hz. The shape of the spectrum is different from the mobile channels. The spectrum for mobile and fixed wireless channels is given below [10].

The above fixed access channel power spectrum density (PSD) is calculated by using the equation given below:

$$S(f) = \begin{cases} 1 - 1.72 f_o^2 + 0.785 f_o^4 & |f_o| \le 1 \\ 0 & |f_o| > 1 \end{cases} \quad \text{where, } f_o = \frac{f}{f_m}. \tag{1.42}$$

Figure 1.24 is based on (1.42) and is the rough approximation of the Doppler PSD, which has the advantage that it is readily available in most existing RF simulators. This Doppler spectrum is representative of fixed mobile (point-to-point) wireless

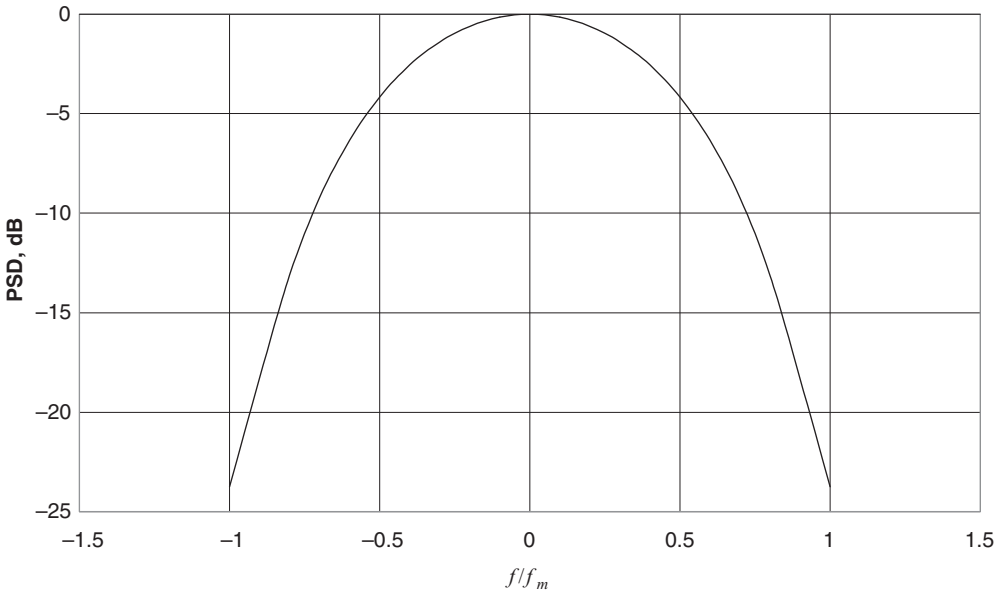

Figure 1.24 Doppler spectrum of a fixed wireless access channel.

channels and so does not represent the Doppler characteristics of a nonstationary mobile wireless channel.

1.5.7 UMTS Mobile Wireless Channel Model

Three different propagation environments for the UMTS mobile wireless channel model are considered: Indoor, Pedestrian and Vehicular. These environments were chosen because they are the main scenarios that cellular mobile radios experience in normal operation.

The Indoor environment model is characterized by having no Doppler frequency shift, as the velocities concerned within the indoor environment are well below walking pace ($\ll$4 miles per hour (MPH)), which produces either zero or a negligible shift in frequency. Within this environment there is no direct line-of-sight path from transmitter to receiver, so the signal propagates via many different paths as the signals are reflected, refracted or diffracted. Only six paths are simulated as the power contained within the strongest six rays are deemed strong enough to be included; the rest are very low power and thus will not affect the results appreciably. If they were included they would just serve to slow the simulation down.

The Pedestrian environment model is characterized by having a small Doppler frequency shift, as the velocities concerned within the pedestrian environment are around walking pace ($\approx$4 MPH), which produces a small Doppler shift in frequency. Within this environment there is limited/no direct line-of-sight path from transmitter to

receiver, so the signal propagates via many different paths as the signals are reflected, refracted or diffracted. Only four paths are simulated as only the power contained within the strongest four rays are deemed strong enough to be included; the rest are very low power and thus will not affect the results appreciably.

The Vehicular environment model is characterized by having a larger Doppler frequency shift as the velocities concerned within the vehicular environment are reasonably large (≈70 MPH), producing a larger Doppler shift in frequency. Within this environment there is no direct line-of-sight path from transmitter to receiver, so the signal propagates via many different paths as the signals are reflected, refracted or diffracted. Only six paths are simulated as only the power contained within the strongest six rays are deemed strong enough to be included; the rest are very low power and thus will not affect the results appreciably.

A Doppler spectrum is given in (1.43), which is the classic spectrum for mobile radio channels, and is used by Jakes [8] and Clarke [7]:

$$S(f) = \begin{cases} \dfrac{E_0}{4\pi f_m} \dfrac{1}{\sqrt{1 - \left(\dfrac{f}{f_m}\right)^2}} & |f| \leq f_m \\ 0 & \text{elsewhere} \end{cases}, \tag{1.43}$$

where E_0 = energy constant. Although this channel is representative of the mobile wireless channels, there are problems with this representation as the power spectral density becomes infinite at $f_C \pm f_m$. In order to find a more realistic (and useable) representation, Aulin [11] described the following:

$$S(f) = \begin{cases} 0 & \forall\ |f| > f_m \\ \dfrac{E_0}{4 \sin \beta_m} \dfrac{1}{f_m}, & \forall\ f_m \cos \beta_m \leq |f| \leq f_m \\ \dfrac{1}{f_m} \left[\dfrac{\pi}{2} - \arcsin \dfrac{2\cos^2 \beta_m - 1 - (f/f_m)^2}{1 - (f/f_m)^2} \right], & \forall\ |f| < f_m \cos \beta_m \end{cases}. \tag{1.44}$$

This claimed to be realistic for small values of β_{m} (angle of signal arrival) and is particularly useful in providing analytic solutions. The problem with this model is that there are sharp discontinuities at $\pm\beta_{\mathrm{m}}$, which causes an unrealistic response. Therefore a model is required which produces the classic Doppler shape for mobile wireless channels but has no infinities or sharp discontinuities. Such a model was proposed by Parsons [6], and the PDF of the angle of arrival $p_\beta(\beta)$ is given by:

$$p_\beta(\beta) = \begin{cases} \dfrac{\pi}{4|\beta_m|} \cos\left(\dfrac{\pi}{2} \cdot \dfrac{\beta}{\beta_m}\right) & |\beta| \leq |\beta_m| \leq \dfrac{\pi}{2} \\ 0 & \text{elsewhere} \end{cases}. \tag{1.45}$$

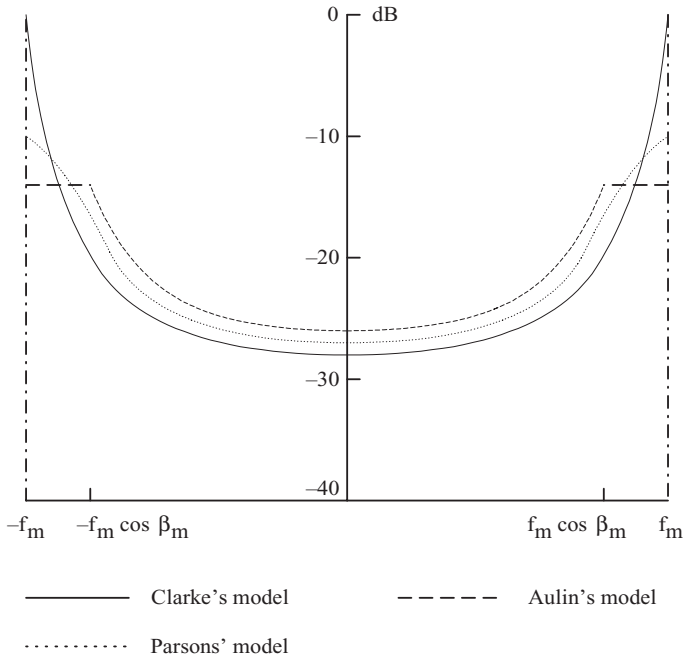

Figure 1.25 Doppler spectra of a mobile wireless access channel.

Using (1.45), the power spectral density can be expressed as:

$$S(f) = \mathrm{FT}\left[\frac{E_0}{2}\int_{-\pi}^{+\pi} J_0\left(2\pi f_m \tau \cos\beta\right) p_\beta\left(\beta\right)\, \mathrm{d}\beta\right], \tag{1.46}$$

where FT is the Fourier transform and J_0 is the zero-order Bessel function of the first kind.

More information on the origin of these equations can be found in [7]. The Doppler spectra of (1.43), (1.44) and (1.46) are shown in Figure 1.25.

1.5.8 Simulating the Fixed Broadband Wireless Access Channel Model

There are different SUI(Stanford University Interim) models presented in [13]. The description of the SUI-3 model is explained here, and the model description and parameters are given in Table 1.2. A three-tap model is described here to represent the multipath scenario. Doppler PSD of fixed wireless is used to model the real environment.

Figure 1.26 shows the delay profile of each tap. The magnitude of the channel coefficients of all the taps, generated and plotted versus time, are shown below in Figure 1.27.

Table 1.2 SUI-3 channel parameters.

SUI-3 Channel				
	Tap 1	Tap 2	Tap 3	Units
Delay	0	0.5	1.0	μs
Power (Omni ant.)	0	-5	-10	dB
K Factor (Omni ant.)	1	0	0	
Power (30° ant.)	0	-11	-22	dB
K Factor (30° ant.)	3	0	0	
Doppler	0.4	0.4	0.4	Hz
Antenna Correlation:	$\rho_{\text{ENV}} = 0.4$			
Gain Reduction Factor:	GRF $= 3$ dB			
Normalization Factor:	$F_{\text{omni}} = -1.5113$ dB, $F_{30} = -0.3573$ dB			

1.5.9 Simulating the UMTS Mobile Radio Channel

The parameters for three different UMTS mobile channel models based on the standards put forward by ETSI, which model an *indoor*, a *pedestrian* and a *vehicular* environment, are given in this chapter. Tables 1.3–1.5 give the relative delays between taps and the average tap power for each scenario [14].

Simulation results evaluating the bit-error rate (BER) performance of QPSK modulation on the indoor, pedestrian and vehicular UMTS channels and BFWA channel are presented in Figure 1.28.

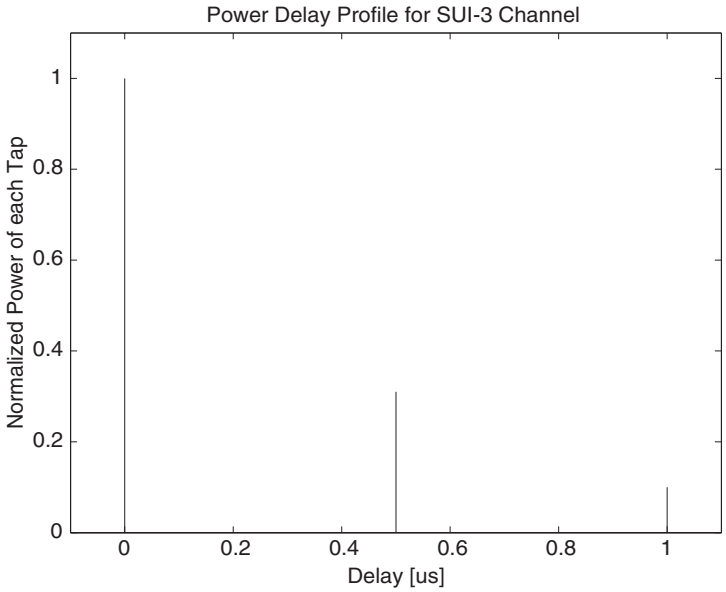

Figure 1.26 Delay profile of SUI-3 channel model.

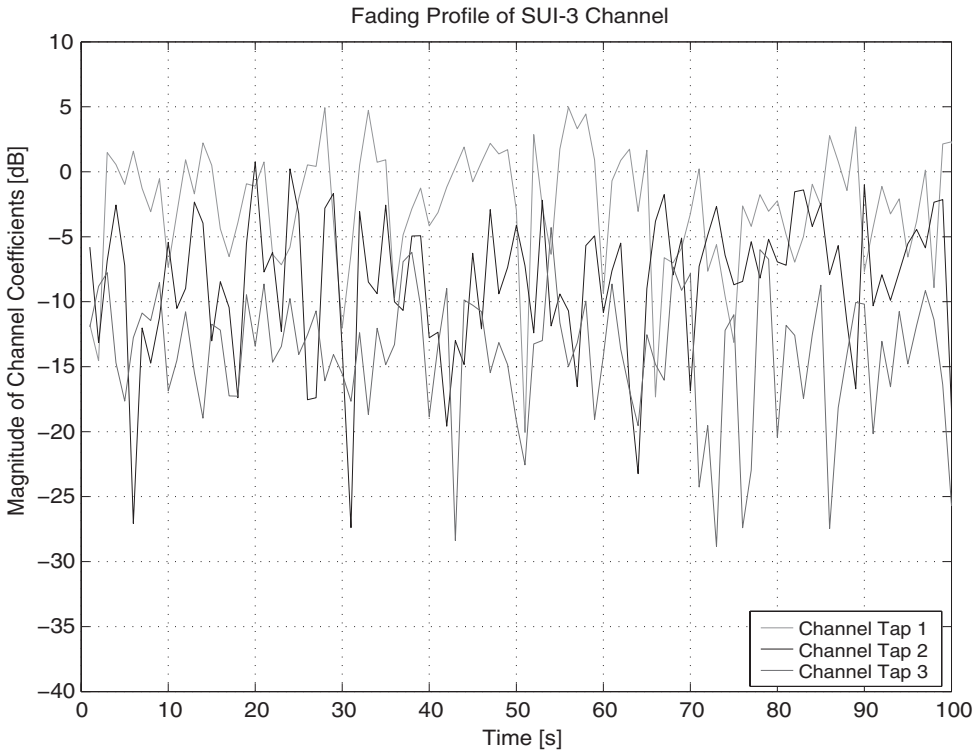

Figure 1.27 Fading profile of SUI-3 channel model.

Table 1.3 Channel parameters for the indoor UMTS channel model.

Tap	Relative delay (ns)	Average power (dB)
1	0	0
2	50	−3.0
3	110	−10.0
4	170	−18.0
5	290	−26.0
6	310	−32.0

Table 1.4 Channel parameters for the pedestrian UMTS channel model.

Tap	Relative delay (ns)	Average power (dB)
1	0	0
2	110	−9.7
3	190	−19.2
4	410	−22.8

Table 1.5 Channel parameters for the vehicular UMTS channel model.

Tap	Relative delay (ns)	Average power (dB)
1	0	0
2	310	−1.0
3	710	−9.0
4	1090	−10.0
5	1730	−15.0
6	2510	−20.0

1.6 (Multiple-Input–Multiple-Output) (MIMO) Channel

The performance analysis of the WiMAX system is extended from a SISO model to the MIMO model. Let us consider a multi-antenna system with N_T transmit and N_R receive antennas characterized by input/output in (1.47); the antenna configuration is shown in Figure 1.29.

$$y_m(t) = \sum_{n=1}^{N_T} \int_{-\infty}^{\infty} h_{mn}(t, \tau)\, x_n(t - \tau)\ \mathrm{d}\tau, \quad m = 1, 2, \ldots, N_R, \tag{1.47}$$

where $x_n(t)$ is the signal transmitted from the nth transmitted antenna, $y_m(t)$ is the signal received at the mth receive antenna and $h_{mn}(t, \tau)$ is the impulse response of the channel between the nth transmitted and mth receive antennas. The above equation

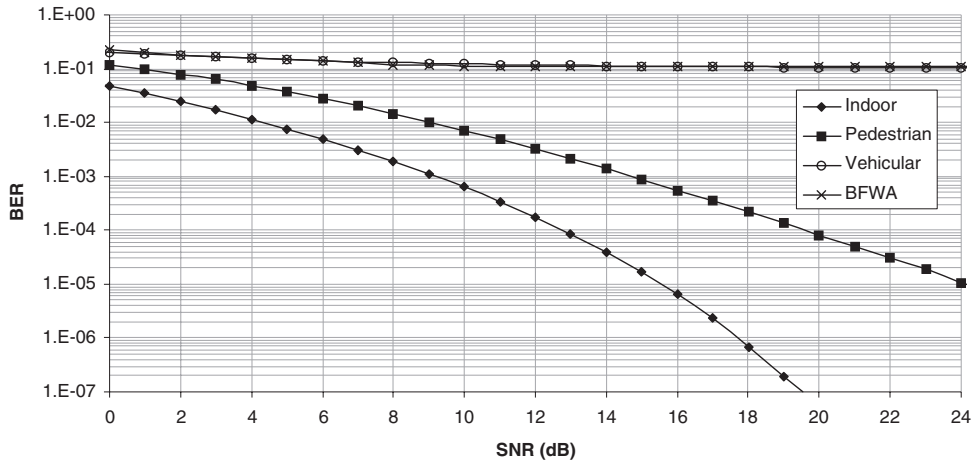

Figure 1.28 BER performance of QPSK modulation on the indoor, pedestrian and vehicular and BFWA channels.

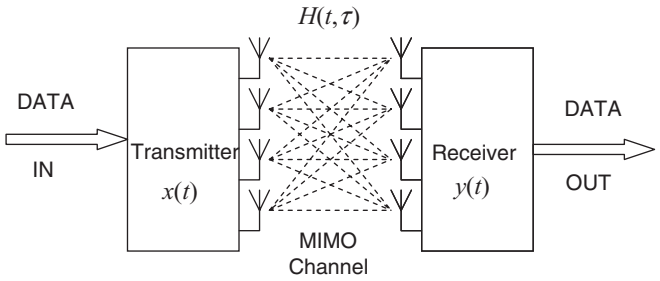

Figure 1.29 General MIMO channel with matrix H.

can be written in matrix form as follows:

$$y(t) = \int_{-\infty}^{\infty} H(t,\tau)\, x(t-\tau)\, \mathrm{d}\tau, \tag{1.48}$$

where $\mathbf{H}(t,\tau)$ is the $N_R \times N_T$ channel matrix and is given as:

$$\mathbf{H}(t,\tau) = \begin{bmatrix} h_{11}(t,\tau) & h_{12}(t,\tau) & h_{13}(t,\tau) & \ldots\ldots\ldots\ldots & h_{1N_T}(t,\tau) \\ h_{21}(t,\tau) & h_{22}(t,\tau) & h_{23}(t,\tau) & \ldots\ldots\ldots\ldots & h_{2N_T}(t,\tau) \\ h_{31}(t,\tau) & h_{32}(t,\tau) & h_{33}(t,\tau) & \ldots\ldots\ldots\ldots & h_{3N_T}(t,\tau) \\ . & . & . & \ldots\ldots\ldots\ldots & . \\ . & . & . & \ldots\ldots\ldots\ldots & . \\ h_{N_R1}(t,\tau) & h_{N_R2}(t,\tau) & h_{N_R3}(t,\tau) & \ldots\ldots\ldots\ldots & h_{N_RN_T}(t,\tau) \end{bmatrix}, \tag{1.49}$$

$\mathbf{y}(t) = \left[y_1(t), y_2(t), \ldots, y_{N_R}(t)\right]$ is the N_R size row vector containing the signals received from N_R antenna and $\boldsymbol{x}(t) = \left[x_1(t), x_2(t), \ldots, x_{N_T}(t)\right]$ is the N_T size row vector containing signals transmitted from N_T antennas. The impulse response in the case of the MIMO channel can be explained using the same criteria as in the SISO case, but the distinguishing feature of MIMO systems is the spatial correlation among the impulse response composed of $\mathbf{H}(t,\tau)$.

The channel matrix **H** defines the input-output relations of the MIMO system and is known as the channel transfer function. Let us consider the two-transmit and two-receive antennae configuration, as shown in Figure 1.30. The system model comprises two transmitting and two receiving antennae and the medium between them is modelled as the MIMO channel.

The antenna correlations among the different paths are calculated and shown in Table 1.6, showing small correlations between the paths of the same channel. Also note that, for example, path 0 of channel A (A0) has correlation factor of 0.4 with path 0 of channels B, C and D (i.e. B0, C0 and D0). However, the same path 0 (A0) has a much lower correlation between other paths of the channels B, C and D.

Similarly, examples of correlation tables for the indoor, pedestrian and vehicular UMTS MIMO channels used in simulations are given in Tables 1.7–1.9.

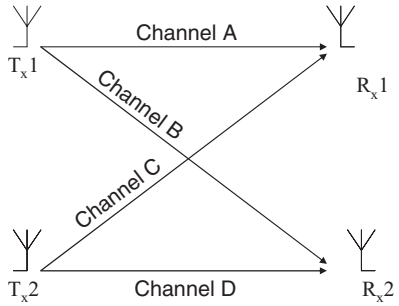

Figure 1.30 2×2 MIMO channel.

Table 1.6 Correlations between four paths, A, B, C and D, with subchannel paths 0, 1, 2 of a 2×2 MIMO system.

	A0	A1	A2	B0	B1	B2	C0	C1	C2	D0	D1	D2
A0	1	0.017	0.008	0.38	0.013	0.011	0.39	0.008	0.005	0.38	0.007	0.004
A1	0.017	1	0.021	0.003	0.4	0.02	0.003	0.39	0.01	0.003	0.403	0.015
A2	0.008	0.021	1	0.009	0.003	0.398	0.012	0.008	0.39	0.01	0.006	0.402
B0	0.38	0.003	0.009	1	0.003	0.008	0.39	0.004	0.006	0.39	0.008	0.002
B1	0.001	0.4	0.003	0.003	1	0.015	0.008	0.39	0.01	0.008	0.39	0.012
B2	0.01	0.02	0.39	0.008	0.015	1	0.016	0.01	0.39	0.01	0.008	0.39
C0	0.39	0.003	0.01	0.39	0.008	0.016	1	0.002	0.008	0.4	0.009	0.0006
C1	0.008	0.39	0.008	0.004	0.39	0.01	0.002	1	0.008	0.007	0.4	0.009
C2	0.005	0.011	0.39	0.006	0.01	0.39	0.008	0.008	1	0.009	0.007	0.401
D0	0.38	0.003	0.01	0.39	0.008	0.01	0.4	0.007	0.009	1	0.008	0.0039
D1	0.007	0.4	0.006	0.008	0.39	0.008	0.009	0.4	0.007	0.008	1	0.012
D2	0.004	0.001	0.4	0.002	0.01	0.39	0.006	0.009	0.4	0.003	0.012	1

Table 1.7 Mean correlation values between each path of the indoor MIMO channel.

	A	B	C	D
A	1	0.006874	0.006467	0.006548
B	0.006874	1	0.005903	0.006568
C	0.006467	0.005903	1	0.006312
D	0.006548	0.006568	0.006312	1

Table 1.8 Mean correlation values between each path of the pedestrian MIMO channel.

	A	B	C	D
A	1	0.049111	0.054736	0.050264
B	0.049111	1	0.056464	0.057746
C	0.054736	0.056464	1	0.062907
D	0.050264	0.057746	0.062907	1

Table 1.9 Mean correlation values between each path of the vehicular MIMO channel.

	A	B	C	D
A	1	0.057106	0.04603	0.052264
B	0.057106	1	0.040664	0.044029
C	0.04603	0.040664	1	0.061777
D	0.052264	0.044029	0.061777	1

In Chapter 7, the UMTS MIMO channels are used to evaluate the performance of space-time ring-TCM codes. The simplest space-time code uses a delay diversity code. Its performance on the indoor, pedestrian and vehicular MIMO channels is given in Figure 1.31.

1.7 Magnetic Storage Channel Modelling

Another application of error-correcting codes can be found in magnetic data storage devices. An error-correcting scheme in this situation must be able to correct long bursts of errors and have an efficient decoding algorithm, minimizing latency, to ensure high data rates. In this section we present a simple channel model for longitudinal magnetic recording, a method of writing data to a magnetic disc that is currently in use. Data is written to the magnetic disc by magnetizing microscopic areas on the disc in one of

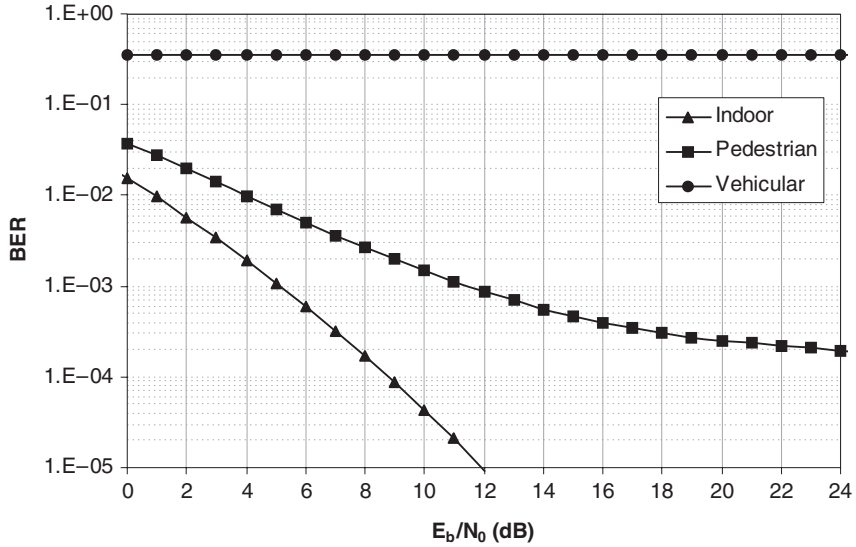

Figure 1.31 Simulation results evaluating the performance of the delay diversity code on the indoor, pedestrian and vehicular 2 × 2 MIMO channels.

two directions, representing either a '1' or a '0'. The data is recovered by passing the *read head* over the disc, which detects changes in the magnetic field corresponding to transitions from '0' to '1' or '1' to '0'. Traditionally, data is written on the plane of the magnetic disc; this is known as *longitudinal recording*. However, we have reached the limits of storage capacity for this particular technique, and now data is written perpendicular to the disc, known as *perpendicular recording*, which allows storage capacity to be further increased. In this chapter, modelling a longitudinal magnetic recording channel is explained, as this is still the most common recording method used for most hard drives.

1.7.1 Longitudinal Recording

A simple linear channel model for a longitudinal magnetic recording system is given in Figure 1.32 [15]. As stated previously, the read head measures the changes in the direction of magnetization, and this can be modelled as a differentiator. The differentiator with transfer function $1 - D$, where D is a memory element, subtracts the previous bit value from the current bit value.

The transition response for longitudinal magnetic recording can be modelled as a *Lorentzian* pulse, given by [15]:

$$h(t) = \frac{1}{1 + \left(\dfrac{2t}{PW50}\right)^2}, \tag{1.50}$$

where $PW50$ is the pulse width of the Lorentzian pulse at half its peak value. Some Lorentzian pulses and the effect of varying $PW50$ are shown in Figure 1.33. It will be shown that increasing $PW50$ increases the level of intersymbol interference (ISI).

A useful measure of ISI is the *recording linear density*, denoted by D_s and given by [15]:

$$D_s = \frac{PW50}{\tau}, \tag{1.51}$$

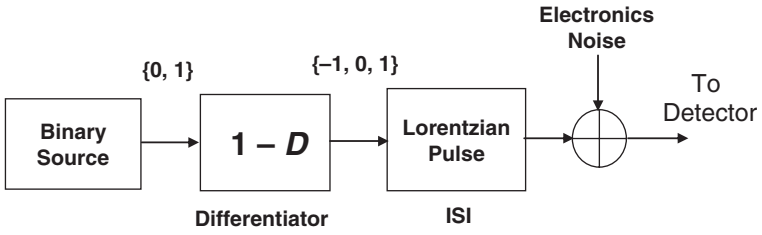

Figure 1.32 Simple channel model for a longitudinal magnetic recording channel.

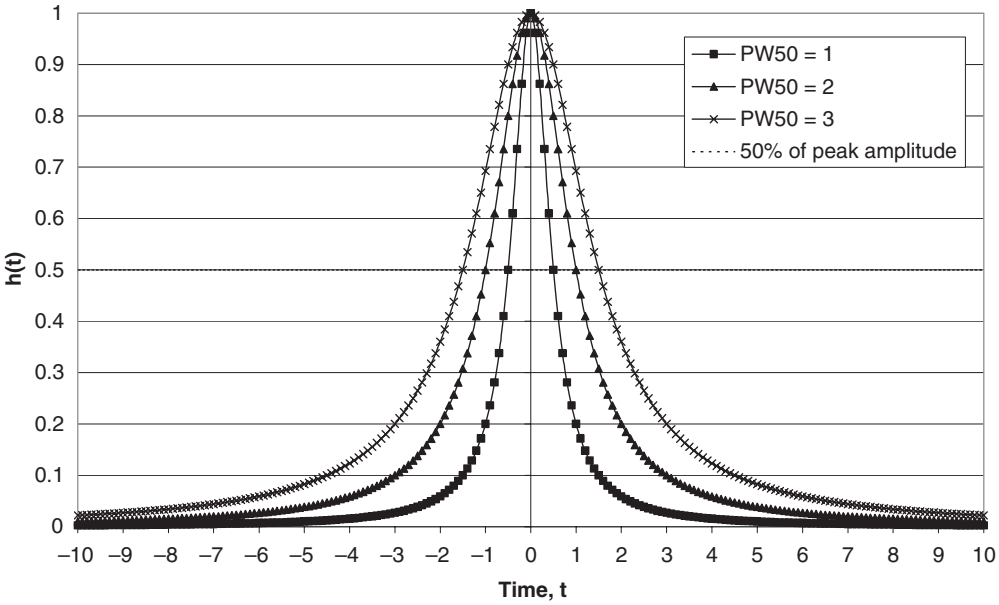

Figure 1.33 Lorentzian pulses with varying values of *PW50*.

where τ is the bit period. It gives the number of bits present within a time interval equal to $PW50$. A higher value of D_s corresponds to a higher level of ISI.

The response of a positive transition from '1' to a '0' is called a *dibit* response, expressed as:

$$d(t) = h(t) - h(t - \tau). \tag{1.52}$$

An example of a dibit response where the Lorentzian pulse has $PW50 = \tau = 1$ is given in Figure 1.34. A positive Lorentzian pulse is initiated at $t = 0$ and is followed by a negative Lorentzian pulse at $t = 1$. The sum of the two pulses gives the dibit response. The peak value of the dibit response is 80% of the peak value of the Lorentzian pulses since $PW50$ is wide enough that both pulses overlap. Increasing $PW50$ reduces the peak value of the dibit response further and this illustrates how ISI occurs in a longitudinal magnetic recording channel.

1.7.2 (Partial Response Maximum Likelihood) (PRML) Detection

Partial response maximum likelihood (PRML) detection is a two-stage process consisting of a transversal FIR filter concatenated with a Viterbi decoder (explained in Chapter 7), as shown in Figure 1.35.

The idea is to design a partial response (PR) equalizer with coefficients that shape the frequency response of the channel output to a predetermined *target response*

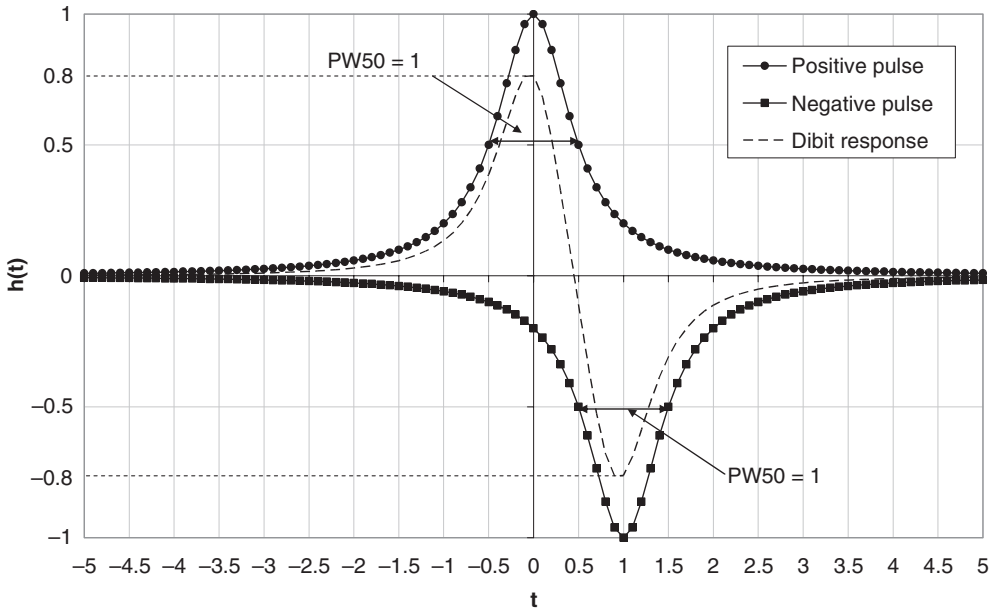

Figure 1.34 A positive Lorentzian pulse at $t = 0$ followed by a negative Lorentzian pulse at $t = 1$, both with $PW50 = 1$. The sum of the two pulses results in the dibit response.

[15, 16]. It is well known that frequency response of the channel output for longitudinal recording with varying values of D_s matches closely with the frequency response of the polynomial of the form:

$$G(D) = (1 - D)(1 + D)^n, \tag{1.53}$$

where D is a delay element and n is a positive integer. For $n = 1$, 2 and 3 the polynomials are known as PR4, EPR4 and E^2PR4 respectively. For example, a longitudinal magnetic recording channel with $D_s = 2$ has a response that matches closely with the response of $G(D) = (1 - D)(1 + D) = 1 - D^2$, known as PR4. Given a binary input

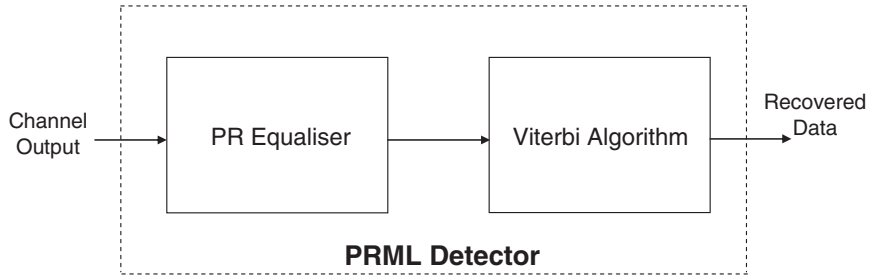

Figure 1.35 Block diagram showing the two processes in PRML detection.

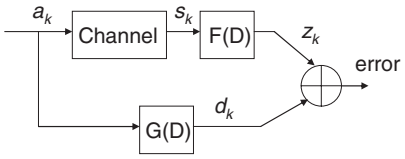

Figure 1.36 PR equalizer design criterion.

a_k, the output z_k of the PR equalizer should closely match the output d_k of the target polynomial $G(D)$, as shown in Figure 1.36 [16].

1.7.2.1 PR Equalizer Design

The PR equalizer polynomial $F(D)$ and target polynomial $G(D)$ are expressed as vectors **F** and **G** respectively [16]:

$$\mathbf{F} = \{-f_K \; -f_{K+1} \cdots f_0 \cdots f_{K-1} \; f_K\}, \tag{1.54}$$

where K is a positive integer and $\pm f_i$, $i = 0, 1, \ldots, K$, are the coefficients of $F(D)$.

$$\mathbf{G} = \{g_0 \; g_1 \cdots g_{L-1}\}, \tag{1.55}$$

where L is a positive integer and g_i, $i = 0, 1, \ldots, L-1$, are the coefficients of $G(D)$. Two further matrices **R** and **T** are defined as:

$$\mathbf{R} = \{R_{i,j}\} = E\{s_{k-i}s_{k-j}\}, \quad -K \leq i, j \leq K, \tag{1.56}$$

where $E\{\}$ is the expectation operator. **R** therefore contains autocorrelation values of the channel output:

$$\mathbf{T} = \{T_{i,j}\} = E\{s_{k-i}a_{k-j}\}, \quad -K \leq i \leq K, 0 \leq j \leq L-1, \tag{1.57}$$

T contains the cross correlation values of the channel output with the binary input. The coefficients of **F** can then be determined by [16]:

$$\mathbf{F} = \mathbf{R}^{-1}\mathbf{T}\mathbf{G}. \tag{1.58}$$

After equalization the output should be similar to the desired output from $G(D)$. A diagram of the PR4 polynomial $G(D) = 1 - D^2$ is given in Figure 1.37.

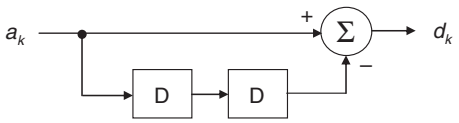

Figure 1.37 The PR4 target.

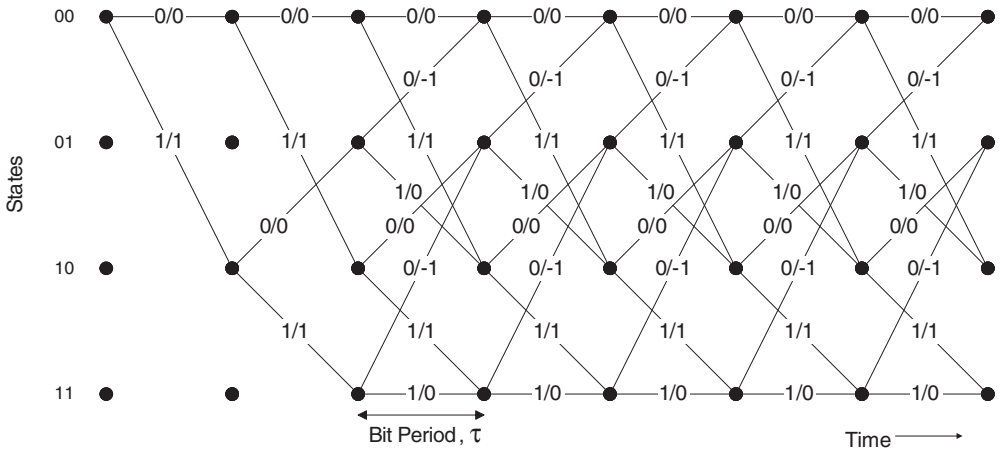

Figure 1.38 Trellis diagram of the PR4 target polyomial.

Since it has two memory elements, this can be expressed as a trellis, as shown in Figure 1.38, with four states and one input and one output for each branch.

The final part of PRML detection involves finding the maximum likelihood path through the trellis corresponding to the original binary input a_k. This can be achieved by using the soft-decision Viterbi algorithm.

1.8 Conclusions

In this chapter, an introduction to information theory and channel modelling was presented. In particular, the capacities of various channel models, such as the AWGN channel and SISO and MIMO fading channels modelling a fixed wireless access channel and UMTS cellular channels were examined. Simulation results were also presented, showing the performance of uncoded schemes on the FWA and UMTS channels for both MIMO and SISO situations. Finally, descriptions of different channel models were presented, allowing the performance for many coding schemes in wireless and magnetic storage channels to be evaluated by computer simulation.

References

[1] Shannon, C.E. (1948) A mathematical theory of communication. *Bell Systems Technical Journal*, **27**, 379–423, 623–56.
[2] Shannon, C.E. (1998) Communications in the presence of noise. *Proceedings of IEEE*, **86**, 447–58.
[3] McEliece, R.J. (1977) *The Theory of Information and Coding*, Addison-Wesley, Massachusetts.
[4] Abramson, N. (1962) *Information Theory and Coding*, McGraw-Hill, New York.
[5] Proakis, J.G. (1989) *Digital Communications*, 2nd edn, McGraw-Hill, New York.
[6] Parsons, J.D. (2000) *The Mobile Radio Propagation Channel*, 2nd edn, John Wiley & Sons Inc.

[7] Clarke, R.H. (1968) A statistical theory of mobile-radio reception. *Bell Systems Technical Journal*, **47**, 957–1000.
[8] Jakes, W.C. (1994) *Microwave Mobile Communications*, John Wiley & Sons, Inc., New York.
[9] Hata, M. (1980) Empirical formula for propagation loss in land mobile radio services. *IEEE Transactions on Vehicular Technology*, **VT-29** (3), 317–25.
[10] IEEE 802.16 (2001) *IEEE Standard for Local and Metropolitan Area Networks: part 1b, Air Interface for Fixed Broadband Wireless Access Systems*, April 8, 2002.
[11] Aulin, T. (1979) A modified model for the fading signal at a mobile radio channel. *IEEE Transactions on Vehicular Technology*, **VT-28** (3), 182–203.
[12] Greenstein, L.J., Erceg, V., Yeh, Y.S. and Clark, M.V. (1997) A new path-gain/delay-spread propagation model for digital cellular channels. *IEEE Transactions on Vehicular Technology*, **46** (2), 477–85.
[13] IEEE 802.11 (1994) *Wireless Access Method and Physical Layer Specification*, New York.
[14] Erceg, V. (1999) An empirically based path loss model for wireless channels in suburban environments. *IEEE Journal in Selected Areas in Communications*, **17** (7), 1205–11.
[15] Vasic, B. and Kurtas, E.M. (eds) (2005) *Coding and Signal Processing for Magnetic Recording Systems*, CRC Press.
[16] Moon, J. and Zeng, W. (1995) Equalization for maximum likelihood detectors. *IEEE Trans. Mag.*, **31** (2), 1083–8.