

Signal Processing and Transmission

Part 1 of Lecture Notes

© Prof. Dr. Heinz Mathis, Fall 2022

Contents

1 Complex Baseband Representation	6
1.1 The Direct-Conversion Receiver	6
1.2 Analytic Signal	6
1.3 Example: Modulated Carrier Signal	11
1.4 I/Q Modulator/Demodulator	12
1.5 Cascaded-Integrator Comb (CIC) Filters	13
1.5.1 Comb Filters	13
1.5.2 Cascaded Integrator-Comb (CIC) Filters	15
1.5.3 Aliasing and Passband Roll-Off of Downsampling CIC Filters	16
2 Detection Theory	18
2.1 Framework	18
2.2 Maximum-Likelihood (ML) Detection	22
2.3 Maximum A-Posteriori (MAP) Detection	23
2.4 Bayes Criterion / Bayes Test	24
2.5 Minimax Criterion	25
2.6 Neyman-Pearson Test	26
2.7 Example: On-Off Keying in an AWGN Channel	26
2.8 Summary	27
3 Propagation	29
3.1 The Wireless Channel	29
3.2 Mechanisms of Wireless Propagation	30
3.2.1 Free-Space Propagation	32
3.2.2 Open-Field Propagation	33
3.2.3 Diffraction	35
3.2.4 Fresnel Zones	38
3.2.5 Scattering	40
3.2.6 Further Losses	41
3.3 Empirical Path-Loss Models	42
3.4 Link Budget	43
3.5 Fading Channels	44
3.5.1 Shadowing	44
3.5.2 Fast Fading	45

3.5.3	Mobile Environment	50
3.5.4	Relationship of Fading Parameters	53
3.5.5	Ways Out	54
3.5.6	Multipath for Low Bandwidth	55
4	Channel Equalization Using Adaptive Filters	56
4.1	Matched Filter Concept	56
4.2	Applications of Adaptive Filters	57
4.3	LMS-Based Adaptive Filters	58
4.3.1	Wiener Filtering	58
4.3.2	The LMS Algorithm	61
4.4	RLS-Based Adaptive Filters	62
4.4.1	The LS Algorithm	62
4.4.2	The RLS Algorithm	63
4.5	Kalman Filtering	64
4.6	Channel Equalization	67
4.6.1	Zero-Forcing Equalizers	70
4.6.2	MMSE Equalizer	71
4.6.3	Decision-Feedback Equalizer	72
4.6.4	Fractionally-Spaced Equalizers	74
4.6.5	Training of an Equalizer	75
4.6.6	Blind Equalizers	76
4.6.7	The Constant-Modulus Algorithm (CMA)	76
4.6.8	An Example	77
5	Diversity	79
5.1	Combining Methods	80
5.1.1	Selection Diversity	80
5.1.2	Switching Diversity	81
5.1.3	Equal-Gain Combining Diversity	82
5.1.4	Maximum-Ratio Combining Diversity	83
5.2	Performance Comparison of Different Methods	86
6	CDMA	88
6.1	Other Channel-Access Systems	88
6.1.1	FDMA	88
6.1.2	TDMA	88
6.1.3	Capacity-Achieving xDMA	89
6.2	Introduction to CDMA	92
6.3	Benefits and Drawbacks	93
6.4	History	93
6.5	Spread-Spectrum Operation	94
6.6	Synchronization	94
6.6.1	Acquisition	96
6.6.2	Tracking	96
6.7	Spreading Codes	97
6.7.1	M-Sequences	99
6.7.2	Gold Codes	100
6.7.3	Walsh Codes	100

6.8	The Near-Far Problem	101
6.9	Examples of CDMA Systems	101
6.9.1	IS-95	102
6.9.2	UMTS	102
6.10	Other CDMA Systems	106
7	Ultrawideband (UWB) Technology	107
7.1	Definition	107
7.2	History	107
7.3	Standards	108
7.4	Applications	108
8	OFDM	109
8.1	Introduction	109
8.2	Mathematical Derivation	109
8.3	Signal-Processing Chain	111
8.3.1	Frame Synchronization	112
8.3.2	Sampling-Time Synchronization	113
8.3.3	Guard Intervals: Zero-Padding and Cyclic Prefix	114
8.3.4	Channel Estimation and Equalization	116
8.3.5	Optimal Energy Allocation (Waterfilling)	116
8.3.6	PAPR Considerations	118
8.4	Applications	119
8.5	Summary	119
9	Channel Coding	121
9.1	Interleaving	122
9.2	Convolutional Codes	123
9.2.1	Construction of Convolutional Codes	123
9.2.2	Decoding of Convolutional Codes	124
9.3	Simulation of Transmission Systems	125
10	Trellis-Coded Modulation (TCM)	126
10.1	History and Motivation	126
10.2	Capacity Computation of Discrete-Alphabet Sources	126
10.2.1	Relationship to Mutual Information and Entropy	126
10.2.2	Continuous Variables	127
10.2.3	Discrete-Input, Discrete-Output Variables	129
10.2.4	Discrete-Input, Continuous-Output Variables	131
10.2.5	Numerical Evaluation	133
10.3	Principles of TCM	134
10.4	Definitions of Parameters	134
10.4.1	Minimum Distance	134
10.4.2	Free Distance	135
10.4.3	Coding Gain	135
10.5	A Simple Introductory Example	135
10.6	Improving the Performance (the Next Two Examples)	138
10.7	Encoding and Signal Mapping	140
10.8	Set Partitioning	140

10.9	Performance	142
10.10	Philosophy of TCM	144
10.11	Extensions of TCM	144
10.11.1	Rotation-Invariant TCM	144
10.11.2	Multidimensional TCM	144
10.12	Real-World Example: Combination of OFDM and TCM	144
10.12.1	Modem Architecture	145
10.12.2	Symbols and Frames	145
10.12.3	Frame Synchronization	146
10.12.4	Channel Estimation and Equalization	146
10.12.5	Waterfilling	146
10.12.6	PAPR Considerations	147
10.12.7	Trellis Coded Modulation over OFDM	147
10.12.8	Conclusions	149
11	Turbo Coding	150
11.1	History and Motivation	150
11.2	Approach	150
11.3	Example	150
Appendices		152
A	Equations and Formulae	152
A.1	Linear Algebra	152
A.2	Analysis	154
A.2.1	Complex Analysis	154
A.3	Statistics	155
A.3.1	Probability Theory	155
A.4	Digital Signals and Systems	155
A.4.1	Convolution	155
A.4.2	Wiener Filtering	156
B	Frequency Usage	157
B.1	General Frequency Bands	157
B.2	RADAR Bands	158
C	Important Tables	159
C.1	The Q-Function	159
C.2	M-Sequences	160
D	Proof on the Capacity Maximization	161

Preface

These lecture notes are covering the first part of the MSE module *Signal Processing and Transmission*. Every book on communication systems has a so-called *Figure 1*. It will always describe the chain from the information source to the information sink.

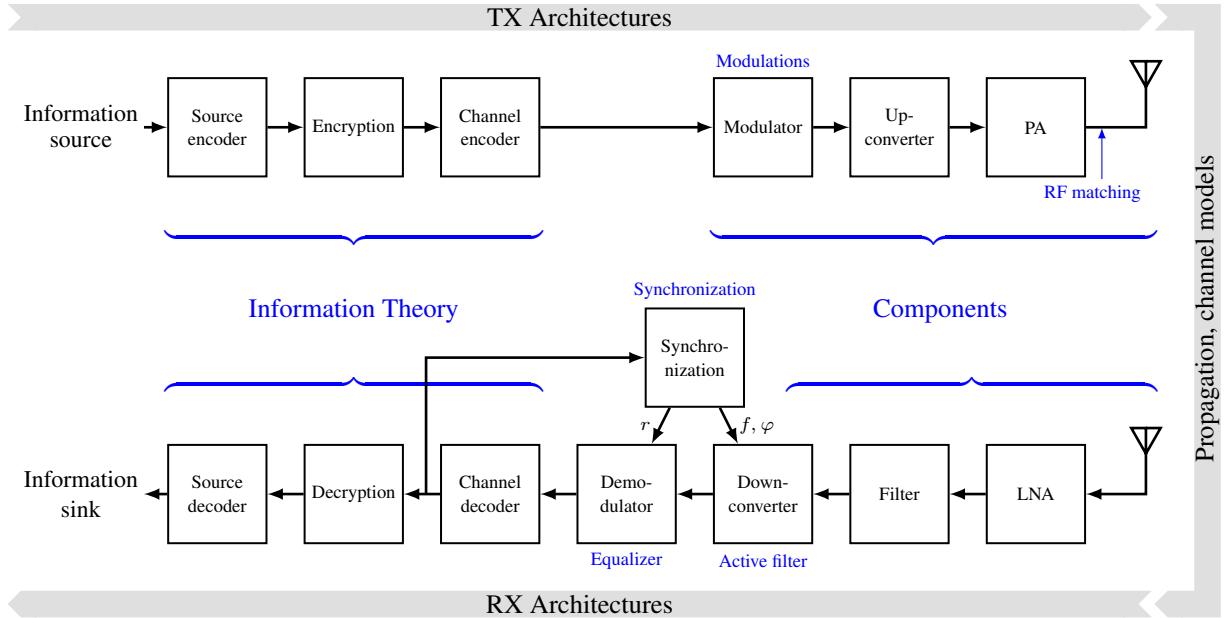


Figure 1 Disciplines involved in the design of a wireless communication system.

During this MSE course we will talk about the following topics: We start with the complex baseband representation to gain some insight into how a modulator or a demodulator is built. In particular, on the receive side, we will learn about signal processing prior to the demodulator, e.g., A-to-D conversion and filtering combined with sample-rate reduction (CIC filter). Once, the signal has been pre-conditioned, optimal detectors are sought after. We will then turn back to pre-conditioning the signal using adaptive filters, which are of great benefit to the channel equalization process.

Acknowledgment

Some parts in the first half have been inherited and adapted from Markus Hufschmid's previous lecture notes.

I would also like to thank my assistants, Marcel Kluser, Daniel Megnet, Nicola Ramagnano, Hans-Dieter Lang, and Patrick Fleischmann for preparing substantial parts of both text and figures within these notes.

1 Complex Baseband Representation

1.1 The Direct-Conversion Receiver

Independent-sideband (ISB) signals as used in analogue and digital wireless systems as well as most modern digital modulation formats in the form of QAM (quadrature-amplitude modulation) produce different upper and lower sidebands (band above and below RF carrier, respectively). If such a spectrum is mixed down to baseband, which means that the carrier is now at DC, negative and positive frequencies are no longer in symmetry. From Fourier theory we know that in order to represent such an asymmetric spectrum, we need a complex signal. Such a complex signal is generated by a direct-conversion receiver as given in Fig. 1.1.

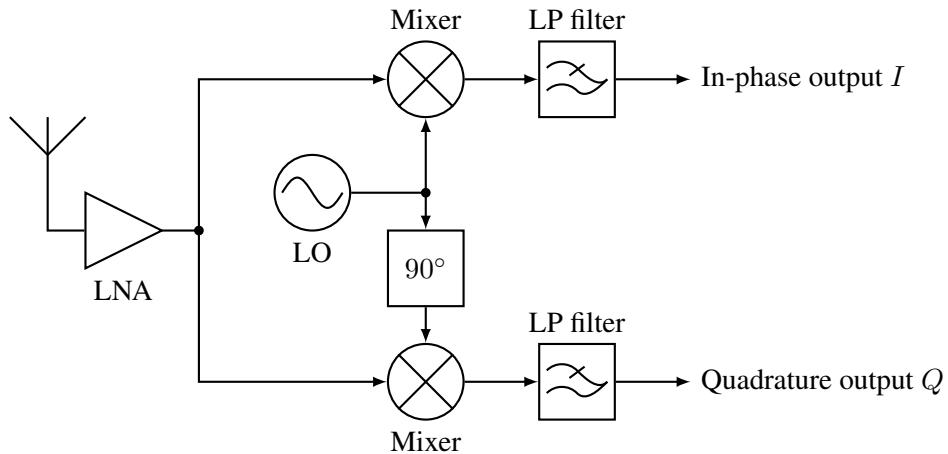


Figure 1.1 Direct-conversion receiver.

The generation of the complex-baseband signal can also be envisaged by multiplying the real-valued RF signal coming through the antenna by an exponential waveform

$$\exp(-j \cdot 2\pi \cdot f_{\text{LO}} \cdot t) = \cos(2\pi \cdot f_{\text{LO}} \cdot t) - j \cdot \sin(2\pi \cdot f_{\text{LO}} \cdot t) \quad (1.1)$$

rather than a real-valued sinusoidal (as for example in superhet receivers). Such a multiplication therefore generates a complex output, consisting of a real and an imaginary part. In communication systems these parts are called the in-phase (real) and the quadrature (imag) component of the signal.

1.2 Analytic Signal

The Fourier transform $X(f)$ of a real signal $x(t)$ is complex with the well-known symmetry property

$$X(-f) = X^*(f). \quad (1.2)$$

In other words, the real part and the imaginary part are even and odd functions, respectively

$$\text{Re}(X(-f)) = \text{Re}(X(f)), \quad (1.3a)$$

$$\text{Im}(X(-f)) = -\text{Im}(X(f)). \quad (1.3b)$$

Thus, there is redundant information in the negative-frequency part of the signal spectrum of a real-valued signal, such that a real-valued signal can be completely recovered from the positive-frequency part.

The analytic signal of a general complex-valued signal $x(t)$ can be generated by masking out the negative-frequency, yielding

$$X_a(f) \stackrel{\Delta}{=} 2 \cdot \varepsilon(f) \cdot X(f) = \begin{cases} 2X(f) & f > 0 \\ X(f) & f = 0 \\ 0 & f < 0 \end{cases} \quad (1.4)$$

$\varepsilon(f)$ is called the Heaviside step function. If the original signal $x(t)$ is real-valued, we can recover $X(f)$ from $X_a(f)$ using Eqs. (1.3a) and (1.3b). The inverse Fourier transform \mathcal{F}^{-1} of Eq. (1.4) is called the analytic signal $x_a(t)$ or the pre-envelope of the signal $x(t)$. The analytic signal lacks the symmetry properties in the Fourier transform (no negative frequencies) and must therefore be complex. We can transform the masking effect in the frequency domain into a convolution in the time domain

$$\begin{aligned} x_a(t) &= \mathcal{F}^{-1}(X_a(f)) \\ &= \mathcal{F}^{-1}(2 \cdot X(f) \cdot \varepsilon(f)) \\ &= 2\mathcal{F}^{-1}(X(f)) * \mathcal{F}^{-1}(\varepsilon(f)) \\ &= 2x(t) * \mathcal{F}^{-1}(\varepsilon(f)). \end{aligned} \quad (1.5)$$

The inverse Fourier transform of the Heaviside step function $\varepsilon(f)$ is given by

$$\mathcal{F}^{-1}(\varepsilon(f)) = \frac{1}{2} \left(\delta(t) + j \cdot \frac{1}{\pi t} \right). \quad (1.6)$$

Thus, we can develop Eq. (1.5) further into

$$\begin{aligned} x_a(t) &= x(t) * \left(\delta(t) + j \cdot \frac{1}{\pi t} \right) \\ &= x(t) + j \cdot \left(x(t) * \frac{1}{\pi t} \right) \\ &= x(t) + j \cdot \mathcal{H}(x(t)) \\ &= x(t) + j \cdot \hat{x}(t), \end{aligned} \quad (1.7)$$

where $\hat{x}(t) \stackrel{\Delta}{=} \mathcal{H}(x(t))$ denotes the Hilbert transform of the signal $x(t)$ given as the convolution with the impulse response

$$h(t) = \frac{1}{\pi t}. \quad (1.8)$$

The time function $h(t)$ can be interpreted as the impulse response of a so-called Hilbert filter. It has a singularity at zero and is slowly decaying (with $1/t$) tails to both sides, see also Fig. 1.2. The Fourier transform of $h(t)$ yields

$$H(f) = -j \cdot \text{sign}(f) = \begin{cases} -j & f > 0 \\ 0 & f = 0 \\ j & f < 0 \end{cases}, \quad (1.9)$$

which is the frequency response of the Hilbert filter. Thus, in order to get the Hilbert transform of a signal, it may be convolved with $h(t)$ in the time domain, or multiplied by $H(f)$ in the frequency domain. Having an infinite impulse response, real Hilbert filters can only be approximated. Furthermore, an additional delay T_0 needs to be inserted to make it causal.

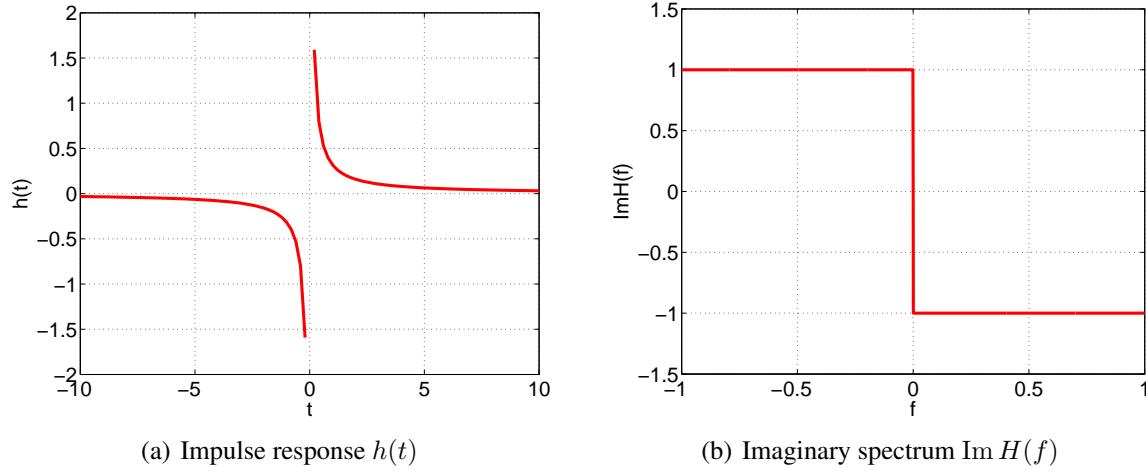


Figure 1.2 Impulse response of the Hilbert filter and its imaginary spectrum.

The Fourier transform of the analytic signal $x_a(t)$ is given by

$$\begin{aligned} X_a(f) &= \mathcal{F}(x_a(t)) = \mathcal{F}(x(t) + j \cdot \hat{x}(t)) \\ &= X(f) + j \cdot \hat{X}(f) \\ &= X(f) + j \cdot H(f) \cdot X(f) \\ &= X(f) + j \cdot (-j \text{sign}(f)) \cdot X(f) \\ &= X(f) \cdot (1 + \text{sign}(f)) = \begin{cases} 2X(f) & f > 0 \\ X(f) & f = 0 \\ 0 & f < 0 \end{cases}, \end{aligned} \quad (1.10)$$

as we have previously postulated, see also Fig. 1.3. If $x(t)$ is a real signal, its Fourier transform shows the above-mentioned symmetry properties and thus we can compute $X(f)$ and also $x(t)$. The analytic signal may be obtained using a circuit according to Fig. 1.4.

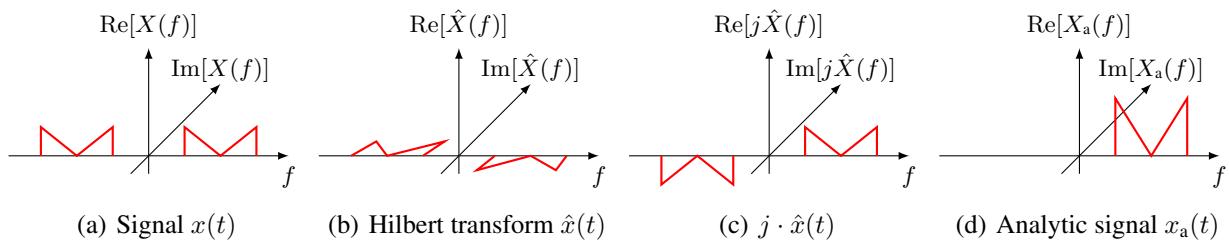


Figure 1.3 Spectra of different views of a signal.

Sometimes it is more practical to represent the analytic signal of Eq. (1.7) in its polar form

$$x_a(t) = A(t) \cdot e^{j\varphi(t)}. \quad (1.11)$$

The function $A(t)$ is called the amplitude envelope, the natural envelope, or simply the envelope of the signal $x(t)$. The argument $\varphi(t)$ of the complex signal $x_a(t)$ is called the instantaneous phase of the signal $x(t)$ and its time derivative is called the instantaneous frequency of the signal $x(t)$. An example is given in Fig. 1.5.

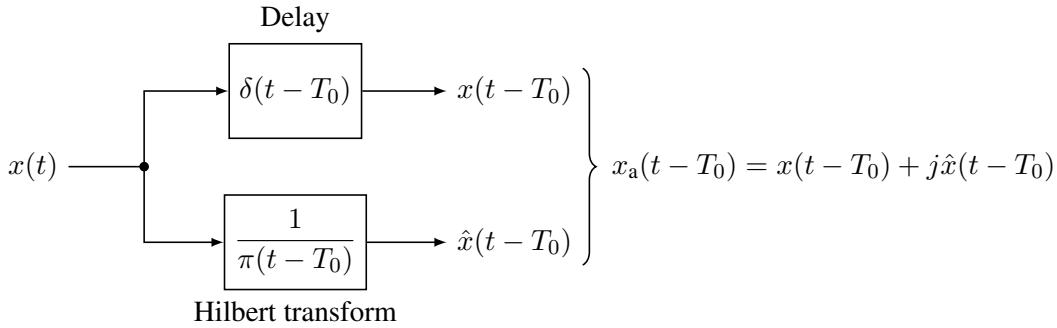


Figure 1.4 Block diagram for generating an analytic signal.

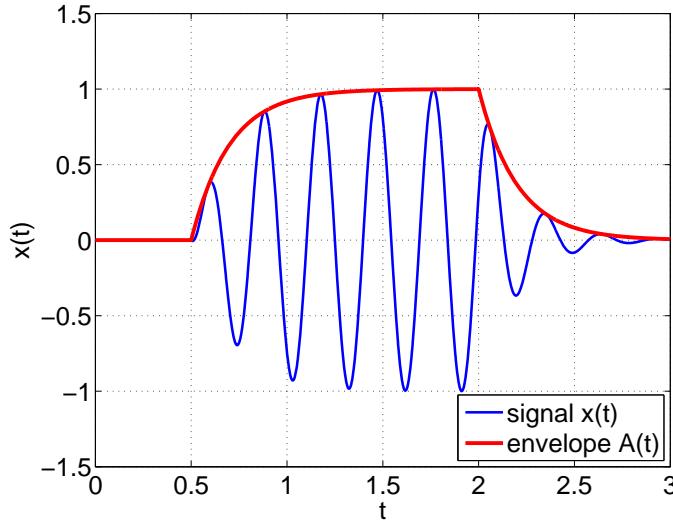


Figure 1.5 Signal and its envelope.

Now, consider a band-pass signal $x_{\text{BP}}(t)$ whose spectrum $X_{\text{BP}}(f)$ is limited to $f_L \leq |f| \leq f_H$. The spectrum is depicted in Fig. 1.6(a). The corresponding analytic signal $x_a(t)$ is complex and contains only non-negative frequencies. Its spectrum is given in Fig. 1.6(b). If we shift the spectrum of this analytic signal to the left by an arbitrary value f_0 , we obtain a complex signal whose spectrum vanishes outside the interval $f_L - f_0 \leq f \leq f_H - f_0$. It is always possible to choose f_0 in such a way that we obtain a (complex) low-pass signal $x_{\text{LP}}(t)$.

This complex low-pass signal, whose spectrum is shown in Fig. 1.6(c), contains exactly the same information as the band-pass signal we started with and is called the complex envelope of $x_{\text{BP}}(t)$. Note that the complex envelope is not unique but rather depends on the choice of f_0 . The following properties hold:

- $x_{\text{LP}}(t)$ depends on the choice of f_0 .
- The spectrum $X_{\text{LP}}(f)$ of $x_{\text{LP}}(t)$ does not exhibit Hermitian symmetry. Therefore, the signal $x_{\text{LP}}(t)$ is complex valued.
- The real part $x_I(t) \triangleq \text{Re}(x_{\text{LP}}(t))$ of the complex envelope is usually called the in-phase component, the imaginary part $x_Q(t) \triangleq \text{Im}(x_{\text{LP}}(t))$ is called the quadrature component.
- If $x(t)$ is a band-pass signal limited to $f_0 - B/2 \leq |f| \leq f_0 + B/2$ then $x_{\text{LP}}(t)$ is a complex low-pass signal limited to $|f| \leq B/2$.

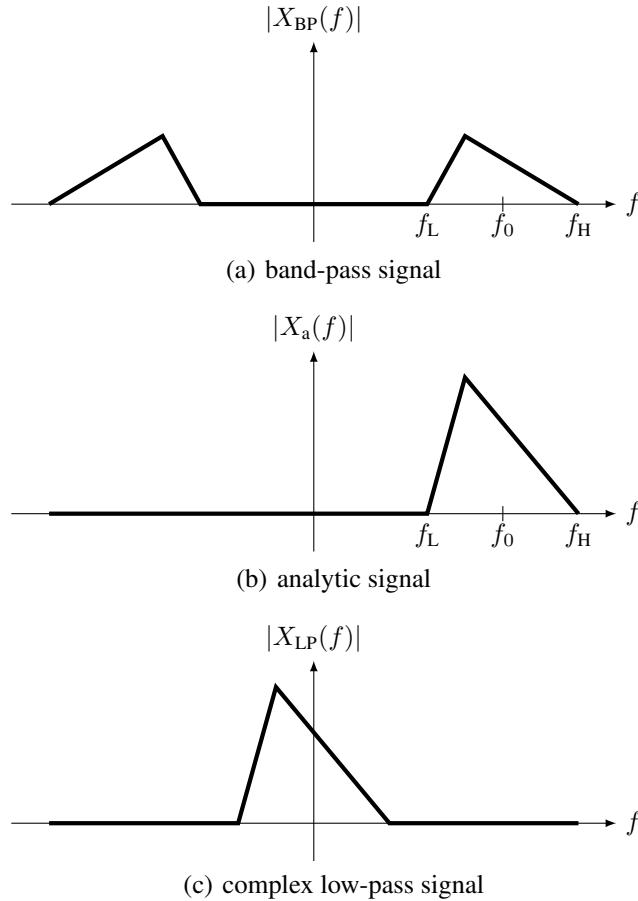


Figure 1.6 Spectra of the band-pass signal and the corresponding analytic and complex low-pass signals.

The left-shift operation in the frequency domain is equal to a multiplication with $e^{-j2\pi f_0 t}$ in the time domain. Thus,

$$x_{LP}(t) = x_a(t) \cdot e^{-j2\pi f_0 t}. \quad (1.12)$$

The analytic signal associated with $x_{BP}(t)$ is given through Eq. (1.7) as

$$x_a(t) = x_{BP}(t) + j \cdot \mathcal{H}(x_{BP}(t)). \quad (1.13)$$

Using Eq. (1.13) in Eq. (1.12) we get

$$x_I(t) = \text{Re}(x_{LP}(t)) = x_{BP}(t) \cdot \cos(2\pi f_0 t) + \mathcal{H}(x_{BP}(t)) \cdot \sin(2\pi f_0 t), \quad (1.14a)$$

$$x_Q(t) = \text{Im}(x_{LP}(t)) = \mathcal{H}(x_{BP}(t)) \cdot \cos(2\pi f_0 t) - x_{BP}(t) \cdot \sin(2\pi f_0 t), \quad (1.14b)$$

where we see that we can evaluate $x_I(t)$ and $x_Q(t)$ if we know $x_{BP}(t)$ and f_0 . The construction of $x_I(t)$ and $x_Q(t)$ can also be seen in Fig. 1.7. Conversely, $x_{BP}(t)$ can easily be computed from $x_I(t)$ and $x_Q(t)$ through

$$x_{BP}(t) = \text{Re} \left(x_{LP}(t) \cdot e^{j2\pi f_0 t} \right) = x_I(t) \cdot \cos(2\pi f_0 t) - x_Q(t) \cdot \sin(2\pi f_0 t). \quad (1.15)$$

We reason that the real band-pass signal $x_{BP}(t)$ and the corresponding complex low-pass signal $x_{LP}(t) = x_I(t) + jx_Q(t)$ are indeed two equivalent descriptions of the same signal. While the band-pass signal usually contains very high frequencies, the complex envelope is a low-pass signal that does not change nearly as fast. Especially during simulations, this may be a huge advantage.

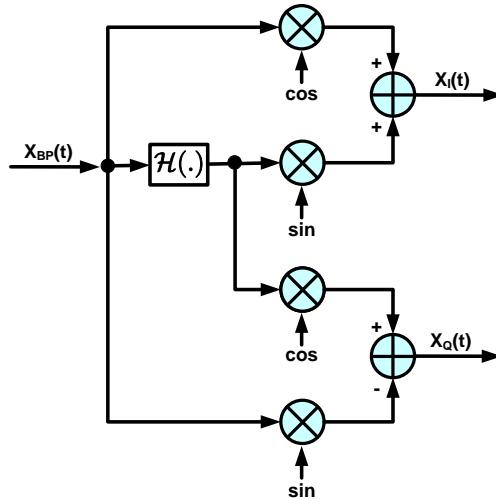


Figure 1.7 Constructing $x_I(t)$ and $x_Q(t)$.

1.3 Example: Modulated Carrier Signal

Assume a band-pass signal

$$x_{BP}(t) = A(t) \cdot \cos(2\pi f_0 t + \phi(t)), \quad (1.16)$$

where either (or both) of $A(t)$ and $\phi(t)$ might carry information. The Hilbert transform of this signal is given by

$$\mathcal{H}(x_{BP}(t)) = A(t) \cdot \sin(2\pi f_0 t + \phi(t)). \quad (1.17)$$

With this, we obtain (after some calculations) the in-phase and quadrature components of the low-pass signal

$$x_I(t) = A(t) \cdot \cos(\phi(t)). \quad (1.18a)$$

$$x_Q(t) = A(t) \cdot \sin(\phi(t)). \quad (1.18b)$$

Once we have in-phase and quadrature components, we can use them to express the complex envelope of the original bandpass signal, the low-pass signal

$$\begin{aligned} x_{LP}(t) &= x_I(t) + jx_Q(t) \\ &= A(t) \cdot (\cos(\phi(t)) + j \sin(\phi(t))) \\ &= A(t) \cdot e^{j\phi(t)}. \end{aligned} \quad (1.19)$$

or the band-pass signal by using Eq. (1.15). Alternatively, we may use the addition theorems for trigonometric functions to decompose $x_{LP}(t)$, yielding

$$\begin{aligned} x_{BP}(t) &= A(t) \cdot \cos(2\pi f_0 t + \phi(t)) \\ &= A(t) \cdot (\cos(2\pi f_0 t) \cdot \cos(\phi(t)) - \sin(2\pi f_0 t) \cdot \sin(\phi(t))) \\ &= \underbrace{A(t) \cdot (\cos(\phi(t)))}_{x_I(t)} \cdot \cos(2\pi f_0 t) - \underbrace{A(t) \cdot (\sin(\phi(t)))}_{x_Q(t)} \cdot \sin(2\pi f_0 t). \end{aligned} \quad (1.20)$$

1.4 I/Q Modulator/Demodulator

Recall that the band-pass signal $x_{\text{BP}}(t)$ can be written in terms of its in-phase and quadrature components, as given by Eq. (1.15). This fact can be exploited to build a modulator. We can generate $x_I(t)$ and $x_Q(t)$ and multiply these signals with $\cos(2\pi f_0 t)$ and $\sin(2\pi f_0 t)$, respectively.

The advantages of this procedure are:

- $x_I(t)$ and $x_Q(t)$ are low-pass signals and can therefore easily be generated employing a digital signal processor (DSP) or an ASIC.
- The modulation process does not depend on the final choice of the carrier frequency f_0 . A change of carrier frequency has no influence on the generation of the complex envelope.

Most of the modern systems employ this sort of modulation technique. A block diagram of an I/Q modulator is given in Fig. 1.8.

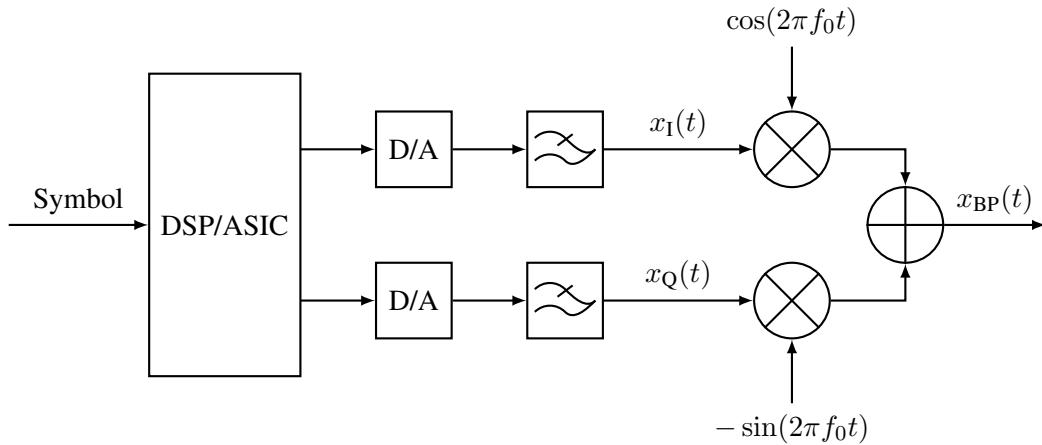


Figure 1.8 I/Q Modulator.

The inverse process is straightforward: By multiplying the last line of Eq. (1.20) with $2 \cos(2\pi f_0 t)$ and $-2 \sin(2\pi f_0 t)$ we get

$$\begin{aligned} A(t) \cdot (\cos(\phi(t)) \cdot 2 \cos^2(2\pi f_0 t) - \sin(\phi(t)) \cdot \sin(2\pi f_0 t) \cdot 2 \cos(2\pi f_0 t)) \\ = A(t) \cdot (\cos(\phi(t)) \cdot (1 + \cos(2\pi \cdot 2f_0 t)) - \sin(\phi(t)) \cdot \sin(2\pi \cdot 2f_0 t)) \end{aligned} \quad (1.21a)$$

and

$$\begin{aligned} A(t) \cdot (\cos(\phi(t)) \cdot \cos(2\pi f_0 t) \cdot (-2) \sin(2\pi f_0 t) - \sin(\phi(t)) \cdot (-2) \sin^2(2\pi f_0 t)) \\ = A(t) \cdot (-\cos(\phi(t)) \cdot \sin(2\pi \cdot 2f_0 t) + \sin(\phi(t)) \cdot (1 - \cos(2\pi \cdot 2f_0 t))), \end{aligned} \quad (1.21b)$$

respectively. Low-pass filters will cut off any components around the carrier f_0 and higher. Thus, the remaining parts of Eqs. (1.21a) and (1.21b) are indeed the low-pass components $x_I(t)$ and $x_Q(t)$. A block diagram of such an I/Q demodulator is shown in Fig. 1.9.

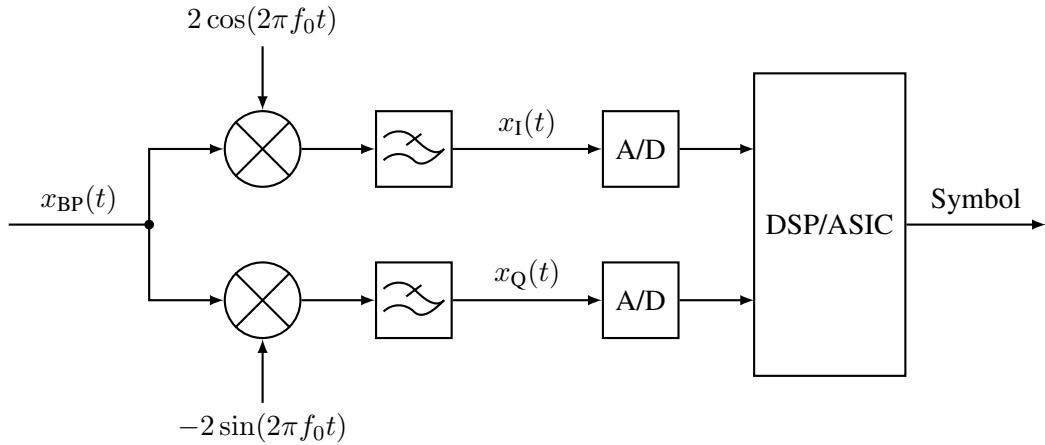


Figure 1.9 I/Q demodulator.

1.5 Cascaded-Integrator Comb (CIC) Filters

Often, in communication systems we start with a rather high sampling rate and would like to reduce the sampling rate to save on memory and power, see Fig. 1.10. The anti-alias filter required for the high-sampling rate might be a relaxed analog filter but might not be adequate for the lower sampling rate. Hence, prior to just leaving out samples we need to refilter the signal, this time with a lower cut-off frequency, adequate for the lower sampling rate. The second filtering process takes place in the digital domain, so in order to keep computational complexity down, we are looking for simple filters, possibly without the use of multiplication operations.

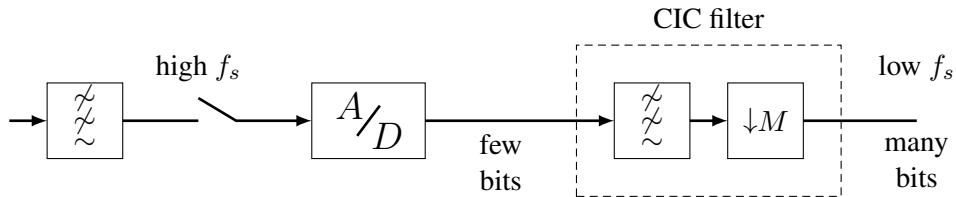


Figure 1.10 Sampling-rate conversion inside a communications receiver.

1.5.1 Comb Filters

The name *comb filters* stems from the analogy to an ordinary hair comb with equally spaced teeth. Depending on the source, this refers to the nulls in the frequency band (Proakis/Manolakis) or to the impulse response of a rake of unit pulses in the time domain. The periodic appearance of the nulls in the frequency domain makes the filter also an ideal tool for the rejection of power-line harmonics. The implementation of such a filter as an FIR filter is particularly simple, since no multipliers are needed. The z -transform of the impulse response is

$$H(z) = \sum_{k=0}^{M-1} z^{-k}, \quad (1.22)$$

essentially a moving-average function¹, whose impulse response is displayed in Fig. 1.11.

¹For the sake of legibility the division by M as would be required by a true averaging function is omitted.

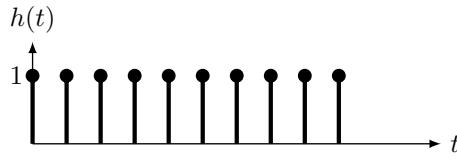


Figure 1.11 Impulse response of comb filter, $M = 10$.

By using $z = e^{j\omega T_s}$ we obtain

$$\begin{aligned}
 H(z) &= \sum_{k=0}^{M-1} e^{-j\omega T_s \cdot k} \\
 &= e^{-j\omega T_s \frac{M-1}{2}} \sum_{k=0}^{M-1} e^{j\omega T_s (\frac{M-1}{2} - k)} \\
 &= e^{-j\omega T_s \frac{M-1}{2}} \sum_{k=1}^{M/2} \left(e^{j\omega T_s \frac{2k-1}{2}} + e^{-j\omega T_s \frac{2k-1}{2}} \right) \\
 &= 2e^{-j\omega T_s \frac{M-1}{2}} \sum_{k=1}^{M/2} \cos \left(\omega T_s \frac{2k-1}{2} \right). \tag{1.23}
 \end{aligned}$$

At first, this does not look particularly useful. However, the FIR impulse response of Eq. (1.22) can also be written in an IIR form (without making the impulse response infinite):

$$H(z) = \sum_{k=0}^{M-1} z^{-k} = \frac{1 - z^{-M}}{1 - z^{-1}}. \tag{1.24}$$

This comes either as the closed form of a geometric series or by noticing that the moving average sums up the same samples except for the first one and the last one, which have to be added additionally, leading to a block diagram as shown in Fig. 1.12.

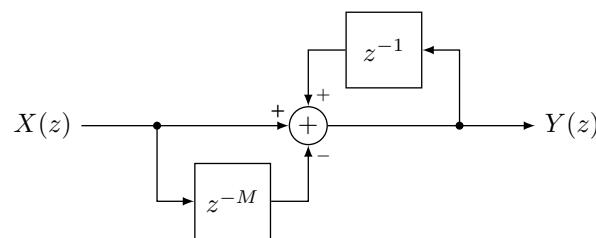


Figure 1.12 Block diagram of a simple comb filter.

We may also reformulate

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{k=0}^{M-1} z^{-k} = \frac{1 - z^{-M}}{1 - z^{-1}} \tag{1.25}$$

into

$$Y(z) = Y(z^{-1}) + X(z) - X(z^{-M}), \tag{1.26}$$

or the recursion equation

$$y_k = y_{k-1} + x_k - x_{k-M}. \tag{1.27}$$

We can now identify three parts in the summation: the accumulated state, the new input and the compensation of the long past input, as shown in Fig. 1.13.

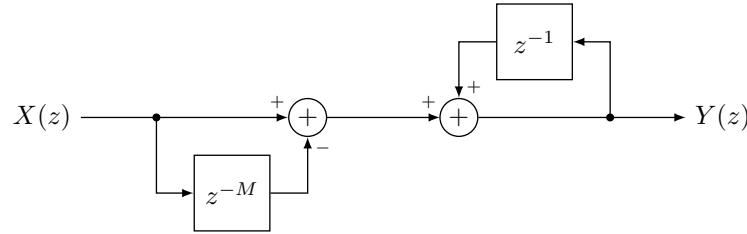


Figure 1.13 Block diagram of a simple comb filter with separate feedforward and recursive parts.

Lastly, by inserting $z = e^{j\omega T_s}$ into Eq. (1.24), we obtain

$$\begin{aligned} H(z) &= \frac{1 - e^{-j\omega T_s \cdot M}}{1 - e^{-j\omega T_s}} \\ &= \frac{e^{-j\frac{\pi f}{f_s} M} \left(e^{j\frac{\pi f}{f_s} M} - e^{-j\frac{\pi f}{f_s} M} \right)}{e^{-j\frac{\pi f}{f_s}} \left(e^{j\frac{\pi f}{f_s}} - e^{-j\frac{\pi f}{f_s}} \right)} \\ &= e^{-j\frac{\pi f}{f_s} (M-1)} \cdot \frac{\sin\left(\frac{\pi f}{f_s} M\right)}{\sin\left(\frac{\pi f}{f_s}\right)}. \end{aligned} \quad (1.28)$$

We can easily see that $H(z)$ has zeros at

$$z = e^{j2\pi \frac{n}{M}}, \quad n \in \mathbb{N} \setminus \{0\}. \quad (1.29)$$

Note that there is no zero at $n = 0$, since this zero is cancelled by a pole. The zeros at equidistant places is what gives the comb filter its name.

1.5.2 Cascaded Integrator-Comb (CIC) Filters

Often enough the comb filter of the previous section is used prior to a downsampling stage. The comb filter then serves as an anti-aliasing filter of the new sampling rate. If the reduction rate of the downsampling process is chosen as M , every operation z^{-M} simply reduces to z^{-1} with respect to the new sampling rate. This is beneficial because it reduces the memory requirements by a factor of M . We can then redraw the setup of Fig. 1.14 to Fig. 1.15

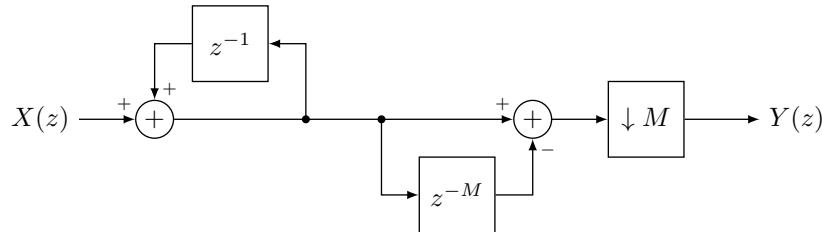


Figure 1.14 Downsampling CIC filter.

A comb filter with the transfer function according to Eq. (1.28) may be cascaded n times to obtain the n th power of the transfer function². If the integrators are collected, followed by the downsampling, followed by

²Note that this n is not the same as used to indicate an arbitrary natural integer number in Eq. (1.29)

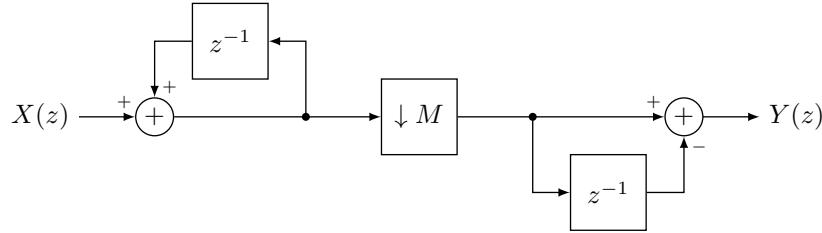


Figure 1.15 Rearranged downsampling CIC filter with reduced memory requirements.

the comb section we obtain the following block diagram of Fig. 1.16(a). Hogenauer was one of the first to realize this in his seminal paper [19]. The transfer function is simply

$$H(z) = \left(\frac{1 - z^{-M}}{1 - z^{-1}} \right)^n. \quad (1.30)$$

Similarly, an upsampling process with interpolation can be implemented using the block diagram according to Fig. 1.16(b).

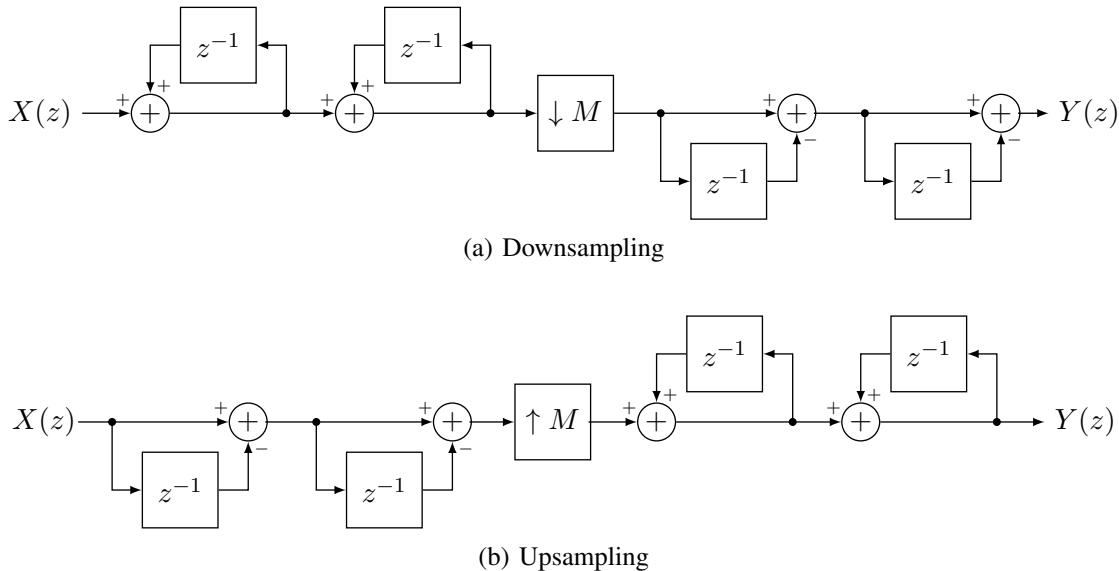


Figure 1.16 Down- and upsampling CIC filters of order $n = 2$.

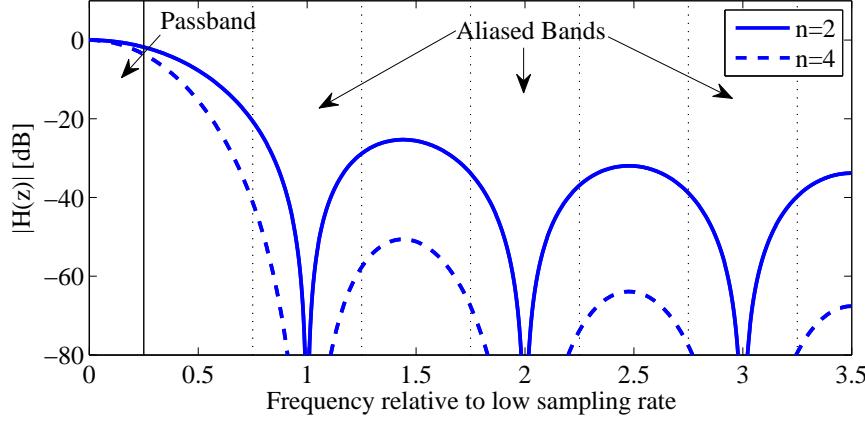
1.5.3 Aliasing and Passband Roll-Off of Downsampling CIC Filters

CIC filters are typically used in modern software defined radios, where A/D and D/A conversions are performed at high data-rates. For example, a downsampling CIC filter is often used as the first signal processing stage after A/D conversion. The following stages are then easier to implement and consume less power because they operate at a lower sampling rate.

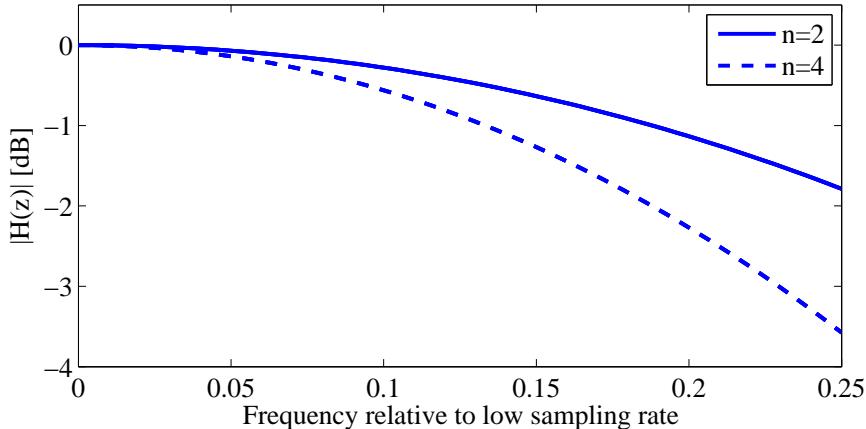
The big advantage of CIC decimators is that they don't require multipliers. Especially in FPGAs, multipliers require much more hardware resources and power than adders. But this advantage comes at a price. The passband of a CIC filter is not flat but has a roll-off that increases with increasing order of the CIC filter.

As an example, Fig. 1.17(a) depicts two frequency responses of downsampling CIC filters with $M = 7$. The amplitude response is normalized to a DC gain of 0 dB. The orders of the filters are $n = 2$ and $n = 4$. The

filter with the higher order exhibits clearly lower aliasing. However, the zoom in Fig. 1.17(b) shows that it also has more roll-off at the passband corner frequency. The corner frequency is arbitrarily defined in this example as $f_c = \frac{1}{4}$, relative to the low sampling rate.



(a) Frequency responses



(b) Passband roll-offs. The decay of the amplitude response at the right-hand side of the frequency axis is called the passband droop.

Figure 1.17 Frequency responses and passband roll-offs of two CIC filters of orders $n = 2$ and $n = 4$ with $M = 7$ and $f_c = \frac{1}{4}$.

The passband roll-off is often compensated by a following FIR filter. This compensation filter requires less computational power because it operates at the low sampling rate. It is also common practice to add further downsampling stages after the compensation filter, since the passband is typically much smaller than half of the lower sampling rate.

When designing a downsampling CIC filter, one typically has to fulfil certain maximum-aliasing and passband roll-off criteria. When the passband and input sampling rate are given, one can only adjust the order and downsampling factor of the CIC filter. Increasing the order reduces aliasing but increases passband roll-off. On the other hand, decreasing the downsampling factor decreases both, the aliasing and the passband roll-off. However, this option increases the computational cost for the following stages since the low sampling rate is increased due to the smaller downsampling factor.

2 Detection Theory

2.1 Framework

In digital communication systems we often have the task to decide which information has been sent. The alphabet size, corresponding to the number of options, is finite, often only consists of two symbols (binary signalling scheme). In contrast to analog systems, where we have to estimate a parameter or a continuous information source, and for which we therefore make use of estimation theory, we speak of detection theory for such digital systems. Detection theory is closely related to hypotheses testing as we will later see. The general task of the detection theory is the description of the methods to decide on the event that occurred based on possibly erroneous observations. The goal is to make the decision such that a certain criterion is optimized, e.g., minimal error rate, minimal cost, maximal detection rate etc. The structure of the optimal detectors thus depends on the knowledge about the process and the quality measure to be optimized. The process can be visualized as in Fig. 2.1.

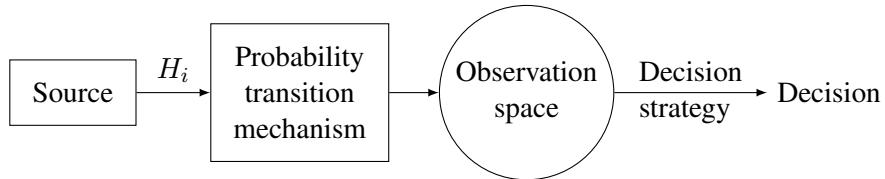


Figure 2.1 Detection process.

The source produces M -ary symbols. Each symbol corresponds to a hypothesis H_i , $i = 0 \dots M - 1$. The probability transition mechanism knows the symbol i produced by the source (and therefore the corresponding true hypothesis H_i) and generates points in the N -dimensional observation space according to a certain, known probability distribution $p(\mathbf{r}|H_i)$. Each point in the observation space corresponds to N observations. Therefore, we use the vector notation $\mathbf{r} = [r_1, r_2, \dots, r_N]^T$. The decision strategy assigns a hypothesis to each point in the observation space. The decision is correct, if $\hat{H}_i = H_i$, otherwise it is wrong. In the case of binary hypotheses we only have H_0 and H_1 .

Option	True hypothesis	Decided hypothesis	Decision	In statistics	Statistics name
1	H_0	\hat{H}_0	correct		Specificity
2	H_0	\hat{H}_1	wrong	Type-I error, α	Statistical significance
3	H_1	\hat{H}_0	wrong	Type-II error, β	
4	H_1	\hat{H}_1	correct		Power, sensitivity

Table 2.1 Decision options for binary hypotheses.

Table 2.1 shows the decision options for binary hypotheses. Note that we often have symmetrical relations (e.g. signalling with ± 1) in communications, whereas in statistics we mostly refer to H_0 as the null hypothesis, indicating there is no effect, while H_1 stands for the hypothesis that there is an effect (which is yet to be proven). Fig. 2.2 shows the decision chain for a binary channel. The source generates binary symbols d_k with the probabilities

$$P_0 = P(d_k = 0), \quad (2.1a)$$

$$P_1 = P(d_k = 1). \quad (2.1b)$$

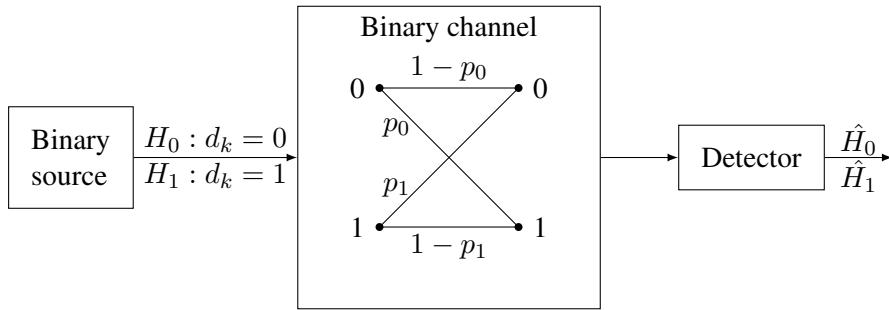


Figure 2.2 Binary hypotheses for a binary channel.

The binary channel, representing the probability transition mechanism, is given by the following conditional probabilities:

$$P(0 \text{ at the output of the channel} | 0 \text{ at the input of the channel}) = 1 - p_0, \quad (2.2a)$$

$$P(1 \text{ at the output of the channel} | 0 \text{ at the input of the channel}) = p_0, \quad (2.2b)$$

$$P(0 \text{ at the output of the channel} | 1 \text{ at the input of the channel}) = p_1, \quad (2.2c)$$

$$P(1 \text{ at the output of the channel} | 1 \text{ at the input of the channel}) = 1 - p_1. \quad (2.2d)$$

The quality of transmission in digital communications is most often expressed in some kind of error probability. Error probability is a very general term. It may refer to the symbol transmitted, to the bits that comprise a symbol, or to a whole frame of data bits. In case of bits, it may concern raw bits (in the form transferred over the channel) or information bits, which are the decoded bits after error decoding. Naturally, the latter bit error rate should be smaller than the raw bit error rate.

If we want to find out about such error rates, there are different methods. One method would be to just build a system, let it run for a while and count the errors occurring. This is the most costly method. Furthermore, it has the disadvantage that very small error rates need extremely long running times to verify. A second method is to simulate systems using so-called Monte-Carlo (MC) simulation methods. They receive their name from their approach of essentially letting a random string of data be disturbed by random noise and measuring the resulting error rate. The advantage is that the system does not actually need to be built, just a good model needs to be available. On the other hand, simulations might not run in real-time. Hence, the simulations of low symbol error rates can last even longer. A third method is to compute error probabilities in an analytic way. This method is described in the following.

The easiest error rate to compute is the symbol error probability. Let us consider the BPSK case, which is the easiest of all signaling schemes. As with any binary signaling schemes, the size of the amplitude values has a direct impact on the error probability. In the absence of interference, no error would ever occur, regardless the distance between the constellation points. However, thermal noise at the input to the receiver is always present. The noise is in the form of an additive Gaussian distributed signal with the power $\sigma^2 = N_0/2$. The signal, on the other hand, has the symbol energy E_S . When a positive amplitude is sent, the combined signal with the noise has a distribution $p(r|s = +1)$ like in Fig. 2.3 (lower), otherwise $p(r|s = -1)$ as in Fig. 2.3 (upper). A detector has to decide which symbol has been sent. Although a negative received value is possible when a positive symbol has been sent (and a strong negative noise signal is added), such a scenario is less likely than an original negative signal. It is therefore reasonable to set the decision level in the middle of the two signaling levels (at zero).

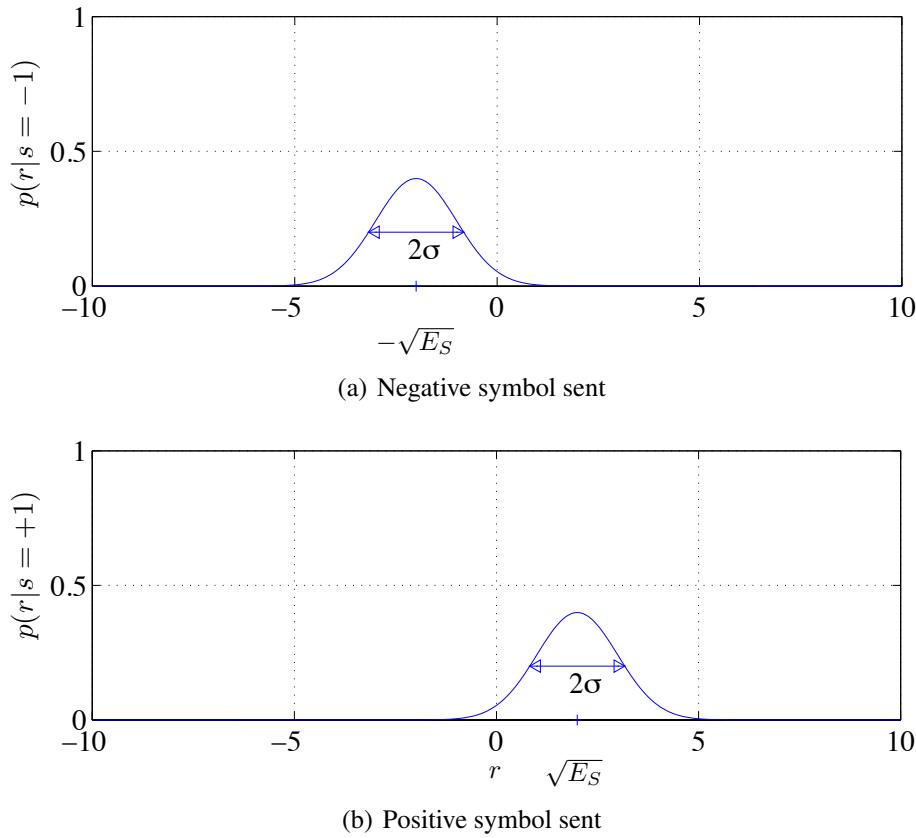


Figure 2.3 Signal distribution after additive noise.

That way, the probability of error is the tail of the Gaussian distribution

$$\begin{aligned}
 P_E &= \int_0^\infty p(r|s = -1) dr \\
 &= \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r+\sqrt{E_S})^2}{2\sigma^2}} dr \\
 &= \int_0^\infty \frac{1}{\sqrt{2\pi}\sqrt{\frac{N_0}{2}}} e^{-\frac{(r+\sqrt{E_S})^2}{N_0}} dr \\
 &= \int_{\sqrt{E_S}}^\infty \frac{1}{\sqrt{2\pi}\sqrt{\frac{N_0}{2}}} e^{-\frac{r^2}{N_0}} dr \\
 &= Q\left(\sqrt{\frac{2E_S}{N_0}}\right). \tag{2.3}
 \end{aligned}$$

$Q(\cdot)$ is called the Q-function or the tail function of the Gaussian distribution. It cannot be written in a direct form, but is often tabulated. Such a table can be found in Appendix C.1. Other symbol error probabilities can be derived using a similar line of argument. The results are collected in Table 2.2 and in Fig. 2.4.

Up to now we have considered symbol error rates. For bit error rates, the computation is often slightly more complicated. For binary signaling, the symbol error rates (SER) and the bit error rates (BER) are identical,

Modulation scheme	Levels/range	Normalized amplitude	SER (AWGN)	Comments
BPSK	$\pm A$	$A = 1$	$Q\left(\sqrt{\frac{2E_s}{N_0}}\right)$	
DPSK	$\pm A$	$A = 1$	$2Q\left(\sqrt{\frac{2E_s}{N_0}}\right)$ $\frac{1}{2}\exp\left(-\frac{E_s}{N_0}\right)$	coherent noncoherent
QPSK	$(\pm 1 \pm j)A$	$A = \frac{1}{\sqrt{2}}$	$1 - \left(1 - Q\left(\sqrt{\frac{E_s}{N_0}}\right)\right)^2$	coherent
DQPSK	$(\pm 1 \pm j)A$	$A = \frac{1}{\sqrt{2}}$	$2\left(1 - \left(1 - Q\left(\sqrt{\frac{E_s}{N_0}}\right)\right)^2\right)$ $\frac{1}{2}\exp\left(-\frac{E_s}{2N_0}\right)$	coherent noncoherent
M -PSK		$\frac{1}{M} \sum_{m=1}^M A \delta\left(x - e^{j(2m-1)\frac{\pi}{M}}\right), \ m \leq M$	$A = 1$ $A = \sqrt{\frac{3}{M^2-1}}$	$\leq 2Q\left(\sqrt{\frac{2E_s}{N_0}} \sin \frac{\pi}{M}\right)$ $2\frac{M-1}{M}Q\left(\sqrt{\frac{E_s}{N_0}} \sqrt{\frac{6}{M^2-1}}\right)$
M -PAM	$\pm(2m-1)A, \ m \leq M/2$			
M -QAM	$(\pm(2m-1) \pm j(2n-1))A, \ m, n \leq \sqrt{M}/2$	$A = \sqrt{\frac{3}{2(M-1)}}$	$1 - \left(1 - 2\frac{\sqrt{M}-1}{\sqrt{M}}Q\left(\sqrt{\frac{E_s}{N_0}} \sqrt{\frac{3}{M-1}}\right)\right)^2$	
2-FSK	$\Delta f = h \cdot \frac{1}{T}, \ h_{\min} = 0.5$ $h_{\min} = 1$		$Q\left(\sqrt{\frac{E_s}{N_0}}\right)$ $\frac{1}{2}\exp\left(-\frac{E_s}{2N_0}\right)$	orthogonal, coherent orthogonal, noncoherent

Table 2.2 Constellations and SERs in an AWGN channel of some digital modulation formats. Most SERs can be found in [28].

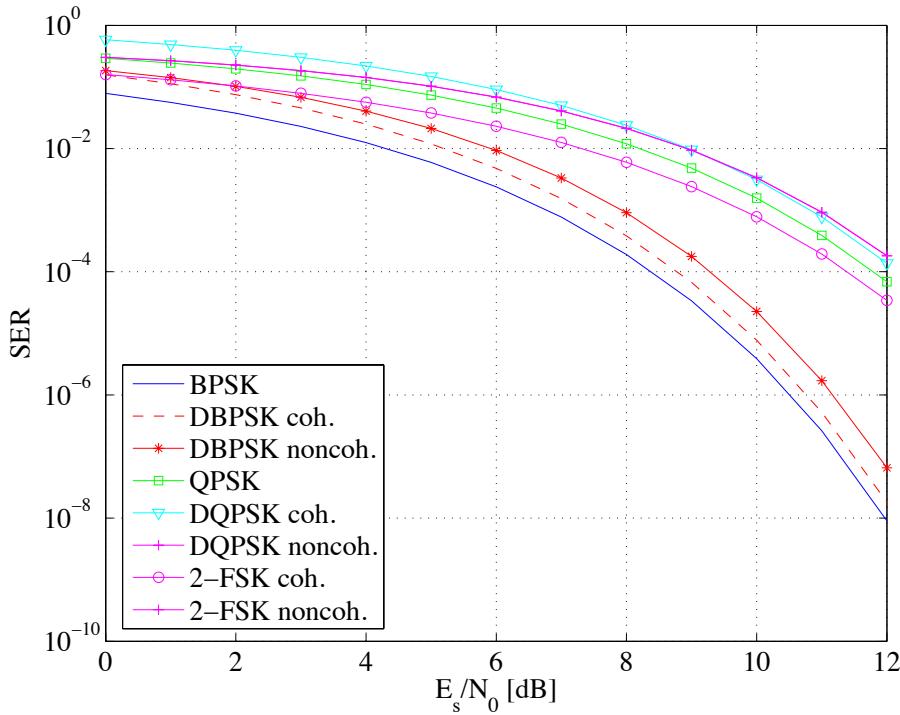


Figure 2.4 SERs of some digital modulation formats.

since one symbol consists of one bit. It is interesting to note that from Table 2.2 we see that

$$\begin{aligned} \text{SER}_{\text{QPSK}} &= 1 - \left(1 - Q \left(\sqrt{\frac{E_S}{N_0}} \right) \right)^2 \\ &= 1 - \left(1 - Q \left(\sqrt{\frac{2E_b}{N_0}} \right) \right)^2 \\ &\approx 2Q \left(\sqrt{\frac{2E_b}{N_0}} \right) \quad , \quad \text{for small } Q(.). \end{aligned} \quad (2.4)$$

If we Gray-code the QPSK bitstream we can get roughly

$$\text{BER}_{\text{QPSK}} = \frac{1}{2} \text{SER}_{\text{QPSK}} = Q \left(\sqrt{\frac{2E_b}{N_0}} \right), \quad (2.5)$$

which corresponds exactly to the BER of BPSK. For the evaluation of the SER, we can easily see that the closer the constellation points are, the more likely errors are occurring.

2.2 Maximum-Likelihood (ML) Detection

Without knowing we have produced the first detection strategy in the previous section, the maximum-likelihood (ML) detection. The ML detectors thus acts after the following rule: For every received signal vector \mathbf{r} (because there are possibly several observations for a symbol transmitted) choose the index i for \hat{d}_i as the one that maximizes the conditional probability $p(\mathbf{r}|H_i)$.

2.3 Maximum A-Posteriori (MAP) Detection

In the case of binary signalling, we have formed our decision rule as the middle between the two signalling points, in other words, the threshold is equal to the mean of the two levels. In the case where the signalling levels have different probabilities we could, however, further minimize the probability of error. In principle, the detection rule is the same as the ML detection rule, but now the *a-priori* probabilities for the symbols are assumed known. We want to minimize the probability of error, or in other words, we want to maximize the probability of a correct decision. This probability can be computed by integrating over the observation space

$$1 - P_E = \int_{-\infty}^{\infty} P(H_d(\mathbf{r})|\mathbf{r}) \cdot p(\mathbf{r}) d\mathbf{r} \quad (2.6)$$

where $d(\mathbf{r})$ is the decision function of the received vector \mathbf{r} . Because $p(\mathbf{r})$ is non-negative and we can choose $d(\mathbf{r})$ individually for all \mathbf{r} , we can make our decision rule such that $P(H_d(\mathbf{r})|\mathbf{r})$ is maximized: Given the observation vector \mathbf{r} , choose the hypotheses \hat{H}_i that was most probably leading to this observation. For binary hypotheses we can mathematically formulate the decision rule as

$$\begin{array}{c} H_1 \\ \uparrow \\ P(H_1|\mathbf{r}) \gtrless P(H_0|\mathbf{r}) \\ \downarrow \\ H_0 \end{array} \quad (2.7)$$

which reads as "choose H_1 if the probability $P(H_1)$ given the observation \mathbf{r} is higher than $P(H_0)$ given the same observation, otherwise choose H_0 ". Very often, the *a-posteriori* probabilities $P(H_i|\mathbf{r})$ are not directly accessible. They can be computed through the Bayesian rule and the *a-priori* probabilities $P_i = P(H_i)$

$$P(H_i|\mathbf{r}) \cdot p(\mathbf{r}) = p(\mathbf{r}|H_i) \cdot P_i \quad (2.8)$$

as

$$P(H_i|\mathbf{r}) = \frac{p(\mathbf{r}|H_i) \cdot P_i}{p(\mathbf{r})}. \quad (2.9)$$

Since $p(\mathbf{r})$ is independent of the hypothesis H_i we can formulate the MAP decision test as

$$\begin{array}{c} H_1 \\ \uparrow \\ p(\mathbf{r}|H_1) \cdot P_1 \gtrless p(\mathbf{r}|H_0) \cdot P_0 \\ \downarrow \\ H_0 \end{array} \quad (2.10)$$

or

$$\begin{array}{c} H_1 \\ \uparrow \\ \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)} > \frac{P_0}{P_1} \\ \downarrow \\ H_0 \end{array} \quad (2.11)$$

In general we have a rule that tells us that we have to compare a likelihood function $\Lambda(\mathbf{r}) \triangleq \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)}$ to a threshold $\eta \triangleq \frac{P_0}{P_1}$, leading to

$$\begin{array}{c} H_1 \\ \uparrow \\ \Lambda(\mathbf{r}) \gtrless \eta. \\ \downarrow \\ H_0 \end{array} \quad (2.12)$$

2.4 Bayes Criterion / Bayes Test

Sometimes, there are additional costs associated with correct or erroneous decisions. The Bayes criterion describes the rule as the decision for the hypothesis that minimizes the expected costs. Assume the costs are given as C_{ij} , where i indicates the detected hypothesis \hat{H}_i , and j the true hypothesis H_j , respectively. The average cost, also called the risk, can be computed as

$$R \stackrel{\Delta}{=} E\{C\} = \sum_i \sum_j C_{ij} \cdot P(\hat{H}_i, H_j). \quad (2.13)$$

From Bayes rule, we know

$$P(\hat{H}_i, H_j) = P(\hat{H}_i | H_j) \cdot P_j. \quad (2.14)$$

In the binary case we know (you have to make your decision)

$$P(\hat{H}_0 | H_0) + P(\hat{H}_1 | H_0) = 1, \quad (2.15a)$$

$$P(\hat{H}_1 | H_1) + P(\hat{H}_0 | H_1) = 1, \quad (2.15b)$$

so that the risk evaluates to

$$\begin{aligned} R &= (C_{00} \cdot P(\hat{H}_0 | H_0) + C_{10} \cdot P(\hat{H}_1 | H_0)) \cdot P_0 \\ &\quad + (C_{01} \cdot P(\hat{H}_0 | H_1) + C_{11} \cdot P(\hat{H}_1 | H_1)) \cdot P_1 \\ &= (C_{00} \cdot P(\hat{H}_0 | H_0) + C_{10} \cdot (1 - P(\hat{H}_0 | H_0))) \cdot P_0 \\ &\quad + (C_{01} \cdot P(\hat{H}_0 | H_1) + C_{11} \cdot (1 - P(\hat{H}_0 | H_1))) \cdot P_1 \\ &= (C_{00} - C_{10}) \cdot P(\hat{H}_0 | H_0) \cdot P_0 + C_{10} \cdot P_0 \\ &\quad + (C_{01} - C_{11}) \cdot P(\hat{H}_0 | H_1) \cdot P_1 + C_{11} \cdot P_1. \end{aligned} \quad (2.16)$$

Because the part $C_{10} \cdot P_0 + C_{11} \cdot P_1$ is not influenced by choosing the decision region, we can form a new risk function as

$$\tilde{R} = (C_{00} - C_{10}) \cdot P(\hat{H}_0 | H_0) \cdot P_0 + (C_{01} - C_{11}) \cdot P(\hat{H}_0 | H_1) \cdot P_1. \quad (2.17)$$

The conditional probabilities used are the pdfs integrated over the decision region

$$P(\hat{H}_0 | H_0) = \int_{Z_0} p(\mathbf{r} | H_0) d\mathbf{r}, \quad (2.18a)$$

$$P(\hat{H}_0 | H_1) = \int_{Z_0} p(\mathbf{r} | H_1) d\mathbf{r}. \quad (2.18b)$$

Now,

$$\begin{aligned} \tilde{R} &= (C_{00} - C_{10}) \cdot \int_{Z_0} p(\mathbf{r} | H_0) d\mathbf{r} \cdot P_0 + (C_{01} - C_{11}) \cdot \int_{Z_0} p(\mathbf{r} | H_1) d\mathbf{r} \cdot P_1 \\ &= \int_{Z_0} \left([(C_{00} - C_{10}) \cdot p(\mathbf{r} | H_0) \cdot P_0] + [(C_{01} - C_{11}) \cdot p(\mathbf{r} | H_1) \cdot P_1] \right) d\mathbf{r} \\ &= \int_{Z_0} \left([(C_{01} - C_{11}) \cdot p(\mathbf{r} | H_1) \cdot P_1] - [(C_{10} - C_{00}) \cdot p(\mathbf{r} | H_0) \cdot P_0] \right) d\mathbf{r}. \end{aligned} \quad (2.19)$$

Under the assumption that costs for correct decisions are always smaller than for wrong decisions, both cost differences in Eq. (2.19) are positive. The square brackets in Eq. (2.19) have therefore always positive values and the risk \tilde{R} is minimized if the integrand is forced to be negative by choosing points to belong to the decision region Z_0 only if

$$[(C_{01} - C_{11}) \cdot p(\mathbf{r}|H_1) \cdot P_1] < [(C_{10} - C_{00}) \cdot p(\mathbf{r}|H_0) \cdot P_0]. \quad (2.20)$$

Hence, the Bayes test is

$$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)} \stackrel{\substack{\uparrow \\ H_1}}{>} \frac{P_0 \cdot (C_{10} - C_{00})}{P_1 \cdot (C_{01} - C_{11})} = \eta. \quad (2.21)$$

Eq. (2.21) is also called a likelihood-ratio test. Since the natural logarithm is a monotonically increasing function and either side of Eq. (2.21) is non-negative, we can also use

$$\ln \Lambda(\mathbf{r}) \stackrel{\substack{\uparrow \\ H_1}}{>} \ln \eta, \quad (2.22)$$

which comes in handy for many expressions involving exponential functions, such as Gaussian distributions.

2.5 Minimax Criterion

Now we observe the case where the costs for correct decisions are zero, the cost for a miss is $C_{01} = C_M$ and the cost for a false-alarm is $C_{10} = C_{FA}$. The different probabilities can be formed as

$$P_D = \int_{Z_1} p(\mathbf{r}|H_1) d\mathbf{r} \quad \text{detection probability or power of a test}, \quad (2.23)$$

$$\begin{aligned} P_M &= \int_{Z_0} p(\mathbf{r}|H_1) d\mathbf{r} && \text{miss probability } (\beta) \\ &= 1 - P_D, \end{aligned} \quad (2.24)$$

$$P_{FA} = \int_{Z_1} p(\mathbf{r}|H_0) d\mathbf{r} \quad \text{false-alarm probability } (\alpha). \quad (2.25)$$

The risk is

$$\begin{aligned} R &= P_M C_M \cdot P_1 + P_{FA} C_{FA} \cdot P_0 \\ &= P_M C_M \cdot (1 - P_0) + P_{FA} C_{FA} \cdot P_0 \\ &= P_M C_M + P_0 \cdot (P_{FA} C_{FA} - P_M C_M). \end{aligned} \quad (2.26)$$

A-priori probabilities are not known. It can be shown that the Minimax criterion leads to

$$P_{FA} C_{FA} = P_M C_M, \quad (2.27)$$

The resulting maximum costs are $P_M C_M$. The minimax criterion, as its name implies, leads to the minimal worst-case costs.

2.6 Neyman-Pearson Test

For the Neyman-Pearson test we are interested in the detection probability and the false-alarm probability. These are conditional probabilities making any knowledge on *a-priori* probabilities as well as costs unessential. The test is now formulated as "Maximize the detection probability P_D under the condition that the false-alarm probability P_{FA} does not exceed a maximal value of α ". The quantity to be maximized is

$$P_D = \int_{Z_1} p(\mathbf{r}|H_1) d\mathbf{r}, \quad (2.28)$$

under the above-mentioned condition that the false-alarm probability

$$P_{FA} = \int_{Z_1} p(\mathbf{r}|H_0) d\mathbf{r} \leq \alpha. \quad (2.29)$$

It can be shown that the condition can be formulated by using a Lagrange multiplier resulting in a Likelihood-Ratio test

$$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)} \stackrel{\substack{\uparrow \\ H_1}}{>} \lambda \quad \text{with } P_{FA}(\lambda) \stackrel{!}{=} \alpha. \quad (2.30)$$

2.7 Example: On-Off Keying in an AWGN Channel

In this example, we assume a source that transmits information by switching on and off a DC-voltage with amplitude m . The detector observes this signal with additive white Gaussian noise with variance σ^2 and zero mean. Furthermore, the detector makes N independent observations of the signal to form the observation vector $\mathbf{r} = [r_1, r_2, \dots, r_N]^T$. Hence, we can define the two hypotheses

$$H_1 : \quad r_i = m + n_i, \quad i = 1, 2, \dots, N \quad (2.31a)$$

$$H_0 : \quad r_i = -n_i. \quad (2.31b)$$

The conditional probabilities are

$$p(\mathbf{r}|H_1) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(r_i - m)^2}{2\sigma^2}\right), \quad (2.32a)$$

$$p(\mathbf{r}|H_0) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{r_i^2}{2\sigma^2}\right). \quad (2.32b)$$

The likelihood-ratio test can now be formulated as

$$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)} = \frac{\exp\left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^N r_i^2 - 2m \sum_{i=1}^N r_i + Nm^2\right)\right]}{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^N r_i^2\right]} \stackrel{\substack{\uparrow \\ H_1}}{>} \frac{P_0}{P_1} = \eta. \quad (2.33)$$

Using the logarithm, we get

$$\ln \Lambda(\mathbf{r}) = \frac{m}{\sigma^2} \sum_{i=1}^N r_i - \frac{Nm^2}{2\sigma^2} \stackrel{\substack{\uparrow \\ H_1}}{>} \stackrel{\downarrow \\ H_0}{<} \ln \eta. \quad (2.34)$$

To perform the test, we can rearrange Eq. (2.34) to

$$\sum_{i=1}^N r_i \stackrel{H_1}{\underset{H_0}{\gtrless}} \frac{\sigma^2}{m} \ln \eta + \frac{Nm}{2}. \quad (2.35)$$

This means we have to sum up the observations and compare the sum to a threshold that depends on the noise variance σ^2 , the amplitude m , and the ratio of *a-priori* probabilities η . Eq. (2.35) shows that we need to know only the sum of all received samples. This sum is therefore called sufficient statistics to indicate that the individual values of r_i are not of importance. It is just as good to know the sum.

2.8 Summary

To sum up and compare the methods, we use again the risk as the expected cost

$$R = \sum_i \sum_j p_j C_{ij} \int_{Z_i} p_{r|H_j} dr. \quad (2.36)$$

Rule	Generic description	Binary test	Known		
			cond. P	<i>a-priori</i> P	costs
Bayes	$\min \sum_{i=1}^I \sum_{j=1}^J C_{ij} P[H_j] p_{R H_j}(r)$	$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r} H_1)}{p(\mathbf{r} H_0)} \stackrel{H_1}{\underset{H_0}{\gtrless}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} = \eta$	✓	✓	✓
MAP	$\max \sum_{j=1}^J P[H_j] p_{R H_j}(r)$	$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r} H_1)}{p(\mathbf{r} H_0)} \stackrel{H_1}{\underset{H_0}{\gtrless}} \frac{P_0}{P_1} = \eta$	✓	✓	—
ML	$\max \sum_{j=1}^J p_{R H_j}(r)$	$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r} H_1)}{p(\mathbf{r} H_0)} \stackrel{H_1}{\underset{H_0}{\gtrless}} 1 = \eta$	✓	—	—
Minimax	minimize maximum cost	$P_F C_F = P_M C_M$	✓	—	✓
Neyman-Pearson	max P_D for $P_{FA} \leq \alpha$	$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r} H_1)}{p(\mathbf{r} H_0)} \stackrel{H_1}{\underset{H_0}{\gtrless}} \lambda = \eta$ with $P_F(\lambda) = \alpha$	✓	—	—

Table 2.3 Comparison of detection methods.

Table 2.3 shows an overview of the different detection methods explained in this chapter. Functions often used to estimate the probability of error are the error function and associated tail functions of Gaussian distributions. The different functions are defined as follows:

$$\phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{\xi^2}{2}\right) d\xi, \quad (2.37)$$

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{\xi^2}{2}\right) d\xi, \quad (2.38)$$

$$\text{erf}(x) \triangleq \frac{2}{\sqrt{\pi}} \int_0^x \exp(-\xi^2) d\xi, \quad (2.39)$$

$$\text{erfc}(x) \triangleq \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-\xi^2) d\xi. \quad (2.40)$$

Their relationships are shown in Tables 2.4 and 2.5.

	$\phi(x)$	$Q(x)$	$\text{erf}(x)$	$\text{erfc}(x)$
$\phi(x) =$	$\phi(x)$	$1 - Q(x)$	$\frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{x}{\sqrt{2}}\right)$	$1 - \frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2}}\right)$
$Q(x) =$	$1 - \phi(x)$	$Q(x)$	$\frac{1}{2} - \frac{1}{2} \text{erf}\left(\frac{x}{\sqrt{2}}\right)$	$\frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2}}\right)$
$\text{erf}(x) =$	$2\phi(\sqrt{2}x) - 1$	$1 - 2Q(\sqrt{2}x)$	$\text{erf}(x)$	$1 - \text{erfc}(x)$
$\text{erfc}(x) =$	$2 - 2\phi(\sqrt{2}x)$	$2Q(\sqrt{2}x)$	$1 - \text{erf}(x)$	$\text{erfc}(x)$

Table 2.4 Relationship of error functions.

	$\phi^{-1}(x)$	$Q^{-1}(x)$	$\text{erf}^{-1}(x)$	$\text{erfc}^{-1}(x)$
$\phi^{-1}(x) =$	$\phi^{-1}(x)$	$Q^{-1}(1-x)$	$\sqrt{2} \text{erf}^{-1}(2x-1)$	$\sqrt{2} \text{erfc}^{-1}(2-2x)$
$Q^{-1}(x) =$	$\phi^{-1}(1-x)$	$Q^{-1}(x)$	$\sqrt{2} \text{erf}^{-1}(1-2x)$	$\sqrt{2} \text{erfc}^{-1}(2x)$
$\text{erf}^{-1}(x) =$	$\frac{1}{\sqrt{2}}\phi^{-1}\left(\frac{1+x}{2}\right)$	$\frac{1}{\sqrt{2}}Q^{-1}\left(\frac{1-x}{2}\right)$	$\text{erf}^{-1}(x)$	$\text{erfc}^{-1}(1-x)$
$\text{erfc}^{-1}(x) =$	$\frac{1}{\sqrt{2}}\phi^{-1}\left(1-\frac{x}{2}\right)$	$\frac{1}{\sqrt{2}}Q^{-1}\left(\frac{x}{2}\right)$	$\text{erf}^{-1}(1-x)$	$\text{erfc}^{-1}(x)$

Table 2.5 Relationship of inverse error functions.

3 Propagation

3.1 The Wireless Channel

The physical medium that transports the signal from the transmitter to the receiver is called the communication channel. In wireless communications, this is simply the atmosphere or free space (satellite communications). Additionally, there are wavelength-dependent phenomena such as earth-guided waves, reflected and diffracted waves etc. The interface between the electric circuit that produces the signal and the medium is the antenna, and as such will be considered as part of the wireless channel, since loss and other closely related phenomena occur in both parts the antenna and the channel.

Any channel (and this is not restricted to the wireless one) modifies the signal in ways more or less unpredictable to the receiver. These modifications are called noise. We distinguish between additive and multiplicative noise, as shown in Fig. 3.1. The additive noise arises from noise sources within the receiver itself, such as thermal noise, flicker noise, shot noise, and also from external noise sources such as atmospheric effects, cosmic radiation, and interference from other transmitters or electric devices.

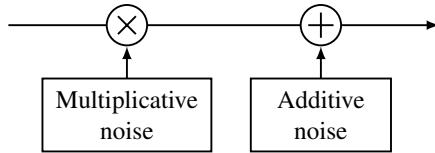


Figure 3.1 Two types of noise in the wireless channel.

The multiplicative noise sources can be found among:

- directional properties of transmitter and receiver antennas
- reflection (ground, walls, and hills)
- absorption (walls, foliage, atmosphere)
- scattering (from rough surfaces such as sea, ground, foliage)
- diffraction (from edges such as buildings and hills)
- refraction (layers of different permittivity)

It is common to further divide the multiplicative processes into three types of fading (see Fig. 3.2): *path loss*, *shadowing* or *slow fading*, and *fast fading* or *multipath fading*. The latter two processes are time-varying and depend heavily on the exact positions of the transmitter and receiver antennas.

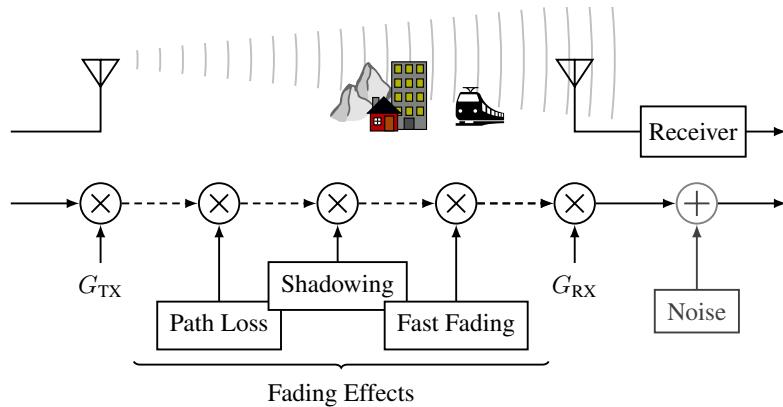


Figure 3.2 Additive and multiplicative noise in the wireless channel.

An example is given by Fig. 3.3. It shows the variation of three multiplicative effects as a mobile receiver moves away from a transmitting base station. The path loss is a deterministic effect, which models the higher attenuation at a larger distance. The shadowing process arises due to different degrees of obstruction in the path, whereas the fast fading process occurs due to constructive and destructive interference between multiple waves reaching the mobile receiver.

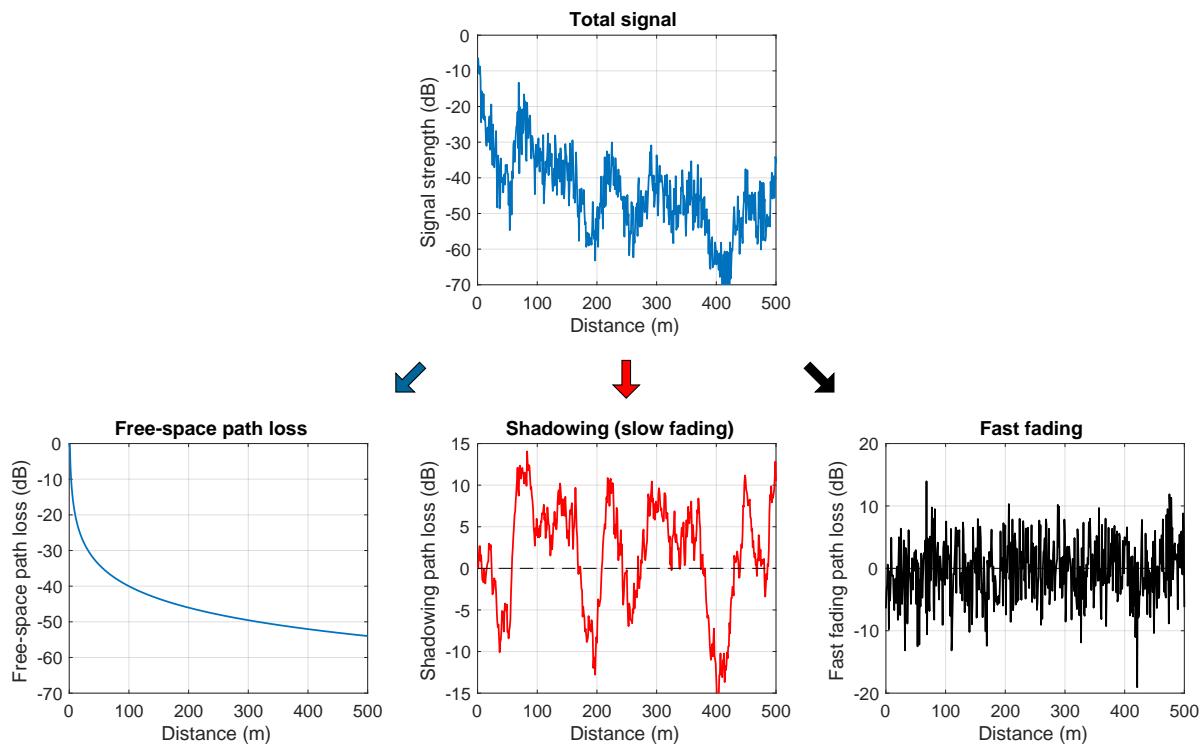


Figure 3.3 Three subtypes of multiplicative noise in the wireless channel. Each signal strength is plotted as a function of distance between transmitter and receiver.

3.2 Mechanisms of Wireless Propagation

In the context of wireless communication, the term *propagation* describes the effects of the advancement of the electromagnetic wave bearing the information to transmit. We have four principle models for wireless propagation to distinguish: ground waves, sky waves, free-space waves, and open-field waves. In order to understand the propagation mechanism we need to have a short glimpse at the layers above the earth surface. They are illustrated in Fig. 3.4.

Ground waves or *surface waves* occur at low frequencies, up to a few MHz, and for vertical polarization (horizontally polarized fields are short-circuited by the earth). Such waves travel over the earth's surface attenuated by the absorption of the conducting earth. Through the curvature of the earth, the waves can make it to points which have no direct line of sight. The higher the frequency, the more the wave gets absorbed by the ground.

Sky waves or *room waves* occur at frequencies between 3 and 30 MHz. They describe the propagation mode by which waves are reflected from the ionosphere, a thin-air layer high above the surface of the earth, a process illustrated in Fig. 3.5(a). The ionosphere is ionized by the sunlight at which point it builds a good reflector for the wavelength indicated above. The reflection quality is therefore heavily dependent on the

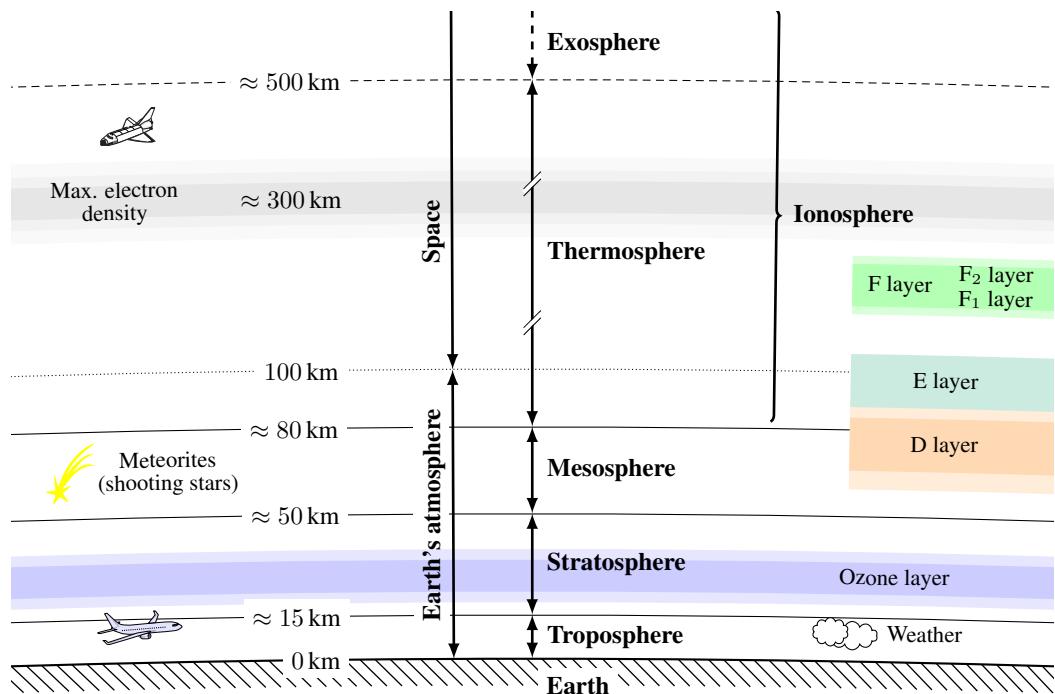
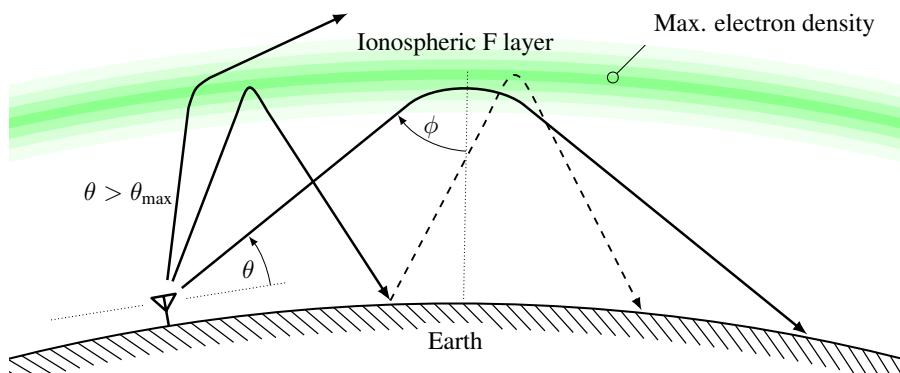
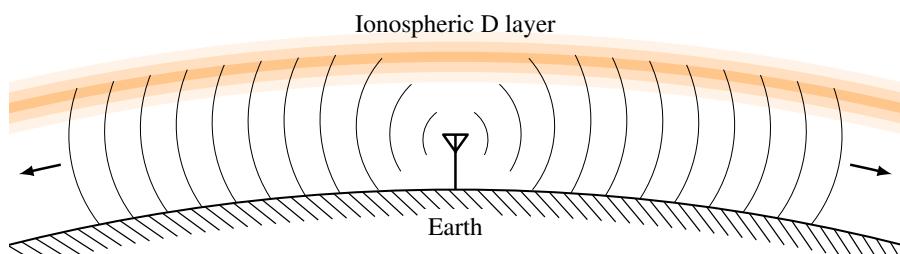


Figure 3.4 Zones around the earth.

time of day, season, longitude on earth, and the sunspot cycle. Using sky waves, transmission virtually around the earth has been reported (amateur radio). The ionosphere may build a waveguide for the signal to propagate according to Fig. 3.5(b).



(a) Sky-wave propagation: reflection on the ionosphere.



(b) Earth-ionosphere waveguide: the ionosphere and ground serve as waveguide walls.

Figure 3.5 Propagation modes making use of the ionosphere.

3.2.1 Free-Space Propagation

The *free-space propagation* is the easiest propagation mechanism to understand. It assumes that no reflections occur. Imagine a setup of two antennas in free space (no obstacles inbetween) as in Fig. 3.6. The

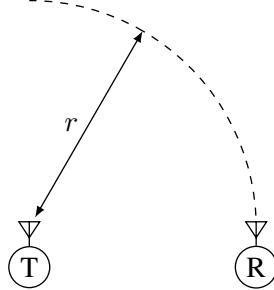


Figure 3.6 Free-space setup with two antennas.

transmitting antenna (marked by 'T') transmits at power P_T . If the directivity was 0 dB, the power would distribute equally over the sphere with radius r , or $S = P_T/(4\pi r^2)$. Due to a larger directivity, we get for the power density at a point r away in the direction of the main lobe

$$S = \frac{P_T G_T}{4\pi r^2}. \quad (3.1)$$

The product $P_T G_T$ is usually called the *equivalent isotropic radiated power* (EIRP), which is essentially the radiated power relative to an isotropic radiator. Sometimes ERP is used, which refers to the power relative to a dipole. The receiving antenna (marked by 'R') now captures the power density over an area A_R , the effective area or aperture. The total received power is therefore

$$P_R = \frac{P_T G_T}{4\pi r^2} A_R. \quad (3.2)$$

Using the area of an antenna

$$A = \frac{\lambda^2 G}{4\pi} \quad (3.3)$$

we get for the power relation in the free-space case

$$\frac{P_R}{P_T} = G_T G_R \left(\frac{\lambda}{4\pi r} \right)^2. \quad (3.4)$$

This is the *Friis transmission formula*. It is valid only if the main lobes of the antennas are adjusted exactly, if the polarizations match, and if both antennas are impedance matched. The path loss is usually given as the inverse of Eq. (3.4) (positive dB numbers), when the antenna gains can be isolated. Thus, the free-space path loss is

$$L = \frac{P_T G_T G_R}{P_R} = \left(\frac{4\pi r}{\lambda} \right)^2 = \left(\frac{4\pi r f}{c} \right)^2. \quad (3.5)$$

In decibels, Eq. (3.5) can be expressed as

$$L = -147.6 \text{ dB} + 20 \log_{10} r + 20 \log_{10} f. \quad (3.6)$$

The free-space path loss increases (gets worse) with increasing frequency because of the smaller capture area of the antenna.

In practice, the free-space model is usually optimistic, because other effects are neglected (e.g., absorption by raindrops etc.). Still, it can serve as a practical minimum of the path loss for a given distance before other, more involved methods are applied.

3.2.2 Open-Field Propagation

A more realistic scenario of propagation is suggested by the open-field model, also called plane-earth model. The model is given by Fig. 3.7. Both transmitting and receiving antennas are situated at a certain height (h_b

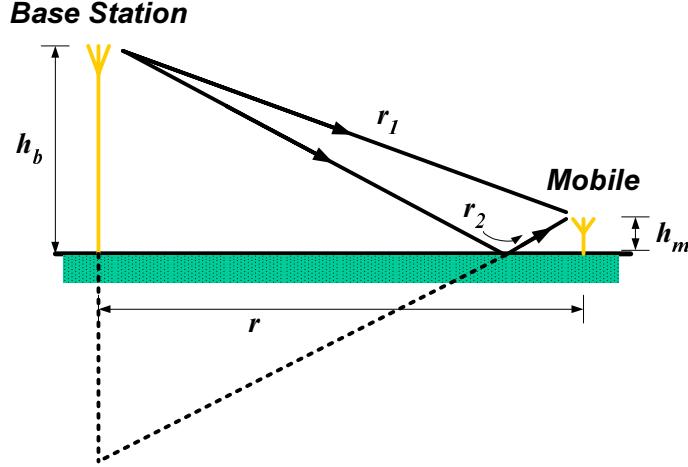


Figure 3.7 Open-field propagation model.

for the base station and h_m for the mobile) above a flat reflecting ground. Propagation now takes place via two paths, a direct path of length

$$r_1 = \sqrt{r^2 + (h_b - h_m)^2}, \quad (3.7)$$

and a reflected path, whose length may be inferred by applying image theory, considering the reflected ray as coming from an image of the transmitter in the ground, as

$$r_2 = \sqrt{r^2 + (h_b + h_m)^2}. \quad (3.8)$$

The interference from the reflection may be constructive, destructive, or part of either. Which one, depends on the path length difference between r_1 and r_2 with respect to the wavelength, hence, we are interested in

$$\begin{aligned} \frac{r_2 - r_1}{\lambda} &= \frac{1}{\lambda} \left(\sqrt{r^2 + (h_b + h_m)^2} - \sqrt{r^2 + (h_b - h_m)^2} \right) \\ &= \frac{r}{\lambda} \left(\sqrt{1 + \left(\frac{h_b + h_m}{r} \right)^2} - \sqrt{1 + \left(\frac{h_b - h_m}{r} \right)^2} \right). \end{aligned} \quad (3.9)$$

For small antenna heights compared to the length r , we can use the first-order approximation

$$(1 + x)^n \approx 1 + nx, \quad x \ll 1. \quad (3.10)$$

Doing so we get

$$\frac{r_2 - r_1}{\lambda} \approx \frac{2h_b h_m}{\lambda r}. \quad (3.11)$$

If we now assume that the direct and reflected waves have approximately the same amplitude (the path lengths are almost the same and the reflection coefficient R is close to minus one, which is strictly true only for low grazing angles $\gamma = \tan^{-1}\left(\frac{h_b+h_m}{d}\right)$), we get for the combined loss the multiplicative factor (in addition to the free-space loss)

$$\left| 1 + R \exp\left(j2\pi \frac{r_2 - r_1}{\lambda}\right) \right| \approx \left| 1 + R \exp\left(j \frac{4\pi h_b h_m}{\lambda r}\right) \right| \approx \left| 1 - \exp\left(j \frac{4\pi h_b h_m}{\lambda r}\right) \right|. \quad (3.12)$$

The minus sign for the reflection coefficient R in the above equations stems from the fact that polarization flips at the reflection (at least for incident angles close to grazing). We get, together with the free-space path loss, the open-field path loss as

$$L = \left(\frac{4\pi r}{\lambda} \right)^2 \left| 1 - \exp \left(j \frac{4\pi h_b h_m}{\lambda r} \right) \right|^{-2} = 4 \left(\frac{\pi r}{\lambda} \right)^2 \sin^{-2} \left(\frac{2\pi h_b h_m}{\lambda r} \right). \quad (3.13)$$

By observing the open-field path in Fig. 3.8, we see that the loss function for the open field has many nulls

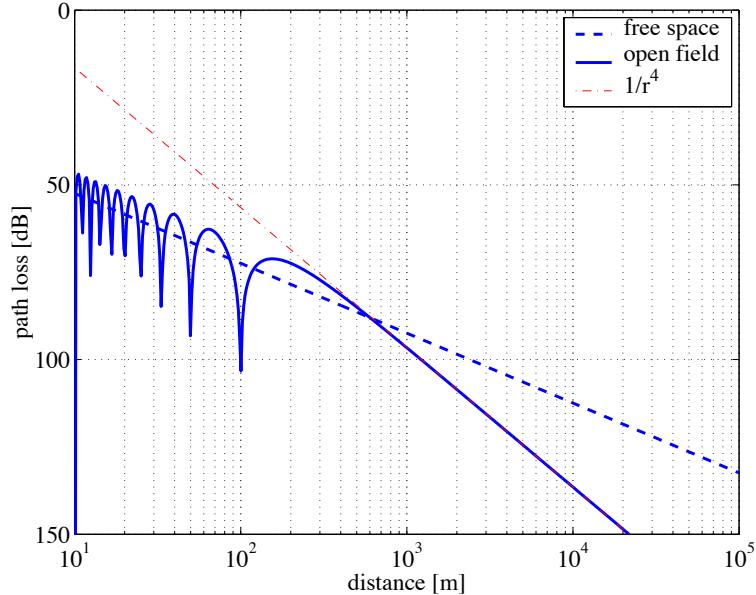


Figure 3.8 Open-field path loss (solid line). In this case $h_b = 20$ m, $h_m = 1.5$ m, and $f = 1$ GHz.

up to a certain point. After that, it decreases much faster than the $20 \log_{10}(.)$ we know from the free-space case. In fact, since exponential functions can be approximated using

$$\exp(x) \approx 1 + x, \quad x \ll 1, \quad (3.14)$$

we get for large distances an open-field path loss of

$$L = \left(\frac{4\pi r}{\lambda} \right)^2 \left(\frac{\lambda r}{4\pi h_b h_m} \right)^2 = \frac{r^4}{h_b^2 h_m^2}. \quad (3.15)$$

As opposed to the free-space loss, which grows with r^2 , the open-field loss grows with r^4 . Furthermore, the loss is now independent of the operating frequency. The point where the curve starts going down with r^4 is roughly at the distance of the last local maximum,

$$r = \frac{4h_b h_m}{\lambda}. \quad (3.16)$$

We have seen that the path-loss exponent in the free-space model was 2 (in logarithmic form $20 \log_{10}(r)$) and in the open-field model 4 (or $40 \log_{10}(r)$). Other, empirically found path-loss exponents are in use, see for example Table 3.1. If we now want to estimate the power P_2 received at a certain point in a certain distance d_2 from the knowledge of the power P_1 at a different distance d_1 , we can write down the relation

$$\frac{P_2}{P_1} = \left(\frac{d_1}{d_2} \right)^n. \quad (3.17)$$

Environment	Path-loss exponent n
Free space	2
Open field (long distance)	4
Cellular radio, urban area	2.7 – 4
Shadowed urban cellular radio	5 – 6
In building, line-of-sight	1.6 – 1.8
In building, obstructed	4 – 6

Table 3.1 Path-loss exponents for different environments. Source: [12, p. 362].

3.2.3 Diffraction

Besides reflexion, which we have met in Section 3.2.2, and absorption, which does not contribute to the models used here, there are other mechanisms of propagation, which we shall shortly describe. Diffraction (*dt. Beugung*) is the explanation to the question of how objects that are not in the direct line of sight can be reached by electromagnetic waves. Shadows are never completely sharp, some energy does always propagate into the shadow region.

We can understand the reason by looking at Huygen's principle¹: First, at a given instant in time, each element of a wavefront can be regarded as the center of a secondary disturbance, giving rise to spherical wavelets. Second, the position of the wavefront at any later time is the envelope of all such wavelets. We can apply this principle to our problem of a plane wavefront hitting a wall-top (so called *knife-edge diffraction*), which is illustrated in Fig. 3.9. Wavefronts impinging on the top of the absorbing screen become curved by the edge in such a way, that waves seem to emerge from points close to the edge, filling the shadow region with diffracted wavelets.

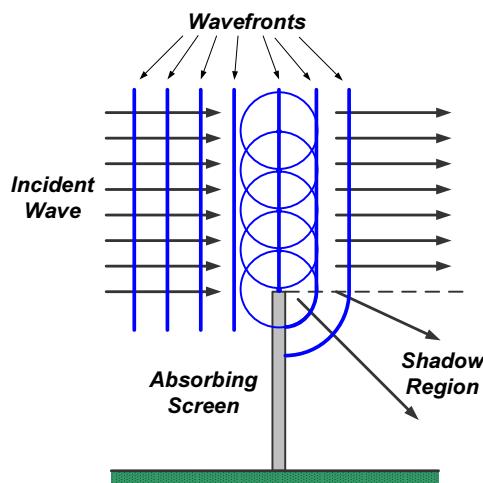


Figure 3.9 Huygen's principle for knife-edge diffraction.

If Huygen's principle is applied in a quantitative way, taking into account the respective amplitudes and phases of an infinite number of secondary sources above the edge, we get the propagation loss inside the shadow region

$$L = 20 \log_{10} \left| \frac{E_i}{E_d} \right| = -10 \log_{10} |F(v)|, \quad (3.18)$$

¹named after the Dutch mathematician and physicist Christiaan Huygens (1629–1695)

where E_i and E_d are the impinging and diffracted fields, respectively, and

$$F(v) = \frac{1+j}{2} \int_v^\infty \exp\left(-j\frac{\pi t^2}{2}\right) dt \quad (3.19)$$

is a function of the Fresnel integrals, see for example [30], and the diffraction parameter given by

$$v = h \sqrt{\frac{2(d_1 + d_2)}{\lambda d_1 d_2}} \quad (3.20)$$

with d_1, d_2 being the respective distances of the transmitter and receiver to the knife edge and h the excess height of the knife edge. In fact, the distances d_1 and d_2 and the height h refer to a coordinate system in which both transmitter and receiver sit on a baseline. Any situation can be transformed into such a setup using a simple coordinate transformation, provided the distances d_1 and d_2 are large compared to h , which is usually the case.

Let us consider three practical situations: when there is a line-of-sight (LOS), meaning there is clearance between line connecting the two antennas and the knife edge case, Fig. 3.10(a), h is negative and diffraction

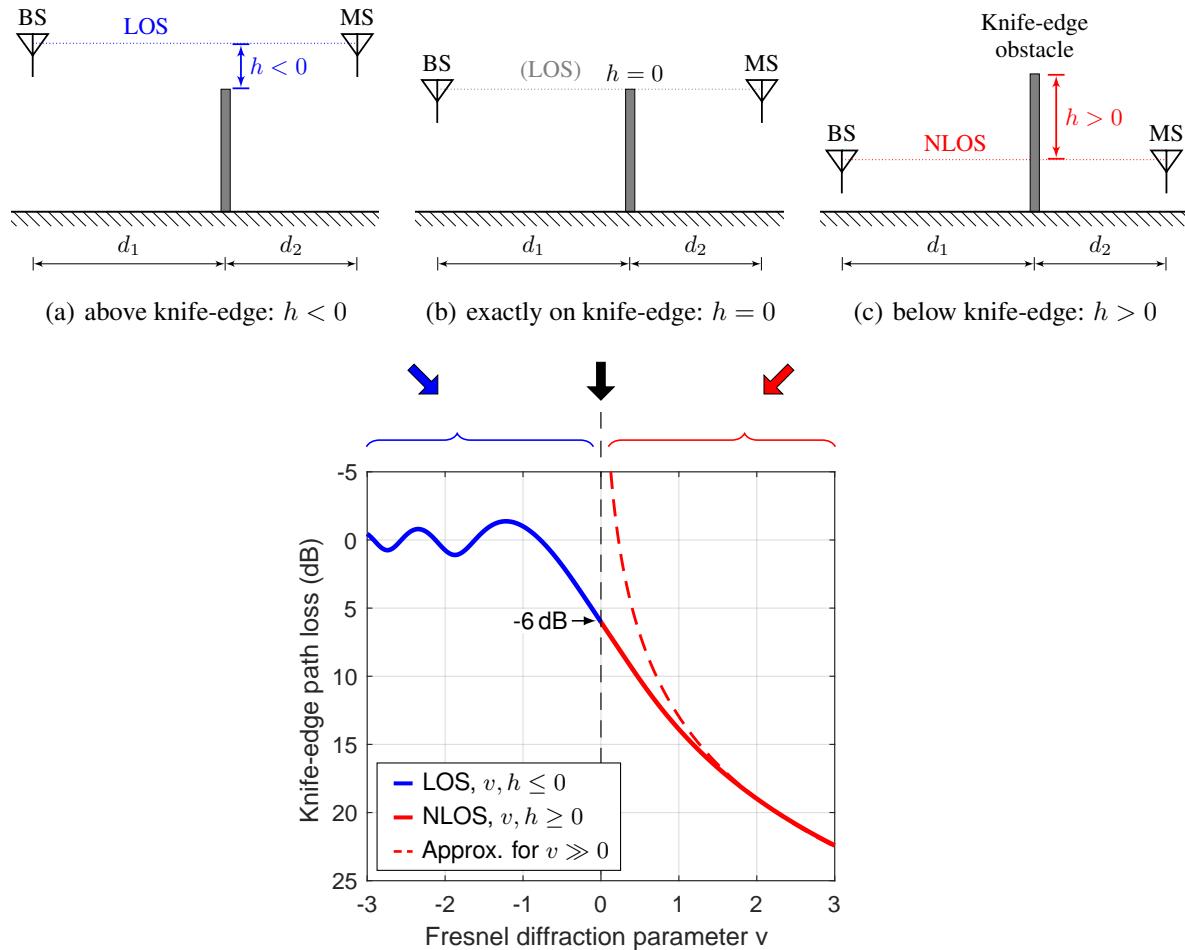


Figure 3.10 Examples of knife-edge propagation due to partial obstruction by a wall and the involved diffraction loss, calculated using exact Fresnel integrals, according to Eq. (3.19), and the approximation for large v Eq. (3.21).

is usually negligible. At the point where the obstacle covers the space up to exactly the LOS line (i.e. the excess height is zero), illustrated in Fig. 3.10(b), a diffraction loss of 6 dB occurs. Lastly, when the obstacle obstructs the LOS line (leading to a non-line-of-sight condition, NLOS), h is positive and increased diffraction loss occurs, as shown in Fig. 3.10(c).

Note that this loss occurs in addition to the free-space loss² for a given distance. For values of $v > 1$, which refers to points well inside the shadow region, Eq. (3.18) may be approximated with accuracy better than 1 dB by

$$L \approx 20 \log_{10}(\sqrt{2}\pi v) \approx 20 \log_{10} \frac{v}{0.225} \quad v > 1, \quad (3.21)$$

or directly as

$$L \approx 20 \log_{10} v + 13 \text{ dB}. \quad (3.22)$$

It has to be emphasized that for all other cases no closed formula can be given and it is best to infer the corresponding loss from Fig. 3.10. For negative values v (negative excess height as seen in Fig. 3.10(a)), which means there is a direct line of sight, there can still be some loss.

To illustrate the dependence of the diffraction loss by the wavelength, Fig. 3.11 shows the result with λ as a parameter. Clearly, when going towards optics, the 'illumination' of shaded areas gets more difficult as would please our intuition.

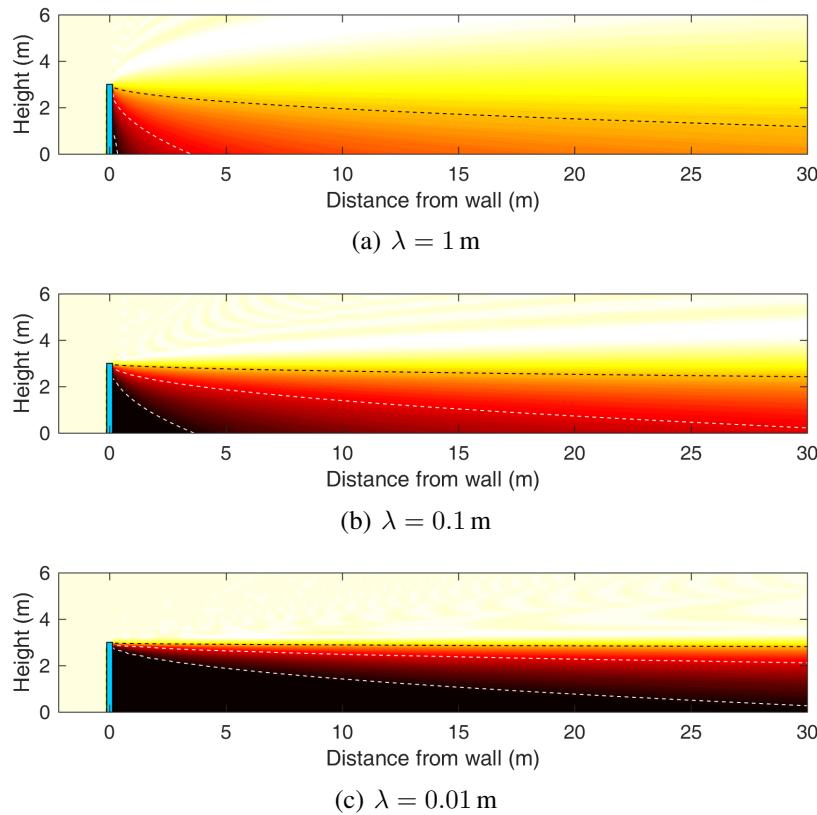


Figure 3.11 Diffraction loss due to a wall (3 m tall) for different wavelengths.

²The diffraction edge prevents any reflections in this case, so that the free-space model is appropriate.

3.2.4 Fresnel Zones

Sometimes, the expression *Fresnel zone* is used to indicate the proximity of obstacles to the direct line of sight. To visualize Fresnel zones, we need to derive ellipses that have the transmitter and the receiver in their focus points. The first Fresnel zone is of special interest and has the following derivation: The ellipse curvature consists of the points through which the path is longer by $\lambda/2$ compared to the direct line. Hence, we can construct the ellipse by choosing the longer and shorter axes as a and b and the distance between transmitter and receiver as $2c$, as seen in Fig. 3.12. Now the problem of computing the ellipse axes of the

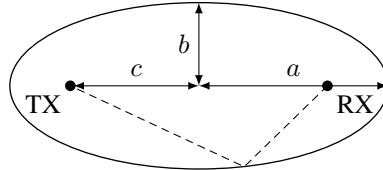


Figure 3.12 Construction of the first Fresnel zone.

first Fresnel zone is working out a and b , given the distance $d = 2c$ and the wavelength λ . For an ellipse, every path via a point on the curve is of the same length, hence also the one via the end points of the ellipse

$$l = c + a + (a - c) = 2a. \quad (3.23)$$

The direct path is $2c$. Thus,

$$2a = 2c + \frac{\lambda}{2} \quad (3.24)$$

or

$$a = c + \frac{\lambda}{4} = \frac{d}{2} + \frac{\lambda}{4}. \quad (3.25)$$

The ellipse equation (linear eccentricity) $c^2 = a^2 - b^2$ delivers

$$b = \sqrt{\frac{c\lambda}{2} + \frac{\lambda^2}{16}}. \quad (3.26)$$

For small wavelengths $\lambda \ll c$, we have

$$b \approx \sqrt{\frac{c\lambda}{2}}. \quad (3.27)$$

We have now derived the major and minor axes of the first Fresnel ellipse. In a wireless transmission system, one usually tries to keep the first Fresnel zone free of obstacles. More generally, the major and minor axes of the n -th order ellipse can be derived as

$$a = \frac{d}{2} + \frac{n\lambda}{4}, \quad (3.28)$$

$$b = \sqrt{\frac{cn\lambda}{2}}. \quad (3.29)$$

Since usually $\lambda \ll d$, we can write Eq. (3.28) as

$$a \approx c = \frac{d}{2}. \quad (3.30)$$

The minor axis b is the clearance height in the middle of the transmitter and receiver. In the following, we want to derive the clearance height at every position between transmitter and receiver, such that the distance

to a potential obstacle is d_1 and d_2 and not necessarily $d/2$, see Fig. 3.13(a). We start by stating that for every pair (x, y) the ellipse equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (3.31)$$

must be satisfied. Solving Eq. (3.31) for y yields

$$y = b \sqrt{\frac{a^2 - x^2}{a^2}}. \quad (3.32)$$

With $d_1 = c + x$ and $d_2 = c - x$ we get

$$d_1 d_2 = c^2 - x^2 = a^2 - b^2 - x^2. \quad (3.33)$$

Using the fact that $b \ll d$ and Eqs. (3.29)-(3.30) in Eq. (3.32), we get

$$y = b \sqrt{\frac{d_1 d_2}{c^2}} = \sqrt{\frac{cn\lambda}{2}} \sqrt{\frac{d_1 d_2}{c^2}} = \sqrt{\frac{n\lambda d_1 d_2}{2c}}. \quad (3.34)$$

Now, c is of course the mean of the two distances, $c = \frac{d_1 + d_2}{2}$, hence

$$y = \sqrt{\frac{n\lambda d_1 d_2}{d_1 + d_2}}. \quad (3.35)$$

Due to the symmetry around the line TX–RX we can speak of the Fresnel radius rather than y . Thus, the n th Fresnel zone is given in approximation by the radius

$$r_n \approx \sqrt{\frac{n\lambda d_1 d_2}{d_1 + d_2}}, \quad (3.36)$$

for which the ellipsoids as shown in Fig. 3.13(a) result. Usually, one tries to avoid the placement of obstacles

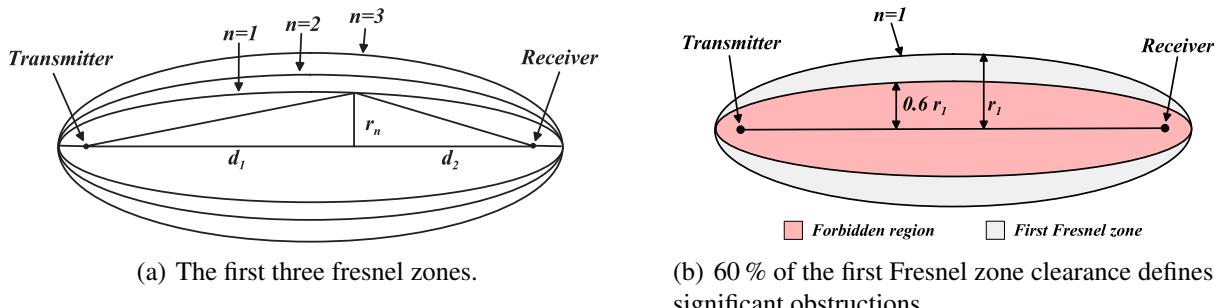


Figure 3.13 Fresnel zones.

in the first Fresnel zone. By expressing the diffraction parameter as a function of the first Fresnel zone, we realize that

$$v = \frac{h}{r_1} \sqrt{2}. \quad (3.37)$$

Now, as long as an obstruction is outside the circle whose radius is given by 60 % of the first Fresnel zone, the diffraction parameter is around $v = -0.85$. Referring to Fig. 3.10, the diffraction loss is then 0 dB. This clearance is often used as the criterion to decide whether an object is a significant obstruction (see also Fig. 3.13(b)). With the exception of the first Fresnel zone, all zones cancel each other in their influence (pairwise). The reception thus stays approximately constant if all but the first zone are shaded.

3.2.5 Scattering

In Section 3.2.2, we have assumed total reflection, so-called *specular reflection*. Frequently, however, when surfaces are not smooth, an impinging wave does not get reflected in one exact direction, but gets scattered, see Fig. 3.14.

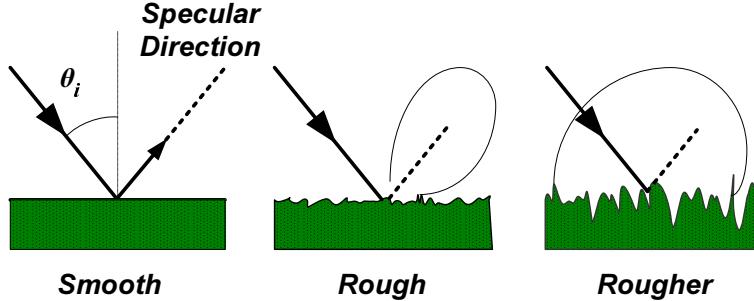


Figure 3.14 Scattering due to surface roughness.

One possible technique to model scattering (*dt. Streuung*) is due to Beckmann and Spizzichino [2], who multiply the reflection coefficient R in Eq. (3.12) by a roughness factor f , which depends on the angle of incidence θ and on the standard deviation of the surface height σ_s . This roughness factor is given as

$$f = \exp\left(-\frac{1}{2}\left(\frac{4\pi\sigma_s \cos \theta}{\lambda}\right)^2\right), \quad (3.38)$$

which results in the plot as provided in Fig. 3.15.

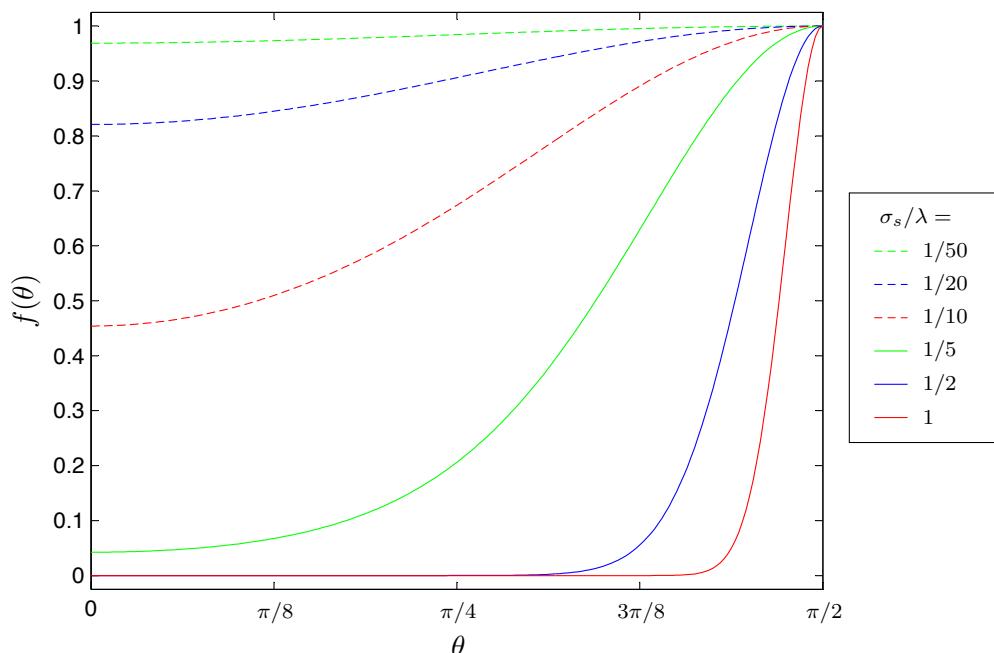


Figure 3.15 Roughness factor.

3.2.6 Further Losses

In addition to the loss factor treated so far, we find further attenuation in the wave propagation due to absorption caused by interaction with gas molecules, most importantly oxygen (O_2) and water vapor (H_2O), as shown in Fig. 3.16. It is well-known that the water molecule possesses an electric dipole moment, which interacts with the electric fields. Lesser known, but equally important, is the fact that the oxygen molecule possesses a permanent magnetic dipole moment associated to aligned electron spins of its constituent atoms. Hence, oxygen molecules interact with magnetic fields. Interestingly, the molecule of nitrogen (N_2), the most abundant gas in earth's atmosphere, neither has a significant electric nor magnetic dipole moment, and, thus, does not interact with electromagnetic waves in the microwave spectrum at all.

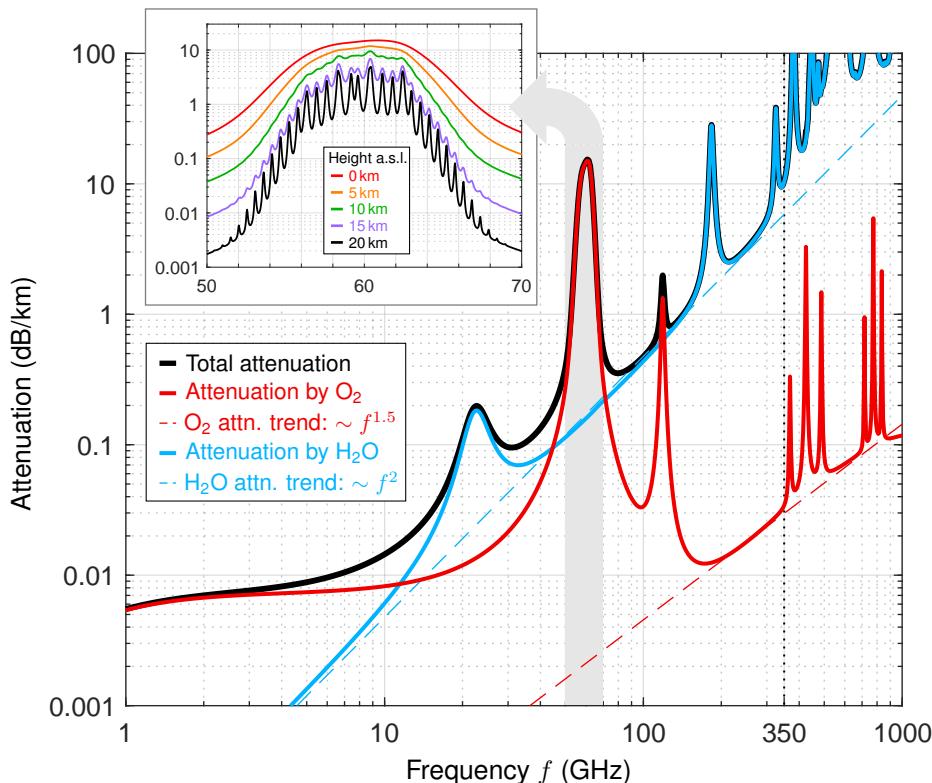


Figure 3.16 Attenuation due to molecular absorption of oxygen and water vapor (at sea level) in the microwave spectrum.

As can be seen by the graphs in Fig. 3.16, the first local maximum of the absorption at microwave frequencies is at 22.2 GHz, due to water. The second local maximum of the absorption is due to a family of 26 resonance lines of the oxygen molecule around 60 GHz; the separate lines only become visible at very high altitudes (above sea level, a.s.l.). At such low gas pressure levels, the interaction between the molecules is reduced considerably, such that the contributions of all energy transitions become separable. The third resonance is due to oxygen again, this time a single sharp resonance line around 119 GHz. After this frequency, water has two more resonances around 183 and 321 GHz, respectively, before both oxygen and water absorption becomes difficult to predict above 350 GHz.

Note that the actual attenuation depends on the amount of water in the atmosphere. Moreover, if undissolved water is in the air (rain), this attenuation has to be added to the free space loss and the two attenuation values

due to water vapor and oxygen. Table 3.2 lists the attenuation due to rain in more detail for some practical frequencies.

Rain type	Intensity	450 MHz	1 GHz	3 GHz	10 GHz
Mizzle (drizzle)	0.25 mm/h	$2.2 \cdot 10^{-8}$	$1.5 \cdot 10^{-6}$	$1.5 \cdot 10^{-4}$	0.02
Light rain	5 mm/h	$1.0 \cdot 10^{-6}$	$2.0 \cdot 10^{-5}$	$1.0 \cdot 10^{-3}$	0.08
Medium rain	12.5 mm/h	$3.0 \cdot 10^{-6}$	$7.0 \cdot 10^{-5}$	$3.0 \cdot 10^{-3}$	0.28
Heavy rain	25 mm/h	$7.5 \cdot 10^{-6}$	$1.5 \cdot 10^{-4}$	$1.0 \cdot 10^{-2}$	0.6
Shower	50 mm/h	$1.0 \cdot 10^{-5}$	$3.0 \cdot 10^{-4}$	$2.0 \cdot 10^{-2}$	1.5

Table 3.2 Attenuation (in dB/km) due to rain. Source: [23].

3.3 Empirical Path-Loss Models

Very often it might not be practical to determine the loss introduced by the wireless channel by the mechanisms covered so far, partly because the environment is not known exactly or because the mathematics get untractable. In these cases it is useful to have empirical path-loss models. The word *empirical* relates to the fact that such models have been extracted using actual measurements without considering the physical justification.

One of the first introduced such model was the one by Egli [8] in his publication 'Radio propagation above 40 MC over irregular terrain'³ published in the Proceedings of the IRE⁴ Egli realised that the open-field model shows a path-loss exponent of four and a dependence on the antenna heights. Additionally, he argued that a dependence on the frequency of f^2 should be reintroduced like in the free-space case. Because in his reasoning there was partly physical justification involved, this approach would also become known as a semi-empirical method. Egli's results of fitting a model with measurements taken around American cities were at the time provided in nomograms, which Delisle *et al.* [6] later translated into an equation (loss in dB) of the form

$$L = 40 \log_{10} R + 20 \log_{10} f - 20 \log_{10} h_b + \begin{cases} 76.3 - 10 \log_{10} h_m & \text{for } h_m < 10 \text{ m} \\ 85.9 - 20 \log_{10} h_m & \text{for } h_m > 10 \text{ m} \end{cases}, \quad (3.39)$$

with a small discontinuity at $h_m = 10$ m. Note that in Eq. (3.39) the unit of R is km and the unit of f is MHz.

The most widely cited model is the Okumura-Hata model, which is a fully empirical model. It is based on measurements taken by Okumura in Tokyo [26] and equations fitted by Hata some years later [17]. By restricting ourselves to frequencies above 400 MHz, a base-station antenna height of 30 to 200 m, a vehicular antenna height below 10 m and a distance R_{km} between 1 and 20 km, which is no real restriction for most modern wireless systems, we can write these equations (loss in dB) as

$$L_p = 69.55 + 26.16 \log_{10} f_{\text{MHz}} - 13.82 \log_{10} h_b + (44.9 - 6.55 \log_{10} h_b) \log_{10} R_{\text{km}} - a(h_m), \quad (3.40)$$

where the correction factor for the mobile antenna height is given as

$$a(h_m) = \begin{cases} 3.2 \cdot (\log_{10} 11.75 h_m)^2 - 4.97 & \text{for large cities} \\ (1.1 \log_{10} f_{\text{MHz}} - 0.7) h_m - (1.56 \log_{10} f_{\text{MHz}} - 0.8) & \text{for medium-small cities.} \end{cases} \quad (3.41)$$

³Note that MC was once used as the abbreviation for Megacycles per second which today is MHz of course.

⁴The IRE (Institute of Radio Engineers) is today known as the IEEE.

For areas outside cities, Hata gives the loss in dB as

$$L = L_p + \begin{cases} -2 \cdot (\log_{10}(f_{\text{MHz}}/28))^2 - 5.4 & \text{for suburban areas} \\ -4.78 \cdot (\log_{10} f_{\text{MHz}})^2 + 18.33 \log_{10} f_{\text{MHz}} - 40.94 & \text{for open areas.} \end{cases} \quad (3.42)$$

Note that in Eqs. (3.40) to (3.42) the units of R_{km} is km and the unit of f_{MHz} is MHz. This equation is the foundation of many commercial network-planning software. Other empiric models are known such as the Lee model, the Ibrahim and Parsons model, and still others. A comparison of different path-loss models is shown in Fig. 3.17, where it can be seen that the deterministic models usually result in too optimistic values as compared to results based on real measurements (empiric models).

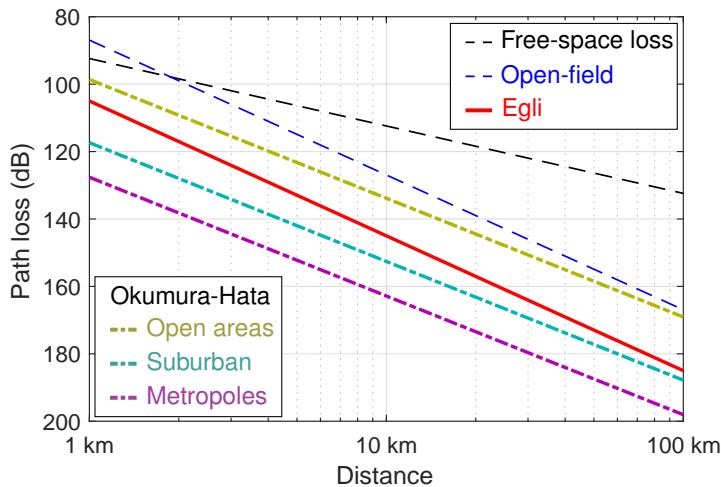


Figure 3.17 Comparison of different deterministic and empiric path-loss models (for $f = 1 \text{ GHz}$, $h_b = 30 \text{ m}$, $h_m = 1.5 \text{ m}$).

For indoor environments, there are empirical models as well. One such model is the COST231 Multi-Wall. It consists of an approximation of the loss in dB by

$$L = L_{\text{free space}} + A + B \cdot n_{w_1} + C \cdot n_{w_2} + D \cdot n^{(\frac{n+2}{n+1}-E)}, \quad (3.43)$$

where n is the number of traversed floors not thicker than 30 cm, n_{w_1} the number of light internal walls, and n_{w_2} the number of concrete/brick internal walls, respectively. The terms for A to E are frequency dependent.

3.4 Link Budget

A task frequently performed by a wireless engineer is the calculation of the link budget of a system. The question is usually one of the following:

- Given all parameters of the transmitter and the operating distance, what sensitivity of the receiver is needed (noise figure)?
- Given the operating distance and the receiver specification what is the minimum transmission power required?
- Given all parameters of the transmitter and receiver, what is the maximum operating range of the system?

The link budget is simply the summation of the dB values of all losses and gains acquired on the way from the transmitter to the receiver.

$$P = N + B + \text{NF} + \text{SNR} + L - G_T - G_R \quad (3.44)$$

with

- N_0 noise power density (i.e. normally $N_0 = -174 \text{ dBm/Hz}$),
- B bandwidth used,
- NF noise figure of the receiver,
- SNR required SNR at the receiver end,
- L total loss in the transmission (e.g. path loss, cable loss, etc.),
- G_T, G_R gains of the transmitter and receiver antenna.

Example: For a system operating at 1 GHz, what is the required transmission power to cover 10 km of range (free-space model assumed) if the receiver needs at least an SNR of 20 dB in a 25 kHz bandwidth with a total receiver chain noise figure of 5 dB, and both transmitter and receiver use half-wave dipoles?

Solution: The gain/loss components are listed in Table 3.3. All we have to do now is add up the figures (with the right signs):

$$P = N_0 + B + \text{NF} + \text{SNR} + L - G_T - G_R = 3.2 \text{ dBm} = 2.1 \text{ mW}.$$

Component	Loss/gain
Noise power density	$N_0 = -174 \text{ dBm}$ (per Hz)
Bandwidth	$B = 10 \log_{10} 25000 = 44 \text{ dB}$
Noise figure	$\text{NF} = 5 \text{ dB}$
Signal-to-noise ratio	$\text{SNR} = 20 \text{ dB}$
Path loss	$L = -10 \log_{10} \left(\frac{\lambda}{4\pi r} \right)^2 = -20 \log_{10} \frac{c}{4\pi r f} = 112.4 \text{ dB}$
Receiver antenna gain	$G_R = 2.1 \text{ dBi}$
Transmitter antenna gain	$G_T = 2.1 \text{ dBi}$

Table 3.3 Gain/loss of each component for link budget example.

3.5 Fading Channels

Until now, we have covered the first of the multiplicative channel effects only, namely path loss, which is a static behaviour. In our analysis so far we have assumed at most two paths, a direct one and a reflected one. In reality, however, there is usually a multitude of paths, making the problem hard to determine. Also, very often the exact surroundings may not be known. The approach sought is often statistical modeling.

3.5.1 Shadowing

A complicated environment with several objects between transmitter and receiver make the outcome of a path-loss analysis highly dependent on the exact location of transmitter and receiver. Placing the receiver

slightly to one side might produce a high difference in path loss, not predicted by our former models. This phenomenon is called *slow fading* or *shadowing*. A typical variation of path loss can be seen in Fig. 3.18.

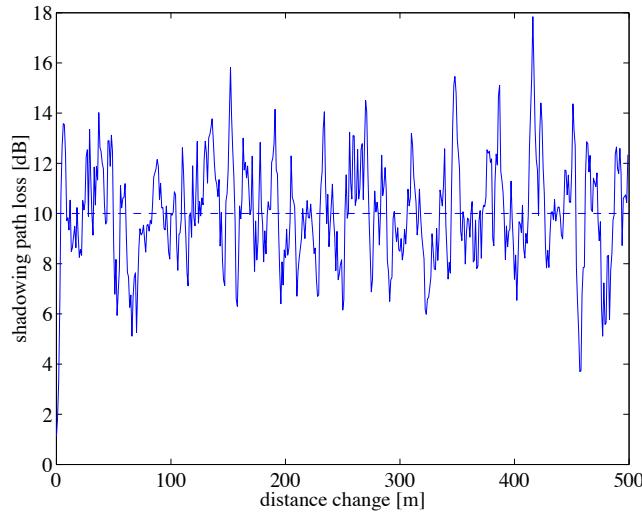


Figure 3.18 Typical shadowing path loss due to location variation of mobile receiver. (Here: median path loss = 10 dB (dashed line); path loss standard deviation (location variability) $\sigma_L = 5 \text{ dB}.$)

The probability distribution of the underlying signal powers is log-normal, i.e., the loss expressed in dB has a normal distribution. The variation of the path loss occurs over distances comparable to object sizes, i.e., widths of buildings and hills and is called *location variability* σ_L . The reason for the log-normal distribution can easily be seen by considering the composition of the path loss by several independent effects on the way. Thus, the loss can be written as

$$L = L_1 \cdot L_2 \cdot L_3 \cdots L_N \quad (3.45)$$

in linear form, or in dB form as

$$L [\text{dB}] = L_1 [\text{dB}] + L_2 [\text{dB}] + L_3 [\text{dB}] + \dots + L_N [\text{dB}]. \quad (3.46)$$

If all of the N contributions are independent random variables, the central limit theorem dictates that $L [\text{dB}]$ is a Gaussian random variable.

Shadowing has some impact on the link budget. If only the median path loss is used to evaluate a certain range covered, the real path loss exceeds the maximal value for reliable communications for 50 % of the cases. Hence, a certain *fade margin* has to be included in the link budget to make sure that most (you can never be 100 % sure but you can get very close to it) path losses incurred lead to successful communications. Thus, the fade margin to be included depends on the location variability and the percentage in desired successful communications.

3.5.2 Fast Fading

An even more dramatic effect than shadowing is *fast fading*, since it occurs on a much smaller time scale or with a much shorter correlation over location change. Since the relative bandwidth of a signal is usually very small, we analyze all fading effects in their complex baseband representation. The fading effects can then be conveniently modeled by complex phasors, indicating that fading changes amplitude and phase rapidly over time.

Flat Fading With *flat fading*, we mean that there is only one path to consider (or at least several paths come in at such a short delay that they can be treated as one). The path loss parameter can therefore be accurately modeled by one time-varying complex scalar (amplitude and phase). Since such a multiplication influences the whole signal spectrum in the same way, flat fading is often also called *frequency non-selective fading*. As with shadowing, the path loss parameter has a statistical distribution. The kind of distribution depends on which one of two main cases we find: First, there is no line-of-sight path and the signal is the composition of a large number of random reflections, or, second, the random reflections are superimposed by a line-of-sight path.

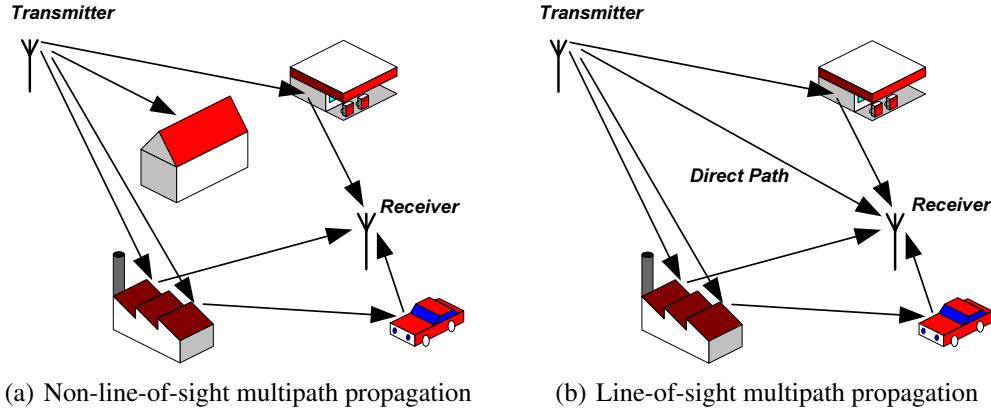


Figure 3.19 The two types of multipath propagation.

Let us consider the first case, which is illustrated by Fig. 3.19(a). Note that as opposed to shadowing, where the composition of one path consisted of several multiplicative effects, we have now several paths adding up to a fluctuating, complex path-loss factor. Since the real part and the imaginary part of the complex path-loss factor (remember we model everything in the baseband) are independent processes, each of the part is a zero-mean Gaussian random variable. It can be shown that the corresponding distribution of the phase, essentially the argument of real and imaginary part, is uniformly distributed over $[0, 2\pi]$ and the amplitude r has a Rayleigh distribution given by its probability-density function (pdf)

$$p(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}. \quad (3.47)$$

The pdf of the Rayleigh distribution is depicted by Fig. 3.20. Such a fading process is therefore termed Rayleigh fading. It concerns the loss factor L in the linear domain (not the dB domain).

In the second case, which is illustrated by Fig. 3.19(b), we have a strong line-of-sight signal in addition to the scatterer. The distribution of the amplitude is a Ricean distribution given as

$$p(r) = \frac{r}{\sigma^2} e^{-(r^2+s^2)/2\sigma^2} I_0\left(\frac{rs}{\sigma^2}\right), \quad (3.48)$$

where $s^2 = m_1^2 + m_2^2$ is the noncentrality parameter consisting of the I - and Q -contribution of the signal present and I_0 is the modified Bessel function of the first kind and degree zero. The pdf of the Rice distribution is also depicted by Fig. 3.20. As opposed to the Rayleigh distribution, the Rice distribution looks much more symmetrical, due to the influence of the direct path. Still, no negative values are possible, so it is slightly asymmetrical. As shown, for large values of s , the Rice distribution can be accurately modeled by a biased Gaussian distribution.

The technique to evaluate the path loss due to fast fading is for either case the same as for shadowing: Specify the needed percentage (probability of successful communication) and evaluate the minimum amplitude via

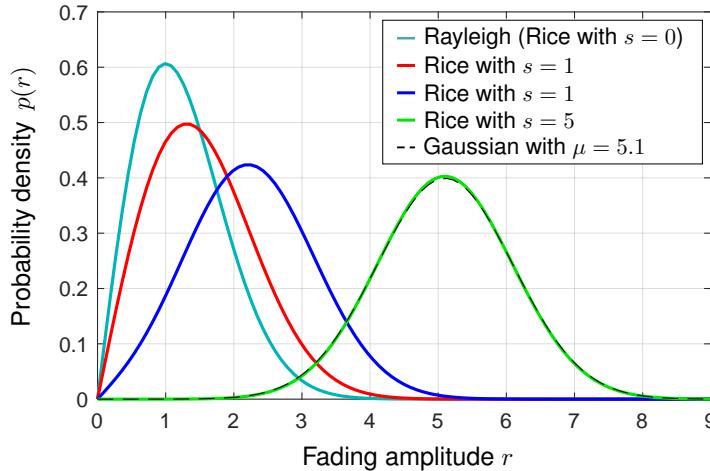


Figure 3.20 Probability density functions of the Rayleigh and the Rice distributions. Both distributions are normalized to $\sigma = 1$.

the cumulative density function (or the tail function). This amplitude then directly converts into a fast fading power loss (or fade margin) to put into the link budget.

One direct application of the Rayleigh distribution given by Eq. (3.47) of a flat fading process is of direct interest to the communications engineer: the BER curve of a BPSK signal in a flat Rayleigh-fading channel. It can be shown that the BER of a BPSK signal is of the form

$$P_E(E_s/N_0) = P_E(\gamma) = Q\left(\sqrt{\frac{2E_s}{N_0}}\right). \quad (3.49)$$

and is clearly a function of the SNR as given by $\gamma = \text{SNR} = E_s/N_0$. The underlying assumption at the point of derivation was that the symbol energy was constant, leading to a constant γ . Now when there is fading, this is no longer true. It can be shown that if the amplitude is Rayleigh distributed, the SNR has a chi-square distribution with two degrees of freedom, which is an exponential distribution given by

$$p(\gamma) = \frac{1}{\Gamma} e^{-\frac{\gamma}{\Gamma}}, \quad (3.50)$$

where Γ is the average SNR of the fading process. We now have to evaluate a tail function, whose argument is exponentially distributed. The BER is then given as the expectation of the different BERs for varying SNR according to their distribution given by Eq. (3.50). Hence,

$$\begin{aligned} P_E &= E\{P_E(\gamma)\} \\ &= \int_0^\infty \frac{1}{\Gamma} e^{-\frac{\gamma}{\Gamma}} Q(\sqrt{2\gamma}) d\gamma \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2\gamma}}^\infty e^{-\frac{x^2}{2}} dx \frac{1}{\Gamma} e^{-\frac{\gamma}{\Gamma}} d\gamma \\ &= \frac{1}{\sqrt{2\pi}\Gamma} \int_0^\infty \int_0^{\frac{x^2}{2}} e^{-\frac{\gamma}{\Gamma}} d\gamma e^{-\frac{x^2}{2}} dx. \end{aligned} \quad (3.51)$$

The last equality is obtained by changing the integration order. The inner integral can now easily be com-

puted. Hence,

$$\begin{aligned}
P_E &= \frac{1}{\sqrt{2\pi}\Gamma} \int_0^\infty \Gamma \left(1 - e^{-\frac{x^2}{2\Gamma}}\right) e^{-\frac{x^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2}} - e^{-\frac{x^2}{2}(1+\frac{1}{\Gamma})} dx \\
&= \frac{1}{2} - \frac{1}{2} \sqrt{\left(1 + \frac{1}{\Gamma}\right)^{-1}} \\
&= \frac{1}{2} \left(1 - \sqrt{\frac{\Gamma}{1 + \Gamma}}\right). \tag{3.52}
\end{aligned}$$

The result can be visualized in Fig. 3.21. The BER of a flat Rayleigh fading channel does not improve as fast as in the AWGN case when increasing the SNR.

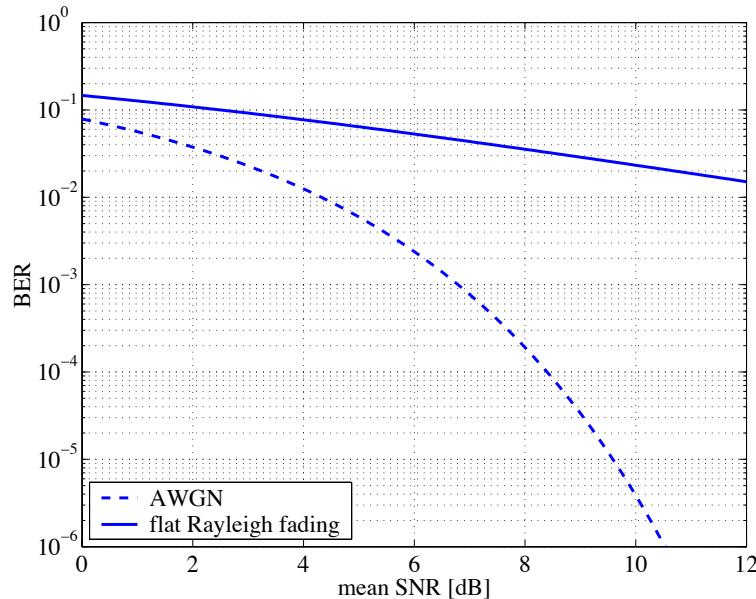


Figure 3.21 Bit-error rates of a BPSK signal in the AWGN case and in a flat Rayleigh-fading channel, respectively.

The above modeling of a channel using Rayleigh and Rice distributions is in fact only half the truth. In addition to this, second-order statistics (essentially how fast the channel changes) have to be incorporated to make a channel model more realistic. The mechanism behind the fact that the fading process is correlated rather than white can be characterized by the *coherence time*, the time during which a channel is considered constant, which is inversely proportional to the *Doppler spread*, a channel characteristics that indicates how wide the frequency occupation of a signal is spread due to a moving transmitter, receiver, or both.

Frequency-Selective Fading As the delays between the individual paths of a multipath signal get larger with respect to the symbol duration, the fading process can no longer be considered flat, or frequency non-selective. In fact, with larger delays between rays, the channel becomes a wideband fast fading channel or a *frequency-selective fading* channel. Such a channel can often be modeled by a baseband FIR filter,

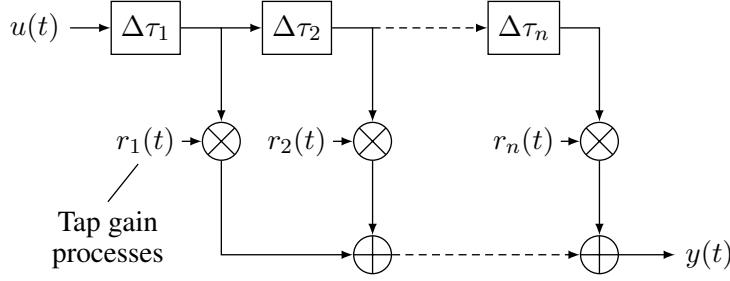


Figure 3.22 Frequency-selective fading channel model.

whose complex-valued taps are independently faded according to the statistics (either Rayleigh or Rice distributed as seen in the last section), see Fig. 3.22.

The relationship between input and output is given by the convolution

$$y(t) = (h * u)(t) = \int h(\tau)u(t - \tau) d\tau, \quad (3.53)$$

where the impulse response of the wireless channel is

$$h(t) = \sum_{k=1}^n r_k(t)\delta(t - t_k). \quad (3.54)$$

The channel impulse response is often abbreviated as CIR in the literature. Its Fourier transform is called the channel transfer function (CTF). Note that the channel taps are both complex-valued and time-variant. The mean powers of the taps

$$P_k = E\{|r_k|^2\} \quad (3.55)$$

are usually given by a channel profile together with the corresponding delays, the so-called *power delay profile* (see Fig. 3.23), which is further characterized by the following parameters:

- *total excess delay*: the delay between the first and the last arriving tap response, essentially the amount by which the duration of a transmitted symbol is extended by the channel,
- *mean delay* defined by

$$\tau_0 = \frac{1}{P_T} \sum_k P_k \tau_k, \quad (3.56)$$

where, compared to Fig. 3.22, we have

$$\tau_k = \sum_{l=1}^k \Delta\tau_l, \quad (3.57)$$

- *RMS delay spread* defined by

$$\tau_{\text{RMS}} = \sqrt{\frac{1}{P_T} \sum_k P_k (\tau_k - \tau_0)^2}, \quad (3.58)$$

where equation total power in the channel is given by

$$P_T = \sum_k P_k. \quad (3.59)$$

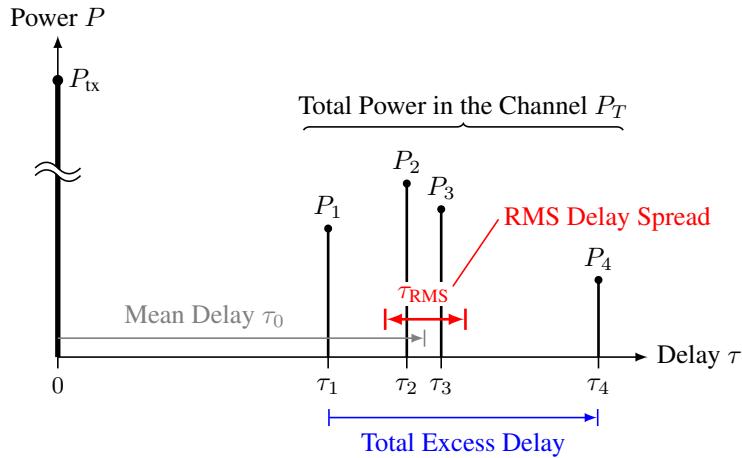


Figure 3.23 Power delay profile.

Eq. (3.58) can be developed into yet another, equivalent form,

$$\begin{aligned}
 \tau_{\text{RMS}} &= \sqrt{\frac{1}{P_T} \sum_k P_k (\tau_k - \tau_0)^2} \\
 &= \sqrt{\frac{1}{P_T} \sum_k P_k \tau_k^2 - \frac{1}{P_T} 2\tau_0 \sum_k P_k \tau_k + \frac{1}{P_T} \sum_k P_k \tau_0^2} \\
 &= \sqrt{\frac{1}{P_T} \sum_k P_k \tau_k^2 - 2\tau_0^2 + \tau_0^2} \\
 &= \sqrt{\frac{1}{P_T} \sum_k P_k \tau_k^2 - \tau_0^2} \\
 &= \sqrt{\bar{\tau}^2 - \tau_0^2}.
 \end{aligned} \tag{3.60}$$

The delay spread, is a good indicator of how frequency selective the channel is. It is inversely proportional to the *coherence bandwidth*. If the delay spread of a channel is much smaller than the symbol duration, the channel may be considered narrowband (flat fading). On the other hand, large delay spreads compared to symbol lengths mean that intersymbol interference (ISI) occurs, i.e., echos of a previous symbol overlap with the current incoming symbol.

3.5.3 Mobile Environment

In a truly mobile communication environment we face two challenges: Firstly, mountains, hills, and buildings lead to excessive delay spread, which in turn introduces frequency-selective fading. Secondly, changing the user position makes the channel response highly variable. An RF receiver usually needs to address both problems, the first one by providing a flexible structure that can essentially invert the channel, and the second one by providing an algorithm to track the channel changes. Such an equalizer structure is shown in Fig. 3.24.

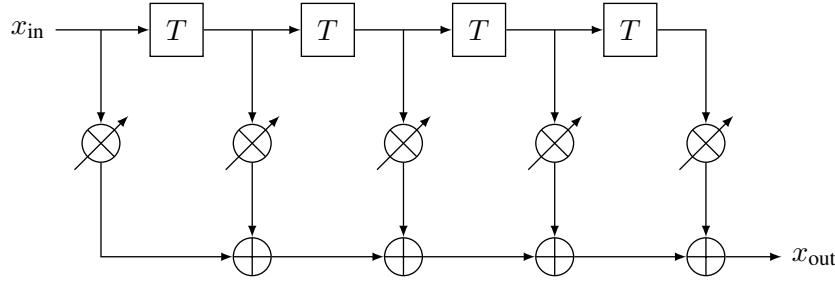


Figure 3.24 Block diagram of a channel equalizer.

Doppler Spectrum The classical Doppler spectrum of the frequency shift experienced by a mobile user driving at an arbitrary direction can be obtained by realising that his angle α towards the base station is uniformly distributed between $-\pi$ and π

$$p(\alpha) = \begin{cases} \frac{1}{2\pi} & -\pi < \alpha < \pi \\ 0 & \text{otherwise} \end{cases}. \quad (3.61)$$

The Doppler shift experienced by the transmission between mobile and base station, or vice versa, is the projection of the speed vector onto the connection line between the mobile and the base station

$$f_d = f \cdot \frac{v}{c} \cdot \cos \alpha = f_m \cdot \cos \alpha, \quad (3.62)$$

where f_m designates the maximum Doppler frequency. Transformation of the pdf from one random variable (in this case α) to the pdf of another random variable (we are interested in f here) takes place using the Jacobian (in this case one-dimensional)

$$\left| \frac{df}{d\alpha} \right| = |-f_m \sin \alpha| = f_m |\sin \alpha|. \quad (3.63)$$

Because the cos function is an even function, the same Doppler shift can be obtained by two different angles $\pm\alpha$. Thus, the power density spectrum of the Doppler shift can be computed using the Jacobian in the following way

$$p(f) = \frac{p(\alpha) + p(-\alpha)}{\left| \frac{df}{d\alpha} \right|} = \frac{\frac{1}{2\pi} + \frac{1}{2\pi}}{\left| \frac{df}{d\alpha} \right|} = \frac{1}{\pi f_m |\sin \alpha|}. \quad (3.64)$$

Substituting

$$\sin \alpha = \sqrt{1 - \cos^2 \alpha} = \sqrt{1 - (f/f_m)^2} \quad (3.65)$$

into Eq. (3.64) results in

$$p(f) = \begin{cases} \frac{1}{\pi f_m \sqrt{1-(f/f_m)^2}} & |f| < f_m \\ 0 & |f| \geq f_m \end{cases}. \quad (3.66)$$

Being a probability density function, Eq. (3.66) integrates to one as can easily be checked,

$$\int_{-f_m}^{f_m} p(f) df = \int_{-f_m}^{f_m} \frac{1}{\pi f_m \sqrt{1 - (f/f_m)^2}} df = \int_{-1}^1 \frac{1}{\pi f_m \sqrt{1 - x^2}} f_m dx = \frac{1}{\pi} \arcsin x \Big|_{-1}^1 = 1. \quad (3.67)$$

The form of the Doppler distribution as given in Eq. (3.66) and seen in Fig. 3.25 explains why the curve is often referred to as a 'bowl shape' or a 'bathtub shape' curve. It is used within many channel simulators. Note that due to its fractional bandwidth being very small and to the momentary Doppler shift being a function of a particular scatterer, the Doppler distribution according to Fig. 3.25 is difficult to observe directly in practice.

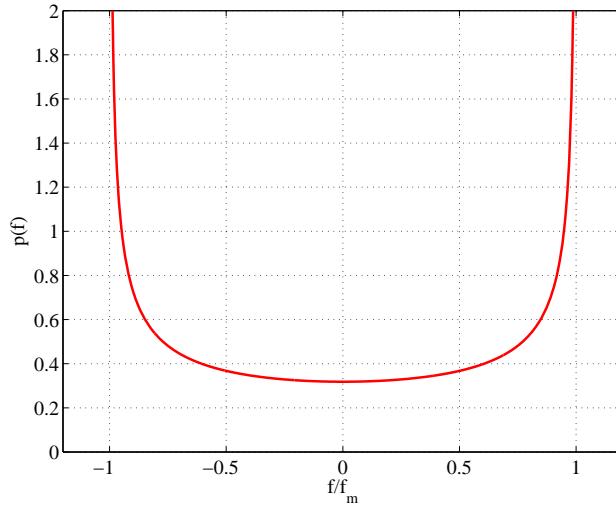


Figure 3.25 Distribution of classical Doppler spectrum.

Rayleigh Fading Parameters Important parameters to characterize the changing aspect of a mobile wireless channel are the level-crossing rate (LCR) and the average fade duration (AFD). For a Rayleigh fading channel they are

$$\text{LCR} = \sqrt{2\pi} \cdot f_d \cdot \rho \cdot e^{-\rho^2} \quad (3.68)$$

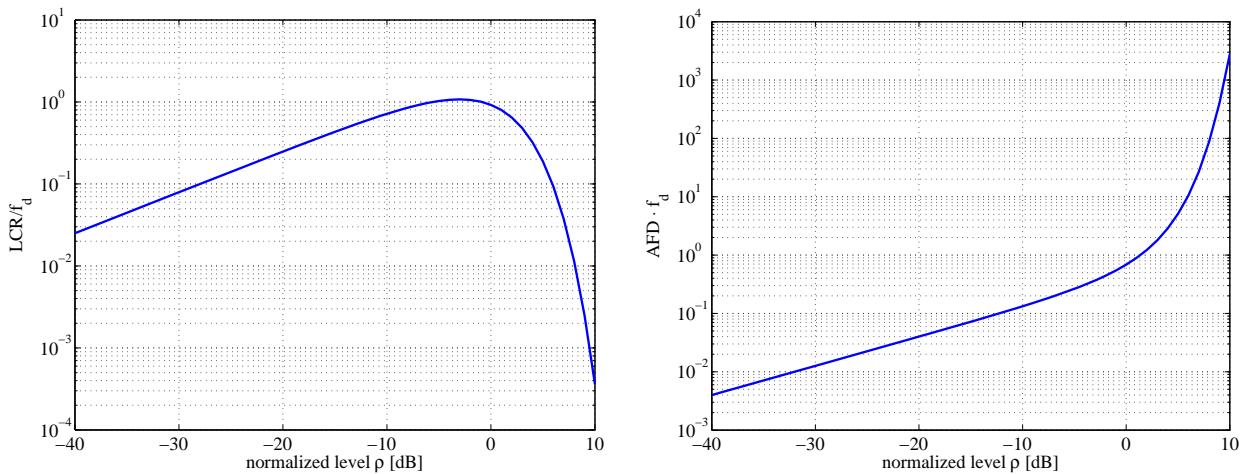
and

$$\text{AFD} = \frac{e^{\rho^2} - 1}{\sqrt{2\pi} \cdot f_d \cdot \rho}, \quad (3.69)$$

where $\rho = R_{\text{threshold}}/R_{\text{RMS}}$ is the normalized threshold level (as an amplitude ratio). Note that similar to other bandwidth/time products, the product of LCR and AFD is a constant given by

$$\text{LCR} \cdot \text{AFD} = 1 - e^{-\rho^2}. \quad (3.70)$$

Fig. 3.26(a) and 3.26(b) show the LCR and the AFD, respectively. The LCR and the AFD are important parameters to consider when choosing error control codes and diversity schemes for wireless systems.



(a) Level crossing rate normalized to Doppler frequency. (b) Average fade duration norm. to Doppler frequency.

Figure 3.26 Rayleigh fading parameters.

3.5.4 Relationship of Fading Parameters

In the last sections, we have encountered different types of fading. These types can be characterized by different parameters, some of which are related to each other.

The coherence time T_{coh} , which characterizes the time over which the channel transfer function stays essentially the same, and the Doppler spread, which describes the spectral broadening, are inversely related, thus

$$T_{coh} \propto \frac{1}{f_d}. \quad (3.71)$$

On the other hand, the coherence bandwidth B_{coh} determines the span over which the channel appears flat. Its inverse parameter, the delay spread τ stands for the time broadening. Again these two parameters relate to each other as

$$B_{coh} \propto \frac{1}{\tau}. \quad (3.72)$$

In summary, for a given communication system with symbol time T and bandwidth B we can now build Table 3.4. Note that a fading channel can be at the same time one of the top two types and one of the bottom two types, hence, any one of four combinations is possible. In fact, the proportionality factors in

Fading type	Condition	Alternative formulation of condition
Flat fading	$B_{coh} \gg B$	$\tau \ll T$
Frequency-selective fading	$B_{coh} < B$	$\tau > T$
Slow fading	$T_{coh} \gg T$	$f_d \ll B$
Fast fading	$T_{coh} < T$	$f_d > B$

Table 3.4 Fading types.

Eqs. (3.71) and (3.72) depend on the exact behavior of the channel impulse response and its variation in time, respectively. Rules of thumb are given by [29] as

$$B_{coh} \approx \frac{1}{50\tau}, \quad (3.73)$$

if the frequency correlation function is above 0.9, and

$$B_{coh} \approx \frac{1}{5\tau}, \quad (3.74)$$

if the frequency correlation function is above 0.5. Sometimes, the relationship

$$B_{coh} \approx \frac{1}{2\pi\tau} \quad (3.75)$$

is also used (leading to yet another value for the correlation function). Very similar rules can be found for the relationship given by Eq. (3.71). Rappaport [29] states that

$$T_{coh} \approx \frac{9}{16\pi f_d} \quad (3.76)$$

for the time correlation to be above 0.5.

3.5.5 Ways Out

If we have to design a wireless communication system in such a hostile environment as mentioned above, we can either accept the situation and deploy a channel equalizer, or try to seek the situation where we do not need one. If we look at a typical snapshot of a channel frequency response, e.g., Fig. 3.27, we realize

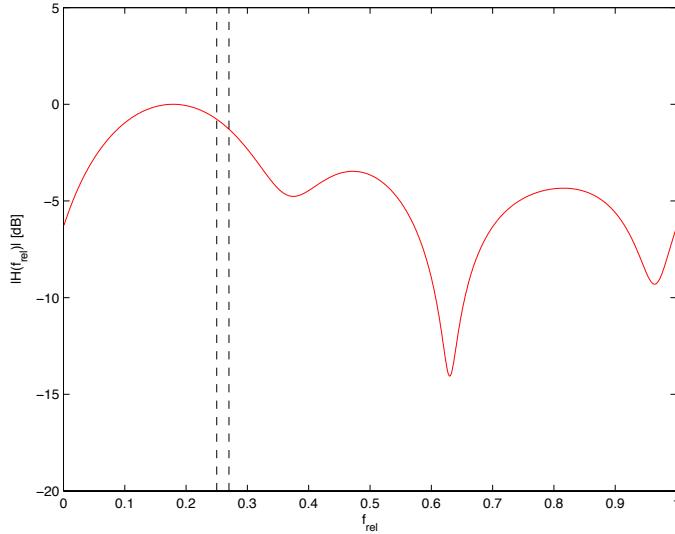


Figure 3.27 Frequency response snapshot of a frequency-selective fading channel.

that the channel looks quite flat as long as we only use a narrow part (subband) of it, indicated by the region between the dashed lines. We then say that the bandwidth (of the subband) is smaller than the coherence bandwidth

$$W_{\text{sub}} \ll B_{\text{coh}}. \quad (3.77)$$

The idea of dividing the available bandwidth into a multitude of subbands that are smaller than the coherence bandwidth was realized by Chang [5] in the 60s. All we need to 'equalize' the channel in each 'flat' subband, is a variable factor, a complex-valued factor, in order to restore amplitude and phase. Why does this factor need to be variable? Well, because each subband may show a different attenuation to be compensated for.

Moreover, the attenuation may change, and so does the complex-valued factor we have to multiply the subband with. We better make sure the situation does not change within the duration of a symbol, otherwise we cannot determine the compensation factor. Hence, we have a second condition on the choice of our system. The duration of a symbol within a subband shall be much shorter than the coherence time

$$T_{\text{sub}} \ll T_{\text{coh}}. \quad (3.78)$$

Dividing a bandwidth W into N_c subbands delivers

$$W_{\text{sub}} = \frac{W}{N_c} \quad (3.79)$$

and

$$T_{\text{sub}} = \frac{1}{W_{\text{sub}}} = \frac{N_c}{W}. \quad (3.80)$$

Eqs. (3.77) and (3.78) now deliver

$$\frac{W}{N_c} \ll B_{\text{coh}} \quad \text{and} \quad \frac{N_c}{W} \ll T_{\text{coh}}. \quad (3.81)$$

Combined into one inequality we finally get

$$\frac{W}{B_{\text{coh}}} \ll N_c \ll WT_{\text{coh}}. \quad (3.82)$$

We thus get lower and upper limits for the choice of the number of subchannels N_c . The division of a problem into small pieces that can be handled more easily has found its application already in ancient times. The Romans were said to govern according to this principle of 'divide and conquer' (*divide et impera*).

3.5.6 Multipath for Low Bandwidth

For a low bandwidth with long symbols, the different paths of a multipath environment are almost simultaneous. We can thus treat the composite signal as

$$\begin{aligned} y &= \sum_k a_k \cos(\omega t + \varphi_k) \\ &= \sum_k a_k (\cos \omega t \cos \varphi_k - \sin \omega t \sin \varphi_k) \\ &= \cos \omega t \underbrace{\sum_k a_k \cos \varphi_k}_{A \cos \Phi} - \sin \omega t \underbrace{\sum_k a_k \sin \varphi_k}_{A \sin \Phi} \\ &= A \cos(\omega t + \Phi), \end{aligned} \quad (3.83)$$

where

$$A = \sqrt{\left(\sum_k a_k \cos \varphi_k \right)^2 + \left(\sum_k a_k \sin \varphi_k \right)^2} \quad (3.84)$$

and

$$\Phi = \arccos \frac{\sum_k a_k \cos \varphi_k}{A}. \quad (3.85)$$

4 Channel Equalization Using Adaptive Filters

4.1 Matched Filter Concept

One of the first steps in the receiver chain is to ensure that not too much noise contaminates the received signal. Thus, the aim of the pre-detector signal processing is to provide the detection circuit with the highest signal-to-noise ratio possible. In order to derive the filter required for this task, we assume the channel signal to have the spectrum $X(f)$. Let the filter transfer function be $W(f)$. The output signal, whose SNR shall be maximized, is then given by the inverse Fourier transform

$$u(t) = \int_{-\infty}^{\infty} W(f)X(f) e^{j2\pi ft} df. \quad (4.1)$$

The power of the signal $u(t)$ sampled at time instant T is

$$\sigma_{u(T)}^2 = |u(T)|^2 = \left| \int_{-\infty}^{\infty} W(f)X(f) e^{j2\pi fT} df \right|^2. \quad (4.2)$$

The additive white Gaussian model assumes uncorrelated noise with spectral density $N_0/2$, so the noise power after the filter is

$$\sigma_n^2 = \frac{N_0}{2} \int_{-\infty}^{\infty} |W(f)|^2 df, \quad (4.3)$$

which is a consequence of Parseval's theorem. The SNR to maximize is now

$$\text{SNR} = \frac{\sigma_{u(T)}^2}{\sigma_n^2} = \frac{\left| \int_{-\infty}^{\infty} W(f)X(f) e^{j2\pi fT} df \right|^2}{\frac{N_0}{2} \int_{-\infty}^{\infty} |W(f)|^2 df}. \quad (4.4)$$

The Schwarz inequality tells us that

$$\left| \int_{-\infty}^{\infty} W(f)X(f) e^{j2\pi fT} df \right|^2 \leq \int_{-\infty}^{\infty} |W(f)|^2 df \int_{-\infty}^{\infty} |X(f)|^2 df. \quad (4.5)$$

Using inequality (4.5) in Eq. (4.4) leads to an inequality for the SNR

$$\text{SNR} \leq \left(\frac{N_0}{2} \right)^{-1} \int_{-\infty}^{\infty} |X(f)|^2 df. \quad (4.6)$$

The maximum of Eq. (4.6) is independent of $W(f)$, but is reached for equality of Eq. (4.5), which in turns occurs if $W(f)$ is of the form

$$W(f) = X^*(f) e^{-j2\pi fT}. \quad (4.7)$$

This condition states that the spectrum of the filter be matched to the signal by taking the complex conjugate of the spectrum of the delayed signal as input filter, hence the name matched filter. This result also explains why a raised-cosine pulse shaping filter is usually equally divided among transmitter and receiver resulting in root raised cosine filters, see for example Fig. 4.1.

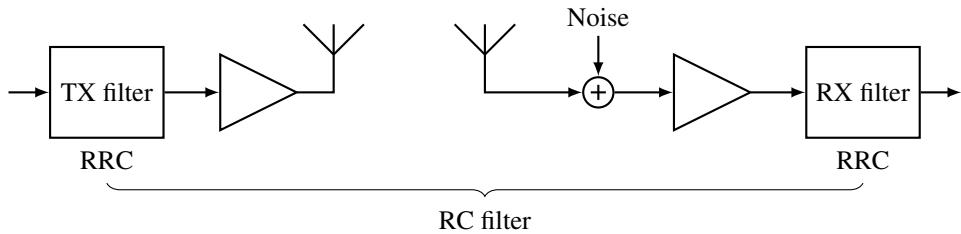


Figure 4.1 Example of a matched-filter concept in a wireless system, where the raised-cosine (RC) filter is constituted by two root-raised-cosine (RRC) filters.

4.2 Applications of Adaptive Filters

Adaptive filters are used mainly to solve one of the following four problems:

- System identification
- Inverse modelling
- Linear prediction
- Noise cancellation

Common to all approaches is the generation of an error signal, which is central to the adaptation process. The error signal $e[k]$ is essentially the difference between a reference $d[k]$ and an estimated signal $y[k]$. The generation of these signals is different for the four application classes, gathered in Fig. 4.2.

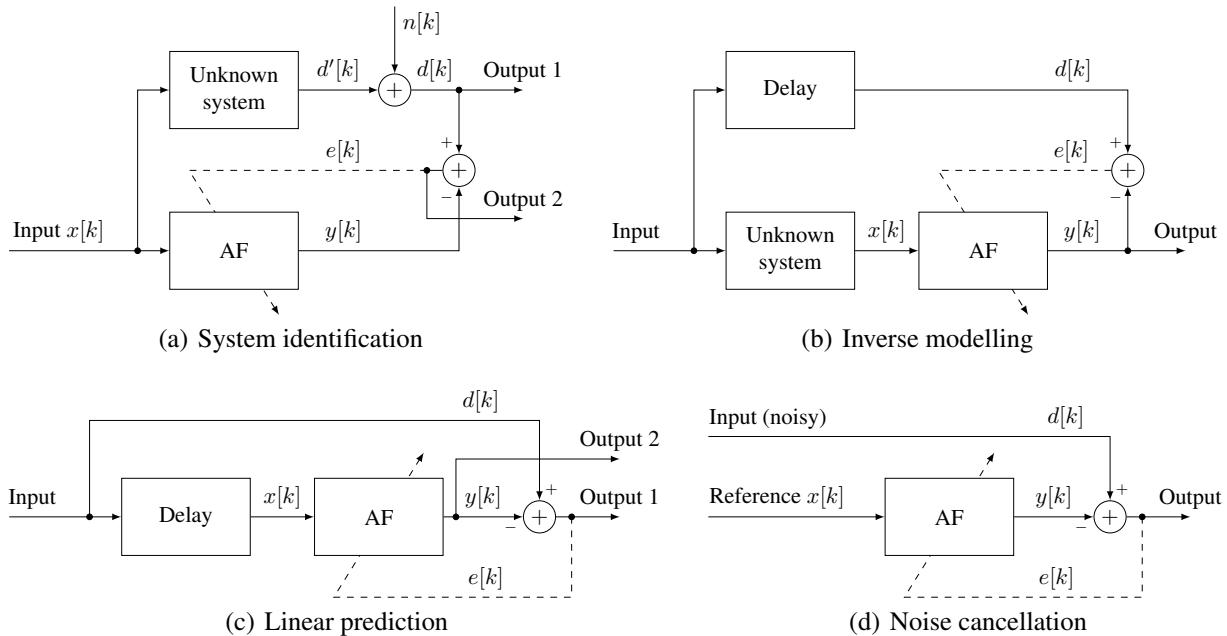


Figure 4.2 The four application classes of adaptive filters.

- The problem of system identification, see Fig. 4.2(a), refers to the task of having to duplicate a system without exactly knowing what is inside.

- The inverse-modelling application, see Fig. 4.2(b), corresponds to the problem often found in a wireless system. We also have an unknown system (the communication channel) as in the previous case. However, we do not attempt to reproduce the output of the channel but are interested in the channel input (the information) instead.
- Linear prediction is very similar to inverse modelling, as can be seen from Fig. 4.2(c). Whereas in inverse modelling we have introduced a delay in the reference signal to give the adaptive filter a better chance to invert the system, in linear prediction the delay is introduced in the path of the adaptive filter, making it even harder to adapt, because it has to predict into the "future".
- In the previous three applications the reference signal was related to the wanted signal, possibly with noise added. In the fourth application the reference signal contains a signal related to the noise, and we try to make the error signal as uncorrelated to the noise signal as possible, see Fig. 4.2(d).

4.3 LMS-Based Adaptive Filters

In this section we assume the structure of an FIR filter, see Fig. 4.3.

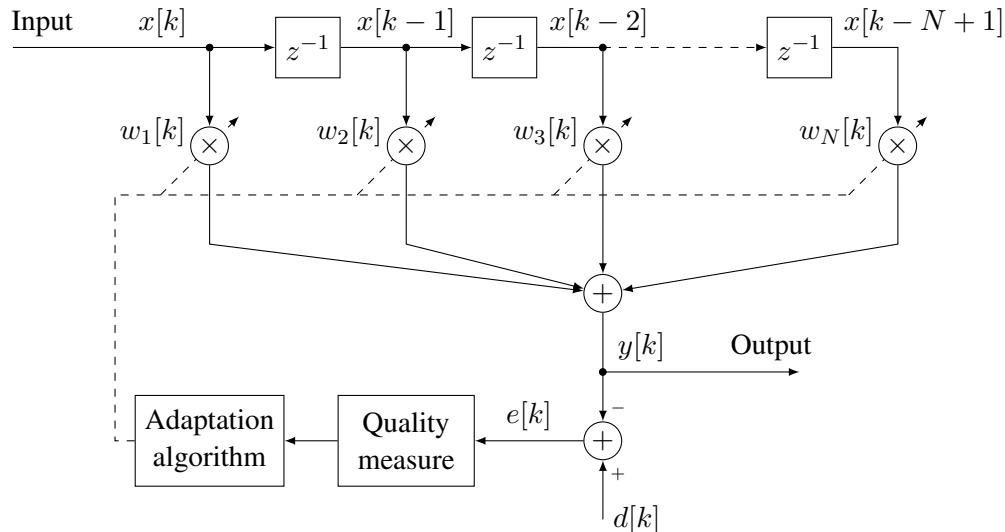


Figure 4.3 FIR-based adaptive filter.

4.3.1 Wiener Filtering

Since the LMS algorithm is based on Wiener filtering, we review the Wiener filter theory. For the following derivation of the Wiener solution for an FIR filter, assume that the output of the FIR filter at time sample k is given by a finite number of weighted past input samples

$$y[k] = \sum_{i=0}^{N-1} w_{i+1} \cdot x[k-i]. \quad (4.8)$$

By using vector definitions $\mathbf{w} \triangleq [w_1, w_2, \dots, w_N]^T$ and $\mathbf{x}[k] \triangleq [x[k], x[k-1], \dots, x[k-N+1]]^T$ we can write above convolution sum as a scalar product

$$y[k] = \mathbf{w}^T \mathbf{x}[k]. \quad (4.9)$$

\mathbf{w} is called the weighting or coefficient vector and $\mathbf{x}[k]$ is called the input signal vector. The difference of the filter output $y[k]$ to a reference signal $d[k]$

$$e[k] = d[k] - y[k]. \quad (4.10)$$

is used to measure the quality of the adaptation process. Using Eq. (4.9) in Eq. (4.10) we get

$$e[k] = d[k] - \mathbf{w}^T \mathbf{x}[k]. \quad (4.11)$$

We are interested in the squared error

$$\begin{aligned} e^2[k] &= (d[k] - \mathbf{w}^T \mathbf{x}[k]) \cdot (d[k] - \mathbf{x}^T[k] \mathbf{w}) \\ &= d^2[k] + \mathbf{w}^T \mathbf{x}[k] \mathbf{x}^T[k] \mathbf{w} - 2d[k] \mathbf{x}^T[k] \mathbf{w}. \end{aligned} \quad (4.12)$$

Note that for now we still regard the coefficient vector as fixed, hence no dependency on k . Our error in Eq. (4.12) is dependant on the input signal, thus time-variant. We are not interested in short-time fluctuations and moreover, the input signal might not always be accessible. Thus, we use the expectation operator to establish the mean-squared error (MSE)

$$E\{e^2[k]\} = E\{d^2[k]\} + \mathbf{w}^T E\{\mathbf{x}[k] \mathbf{x}^T[k]\} \mathbf{w} - 2E\{d[k] \mathbf{x}^T[k]\} \mathbf{w}. \quad (4.13)$$

For ease of use we introduce the auto-correlation matrix

$$\begin{aligned} \mathbf{R} &\stackrel{\triangle}{=} \mathbf{R}_{xx} \stackrel{\triangle}{=} E\{\mathbf{x}[k] \mathbf{x}^T[k]\} \\ &= \begin{pmatrix} E\{\mathbf{x}[k] \mathbf{x}[k]\} & E\{\mathbf{x}[k] \mathbf{x}[k-1]\} & \cdots & E\{\mathbf{x}[k] \mathbf{x}[k-N+1]\} \\ E\{\mathbf{x}[k-1] \mathbf{x}[k]\} & E\{\mathbf{x}[k-1] \mathbf{x}[k-1]\} & \cdots & E\{\mathbf{x}[k-1] \mathbf{x}[k-N+1]\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{\mathbf{x}[k-N+1] \mathbf{x}[k]\} & E\{\mathbf{x}[k-N+1] \mathbf{x}[k-1]\} & \cdots & E\{\mathbf{x}[k-N+1] \mathbf{x}[k-N+1]\} \end{pmatrix} \\ &= \begin{pmatrix} E\{\mathbf{x}^2[0]\} & E\{\mathbf{x}[1] \mathbf{x}[0]\} & \cdots & E\{\mathbf{x}[N-1] \mathbf{x}[0]\} \\ E\{\mathbf{x}[0] \mathbf{x}[1]\} & E\{\mathbf{x}^2[0]\} & \cdots & E\{\mathbf{x}[N-2] \mathbf{x}[0]\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{\mathbf{x}[0] \mathbf{x}[N-1]\} & E\{\mathbf{x}[0] \mathbf{x}[N-2]\} & \cdots & E\{\mathbf{x}^2[0]\} \end{pmatrix} \end{aligned} \quad (4.14)$$

and the cross-correlation vector

$$\begin{aligned} \mathbf{p} &\stackrel{\triangle}{=} E\{d[k] \mathbf{x}[k]\} \\ &= \begin{pmatrix} E\{d[k] \mathbf{x}[k]\} \\ E\{d[k] \mathbf{x}[k-1]\} \\ \vdots \\ E\{d[k] \mathbf{x}[k-N+1]\} \end{pmatrix} \\ &= \begin{pmatrix} E\{d[0] \mathbf{x}[0]\} \\ E\{d[0] \mathbf{x}[1]\} \\ \vdots \\ E\{d[0] \mathbf{x}[N-1]\} \end{pmatrix}. \end{aligned} \quad (4.15)$$

Strictly speaking, the last equalities of Eqs. (4.14) and (4.15) are only valid for wide-sense stationary real signals. Thus, \mathbf{R} is a Toeplitz matrix (diagonal structure) and for real signals also symmetric. Now, Eq. (4.13) becomes

$$E\{e^2[k]\} = E\{d^2[k]\} + \mathbf{w}^T \mathbf{R} \mathbf{w} - 2\mathbf{p}^T \mathbf{w}. \quad (4.16)$$

The minimization of the mean-squared error according to Eq. (4.16) can be achieved by setting the differentiation of this quadratic equation with respect to the coefficient vector \mathbf{w} equal to zero,

$$\nabla_{\mathbf{w}} \{E\{e^2[k]\}\} = 2\mathbf{R} \mathbf{w} - 2\mathbf{p}. \quad (4.17)$$

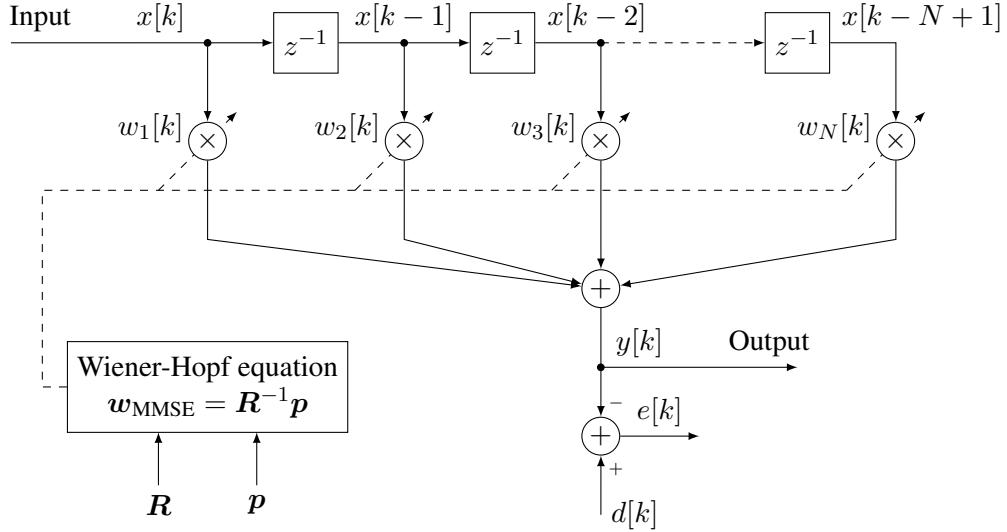


Figure 4.4 Wiener-based FIR filter.

The solution of Eq. (4.17)

$$\mathbf{w}_{\text{MMSE}} = \mathbf{R}^{-1} \mathbf{p} \quad (4.18)$$

is called the Wiener filter or the MMSE filter. Using the above-derived criterion we can build an adaptive filter according to Fig. 4.4. Interestingly, Eq. (4.18) can also be derived using the orthogonality principle, which is illustrated for a two-tap FIR filter in Fig. 4.5. In this context, the signals are interpreted as geometric

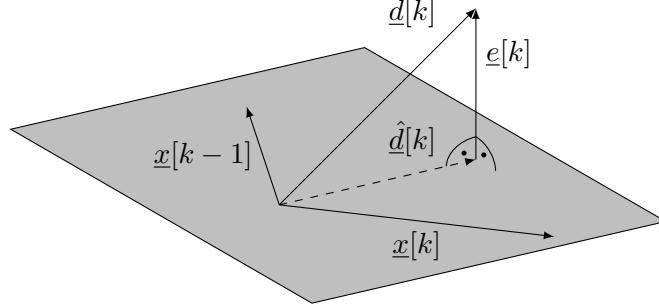


Figure 4.5 Orthogonality principle shown for a two-tap FIR filter.

vectors and therefore underlined. The problem of optimal filtering now corresponds to the problem of estimating the true vector $\underline{d}[k]$ by an estimate $\hat{\underline{d}}[k]$ that is the weighted linear combination of two input vectors. The estimate $\hat{\underline{d}}[k]$ will always be constrained to the plane spanned by the input vectors $\underline{x}[k]$ and $\underline{x}[k-1]$. The true value $\underline{d}[k]$ is thus best approximated, if the error $\underline{e}[k]$ is orthogonal to all input vectors. Using the scalar product $\langle \cdot, \cdot \rangle$, the orthogonality principle can be expressed as

$$\langle \underline{x}[k-i], \underline{e}[k] \rangle = 0 \quad \forall i \in [0, N-1]. \quad (4.19)$$

Coming back from the geometric interpretation to the signals, the orthogonality principle says that

$$E\{\mathbf{x}[k]\mathbf{e}[k]\} = \mathbf{0}. \quad (4.20)$$

Eq. (4.11) used in Eq. (4.20) leads to

$$E\{\mathbf{x}[k](d[k] - \mathbf{w}^T \mathbf{x}[k])\} = \mathbf{0} \quad (4.21)$$

or

$$\begin{aligned} E\{\mathbf{x}[k]d[k]\} - E\{\mathbf{x}[k](\mathbf{w}^T \mathbf{x}[k])\} &= \\ E\{\mathbf{x}[k]d[k]\} - E\{\mathbf{x}[k](\mathbf{x}^T[k]\mathbf{w})\} &= \\ E\{\mathbf{x}[k]d[k]\} - E\{\mathbf{x}[k]\mathbf{x}^T[k]\}\mathbf{w} &= \mathbf{0}. \end{aligned} \quad (4.22)$$

Thanks to the definitions above we can write

$$\mathbf{p} - \mathbf{R}\mathbf{w} = \mathbf{0}, \quad (4.23)$$

which will lead to the same solution as obtained in Eq. (4.18).

4.3.2 The LMS Algorithm

The solution of the last section has three significant drawbacks. Firstly, the inversion of a matrix is not always feasible in real time. Secondly, the information leading to the computation of \mathbf{R} and \mathbf{p} might not be available. And finally, the statistics of \mathbf{R} and \mathbf{p} might change over time. The first drawback can be avoided by applying the method of steepest descent to find the filter coefficients. The one-time calculation of Eq. (4.18) is replaced by the iterative computation

$$\mathbf{w}[k+1] = \mathbf{w}[k] - c\nabla_{\mathbf{w}[k]}\{E\{e^2[k]\}\}, \quad (4.24)$$

where c is a step-size small enough in order not to overshoot the minimum. As mentioned, the statistics might not be available. We thus need an estimation for the statistics. For ergodic processes we may substitute the mean of a random variable by its time average, in our case $E\{e^2[k]\}$ by

$$\widehat{\text{MSE}} = \frac{1}{M} \sum_{l=k-M+1}^k e^2[l] \quad (4.25)$$

Eq. (4.24) used with the gradient estimate of Eq. (4.25) is called the block-LMS algorithm. Its special form if $M = 1$ is called the LMS algorithm for which we can express the gradient as

$$\begin{aligned} \nabla_{\mathbf{w}[k]}\{e^2[k]\} &= \nabla_{\mathbf{w}[k]}\{(d[k] - \mathbf{w}^T \mathbf{x}[k])^2\} \\ &= -2\mathbf{x}[k](d[k] - \mathbf{w}^T \mathbf{x}[k]) \\ &= -2\mathbf{x}[k]e[k]. \end{aligned} \quad (4.26)$$

Using another step-size parameter $\mu = 2c$ we obtain the LMS update rule as

$$\mathbf{w}[k+1] = \mathbf{w}[k] + \mu \cdot e[k] \cdot \mathbf{x}[k]. \quad (4.27)$$

The LMS algorithm as described by Eq. (4.27) and implemented in Fig. 4.6 is due to Widrow and Hoff as far back as 1960, but is still used today thanks to its simplicity. Note that when used for complex-valued signals (and possibly complex coefficients) we have [41]

$$\mathbf{w}(t+1) = \mathbf{w}(t) + 2\mu e\mathbf{x}^*, \quad (4.28)$$

$$\mathbf{w}[k+1] = \mathbf{w}[k] + \mu \cdot e[k] \cdot \mathbf{x}^*[k]. \quad (4.29)$$

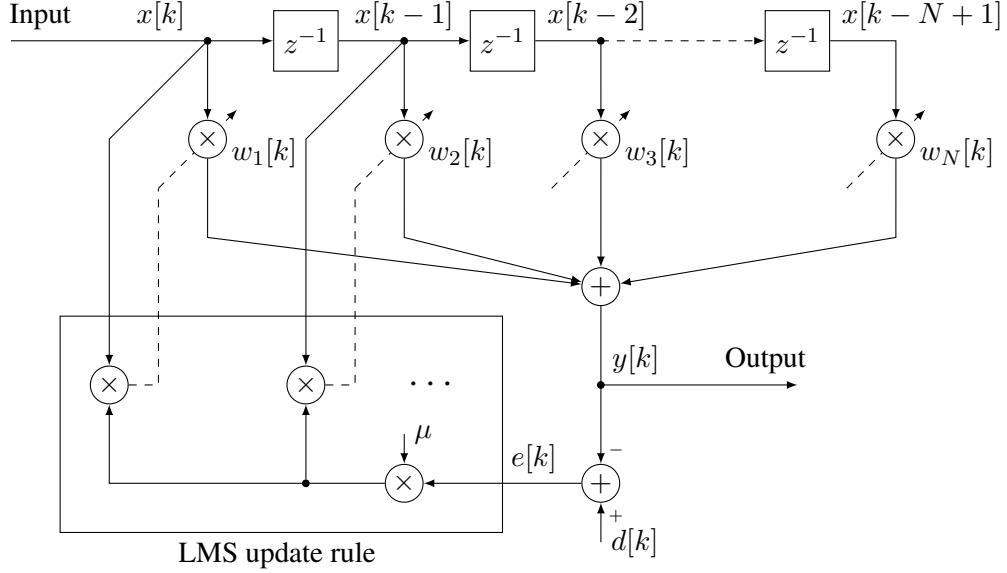


Figure 4.6 LMS algorithm to find the coefficients of an adaptive filter.

4.4 RLS-Based Adaptive Filters

4.4.1 The LS Algorithm

If we look back at Eq. (4.25), where we replaced the mean error by a moving-average of some deterministic values, we realize that we could compute all statistics needed by deterministic sums. We start by computing the deterministic auto-correlation matrix at time point k , for which we assume that we use a block of length $k - k_0 + 1$

$$\mathcal{R}_k \triangleq \sum_{l=k_0}^k \mathbf{x}[l] \mathbf{x}^T[l] = \begin{bmatrix} \sum_{l=k_0}^k x[l]x[l] & \sum_{l=k_0}^k x[l]x[l-1] & \dots & \sum_{l=k_0}^k x[l]x[l-N+1] \\ \sum_{l=k_0}^k x[l-1]x[l] & \sum_{l=k_0}^k x[l-1]x[l-1] & \dots & \sum_{l=k_0}^k x[l-1]x[l-N+1] \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{l=k_0}^k x[l-N+1]x[l] & \sum_{l=k_0}^k x[l-N+1]x[l-1] & \dots & \sum_{l=k_0}^k x[l-N+1]x[l-N+1] \end{bmatrix}. \quad (4.30)$$

Similarly, the deterministic cross-correlation vector is

$$\mathcal{P}_k \triangleq \sum_{l=k_0}^k d[l] \mathbf{x}[l] = \begin{bmatrix} \sum_{l=k_0}^k d[l]x[l] \\ \sum_{l=k_0}^k d[l-1]x[l] \\ \vdots \\ \sum_{l=k_0}^k d[l-N+1]x[l] \end{bmatrix}. \quad (4.31)$$

Recalling Eq. (4.18) we can now write

$$\mathbf{w}_{\text{LS}} = \mathcal{R}_k^{-1} \mathcal{P}_k, \quad (4.32)$$

which we call the least-square (LS) solution.

We can usually find the inverse of above deterministic auto-correlation matrix. If we define the error sum over more error elements than the number of coefficients we want to find, the problem might be formulated

using a non-square matrix \mathcal{R} . In this case, the system is overdetermined and the coefficients might be found using the so-called Moore-Penrose pseudoinverse

$$\mathbf{w}_{\text{LS}} = (\mathcal{R}^T \mathcal{R})^{-1} \mathcal{R}^T \mathcal{P}. \quad (4.33)$$

4.4.2 The RLS Algorithm

The deterministic auto-correlation matrix can be written in a recursive way

$$\mathcal{R}_k = \mathcal{R}_{k-1} + \mathbf{x}[k] \mathbf{x}^T[k]. \quad (4.34)$$

and the recursive description of the cross-correlation vector yields

$$\mathcal{P}_k = \mathcal{P}_{k-1} + d[k] \mathbf{x}[k]. \quad (4.35)$$

With this, we can formulate the recursive least-square algorithm as:

1. Update the auto-correlation matrix according to Eq. (4.34)
2. Update the cross-correlation matrix according to Eq. (4.35)
3. Invert \mathcal{R}_k
4. Compute the new optimal coefficient vector $\mathbf{w}_{\text{RLS},k}$ using Eq. (4.32)

The inversion of the matrix R_k in each iteration is computationally inefficient. The Matrix-Inversion Lemma

$$(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{D}\mathbf{A}^{-1} \quad (4.36)$$

avoids the recalculation of an inverse after every step. By setting

$$\mathbf{A} \triangleq \mathcal{R}_{k-1} \quad (4.37a)$$

$$\mathbf{B} \triangleq \mathbf{x}[k] \quad (4.37b)$$

$$\mathbf{C} \triangleq 1 \quad (4.37c)$$

$$\mathbf{D} \triangleq \mathbf{x}^T[k] = \mathbf{B}^T \quad (4.37d)$$

the middle bracket on the right-hand side of Eq. (4.36) is then

$$(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1} = (\mathbf{x}^T[k]\mathcal{R}_{k-1}^{-1}\mathbf{x}[k] + 1)^{-1} = \frac{1}{1 + \mathbf{x}^T[k]\mathcal{R}_{k-1}^{-1}\mathbf{x}[k]}. \quad (4.38)$$

The whole Matrix-Inversion Lemma now translates to

$$(\mathcal{R}_{k-1} + \mathbf{x}[k]\mathbf{x}^T[k])^{-1} = \mathcal{R}_{k-1}^{-1} - \frac{\mathcal{R}_{k-1}^{-1}\mathbf{x}[k]\mathbf{x}^T[k]\mathcal{R}_{k-1}^{-1}}{1 + \mathbf{x}^T[k]\mathcal{R}_{k-1}^{-1}\mathbf{x}[k]}. \quad (4.39)$$

But the left-hand side of Eq. (4.39) is \mathcal{R}_k^{-1} , hence

$$\mathcal{R}_k^{-1} = \mathcal{R}_{k-1}^{-1} - \frac{\mathcal{R}_{k-1}^{-1}\mathbf{x}[k]\mathbf{x}^T[k]\mathcal{R}_{k-1}^{-1}}{1 + \mathbf{x}^T[k]\mathcal{R}_{k-1}^{-1}\mathbf{x}[k]}. \quad (4.40)$$

Sometimes, older data becomes obsolete and should be weighted by a forgetting factor $\lambda < 1$. Since we compute the inverse of the auto-correlation matrix, the exponential weighting happens in the following fashion

$$\mathcal{R}_k^{-1} = \lambda^{-1} \mathcal{R}_{k-1}^{-1} - \frac{\lambda^{-1} \mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k] \lambda^{-1} \mathcal{R}_{k-1}^{-1}}{1 + \lambda^{-1} \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]}. \quad (4.41)$$

We can now write

$$\begin{aligned} \mathbf{w}_{\text{RLS},k} &= \mathcal{R}_k^{-1} \cdot \mathcal{P}_k \\ &= \left(\mathcal{R}_{k-1}^{-1} - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1}}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} \right) \cdot (\mathcal{P}_{k-1} + d[k] \mathbf{x}[k]) \\ &= \mathcal{R}_{k-1}^{-1} \mathcal{P}_{k-1} + \mathcal{R}_{k-1}^{-1} d[k] \mathbf{x}[k] - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1}}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} \mathcal{P}_{k-1} - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1}}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} d[k] \mathbf{x}[k] \\ &= \mathbf{w}_{\text{RLS},k-1} + \mathcal{R}_{k-1}^{-1} d[k] \mathbf{x}[k] - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k]}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} \mathbf{w}_{\text{RLS},k-1} - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1}}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} d[k] \mathbf{x}[k]. \end{aligned} \quad (4.42)$$

We can further simplify Eq. (4.42) by using $\mathbf{x}^T \mathcal{R}^{-1} \mathbf{x} \mathbf{x} = \mathbf{x} \mathbf{x}^T \mathcal{R}^{-1} \mathbf{x}$

$$\begin{aligned} \mathbf{w}_{\text{RLS},k} &= \mathbf{w}_{\text{RLS},k-1} - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k]}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} \mathbf{w}_{\text{RLS},k-1} + \mathcal{R}_{k-1}^{-1} d[k] \mathbf{x}[k] - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1}}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} d[k] \mathbf{x}[k] \\ &= \mathbf{w}_{\text{RLS},k-1} - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k]}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} \mathbf{w}_{\text{RLS},k-1} \\ &\quad + \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] + \mathcal{R}_{k-1}^{-1} \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}[k] - \mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} d[k] \\ &= \mathbf{w}_{\text{RLS},k-1} - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k]}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} \mathbf{w}_{\text{RLS},k-1} + \frac{\mathcal{R}_{k-1}^{-1}}{1 + \mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k]} d[k] \mathbf{x}[k] \\ &= \mathbf{w}_{\text{RLS},k-1} + \frac{\mathcal{R}_{k-1}^{-1}}{\mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k] + 1} d[k] \mathbf{x}[k] - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k] \mathbf{x}^T[k]}{\mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k] + 1} \mathbf{w}_{\text{RLS},k-1} \\ &= \mathbf{w}_{\text{RLS},k-1} + \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k]}{\mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k] + 1} d[k] - \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k]}{\mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k] + 1} \mathbf{x}^T[k] \mathbf{w}_{\text{RLS},k-1} \\ &= \mathbf{w}_{\text{RLS},k-1} + \frac{\mathcal{R}_{k-1}^{-1} \mathbf{x}[k]}{\mathbf{x}^T[k] \mathcal{R}_{k-1}^{-1} \mathbf{x}[k] + 1} (d[k] - \mathbf{x}^T[k] \mathbf{w}_{\text{RLS},k-1}), \end{aligned} \quad (4.43)$$

which shows a clear similarity to the LMS update rule for the adaptive FIR filter, Eq. (4.27).

4.5 Kalman Filtering

The Kalman filter owns its name to Rudolf E. Kálmán (1930–2016), who first came up with the idea of expressing prediction values in terms of statistical properties of observations [20].

For the following derivation of the Kalman filter we assume a process in its state-space description of a time-discrete system, see Fig. 4.7. The process has multiple inputs, multiple states and a single output. The

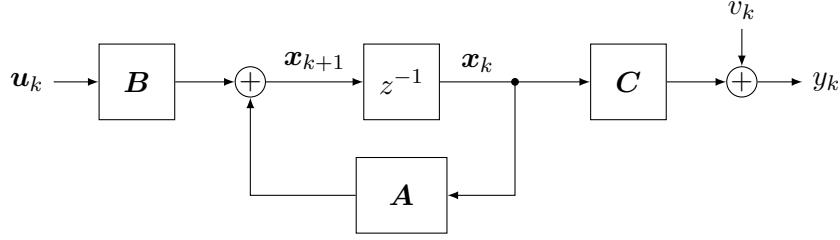


Figure 4.7 State-space model.

state variables are conveniently collected in the column vector \mathbf{x}_k . The state equations to describe the above model are

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \quad (4.44)$$

$$y_k = \mathbf{C}\mathbf{x}_k + v_k. \quad (4.45)$$

Note that in contrast to many state-space descriptions with a common driving force, the noise processes for Eqs. (4.44) and (4.45) are different. While the square matrix \mathbf{A} describes the feedback, i.e., how one state leads to the next, the matrices \mathbf{B} and \mathbf{C} describe the input transfer and the output transfer, respectively. At the output, there is an additional noise term, the measurement noise v_k , modeled as a zero-mean Gaussian white noise sequence characterized by

$$v_k \sim \mathcal{N}(0, R_k), \quad (4.46)$$

where $R_k = E\{v_k^2\}$ is the measurement-noise variance. Note that we have deliberately chosen a single-output variable, hence y_k and v_k are scalars and \mathbf{C} is a row vector. The process noise \mathbf{u}_k stands at the beginning of the whole plant and may be looked at as the initiating force. It is modeled as a zero-mean Gaussian white noise sequence characterized by

$$\mathbf{u}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k), \quad (4.47)$$

where $\mathbf{Q}_k = E\{\mathbf{u}_k \mathbf{u}_k^T\}$ is the $N \times N$ process-noise covariance matrix. The goal is now to predict or estimate the state vector together with other stochastic parameters such as the covariance matrices. We are interested in predicting the output at time step k under the assumption that we observe the output up to time step $k - 1$. We call this $\hat{y}_{k|k-1}$ and the associated prediction error

$$\begin{aligned} e_{k|k-1} &= y_k - \hat{y}_{k|k-1} \\ &= \mathbf{C}\mathbf{x}_k + v_k - \hat{y}_{k|k-1}. \end{aligned} \quad (4.48)$$

Because the measurement noise is uncorrelated and of zero mean, we can use Eq. (4.45) to write

$$\hat{y}_{k|k-1} = \mathbf{C}\hat{\mathbf{x}}_{k|k-1}, \quad (4.49)$$

and the error can now be expressed as

$$e_{k|k-1} = \mathbf{C}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + v_k. \quad (4.50)$$

Similarly to the orthogonality reasoning in the case of the Wiener Filter, we wish our prediction error to be orthogonal to the predicted state $\hat{\mathbf{x}}_{k|k-1}$ and hence to the output prediction, essentially being a linear combination of the state:

$$E\{e_{k|k-1} \hat{\mathbf{x}}_{k|k-1}\} = \mathbf{0}, \quad (4.51)$$

$$E\{e_{k|k-1} \hat{y}_{k|k-1}\} = 0. \quad (4.52)$$

This prediction error is called the innovation, which we use to come up with a still better estimate of the state

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{k}_k e_{k|k-1}, \quad (4.53)$$

where the column vector \mathbf{k}_k is the Kalman gain yet to be determined. The remaining difference $\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}$ cannot be determined by the innovation $e_{k|k-1}$. Again the orthogonality principle rules that

$$E\{e_{k|k-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})\} = \mathbf{0}. \quad (4.54)$$

We can bring this into another form,

$$E\{e_{k|k-1}\mathbf{x}_k\} = E\{e_{k|k-1}\hat{\mathbf{x}}_{k|k}\}. \quad (4.55)$$

By subtracting a common term, we get

$$E\{e_{k|k-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\} = E\{e_{k|k-1}(\hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k-1})\}. \quad (4.56)$$

Eq. (4.53) written differently,

$$\mathbf{k}_k e_{k|k-1} = \hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k-1}, \quad (4.57)$$

can be used for the right-hand side of Eq. (4.56), such that

$$E\{e_{k|k-1}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\} = E\{\mathbf{k}_k e_{k|k-1}^2\}. \quad (4.58)$$

Developing simultaneously the left-hand side and the right-hand side of Eq. (4.58) we get

$$\begin{aligned} E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})e_{k|k-1}\} &= \mathbf{k}_k E\{e_{k|k-1}^2\} \\ E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})((\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T \mathbf{C}^T + v_k)\} &= \mathbf{k}_k E\{\mathbf{C}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T \mathbf{C}^T + v_k^2\} \\ \mathbf{P}_{k|k-1} \mathbf{C}^T &= \mathbf{k}_k (\mathbf{C} \mathbf{P}_{k|k-1} \mathbf{C}^T + R_k). \end{aligned} \quad (4.59)$$

where we have used the error covariance matrix

$$\mathbf{P}_{k|k-1} \triangleq E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\}. \quad (4.60)$$

Finally, we can solve Eq. (4.59) for the Kalman Gain

$$\mathbf{k}_k = \mathbf{P}_{k|k-1} \mathbf{C}^T (\mathbf{C} \mathbf{P}_{k|k-1} \mathbf{C}^T + R_k)^{-1}. \quad (4.61)$$

The estimation error covariance matrix can be written into a recursive form

$$\begin{aligned} \mathbf{P}_{k|k} &= E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T\} \\ &= E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{k}_k e_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{k}_k e_{k|k-1})^T\} \\ &= E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T - (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \mathbf{k}_k^T e_{k|k-1} \\ &\quad - \mathbf{k}_k e_{k|k-1} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T + \mathbf{k}_k e_{k|k-1}^2 \mathbf{k}_k^T\} \\ &= E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T - 2\mathbf{k}_k e_{k|k-1} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T + \mathbf{k}_k e_{k|k-1}^2 \mathbf{k}_k^T\}. \end{aligned} \quad (4.62)$$

Using Eq. (4.58) we see that the last two terms of Eq. (4.62) can be combined leading to

$$\begin{aligned} \mathbf{P}_{k|k} &= E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T - \mathbf{k}_k e_{k|k-1} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\} \\ &= \mathbf{P}_{k|k-1} - \mathbf{k}_k \mathbf{C} \mathbf{P}_{k|k-1} \\ &= (\mathbf{I} - \mathbf{k}_k \mathbf{C}) \mathbf{P}_{k|k-1}. \end{aligned} \quad (4.63)$$

We can now develop the prediction error covariance matrix as

$$\begin{aligned} \mathbf{P}_{k+1|k} &= E\{(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})^T\} \\ &= E\{\mathbf{A}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T \mathbf{A}^T + \mathbf{u}_k \mathbf{u}_k^T\} \\ &= \mathbf{A} \mathbf{P}_{k|k} \mathbf{A}^T + \mathbf{Q}_k. \end{aligned} \quad (4.64)$$

Fig. 4.8 shows the block diagram of the Kalman filter structure.

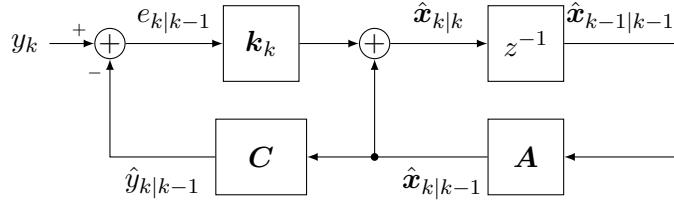


Figure 4.8 Kalman filter.

In summary, the Kalman filter is the following procedure:

1. Set $k = 0$, enter the prior estimate $\hat{x}_{0|0}$ and an initial value for the error covariance matrix $P_{0|0}$
 2. Compute the Kalman gain: $k_k = P_{k|k-1}C^T(CP_{k|k-1}C^T + R_k)^{-1}$
 3. Update the estimate with the new measurement y_k : $\hat{x}_{k|k} = \hat{x}_{k|k-1} + k_k(y_k - C\hat{x}_{k|k-1})$
 4. Compute the error covariance matrix: $P_{k|k} = (I - k_k C)P_{k|k-1}$
 5. Project ahead: $\hat{x}_{k+1|k} = A\hat{x}_{k|k}$ and $P_{k+1|k} = AP_{k|k}A^T + Q_k$
 6. Increment k and return to step 2.

In practice, determining the measurement-noise variance R_k is generally possible by taking an offline sample measurement. Determining the process-noise covariance matrix Q_k is more difficult, since the process is not directly observable. Even if the parameters Q_k and R_k cannot be determined on a rational basis, they can be *tuned* for best filter performance. This is often done in an offline simulation. If the measurement and process noise are stationary, R_k and Q_k are constant and P_k and k_k will quickly converge to a constant value. In this case, P_k and k_k can be precomputed offline [40].

On another note, if we reduce above form to the so-called Kalman filter of unforced dynamics, where we have $A = I$, $Q = 0$ and $R_k = 1$, we have Step 2 and Step 3 as:

$$k_k = \frac{P_{k|k-1} C^T}{C P_{k|k-1} C^T + 1}, \quad (4.65)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + k_k(y_k - C\hat{x}_{k|k-1}). \quad (4.66)$$

Together they result in

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \frac{P_{k|k-1}C^T}{CP_{k|k-1}C^T + 1}(y_k - C\hat{x}_{k|k-1}), \quad (4.67)$$

which looks very much like the RLS algorithm in Eq. (4.42), if the coefficient vector \mathbf{w} is assigned as the state vector.

4.6 Channel Equalization

One of the more challenging tasks within the signal processing of a wireless receiver is the equalization of the channel impulse response. A mobile communication channels experiences many undesired effects. In addition to fast and slow fading it may also show multipath propagation due to reflections on mountains and

buildings. Mathematically, the transmitted signal samples s get convolved by the impulse response h of the channel

$$x(kT) = \sum_{i=0}^{L_{\text{ch}}-1} h_i \cdot s((k-i)T). \quad (4.68)$$

In order to undo the channel distortion, we need to equalize or deconvolve the channel. Fig. 4.9 shows a configuration of such a system, where the source signal s is filtered by the transmission channel designated by $h(z)$. Some noise may be added at this point. The convolved signal x is then deconvolved by the deconvolution filter $w(z)$ to recreate the source signal. If $w(z)$ can invert $h(z)$, u is close to the original source signal s . The settling of the equalizer coefficients $w(z)$ is usually done in an adaptive way, often using a stochastic-gradient method to minimize some cost function.

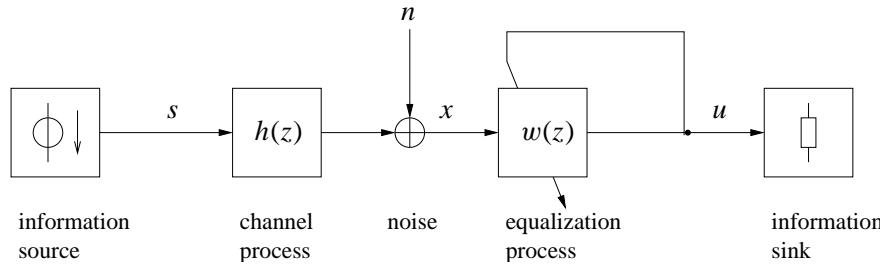


Figure 4.9 Adaptive equalization system.

In vector notation, an equalization system can be described as follows. Let the vector $\mathbf{h} \triangleq [h_0, h_1, \dots, h_{L-1}]^T$ denote the coefficients of the channel filter of length L . We will assume this channel to be constant, but unknown. By filling up the time series s into a vector of the form $\mathbf{s}_t \triangleq [s_t, s_{t-1}, \dots, s_{t-L+1}]^T$, we can describe the input to the equalizer as

$$\mathbf{x}_t = \mathbf{h}^T \mathbf{s}_t + \mathbf{n}_t, \quad (4.69)$$

where the noise vector is designated as \mathbf{n}_t . Likewise, we have $\mathbf{w}_t \triangleq [w_{-N,t}, \dots, w_{0,t}, \dots, w_{N,t}]^T$ denoting the coefficients of the deconvolution filter of length $2N + 1$. Using a symmetric configuration we imply an initialization of the filter coefficients of $w_{k,0} = \delta(k)$, essentially a center-tap initialization, which allows the causal development of inverse filters for nonminimum-phase systems. Since the equalizer will be an adaptive filter, we have the additional parameter t , indicating that \mathbf{w}_t is the set of coefficients at time t . Defining the vector of channel outputs as $\mathbf{x}_t \triangleq [x_{t+N}, \dots, x_t, \dots, x_{t-N}]^T$, we get the signal after the equalizer as

$$u_t = \mathbf{w}_t^T \mathbf{x}_t. \quad (4.70)$$

The first part of the equalization task is now to find vector \mathbf{w} such that the signal u_t is essentially the same as the source signal s_t . By ‘‘essentially’’ we mean up to a delay term plus terms relating to the noise \mathbf{n}_t .

Alternatively, the formulation of the above task can be carried out in the z -domain. By assigning $s(z)$, $n(z)$, $x(z)$, and $u(z)$ the z -transforms of the source signal, the noise signal, the channel output, and the equalizer output, respectively, and

$$h(z) = \sum_{k=0}^L h_k z^{-k}, \quad (4.71)$$

$$w_t(z) = \sum_{k=-N}^N w_{k,t} z^{-k}. \quad (4.72)$$

the z -transforms of the channel filter and the equalizer, respectively, we can write

$$x(z) = h(z)s(z) + n(z) \quad (4.73)$$

and

$$\begin{aligned} u(z) &= w_t(z)x(z) \\ &= w_t(z)h(z)s(z) + w_t(z)n(z) \end{aligned} \quad (4.74)$$

for the input to and the output of the equalizer, respectively. As an example, the effect of intersymbol interference on the eye diagram of a 4-PAM signal is shown in Fig. 4.10.

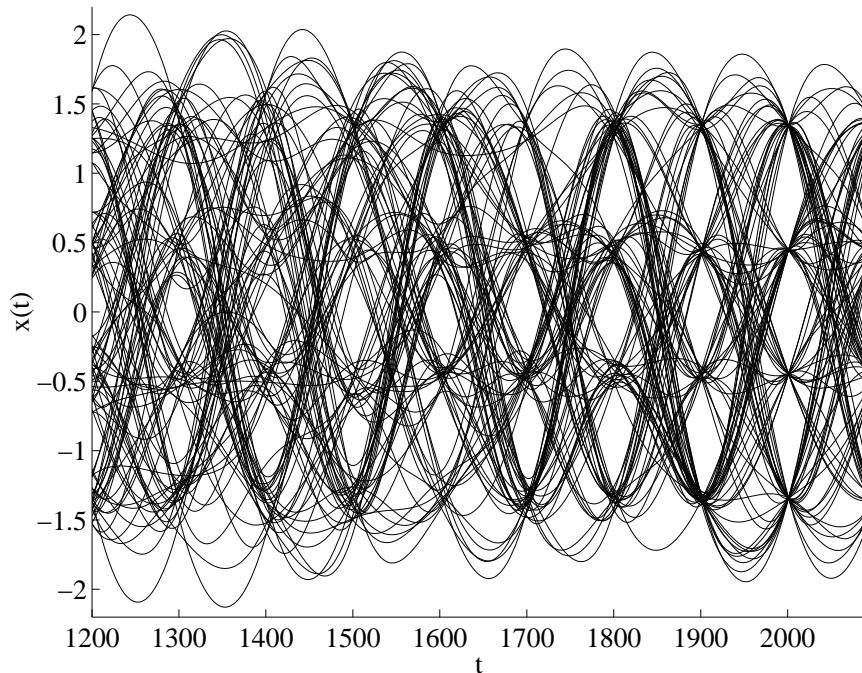


Figure 4.10 Eye diagram starting with heavily ISI and constantly changing over to an ISI-free environment.

Equalizers can be classified according to a variety of criteria, which are given in the following.

- **Optimization criterion:** Zero-forcing, MMSE
- **Sampling:** Baud spaced, fractionally spaced
- **Feedback structure:** Linear (no feedback), decision feedback
- **Training:** TS based, decision-directed, blind
- **Adaptation:** LMS, RLS, Kalman Filter approach
- **Domain:** time, frequency, mixed

In Section 4.1 it was stated that the input filter should be matched to the signal of interest. The signal can be distorted by an unknown channel, so an approach often employed is to fix the matched filter and mitigate the distortion by an equalizer. Because timing offsets would destroy information through aliasing in critically sampled systems, the matched filter works on an oversampled basis. If timing offsets are compensated either

within the matched filter or subsequently but prior to further equalization, the equalizer may be symbol spaced (also called baud spaced). In the following we give a brief overview of the most common types of equalizers.

4.6.1 Zero-Forcing Equalizers

In the following, the variables L_{ch} and L_{eq} designate the length of the channel and the equalizer length in number of taps, respectively. The following linear equalizer minimizes the peak distortion and is often referred to as the zero-forcing equalizer. It tries to exactly equalize a channel, regardless of any possible noise enhancement due to channel nulls near the unit circle. Strictly speaking, the name zero-forcing applies to infinite-length equalizers, because with a finite number of coefficients, complete inversion of the channel might not be possible.

The aim is to find an equalizer vector \mathbf{w} , which when applied to the channel response, collected in the channel transfer matrix \mathbf{H} , leads to a unit vector $\boldsymbol{\delta}$ (meaning a perfectly equalized unit impulse response of the system, with the location of the 1 corresponding to the delay after equalization):

$$\mathbf{H}\mathbf{w} = \boldsymbol{\delta} \Leftrightarrow \begin{bmatrix} h_0 & 0 & \dots & 0 & 0 \\ h_1 & h_0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{L_{\text{ch}}-1} & h_{L_{\text{ch}}-2} & \dots & h_0 & 0 \\ 0 & h_{L_{\text{ch}}-1} & \dots & h_1 & h_0 \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & h_{L_{\text{ch}}-1} & h_{L_{\text{ch}}-2} \\ 0 & 0 & \dots & 0 & h_{L_{\text{ch}}-1} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{L_{\text{eq}}-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (4.75)$$

In order to get a solution for the equalizer vector $\mathbf{w} = [w_0 \dots w_{L_{\text{eq}}-1}]^T$, we need to invert the channel transfer matrix \mathbf{H} . Unfortunately, except for the trivial case of a single-path channel, the dimension of the channel transfer matrix \mathbf{H} is always

$$\dim(\mathbf{H}) = (L_{\text{ch}} + L_{\text{eq}} - 1) \times L_{\text{eq}}. \quad (4.76)$$

However, an optimal solution in terms of minimal peak distortion might still be formulated. Suppose a channel with the finite-length impulse response packed into vector $\mathbf{h} = [h_0 \dots h_{L_{\text{ch}}-1}]^T$. We now want to choose the equalizer coefficients $\mathbf{w} = [w_0 \dots w_{L_{\text{eq}}-1}]^T$, such that the combined result contains one 1 and as many zeros as possible. With L_{eq} degrees of freedom we can form L_{eq} equations

$$\begin{bmatrix} h_0 & 0 & \dots & 0 & 0 \\ h_1 & h_0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{L_{\text{eq}}-2} & h_{L_{\text{eq}}-3} & \dots & h_0 & 0 \\ h_{L_{\text{eq}}-1} & h_{L_{\text{eq}}-2} & \dots & h_1 & h_0 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{L_{\text{eq}}-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (4.77)$$

or in short

$$\mathbf{w} = \mathbf{H}^{-1} \boldsymbol{\delta}. \quad (4.78)$$

The d leading zeros in δ are useless, because they force w_0 to w_{d-1} to zero, hence reduce the degree of freedom. δ should therefore start with a one for maximum flexibility. w_0 to $w_{L_{\text{eq}}-1}$ can then be solved for easily using back-substitution

$$\begin{aligned} w_0 &= \frac{1}{h_0}, \\ w_k &= \frac{-\sum_{l=0}^{k-1} h_{k-l} w_l}{h_0} \quad k = 1 \dots L_{\text{eq}} - 1. \end{aligned} \quad (4.79)$$

This 'naive' approach does not work well for non-minimum-phase systems. Alternatively, (4.75) can be solved in the MMSE sense, using

$$\mathbf{w} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \boldsymbol{\delta}, \quad (4.80)$$

where in this case $\boldsymbol{\delta}$ may contain leading zeros. The expression $(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ is called the Moore-Penrose pseudoinverse.

4.6.2 MMSE Equalizer

The previously explained zero-forcing equalizer has one major drawback in the case of low SNR. If zeros in the transfer function are placed close to the unit circle, the zero-forcing equalizer tries to invert them amplifying any noise present. Often, the source of the deviation from the ideal signal does not matter, so a minimum mean square error criterion is applied to find the equalizer filter coefficients. For this we have to solve the matrix form of the Wiener-Hopf equation. Let us denote the input to the equalizer as the convolution of the channel impulse response \mathbf{h} with the symbols s_k plus a noise term n_k

$$x_k = \sum_{l=0}^{L_{\text{ch}}-1} h_l s_{k-l} + n_k. \quad (4.81)$$

The output of the equalizer is then

$$u_k = \sum_{i=0}^{L_{\text{eq}}-1} w_i x_{k-i} = \sum_{i=0}^{L_{\text{eq}}-1} w_i \left(\sum_{l=0}^{L_{\text{ch}}-1} h_l s_{k-i-l} + n_{k-i} \right). \quad (4.82)$$

The error signal is then formulated as

$$e_k = s_{k-d} - u_k \quad (4.83)$$

The criterion to minimize is the mean squared error

$$J = E \{ |e_k|^2 \} = E \{ e_k e_k^* \}. \quad (4.84)$$

which can be reached if the error signal is orthogonal to the input at the taps of the filter

$$E \{ e_k^* x_{k-m} \} = 0, \quad m = 0 \dots L_{\text{eq}} - 1. \quad (4.85)$$

Using Eq. (4.83) in Eq. (4.85) we obtain

$$E \left\{ \left(s_{k-d}^* - \sum_{i=0}^{L_{\text{eq}}-1} w_i^* \left(\sum_{l=0}^{L_{\text{ch}}-1} h_l^* s_{k-i-l}^* + n_{k-i}^* \right) \right) \left(\sum_{\tilde{l}=0}^{L_{\text{ch}}-1} h_{\tilde{l}} s_{k-\tilde{l}-m} + n_{k-m} \right) \right\} = 0, \\ m = 0 \dots L_{\text{eq}} - 1. \quad (4.86)$$

4.6.3 Decision-Feedback Equalizer

The linear equalizer given above might not yield acceptable performance for channel zeros near the unit circle, because when the equalizer tries to equalize frequencies around this zero it enhances the noise present around this frequency considerably. If at the time of the detection of a certain symbol, previous symbols are known, the influence of these symbols on the present symbol under detection might be subtracted. The decision-feedback equalizer (see Fig. 4.11) with the coefficients $\mathbf{b} = [b_1 \dots b_{L_{\text{dfe}}}]^T$ is therefore able to

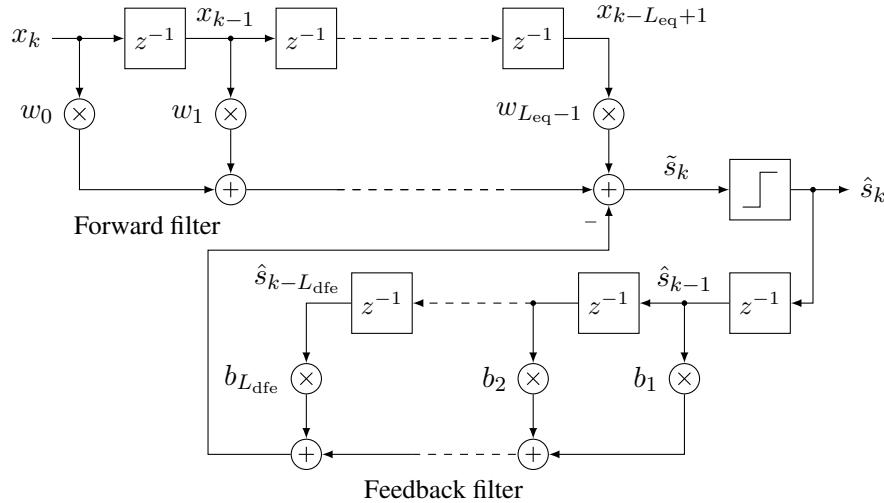


Figure 4.11 Decision-feedback equalizer.

cancel postcursor ISI, while the precursor ISI has to be cancelled in the conventional way using a linear forward filter. We introduce the signal after the summation point of the forward filter and the feedback filter

$$\tilde{s}_k = u_k - \sum_{n=1}^{L_{\text{dfe}}} b_n \hat{s}_{k-n} = \sum_{i=0}^{L_{\text{eq}}-1} w_i \left(\sum_{l=0}^{L_{\text{ch}}-1} h_l s_{k-i-l} + n_{k-i} \right) - \sum_{n=1}^{L_{\text{dfe}}} b_n \hat{s}_{k-n}. \quad (4.87)$$

The error signal can be expressed similarly to Eq. (4.83)

$$e_k = s_{k-d} - \tilde{s}_k. \quad (4.88)$$

The two filters have different input signals. Hence, the error signal has to be orthogonal to both the input to the forward and the input to the feedback filters

$$E \{ e_k^* x_{k-m} \} = 0, \quad m = 0 \dots L_{\text{eq}} - 1, \quad (4.89)$$

$$E \{ e_k^* \hat{s}_{k-m} \} = 0, \quad m = 1 \dots L_{\text{dfe}}. \quad (4.90)$$

If incorrect decisions are disregarded, Eq. (4.90) can be written as

$$E \{ e_k^* s_{k-d-m} \} = 0, \quad m = 1 \dots L_{\text{dfe}}, \quad (4.91)$$

where the decisions \hat{s}_{k-m} have been replaced by the true but delayed symbols s_{k-d-m} . After further modifications this leads to

$$b_m = - \sum_{i=0}^{L_{\text{eq}}-1} w_i h_{d+m-i}, \quad m = 1 \dots L_{\text{dfe}}. \quad (4.92)$$

Eq. (4.89) leads to

$$E_s h_{d-m} - E_s \sum_{i=0}^{L_{\text{eq}}-1} w_i^* \sum_{l=0}^{L_{\text{ch}}-1} h_l^* h_{i+l-m} + E_s \sum_{n=1}^{L_{\text{dfe}}} b_n^* h_{n+d-m} - w_m^* \frac{N_0}{2} = 0, \quad (4.93)$$

where $m = 0 \dots L_{\text{eq}} - 1$. Using Eq. (4.92) in Eq. (4.93) and substituting $k = i_1 + d - i$ results in

$$\sum_{i=0}^{L_{\text{eq}}-1} w_i^* R(i, m) = h_{d-m}^*, \quad m = 0 \dots L_{\text{eq}} - 1. \quad (4.94)$$

with

$$R(i, m) = \sum_{l=0}^{L_{\text{ch}}-1} h_l^* h_{l+i-m} - \sum_{l=d-i+1}^{L_{\text{dfe}}+d-i} h_k^* h_{l+i-m} + \frac{N_0}{2E_s} \delta_{i-m}. \quad (4.95)$$

Under the assumption that $L_{\text{dfe}} + d - i \geq L_{\text{ch}}$ we can write Eq. (4.95) as

$$R(i, m) = \sum_{l=0}^{d-i} h_l^* h_{l+i-m} + \frac{N_0}{2E_s} \delta_{i-m}. \quad (4.96)$$

Note that $R(i, m)$ is no longer Toeplitz but still Hermitian.

Two effects make the concept of a feedback filter work so well. First, it ignores the noise due to orthogonality principle, and second, there is no need for long filters, even for channel zeros near the unit circle. These zeros are simply cancelled by zeros themselves.

The biggest disadvantage of decision-feedback equalizers is error propagation. To demonstrate the error propagation problem, consult Fig. 4.12. The parameters chosen for this illustration are a 4-PAM signal with an equalizer of 24 taps in the forward path (causal delay $d = 10$) and 10 taps in the feedback path. The channel model is a moderate ISI channel.

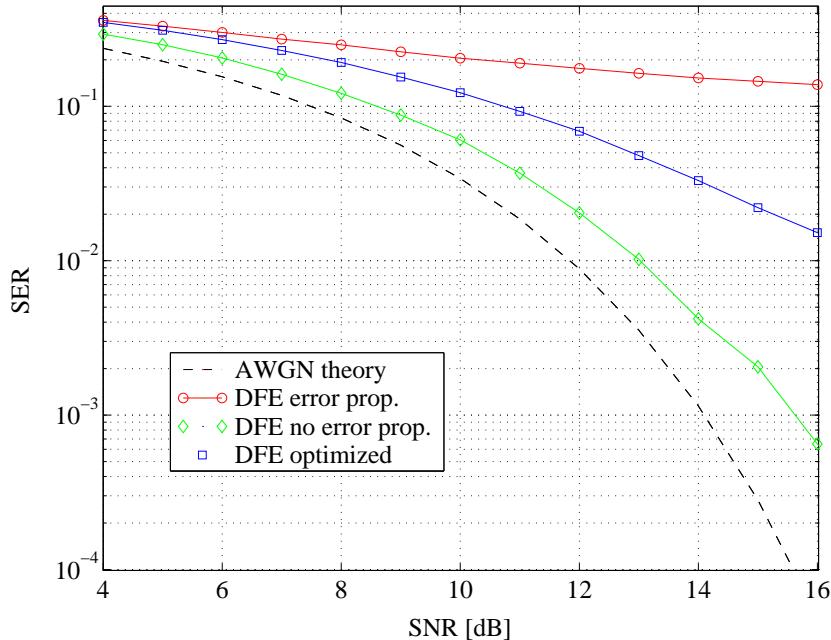


Figure 4.12 BER performance of a DFE with error propagation

4.6.4 Fractionally-Spaced Equalizers

There are at least two good reasons why FSE are used, namely

- insensitivity to timing errors (no aliasing), and
- full equalizability even for FIR responses.

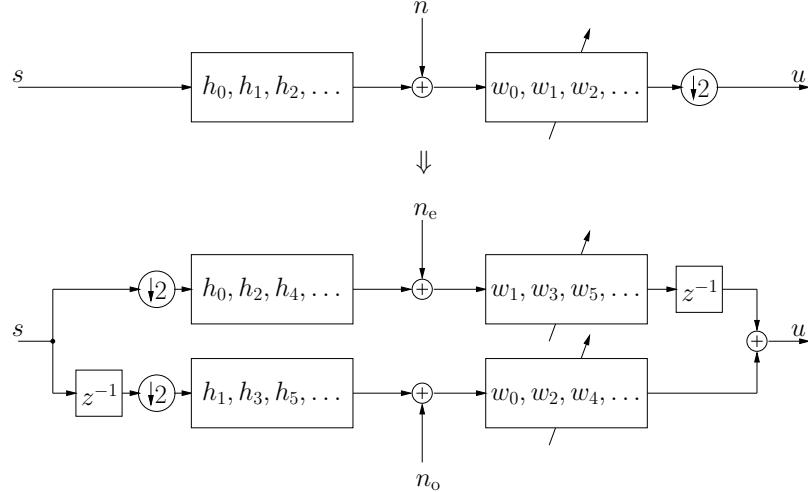


Figure 4.13 Equivalent model for fractionally spaced equalizer.

Fractionally-spaced systems can be modeled at different phases within the symbol sampling instance as depicted in Fig. 4.13. The channel is fed by the signal sequence s . The respective channel responses of the subchannels are \mathbf{h}_1 and \mathbf{h}_2 . The output sequence at the two sample phases are then given by the convolutions

$$x_1 = s * \mathbf{h}_1 \quad (4.97)$$

$$x_2 = s * \mathbf{h}_2. \quad (4.98)$$

Such systems are often referred to as multiphase systems. A solution of Eqs. (4.97) and (4.98) for $(\mathbf{h}_1, \mathbf{h}_2)$ can then be found by solving

$$x_2 * \mathbf{h}_1 - x_1 * \mathbf{h}_2 = 0 \quad (4.99)$$

To see the conditions of the number of coefficients of the equalization filter, we have to observe the channel response matrix which is given for a T -spaced design and for a FSE below. For a system of equations to be determined we need at least as many coefficients as conditions to satisfy. This can never be the case for a T -spaced design, but for a $T/2$ spaced design it is sufficient, if the equalizer is of the same length as the channel. Remember, the system of equations for a T -spaced equalizer ($L_{\text{eq}} = L_{\text{ch}}$) is

$$\begin{bmatrix} h_0 & 0 & \dots & 0 & 0 \\ h_1 & h_0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_{L_{\text{ch}}-2} & h_{L_{\text{ch}}-3} & \dots & h_0 & 0 \\ h_{L_{\text{ch}}-1} & h_{L_{\text{ch}}-2} & \dots & h_1 & h_0 \\ 0 & h_{L_{\text{ch}}-1} & \dots & h_2 & h_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & h_{L_{\text{ch}}-1} & h_{L_{\text{ch}}-2} \\ 0 & 0 & \dots & 0 & h_{L_{\text{ch}}-1} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{L_{\text{eq}}-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (4.100)$$

On the other hand, the system of equations for a $T/2$ -spaced equalizer ($L_{\text{eq}} = L_{\text{ch}}$) equates to

$$\left[\begin{array}{cccc} h_0 & 0 & \dots & 0 \\ h_2 & h_0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{2L_{\text{ch}}-2} & h_{2L_{\text{ch}}-4} & \dots & h_0 \\ 0 & h_{2L_{\text{ch}}-2} & \dots & h_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h_{2L_{\text{ch}}-2} \end{array} \right] \left[\begin{array}{cccc} h_1 & 0 & \dots & 0 \\ h_3 & h_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{2L_{\text{ch}}-1} & h_{2L_{\text{ch}}-3} & \dots & h_1 \\ 0 & h_{2L_{\text{ch}}-1} & \dots & h_3 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h_{2L_{\text{ch}}-1} \end{array} \right] \left[\begin{array}{c} w_1 \\ w_3 \\ \vdots \\ w_{2L_{\text{eq}}-1} \\ w_0 \\ w_2 \\ \vdots \\ w_{2L_{\text{eq}}-2} \end{array} \right] = \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{array} \right]. \quad (4.101)$$

By oversampling with factor N , the channel response matrix (see Section 4.6.1) now shows the dimensions

$$\dim(\mathbf{H}) = (L_{\text{ch}} + L_{\text{eq}} + 1) \times N(L_{\text{eq}} + 1) \quad (4.102)$$

which by choosing N appropriately can yield more variables than conditions, so that a solution exists. In general, for the system to be well determined we must therefore ensure that

$$N(L_{\text{eq}} + 1) \geq (L_{\text{ch}} + L_{\text{eq}} + 1), \quad (4.103)$$

or

$$L_{\text{eq}} \geq \frac{L_{\text{ch}} - (N - 1)}{N - 1} = \frac{L_{\text{ch}}}{N - 1} - 1. \quad (4.104)$$

4.6.5 Training of an Equalizer

For the training of an adaptive equalizer there are basically three different possibilities. If a certain part of the sequence sent over the channel is known we have a reference signal to evaluate the error of the equalizer. But this inclusion of a training sequence decreases the throughput, so should be restricted for capacity reasons as much as possible. The second way of training is the decision-directed approach. The output of the equalizer is passed to a detector, after which we have estimates of the true discrete values. For severe channel distortion this approach is not very robust, since the initially unequalized signal produces many errors in the detector leading to a distorted error signal. Traditionally, these two approaches are combined leading to an equalizer which initially adjusts itself using a training sequence, after which decisions are reliable. The equalizer can then switch (see Fig. 4.14) over to the decision-directed mode to track further

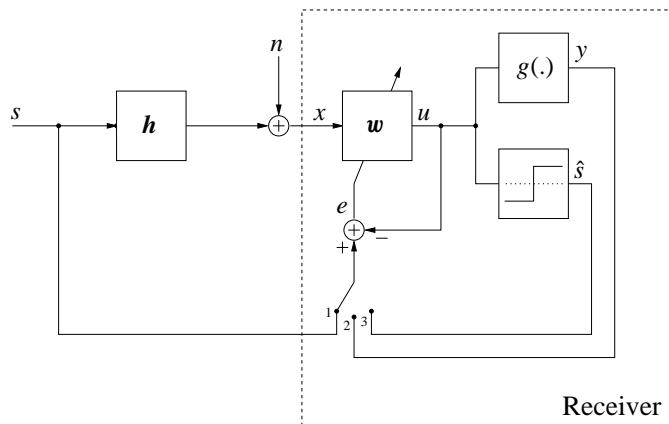


Figure 4.14 Training modes of an equalizer: training-sequence based (1), blind using some nonlinearity (2), and decision-directed (3).

changes in the channel. When decisions are too unreliable and no training sequence is available, a blind equalizer, also called unsupervised or self-recovering, may be employed.

To see how related traditional equalizers and blind equalizer may be, this figure shows an approach which is, depending on the switch position, TS based, decision directed or blind. The switch is not static but can, e.g., after an initial training sequence or blind start-up (opening the eye of an eye diagram), be moved over to decision-directed.

The derivation of the update equation for a complex adaptive linear equalization is given in [41] and results in

$$\mathbf{w}(t+1) = \mathbf{w}(t) + 2\mu e \mathbf{x}^*. \quad (4.105)$$

This is the complex case of the famous LMS algorithm. Along similar lines of argument as in [41], we can derive the update equations for the adaptive feedback equalizer as

$$\mathbf{w}(t+1) = \mathbf{w}(t) + 2\mu_1 e \mathbf{x}^*, \quad (4.106)$$

$$\mathbf{b}(t+1) = \mathbf{b}(t) - 2\mu_2 e \mathbf{s}^*, \quad (4.107)$$

where μ_1 and μ_2 are the not necessarily equal step sizes of the forward and feedback filters, respectively.

4.6.6 Blind Equalizers

As long as communication systems such as GSM can devote up to 22 % of their transmission time to pilot tones, there is a huge potential for blind equalizers, which do not involve any training sequence at all. Moreover, propagation environments and system considerations may make the use of training sequences impractical. Besides, system designs involving heavy coding, thus working in low raw BER environments, prohibit the use of data-aided equalizers. An example of such a system is the new standard¹ for HDTV in the US.

4.6.7 The Constant-Modulus Algorithm (CMA)

The CMA is the most often used algorithm for blind equalization. It is a special case of Godard's algorithm and was independently of Godard discovered by Treichler and Agee (1983). The cost function of the CMA is

$$J(u) = E\{(|u|^p - R_p)^2\}, \quad \text{with } R_p = \frac{E\{|s|^{2p}\}}{E\{|s|^p\}} \quad (4.108)$$

for $p = 2$. Differentiation of this cost function and applying a steepest-ascent algorithm results in the following update equation

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \mu u(|u|^2 - R_2) \mathbf{x}^*. \quad (4.109)$$

Note the similarity to the LMS algorithm, which can be written as

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \mu(u - \hat{s}) \mathbf{x}^*. \quad (4.110)$$

The CM criterion for 16-QAM is illustrated in Fig. 4.15. The circle is the reference curve, from which any deviation will add up to the CM measure J_{CMA} . Despite the fact that 16-QAM is not a constant-modulus constellation, the CM measure J_{CMA} still maintains its minimum (> 0) at perfect equalization, and grows with increasing distortion. As can be further seen, the CM measure is phase insensitive, since the cost function only involves absolute values.

¹Although the standard as such does not specify the receiver equalizer, the system specification is such that blind equalizers are clearly favored.

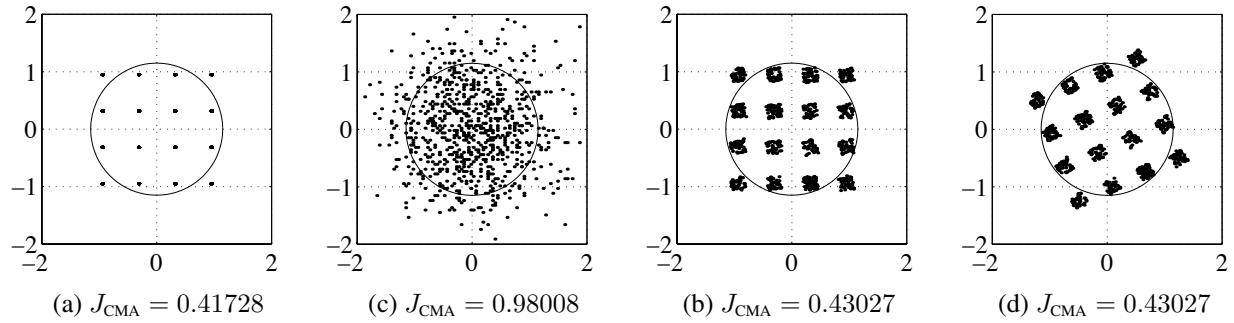


Figure 4.15 Constellation and CMA cost function of some distorted 16-QAM signals: (a) original signal; (b) heavily distorted; (c) equalized with low residual distortion; (d) equalized but with remaining phase error.

4.6.8 An Example

An example shall illustrate the convergence behavior of different algorithms. Assume we have a system with the a channel according to Fig. 4.16. The source is a 16-QAM single source in an AWGN channel, resulting

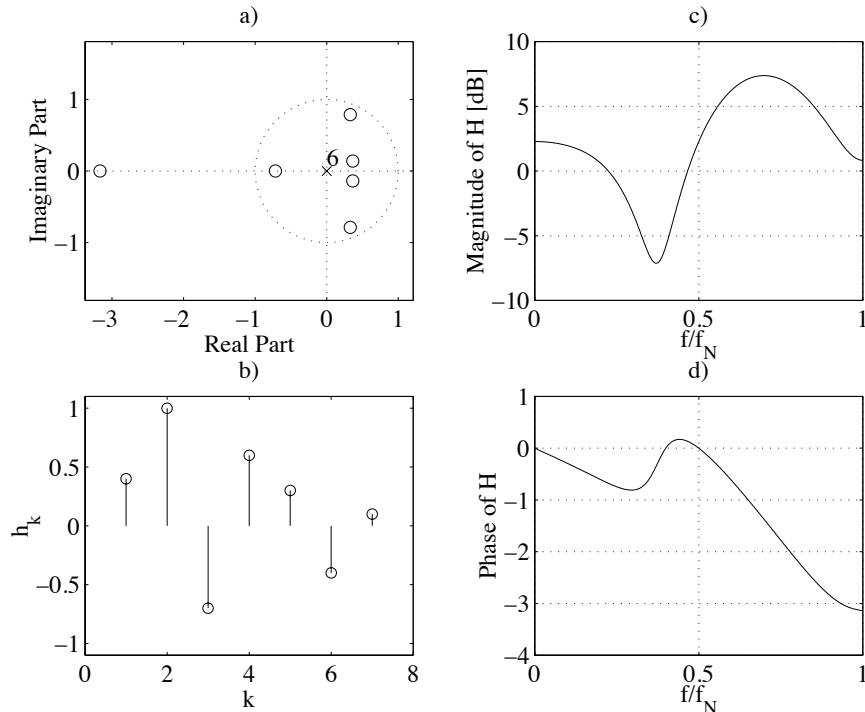


Figure 4.16 Channel model for simulation of adaptive equalization algorithms; a) zero-pole plot, b) impulse response, c) frequency response (magnitude), d) frequency response (phase).

in an SNR=16 dB. The equalizer is assigned 20 taps (4 non-causal taps). Four different training methods are used and their respective convergence times are compared in Fig. 4.17. As can be seen, the decision-directed does not converge at all. Thus, the decision-directed equalizer fails to work in an environment with large initial ISI. The fastest algorithm is, of course, the trained case at the prize of having to provide a training sequence. Stable but slow performance can be expected from the pure CMA. A mixed CMA/decision-directed approach combines stability and fast convergence. Here, the equalizer initially works blindly and

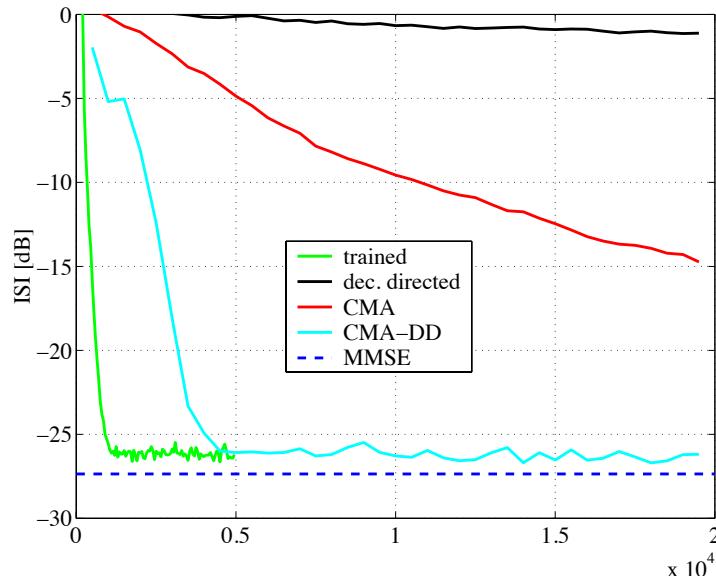


Figure 4.17 Convergence of different equalizer types.

switches over to a decision-directed mode after an initial stage. In the configuration used the switch-over process is done by the algorithm itself.

At convergence the impulse responses of the equalizer and the global impulse response will look like the ones given in Fig. 4.18.

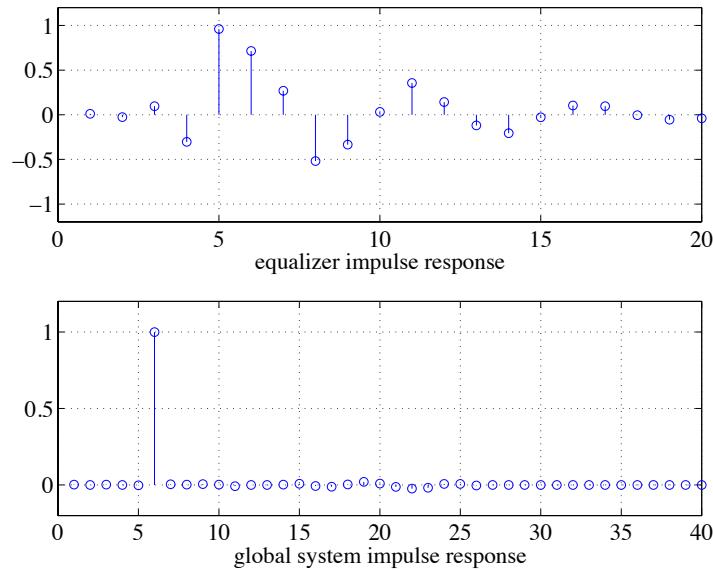


Figure 4.18 Equalizer impulse response and overall impulse response for the channel according to Fig. 4.16.

5 Diversity

Diversity exploits the fact that different means (branches) of propagation are often uncorrelated. It is highly unlikely that two independent propagation paths are in a deep fade at exactly the same time. There are several parameters that differ for a multi-branch transmission. We distinguish the following possibilities:

- space (antenna) diversity
- polarization diversity
- frequency diversity
- time diversity

Fig. 5.1 shows the benefit of diversity in a simple situation. Whereas the single branches show deep fades, the combined amplitude (in this case at any given time the stronger signal is picked) shows much less fluctuations.

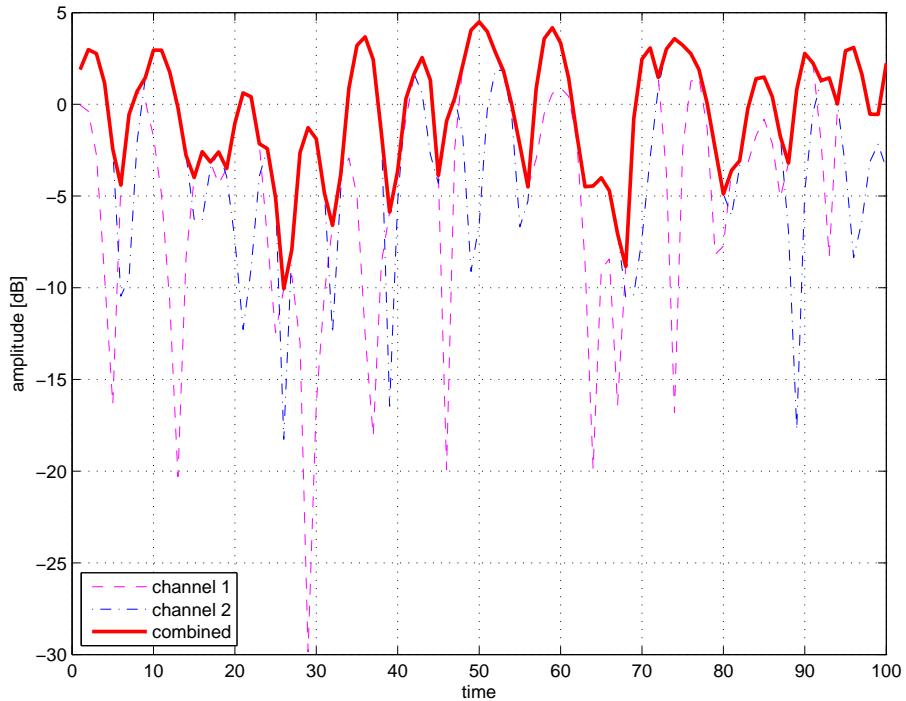


Figure 5.1 Amplitudes in a fading situation, single-branch and diversity combining.

While microscopic diversity techniques help overcome fast fading effects, macroscopic diversity improves large-scale fading due to shadowing. In the following we address the microscopic diversity. In a Rayleigh-fading channel, the pdf of the instantaneous SNR γ_i of the i th branch is given by the χ^2 distribution

$$p(\gamma_i) = \frac{1}{\Gamma} e^{-\frac{\gamma_i}{\Gamma}} \quad \gamma_i \geq 0, \quad (5.1)$$

where Γ is the average SNR of a single branch (thus the subscript 1), as can be seen by evaluating the integral (using partial integration)

$$\begin{aligned}\overline{\text{SNR}}_1 &= \int_0^\infty \gamma_i \frac{1}{\Gamma} e^{\frac{-\gamma_i}{\Gamma}} d\gamma_i \\ &= \int_0^\infty e^{\frac{-\gamma_i}{\Gamma}} d\gamma_i \\ &= -\Gamma e^{\frac{-\gamma_i}{\Gamma}} \Big|_0^\infty = \Gamma.\end{aligned}\quad (5.2)$$

There are three measures (figures of merit) to compare diversity gains:

- average overall SNR
- probability of overall SNR to fall below a threshold (fade-event rate)
- bit-error rate (BER) for a certain digital modulation scheme

The last measure is of course depending on the modulation scheme applied, so we restrict ourselves to the first two measures. For a single-branch situation (no diversity) the fade-event rate, i.e., the probability of a single branch being less than a certain threshold value γ_ϑ is

$$\begin{aligned}\text{Prob}(\gamma_i \leq \gamma_\vartheta) &= \int_0^{\gamma_\vartheta} p(\gamma_i) d\gamma_i \\ &= \int_0^{\gamma_\vartheta} \frac{1}{\Gamma} e^{\frac{-\gamma_i}{\Gamma}} d\gamma_i \\ &= 1 - e^{\frac{-\gamma_\vartheta}{\Gamma}}.\end{aligned}\quad (5.3)$$

In the following we compare the average overall SNR and the fade-event rate for different combining methods of the branches. Note that in order for these values to be true, the branch statistics need to be as uncorrelated as possible, although already correlation values of 0.3 and lower have only little effect.

5.1 Combining Methods

5.1.1 Selection Diversity

Selection diversity refers to the method of picking the branch with the best instantaneous SNR. A block diagram of a possible receiver employing selection diversity is shown in Fig. 5.2. The average overall SNR

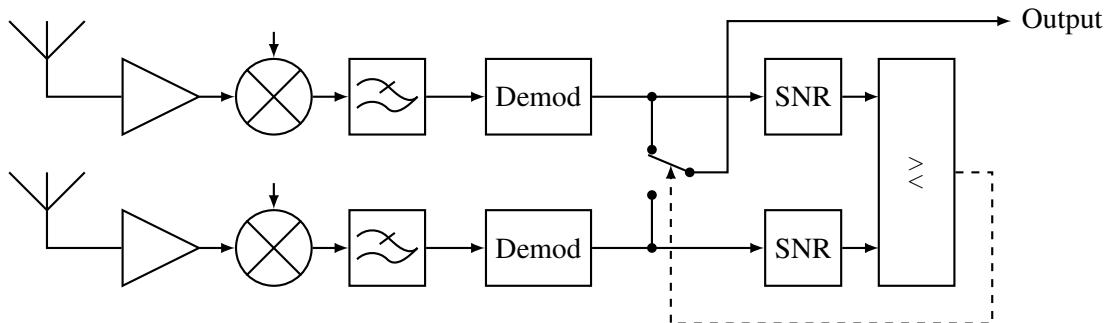


Figure 5.2 Selection-diversity receiver.

of a two-branch selection-diversity scheme in the case where $\Gamma_1 = \Gamma_2 = \Gamma$ can be evaluated according to

$$\begin{aligned}\overline{\text{SNR}}_{2,\text{sel}} &= \int_0^\infty \int_0^\infty \max(\gamma_1, \gamma_2) \frac{1}{\Gamma^2} e^{-\frac{\gamma_1}{\Gamma}} e^{-\frac{\gamma_2}{\Gamma}} d\gamma_1 d\gamma_2 \\ &= \frac{1}{\Gamma^2} \int_0^\infty e^{-\frac{\gamma_2}{\Gamma}} \int_0^\infty \max(\gamma_1, \gamma_2) e^{-\frac{\gamma_1}{\Gamma}} d\gamma_1 d\gamma_2 \\ &= \frac{1}{\Gamma^2} \int_0^\infty e^{-\frac{\gamma_2}{\Gamma}} \left[\int_0^{\gamma_2} \gamma_2 e^{-\frac{\gamma_1}{\Gamma}} d\gamma_1 + \int_{\gamma_2}^\infty \gamma_1 e^{-\frac{\gamma_1}{\Gamma}} d\gamma_1 \right] d\gamma_2.\end{aligned}\quad (5.4)$$

The expression inside the large brackets can be written as (using partial integration for the second integral)

$$\begin{aligned}\int_0^{\gamma_2} \gamma_2 e^{-\frac{\gamma_1}{\Gamma}} d\gamma_1 + \int_{\gamma_2}^\infty \gamma_1 e^{-\frac{\gamma_1}{\Gamma}} d\gamma_1 &= -\Gamma \gamma_2 e^{-\frac{\gamma_1}{\Gamma}} \Big|_0^{\gamma_2} - \Gamma^2 e^{-\frac{\gamma_1}{\Gamma}} \Big|_{\gamma_2}^\infty - \Gamma \gamma_1 e^{-\frac{\gamma_1}{\Gamma}} \Big|_{\gamma_2}^\infty \\ &= -\Gamma \gamma_2 e^{-\frac{\gamma_2}{\Gamma}} + \Gamma \gamma_2 + \Gamma^2 e^{-\frac{\gamma_2}{\Gamma}} + \Gamma \gamma_2 e^{-\frac{\gamma_2}{\Gamma}} \\ &= \Gamma \gamma_2 + \Gamma^2 e^{-\frac{\gamma_2}{\Gamma}}.\end{aligned}\quad (5.5)$$

Finally, we get

$$\begin{aligned}\overline{\text{SNR}}_{2,\text{sel}} &= \frac{1}{\Gamma^2} \int_0^\infty e^{-\frac{\gamma_2}{\Gamma}} \left[\Gamma \gamma_2 + \Gamma^2 e^{-\frac{\gamma_2}{\Gamma}} \right] d\gamma_2 \\ &= \frac{1}{\Gamma} \int_0^\infty \gamma_2 e^{-\frac{\gamma_2}{\Gamma}} d\gamma_2 + \int_0^\infty e^{-\frac{2\gamma_2}{\Gamma}} d\gamma_2 \\ &= \Gamma + \frac{\Gamma}{2} = \frac{3\Gamma}{2}.\end{aligned}\quad (5.6)$$

Generally, it can be shown that the selection diversity with M branches leads to

$$\overline{\text{SNR}}_{M,\text{sel}} = \Gamma \sum_{k=1}^M \frac{1}{k}. \quad (5.7)$$

The average overall SNR in the selection-diversity setup is shown in Fig. 5.3.

The evaluation of the fade-event rate is easy by considering the fact that all M branches are independent and must all have an SNR below a certain threshold value γ . According to Eq. (5.3) we can thus write

$$\text{Prob}(\gamma_{\max} \leq \gamma_\theta) = \left(1 - e^{-\frac{\gamma_\theta}{\Gamma}}\right)^M. \quad (5.8)$$

The fade-event rate in the selection-diversity setup is shown in Fig. 5.4.

5.1.2 Switching Diversity

Although selection diversity is the easiest form of diversity, it has the drawback that we need to monitor all M branches simultaneously. If instead we want to stay with one branch until the performance drops below a certain threshold and then switch, we have a diversity scheme called switching diversity. Its analysis is depending on the threshold and shall be omitted here, a block diagram is shown in Fig. 5.5.

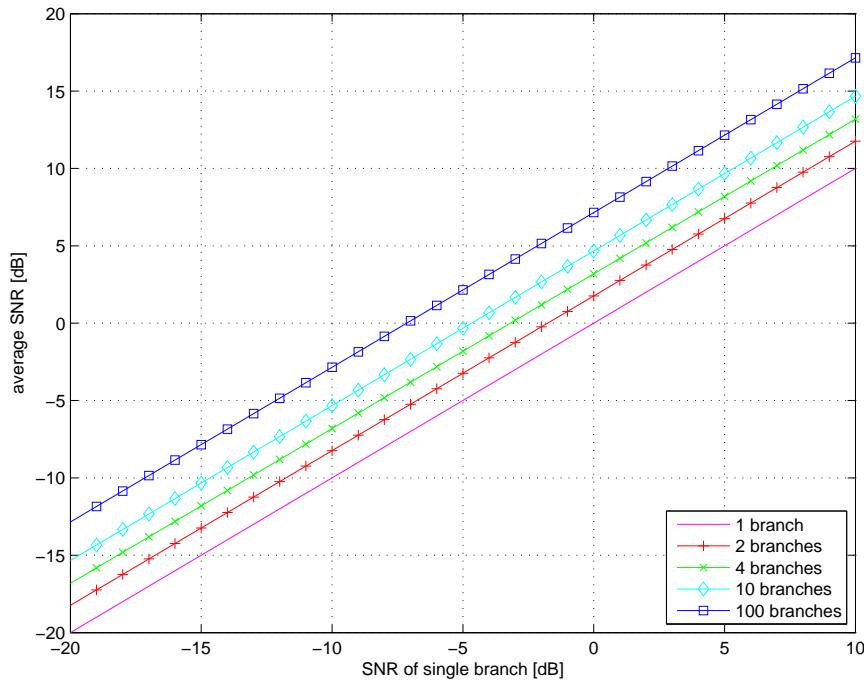


Figure 5.3 Average SNR for selection diversity.

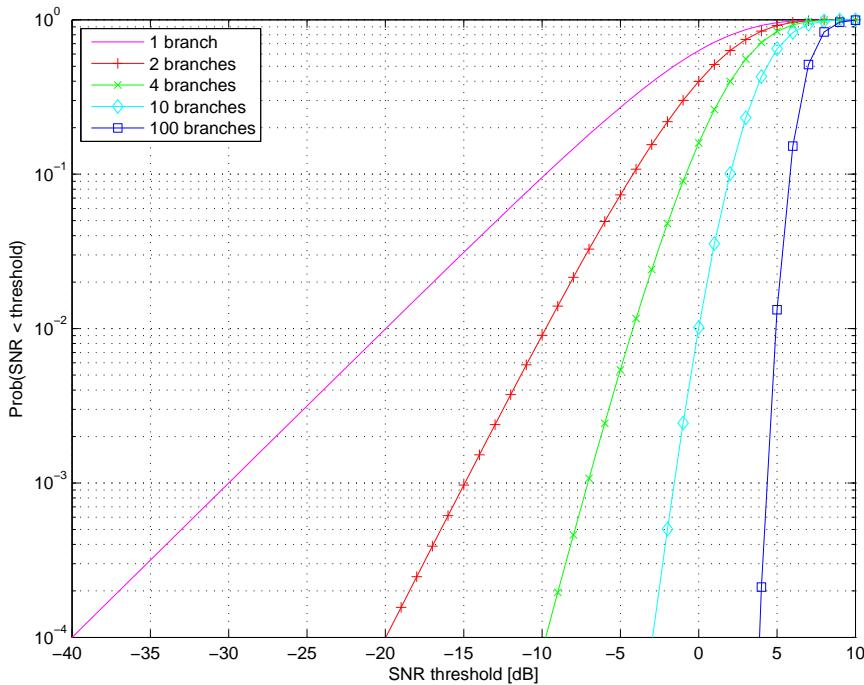


Figure 5.4 Fade-event rate for selection diversity. Average SNR of a single branch $\Gamma = 0$ dB.

5.1.3 Equal-Gain Combining Diversity

As another combining scheme, see Fig. 5.6 for a block diagram, we might just add the signals of the different branches. In a first approach we do this using an equal-gain weight factor for every branch. It can be shown

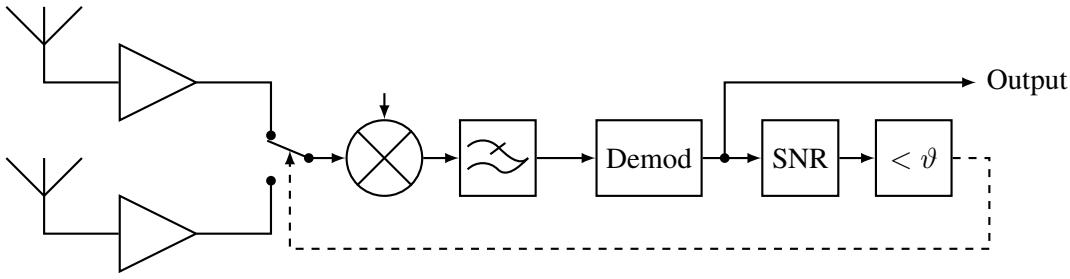


Figure 5.5 Switching-diversity receiver.

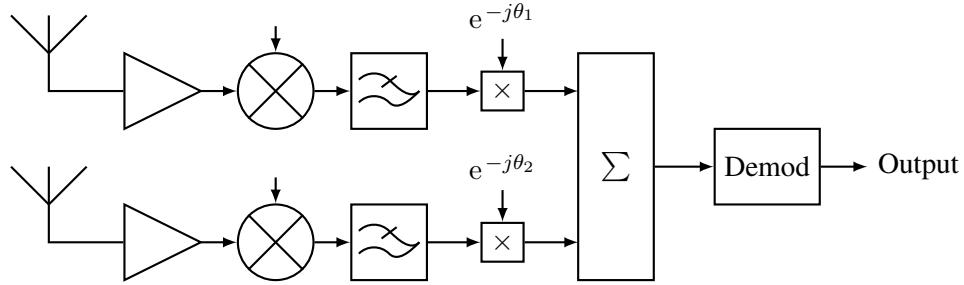


Figure 5.6 Equal-gain combining receiver.

that the instantaneous overall SNR of equal-gain combining for two branches is given by [30]

$$\gamma_c = \frac{\gamma_1 + \gamma_2 + \sqrt{\gamma_1 \gamma_2}}{2}. \quad (5.9)$$

The average overall SNR is thus

$$\begin{aligned} \overline{\text{SNR}}_{2,\text{egc}} &= \int_0^\infty \int_0^\infty \gamma_c \cdot \frac{1}{\Gamma^2} \cdot e^{-\frac{\gamma_1}{\Gamma}} e^{-\frac{\gamma_2}{\Gamma}} d\gamma_1 d\gamma_2 \\ &= \Gamma + \frac{1}{\Gamma^2} \int_0^\infty \int_0^\infty \sqrt{\gamma_1 \gamma_2} \cdot e^{-\frac{\gamma_1}{\Gamma}} e^{-\frac{\gamma_2}{\Gamma}} d\gamma_1 d\gamma_2 \\ &= \Gamma + \frac{1}{\Gamma} \int_0^\infty \sqrt{\gamma_1} \cdot e^{-\frac{\gamma_1}{\Gamma}} d\gamma_1 \cdot \frac{1}{\Gamma} \int_0^\infty \sqrt{\gamma_2} \cdot e^{-\frac{\gamma_2}{\Gamma}} d\gamma_2. \end{aligned} \quad (5.10)$$

Using the substitution $\gamma \triangleq x^2$ we can write the integrals as the one-sided variance expression of a Gaussian distribution

$$\frac{1}{\Gamma} \int_0^\infty \sqrt{\gamma} \cdot e^{-\frac{\gamma}{\Gamma}} d\gamma = \frac{1}{\Gamma} \int_0^\infty 2x^2 e^{-\frac{x^2}{\Gamma}} dx = \frac{\sqrt{\pi}}{2} \cdot \sqrt{\Gamma}, \quad (5.11)$$

resulting in

$$\overline{\text{SNR}}_{2,\text{egc}} = \left(1 + \frac{\pi}{4}\right) \Gamma. \quad (5.12)$$

5.1.4 Maximum-Ratio Combining Diversity

Equal-gain combining diversity is not ideal for the case where one or several branches are in a deep fade. It is optimal to weight the branches proportionally to their amplitude, see also Fig. 5.7, resulting in an overall SNR of [29]

$$\gamma_c = \sum_{k=1}^M \gamma_k. \quad (5.13)$$

The distribution of the overall SNR is thus a simple convolution, which in the case of two branches can be written as

$$\begin{aligned} p(\gamma_c) &= \frac{1}{\Gamma^2} \int_0^{\gamma_c} e^{-\frac{\gamma_1}{\Gamma}} e^{-\frac{(\gamma_c-\gamma_1)}{\Gamma}} d\gamma_1 \\ &= \frac{1}{\Gamma^2} \int_0^{\gamma_c} e^{-\frac{\gamma_c}{\Gamma}} d\gamma_1 \\ &= \frac{\gamma_c}{\Gamma^2} e^{-\frac{\gamma_c}{\Gamma}}, \quad \gamma_c \geq 0. \end{aligned} \quad (5.14)$$

The probability of a fading event can then be evaluated using partial integration

$$\text{Prob}(\gamma \leq \gamma_\vartheta) = \frac{\gamma}{\Gamma^2} \int_0^{\gamma_\vartheta} \gamma \cdot e^{-\frac{\gamma}{\Gamma}} = 1 - e^{-\frac{\gamma_\vartheta}{\Gamma}} \cdot \left(1 + \frac{\gamma_\vartheta}{\Gamma}\right). \quad (5.15)$$

For the general case of M branches, Rappaport [29] states the distribution of the overall SNR as the Chi-square distribution

$$p(\gamma_c) = \frac{\gamma_c^{M-1} e^{-\frac{\gamma_c}{\Gamma}}}{\Gamma^M (M-1)!}, \quad \gamma_c \geq 0. \quad (5.16)$$

Since the distributions are independent and the instantaneous SNRs sum up, the average overall SNR is also the sum of the individual averages

$$\overline{\text{SNR}}_{M,\text{mrc}} = M\Gamma. \quad (5.17)$$

The average overall SNR for maximum-ratio combining diversity is shown in Fig. 5.8.

The probability of γ_c being less than some SNR value γ_ϑ is [29]

$$\text{Prob}(\gamma_c \leq \gamma_\vartheta) = 1 - e^{-\frac{\gamma_\vartheta}{\Gamma}} \sum_k^M \frac{\gamma_\vartheta^{k-1}}{\Gamma^{k-1}(k-1)!}. \quad (5.18)$$

The fade-event rate for maximum-ratio combining diversity is shown in Fig. 5.9.

As an example for the case of two branches ($M = 2$) we compute the weights w_1 and w_2 for the maximum-ratio combining receiver. Assume that the corresponding channel weight parameters are h_1 and h_2 . Then, the signal power can be computed as

$$s = (w_1 \cdot h_1 + w_2 \cdot h_2)^2. \quad (5.19)$$

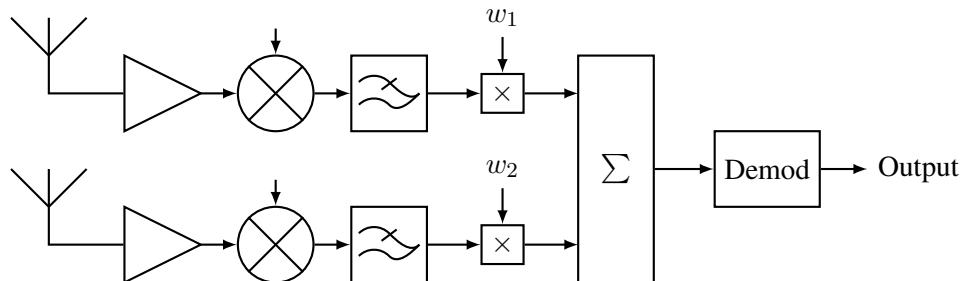


Figure 5.7 Maximum-ratio combining receiver.

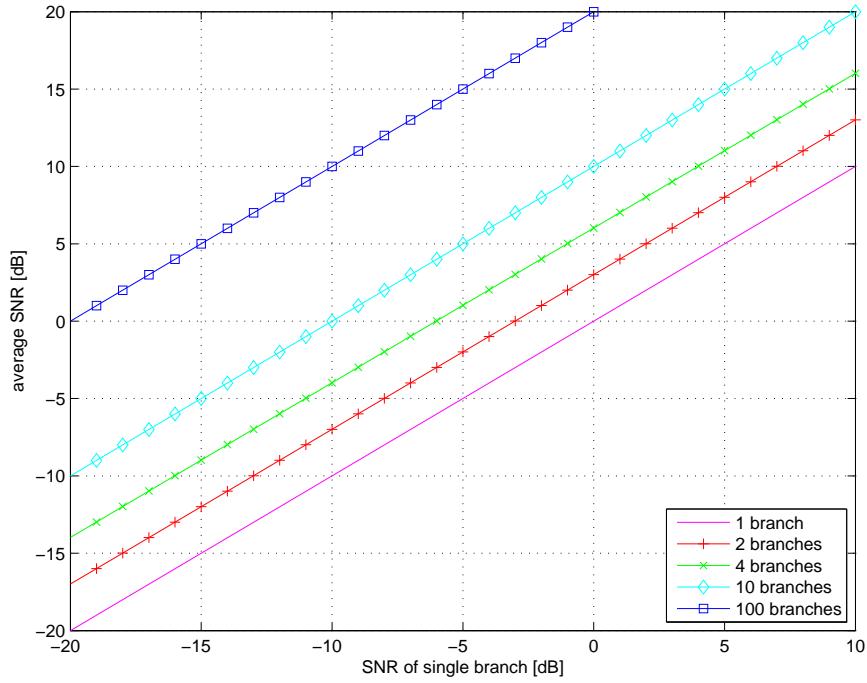


Figure 5.8 Average overall SNR for maximum-ratio combining diversity.

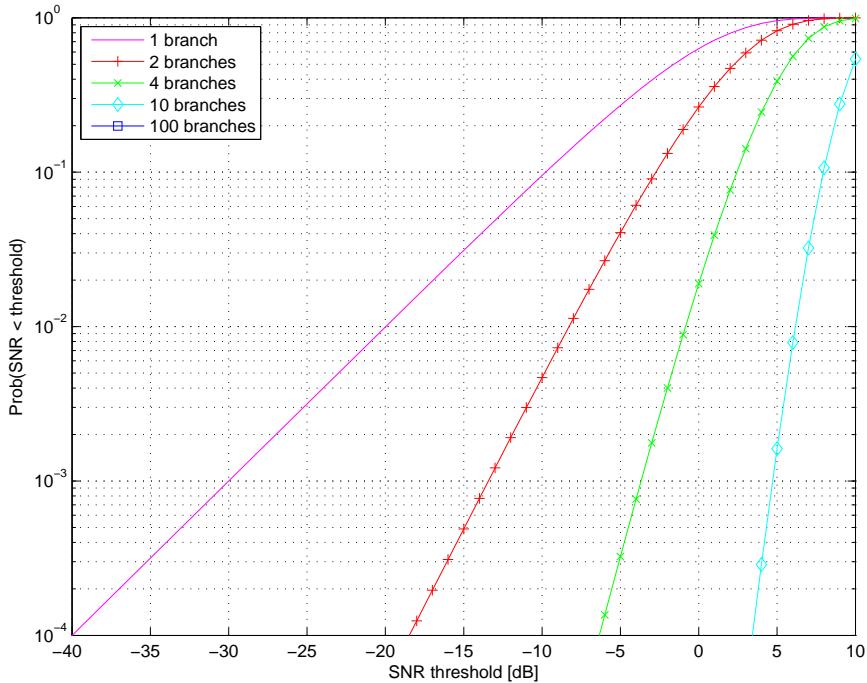


Figure 5.9 Fade-event rate for maximum-ratio combining. Average SNR of a single branch $\Gamma = 0$ dB.

The noise power, originating at the input of the receiver depends on the filtering weights only,

$$n = w_1^2 + w_2^2. \quad (5.20)$$

We now have to maximize

$$\frac{s}{n} = \frac{(w_1 \cdot h_1 + w_2 \cdot h_2)^2}{w_1^2 + w_2^2}. \quad (5.21)$$

Since only the relative values of the weights matter, we can arbitrarily set one weight to one. We choose $w_1 = 1$. Thus, we have the problem of maximizing above equation for the second weight

$$w_{2,\text{MRC}} = \arg \max_{w_2} = \frac{(h_1 + w_2 \cdot h_2)^2}{1 + w_2^2}. \quad (5.22)$$

By differentiating Eq. (5.22) with respect to w_2 and setting it equal to zero we finally get

$$w_{2,\text{MRC}} = \frac{h_2}{h_1}. \quad (5.23)$$

5.2 Performance Comparison of Different Methods

In the previous section two objective functions (average overall SNR and fade-event rate) have been compared with respect to the number of (uncorrelated) branches and plotted for one method at a time. It might be instructive to compare the different methods with each other. In order not to overload the figures, only two branches are considered for each method. Naturally, some diversity methods gain more by increasing the number of branches. The average overall SNR for different diversity methods is shown in Fig. 5.10. The

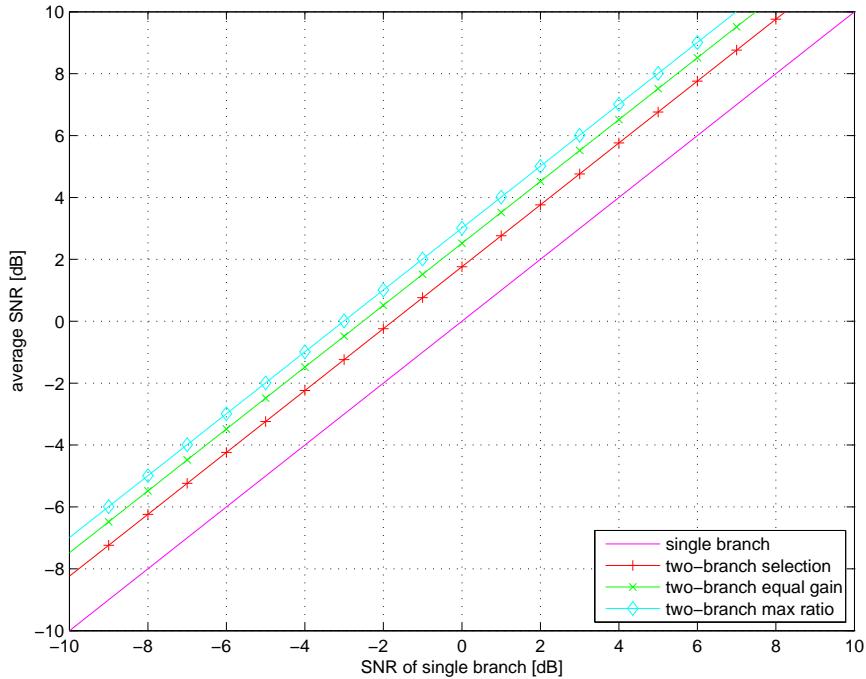


Figure 5.10 Average overall SNR for different diversity methods.

fade-event rate for different diversity methods is shown in Fig. 5.11.

One final comment is appropriate here: Our observations and results are valid in a Rayleigh-fading channel. The results might be different using a Ricean- or Nakagami-fading model.

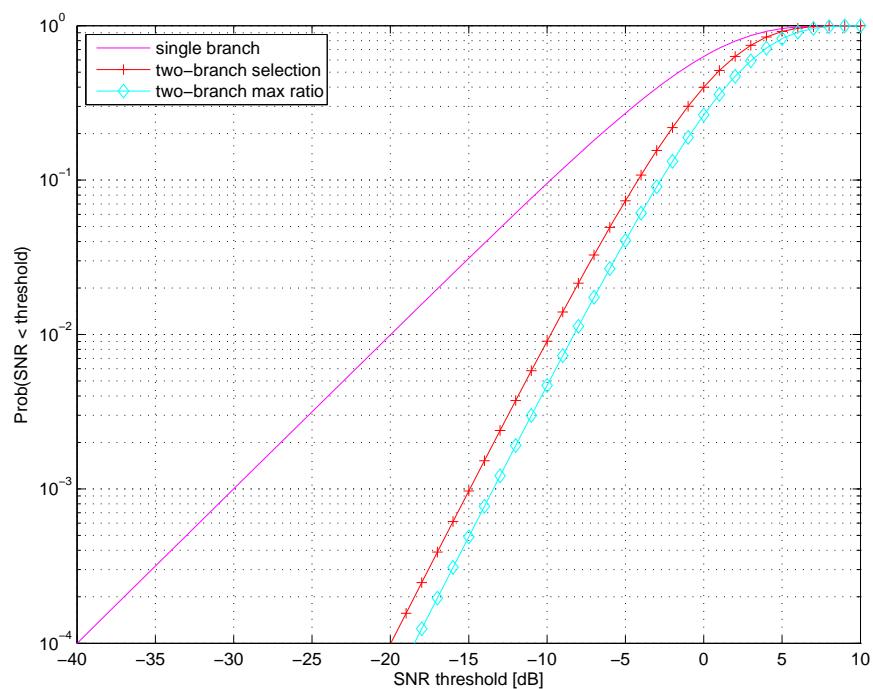


Figure 5.11 Fade-event rate for different diversity methods. Average SNR of a single branch $\Gamma = 0$ dB.

6 CDMA

6.1 Other Channel-Access Systems

For systems where we only have one single transmitter and one single receiver, channel access is not an issue. When several users have to share the medium, however, we need to think about the best use of the radio channel. In the wireless-communication world, past government auctions of frequency bands have clearly indicated the need to make optimum use of the bandwidth available. Traditional approaches use TDMA, FDMA, and CDMA, which shall be explained in due course. Multiple use of common channels is also possible using space division multiple access (SDMA), where the fact is exploited that no two transmitters or receivers have the same physical location. Finally, multiplexing of the channel may be achieved using different polarization schemes (e.g., horizontal and vertical). Such an access scheme is sometimes referred to by the name *polarization division multiple access* (PDMA).

The basic configuration of a wireless communication system with several transmitters sharing the same channels can be visualized by Fig. 6.1. Many variants of this picture are possible: several receivers may share the medium, several transmitters and receivers share the medium etc. In order to avoid collisions and use the channel efficiently, we need to define an access protocol.

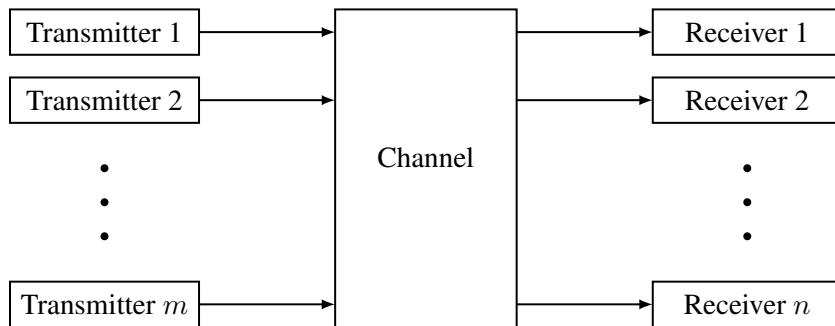


Figure 6.1 Multichannel wireless communication system.

6.1.1 FDMA

In the first days of telephony, an individual pair of wires was needed for every connection between the different switching centers. In the early 1900s, FDMA (frequency division multiple access) was successfully applied to reduce the number of wires. Many radio systems employ some type of FDMA. This is particularly true for the old analog FM systems. But even systems such as GSM (see Fig. 6.2), which are categorized under TDMA, apply some form of FDMA on top of that in both the uplink and the downlink. This means that some channels differ in their frequencies, some in their assigned time slots, and some in both frequency and time slot.

6.1.2 TDMA

Time division multiple access (TDMA) is what people do (if they are polite) during a conversation. At some periods they talk, at others they listen to other people talking. Only one person is accessing the channel at one given time, so to speak.

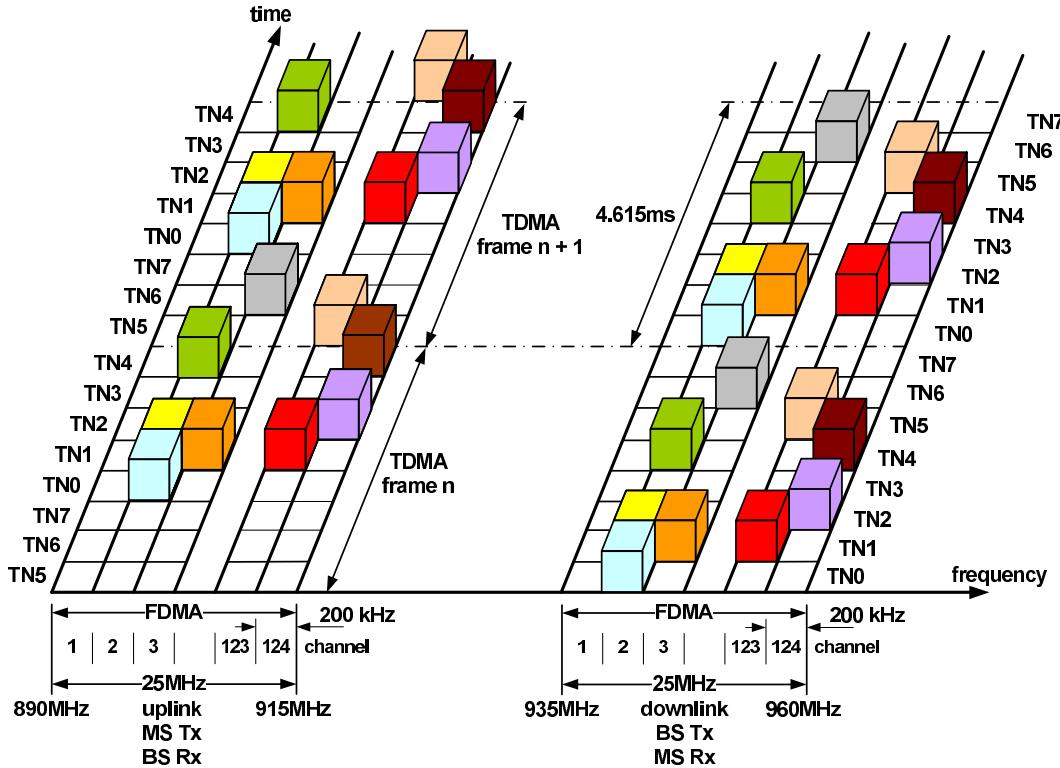


Figure 6.2 TDMA and FDMA as used in GSM.

6.1.3 Capacity-Achieving xDMA

If a medium has to be shared among different users, we can of course divide it up using different schemes (xDMA), where x stands for the domain to be shared. That means that the capacity of the single channel is shared among the users, which is not optimal. Better exploitation can be made if every user has the whole channel at his disposal. What does this mean in terms of capacity or achievable rates for the different users? To see this, let us consider the two-user case for an uplink AWGN channel. If we allow the users to access the channel and detect the signals simultaneously, we have the following rate constraints (remember, for successful transmission the rate needs to be smaller than the capacity):

$$R_1 < \log_2 \left(1 + \frac{P_1}{N_0} \right), \quad (6.1)$$

$$R_2 < \log_2 \left(1 + \frac{P_2}{N_0} \right), \quad (6.2)$$

$$R_1 + R_2 < \log_2 \left(1 + \frac{P_1 + P_2}{N_0} \right). \quad (6.3)$$

The capacity region illustrating the three conditions above is the pentagon shown in Fig. 6.3. The corner points A and B can easily be found by using Eq. (6.3) at point B

$$R_1|_B = \log_2 \left(1 + \frac{P_1 + P_2}{N_0} \right) - R_2|_B. \quad (6.4)$$

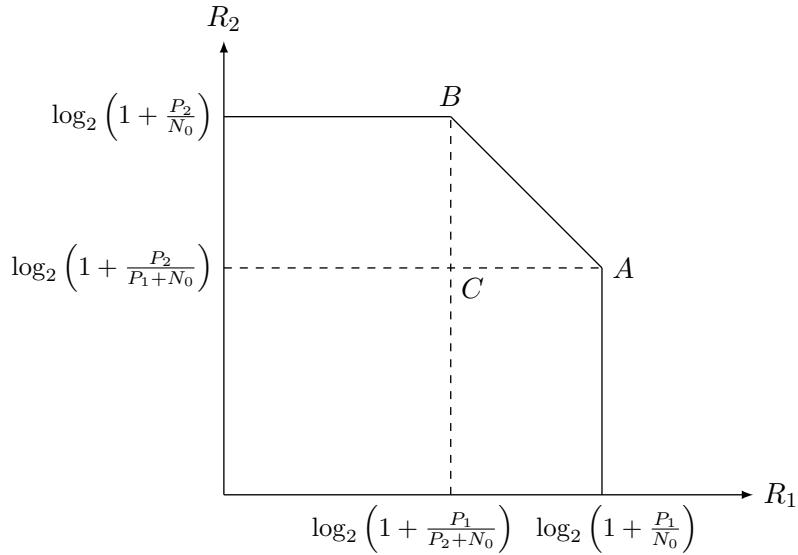


Figure 6.3 Capacity region of a two-user AWGN channel.

When using $R_2|_B$ as the limit of Eq. (6.2) we get

$$\begin{aligned}
 R_1|_B &= \log_2 \left(1 + \frac{P_1 + P_2}{N_0} \right) - \log_2 \left(1 + \frac{P_2}{N_0} \right) \\
 &= \log_2 \frac{1 + \frac{P_1 + P_2}{N_0}}{1 + \frac{P_2}{N_0}} \\
 &= \log_2 \left(1 + \frac{P_1}{P_2 + N_0} \right). \tag{6.5}
 \end{aligned}$$

For symmetry reasons

$$R_2|_A = \log_2 \left(1 + \frac{P_2}{P_1 + N_0} \right). \tag{6.6}$$

The region limited by A, B indicates that we can not have the maximal capacity for the single-user channel simultaneously for both users. They have to back-off with respect to their data rate. The conventional CDMA receiver operates at a point C , which is clearly not optimal. Interpreting Eq. (6.5), we can see that the rate R_1 drops because P_2 is essentially treated as additional noise (it is in the denominator), and vice versa for R_2 and P_1 . In order to maximize the system throughput (sum capacity), the operating point should lie on the AB segment.

However, this cannot be achieved by individually receiving the two channels and regarding the undesired one as additional noise. We can only get the maximum sum capacity if we apply so-called successive interference cancellation (SIC), which works as follows. If we want to target our capacity at point A , user 2 is detected first, with the signal of user 1 acting as additional noise and thus limiting R_2 . After detecting user 2's signal, it might be subtracted from the joint signal, leaving the detection of user 1 with only the N_0 noise contribution. User 1 can therefore operate at its single-user bound and the sum capacity (being at point A) lies on the AB segment, the maximum for the sum capacity. Nothing is wasted. We can reach other points on the AB segment, depending on the weights we put on each user's rate. Not every weight might be equally fair to the users, but they all achieve the sum capacity. We call such a receiver also a multi-user receiver, because it is 'aware' of other users.

If we have unequal single-channel bounds, essentially because $P_1 \neq P_2$, the situation is slightly different, see Fig. 6.4. Operating at point C , again, the conventional single-user CDMA receiver does not maximize the sum capacity, this time with a heavy disadvantage to the user 1. This is because the power of user 2 is so high. Backing off with P_2 helps user 1 as can be seen, because we can now operate at the symmetric point D . However, if we 'solve' the near-far problem by attenuating the strong signal (with the user whose single-channel bound would suggest a high capacity), we lose in the sum capacity, because we move to the point D . Again, using a multi-user receiver we are much better off for both users by aiming at point A .

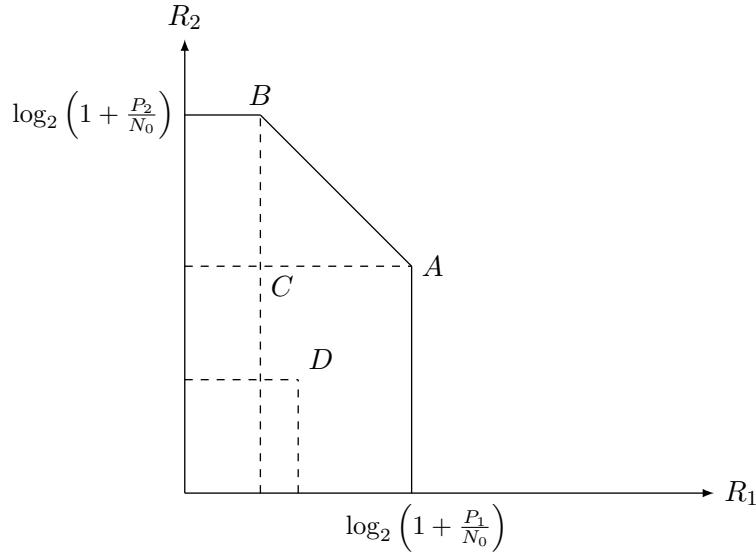


Figure 6.4 Capacity region of an unequal two-user AWGN channel.

The question remains as to whether it is at all possible to achieve the maximum sum capacity for an orthogonal multiple-access scheme (no matter whether CDMA, FDMA, or TDMA) with single-user detection only. To see this, we carry out a numerical simulation, where we weight the access to the channel by the parameter α , which means that for a proportion of α user 1 has access, and for a proportion of $(1 - \alpha)$ user 2 has access to the channel. Each user might therefore concentrate his power to the time of access but only gets partial access, which means that the capacity expressions of the users become

$$C_1 = \alpha \log_2 \left(1 + \frac{P_1}{\alpha N_0} \right), \quad (6.7)$$

$$C_2 = (1 - \alpha) \log_2 \left(1 + \frac{P_2}{(1 - \alpha) N_0} \right). \quad (6.8)$$

We carry out the simulation twice, once with equal SNR for both users and once with a poorer SNR for user 1. The results in Fig. 6.5(a) confirm what we have elaborated upon just previously. A power-control scheme does not change the situation because the power are the same initially. If we start with different channels, we have to consider Fig. 6.5(b). Here, we see that in the single-user detection, the situation for user 1 is really poor. It gets better for user 1 once we apply power control. This, however, makes the sum capacity the lowest of all modes. Orthogonal access improves the situation. As we can see from comparing the dashed curves to the capacity boundary, capacity is achieved only at one point, namely where

$$\alpha = \frac{P_1}{P_1 + P_2}. \quad (6.9)$$

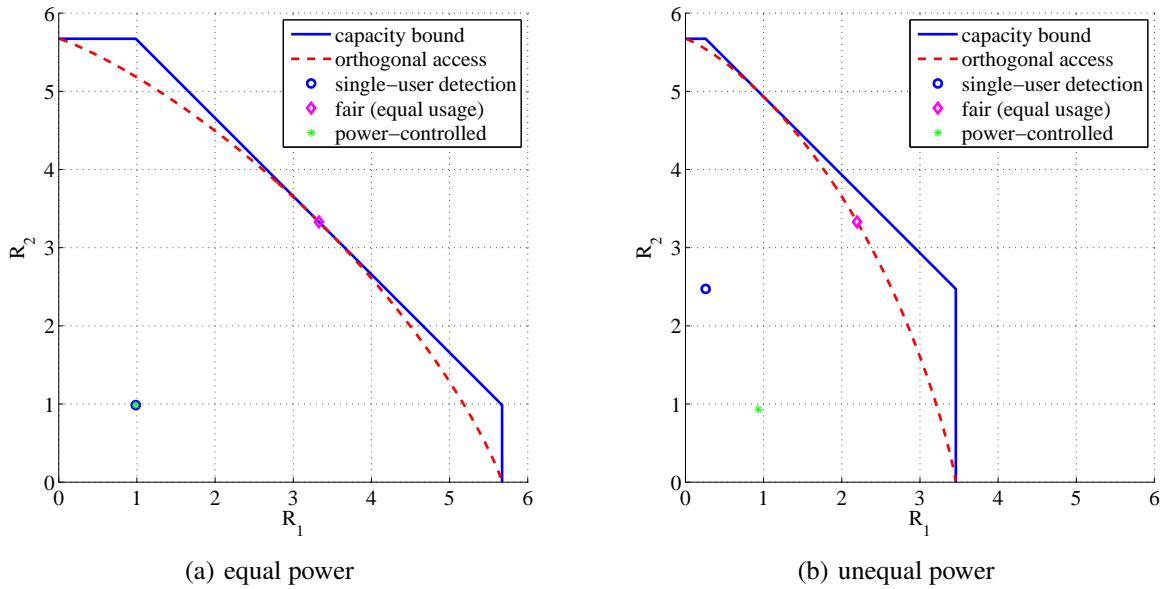


Figure 6.5 Simulation of capacity for a two-user AWGN channel.

This point is highly unfair to the user with a low SNR, but maximizes the capacity of the system (sum capacity). Here, a successive interference cancellation (SIC) system can be made fair and achieves capacity at the same time.

6.2 Introduction to CDMA

CDMA (code division multiple access) is a scheme where all transmitters use the same time and frequency allocation. To distinguish the different users, a code is assigned to each of them. The code is then used to spread a small information bandwidth over the much wider channel bandwidth. CDMA is thus a spread-spectrum (SS) technique. To spread the signal, we can either hop on different frequencies, see Fig. 6.6(a), where we speak of frequency-hopping spread spectrum (FH-SS), or we can directly multiply the data sequence with a much higher chip sequence, see Fig. 6.6(b), which we call direct-sequence spread spectrum (DSSS). Sometimes, chirp spread-spectrum, where the carrier frequency is swept over a certain range, is considered as a third form of spread spectrum.

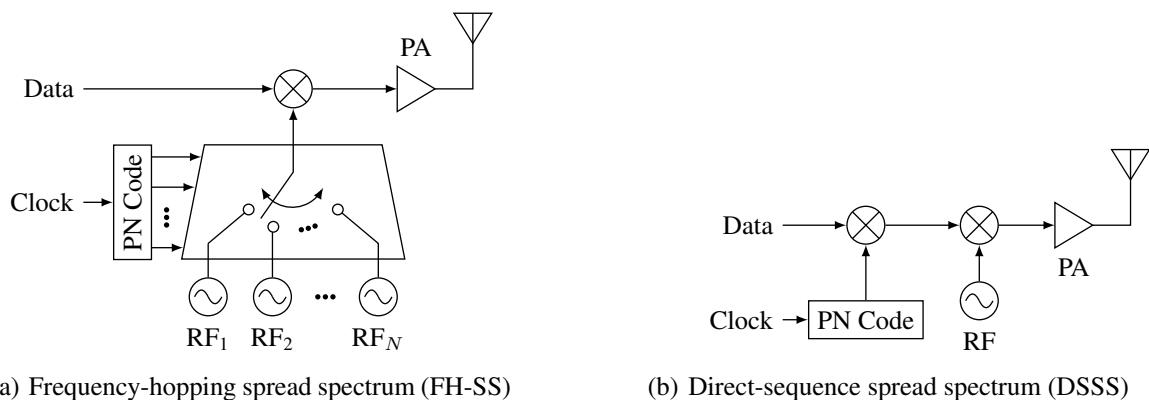


Figure 6.6 Spread spectrum architectures.

6.3 Benefits and Drawbacks

There is an ongoing argument whether CDMA systems provide more space to transfer information compared to TDMA and FDMA. Whereas the fundamental quantities that define the channel capacity are independent of the access scheme, there are some differences. If few users use a channel, those who do use the channel get a very good SNR performance in CDMA. In TDMA and FDMA the SNR is independent of the network load. On the other hand, in CDMA, if users are at different distances from a base station, the channel is predominantly assigned to the strong user. The weak-signal user gets a very low SNR. This scenario is called the near-far problem associated with spread spectrum systems. On the other hand, spread-spectrum techniques are suitable for 'blind' assignment of frequency resources to many different users in an uncoordinated way. They do not need assigned time slots and the like. An example is the popular ISM band at 2.4 GHz, where systems like Bluetooth and W-LAN coexist. Further advantages of CDMA are:

- Privacy: Interception is not possible without knowing the code.
- Fading immunity: The resistance against deep fades in the frequency spectrum is much higher.
- Jamming resistance: Narrowband interferer are spread over a wide band at the receiver where only the wanted band will be concentrated back to the signal bandwidth.
- Low spectral density: At no frequency there will be strong emissions. Moreover, the signal may be hidden in thermal noise.
- Higher positioning accuracy: The wider bandwidth allows a much more accurate localization of a signal.

At the downside we often have higher complexity in the receiver, since synchronization is a more important issue. Furthermore, we need a very sophisticated power-control algorithm, as we will see. In order to collect all energy in a channel with echo (multipath propagation), very often a RAKE receiver as shown in Fig. 6.7 is employed in order to receive a CDMA signal. Hence, the approach of trying to invert the channel is turned into one of exploiting every path contribution.

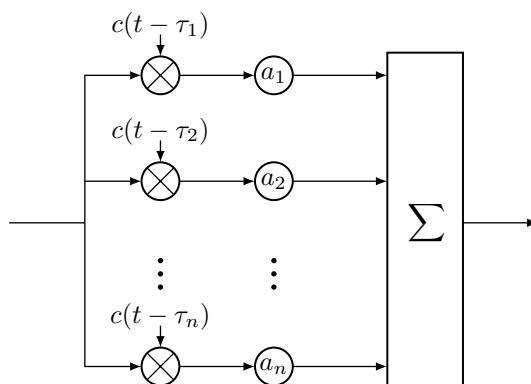


Figure 6.7 RAKE receiver structure.

6.4 History

Spread-spectrum technologies go back as far as to Tesla, who obtained two patents in the field around 1900. Both patents were based around changing the carrier frequency. One of the rather colorful stories in the history of spread-spectrum technology is the fact that in 1942 the actress Hedy Lamarr and the composer George Antheil received a patent called "Secret Communication System". In the middle of the last century, progress in the field was mainly driven by the US army, who tried to hide signals in the spectrum from

eaves-dropping. Later in the 70s, the book by Dixon [7] caught interest. The San Diego based company Qualcomm is certainly one of the big proponents of CDMA in the commercial market. In 1993, IS-95 was standardized. For Europe, the most important CDMA system is the 3rd generation cellular system, UMTS.

6.5 Spread-Spectrum Operation

The three possibilities of spreading the spectrum have been mentioned in the previous chapter. It is important to note that the spreading signal is independent of the information signal, hence, FM is not considered a spread-spectrum technique, although the spectrum can be considerably wider than its original audio spectrum.

CDMA is a spread-spectrum technique. Spread-spectrum systems spread the signal power over a much larger bandwidth than the modulated spectrum alone would do, resulting in a lower power spectral density, sometimes even below the thermal noise floor. The basic configuration of transmitters and receivers operating in spread-spectrum techniques can be seen in Fig. 6.8. The resulting signals can be viewed in Fig. 6.9.

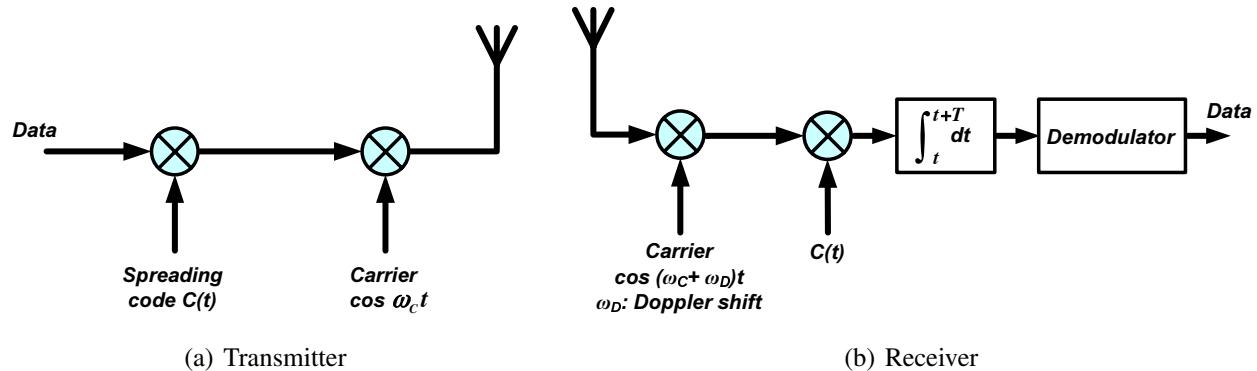


Figure 6.8 Spread-spectrum transmitter and receiver configurations.

An important parameter of a CDMA system is the so-called *spreading gain* G , which relates the SNR after the spreading block to the SNR before it. The spreading gain is depending on the ratio of the respective bandwidth, which are inversely proportional to the respective bit/chip¹ duration T_d and T_c , respectively. In mathematical terms,

$$G_{\text{dB}} = 10 \log_{10} \frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}} = 10 \log_{10} \frac{B_{\text{ss}}}{B_d} = 10 \log_{10} \frac{T_d}{T_c}, \quad (6.10)$$

where B_{ss} and B_d are the spreaded bandwidth and the original bandwidth of the data stream, respectively.

6.6 Synchronization

One of the hard jobs within a CDMA receiver is the synchronization, i.e., finding the right moment in time to despread the signal. Most CDMA receivers operate in two different modes: the acquisition mode, where a coarse synchronization process searches for the signal, and a tracking mode, in which the signal found is followed in order to track fluctuations of carrier and code phase.

¹Since the spreading code does not bear any information, the smallest time unit is usually called a chip, rather than a bit.

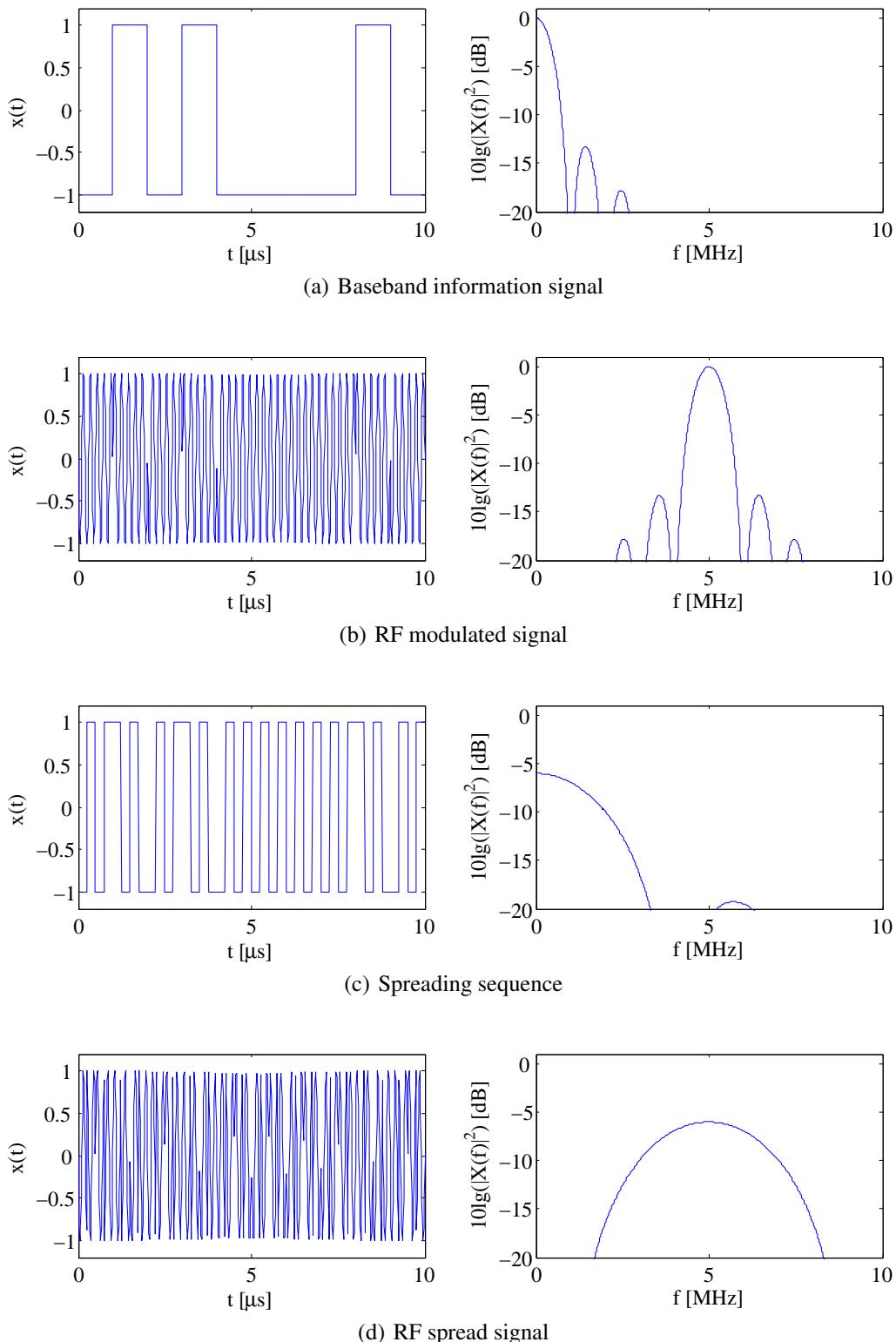


Figure 6.9 Time-domain (left) and frequency-domain signals (right, in dB) at different stages of a spread-spectrum system. Note that the spreading factor in this chosen example is 4, thus there is a decrease of 6 dB in the power density maximum.

6.6.1 Acquisition

The acquisition is the tiresome job of searching a three-dimensional space for the correlation peak. One dimension is the user code (at least if the receiver has no *a-priori* knowledge). The second dimension is the frequency space, since an unknown Doppler shift due to user movements and an inaccurate frequency reference in the receiver produce a high frequency offset uncertainty. The third dimension is the code phase. The generic correlator channel is shown in Fig. 6.10. Note that for both the I and the Q channel, three equidistantly spaced code phases are correlated. They are denoted by E (early), P (prompt), and L (late), respectively. Later in the explanation of the tracking mode, we will see that the code phase is adjusted such as to give equal early and late correlation and maximum prompt correlation.

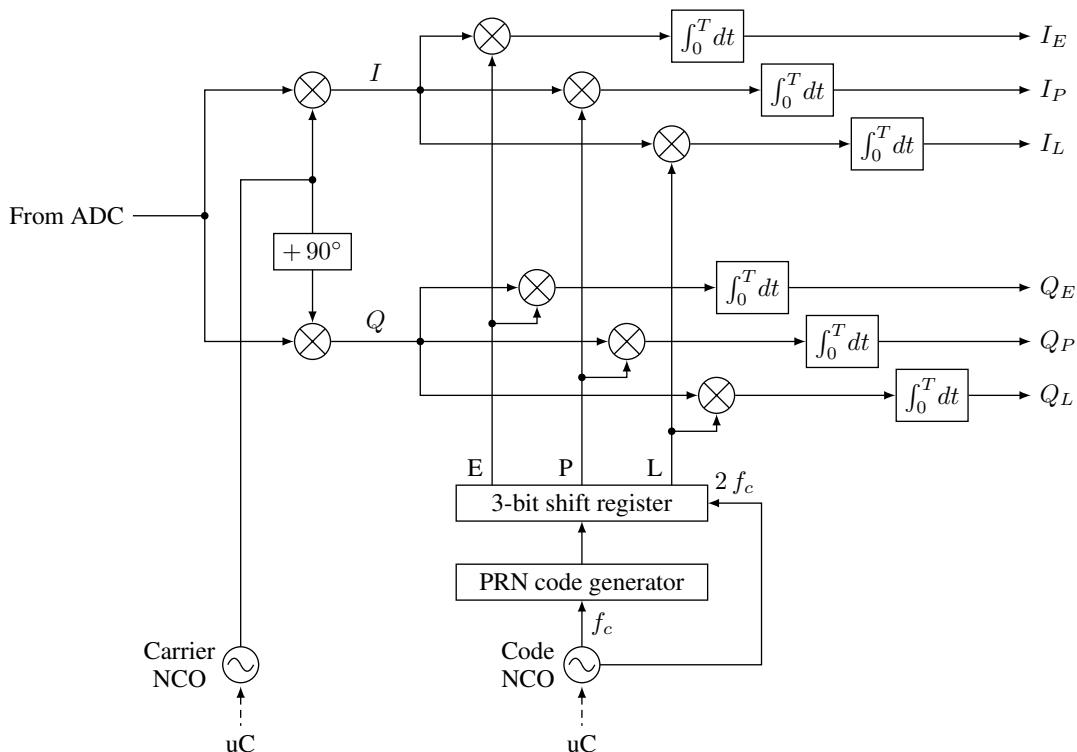


Figure 6.10 Generic correlator channel.

6.6.2 Tracking

After the frequency and code phase parameters have been obtained from the acquisition process, they are passed on to the tracking mode. Three basic parameters are tracked, the code phase, the carrier frequency, and the carrier phase. The latter two are often tracked using a PLL, whereas the code phase is tracked using an independent loop, usually implemented with a DLL (delay-locked loop), as shown in Fig. 6.11. Let us have a look at the DLL in more detail. The DLL is based around the correlation triangle (at least for binary sequences, which is usually the case), whose peak occurs for exact alignment of the code sequences. Since the peak's height itself depends on the signal strength, one correlation result alone can never reveal the direction of misalignment. In fact, the correlation result is evaluated at two positions before and after the main peak, usually spaced by one chip duration (denoted by EL spacing in Fig. 6.12). The correlation window are referred to as early and late-gate, respectively. By means of a control loop, the results from early and late-gate correlation are balanced by shifting all three correlator times (early, late, and prompt) together forth and back. When early and late-gate correlations are equal, the prompt correlation will be maximum,

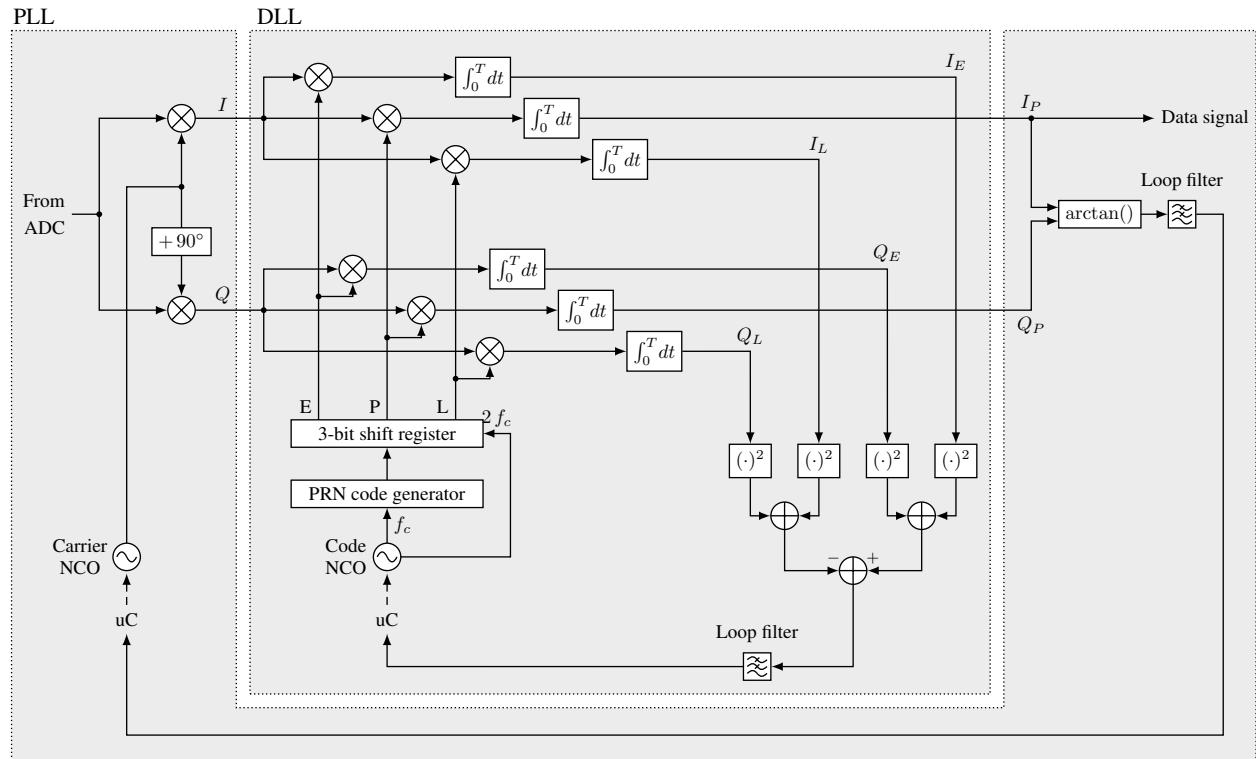


Figure 6.11 Code phase tracking loop (DLL) and carrier phase tracking loop (PLL) of a CDMA tracking receiver.

and its timing alignment indicates the correct PRN position. Fig. 6.12 illustrates this behavior. In subfigure (a) correlation occurs too early, i.e., the result of the early correlation (E) is smaller than that of the late correlation (L). As a result, the late correlation result is higher and will retard the correlation times of all three positions (E, P, and L). The contrary is the case in subfigure (b). If the correlation time is adjusted correctly, the correlation results of E and L are equal as can be seen in subfigure (c) and no further changing of the correlation time is necessary.

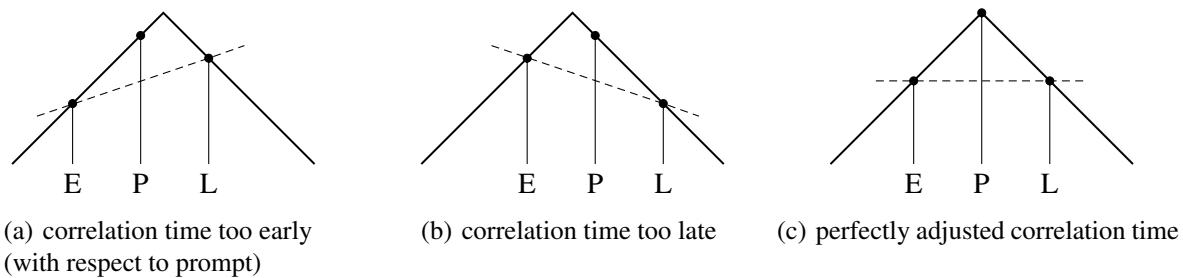


Figure 6.12 Early/late-gate correlation of the DLL.

6.7 Spreading Codes

Spreading codes should ideally be orthogonal. They are usually binary in practical systems. Fig. 6.13 illustrates the principle of superposition of two messages and the following separation using two orthogonal codes. There are a lot of different codes that are typically used for CDMA operation. They all have

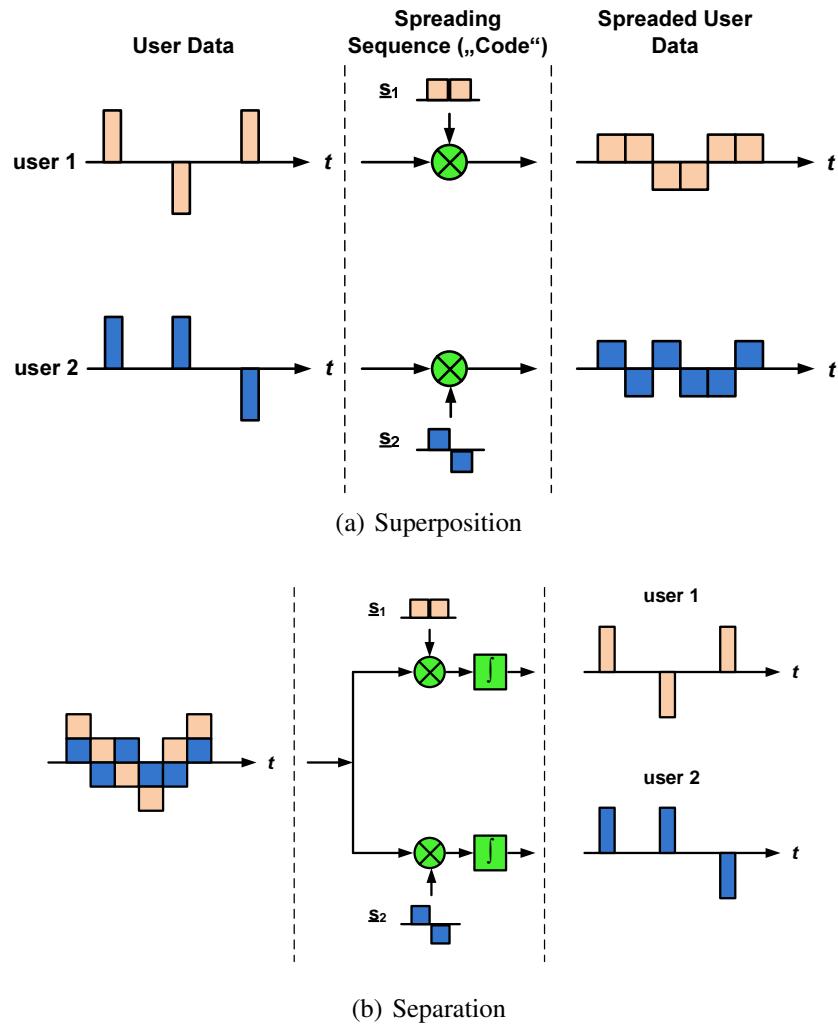


Figure 6.13 Superposition and separation of two independent messages in a CDMA receiver.

advantages and drawbacks. The requirements to the spreading codes are listed below:

- easy to generate
- look random
- long enough to give high spreading gain
- difficult to decode (military)

The following codes are useful within a CDMA system:

- m-sequences (e.g. Glonass)
- Gold codes (e.g. GPS)
- Kasami codes (e.g. scrambling code in UMTS)
- Barker codes (e.g. 802.11b)
- Walsh codes (e.g. IS-95)

6.7.1 M-Sequences

Maximum-length sequences have very interesting autocorrelation properties: they have one peak for exact alignment, and the same low level for misalignment of one or more chips. Since the sequences always consist of an odd number of chips, the autocorrelation functions can never be zero. If the autocorrelation function is taken for a relative shift of below ± 1 chip (just a slight misadjustment), we can observe the correlation triangle. In summary, for a code period of $2^n - 1$, for perfect alignment we have a peak of $2^n - 1$, and for misalignment outside one chip an autocorrelation value of -1 . For an example of an m-sequence of length 1023, the autocorrelation function is shown in Fig. 6.14.

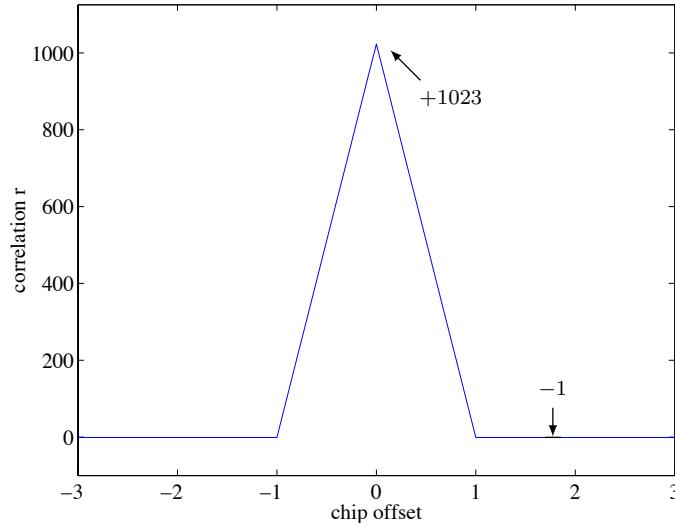


Figure 6.14 Autocorrelation function of m-sequence ($n = 10$).

Similar to different but equivalent architectures of IIR filters, there are two well-known forms of linear feedback shift registers: the Fibonacci form and the Galois form. They are shown in Fig. 6.15.

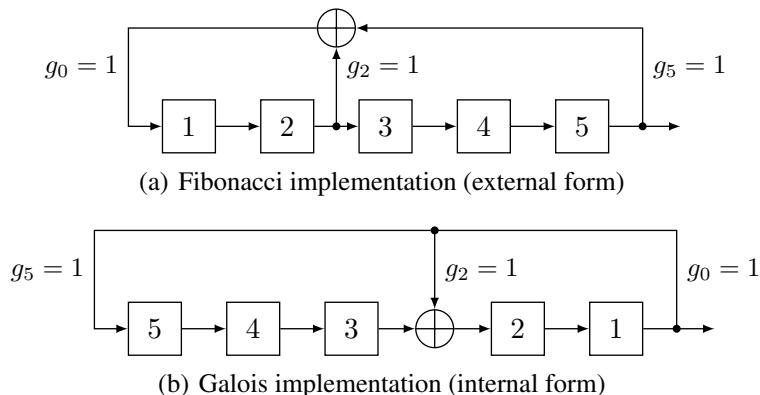


Figure 6.15 Example of equivalent forms corresponding to a feedback polynomial of $x^5 + x^2 + 1$.

The description of the feedback takes different forms in the literature. The feedback used in Fig. 6.15 can be written as $[5, 2]$ or as a polynomial $x^5 + x^2 + 1$. Note that some authors reverse the order of the feedback shift register, so that the feedback polynomial needs a different interpretation. In order to generate a maximum-length sequence, the feedback polynomial needs to be irreducible. Also, the number of feedback taps is always even.

Although the autocorrelation properties of an m-sequence code are very good, only few available distinct codes of a certain length usually prohibit the use of such a code in practical systems. When two or more m-sequences are combined into one new code we get new families of codes (Gold or Kasami codes) at the prize of suboptimal correlation properties.

6.7.2 Gold Codes

Gold codes are the composition of two maximum-length sequences of the same length. Compared to maximum-length sequences alone, Gold codes have only limited autocorrelation and crosscorrelation capabilities. Apart from the main peak (which is 1023 for this example), the correlation functions show rather high side-peaks of +63 and -65, respectively, which results in a dynamic range of

$$20 \log_{10} \frac{1023}{65} = 23.9 \text{ dB}. \quad (6.11)$$

Autocorrelation and crosscorrelation functions are shown in Fig. 6.16. False locking to sidepeaks of the correlation function can now occur more frequently. However, there are many more Gold codes of a certain length than there are m-sequences. Hence, the multiple-access problem might be more easily solved by these code family. Fig. 6.17 shows a typical Gold code generator.

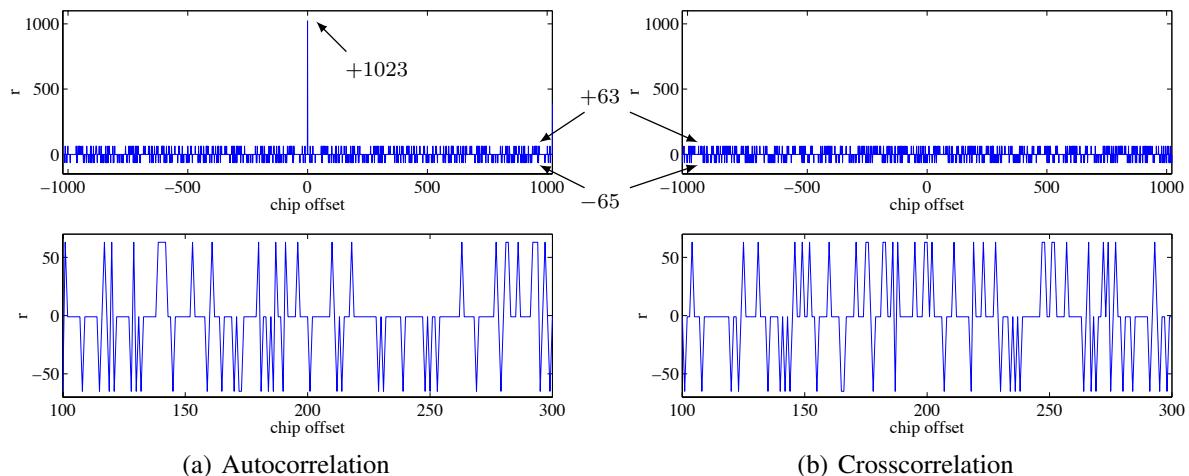


Figure 6.16 Typical properties of the Gold codes with two different zoom sizes.

6.7.3 Walsh Codes

Of importance are Walsh codes, codes which can be constructed using Hadamard matrices. Hadamard matrices, also called Walsh matrices, date back to Sylvester (1867), are square, and contain +1 and -1 only in such a way that all their rows are mutually orthogonal. Hence, a $n \times n$ Hadamard matrix suffices

$$H_n H_n^T = n I_n. \quad (6.12)$$

The most simple Hadamard matrix is

$$H_1 = [1]. \quad (6.13)$$

The next larger Hadamard matrix is

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (6.14)$$

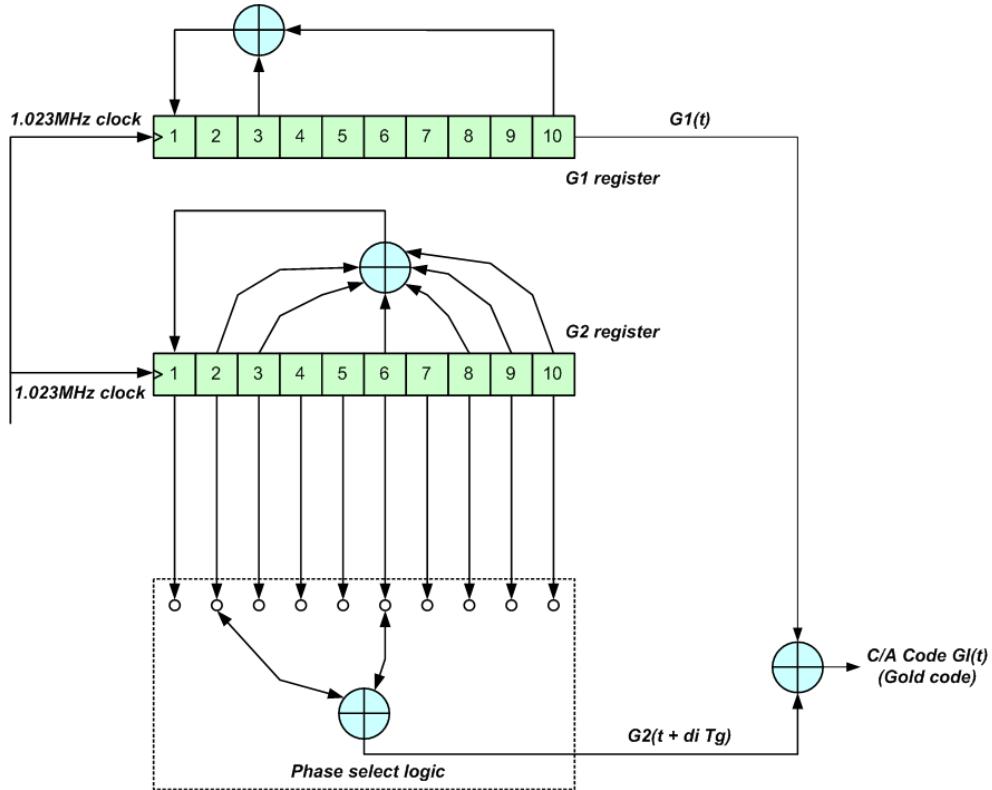


Figure 6.17 Gold code generator.

Hadamard matrices of order $2n$ can be constructed using

$$H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}. \quad (6.15)$$

One of the problems is that orthogonality of such codes is only preserved at perfect synchronization.

6.8 The Near-Far Problem

Whereas in TDMA and FDMA systems the different levels received from users sitting at different distances from the base stations is not considered a problem, this circumstance produces some problems in a CDMA system. Since every user is received as yet another noise contribution, those who are close to the station come in very 'loud' compared with users further away. This situation is called the *near-far problem* and has to be accounted for by using fast power control². Power control is mandatory in order not to allow the situation that one strong user masks a weak user. By employing power control, the system enforces the signals arriving at the base station to an equal power for every user, hence minimizing the interference within the system.

6.9 Examples of CDMA Systems

CDMA has been the system of choice for many communication systems, particularly for cellular mobile phone systems. In Europe, the most prominent CDMA communication system is UMTS, the *Universal*

²In UMTS, for example, the uplink signal power level is changed at a frequency of 1500 Hz!

Mobile Telecommunications System, also known as 3G, as it is the third generation of cellular mobile communication system based on the GSM standard, often referred to as 2G. Extensions to the GSM system, such as HSCSD, GPRS, or EDGE are often called 2.5G (and 2.75G) systems, since they should bridge from 2G to 3G systems.

6.9.1 IS-95

Long before UMTS came along, the US already had a CDMA-based mobile communication system in place. *Interim Standard 95*, better known simply as IS-95 or by its proprietary name *cdmaOne*, was the first ever CDMA-based digital cellular technology. It was developed by Qualcomm and later adopted as a standard by the Telecommunications Industry Association (TIA) in 1995.

IS-95 is a 2G mobile telecommunications standard, which similar to the original GSM in Europe, was used as the default wireless system in the 800 and 900 MHz bands. The subbands are 1.25 MHz wide. To obtain this bandwidth, a 307.2 kbps bitstream is exored with a PRN code running at 1.228 Mbps and a period length of $2^{42} - 1$.

Eventually, it was supplanted by IS-2000, better known as *CDMA2000*, the 3G standard developed for North America and South Korea, in a parallel evolution path to UMTS in Europe, see Fig. 6.18. CDMA2000 combines CDMA and TDMA techniques for enhanced data rates (then called Evolution-Data Optimized, EV-DO) and has been adopted with similar if not greater acceptance as UMTS (or HSPA, its evolved packet access for high-speed data rates) in Europe. The planned fourth-generation (4G) successor UMB (Ultra Mobile Broadband) was abandoned by Qualcom in 2008, opening the way for LTE to become the first global communication standard.

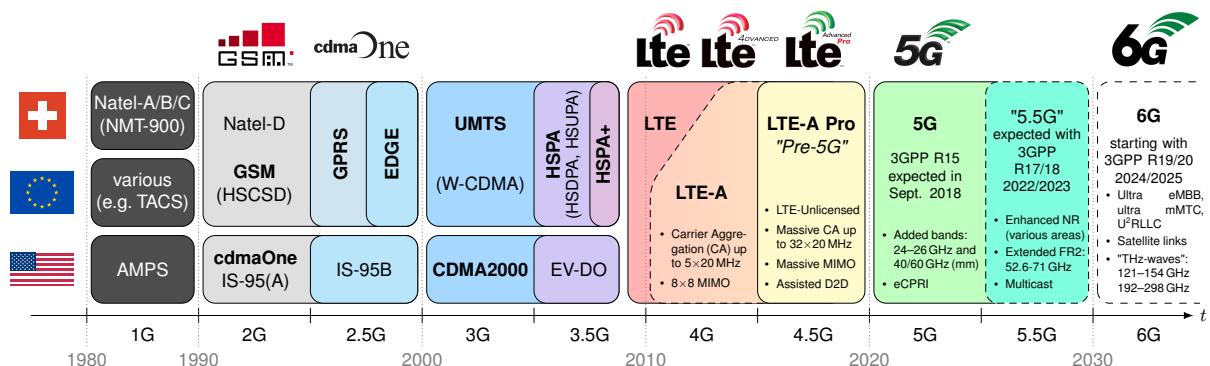


Figure 6.18 Evolution paths of the cellular mobile communication technologies³ (generations).

6.9.2 UMTS

The most prominent example of a CDMA-based cellular communication system in Europe is UMTS (Universal Mobile Telecommunications System), launched by the European Commission in 1989. Originally

³eCPRI stands for *Evolved Common Public Radio Interface* and denotes the highly desired standard interface to all communication systems allowing direct interconnections between systems of different manufacturers (e.g. Ericsson, Nokia, Huawei, and NEC), which is a fundamental requirement for future broadband communication technologies. The terms NaaS and XaaS stand for *Network as a Service* and *Everything as a Service*, respectively, and referring to high-level service models for general connectivity. NaaS involves the optimization of resource allocations by considering network and computing resources as a unified whole; XaaS goes one step further and also incorporates infrastructure and software (thus, *everything*) as parts of the unified system. What these ideas actually entail is still quite fuzzy and will have to be defined as these evolved communication standards are developed.

planned to cover the entire globe, it has mostly been adopted in Europe, Australia and Africa as well as in parts of Asia. Since different 2G systems existed, the migration scenarios to 3G were different depending on the region of the world. The possible evolutions into 3G standards are shown in Fig. 6.18. The current 3rd-generation system in Europe is UMTS, which consists of two modes:

- W-CDMA, which originally occupied the frequency bands from 1920–1980 and 2110–2170 MHz in FDD (Frequency Division Duplex) for uplink and downlink.
- TD-CDMA, which was planned to occupy the frequency bands from 1900 MHz–1920 and 2010–2025 MHz in TDD (Time Division Duplex).

Some key parameters of UMTS are given in Table 6.1.

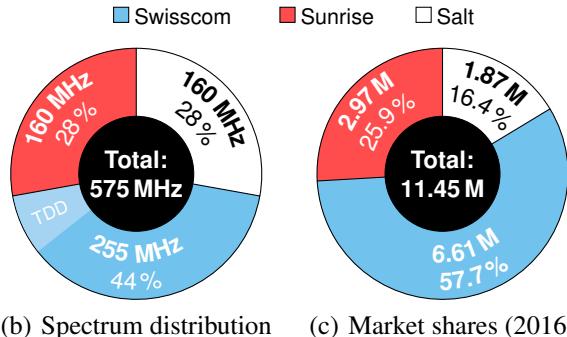
Standard	Chip rate	Modulation scheme	Bandwidth	Tx filter
WCDMA	3.84 Mchip/s	QPSK	5 MHz	RRC, $\rho = 0.22$

Table 6.1 Modulation parameters of UMTS.

Table 6.2(a) lists the spectrum allocations of all three cellular communication providers in Switzerland, which resulted from the spectrum auction in early 2012 and are valid up to 2028. It is interesting to note (albeit not entirely surprising) that the percentages of the allocated spectra do not directly correspond to the market share percentages of each mobile network provider, as the diagrams in (b) and (c) reveal.

Band	Duplex	Swisscom	Sunrise	Salt
800	FDD	20 MHz	20 MHz	20 MHz
900	FDD	30 MHz	30 MHz	10 MHz
1800	FDD	60 MHz	40 MHz	50 MHz
2100	FDD	60 MHz	20 MHz	40 MHz
2100	TDD	—	—	—
2600	FDD	40 MHz	50 MHz	40 MHz
2600	TDD	45 MHz	—	—

(a) Spectrum allocation



(b) Spectrum distribution

(c) Market shares (2016)

Table 6.2 Spectrum allocation of the three mobile communication providers in Switzerland resulting from the auction of February 2012 (valid until 2028).

Unlike in previous years, the awarded licenses are now technology-neutral. Each provider can decide which communication technology to use within each band separately and dynamically. However, on the basis of international developments, it is expected that most bands will witness rapid deployment of the newest technology for high-speed data transfer, currently LTE (Long Term Evolution), the successor to UMTS. Table 6.3 shows which of the communication standards typically are used by each of the three network providers in Switzerland, within the specific bands. Many mobile network operators are abandoning 2G services within the next few years. For example, all of Japan's operators have done so already; Canadian operators have mostly phased out their 2G services as well and the US is expected to follow soon. Vodafone switched off 2G in Australia but is waiting with similar steps in Europe. In most European countries, the operators have phase-out plans for the years 2020 to 2025. In Switzerland, Sunrise has announced phasing out 2G by end of 2018 (they previously announced to cease 2G service by 2017, but did not follow through). Swisscom and Salt have announced plans to turn off GSM by 2020.

Band	GSM/EDGE (2G)	UMTS/HSPA(+) (3G)	LTE/LTE-A (4G)	
800	—	—	■ ■ □	Swisscom
900	■ ■ □	■ ■ □	—	Sunrise
1800	□	—	■ ■ □	Salt
2100	—	■ ■ □	■	
2600	—	—	■ ■ □	

Table 6.3 Typical usage of the spectrum for the different mobile communication technologies in Switzerland (as of 2012, the awarded licenses are technology-neutral).

Fig. 6.19 illustrates the current spectrum in and around the cellular communication bands in Switzerland. Additional resources in the 700 and 1400 MHz bands as well as in the range of 3.4 to 3.8 GHz are planned to become available and be auctioned off within the next few years. A few additional bands are currently being investigated for future use by CEPT (the European Conference of Postal and Telecommunications Administrations). However, as can be seen, the potential increase in spectrum is very small. Thus, expanding into frequencies above 6 GHz (most importantly in the 24 to 26 GHz range, but also in the 40 and 60 GHz bands) is being considered as well.

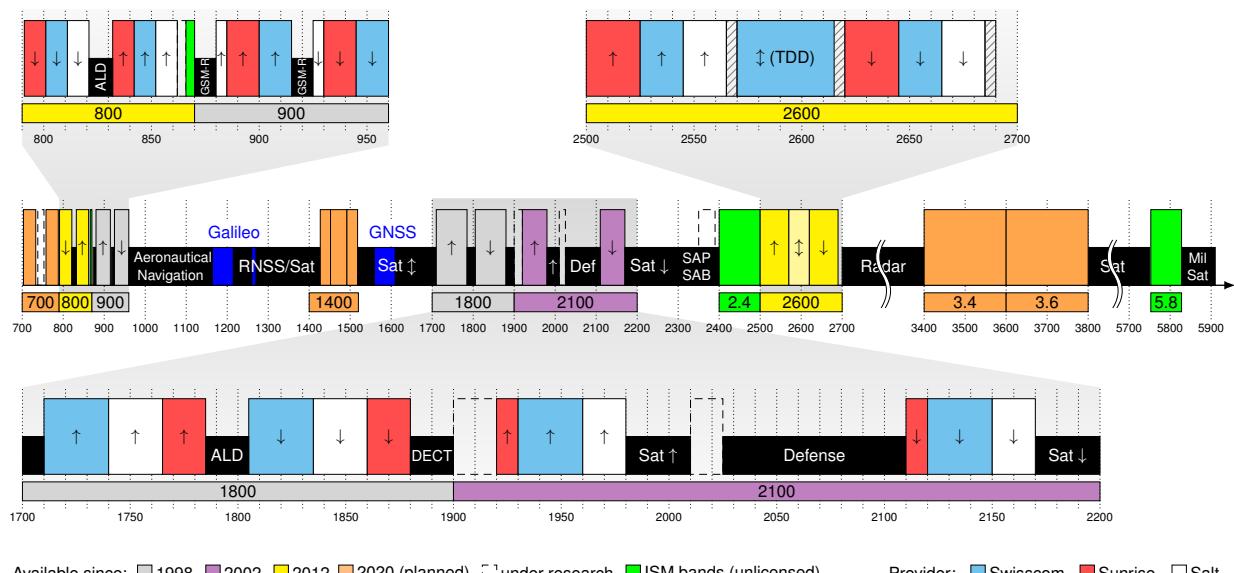


Figure 6.19 Illustration of the current frequency allocation plan for cellular communication systems and other systems in spectral proximity.

Soft Handover As opposed to GSM and other TDMA-based cellular systems, where one user has a link to one cell only at one given time, a UMTS user can use signals from two base stations. They send with the same user code, such that the signal looks as if it was sent through a multipath environment. Thus, we do not have a hard handover when changing from one cell to another, but rather a smooth, so-called *soft handover*, where the user may have a simultaneous link to two base stations during a transition phase, see Fig. 6.20.

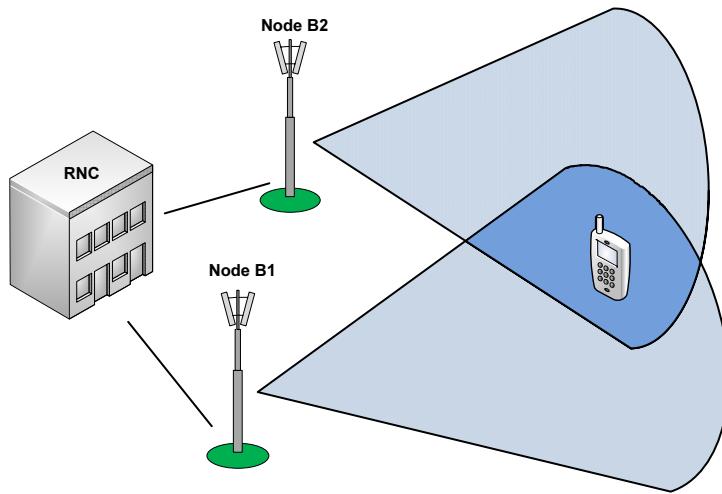


Figure 6.20 Soft handover as a special feature of UMTS.

Orthogonal Variable Spreading Factors The spreading factor is an instrument to control the amount of capacity a user gets. If one user has a low data rate, but his signal is spread using a higher spreading factor to get the final bandwidth, we may accommodate many such users in one given channel. On the other hand, if one user's data rate is already quite high, only a small spreading factor is needed to spread the signal to the final bandwidth. In a UMTS cell, many users using different data rates are served simultaneously. Still, the codes need to be orthogonal. This can be achieved using *orthogonal variable spreading factors*. For an illustration of how to choose the codes, see Fig. 6.21. Once a node is picked, all other codes later and earlier

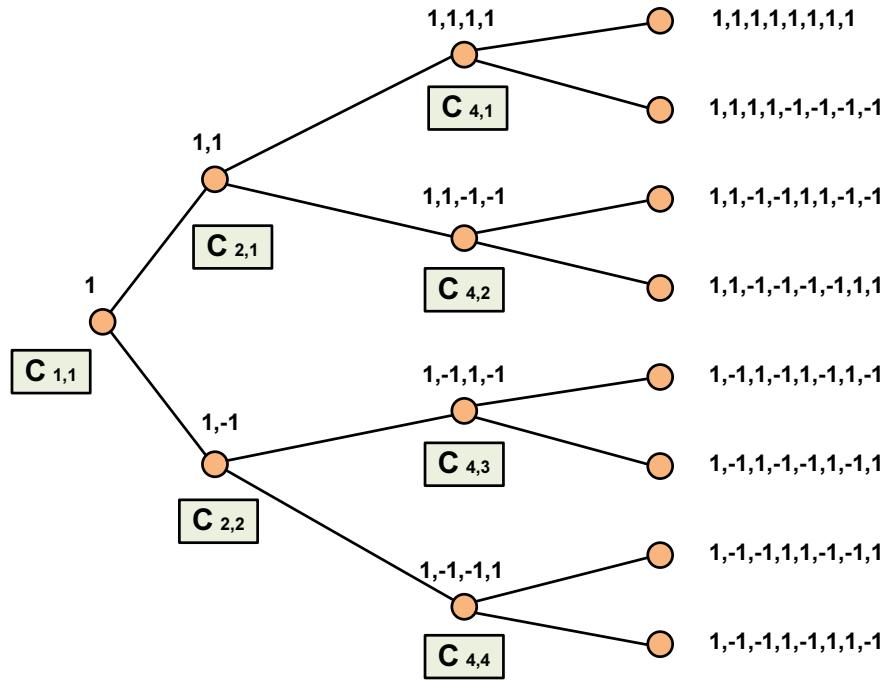


Figure 6.21 OSVF in UMTS to adapt the data rate to user need and capacity.

in the branch are no longer available. Usually, more than one code is applied, e.g., in order to distinguish the user (channelization) and the region (scrambling), see Fig. 6.22.

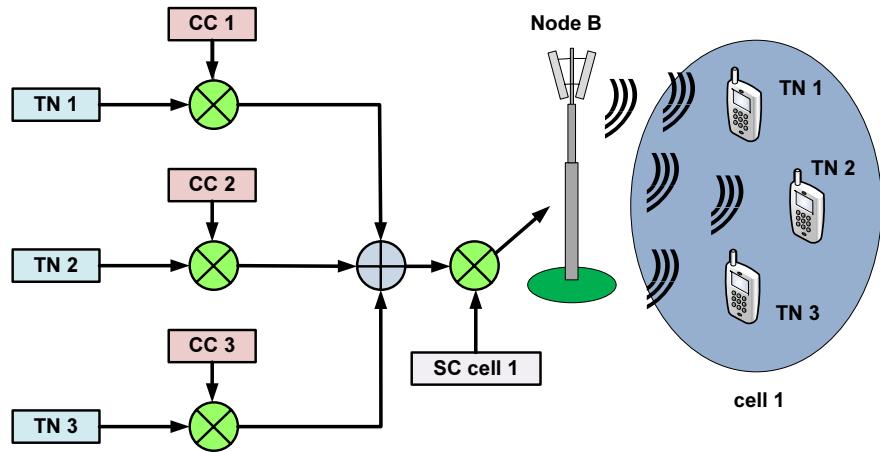


Figure 6.22 Messages from different users are spread using different spreading codes.

Cell Breathing The cell planning of 2G systems is usually a static process in the sense that during operation, cell sizes are constant. For special events involving many people at a certain location, network planners usually extend the number of channels for a basestation by adding hardware modules. A 3G cell in UMTS is much more dynamic. The 'bandwidth'⁴ may be shared among few users with high data rates or among many users with much lower data rates, depending on the user behavior. High data rates can only be provided at a short distance to the base station (due to the lower spreading gain), hence the radius of the cell changes to adapt to the situation. This process is called *cell breathing*.

6.10 Other CDMA Systems

Satellite-Based Navigation Systems Global Navigation Satellite Systems (GNSS) such as GPS or Galileo apply CDMA for two reasons: Firstly, to bring the signal above the thermal noise using the spreading gain such a CDMA system provides, and secondly, to distinguish between the satellites. The Russian counterpart Glonass, however, although using spread-spectrum technology to benefit from the first advantage, distinguishes its satellites using an FDMA technology.

802.11b Although not used so much anymore, one of the earlier standards of wireless LAN deployed direct sequence spread spectrum (DSSS). It was, however, restricted to data rates below 11 Mbit/s. Today, all WLAN standards are based on OFDM.

⁴We set quotes here since bandwidth is not meant in the literal sense, but more in the meaning of capacity.

7 Ultrawideband (UWB) Technology

7.1 Definition

In general, a signal is considered ultrawideband if its fractional bandwidth ν

$$\nu = \frac{f_h - f_l}{f_c} = 2 \frac{f_h - f_l}{f_h + f_l} \quad (7.1)$$

is more than 20 %, where f_h , f_l , and f_c are the higher cutoff frequency, the lower cutoff frequency, and the center frequency, respectively. For the regulatory body (in the States the FCC), a system must use at least 500 MHz to count as ultrawideband. The ultrawideband approach is based on the fact that as long as signal emissions in a certain frequency band are below the unintentional emissions that every electronic device may produce, limited by EMC regulations, no interferences are caused by using an extremely wide spectrum. A total power of less than half a milliwatt is essentially spread over 7 GHz of bandwidth. The L-band around GPS frequencies is exempt. While initially the generation of such a signal fitting the spectral mask of Fig. 7.1 was planned to result from pulse-shaping very short pulses (carrier-less radio), more recently OFDM approaches have been proposed.

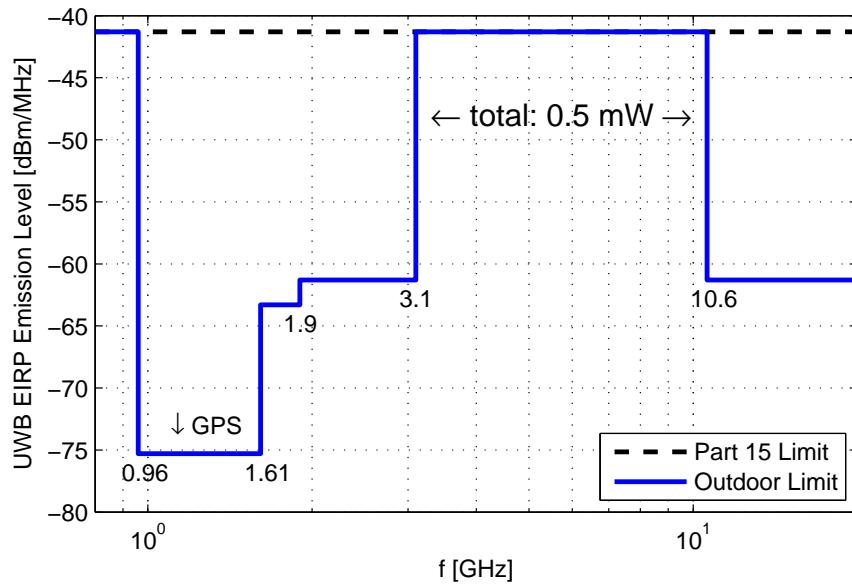


Figure 7.1 Spectral mask for handheld ultrawideband (UWB) systems for indoor and outdoor applications. Defined in FCC Order 02-48.67 and FCC Part 15.209.

7.2 History

If we regard UWB as a carrier-less radio transmission (which it initially was envisaged as), then certainly Hertz, Tesla, and Marconi would have their respective place early in the history of UWB. At a time—more than a hundred years ago—when nobody bothered about spectrum resources or even remotely thought about a carrier frequency, wireless transmission produced by a spark-gap emitter found its own way into the spectrum, or rather, was spread over an undefined band. The idea of generating very short pulses, so-called *impulse radio* is of the same nature.

Several patents on impulse radio were filed in the period between 1942 to 1945. Due to the Second World War, they were granted only more than ten years later. The term UWB itself seems to have emerged around 1990.

The practical generation of sharp pulses is difficult. It is much easier to generate a similar wideband spectrum using conventional technology. Moreover, a spectrum may be shaped using filters to comply with regulations, see for example Fig. 7.1. Still, in a publication in 2003 [27], UWB was described as a technique based on the generation of extremely short pulses in the sub-nanosecond range. This approach was called the direct-sequence approach (DS).

In 2004, however, a group around Texas Instruments put forward a proposal to the IEEE 802.15.3a group, which involved an OFDM approach to generate the UWB signal. That motion split the world into two parts a heavy dispute broke out between the opponents: On the one hand there was the DS "camp" consisting of Freescale/Motorola, which hold on tight to the original approach, and on the other hand there was the TI camp, allied by companies of high importance such as Intel, Philips, and ST. Interestingly enough, although Freescale has had a chip solution on the market for some years, both technologies that would benefit from an increase in data rate, wireless-USB and Bluetooth (future extensions), have chosen the OFDM approach.

7.3 Standards

The IEEE 802.15.3a task group, originally the driving force behind UWB, was dissolved in 2006 due to a dead-lock situation with the two manufacturer consortia. The OFDM approach is now standardized under ECMA368 by the UWB WiMedia Alliance. The consortium is now also known as the MBOA-SIG (Multiband OFDM Alliance SIG). UWB-OFDM uses 14 bands (organized in four 3-band and one 2-band groups) of 528 MHz OFDM signals, 8 of which are above 6 GHz. Each 528-MHz band is divided into 128 subcarriers. Using this approach, a data rate of up to 480 Mbit/s can be achieved.

Similarly to what happened in history with video-recording and CDROM standards, the market will have to decide upon a winner. Proponents of the DS-UWB approach (impulse radio) push the CableFree-USB application, whereas the WiMedia Alliance develops the Certified Wireless USB.

7.4 Applications

Regulations as mentioned above make UWB systems predominantly usable for indoor applications. Applications of UWB systems envisage the wireless connection of PC and monitor, i.e., a high data rate over a very short distance. There are groups that propose to get rid of the SCART-cable between a set-top box and a television display. However, since this becomes more and more a multi-Gbit/s job, UWB seems unlikely to replace the SCART cable with every year that passes by.

Since the UWB signals are wideband in nature, they lead themselves easily to positioning system. In particular, people devising indoor positioning systems, have high hope in UWB. Impulse radio produces such short pulses that multipath is not considered a problem, since very often the reflection is delayed longer than the pulse duration, avoiding thus any overlapping of pulses.

8 OFDM

8.1 Introduction

Orthogonal frequency division multiplexing (OFDM) has become the signaling scheme of choice for many standards, both in the wireless domain and for wired services. Wireless standards such as the most current 802.11 derivates, 802.16 (WiMAX), DAB, DVB-T, DRM, some Satellite Radio standards, the UWB WiMedia standard, and many more all employ some form of OFDM. The two most recent mobile communication standards, 4G and 5G, also called LTE and NR, respectively, are both based on OFDM. Baseband transmission schemes on wire-line connections such as xDSL benefit from OFDM, too. In the literature, OFDM is often referred to as yet another modulation scheme. Instead, we like to think of OFDM as an access scheme, since the choice of modulation remains the same for the input signal to the OFDM block.

With three common letters out of four, OFDM bears a certain similarity to FDMA (frequency division multiple access). FDMA has its roots way back in the realm of analog systems such as AM, PM, FM, etc. An FDMA-based radio system allocates different frequency channels to individual users. In order for the frequency-offset transmissions of different users not to interfere with each other, channel transfer functions must not overlap. The frequency responses are said to be orthogonal, since the integral of the product of two channels is zero, if the channels do not overlap. Since sharp filters are difficult to produce, an FDMA system requires guard bands between the individual channels, i.e., orthogonality comes at a price.

TDMA (time division multiple access) has been considered the solution to this problem for quite a while, since the orthogonality property may easily be achieved by securing the simultaneous arrival of signals from different users. Mobile communication systems (such as GSM) accomplish this using techniques such as timing advance. Rather interestingly, this creates a trade-off, since TDMA channels are wider in bandwidth and thus exposed to frequency-selective fading. In one prominent example, this question has divided Europe into countries employing a wider-band TDMA and countries using a narrower-band FDMA approach for their private mobile radio systems used by the Police and the like (Tetra vs. Tetrapol). The former approach is more efficient in its frequency usage but needs an equalizer to tackle the frequency-selective fading situation, whereas the latter can cope without an equalizer but wastes more bandwidth due to guard bands.

OFDM divides the spectrum into small channels, called subcarriers, such that frequency-selective fading is no longer a problem. Yet, no spectrum is wasted, since a non-overlapping situation is not aimed at. Rather, the overlapping of adjacent subcarriers is done in such a controlled way that orthogonality is achieved all the same.

In OFDM, the data is transmitted on parallel subcarriers using frequency-division multiplexing. In some applications, this method is also called *discrete multitone modulation* (DMT). The carrier spacing depends on the symbol period and is selected such that each subcarrier is orthogonal to any other subcarrier over a symbol period T_S .

In order to avoid ICI (interchannel interference), guard intervals are inserted that have the form of a *cyclic prefix*. The length of the cyclic prefix must be larger than or equal to the channel impulse response length (echo duration) in order to maintain the orthogonality between subcarriers, so that no ICI arises.

8.2 Mathematical Derivation

Mixing each of the subband signals onto a different carrier, we obtain a situation as illustrated in Fig. 8.1. Rather than using sine or cosine carriers, we will use exponential functions ($\cos + j \sin$) in order to retain

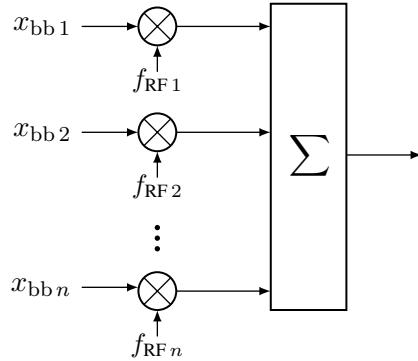


Figure 8.1 Construction of a multicarrier signal.

the complex-valued signal of a asymmetric baseband spectrum¹. The constructed signal can thus be written as

$$y = x_{bb1} \cdot \exp(j\omega_1 t) + x_{bb2} \cdot \exp(j\omega_2 t) + x_{bb3} \cdot \exp(j\omega_3 t) + \dots , \quad (8.1)$$

where the frequencies $f_k = \frac{\omega_k}{2\pi}$ are the baseband subcarrier frequencies. Measured over a symbol period T_S , all subcarriers need to be orthogonal with respect to each other,

$$\frac{1}{T_S} \int_0^{T_S} \exp(j2\pi f_i t) \cdot (\exp(j2\pi f_j t))^* dt = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \quad (8.2)$$

The above-mentioned orthogonality can be achieved by choosing $f_k = k/T_S$. This can be seen by observing a series of cosines for different indices k as shown in Fig. 8.2(a). In Fig. 8.2(b) the corresponding frequency spectra of each cosine (amplitude only) shows that each subcarrier aligns with all the other subcarriers' spectral zero-crossing points. Although the subcarriers spectrally overlap, they do not interfere with each other if the signal is correctly sampled.

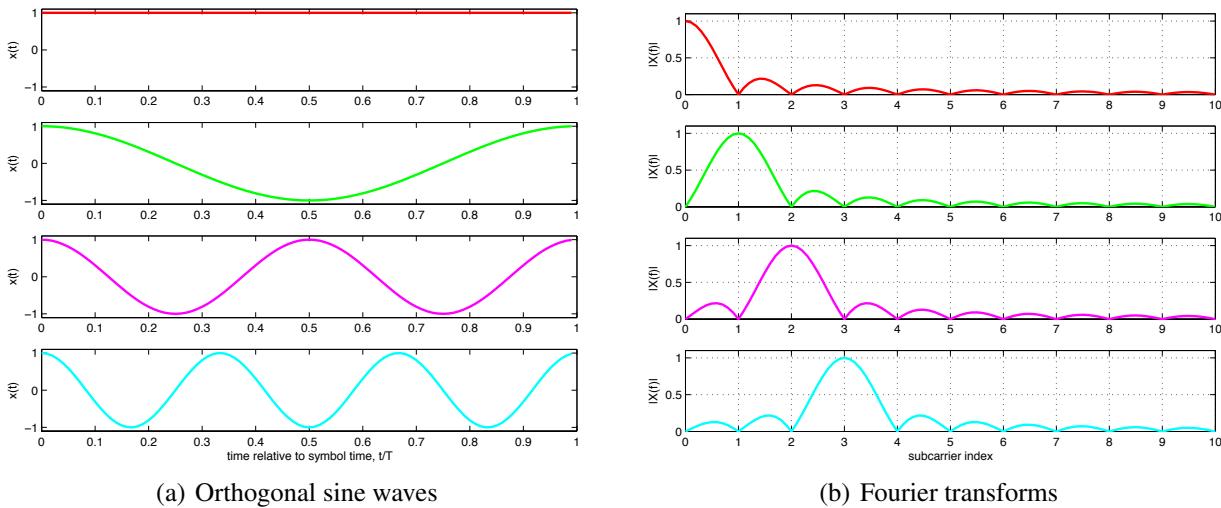


Figure 8.2 Series of sine waves that are orthogonal to each other and their corresponding Fourier transforms.

¹For illustration purposes we may occasionally show the real part or the imaginary part only.

If all subcarriers are present, see Fig. 8.3, they still do not interfere with each other at the exact subcarrier positions, indicated by the circles. When the subcarriers bear information, their amplitudes or phases can

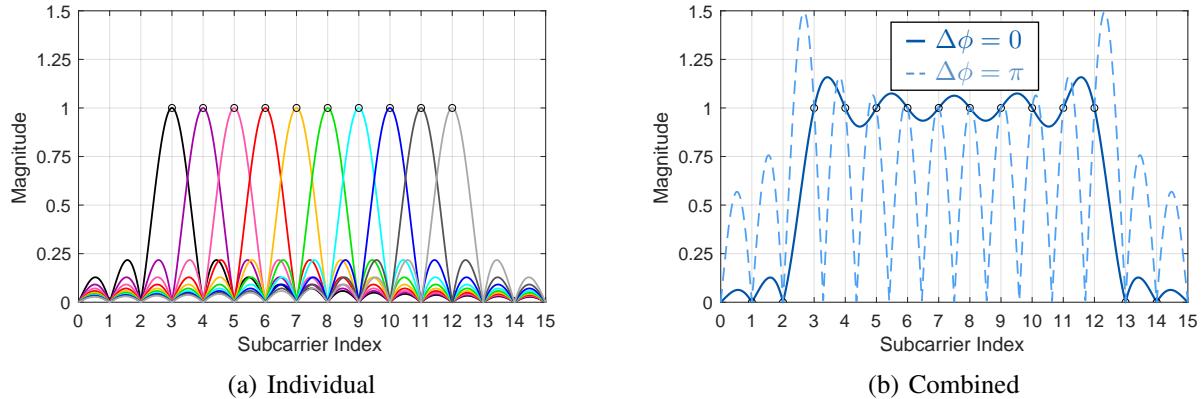


Figure 8.3 Series of subcarriers with identical amplitude/phase information as individual carriers and in its combined form.

have different values, depending on the modulation scheme chosen. A situation with four different amplitudes is shown in Fig. 8.4. When arbitrary phase-modulated subcarriers are averaged, we get a spectrum as shown using the solid line in Fig. 8.3(b).

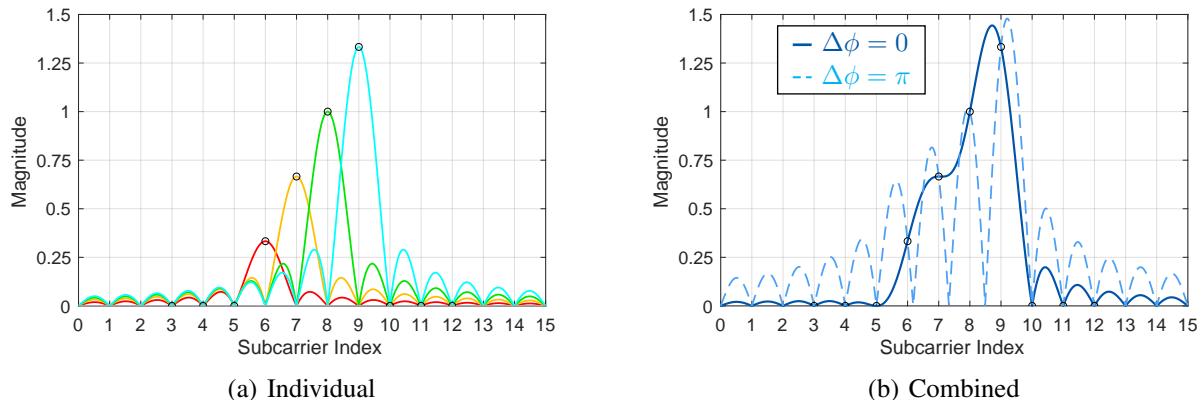


Figure 8.4 Series of subcarriers with different amplitude information as individual carriers and in its combined form.

8.3 Signal-Processing Chain

The basic block diagram of an OFDM system is given in Fig. 8.5. In OFDM, the modulator can be implemented as an N -point *inverse discrete Fourier transform* (IDFT) on a block of N data symbols followed by a *digital-to-analog converter* (DAC) on the IDFT samples. For the following we assume a sequence $\{s_k, k = 1, \dots, N\}$, of complex data symbols s_k chosen from any modulation scheme, e.g., *Quadrature*

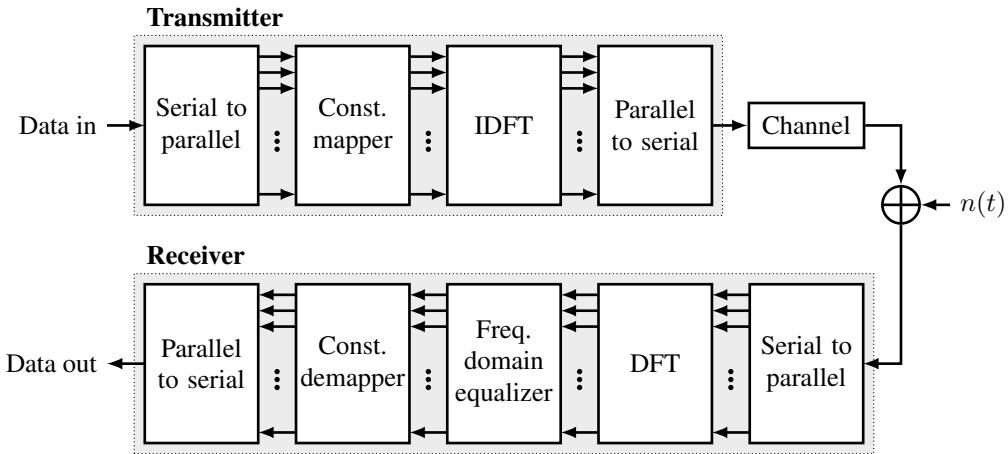


Figure 8.5 Components of an OFDM system.

Amplitude Modulation (QAM). Carrying out the IDFT on the data block, we get

$$S_n = \text{IDFT}\{s_k\} = \frac{1}{N} \sum_{k=0}^{N-1} s_k e^{\frac{j2\pi nk}{N}}, \quad n = 0, 1, \dots, N-1. \quad (8.3)$$

Contrariwise, the OFDM demodulator samples the received time-domain signal S_n with an *analog-to-digital converter* (ADC). Afterwards, the demodulation is realized with the N -point *discrete Fourier transform* (DFT)

$$s_k = \text{DFT}\{S_n\} = \sum_{n=0}^{N-1} S_n e^{-\frac{j2\pi nk}{N}}, \quad k = 0, 1, \dots, N-1. \quad (8.4)$$

In theory, where the exact sample index and sampling time is known, this inverse signal processing task seems easy and logical. In practice, however, the information as to when the signal has to be sampled has to be obtained from the signal itself. This process is called synchronization and shall be described in the following.

8.3.1 Frame Synchronization

The frame synchronization is the process of finding the beginning and the end position of a frame. It is usually done in one of two ways. Usually, a compromise between complexity, reliability and minimum latency has to be found. One possibility is to insert a null symbol into every frame, see Fig. 8.6(a). The receiver simply has to detect a drop in the power of the received signal. This can be done using the filtered output of an envelope detector. This situation is illustrated in Fig. 8.7.

The other possibility is to look for the *training symbol*, which is part of every OFDM frame, see Fig. 8.6(b). The data stream consists of a series of single frames. Every frame starts with a so-called *training symbol*. This symbol consists of an OFDM symbol, whose content is known by the receiver. The purpose of that symbol is to have both a reference for every carrier used in the OFDM transmission in order to equalize the channel distortion and a reference to synchronize with the start of a frame. The rest of the frame consists of an array of OFDM symbols that contains the data.

The frame synchronization searches the training symbol by cross-correlating them with the incoming signal. At the point of perfect match, the correlation results in a high peak as can be seen in Fig. 8.8. The training symbols are often not only used for the symbol and frame synchronization, but can also serve for the channel

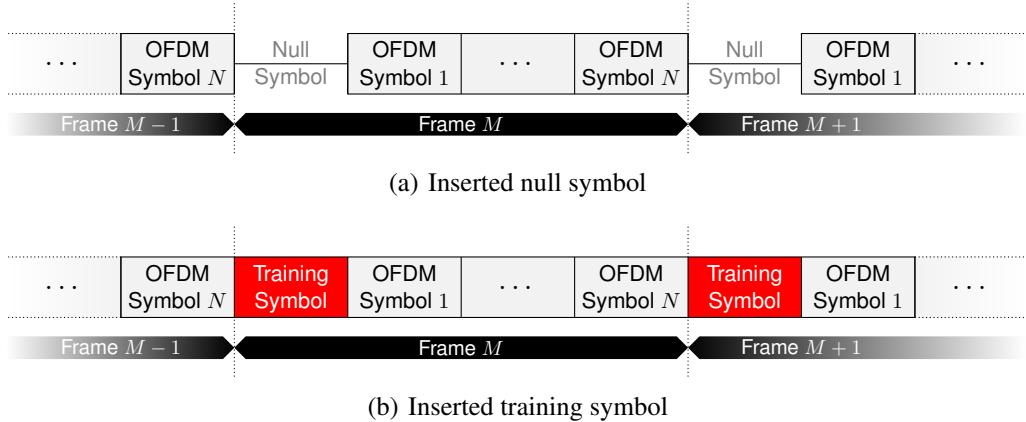


Figure 8.6 Frame structures of OFDM systems using a null symbol or a training symbol.

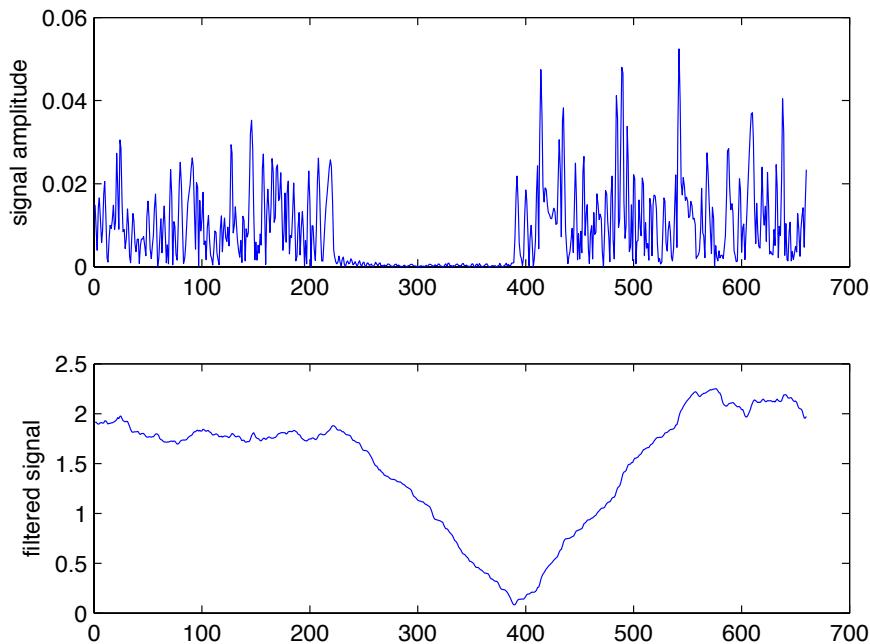


Figure 8.7 Frame synchronization by envelope detection.

estimation of the equalization process as will be seen in the next section. A possible placement of the pilots over time across the different subcarrier in order to get a representative picture of the channel is shown in Fig. 8.9.

8.3.2 Sampling-Time Synchronization

During the demodulation of the OFDM symbols in a frame, there is another effect that must be considered. In any communication system, the sampling clock at the receiver is often different from that at the transmitter. The sampling clock difference introduces interchannel interference because the subcarriers lose their perfect orthogonality. If the difference remains small, the deviation only consists of a subchannel-dependent

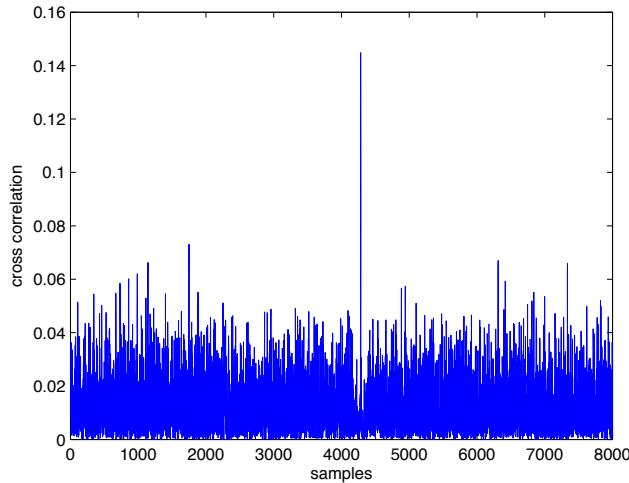


Figure 8.8 Frame synchronization by cross-correlation.

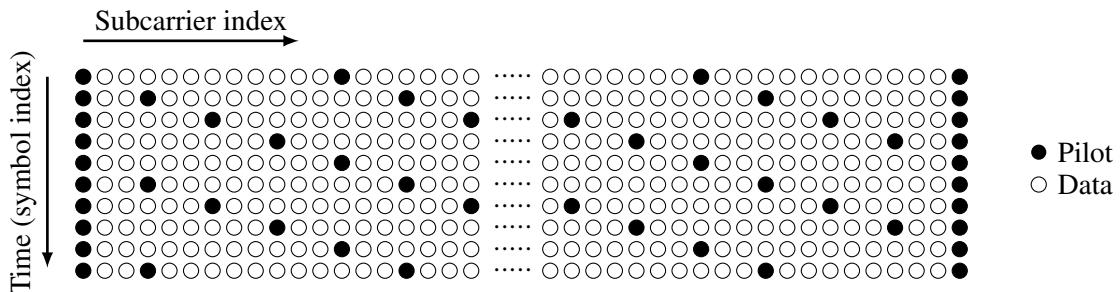


Figure 8.9 Pilot-tone placement.

phase rotation over time, according to

$$f(nT_s \pm \tau) \circ \longrightarrow F(jk\omega_0) \cdot e^{\pm jk\omega_0\tau}. \quad (8.5)$$

One approach to compensate that phase rotation is to use the *pilot carrier* to estimate and compensate them. The phase of the pilot carrier continuously grows after every OFDM symbol by a constant increment.

8.3.3 Guard Intervals: Zero-Padding and Cyclic Prefix

Since OFDM is based on the DFT and the IDFT applied on a certain block length at a time, this block is inherently assumed to be repetitive or cyclic. However, the convolution with the channel response destroys that 'illusion' since echo in the early part of the block originates from a later part of the previous block. We also say that one block 'bleeds' into the next one. Since a block corresponds to a symbol, we have a situation of intersymbol interference (ISI). Moreover, because we apply the Fourier transform to this symbol, the destruction of the cyclic property of the symbol will turn ISI into interchannel interference (ICI), i.e., the subcarriers are no longer orthogonal.

ISI can be avoided by inserting a guard interval in front of a symbol. Two types of guard intervals are commonly used:

- zero-padding: an empty guard interval,
- cyclic prefix: a guard interval consisting of a partial copy of the symbol.

Inserting a few zeros in front of the symbol is simple, as shown in Fig. 8.10(a) and ISI is avoided with this method. However, since the orthogonality is not preserved, we still have ICI. The remedy for this situation is to make the block (symbol) cyclic again, by copying a part of one end of the block in front of the other end, see 8.10(b).

Depending on the place where we put the extension, we speak of cyclic prefix or cyclic postfix (more commonly known as suffix), as illustrated in Fig. 8.11. The length of the cyclic prefix must be larger than or equal to the channel impulse response length (echo duration) in order to maintain the orthogonality between subcarriers, so that no ICI arises.

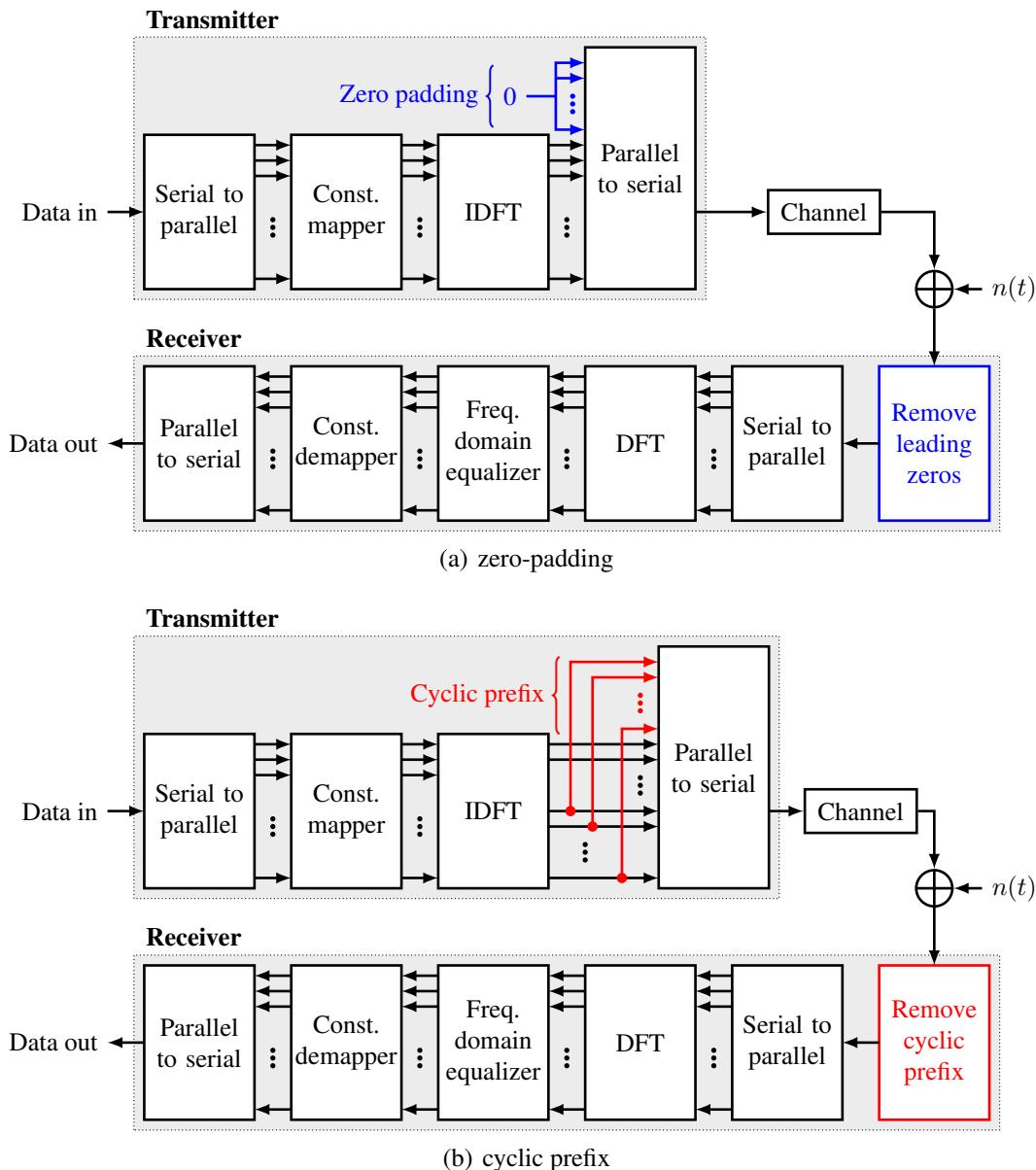


Figure 8.10 Block diagrams illustrating methods to reduce ISI in OFDM systems. (Note that constellation mapper/demapper and an equalizer have also been added.)

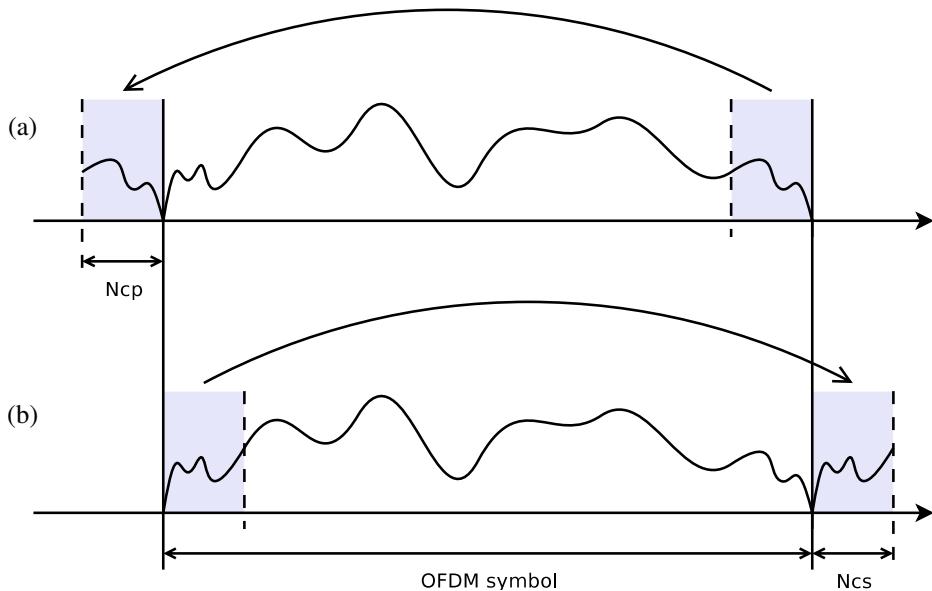


Figure 8.11 OFDM signal: (a) with cyclic prefix, (b) with cyclic suffix.

8.3.4 Channel Estimation and Equalization

The received OFDM signal is distorted by the channel and has to be equalized prior to be demodulated. In OFDM, the whole frequency bandwidth is divided into N subcarriers. If N is high enough, then the transfer function $H(f)$ does not change significantly over the subchannel and it can be regarded as flat over a single narrow subcarrier. The estimation of the channel transfer function is based on the knowledge of the training symbol, which serves as a reference word. The insertion of a training symbol into the frame structure was already shown in Fig. 8.6(b). The approximation of the transfer function for a single subcarrier reduces to a single complex coefficient $H_n = H(f_n)$. Therefore, the equalization of a single subcarrier can be simply performed by multiplying the received sample with the inverse of the generally complex-valued H_n . The equalization coefficients w_n can be computed by the element-wise division of the original FFT coefficients z_n of the training symbol by those of the received training symbol \hat{z}_n

$$w_n = \frac{z_n}{\hat{z}_n} = \frac{1}{H_n}. \quad (8.6)$$

All further OFDM symbols in the same frame can now be equalized by multiplying the received OFDM symbol in the frequency domain with the equalization coefficients w_n . These complex coefficients also include the linear phase caused by the timing error in the sampling device.

8.3.5 Optimal Energy Allocation (Waterfilling)

One of the main advantages of OFDM is that every subcarrier can be individually parameterized in terms of constellation size, transmit energy, and even error protection. This property provides an opportunity to improve the performance of the OFDM signal. Remember that the Shannon capacity of the channel is given by

$$C = W \log_2 \left(1 + \frac{S}{N} \right) = W \log_2 \left(1 + \frac{S}{\eta W} \right), \quad (8.7)$$

where W is the total bandwidth available, S is the total signal power (arriving at the receiver in the N_c different subchannels), and η is the noise power density. Adapting Eq. (8.7) to the case of a subcarrier

system, we have

$$C = \sum_{n=1}^{N_c} \frac{W}{N_c} \log_2 \left(1 + \frac{S_n}{\eta W/N_c} \right), \quad (8.8)$$

where S_n is the received power in the n th subchannel. This partial power is given by the attenuation for that particular subchannel

$$S_n = P_n \cdot |H_n|^2, \quad (8.9)$$

where P_n is the power allocated to the n th subchannel. We thus have

$$C = \sum_{n=1}^{N_c} \frac{W}{N_c} \log_2 \left(1 + \frac{P_n |H_n|^2}{\eta W/N_c} \right). \quad (8.10)$$

It can now be shown that this capacity can be achieved if the the power allocation is chosen as

$$P_n = \beta - \frac{\eta W/N_c}{|H_n|^2}, \quad n = 1 \dots N_c, \quad (8.11)$$

where β must be chosen such that the power constraint

$$\sum_{n=1}^{N_c} P_n = P \quad (8.12)$$

is satisfied. The value of β can now be worked out using

$$P = \sum_{n=1}^{N_c} P_n = N_c \beta - \eta W/N_c \sum_{n=1}^{N_c} \frac{1}{|H_n|^2}, \quad (8.13)$$

and hence

$$\beta = \frac{P + \eta W/N_c \sum_{n=1}^{N_c} \frac{1}{|H_n|^2}}{N_c}. \quad (8.14)$$

In words, the sum of the channel noise power (normalized by the channel attenuation) and the transmit power at each frequency must be kept constant, see also [22]. This procedure is also known as *water-filling*. With the aid of the water-filling algorithm, the best bit and energy distribution over the subcarriers can be found in an OFDM system. Water-filling energy allocation achieves capacity. If we let the noise-to-signal ratio (NSR) in each subcarrier be

$$\text{NSR}_n = \frac{\eta W/N_c}{|H_n|^2}, \quad n = 1 \dots N_c, \quad (8.15)$$

then the sum of NSR_n and P_n will be constant for any² n . An example of a spectrum of NSR and the correspondings allocated power values P_n (for all $n = 1 \dots N_c$) for different amounts of total power P is given in Fig. 8.12(a) to 8.12(c).

Besides the calculation of the subcarrier power values, we can also compute the optimum number of bits to transmit on each subcarrier. Generally, such an algorithm produces bit distributions where the single bit loads can be any real number and the realization of these non-integer bit values can be difficult. Taking into account this special need, the so-called bit-loading problem, algorithms such as the one proposed by Levin and Campello [21] are utilized because they allow the computation of bit distributions with integer bit loads.

²This is true as long as there is enough power. Note that P_n cannot become negative, in this case P_n will be zero and NSR_n for that particular n might exceed the sum in other subcarriers.

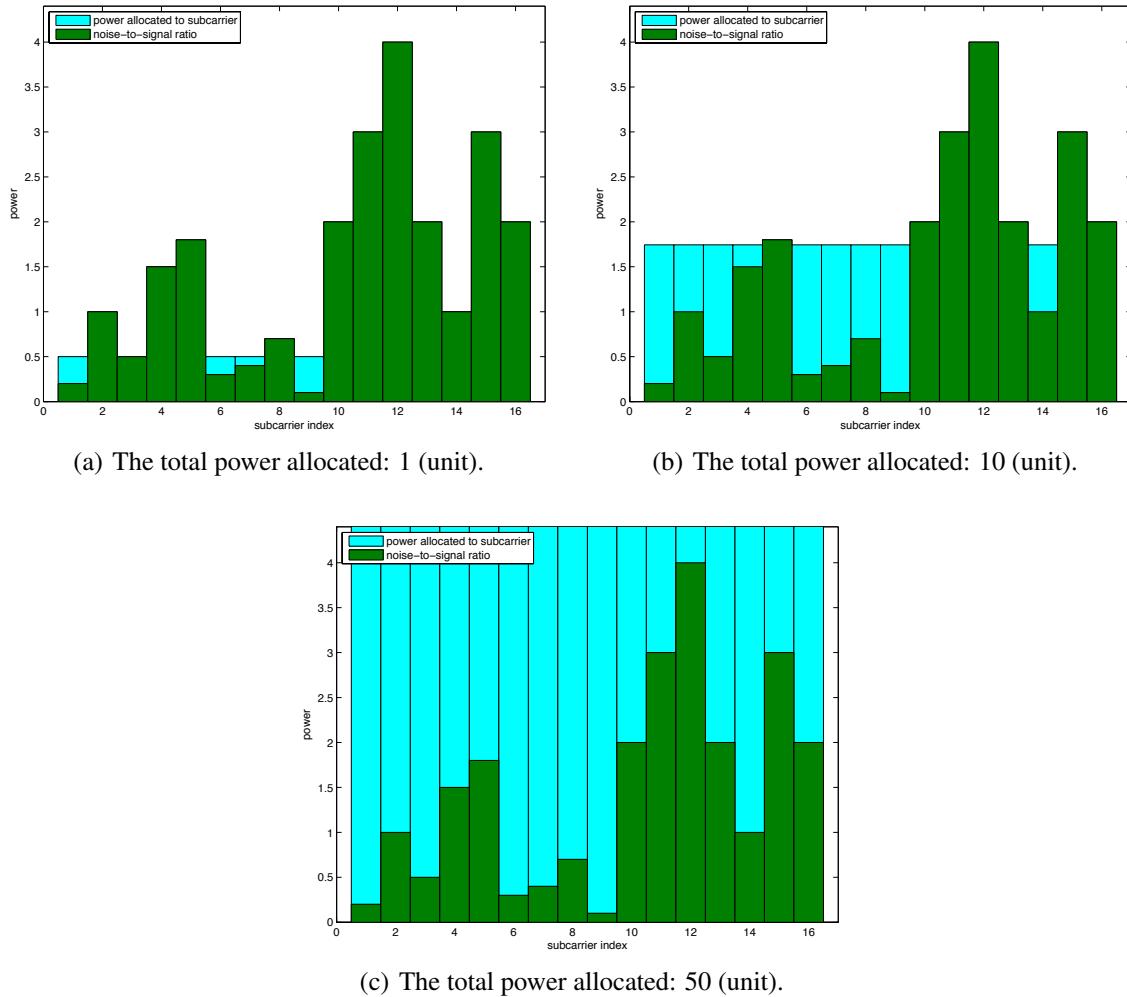


Figure 8.12 Example of water-filling spectrum.

8.3.6 PAPR Considerations

One of the main drawbacks of OFDM is the high *peak-to-average-power ratio* (PAPR). It is the result of the superposition of a large number of (usually statistically independent) subcarriers that can constructively sum up and temporarily result in a total signal power which is much higher than the average power. When N_c signals are added together with the same phase, they produce a peak power of

$$P_{\text{peak}} = \left| \sum_n x_n \right|^2 \xrightarrow{x_n=x} |N_c x|^2 = N_c^2 |x|^2, \quad (8.16)$$

whereas the average power is

$$P_{\text{avg}} = \frac{1}{N_c} \left| \sum_n x_n \right|^2 \xrightarrow{x_n=x} N_c |x|^2. \quad (8.17)$$

Thus, in this worst-case, the peak-to-average power ratio is

$$\text{PAPR} = \frac{P_{\text{peak}}}{P_{\text{avg}}} = \frac{N_c^2 |x|^2}{N_c |x|^2} = N_c. \quad (8.18)$$

Hence, the PAPR problem of OFDM generally increases with the number of subcarriers used.

If the peak signal power exceeds the transmitter capabilities, the signal is clipped, resulting in intermodulation interference and bit-errors, thereby lowering the achievable data rate. To avoid clipping, the power of the transmitted signal must be lowered, so as to ensure that it always fits in the transmitter's input range. However, this so-called *power back-off* leads to a lower average signal power, and hence to a lower SNR available for data transmission, which can diminish the data rate as well.

Various techniques to reduce the PAPR have been proposed [15]; among these are:

- backing off amplifier (sacrifice power efficiency)
- scrambling (let the statistics work for you)
- redundant selection (sacrifice overhead efficiency)
- coding techniques (sacrifice computational efficiency)
- soft clipping
- tone reservation (sacrifice subcarriers)
- constellation expansion (allow the highest signal to be still higher)
- use of OFDM in downlink only (e.g. LTE).

8.4 Applications

Some applications of OFDM are given in the following list:

- ADSL, VDSL
- DAB (digital audio broadcasting)
- DVB-T, DVB-H (but not -S and -C), DRM
- 4G cellular (LTE)
- 802.11a, 802.11g, 802.11n
- 802.16, 802.16a, 802.16b (WiMAX)
- UWB
- Satellite systems

8.5 Summary

Whereas the usual access to the channel using FDMA requires guard bands between the individual information bands, OFDM makes such a waste of bandwidth unnecessary by employing small bands whose access exploits the property of orthogonality with respect to the subcarriers. This new scheme which is in some way an access scheme as well as a modulation scheme is called *orthogonal frequency division multiplexing* (OFDM).

In OFDM, the data are transmitted on parallel subcarriers using frequency-division multiplexing. The carrier spacing depends on the symbol period and is selected such that each subcarrier is orthogonal to any other subcarrier over a symbol period T_S .

The division of a band into several subbands has many advantages. If the bandwidth is small enough, the characteristics of each subchannel can be assumed ideal. The subchannel may be approximated as a channel

with a flat transfer function across the whole subband, and the noise in all subbands can be modeled as AWGN. A channel equalizer is essentially one complex tap per subcarrier and this multiplication can be carried out right after the DFT in the OFDM receiver.

The number of subcarriers of an OFDM system is limited by two considerations. Firstly, the number of subcarriers should be chosen high enough to make the symbol period much longer than the maximum delay of the channel, or, in other words, to make the fading process frequency non-selective (within the subcarrier). The subcarrier bandwidth W/N_c is to be smaller than the coherence bandwidth B_{coh} . Secondly, the subcarrier symbol period N_c/R , where R is the symbol rate of the whole band needs to be short enough in order not to be exposed to the time invariance of the channel, which is characterized by the coherence time T_{coh} . A further advantage of OFDM technology is the bit-loading of subcarriers according to their individual SNR (adaptive loading), thereby making optimum usage of the spectrum.

Introductory books treating OFDM in great details are, amongst many others, [33, 16, 22, 9].

9 Channel Coding

The main problem of a receiver in the digital transmission of information is to reliably decide what was sent by the transmitter. 1948 Claude Shannon showed in one of the most important publications in communication theory [34], his treatise on the mathematical foundations of communication theory, that, as long as the information rate does not exceed the channel capacity, transmission errors can be made arbitrarily small. The means for this are error-correcting codes. By adding redundancy to the information bits, the robustness against transmission errors is increased. For the following considerations we refer to Fig. 9.1, where a block diagram of a typical transmission system¹ is shown.

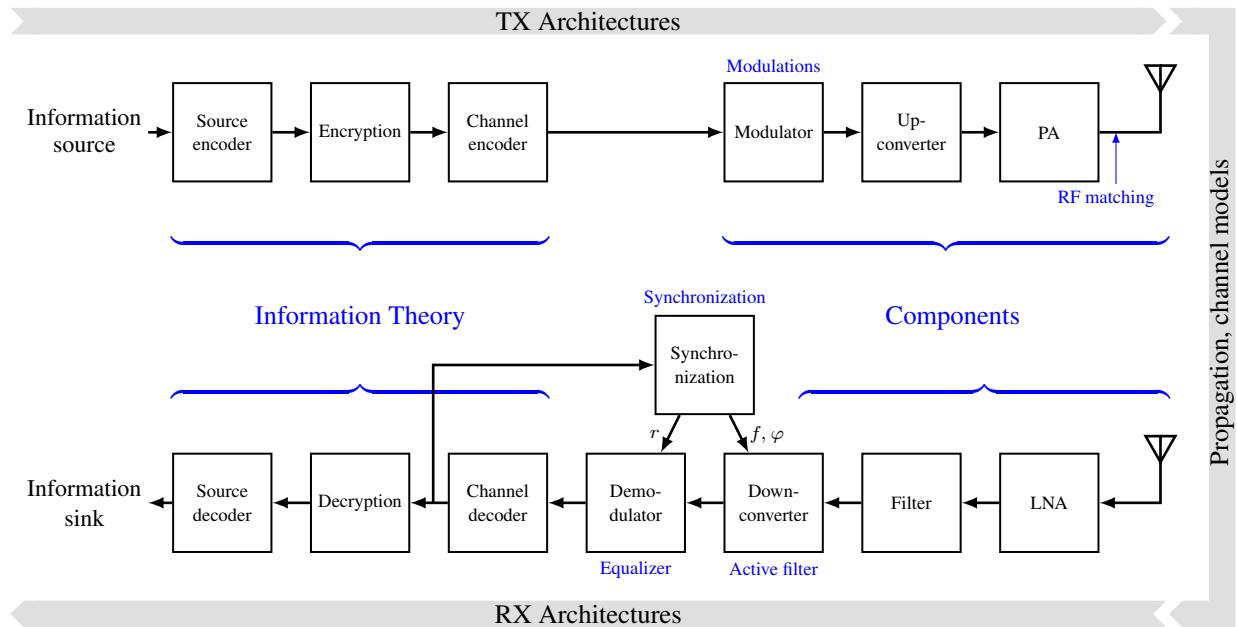


Figure 9.1 Disciplines involved in the design of a wireless communication system.

The information source produces the information packets, which, in turn, are passed on to the next blocks, some of which are optional. The source encoder tries to compress the signal with respect to information content. That can be accomplished by compression algorithms such as Huffman-Coding or—particularly in mobile communications—voice coders, which, rather than sending sampled speech data, extract parameters and transmit them over the channel.

The next block *channel encoding* and the corresponding block *channel decoding* we will have a closer look at during this chapter. Channel coding adds redundancy to the signal and protects it thus from the "attacks" in the transmission channel. The modulator performs the channel adaptation, i.e., for wireless radio channels the information is modulated onto an RF carrier². For radio channels, the transmission channel means the air distance between transmitter and receiver, for wireline systems cable, converters, and repeaters. Channel models for the former medium are generally more complicated due to their time-varying nature and multipath propagation. Moreover, there is additive noise³.

¹With only minor changes in the notation, the block diagram can also be used for digital storage systems. Modulator and demodulator become write and read mechanism, respectively, and the transmission channel becomes the storage medium.

²For wireline systems the channel signal can be an RF or a baseband signal.

³Strictly speaking, the larger part of the noise stems from thermal noise at the input to the receiver, but for reasons of simplicity noise is usually modeled together with the channel.

In the receiver, the same chain of blocks is passed in reverse order. The demodulator carries out the inverse operation of the modulator. The channel decoder tries to correct corrupted symbols. Depending on the kind of error-correcting code, more or fewer erroneous symbols can be corrected. Some errors might be detected but cannot be corrected.

9.1 Interleaving

In practice, the situation often occurs that transmission errors are not independent of each other, but occur in bursts. Examples of this situation are shadowing by buildings in mobile radio channels, switching transients in wireline transmission, scratches on the CD etc. We then say that the channel has memory and express thereby that the probability of a symbol error is much higher if the symbol immediately preceding was already in error. Although there are some codes that are well suited for the correction of bursty errors, the codes are generally designed for the correction of only a few, independent errors. With channels with memory we often have the absurd situation that for long periods there are no errors—the decoder is under-employed not having to correct anything—only to find a sudden burst of errors, which, in turn, exceeds the capability of the code. It would be much wiser to spread these errors equally among all codewords.

To this end we introduce a further block in Fig. 9.1 between the channel encoding and the modulator, the so-called *interleaver*. On the receiver side, consequentially, the inverse operation, the so-called *deinterleaver* is introduced between the demodulator and the channel decoding. The interleaver can be seen as a table that is written to in a column-wise manner, and read from in a row-wise manner, as exemplified in Fig. 9.2.

to modulator							
↑							
from encoder →							
1	7	13	19	25	31	37	43
2	8	14	20	26	32	38	44
3	9	15	21	27	33	39	45
4	10	16	22	28	34	40	46
5	11	17	23	29	35	41	47
6	12	18	24	30	36	42	48

Figure 9.2 Block interleaver.

In our example, a block consists of 48 symbols. The symbols from the encoder are written in their natural order into the interleaver. Six subsequent symbols fill a column. Then, the writing is progressed to the next column. The memory content is now read out row-wise and passed to the modulator, i.e., in our example the indices of subsequent symbols differ at least by six (more after a row change).

For the following demonstration example of a bursty transmission error, we suppose that each column of the interleaver corresponds to a codeword of a code that can correct one error. Now we have a bursty error of four subsequent errors. Without the interleaver there is always at least one uncorrectable codeword, since at least one codeword contains more than one error, as can be easily shown. Fig. 9.3 shows the symbols in the order of transmission, where the erroneous symbols are framed.

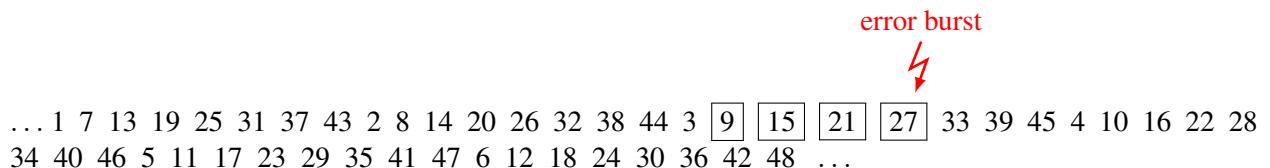


Figure 9.3 Order of the symbols in the channel with error burst.

Fig. 9.4 shows the order of the symbols after the deinterleaver immediately before the decoding process. As can be easily verified, there is at most one error per codeword, a situation that can usually be handled by the chosen code.

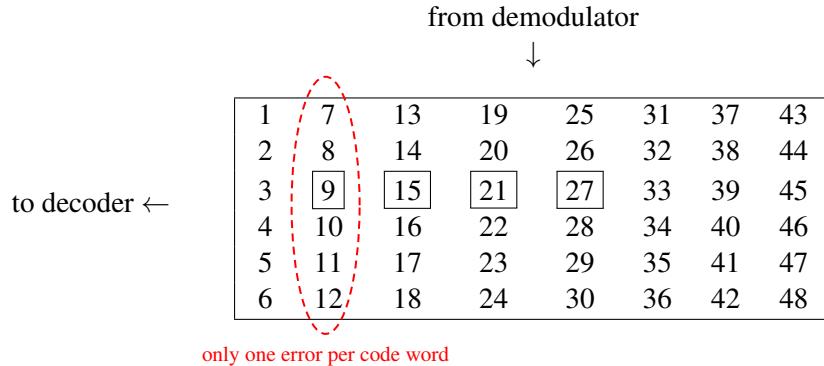


Figure 9.4 Distribution of errors after deinterleaver.

9.2 Convolutional Codes

9.2.1 Construction of Convolutional Codes

One group of error-correcting codes beside the block codes are the so-called *convolutional codes*. While the operations of a block code are always within the length of a codeword of a given length, the encoding process for convolutional codes is a continuous one. The coding recipe is thus normally not given as a generator matrix but either as a set of generator sequences or as a circuit diagram of the encoder. Encoders of convolutional codes can be very simple as Fig. 9.5 shows. The generator vectors are $g_1 = [1, 1, 1]$ and $g_2 = [1, 0, 1]$, corresponding to the octal representation $(7, 5)_8$.

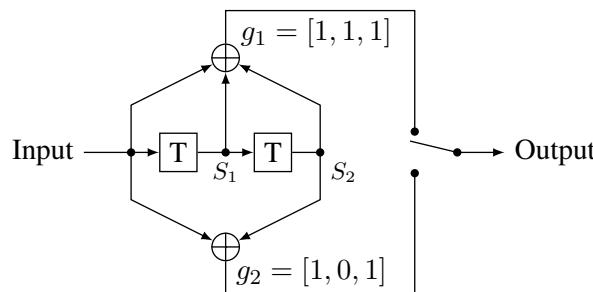


Figure 9.5 Implementation of a convolutional encoder. The state is described by S_1, S_2 .

As can be seen, the overall structure of Fig. 9.5 is similar to an FIR filter with binary coefficients, where addition is not carried out in \mathbb{R} , but modulo 2, i.e., as an EXOR function. The ratio of the number of output symbols to the number of input symbols depends on the desired "protection". In the given example, there are two output symbols per input symbol. Similarly to block codes, there are also systematic codes. Since the process is not a blockwise one, the information bits are not subsequent but isolated. Systematic codes mean in this case that one of the generator polynomials has one term only, e.g., $g_i(X) = 1$ or $g_i(X) = X^n$ with $n \in \mathbb{N}$. If one is willing to give up the correction capability of the codes one can easily find the information bits in the receiver through extraction of the corresponding symbols from the

data stream without further processing. Since convolutional codes are not used for error detection but only for error correction, systematic convolutional codes are of minor interest.

A further design parameter is the so-called constraint length, which describes how long an input symbol influences the output of the encoder. The constraint length is simply the number of shift registers plus one. In our example the constraint length is three. An increase of the constraint length generally improves the correction capability of the code, however, at the price of increased decoding burden. The complexity of the decoding process grows exponentially with the constraint length.

9.2.2 Decoding of Convolutional Codes

The decoding of convolutional codes is very different from that of block codes. The main task at hand in the receiver is to compare the received sequence with all possible sequences (given the generator polynomial) and choose the most likely one. Such a receiver is called a *maximum likelihood* (ML) receiver. If we deal with an AWGN signal, i.e., in the transmission channel we only have additive white Gaussian noise with the probability density

$$p_{\text{AWGN}}(v) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{v^2}{2\sigma^2}}, \quad (9.1)$$

the most likely sequence \hat{x} is the one that has the smallest sum of the Euclidean distances

$$\sum_i (x_i - \hat{x}_i)^2 \quad (9.2)$$

to the received signal x .

Since the number of possible sequences grows exponentially with the length of the sequence, it becomes clear that the comparison of the received sequence with every possible sequence is close to impossible. However, this is not necessary as we shall see. The Viterbi algorithm facilitates a very efficient decision of the most probable sequence. In order to understand this important algorithm of communication theory, let us have a look at the code of Fig. 9.5. If we regard the outputs of the shift register as state variables of a finite state machine (FSM), we can represent the code by Fig. 9.6.

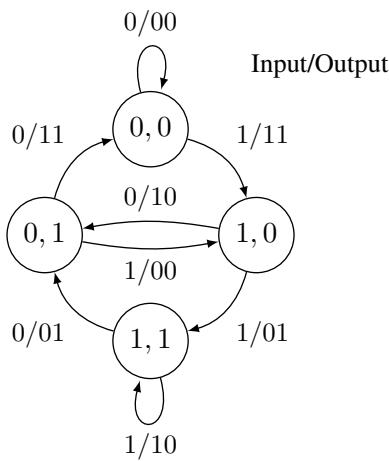


Figure 9.6 State diagram of a convolutional encoder. The state is described by S_1, S_2 .

As can easily be seen, only half of the state transitions are possible as a consequence from the fact that the state bit S_1 becomes state bit S_2 in the next step. If we now convert this state diagram into a so-called *trellis* we get Fig. 9.7. The reader realizes that a trellis is nothing but the development of a state diagram with time.

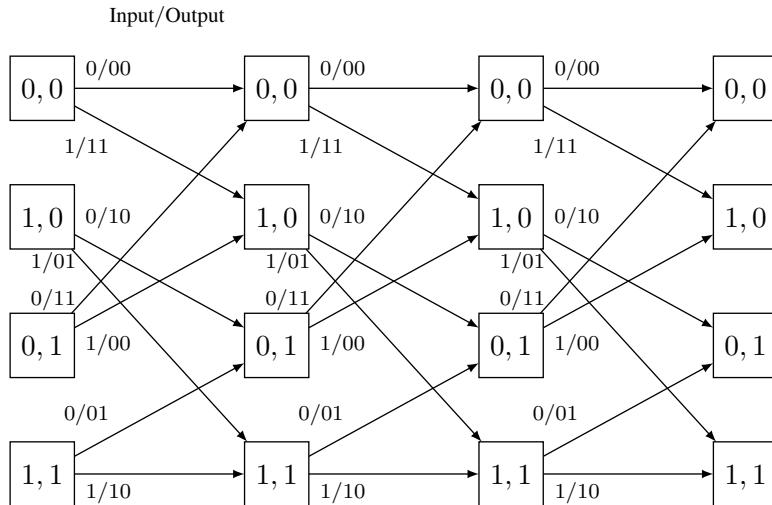


Figure 9.7 Trellis of a convolutional code. The state is described by S_1, S_2 .

The task of decoding is now to find the right path in the trellis from the first to the last square. A possible path may only be drawn on the arrows given. The information sequence can be read from the notation on the transitions. By observing some different possible sequences through the trellis, we realize that many sequences cross. For the comparison of the likelihood of two sequences that cross, we can split them into two parts each (one before and one after the cross). Independent of the second part we can compare their first respective part with respect to their likelihood. The part sequence with the smallest Euclidean distance is stored. If at the end it is decided that the most likely sequence runs through the state of the crossing we can come back to this stored part sequence. This way, the Viterbi algorithm walks through the trellis step by step. For every intermediate step the part sequence with the smallest Euclidean distance leading into each state is stored. This means that only as many paths and distance sums as there are states (in the following given by N) need to be stored. In the end there is a comparison of N path distances. In reality there are a few more subtleties to follow such that not the whole path history has to be stored. This however, shall not be detailed here.

9.3 Simulation of Transmission Systems

In order to prove the efficacy of a certain code, we need an objective criterion. Most often this is the so-called *bit error rate* (BER) after the decoding step, which means the ratio of bits⁴ in error to the total number of bits. The range of the BER usually lies between 0 and 0.5, where the latter is, of course, unusable. 50% bit error means that we may as well toss a coin to come up with the bits in the receiver. In order to test a code, the BER is being drawn as a function of the signal-to-noise ratio (SNR). Alternatively, the quotient of bit energy and noise power density E_b/N_o is used, which for a given system differs by a constant factor from the SNR. Both axes are normally scaled logarithmically. In comparison to the BER of the uncoded signal the *coding gain* can now be determined. To this end the SNR needed for a certain BER in the original BER curve (i.e., without coding) is compared against the SNR needed in the coded case. The latter is smaller for a suitable code and the difference is called the coding gain.

Finally, it shall be mentioned that very often not only single bit and symbol errors are of interest, but also whether a message can be detected without errors. Correspondingly, we talk of a message-error rate (MER).

⁴For non-binary codes one is sometimes interested in the *symbol error rate* (SER).

10 Trellis-Coded Modulation (TCM)

10.1 History and Motivation

Jim Massey proclaimed in his 1974 Zurich paper [24] that "modulation and coding have to be designed as an entity". Trellis-coded modulation (TCM) itself was first suggested by Gottfried Ungerboeck in 1976 and introduced in 1982 by [36]. Subsequently, more detailed papers about TCM were published by Ungerboeck in 1987, ([37] and [38]). TCM is a coding technique which improves the error performance of a data link without reducing the data rate or requiring more bandwidth. The difference between TCM and a classical digital communication system is that in a classical system the modulation and error-correction coding are separated, whereas TCM appears as the marriage of coding and modulation.

Data throughput in voiceband modems (for analog telephone lines) highly increased to 14 kbps and more after the year 1984 thanks to the deployment of TCM. The technologies used over the years are shown in Table 10.1. Prior to TCM deployment, there was a lack of understanding of the gap between the Shannon capacity of around 34 kbps (for the telephone channel with the SNR available) and the rates achieved using QAM only.

Standard	Date (ratified)	Speed (bps)	Modulation
v.21	1964	200	FSK
v.22	1980	1200	PSK
v.22 bis	1984	2400	QAM
v.23	1964	1200	FSK
v.26	1968	2400	PSK
v.26 bis	1972	2400	PSK
v.26 ter	1984	2400	PSK
v.27	1972	4800	PSK
v.27 bis	1976	4800	PSK
v.27 ter	1976	4800	PSK
v.29	1976	9600	QAM
v.32	1984	9600	QAM
v.32 bis	1991	14400	TCM
v.32 ter		19200	TCM
v.34 (v.fast)	1994	28800	TCM
v.24 bis	1995	33600	TCM
v.90	1998	56000	baseband

Table 10.1 Overview of voiceband modem standards.

10.2 Capacity Computation of Discrete-Alphabet Sources

10.2.1 Relationship to Mutual Information and Entropy

When dealing with the capacity of a channel, we must make clear we understand the conditions, i.e., the alphabet we want to send our information with. Usually we assume a discrete alphabet of a finite size, essentially symbols we want to map our information bits into. Very often, however, we start with a finite

alphabet, but at the end of the channel, the output variable looks continuous, due to additive noise. Certainly, we make decisions at the receiver, in order to retain the original information. By doing so, we might lose information as we shall see in due course. Hence, the capacity might depend on the condition that we make a decision or not.

Generally, the capacity is given as the mutual information between the channel input and the channel output

$$C = I(X; Y) = H(Y) - H(Y|X). \quad (10.1)$$

Using the definitions for the entropies

$$H(Y) \triangleq - \int_{-\infty}^{\infty} p(y) \log_2 p(y) dy \quad (10.2)$$

$$H(Y|X) \triangleq - \int_{-\infty}^{\infty} p(x) \int_{-\infty}^{\infty} p(y|x) \log_2 p(y|x) dy dx. \quad (10.3)$$

Now, using Eq. (10.1) the capacity can be written as

$$\begin{aligned} C &= - \int_{-\infty}^{\infty} p(y) \log_2 p(y) dy + \int_{-\infty}^{\infty} p(x) \int_{-\infty}^{\infty} p(y|x) \log_2 p(y|x) dy dx \\ &= - \int_{-\infty}^{\infty} p(x) \int_{-\infty}^{\infty} p(y|x) \log_2 p(y) dy dx + \int_{-\infty}^{\infty} p(x) \int_{-\infty}^{\infty} p(y|x) \log_2 p(y|x) dy dx \\ &= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(y|x) p(x) (\log_2 p(y) - \log_2 p(y|x)) dy dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(y|x) p(x) \log_2 \frac{p(y|x)}{p(y)} dy dx. \end{aligned} \quad (10.4)$$

Strictly speaking, the capacity is achieved if the expression is maximized over $p(x)$.

10.2.2 Continuous Variables

It can be shown, see App. D, that the maximization of the above defined capacity is obtained for a Gaussian distributed variable. The noise on the channel is Gaussian distributed, too. Thus, we can write

$$p(x) = \frac{1}{\sqrt{2\pi E_s}} e^{-\frac{x^2}{2E_s}}, \quad (10.5)$$

$$p(y|x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-x)^2}{2\sigma^2}}, \quad (10.6)$$

with E_s being the energy of the source and σ^2 being the variance of the noise. Since the sum of two independent Gaussian random variables is again Gaussian distributed with summed-up variances, we get

$$p(y) = \frac{1}{\sqrt{2\pi(E_s + \sigma^2)}} e^{-\frac{y^2}{2(E_s + \sigma^2)}}. \quad (10.7)$$

Using Eqs. (10.6) and (10.7) in Eq. (10.4) we can write

$$\begin{aligned}
 C &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(y|x) p(x) \log_2 \left(\frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-x)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi(E_s+\sigma^2)}} e^{-\frac{y^2}{2(E_s+\sigma^2)}}} \right) dy dx \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(y|x) p(x) \log_2 \left(\sqrt{\frac{E_s + \sigma^2}{\sigma^2}} e^{-\frac{(y-x)^2}{2\sigma^2} + \frac{y^2}{2(E_s + \sigma^2)}} \right) dy dx \\
 &= \int_{-\infty}^{\infty} p(x) \int_{-\infty}^{\infty} p(y|x) \left(\log_2 \left(\sqrt{\frac{E_s + \sigma^2}{\sigma^2}} \right) + \left(-\frac{(y-x)^2}{2\sigma^2} + \frac{y^2}{2(E_s + \sigma^2)} \right) \log_2 e \right) dy dx \\
 &= \frac{1}{2} \log_2 \left(\frac{E_s + \sigma^2}{\sigma^2} \right) - \frac{\log_2 e}{2\sigma^2} \int_{-\infty}^{\infty} p(x) \underbrace{\int_{-\infty}^{\infty} p(y|x) (y-x)^2 dy}_{\sigma^2} dx \\
 &\quad + \frac{\log_2 e}{2(E_s + \sigma^2)} \int_{-\infty}^{\infty} p(x) \underbrace{\int_{-\infty}^{\infty} p(y|x) y^2 dy}_{E_s + \sigma^2} dx \\
 &= \frac{1}{2} \log_2 \left(1 + \frac{E_s}{\sigma^2} \right), \tag{10.8}
 \end{aligned}$$

which we recognize as the Shannon capacity [34]. Note that this applies for a bandwidth of one Hertz. The factor 1/2 stems from the fact that we create a double-sideband signal, since binary signaling uses a real signal only.

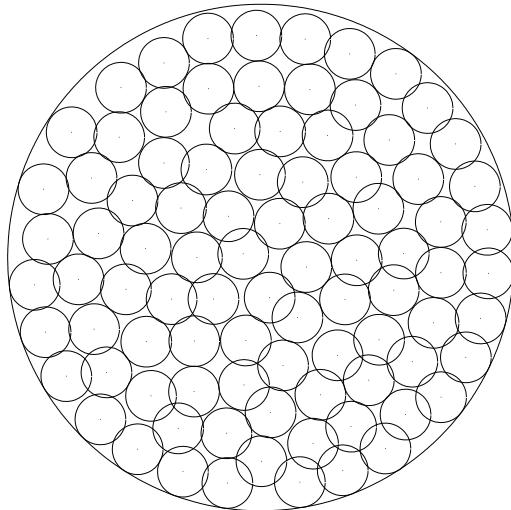


Figure 10.1 Sphere-packing illustration. The radii of the small spheres are $\sqrt{N\sigma^2}$, the radius of the containing sphere is $\sqrt{N(E_s + \sigma^2)}$.

The previous result can be intuitively understood, see also [35], by observing the reliable communication of some kind of coding scheme with K codewords producing dots in an N -dimensional sphere, where N is large. By the law of large numbers the received signal lies with high probability inside an N -dimensional sphere of radius $\sqrt{N(E_s + \sigma^2)}$, which corresponds to some sort of average cloud. More specifically, for

each possible symbol represented by a dot inside the overall sphere, the signal ends up near the surface of a noise sphere of radius $\sqrt{N\sigma^2}$ around the transmitted dot. Hence, in the N -dimensional space, the volumes of the K noise spheres should roughly correspond to the volume of the overall sphere, see Fig. 10.1,

$$K = \frac{V_{\text{overall sphere}}}{V_{\text{noise sphere}}} = \frac{\left(\sqrt{N(E_s + \sigma^2)}\right)^N}{\left(\sqrt{N\sigma^2}\right)^N}. \quad (10.9)$$

Hence, the capacity per channel use (we have to 'invest' N channel bits to get $\log_2 K$ info bits) is

$$C = \frac{1}{N} \log_2 K = \frac{1}{N} \log_2 \left(\frac{\left(\sqrt{N(E_s + \sigma^2)}\right)^N}{\left(\sqrt{N\sigma^2}\right)^N} \right) = \frac{1}{2} \log_2 \left(1 + \frac{E_s}{\sigma^2} \right). \quad (10.10)$$

10.2.3 Discrete-Input, Discrete-Output Variables

If both channel input and channel output are considered discrete, we talk about a discrete memoryless channel (DMC). The capacity of a DMC is given by [11], [28]

$$C_{\text{DMC}} = \max_{p(k)} \sum_{k=1}^K \sum_{j=1}^J p(j|k) p(k) \log_2 \frac{p(j|k)}{p(j)}. \quad (10.11)$$

For $M = 2$ this reduces to the case of the so-called binary-source channel (BSC), the transfer diagram of which is shown in Fig. 10.2.

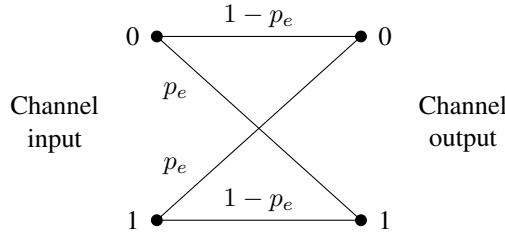


Figure 10.2 DMC for binary-sources, a so-called BSC.

By simplifying Eq. (10.11) and for the sake of symmetry assuming $p(1) = p(2) = \frac{1}{2}$ we get

$$C_{\text{BSC}} = \frac{1}{2} \left(p(1|1) \log_2 \frac{p(1|1)}{\frac{1}{2}} + p(2|1) \log_2 \frac{p(2|1)}{\frac{1}{2}} + p(1|2) \log_2 \frac{p(1|2)}{\frac{1}{2}} + p(2|2) \log_2 \frac{p(2|2)}{\frac{1}{2}} \right). \quad (10.12)$$

Using the bit-error probability $p_e = p(1|2) = p(2|1)$ and $p(1|1) = p(2|2) = 1 - p_e$ we get

$$\begin{aligned} C_{\text{BSC}} &= (1 - p_e)(\log_2(1 - p_e) + 1) + p_e(\log_2(p_e) + 1) \\ &= (1 - p_e)\log_2(1 - p_e) + (1 - p_e) + p_e\log_2(p_e) + p_e \\ &= 1 + (1 - p_e)\log_2(1 - p_e) + p_e\log_2(p_e) \\ &= 1 - h(p_e), \end{aligned} \quad (10.13)$$

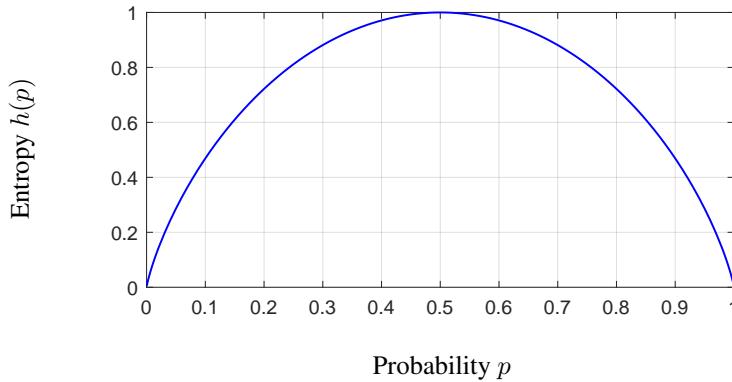


Figure 10.3 Binary entropy function.

where the binary entropy function is given by

$$h(p) = -p \log_2 p - (1-p) \log_2(1-p), \quad (10.14)$$

displayed in Fig. 10.3.

The error probability itself results from the evaluation of the tail function $Q(\cdot)$ of the Gaussian distribution

$$p_e = Q\left(\sqrt{\frac{E_s}{\sigma^2}}\right). \quad (10.15)$$

For $M = 4$ not all error probabilities are equal since the distances between diagonal symbols are different from the ones between neighbouring symbols, see Fig. 10.4. The error probability between diagonal symbols is obtained when considering enough noise in both the real and the quadrature component (in the complex baseband representation), hence

$$p_{ee} = Q^2\left(\sqrt{\frac{E_s}{2\sigma^2}}\right). \quad (10.16)$$

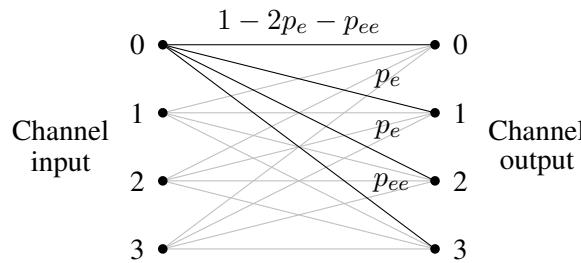


Figure 10.4 DMC for QPSK.

The error probability for neighbouring symbols, on the other hand, is obtained by multiplying the independent probability that an error occurs in one of the two dimensions but not in the other,

$$p_e = Q\left(\sqrt{\frac{E_s}{2\sigma^2}}\right) \cdot \left(1 - Q\left(\sqrt{\frac{E_s}{2\sigma^2}}\right)\right). \quad (10.17)$$

The capacity of a QPSK signal is now

$$C_{\text{QPSK}} = \log_2(4) + 2p_e \log_2(p_e) + p_{ee} \log_2(p_{ee}) + (1 - 2p_e - p_{ee}) \log_2(1 - 2p_e - p_{ee}). \quad (10.18)$$

10.2.4 Discrete-Input, Continuous-Output Variables

Now we leave the input alphabet the same but consider continuous noise. If we estimate the capacity prior to making a decision on the receiver side, we get other expressions for the capacity. Usually we assume a regular discrete-input alphabet with $p(x)$ being the same for all discrete points. In this case we can replace the integration over x by a sum, yielding

$$\begin{aligned}
 C &= \frac{1}{M} \sum_{m=1}^M \int_{-\infty}^{\infty} p(y|x_m) \log_2 \frac{p(y|x_m)}{p(y)} dy \\
 &= \frac{1}{M} \sum_{m=1}^M \int_{-\infty}^{\infty} p(y|x_m) \log_2 \frac{p(y|x_m)}{\frac{1}{M} \sum_{m=1}^M p(y|x_m)} dy \\
 &= \frac{1}{M} \sum_{m=1}^M \int_{-\infty}^{\infty} p(y|x_m) \left(\log_2 M + \log_2 \frac{p(y|x_m)}{\sum_{m=1}^M p(y|x_m)} \right) dy \\
 &= \log_2 M + \frac{1}{M} \sum_{m=1}^M \int_{-\infty}^{\infty} p(y|x_m) \log_2 \frac{p(y|x_m)}{\sum_{m=1}^M p(y|x_m)} dy. \tag{10.19}
 \end{aligned}$$

For $M = 2$ (binary source) we have

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma} \left(e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} + e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \right) \tag{10.20}$$

and

$$p(y|x = \sqrt{E_s}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}}, \tag{10.21}$$

$$p(y|x = -\sqrt{E_s}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}}. \tag{10.22}$$

All distributions above are illustrated in Fig. 10.5.

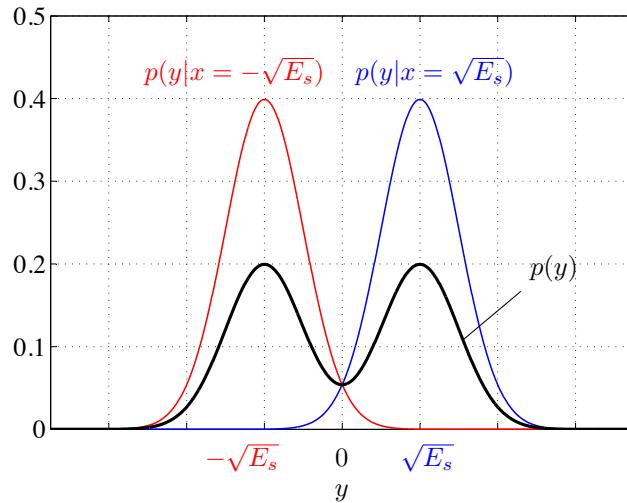


Figure 10.5 Binary source probability distributions.

Using the above distributions we get for the capacity of a binary input, continuous output channel

$$\begin{aligned}
 C &= \log_2 2 + \frac{1}{2} \left(\int_{-\infty}^{\infty} p(y|x = \sqrt{E_s}) \log_2 \frac{p(y|x = \sqrt{E_s})}{p(y|x = \sqrt{E_s}) + p(y|x = -\sqrt{E_s})} dy \right. \\
 &\quad \left. + \int_{-\infty}^{\infty} p(y|x = -\sqrt{E_s}) \log_2 \frac{p(y|x = -\sqrt{E_s})}{p(y|x = \sqrt{E_s}) + p(y|x = -\sqrt{E_s})} dy \right) \\
 &= 1 + \frac{1}{2} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} \log_2 \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi}\sigma} \left(e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} + e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \right)} dy \right. \\
 &\quad \left. + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \log_2 \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi}\sigma} \left(e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} + e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \right)} dy \right) \\
 &= 1 + \frac{1}{2} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} \log_2 \frac{e^{-\frac{y^2}{2\sigma^2}} e^{\frac{y\sqrt{E_s}}{\sigma^2}} e^{-\frac{E_s}{2\sigma^2}}}{e^{-\frac{y^2}{2\sigma^2}} e^{-\frac{E_s}{2\sigma^2}} \left(e^{\frac{y\sqrt{E_s}}{\sigma^2}} + e^{-\frac{y\sqrt{E_s}}{\sigma^2}} \right)} dy \right. \\
 &\quad \left. + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \log_2 \frac{e^{-\frac{y^2}{2\sigma^2}} e^{-\frac{y\sqrt{E_s}}{\sigma^2}} e^{-\frac{E_s}{2\sigma^2}}}{e^{-\frac{y^2}{2\sigma^2}} e^{-\frac{E_s}{2\sigma^2}} \left(e^{\frac{y\sqrt{E_s}}{\sigma^2}} + e^{-\frac{y\sqrt{E_s}}{\sigma^2}} \right)} dy \right) \\
 &= 1 + \frac{1}{2} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} \log_2 \frac{e^{\frac{y\sqrt{E_s}}{\sigma^2}}}{e^{\frac{y\sqrt{E_s}}{\sigma^2}} + e^{-\frac{y\sqrt{E_s}}{\sigma^2}}} dy \right. \\
 &\quad \left. + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \log_2 \frac{e^{-\frac{y\sqrt{E_s}}{\sigma^2}}}{e^{\frac{y\sqrt{E_s}}{\sigma^2}} + e^{-\frac{y\sqrt{E_s}}{\sigma^2}}} dy \right) \\
 &= 1 + \frac{1}{2} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} \frac{y\sqrt{E_s}}{\sigma^2} \log_2 e dy - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} \log_2 \left(e^{\frac{y\sqrt{E_s}}{\sigma^2}} + e^{-\frac{y\sqrt{E_s}}{\sigma^2}} \right) dy \right. \\
 &\quad \left. - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \frac{y\sqrt{E_s}}{\sigma^2} \log_2 e dy - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \log_2 \left(e^{\frac{y\sqrt{E_s}}{\sigma^2}} + e^{-\frac{y\sqrt{E_s}}{\sigma^2}} \right) dy \right) \\
 &= 1 + \frac{E_s}{\sigma^2} \log_2 e - \frac{1}{2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \left(e^{-\frac{(y-\sqrt{E_s})^2}{2\sigma^2}} + e^{-\frac{(y+\sqrt{E_s})^2}{2\sigma^2}} \right) \log_2 \left(e^{\frac{y\sqrt{E_s}}{\sigma^2}} + e^{-\frac{y\sqrt{E_s}}{\sigma^2}} \right) dy \\
 &= 1 + \frac{E_s}{\sigma^2} \log_2 e - \frac{1}{2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} e^{-\frac{E_s}{2\sigma^2}} \left(2 \cosh \frac{y\sqrt{E_s}}{\sigma^2} \right) \log_2 \left(2 \cosh \frac{y\sqrt{E_s}}{\sigma^2} \right) dy \\
 &= 1 + \frac{E_s}{\sigma^2} \log_2 e - \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{E_s}{2\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} \left(\cosh \frac{y\sqrt{E_s}}{\sigma^2} \right) \log_2 \left(2 \cosh \frac{y\sqrt{E_s}}{\sigma^2} \right) dy \\
 &= \frac{E_s}{\sigma^2} \log_2 e - \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{E_s}{2\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} \left(\cosh \frac{y\sqrt{E_s}}{\sigma^2} \right) \log_2 \left(\cosh \frac{y\sqrt{E_s}}{\sigma^2} \right) dy. \tag{10.23}
 \end{aligned}$$

10.2.5 Numerical Evaluation

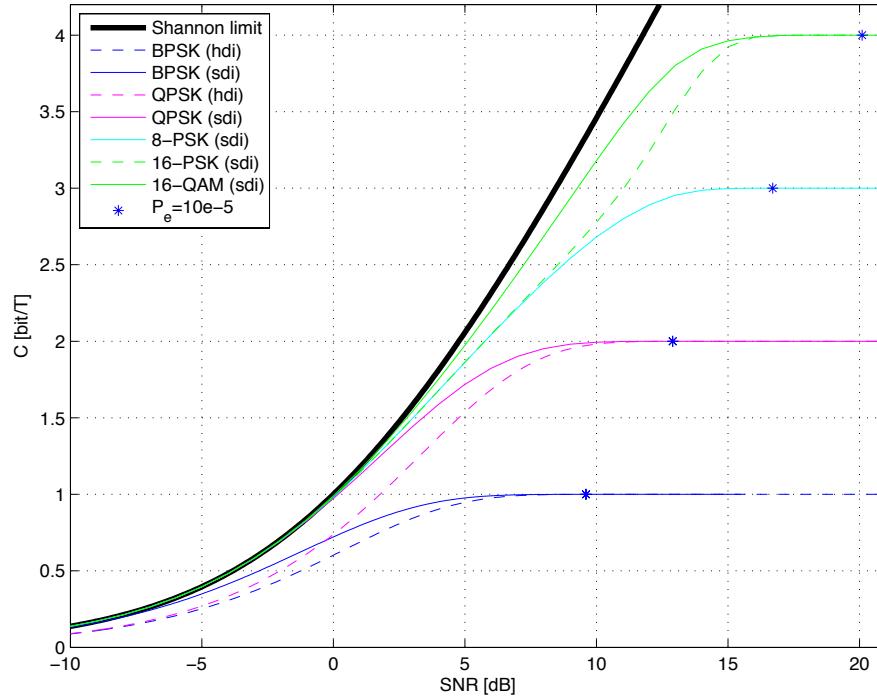


Figure 10.6 Capacity of different digital modulation schemes.

The general expression given by Eq. (10.19) is now evaluated numerically. The result can be seen in Fig. 10.6. It is interesting to discuss a couple of items in more detail.

For either BPSK and QPSK, the capacity was worked out for discrete and continuous channel output. For either one we can see a benefit of around 1 dB if the channel output is continuous. This tells us that in the noisy case it might be of advantage not to make decisions too early. Obviously we lose some information when making hard decisions as opposed to soft decision, for which we carry some reliability information to a higher layer.

Another interesting fact of the curves shown in Fig. 10.6 to look at is the difference in capacity between 16-PSK and 16-QAM. Not restricting ourselves to a unit amplitude, we have one more degree of freedom for the latter constellation. The increase in capacity is therefore not unexpected.

Finally, for high SNR values, the capacity of low alphabet sizes does not exploit a channel well, since the curves all flatten to the right, far away from Shannon's capacity. For low SNR values, on the other hand, only the soft-decision curves approach the Shannon limit closely. In summary, for good SNR values, capacity is optimized if the constellation is chosen high enough. Likewise, for poor SNR situations, capacity is achieved regardless of the alphabet size, as long as the channel output stays continuous (soft-decision information).

Ungerboeck [36] now noticed the following: Starting out with QPSK at the point where a BER of 10^{-5} is obtained, roughly around 12.9 dB of SNR, we can back off the SNR to 5.9 dB and still get the same capacity in theory (thus moving to the left on a horizontal line) when introducing 8-PSK. Still higher constellations would not increase the benefit much, however. Ungerboeck then concluded that by doubling the alphabet size, almost all the channel capacity that is achievable is gained.

10.3 Principles of TCM

The conventional approach to improved robustness against channel errors is to use channel coding. To this end, block codes or convolutional codes insert redundancy in the form of parity bits into the bitstream. The price for this is a decrease in data rate. In order to compensate for the lower net data rate, either the power must be increased (and a larger constellation used) or the bandwidth needs be expanded, or both. Thus, using channel coding, a lower bit error rate (BER) is possible at the price of lower bandwidth efficiency.

The principle of TCM is to expand the signal constellation to provide redundancy for coding. The constellation is typically doubled, i.e., that an extra bit at the input is needed. The input bits are encoded with a special convolutional code. The output bits of the encoder are then fed to the signal-mapping block, where the constellation has been partitioned in a special manner known as *set-partitioning*. The concept of set partitioning is of central significance for TCM schemes. Set partitioning divides a signal set successively into smaller subsets with maximally increasing the smallest intra-set distances. That way, the minimum Euclidean distance between coded signal sequences is maximized. Note that neither the power nor the bandwidth has been increased.

At the receiving end, the noisy signals are decoded by a soft-decision maximum-likelihood sequence decoder. In practice, a Viterbi decoder is used because it is an optimal, recursive solution to the problem of estimating the encoded sequence [10]. Thus, with a simple four-state Ungerboeck code, the robustness of a digital transmission against additive noise can be improved by 3 dB compared to conventional uncoded modulation. With more complex codes, the coding gain can reach 6 dB or more.

The way the Viterbi decoder works, there is always a minimum decoding delay (trace-back depth) needed that amounts to five times the constraint length of the convolutional code used, so that the Viterbi decoder can determine the output. Thus, for a four-state code a minimum delay of $5 \cdot 3 = 15$ TCM symbols has to be considered. The trace-back path must always start and end in a known state (e.g., all-zero state). This means that the TCM encoder must insert enough zeros at the end of the binary data. When using the four-state code, the last two TCM symbols must contain redundancy only. Fig. 10.7 shows the block diagram of a TCM system. The decoding process is similar to the decoding of a convolutional code: The

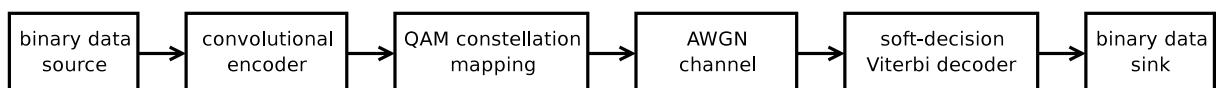


Figure 10.7 Block diagram of a TCM system.

received sequence is compared with all possible (and valid) paths through the trellis. The most likely one is then picked (maximum likelihood decoding). This is the one with the smallest distance to the received sequence.

10.4 Definitions of Parameters

10.4.1 Minimum Distance

While in coding theory the Hamming distance between code words is of importance and of interest, with respect to TCM we are interested in the so-called Euclidean distance $d(k, l)$, or rather the squared Euclidean distance $d^2(k, l)$, between two constellation points. In the complex constellation plane, this measure can be expressed as a sum of squared differences of the real and the imaginary parts of two constellation points P_k and P_l ,

$$d^2(k, l) = (\text{Re}(P_k) - \text{Re}(P_l))^2 + (\text{Im}(P_k) - \text{Im}(P_l))^2. \quad (10.24)$$

The minimum squared Euclidean distance d_{\min}^2 , sometimes also written as Δ_0^2

$$d_{\min}^2 = \min_{k,l} d^2(k, l) \quad (10.25)$$

is the squared distance between two neighboring constellation points. For 8-PSK, e.g., the minimum squared Euclidean distance can be derived by the law of cosines as

$$d_{\min}^2 = 1 + 1 - 2 \cos \alpha = 2 - 2 \cos 45^\circ = 2 - \sqrt{2} = 0.586, \quad (10.26)$$

if the average energy of a symbol is normalized to one, as shown in Fig. 10.8.

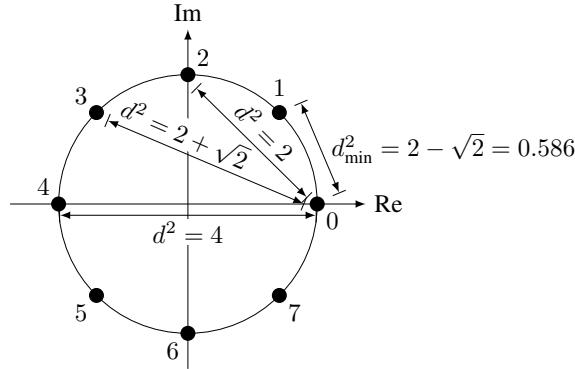


Figure 10.8 Minimum (squared) Euclidean distance in a constellation plot (unit circle) for 8-PSK.

10.4.2 Free Distance

We are also interested in another parameter: the Euclidean free distance d_{free}^2 is the smallest distance between two valid sequences. As such, it is the sum of the Euclidean distances between constellation points corresponding to the states belonging to different paths that diverge at one point and converge at a later one through the trellis. A mathematical definition for the free distance can be given according to Eq. (10.25)

$$d_{\text{free}}^2 = \min_{\{a_k\} \neq \{a'_k\}} \sum_k d^2(a_k, a'_k), \quad (10.27)$$

where the sequences $\{a_k\}$ and $\{a'_k\}$ represent all possible combinations the encoder can produce.

10.4.3 Coding Gain

The asymptotic coding gain G is defined as the reduction in the SNR that a TCM signal can allow in order to achieve the same BER as an uncoded signal. Hence,

$$G = \frac{d_{\text{free,coded}}^2}{d_{\min,\text{uncoded}}^2}. \quad (10.28)$$

10.5 A Simple Introductory Example

Imagine we have a QPSK system with a symbol period of T seconds. The power spectral density is then proportional to

$$\frac{\sin^2(\pi f T)}{(\pi f T)^2}, \quad (10.29)$$

with a null-to-null bandwidth of $B = 2/T$ and a bit rate of $2/T$. If, for some reason, the SNR is not sufficient to deliver the BER required, we can use a convolutional code with rate $2/3$ prior to the QPSK modulator. Since the number of channel bits (as opposed to information bits) is increased by 50%, the bandwidth required is also increased by 50%. This might not always be possible.

If however, and this is crucial to the design of a TCM system, that same convolutional code as mentioned above is followed by an 8-PSK modulator, the bandwidth required remains the same. At first sight, the constellation used is denser and might be prone to channel symbol errors more easily. The coding gain of the convolutional code just prior to the modulation stage has to compensate this reduction in distance of the constellation points. In reality, this is done by not bothering about the minimum Euclidean distance between

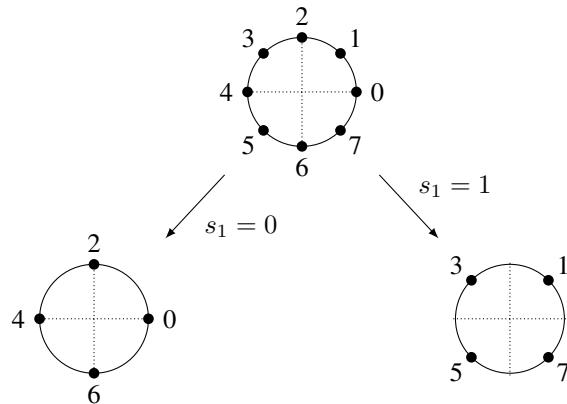


Figure 10.9 Set partitioning of an 8-PSK constellation into 2 sets.

two neighboring constellation points, but by concentrating instead on the free Euclidean distance between two valid sequences produced by the convolutional code chosen.

In our example we have the information bits x_2 and x_1 , which are encoded into the channel bits y_2 , y_1 and y_0 , respectively. Note that y_2 is the same as x_2 . Because y_2 is 'uncoded', i.e., it is independent of other channel bits, the distance in the constellation needs to be increased by letting y_2 choose among points that are not neighbors. Thus, we make two sets of 'half-constellations', such that the minimum distance in each is maximized, see Fig. 10.9.

The choice of the constellation alphabet can now be described as follows. We define a state machine with two states. If the state is 0, then alphabet $\{0,2,4,6\}$ is used. If the state is 1, then alphabet $\{1,3,5,7\}$ is used. We can quickly draw the state diagram of this procedure and come up with something like Fig. 10.10. If the

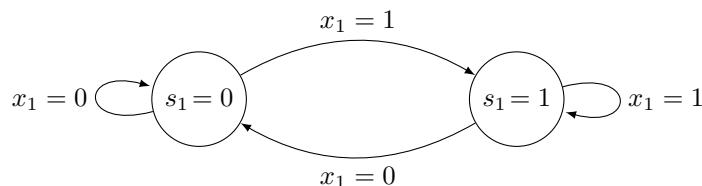


Figure 10.10 State diagram for a two-state TCM. Note that a rate 1/2 or a rate 2/3 TCM will result in the same state diagram.

state diagram is deconvolved in time, we get a trellis. Very often a trellis is assigned all kind of numbers and the reader gets confused easily. In order not to overload the diagram, we only show the output variable $y = y_1y_0$, see Fig. 10.11. The encoder that produces exactly that is rather simple, see Fig. 10.12. The value

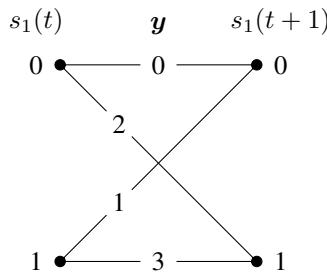


Figure 10.11 Trellis for a two-state TCM, $\mathbf{y} = y_1 y_0$, whereas y_2 is not shown. Note that rate-1/2 and rate-2/3 TCM result in the same trellis diagram, as long as no parallel transitions are shown.

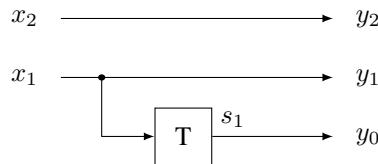


Figure 10.12 Encoder for the constellation index for a two-state rate 2/3 TCM.

of the register (one bit only) is referred to as the state s_1 . Additionally to the two bits y_1 and y_0 , we also have y_2 . This bit is now taken to determine whether a certain constellation point or the one sitting on the opposite side of the constellation is chosen. Together, the triplet y_2, y_1, y_0 is the binary code for the point in the original 8-PSK constellation at the top of Fig. 10.9. This leads to parallel transitions in the trellis, see Fig. 10.13. Now we want to determine what this simple TCM has brought us in terms of gain. To this end

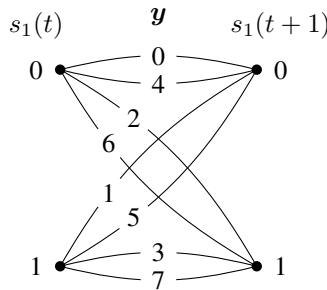


Figure 10.13 Trellis for a two-state rate 2/3 TCM including parallel transitions, $\mathbf{y} = y_2 y_1 y_0$.

we must look for the smallest distance between two alternative routes starting and ending in the same nodes. Thus, we redraw the trellis as in Fig. 10.14.

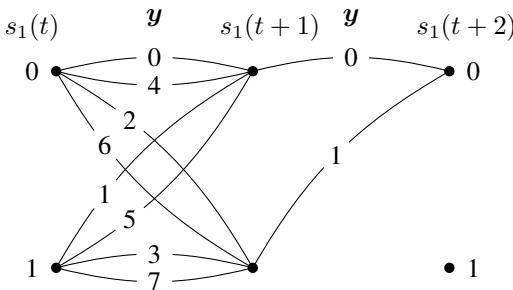


Figure 10.14 Trellis for a two-state rate 2/3 TCM for one example of alternative routes, $\mathbf{y} = y_2 y_1 y_0$.

The free Euclidean distance is now given as

$$d_{\text{free}}^2 = d^2(0, 2) + d^2(0, 1) = 2 + (2 - \sqrt{2}) = 2.5858. \quad (10.30)$$

Note that we also have to check the parallel transitions. In this case we have $d^2(0, 4) = 4$, which is not limiting. Compared to the QPSK case, where $d_{\text{min}}^2 = 2$ we obtain a coding gain of

$$G = \frac{2.5858}{2} = 1.2929 = 1.12 \text{ dB}. \quad (10.31)$$

10.6 Improving the Performance (the Next Two Examples)

The coding gain obtained for this simple example is not exhilaratingly high. We can do better by introducing more memory into the encoder. The extended encoder is shown in Fig. 10.15, for the state diagram as depicted in Fig. 10.16.

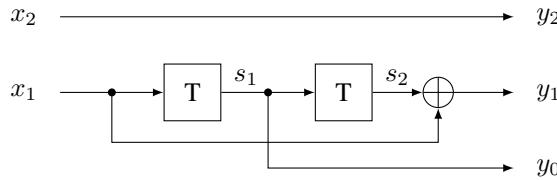


Figure 10.15 Encoder for a four-state rate 2/3 TCM.

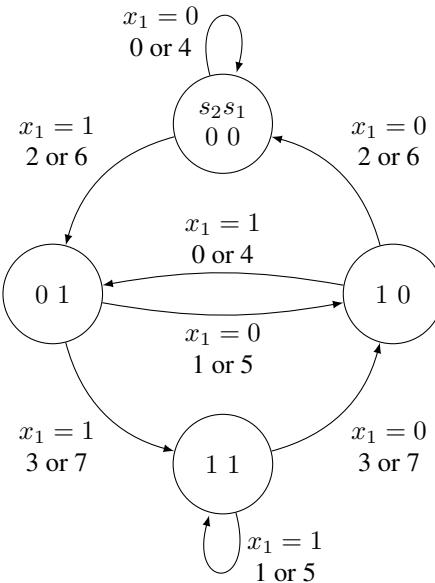


Figure 10.16 State diagram for a four-state TCM.

We can now interpret the signal y_2 as the index within each of the four alphabets, namely $\{0,4\}$, $\{2,6\}$, $\{1,5\}$ and $\{3,7\}$. Which of the four alphabets is chosen, is determined by y_1 and y_0 . The set partitioning of the 8-PSK constellation can be seen in Fig. 10.17. The input signal x_2 is directly connected to the signal mapping block. Thus, again *parallel transitions* are generated, as can be seen in Fig. 10.18. This time, the

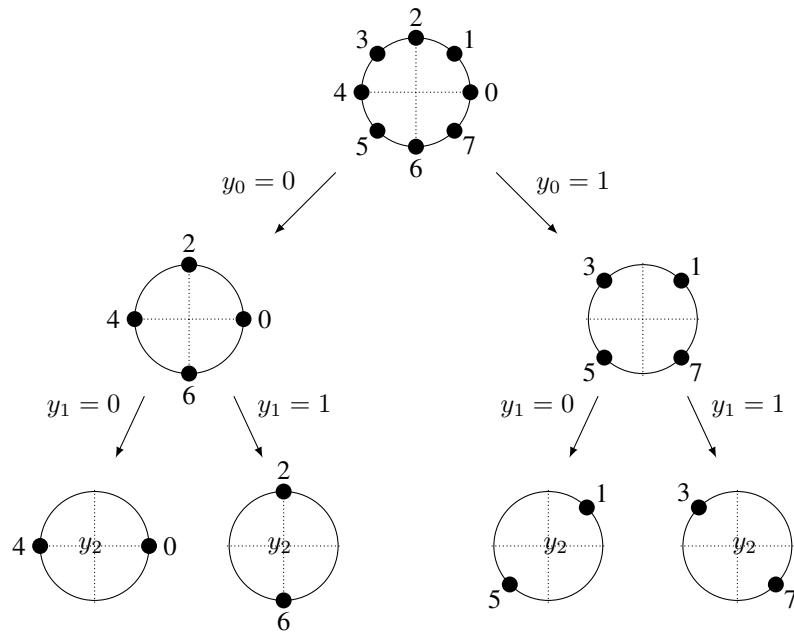


Figure 10.17 Set partitioning of an 8-PSK constellation into 4 sets.

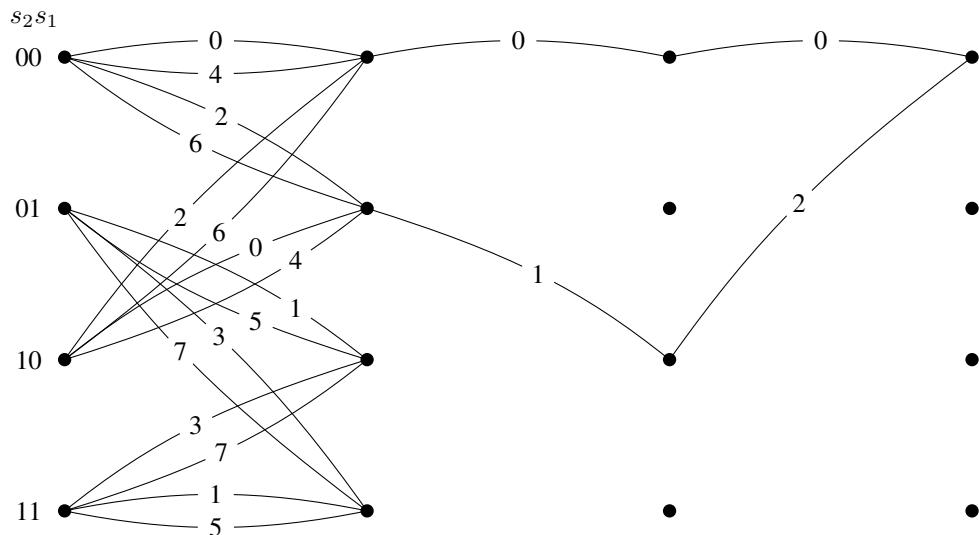


Figure 10.18 Trellis for a four-state rate 2/3 TCM.

parallel transitions are the limiting case, because $d^2(0, 4) < d^2(0, 2) + d^2(0, 1) + d^2(0, 2) = 4.5858$ and we have

$$d_{\text{free}}^2 = d^2(0, 4) = 4, \quad (10.32)$$

$$G = \frac{4}{2} = 2 = 3 \text{ dB}. \quad (10.33)$$

In order to obtain a still higher gain in the case of a 8-PSK TCM, we must get rid of parallel transitions. That can be achieved using eight states. The corresponding trellis is shown in Fig. 10.19. A circuit diagram producing this trellis is no longer straightforward and shall be omitted here. We get the following free

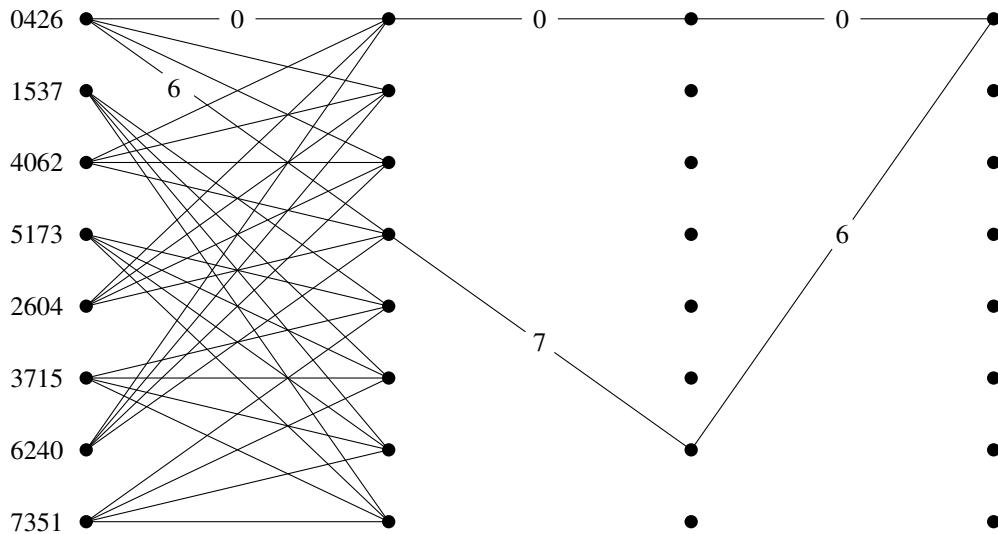


Figure 10.19 Trellis for an eight-state rate 2/3 TCM. For ease of reading the constellation indices are not written on the branches as previously but to the left of the node in the order from top to bottom.

distance and asymptotic gain:

$$d_{\text{free}}^2 = d^2(0, 6) + d^2(0, 7) + d^2(0, 6) = 2 + (2 - \sqrt{2}) + 2 = 4.586, \quad (10.34)$$

$$G = \frac{4.586}{2} = 3.6 \text{ dB}. \quad (10.35)$$

10.7 Encoding and Signal Mapping

As a general procedure, see Fig. 10.20, we can say that of the m bits that define the original symbol, $m - \tilde{m}$ select a signal within the subset. These $m - \tilde{m}$ bits generate parallel transitions. The remaining \tilde{m} bits select the subset. Usually, \tilde{m} is 1 or 2. These \tilde{m} bits are used in a convolutional encoder¹ of rate $\tilde{m}/(\tilde{m} + 1)$. Hence, $m + 1$ bits are passed to the signal-mapping block, so that the constellation is doubled as compared with the original constellation defined by m bits.

We can now understand the efficacy of a TCM system. The uncoded part of the bits are mapped to signals within a subset that are as distant of each other as possible. In order to protect the subset chosen, redundancy is introduced in the form of a convolutional code prior to choosing the subset.

10.8 Set Partitioning

Set partitioning, due to Ungerboeck [36] is the process of defining the signal mapping after the convolutional encoder. At each step, the constellation is halved into two subsets with increasing minimum distance. Fig. 10.21 serves as an example of this process. The following three rules are called the Ungerboeck rules:

- To parallel transitions are assigned members of the same partition.

¹If the convolutional encoder is linear and of rate $\tilde{m}/(\tilde{m} + 1)$ we have a so-called Ungerboeck code.

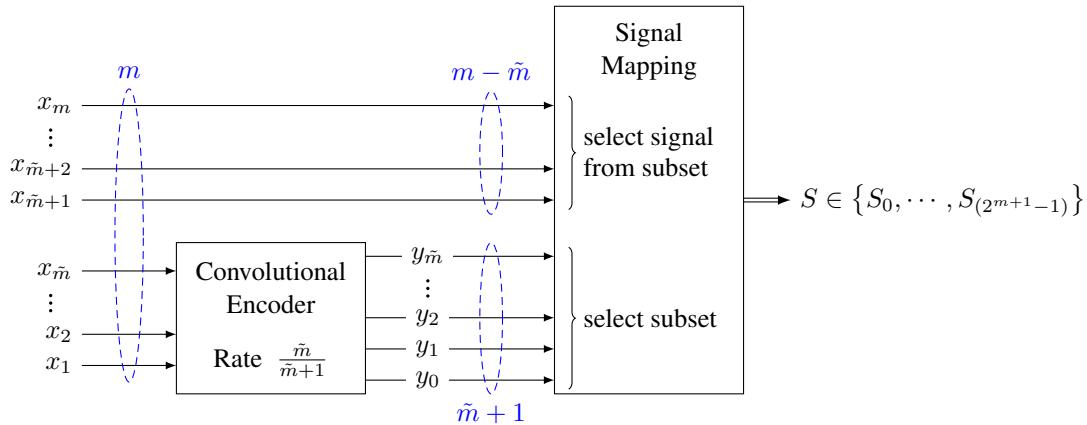


Figure 10.20 Encoding and signal mapping for TCM.

- To adjacent transitions² are assigned members of the next larger partition (because it takes at least two cycles to get an alternative path).
- All signals are used equally often.

Note that for a two-state rate 2/3 TCM as given above, the second Ungerboeck rule can only be satisfied for the branches emanating from the same state but not for the paths ending in the same state (they are from different partitions). Thus, this TCM does not follow all the Ungerboeck rules for a good code and the coding gain is rather limited with a coding gain of only 1.1 dB. Some examples of TCM schemes and their corresponding asymptotic gains are listed in Table 10.2.

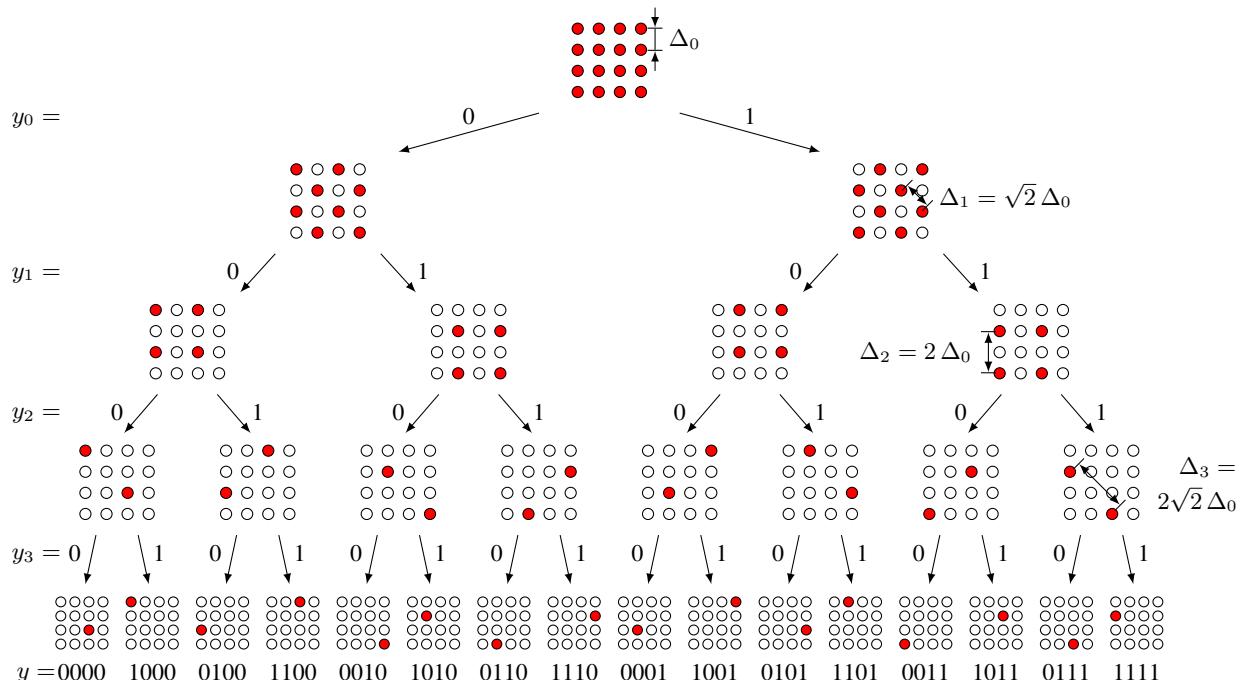


Figure 10.21 Set partitioning for 16-QAM.

²Adjacent transitions are called paths diverging from the same state or paths terminating in the same state.

\tilde{m}	Number of states	m	Rate	Constellation	Asympt. gain G
1	2	1	1/2	BPSK/QPSK	1.76 dB
1	2	2	2/3	QPSK/8-PSK	1.1 dB
1	4	2	2/3	QPSK/8-PSK	3.0 dB
2	8	2	2/3	QPSK/8-PSK	3.6 dB
2	8	3	3/4	8-PSK/16-PSK	3.98 dB

Table 10.2 Asymptotic gain of some selected TCM schemes.

10.9 Performance

The performance of a TCM system is directly depending on the memory depth of the convolutional encoder and the constellation size. The complexity of the Viterbi decoder grows exponentially with the memory depth. For a four-state TCM system, the performance is compared with conventional QAM in Fig. 10.22.

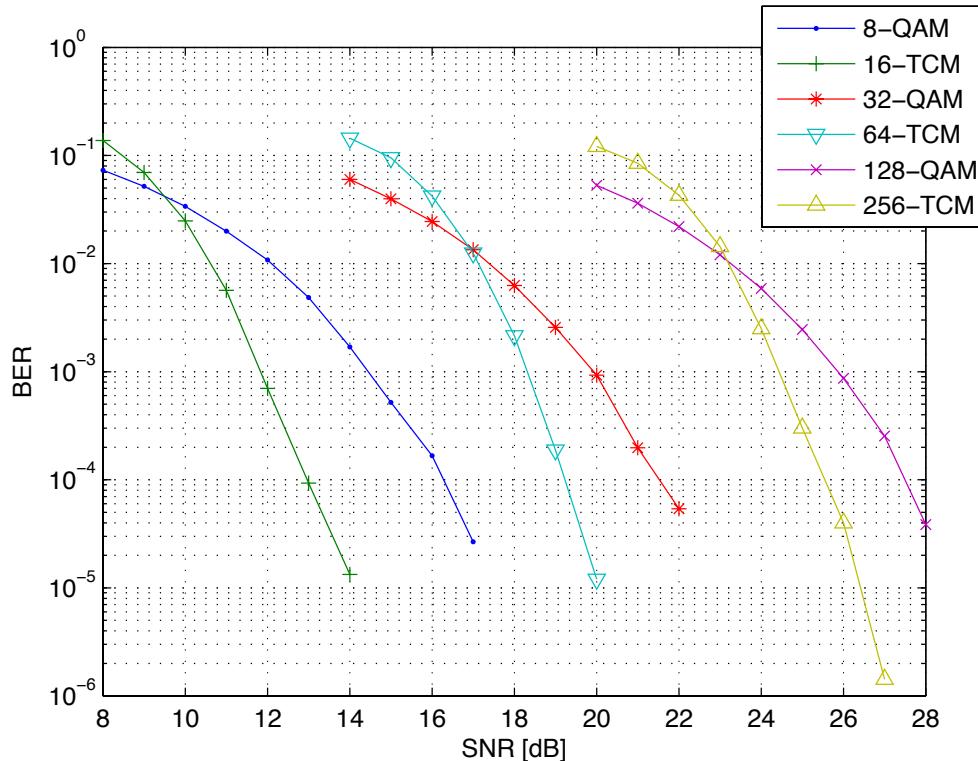


Figure 10.22 Performance of TCM (4 states) vs. conventional QAM.

The SNR values at a BER of 10^{-5} are in the following compared against the channel capacity, which was given previously as

$$C = W \cdot \log_2 \left(1 + \frac{S}{N} \right). \quad (10.36)$$

Using $S = E_b \cdot C$ and $N = N_0 \cdot W$, we can write

$$\frac{C}{W} = \log_2 \left(1 + \frac{E_b}{N_0} \cdot \frac{C}{W} \right), \quad (10.37)$$

or

$$2^{\frac{C}{W}} = 1 + \frac{E_b}{N_0} \cdot \frac{C}{W}, \quad (10.38)$$

which can be reformulated³ into

$$\frac{E_b}{N_0} = \frac{2^{\frac{C}{W}} - 1}{\frac{C}{W}}. \quad (10.39)$$

This expression and some results for QAM and for TCM are drawn in Fig. 10.23. Note the subtle differences to Fig. 10.6. It can clearly be seen, that TCM closes almost half the gap between uncoded modulation and the theoretical capacity.

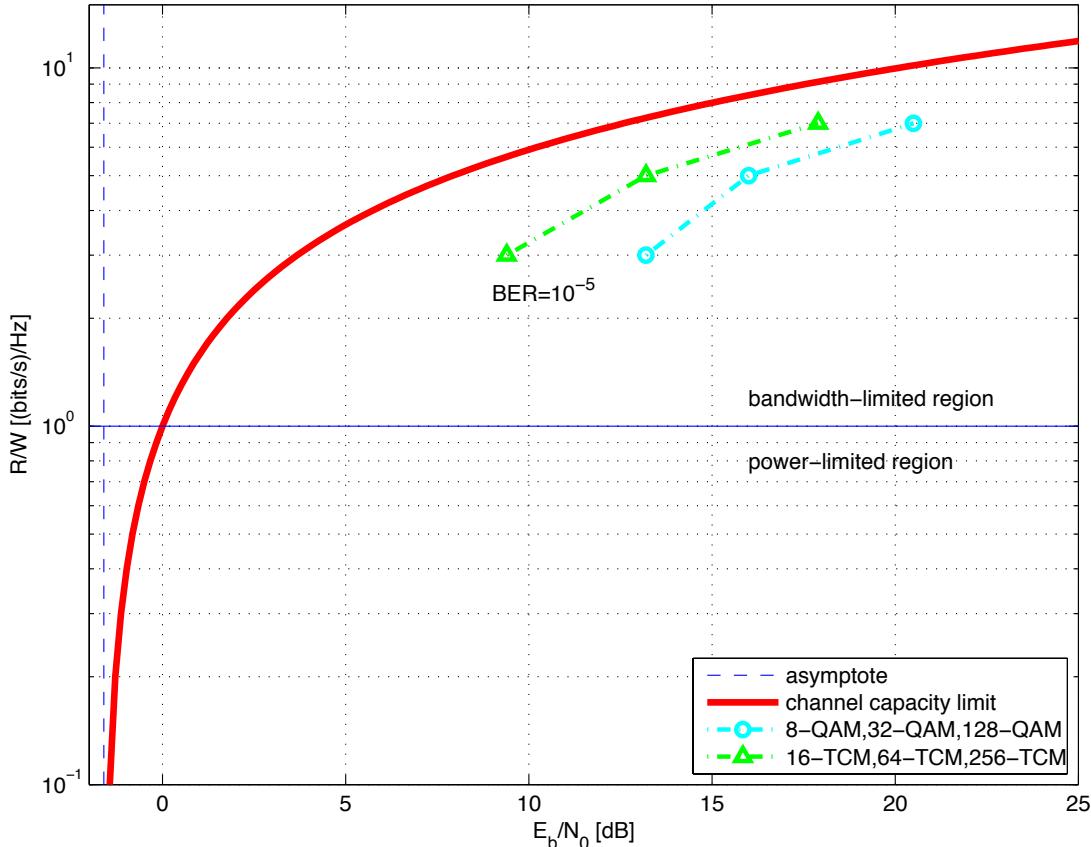


Figure 10.23 Comparison of the gaps to channel capacity.

³By the way, Eq. (10.39) also yields the absolute limit (asymptote) of the $\frac{E_b}{N_0}$, for which transmission is possible at all:

$$\lim_{\frac{C}{W} \rightarrow 0} \frac{E_b}{N_0} = \ln 2 = -1.6 \text{ dB}. \quad (10.40)$$

10.10 Philosophy of TCM

In summary, we note the procedure for designing a TCM system:

- Use a special convolutional code with rate $R = k/(k + 1)$ to add redundancy.
- The constellation is doubled and numbered using *set-partitioning*.
- Only one half of the potential constellation set is used during a single transmission.
- Using a soft-decision Viterbi decoder the original bits are reconstructed.

Books that contain chapters on TCM are [31], [32]. A very detailed book devoted entirely to TCM is the one by Biglieri *et al.* [4].

10.11 Extensions of TCM

10.11.1 Rotation-Invariant TCM

TCM codes are not automatically invariant to phase rotations as they can happen in a wireless fading channel. A constellation in baseband often looks the same regardless if it is at its original phase or at multiples of 90° . The sequences through the trellis, however, only works if the constellation is at its right position, unless special care has been taken to produce a rotation-invariant TCM. To do so, nonlinear elements have to be incorporated into the encoder.

10.11.2 Multidimensional TCM

Similar to the extension from real-valued to complex-valued constellations, a further extension to multidimensional constellations can be of advantage in TCM. But how, beside going to complex numbers, can more dimensions be gained? Dimensions can be obtained by using different channels or by using orthogonal field polarization.

10.12 Real-World Example: Combination of OFDM and TCM

Many wireless audio systems used in conjunction with hearing aids are narrowband analog FM links running on 169 MHz with an RF bandwidth of 50 kHz. For certain applications it is of interest to use the link for a wireless digital data transfer. The same spectral RF requirements and transmit restrictions apply as with the analog FM link. Hence, the problem at hand is one of designing a wireless data link on top of an FM modem using the highest possible QAM constellation in order to maximize throughput. There is no feedback on the channel and latency times shall be minimized. Moreover, computational complexity should not exceed the possibilities of low-priced microcontrollers or DSPs, possibly enhanced by an FFT engine.

A first measurement of the FM modem channel revealed an SNR of around 30 dB, when considering the power constraint due to “back-off”, yet to be introduced. With a usable submodulation bandwidth of 5.5 kHz, the channel capacity is roughly 55 kbit/s. This number serves as a very loose upper bound on the data rate to expect.

10.12.1 Modem Architecture

The basic architecture of the data modem designed can be seen in Fig. 10.24. In OFDM, the modulator can be implemented as an N -point *inverse discrete Fourier transform* (IDFT) on a block of N data symbols followed by a *digital-to-analog converter* (DAC) on the IDFT samples [39]. For the following we assume

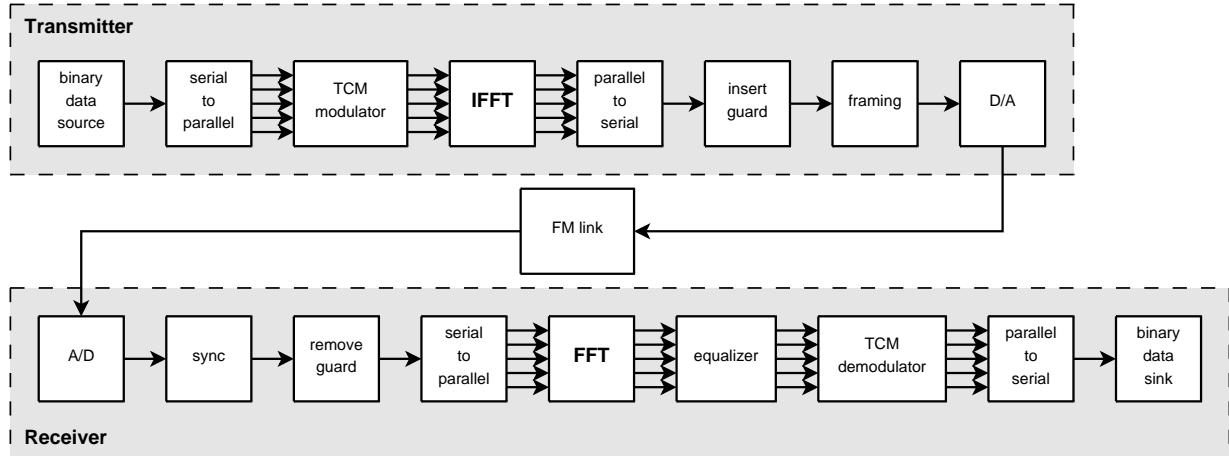


Figure 10.24 Block diagram of the OFDM/TCM modem for an FM link.

a sequence $\{s_k, k = 1, \dots, N\}$, of complex data symbols s_k chosen from any modulation scheme, e.g., *Quadrature Amplitude Modulation* (QAM). Carrying out the IDFT on the data block, we get

$$S_n = \text{IDFT}\{s_k\} = \frac{1}{N} \sum_{k=0}^{N-1} s_k e^{\frac{j2\pi nk}{N}}, \quad n = 0, 1, \dots, N-1. \quad (10.41)$$

In order to be able to transmit that sequence over the FM channel, it has to be made real-valued. This can be easily achieved by modulating the positive half of the subcarriers only and mirroring the other half of the spectrum with the respective complex conjugate.

Contrariwise, the OFDM demodulator samples the received time-domain signal S_n with an *analog-to-digital converter* (ADC). Afterwards, the demodulation is realized with the N -point *discrete Fourier transform* (DFT)

$$s_k = \text{DFT}\{S_n\} = \sum_{n=0}^{N-1} S_n e^{-\frac{j2\pi nk}{N}}, \quad k = 0, 1, \dots, N-1. \quad (10.42)$$

Hardware considerations led to FFT symbols of length $N = 128$.

10.12.2 Symbols and Frames

The cyclicly extended OFDM symbols (the cyclic extension applied is 30 samples) can actually be serially transmitted over the channel. But, in order for the receiver to synchronize to the OFDM symbol start and to equalize the signal distortion caused by the channel response, a sophisticated equalization and synchronization technique has to be adopted. For the OFDM modem developed, a compromise between complexity, reliability and minimum latency had to be found. A frame structure as the one given in Fig. 8.6 has been deployed. The data stream consists of a series of single frames. Every frame starts with a so called *Training symbol*. This symbol consists of an OFDM symbol, whose content is known by the receiver. The purpose of that symbol is to have both a reference for every carrier used in the OFDM transmission in order to equalize

the channel distortion and a reference to synchronize with the start of a frame. The rest of the frame consists of an array of OFDM symbols that contains the data.

10.12.3 Frame Synchronization

The frame synchronization searches the Training symbol by cross-correlating them with the incoming signal. At the point of perfect match, the correlation results in a high peak.

The training symbols are not only used for the symbol and frame synchronization, but also for the channel estimation of the equalization process as will be seen in the next section. The choice of the 16-PSK symbols (no amplitude variation) is further optimized by *peak-to-average-power ratio* considerations in order to minimize distortion.

10.12.4 Channel Estimation and Equalization

The received OFDM signal is distorted by the channel and has to be equalized prior to be demodulated. The transmission channel of the FM modem can be widely regarded as a channel with linear distortions and additive white Gaussian noise. Some modifications are also introduced by the wireless link which can be regarded as multiplicative noise. That fading noise is mainly caused by directional properties of the transmitter and receiver antenna, or by reflections and absorptions of the signal on the ground or walls.

After the frame synchronization, the receiver uses the Training symbol to perform the channel equalization. In OFDM, the whole frequency bandwidth is divided into N subcarriers. If N is high enough, then the transfer function $H(f)$ does not change significantly over the subchannel and it can be regarded as flat over a single narrow subcarrier. The approximation of the transfer function for a single subcarrier reduces to a single complex coefficient $H_n = H(f_n)$. Therefore, the equalization of a single subcarrier can be simply performed by multiplying the received sample with the inverse of H_n [39]. The equalization coefficients w_n can be computed by the element-wise division of the original FFT coefficients z_n of the training symbol by those of the received training symbol \hat{z}_n . All further OFDM symbols in the same frame can now be equalized by multiplying the received OFDM symbol in the frequency domain with the equalization coefficients w_n . Those complex coefficients also include the linear phase caused by the timing error.

During the demodulation of the OFDM symbols in a frame there is another effect that must be considered. In any communication system, the sampling clock at the receiver is often different from that at the transmitter. The sampling clock difference introduces interchannel interference because the subcarriers lose their perfect orthogonality. If the difference remains small, the deviation only consists of a subchannel-dependent phase rotation over time.

10.12.5 Waterfilling

With the aid of a water-filling algorithm, the best bit and energy distribution over the subcarriers can be found. Generally, such an algorithm produces bit distributions where the single bit loads can be any real number and the realization of these non-integer bit values can be difficult. In this work, the algorithm proposed by Levin and Campello [21] has been utilized because it allows the computation of bit distributions with integer bit loads.

In the current example, five different constellations from QPSK up to 256-QAM have been used depending on the partial SNR of the subcarrier, see Fig. 10.25. Such high N -QAM seem unusual for wireless links, however, in this example the amplitude-dependent modulation is 'protected' by the FM link.

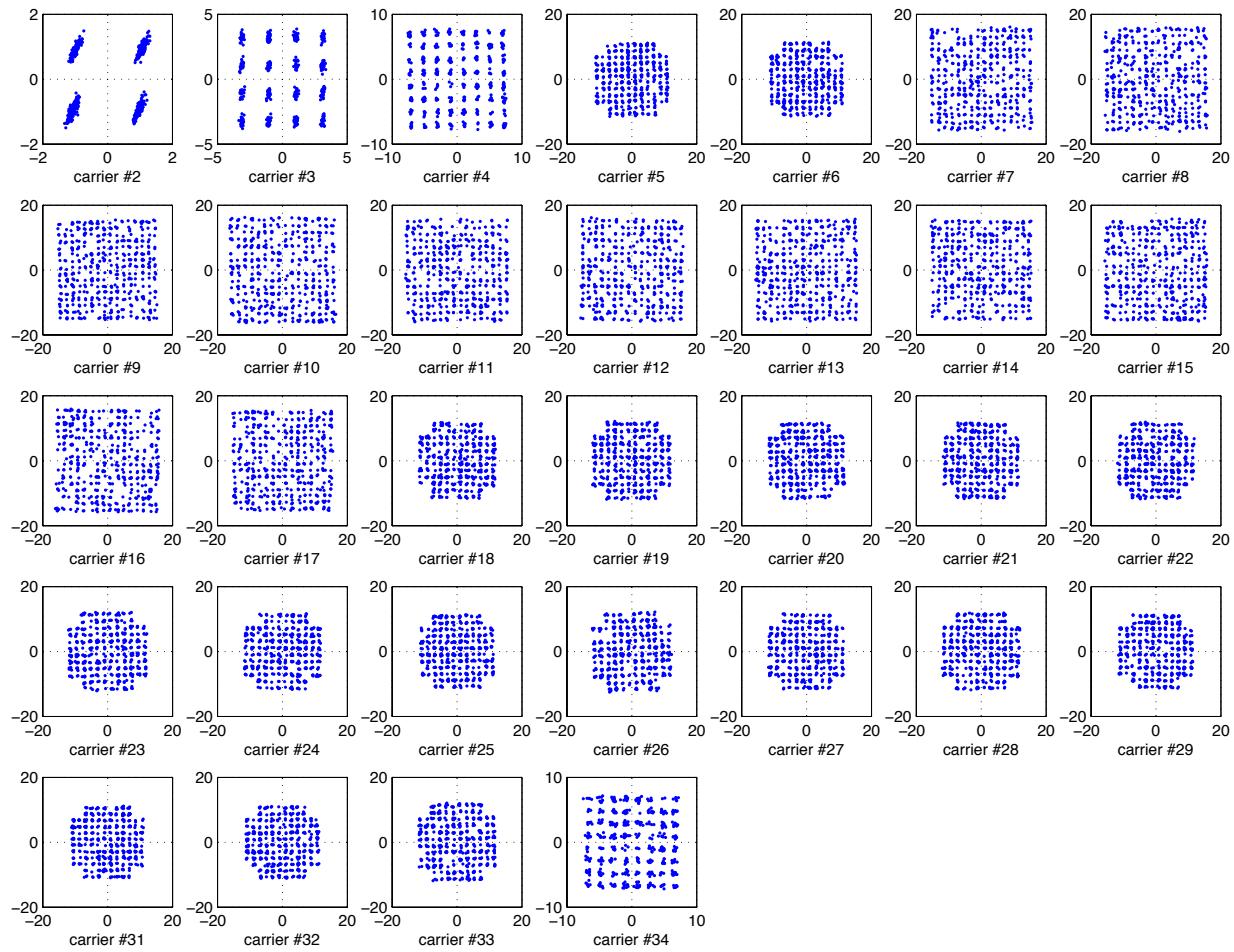


Figure 10.25 Different modulation schemes are used for the individual subcarriers.

10.12.6 PAPR Considerations

In the current approach, as is often the case in OFDM, the PAPR is not tried to reduce, but rather a compromise of noise-limited and distortion-limited operation is sought after. Generally, in OFDM, very high signal peaks occur rarely. The actual PAPR does not exceed 12 dB in 0.9999% of the cases, where the maximum PAPR would be approximately 18 dB. In Fig. 10.26 the amplitude density of an OFDM signal with over 30 subcarriers is plotted. The actual PAPR does not exceed 12 dB in 0.9999% of the cases, where the maximum PAPR would be approximately 18 dB.

With respect to the situation described, the optimum signal power has been evaluated for the OFDM signal. The results are displayed in Fig. 10.27. It can be seen that there is an optimum at about $4 \text{ mV}_{\text{rms}}$. When increasing the amplitude, the signal is clipped more and more and the bit error rate increases as well. Below the optimum amplitude the noise becomes dominant.

10.12.7 Trellis Coded Modulation over OFDM

As mentioned above, for a four-state code a minimum delay of $5 \cdot 3 = 15$ TCM symbols has to be considered. If TCM were applied onto the single OFDM subcarriers sequentially in the time domain, the decoding delay

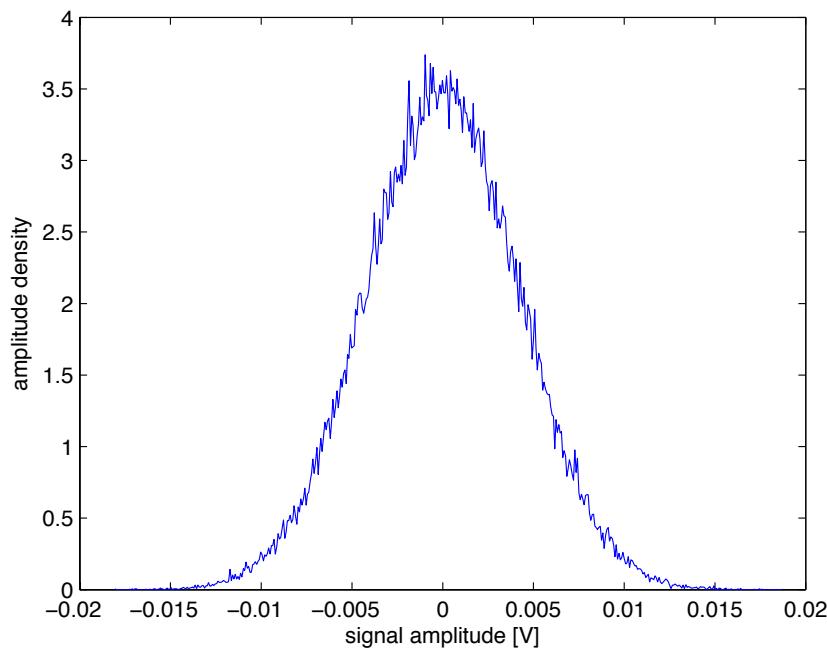


Figure 10.26 OFDM amplitude density $V_{\text{rms}}=4\text{mV}$.

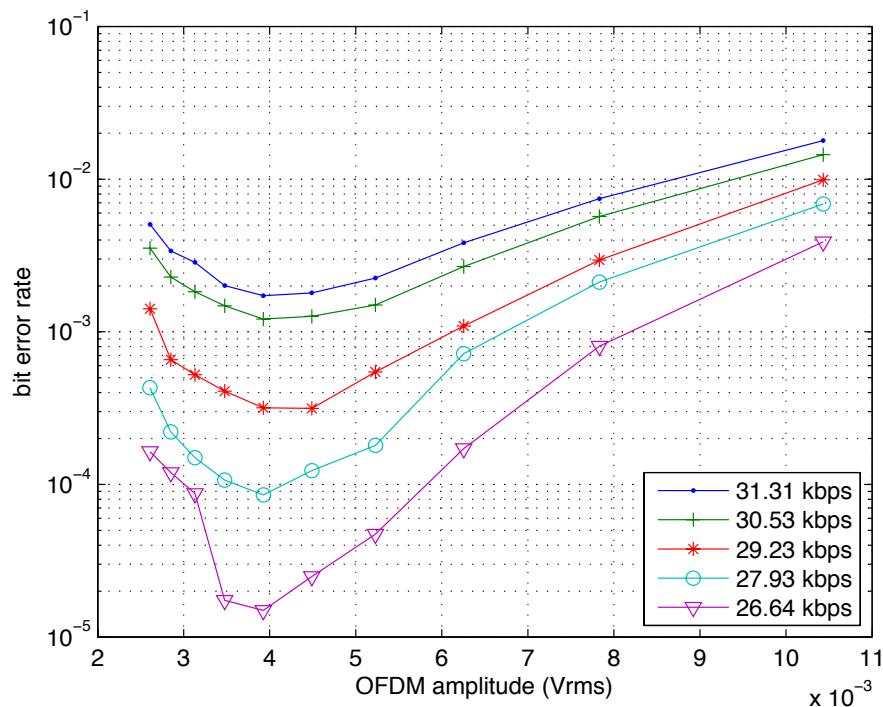


Figure 10.27 Impact of the signal clipping on the BER.

would increase the latency by the same amount and the implementation would result infeasible for a large number of use cases.

In order to avoid the introduction of such a large additional latency in the communication channel through the use of TCM, the data can be encoded/modulated with TCM on the OFDM subcarriers in the frequency domain. This means that the decoding process uses only the data of one OFDM symbol at a time. That way, every OFDM symbol can be decoded independently and no additional latency is introduced, but some restrictions occur. On the one hand, the subcarriers used by TCM must transmit the same constellation size and on the other hand, redundancy must be inserted in order to decode all data correctly. More precisely: the trace-back path must always start and end in a known state (e.g., all-zero state). This means that the TCM encoder must insert enough zeros at the end of the binary data in every OFDM symbol. When using the four-state code, the last two TCM symbols must contain redundancy only.

Using TCM (256-QAM) over 30 of a total of 36 subcarriers (the remaining 6 subcarriers use a lower constellation), the data rate (at a BER of 10^{-4}) improved from 28 kbit/s to 32 kbit/s.

10.12.8 Conclusions

A stacked modulation scheme using FM as the lowest layer, OFDM in between, and TCM over the subcarriers has been proposed for a narrowband wireless channel. The equalization and synchronization processes are computationally economic and allow the deployment of microcontrollers accompanied by an FFT accelerator. The application of the TCM over frequency bins rather than over time keeps latency times at a minimum and further increases the data rate for the given channel conditions.

11 Turbo Coding

11.1 History and Motivation

People have looked for methods to come close to the Shannon capacity shown in Fig. 10.23 for more than half a century. Combining modulation and coding as in TCM shown in the last chapter is just one possibility. Yet another approach are Turbo Codes, which have their roots in the seminal paper by Berrou *et al.* [3]. Today, Turbo Codes are still a research topic with a lot of open issues remaining. It is, however, no longer just a theoretical concept, but has many interesting applications, e.g., in 3G mobile communications, deep-space satellite communications, and power-line communications (PLC), just to name a few. The name *Turbo* does not refer to a marketing preword (as is often the case in IT and for programming languages to indicate speed), but rather to its fundamental concept, which resembles a turbocharger in the way that exhaust gas is reused.

11.2 Approach

Two classical approaches exist: serial concatenation schemes and parallel concatenated convolutional codes. Serial concatenation schemes may consist of traditional inner and outer codes as long as both decoders use soft-decision output information, so that we can iterate between inner and outer code. Parallel concatenated convolutional codes, on the other hand, may be regarded as a two-dimensional table of bits which are first encoded row by row and in a second go column by column. On the receiver side, the decoder can now switch between processing the rows, use this information to process the columns, use the resulting information to go back and reprocess the rows, and so on. We start with possibly poor information on the bits, but improve with every iteration.

11.3 Example

In the following we build a Turbo Code example using a so-called RSC (recursive systematic convolutional) code with the generator expression

$$G(D) = \left[1 \quad \frac{1+D+D^2}{1+D^2} \right]. \quad (11.1)$$

The encoder for this system is shown in Fig. 11.1. As can be seen, the same RSC code is used twice, once in ordinary bit indexing order resulting in stream b_1 and once after an interleaver resulting in stream b_2 (corresponding to the idea of rowwise and columnwise), and thus corresponding to the idea of parallel concatenation. It is a systematic code, hence we also have the original uncoded information bit stream b_0 . The Viterbi Algorithm might be used to decode the received RSC streams. However, the Viterbi Algorithm produces so-called hard decisions, losing all reliability information on corrected bits. There are extended versions available, e.g., the Soft-Output Viterbi Algorithm (SOVA) carries over likelihood information, see for example [14], [25]. An alternative is the so-called BCJR algorithm [1], also called the MAP algorithm or forward-backward algorithm. Thus, at the receiver we decode the two encoded bit streams b_1 and b_2 together with the uncoded bit stream b_0 using BCJR decoder 1 and BCJR decoder 2, respectively, see also Fig. 11.2. After each decoding step, the decoder produces some intrinsic information $p(m_x)$, which is used in the next decoding step of the other decoder.

A user contributed Matlab file (Matlab Central File Exchange) models a transmitter/receiver chain for a BPSK signal under AWGN conditions. The result can be seen in Fig. 11.3.

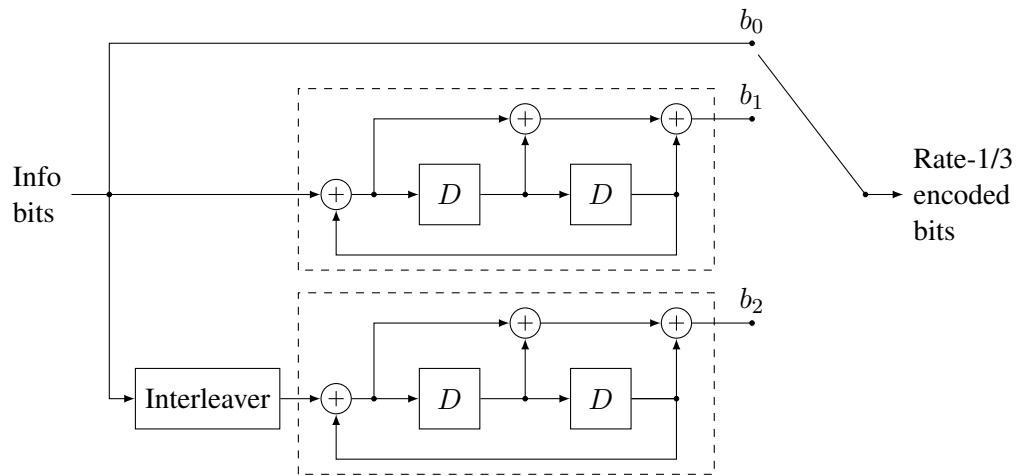


Figure 11.1 Turbo encoder. The dashed boxes produce the recursive convolutional codes.

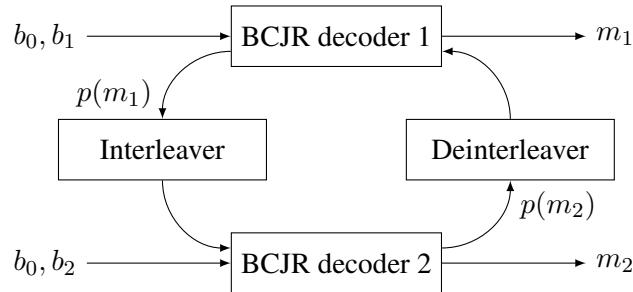


Figure 11.2 Turbo decoder.

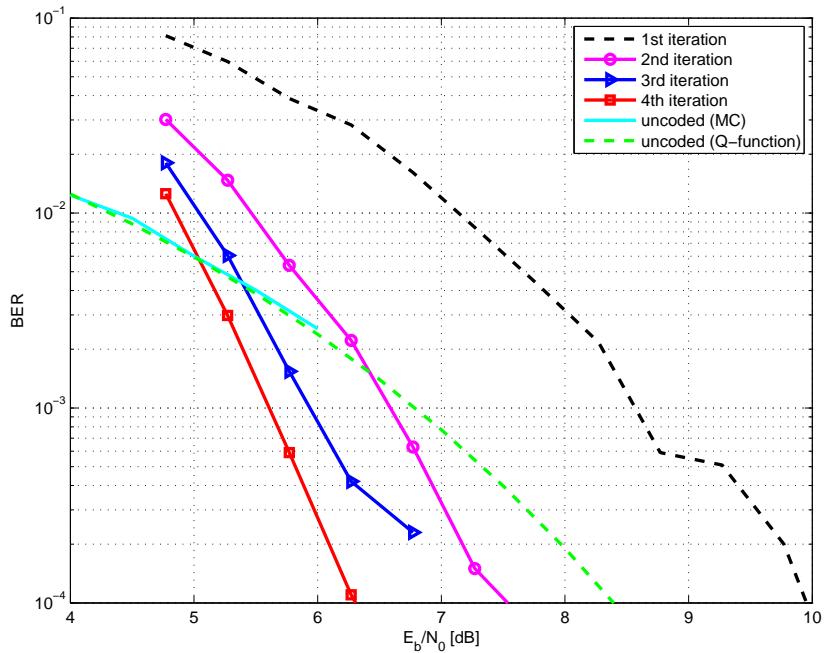


Figure 11.3 Performance of the Turbo code after several iterations.

Appendix

A Equations and Formulae

A.1 Linear Algebra

General rules for matrix multiplication:

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}, \quad (\text{A.1})$$

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$$

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) \quad (\text{A.2})$$

$$(\mathbf{AB})^T = (\mathbf{B})^T(\mathbf{A})^T \quad (\text{A.3})$$

$$(\mathbf{AB})^{-1} = (\mathbf{B})^{-1}(\mathbf{A})^{-1} \quad (\text{A.4})$$

symmetric matrix

$$\mathbf{A} = \mathbf{A}^T \quad (\text{A.5})$$

skew-symmetric matrix

$$\mathbf{A} = -\mathbf{A}^T = \begin{pmatrix} 0 & a_{12} & a_{13} \\ -a_{12} & 0 & a_{23} \\ -a_{13} & -a_{23} & 0 \end{pmatrix} \quad (\text{A.6})$$

Hermitian matrix

$$\mathbf{A} = \mathbf{A}^H = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12}^* & a_{22} & a_{23} \\ a_{13}^* & a_{23}^* & a_{33} \end{pmatrix} \quad (\text{A.7})$$

skew-Hermitian matrix

$$\mathbf{A} = -\mathbf{A}^H = \begin{pmatrix} -0 & a_{12} & a_{13} \\ -a_{12}^* & 0 & a_{23} \\ -a_{13}^* & -a_{23}^* & 0 \end{pmatrix} \quad (\text{A.8})$$

unitary matrix¹

$$\mathbf{AA}^H = \mathbf{A}^H\mathbf{A} = \mathbf{I} \quad (\text{A.9})$$

normal matrix

$$\mathbf{AA}^H = \mathbf{A}^H\mathbf{A} \quad (\text{A.10})$$

inverse matrix (2 by 2)

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \quad (\text{A.11})$$

Matrix inversion lemma

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{D}\mathbf{A}^{-1} \quad (\text{A.12})$$

Matrix inversion lemma for block-partitioned matrices

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C}^T & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{E}\Delta^{-1}\mathbf{F} & -\mathbf{E}\Delta^{-1} \\ -\Delta^{-1}\mathbf{F} & \Delta^{-1} \end{bmatrix} \quad (\text{A.13})$$

¹A unitary matrix is used for complex-valued data, whereas orthogonal matrix is used for real-valued data, i.e. $\mathbf{AA}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}$

where

$$\Delta = D - C^T A^{-1} B \quad (\text{A.14})$$

$$E = A^{-1} B \quad (\text{A.15})$$

$$F = C^T A^{-1} \quad (\text{A.16})$$

Derivatives of vectors

$$\nabla_{\mathbf{v}} \{\mathbf{u}^T \mathbf{v}\} = \nabla_{\mathbf{v}} \{\mathbf{v}^T \mathbf{u}\} = \mathbf{u} \quad (\text{A.17})$$

$$\nabla_{\mathbf{u}} \{\mathbf{u}^T \mathbf{u}\} = 2\mathbf{u} \quad (\text{A.18})$$

$$\nabla_{\mathbf{v}} \{\mathbf{v}^T \mathbf{A}\} = \mathbf{A} \quad (\text{A.19})$$

$$\nabla_{\mathbf{v}} \{\mathbf{v}^T \mathbf{A} \mathbf{v}\} = \mathbf{A} \mathbf{v} + \mathbf{A}^T \mathbf{v} \quad (\text{A.20})$$

For symmetric matrices \mathbf{R}

$$\nabla_{\mathbf{v}} \{\mathbf{v}^T \mathbf{R} \mathbf{v}\} = 2\mathbf{R} \mathbf{v} \quad (\text{A.21})$$

Derivatives of matrix functions (scalars with respect to a matrix), see also [13]

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}) = \mathbf{I} \quad (\text{A.22})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A} \mathbf{X}) = \mathbf{A}^T \quad (\text{A.23})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A} \mathbf{X}^T) = \mathbf{A} \quad (\text{A.24})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A} \mathbf{X} \mathbf{B}) = \mathbf{A}^T \mathbf{B}^T \quad (\text{A.25})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A} \mathbf{X}^T \mathbf{B}) = \mathbf{B} \mathbf{A} \quad (\text{A.26})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X} \mathbf{X}) = 2\mathbf{X}^T \quad (\text{A.27})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X}) = (\mathbf{A} + \mathbf{A}^T) \mathbf{X} \quad (\text{A.28})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X} \mathbf{A} \mathbf{X}^T) = \mathbf{X} (\mathbf{A} + \mathbf{A}^T) \quad (\text{A.29})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A} \mathbf{X} \mathbf{B} \mathbf{X}) = \mathbf{A}^T \mathbf{X}^T \mathbf{B}^T + \mathbf{B}^T \mathbf{X}^T \mathbf{A}^T \quad (\text{A.30})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A} \mathbf{X} \mathbf{B} \mathbf{X}^T) = \mathbf{A} \mathbf{X} \mathbf{B} + \mathbf{A}^T \mathbf{X} \mathbf{B}^T \quad (\text{A.31})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A} \mathbf{X} \mathbf{X}^T \mathbf{B}) = (\mathbf{A}^T \mathbf{B}^T + \mathbf{B} \mathbf{A}) \mathbf{X} \quad (\text{A.32})$$

$$\frac{\partial}{\partial \mathbf{X}} \det(\mathbf{X}) = \det(\mathbf{X}) \mathbf{X}^{-T} \quad (\text{A.33})$$

Derivatives with respect to a complex matrix

$$\left. \begin{array}{l} \frac{\partial}{\partial \mathbf{X}} \\ \frac{\partial}{\partial \mathbf{X}^*} \end{array} \right\} = \frac{1}{2} \begin{pmatrix} \frac{\partial}{\partial x_{11}^{\text{Re}}} \mp j \frac{\partial}{\partial x_{11}^{\text{Im}}} & \frac{\partial}{\partial x_{12}^{\text{Re}}} \mp j \frac{\partial}{\partial x_{12}^{\text{Im}}} \\ \frac{\partial}{\partial x_{21}^{\text{Re}}} \mp j \frac{\partial}{\partial x_{21}^{\text{Im}}} & \frac{\partial}{\partial x_{22}^{\text{Re}}} \mp j \frac{\partial}{\partial x_{22}^{\text{Im}}} \end{pmatrix} \quad (\text{A.34})$$

After Haykin [18], the complex gradient can be defined as

$$\nabla_{\mathbf{X}} = 2 \frac{\partial}{\partial \mathbf{X}^*} \quad (\text{A.35})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A}\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}\mathbf{A}) = \mathbf{A}^T \quad (\text{A.36})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A}\mathbf{X}^H) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}^H\mathbf{A}) = \mathbf{0} \quad (\text{A.37})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}\mathbf{X}^H) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}^H\mathbf{X}) = \mathbf{X}^* \quad (\text{A.38})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}\mathbf{X}^*) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}^*\mathbf{X}) = \mathbf{X}^H \quad (\text{A.39})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}^2) = 2\mathbf{X}^T \quad (\text{A.40})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A}\mathbf{X}^H\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}\mathbf{A}\mathbf{X}^H) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}^H\mathbf{X}\mathbf{A}) = \mathbf{X}^* \mathbf{A}^T \quad (\text{A.41})$$

$$\frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{A}\mathbf{X}\mathbf{X}^H) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}^H\mathbf{A}\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}} \text{tr}(\mathbf{X}\mathbf{X}^H\mathbf{A}) = \mathbf{A}^T \mathbf{X}^* \quad (\text{A.42})$$

$$\frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{A}\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}\mathbf{A}) = \mathbf{0} \quad (\text{A.43})$$

$$\frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{A}\mathbf{X}^H) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}^H\mathbf{A}) = \mathbf{A} \quad (\text{A.44})$$

$$\frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}\mathbf{X}^H) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}^H\mathbf{X}) = \mathbf{X} \quad (\text{A.45})$$

$$\frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}\mathbf{X}^*) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}^*\mathbf{X}) = \mathbf{X}^T \quad (\text{A.46})$$

$$\frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}^2) = \mathbf{0} \quad (\text{A.47})$$

$$\frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{A}\mathbf{X}^H\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}\mathbf{A}\mathbf{X}^H) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}^H\mathbf{X}\mathbf{A}) = \mathbf{X}\mathbf{A} \quad (\text{A.48})$$

$$\frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{A}\mathbf{X}\mathbf{X}^H) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}^H\mathbf{A}\mathbf{X}) = \frac{\partial}{\partial \mathbf{X}^*} \text{tr}(\mathbf{X}\mathbf{X}^H\mathbf{A}) = \mathbf{A}\mathbf{X} \quad (\text{A.49})$$

A.2 Analysis

A.2.1 Complex Analysis

If $z = I + jQ$ then

$$z + z^* = 2 \operatorname{Re}(z) \quad (\text{A.50})$$

$$z - z^* = 2 \operatorname{Im}(z) \quad (\text{A.51})$$

$$z^2 = I^2 - Q^2 + 2jIQ \quad (\text{A.52})$$

$$\operatorname{Re}(z^2) = I^2 - Q^2 \quad (\text{A.53})$$

$$|z|^2 = z \cdot z^* = I^2 + Q^2 \quad (\text{A.54})$$

$$|z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + 2 \operatorname{Re}(z_1^* z_2) \quad (\text{A.55})$$

$$|z|^4 = |z^4| = (z \cdot z^*)^2 = (I^2 + Q^2)^2 \quad (\text{A.56})$$

$$\left(\frac{a+jb}{c+jd} \right)^* = \frac{a-jb}{c-jd} \quad (\text{A.57})$$

$$\left| \frac{a+jb}{c+jd} \right| = \sqrt{\frac{a^2+b^2}{c^2+d^2}} \quad (\text{A.58})$$

$$\operatorname{Re} \left(\frac{a+jb}{c+jd} \right) = \frac{ac+bd}{c^2+d^2} \quad (\text{A.59})$$

$$\operatorname{Im} \left(\frac{a+jb}{c+jd} \right) = \frac{bc-ad}{c^2+d^2} \quad (\text{A.60})$$

A.3 Statistics

A.3.1 Probability Theory

Variance

$$\operatorname{Var}[X] = E[(X - m_X)^2] = E[X^2] - (E[X])^2 \quad (\text{A.61})$$

Covariance

$$\operatorname{Cov}(X, Y) = E[(X - m_X)(Y - m_Y)] \quad (\text{A.62})$$

Independent events

$$P(A, B) = P(A) \cdot P(B) \quad (\text{A.63})$$

Conditional probability

$$P(A|B) = \frac{P(A, B)}{P(B)} \quad (\text{A.64})$$

Bayes' theorem

$$P(A|B) = \frac{P(A|B) \cdot P(A)}{P(B)} \quad (\text{A.65})$$

Chain rule for probabilities

$$P(ABC) = P(A|BC) \cdot P(B|C) \cdot P(C) \quad (\text{A.66})$$

A.4 Digital Signals and Systems

A.4.1 Convolution

Continuous-time convolution

$$y(t) = x(t) * y(t) = \int_{\tau=-\infty}^{\infty} x(\tau)g(t-\tau)d\tau \quad (\text{A.67})$$

Discrete-time convolution

$$y[k] = x[k] * y[k] = \sum_{m=-\infty}^{\infty} x[m]g[k-m] \quad (\text{A.68})$$

A.4.2 Wiener Filtering

Auto-correlation function

$$r_{XY}[k] = E[X[i+k]Y[i]] \quad (\text{A.69})$$

$$r_{YX}[k] = E[X[i]Y[i+k]] \quad (\text{A.70})$$

Wide-sense stationarity (WSS)

$$\begin{aligned} m[k] &= m \\ r[k, k-i] &= r[i] \\ c[k, k-i] &= c[i] \end{aligned} \quad (\text{A.71})$$

Wiener-Kinchine relation

$$S_{YX}(z) = H(z) \cdot S_X(z) \quad (\text{A.72})$$

$$S_{XY}(z) = H(z^{-1}) \cdot S_X(z) \quad (\text{A.73})$$

$$S_Y(z) = H(z)H(z^{-1}) \cdot S_X(z) \quad (\text{A.74})$$

FIR Wiener filter

$$\mathbf{R}_Y \cdot \mathbf{h} = \mathbf{R}_{YX} \quad (\text{A.75})$$

$$\begin{pmatrix} r_Y[0] & r_Y[1] & \cdots & r_Y[M] \\ r_Y[1] & r_Y[0] & \cdots & r_Y[M-1] \\ \vdots & \vdots & \ddots & \vdots \\ r_Y[M] & r_Y[M-1] & \cdots & r_Y[0] \end{pmatrix} \cdot \begin{pmatrix} h[0] \\ h[1] \\ \vdots \\ h[M] \end{pmatrix} = \begin{pmatrix} r_{YX}[-d] \\ r_{YX}[1-d] \\ \vdots \\ r_{YX}[M-d] \end{pmatrix} \quad (\text{A.76})$$

The MSE of the Wiener filter is

$$\text{MSE} = r_X[0] - \sum_{i=0}^M h[i] \cdot r_{YX}[i-d] \quad (\text{A.77})$$

Non-causal Wiener filter

$$H(z) = \frac{z^{-d} S_{XY}(z)}{S_Y(z)} \quad (\text{A.78})$$

Causal Wiener filter

If $G(z)$ is a whitening filter of the signal Y , so that

$$S_Y(z) \cdot G(z) \cdot G(z^{-1}) = 1 \quad (\text{A.79})$$

then the causal Wiener filter is

$$H(z) = \left| z^{-d} \cdot G(z^{-1}) \cdot S_{XY}(z) \right|^+ \cdot G(z) \quad (\text{A.80})$$

B Frequency Usage

B.1 General Frequency Bands

Band	Wavelength	Frequency range	Propagation	Examples
ELF	10^5 km - 10^4 km	<30 Hz		
SLF	10^4 km - 10^3 km	30 - 300 Hz		
ULF	1000 km - 100 km	300 Hz - 3 kHz		
VLF	100 km - 10 km	3 - 30 kHz	ground and room wave	heart-beat monitor
LF	10 km - 1 km	30 - 300 kHz	ground wave > 1000 km room wave at night	Long wave radio (150 - 300 kHz)
MF	1 km - 100 m	300 kHz - 3 MHz	ground wave > 100 km room wave at night	Medium wave radio (500 - 1500 kHz)
HF	100 m - 10 m	3 MHz - 30 MHz	sky wave reflected by ionosphere	Short wave radio (5 - 30 MHz)
VHF	10 m - 1 m	30 MHz - 300 MHz	free-space openfield	FM radio (30 - 300 MHz)
UHF	1 m - 10 cm	300 MHz - 3 GHz	free-space openfield	Consumer Communications, GPS
SHF	10 cm - 1 cm	3 GHz - 30 GHz	free-space openfield	Radar, Satellite
EHF	1 cm - 1 mm	30 GHz - 300 GHz	free-space	Radar, Satellite

Table B.1 Frequency bands.

SHF is also called microwaves and EHF is called millimeter waves.

Frequency	Wavelength	Band (old)
1 – 2 GHz	30 – 15 cm	L
2 – 4 GHz	15 – 7.5 cm	S
4 – 8 GHz	7.5 – 3.75 cm	C
8 – 12 GHz	3.75 – 2.5 cm	X
12 – 18 GHz	2.5 – 1.67 cm	Ku
18 – 27 GHz	1.67 – 1.11 cm	K
27 – 40 GHz	11.1 – 7.5 mm	Ka
40 – 75 GHz	7.5 – 4 mm	V
75 – 110 GHz	4 – 2.72 mm	W
110 – 300 GHz	2.72 – 1 mm	mm

Table B.2 IEEE frequency band designations.

B.2 RADAR Bands

Frequency	Band (old)
0.22 - 0.3 GHz	P
33 - 50 GHz	Q
40 - 60 GHz	U
60 - 90 GHz	E
90 - 140 GHz	F
110 - 170 GHz	D
140 - 220 GHz	G
170 - 260 GHz	Y
220 - 325 GHz	J

Table B.3 IEEE RADAR band designations.

There is yet another set of letters in use, which matches the alphabetical order to the increasing frequencies in succeeding bands. Although this set is newer and hence more modern, it has not found wide acceptance and is therefore left out here.

C Important Tables

C.1 The Q-Function

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0+	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1+	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2+	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3+	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4+	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5+	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6+	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7+	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8+	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9+	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0+	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1+	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2+	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3+	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4+	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5+	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6+	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7+	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8+	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9+	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0+	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1+	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2+	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3+	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4+	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5+	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6+	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7+	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8+	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9+	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0+	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1+	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2+	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3+	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.4+	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
3.5+	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
3.6+	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

Table C.1 Numerical evaluations of some values of $Q(x)$.

Table C.1 has been produced by Matlab. In Matlab, the Q -function as given by

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \quad (\text{C.1})$$

is not directly accessible. It can, however, be easily computed using the `erfc` command, since

$$Q(x) = \frac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right). \quad (\text{C.2})$$

C.2 M-Sequences

Stages	Possible feedback taps	Stages	Possible feedback taps
2	[2, 1]		[9, 8, 7, 6, 3, 1]
3	[3, 2]		[9, 8, 7, 6, 2, 1]
4	[4, 3]		[9, 8, 7, 5, 4, 3]
5	[5, 3]		[9, 8, 7, 5, 4, 2]
	[5, 4, 3, 2]		[9, 8, 6, 5, 4, 1]
	[5, 4, 3, 1]		[9, 8, 6, 5, 3, 2]
6	[6, 5]		[9, 8, 6, 5, 3, 1]
	[6, 5, 4, 1]		[9, 7, 6, 5, 4, 3]
	[6, 5, 3, 2]		[9, 7, 6, 5, 4, 2]
7	[7, 6]		[9, 8, 7, 6, 5, 4, 3, 1]
	[7, 4]	10	[10, 7]
	[7, 6, 5, 4]		[10, 9, 8, 5]
	[7, 6, 5, 2]		[10, 9, 7, 6]
	[7, 6, 4, 2]		[10, 9, 7, 3]
	[7, 6, 4, 1]		[10, 9, 6, 1]
	[7, 5, 4, 3]		[10, 9, 5, 2]
	[7, 6, 5, 4, 3, 2]		[10, 9, 4, 2]
	[7, 6, 5, 4, 2, 1]		[10, 8, 7, 5]
8	[8, 7, 6, 1]		[10, 8, 7, 2]
	[8, 7, 5, 3]		[10, 8, 5, 4]
	[8, 7, 3, 2]		[10, 8, 4, 3]
	[8, 6, 5, 4]		[10, 9, 8, 7, 5, 4]
	[8, 6, 5, 3]		[10, 9, 8, 7, 4, 1]
	[8, 6, 5, 2]		[10, 9, 8, 7, 3, 2]
	[8, 7, 6, 5, 4, 2]		[10, 9, 8, 6, 5, 1]
	[8, 7, 6, 5, 2, 1]		[10, 9, 8, 6, 4, 3]
9	[9, 5]		[10, 9, 8, 6, 4, 2]
	[9, 8, 7, 2]		[10, 9, 8, 6, 3, 2]
	[9, 8, 6, 5]		[10, 9, 8, 6, 2, 1]
	[9, 8, 5, 4]		[10, 9, 8, 5, 4, 3]
	[9, 8, 5, 1]		[10, 9, 8, 4, 3, 2]
	[9, 8, 4, 2]		[10, 9, 7, 6, 4, 1]
	[9, 7, 6, 4]		[10, 9, 7, 5, 4, 2]
	[9, 7, 5, 2]		[10, 9, 6, 5, 4, 3]
	[9, 6, 5, 3]		[10, 8, 7, 6, 5, 2]
	[9, 8, 7, 6, 5, 3]		[10, 9, 8, 7, 6, 5, 4, 3]
	[9, 8, 7, 6, 5, 1]		[10, 9, 8, 7, 6, 5, 4, 1]
	[9, 8, 7, 6, 4, 3]		[10, 9, 8, 7, 6, 4, 3, 1]
	[9, 8, 7, 6, 4, 2]		[10, 9, 8, 6, 5, 4, 3, 2]
	[9, 8, 7, 6, 3, 2]		[10, 9, 7, 6, 5, 4, 3, 2]

Table C.2 Feedback taps that result in m-sequences for up to ten stages.

Source: <http://www.newwaveinstruments.com>

D Proof on the Capacity Maximization

In the following we want to determine the pdf that maximizes capacity in a continuous channel. To do so we write down the constraints for the input distribution $p_X(x)$ to be a probability density function of zero mean and variance σ^2 :

$$\int_{-\infty}^{\infty} p_X(x)dx = 1, \quad (\text{D.1})$$

$$\int_{-\infty}^{\infty} xp_X(x)dx = 0, \quad (\text{D.2})$$

$$\int_{-\infty}^{\infty} x^2 p_X(x)dx = \sigma^2. \quad (\text{D.3})$$

We now want to maximize

$$H(X) = \int_{-\infty}^{\infty} p_X(x) \ln p_X(x)dx. \quad (\text{D.4})$$

The constraints may be expressed using Lagrange multipliers

$$L(X, p_X(x)) = -p_X(x) \ln p_X(x) + \lambda_1(p_X(x) - 1) + \lambda_2(xp_X(x) - 0) + \lambda_3(x^2 p_X(x) - \sigma^2). \quad (\text{D.5})$$

The partial derivative

$$\frac{\partial L(X, p_X(x))}{\partial p_X(x)} = -\ln p_X(x) - 1 + \lambda_1 + \lambda_2 x + \lambda_3 x^2 \quad (\text{D.6})$$

set equal to zero leads to

$$\ln p_X(x) = (\lambda_1 - 1) + \lambda_2 x + \lambda_3 x^2 \quad (\text{D.7})$$

or

$$p_X(x) = e^{(\lambda_1 - 1) + \lambda_2 x + \lambda_3 x^2}, \quad (\text{D.8})$$

with the constraints

$$\int_{-\infty}^{\infty} e^{(\lambda_1 - 1) + \lambda_2 x + \lambda_3 x^2} dx = 1, \quad (\text{D.9})$$

$$\int_{-\infty}^{\infty} x e^{(\lambda_1 - 1) + \lambda_2 x + \lambda_3 x^2} dx = 0, \quad (\text{D.10})$$

$$\int_{-\infty}^{\infty} x^2 e^{(\lambda_1 - 1) + \lambda_2 x + \lambda_3 x^2} dx = \sigma^2. \quad (\text{D.11})$$

The second constraint, zero mean, is achieved for $\lambda_2 = 0$. The other two constraints can be reformulated into

$$e^{(\lambda_1 - 1)} \int_{-\infty}^{\infty} e^{\lambda_3 x^2} dx = 1, \quad (\text{D.12})$$

$$e^{(\lambda_1 - 1)} \int_{-\infty}^{\infty} x^2 e^{\lambda_3 x^2} dx = \sigma^2. \quad (\text{D.13})$$

By comparing the coefficients of a Gaussian variable with $e^{(\lambda_1 - 1)}$ and e^{λ_3} , respectively, it can be seen that

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad (\text{D.14})$$

indeed.

References

- [1] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Transactions on Information Theory*, vol. 20, no. 2, March 1974, pp. 284–287.
- [2] P. Beckmann and A. Spizzichino, *The scattering of Electromagnetic Waves from Rough Surfaces*. Macmillan, New York, 1963.
- [3] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near Shannon limit error-correcting coding and decoding: Turbo codes,” in *Proc. IEEE Int. Conf. Commun. ICC ’93*, 1993, pp. 1064–1070.
- [4] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation With Applications*. Macmillan Pub Co, 1991.
- [5] R. W. Chang, “Orthogonal frequency division multiplexing,” *U.S. Patent 3'488'445*, 1970.
- [6] G. Delisle, J.-P. Lefèvre, M. Lecours, and J.-Y. Chouinard, “Propagation loss prediction: a comparative study with application to the mobile radio channel,” *IEEE Trans. on Vehicular Technology*, vol. 34, no. 2, May 1985, pp. 86–96.
- [7] R. C. Dixon, *Spread Spectrum Systems*. John Wiley & Sons, 1976.
- [8] J. J. Egli, “Radio propagation above 40 mc over irregular terrain,” *Proceedings of the IRE*, vol. 45, no. 10, Oct. 1957, pp. 1383–1391.
- [9] M. Engels, *Wireless OFDM Systems: How to make them work? (The Springer International Series in Engineering and Computer Science)*. Springer, 2002.
- [10] G. D. Forney, “The Viterbi algorithm,” *Proc. IEEE*, vol. 61, March 1973, pp. 268–278.
- [11] R. G. Gallager, ed., *Information theory and reliable communication*. John Wiley & Sons, 1968.
- [12] J. Gibson, ed., *The Mobile Communications Handbook*. CRC Press, 1996.
- [13] D. Graupe, *Identification of Systems*. Robert E. Krieger Publishing Company, 1975.
- [14] J. Hagenauer and P. Hoeher, “A Viterbi Algorithm with Soft-Decision Outputs and its Applications,” in *Conf. proceedings IEEE Globecom ’89*, IEEE, 1989, pp. 1680–1686.
- [15] S. H. Han and J. H. Lee, “An overview of peak-to-average power ratio reduction techniques for multi-carrier transmission,” *IEEE Wireless Comm.*, vol. 12, no. 2, Apr. 2005, pp. 56–65.
- [16] L. Hanzo and T. Keller, *OFDM and MC-CDMA: A Primer*. Wiley-IEEE Press, 2006.
- [17] M. Hata, “Empirical formula for propagation loss in land mobile radio services,” *IEEE Trans. on Vehicular Technology*, vol. 29, no. 3, Aug. 1980, pp. 317–325.
- [18] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, 3rd ed., 1996.
- [19] E. Hogenauer, “An economical class of digital filters for decimation and interpolation,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 2, April 1981, pp. 155–162.
- [20] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Transactions of the ASME-Journal of Basic Engineering*, vol. 82, no. Series D, no. Series D, 1960, pp. 35–45.
- [21] H. E. Levin, “A complete and optimal data allocation method for practical discrete multitone systems,” in *Global Telecomm. Conf. (GLOBECOM)*, vol. 1, 2001, pp. 369–374.

- [22] Y. G. Li and G. L. Stuber, *Orthogonal Frequency Division Multiplexing for Wireless Communications (Signals and Communication Technology)*. Springer, 2006.
- [23] W. Mansfeld, ed., *Satellitenortung und Navigation*. Vieweg & Sohn Verlagsgesellschaft mbH, 1998.
- [24] J. L. Massey, “Coding and modulation in digital communications,” in *Proc. Int. Zurich Seminar on Digital Communications*, 1974, pp. E2(1–4).
- [25] H. Mathis, “Differential detection of GMSK signals with low $B_t T$ using the SOVA,” *IEEE Transactions on Communications*, vol. 46, no. 4, April 1998, pp. 428–430.
- [26] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, “Field strength and its variability in vhf and uhf land mobile radio service,” *Review of the Elec. Comm. Lab*, vol. 16, no. 9–10, May 1968, pp. 825–873.
- [27] B. Pattan, “A brief exposure to ultra-wideband signaling,” *Microwave Journal*, Dec. 2003, pp. 104–110.
- [28] J. G. Proakis, *Digital Communications*. McGraw-Hill, 3rd ed., 1995.
- [29] T. S. Rappaport, *Wireless Communications*. Prentice Hall, 2nd ed., 2002.
- [30] S. R. Saunders and A. Aragon-Zavala, *Antennas and Propagation for Wireless Communication Systems*. Wiley, 2nd ed., 2007.
- [31] C. Schlegel, *Trellis Coding*. Wiley-IEEE Press, 1997.
- [32] C. Schlegel and L. Perez, *Trellis and Turbo Coding*. Wiley-IEEE Press, 2003.
- [33] H. Schulze and C. Lueders, *Theory and Applications of OFDM and CDMA: Wideband Wireless Communications*. Wiley, 2005.
- [34] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, 1948.
- [35] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.
- [36] G. Ungerboeck, “Channel coding with multilevel/phase signals,” *IEEE Trans. on Information Theory*, vol. 28, no. 1, Jan. 1982, pp. 55–67.
- [37] G. Ungerboeck, “Trellis-coded modulation with redundant signal sets, Part I: Introduction,” *IEEE Communications Magazine*, vol. 25, no. 2, Feb. 1987, pp. 5–11.
- [38] G. Ungerboeck, “Trellis-coded modulation with redundant signal sets, Part II: State of the art,” *IEEE Communications Magazine*, vol. 25, no. 2, Feb. 1987, pp. 12–21.
- [39] S. Weinstein and P. Ebert, “Data transmission by frequency-division multiplexing using the discrete Fourier transform,” *IEEE Trans. on Comm.*, vol. 19, no. 5, Oct. 1971, pp. 628–634.
- [40] G. Welch and G. Bishop, “An introduction to the Kalman filter,” 1995.
- [41] B. Widrow, J. McCool, and M. Ball, “The complex LMS algorithm,” *Proceedings of the IEEE*, April 1975, pp. 719–720.